B

Bayes's Theorem

Konstantinos Modis
School of Mining and Metallurgical Engineering, National Technical University of Athens, Athens, Greece

Synonyms

Bayes's law; Bayes's rule

Definition

In probability theory, Bayes's theorem can be formally stated according to the following equation:

$$P(\mathcal{A}|\mathcal{B}) = \frac{P(\mathcal{B}|\mathcal{A})P(\mathcal{A})}{P(\mathcal{B})} \tag{1}$$

where $\mathcal{A}$ and $\mathcal{B}$ are events, $P(\mathcal{A})$ and $P(\mathcal{B})$ are the probabilities of these events to occur with $P(\mathcal{B}) \neq 0$, and $P(\mathcal{A}|\mathcal{B})$ and $P(\mathcal{B}|\mathcal{A})$ are the conditional probabilities that $\mathcal{A}$ given $\mathcal{B}$ and $\mathcal{B}$ given $\mathcal{A}$ occur, respectively (Papoulis and Pillai 2002).

Application to Mineral Exploration or a "Mineral Adventure"

In mineral exploration, managers are often called upon to take decisions in extremely uncertain environments. As an example, consider the CEO of a small company that has been appointed to investigate an area where indications show that an economically interesting orebody may occur. On this account, the company is about to start the mineral exploration activities from the preliminary phase up to a pilot mine excavation and trial exploitation of the mineralization (Fig. 1). Moreover, since this is a step-by-step procedure, there is always the option to terminate the development of the project in the case it is no longer cost-effective, based on evidence.

An overview of the recorded experience on similar setups and a consultant geologist's advice says that the à priori probability of finding an economic mineralization may be up to 55%. Because this probability is not satisfactory and in order to increase the level of confidence on the presence of the deposit, the CEO considers the option to conduct first a geophysical survey of the subsurface, which, as also shown by experience, has a reliability of 80%. After reasonable thinking, he chooses to proceed on the exploration. This judgment pays back with a positive result.

At this point, the company has to decide whether to proceed to the next step of the exploration, which is to start a drilling campaign in order to collect real samples from the subsurface. But this decision is difficult because drilling might be very expensive. The CEO has to convince the board of directors that the new evidence is sufficient to support this decision (or is not?). How will he estimate the updated probability of the existence of the orebody? How can the new information be incorporated into the already accumulated knowledge? Bayes's theorem can solve this problem.

Consider translating event $\mathcal{A}$ from Eq. 1 as the existence of an orebody in the area under exploration and event $\mathcal{B}$ as the result of the geophysical campaign (positive or negative for simplicity). Then, $P(\mathcal{A})$ is the à priori or prior probability of existence (which means before the present study), and $P(\mathcal{A}|\mathcal{B})$ is the updated (with the geophysical measurements) or the posterior probability of existence. Furthermore, $P(\mathcal{B}|\mathcal{A})$ can be conveniently translated as the likelihood of the data (geophysical measurements) and, finally, $P(\mathcal{B})$ as the total probability that geophysical exploration will give a positive result, that is, a constant number that ensures that the output will always be between 0 and 1. In this line, Bayes's theorem can be written as

B. S. Daya Sagar et al. (eds.), *Encyclopedia of Mathematical Geosciences*, Encyclopedia of Earth Sciences Series,
https://doi.org/10.1007/978-3-030-26050-7_440-1

Bayes's Theorem, Fig. 1 In the Mineral Exploration context, Bayes's theorem is used to enhance our knowledge on the mineralization by gradually assimilating new information coming from indirect or inexact sources. The screenshot is from "Mineral Adventure," a simulation game developed by National Technical University of Athens in the framework of EU-funded "Virtual Mine" project (https://virtualmine.net/). You can practice strategic decision-making with Bayes's theorem in Mineral Exploration by playing this game in http://geostatistics.eu/exercise-en.php#

$$\text{Posterior probability} = \frac{\text{Prior probability} \times \text{Likelihood of the data}}{\text{Normalization constant}} \tag{2}$$

According to the above, the normalization constant $P(\mathcal{B})$ has to be estimated first. By the total probability theorem (Papoulis and Pillai 2002),

$$P(\mathcal{B}) = P(\mathcal{B}|\mathcal{A})P(\mathcal{A}) + P(\mathcal{B}|\mathcal{A}')P(\mathcal{A}') \tag{3}$$

where $\mathcal{A}'$ is the complementary event of $\mathcal{A}$, i.e., a deposit does not exist.

After substitution in Eq. 3, we get

$$P(\mathcal{B}) = 0.8 \times 0.55 + 0.2 \times 0.45 = 0.53 \tag{4}$$

Inserting this probability into Bayes's theorem, we finally obtain the updated probability of existence:

$$P(\mathcal{A}|\mathcal{B}) = \frac{0.8 \times 0.55}{0.53} = 0.83 \tag{5}$$

What unveils here is that Bayes's theorem is a valuable tool to assimilate the newly gained information into the knowledge base. In this example, since the geophysical method used is reliable, the belief that a deposit may exist is substantially improved, and now it is reasonable to proceed to the drilling campaign.

