

A Bayesian Approach to the Brain

Florent Meyniel, Ph.D.

Cognitive Neuroimaging Unit, Neurospin, CEA, University Paris-Saclay, France

Bayesian concepts are appealing to many researchers in fundamental and applied research, including neuroscience. Bayesian tools, part of probability theory, are useful whenever quantitative analysis is needed, such as in statistics, data mining, or forecasting. However, Bayesian concepts have much further reaching implications in neuroscience. They are essential