Application to Geological Systems Modeling or "Like Oil and Water"

Modeling of geological systems like groundwater tables or petroleum reservoirs is typical in geosciences (Pyrcz and Deutsch 2014). The scope of modeling (or simulation or forward problem) is to predict the response of the system to certain external stimuli. For example, in a petroleum reservoir, we would like to know the flow characteristics of oil in a certain production well. The usual difficulty in modeling a geological system is the inadequate knowledge of the parameters of the physical law representing its response to the stimulation. This is because these parameters are in general some regionalized quantities referring to soil or rock mass properties, and thus they may present large variability from site to site. In the inverse problem, the results of a set of measurements are used in order to infer the actual values of the parameters characterizing the geological system. But, although the forward problem normally has a unique solution, the inverse problem has not. This characteristic is referred to as "ill-posedness" and is the main reason why stochastic methods are preferable in order to solve the inverse problem (Tarantola 2005). Instead of focusing on exact numbers, by representing the spatial parameter field with a random function, the idea is to estimate intervals containing the real values of the parameters, given some uncertain measurements of system response.

Summarizing the above in a mathematical notation, the model of a physical system g with boundary conditions $\mathbf{u}$ can be described by its parameter set $\mathbf{m} = \{m_1, \ldots, m_n\}$ and their spatial coordinates $\mathbf{x}$ over a period t. Then, the forward solution of the system can be written as

$$\mathbf{d} = g(\mathbf{x}, t, \mathbf{u}; \mathbf{m}) \tag{6}$$

where $\mathbf{d} = \{d_1, \ldots, d_p\}$ is the response of the system.

Since interest is now focused on the parameters, Eq. 6 can be written for simplicity as

$$\mathbf{d} = g(\mathbf{m}) \tag{7}$$

As previously stated, the objective of the inverse problem in geosciences is to estimate the spatiotemporal distribution of the parameters $\mathbf{m}$ by using all the available information such as physical laws, measurements and moments concerning the parameters, and also measurements of the response variables (e.g., head measurements). Under the probabilistic formalism, the response variable $\mathbf{d}$ and the parameter set of the system are considered as random variables. The vector of observations $\mathbf{d}_{\text{obs}}$ is considered as the response from a particular parameter set, which has to be determined. In order to define a statistical distribution for $\mathbf{d} \mid \mathbf{m}$, a known error distribution of the observations has to be assumed (likelihood of data). Also, it is important to consider a known distribution $f(\mathbf{m})$ of the parameters (prior knowledge). Then, according to Bayes's theorem, the a posteriori distribution of the parameters is defined as

$$f(\mathbf{m}|\mathbf{d}) = \frac{f(\mathbf{d}|\mathbf{m})f(\mathbf{m})}{f(\mathbf{d})} \tag{8}$$

where $f(\mathbf{d}|\ \mathbf{m})$ is the likelihood of data and $f(\mathbf{m})$ is the prior distribution of the parameter set (see derivation below). Repeating the calculation of Eq. 8 for different parameter sets derived from the prior distribution and producing different $\mathbf{d}_{\text{obs}}$ sets, we can obtain the a posteriori distribution of $\mathbf{m}$. Then, the best estimator of $\mathbf{m}$ is the one that maximizes the a posteriori probability $f(\mathbf{m}|\ \mathbf{d})$. This methodology is known as the maximum à posteriori probability (MAP) (Oliver et al. 2008).

Derivation

In its simple form for events as in Eq. 1, Bayes's theorem is derived from the definition of conditional probability:

$$P(\mathcal{A}|\mathcal{B}) = \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{B})}, \text{with } P(\mathcal{B}) \neq 0 \text{ and}$$

$$P(\mathcal{B}|\mathcal{A}) = \frac{P(\mathcal{B} \cap \mathcal{A})}{P(\mathcal{A})}, \text{with } P(\mathcal{A}) \neq 0$$

where $P(\mathcal{A} \cap \mathcal{B})$ is the joint probability of events $\mathcal{A}$ and $\mathcal{B}$.

$$\text{Since } P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{B} \cap \mathcal{A}) \Rightarrow P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A}|\mathcal{B})P(\mathcal{B}) = P(\mathcal{B}|\mathcal{A})P(\mathcal{A})$$

$$\Rightarrow P(\mathcal{A}|\mathcal{B}) = \frac{P(\mathcal{B}|\mathcal{A})P(\mathcal{A})}{P(\mathcal{B})}, \text{if } P(\mathcal{B}) \neq 0$$

In the continuous case of random variables as used in Eq. 8, the theorem can be derived in an analogous manner from the definition of conditional probability (Fig. 2):

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \text{ and}$$

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

Hence,

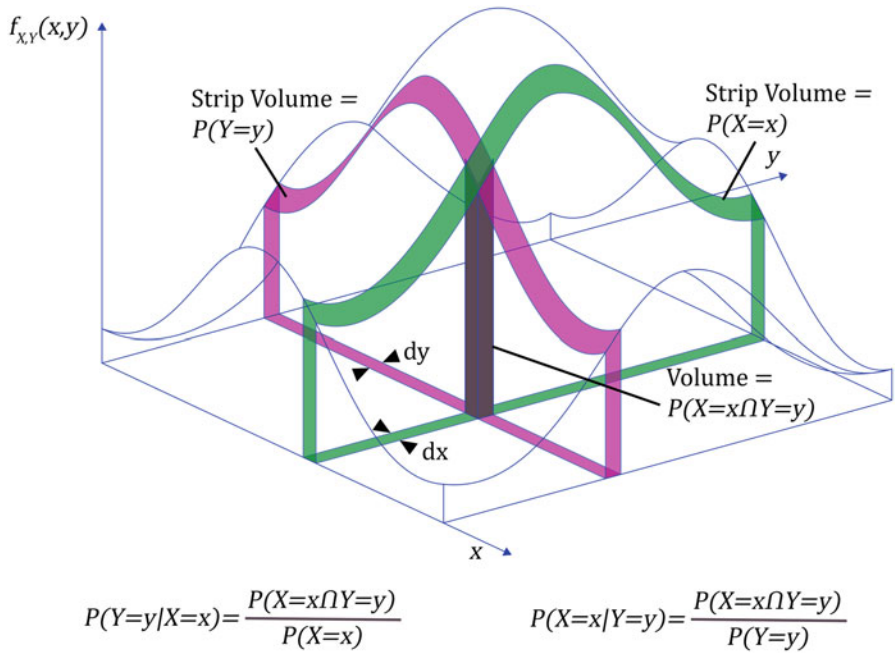
$$f_{X|Y=y}(x) = \frac{f_{Y|X=x}(y)f_X(x)}{f_Y(y)} \tag{9}$$

Bayesian Statistics

Stochastic methods in geosciences do not adopt in general the classic statistical but rather the Bayesian view, where probabilities represent states of knowledge or accessible information. The main point in this view is that the variable of interest over an area is modeled as a random function because there is inadequate information to enable the use a deterministic model, rather than due to the "law of large numbers," that is, because repeated observations indicate a statistical regularity (Kitanidis 2007).

Thus, in Bayesian statistics, probability is used as a measure of uncertainty, or "belief," or available information about an event. In the context of spatial estimation, which is of main concern in geosciences, the essence is to comply with two requirements: Include all previous knowledge about spatial variability in a highly informative prior distribution and then use the observations to infer a maximized posterior distribution through Bayes's relation (Christakos 1990; Valakas and Modis 2016). Classical Kriging estimators can also derive under this generalized perspective: Since prior information

Bayes's Theorem, Fig. 2 On the proof of continuous case of Bayes's theorem: Conditional probabilities applied to the event space of random variables X and Y. By equating these probabilities in the infinitesimal prism formed by the intersection of the two stripes, an instance of Bayes's theorem is revealed for each point in the domain. Integrating these instances for x and y, Eq. 9 is derived



is limited to the mean and covariance functions and only exact measurements are considered, maximization of entropy leads to Gaussian prior and posterior distributions. Then, the minimum mean squared error estimator equals the conditional mean according to Bayes's theorem (Christakos 2000).

Conclusions

Bayes's theorem relates events or random variables by avoiding using their joint probability in favor of conditional and marginal ones. In many cases, these probabilities are easier to define.

In mineral exploration, it provides the tools to update our knowledge on the subsurface with data from different sources of fuzzy information.

In geological modeling, Bayes's theorem is used to transform the prior distribution of the system's parameters into a narrower posterior distribution by using a set of uncertain measurements of the response variables.

Cross-References

- ▶ Bayesian Inversion
- ▶ Bayesian Maximum Entropy
- ▶ Ensemble Kalman Filtering
- ▶ Exploratory Data Analysis
- ▶ Inversion Theory in Geosciences
- ▶ Petroleum Exploration and Recovery
- ▶ Random Variable

Bibliography

Christakos G (1990) A Bayesian/maximum-entropy view to the spatial estimation problem. Math Geol 22:763–777

Christakos G (2000) Modern spatiotemporal geostatistics. Oxford University Press, New York. 288 p

Kitanidis P (2007) On stochastic inverse modeling. Geophys Monogr Ser 171:19–30

Oliver D, Reynolds A, Liu N (2008) Inverse theory for petroleum reservoir characterization and history matching. Cambridge University Press, New York. 380 p

Papoulis A, Pillai U (2002) Probability, random variables and stochastic processes. McGraw Hill, New York. 852 p

Pyrcz M, Deutsch C (2014) Geostatistical Reservoir Modeling. Oxford University Press, New York. 433 p

Tarantola A (2005) Inverse problem theory and methods for model parameter estimation. Siam, Philadelphia. 342 p

Valakas G, Modis K (2016) Using informative priors in facies inversion: the case of C-ISR method. Adv Water Resour 94:23–30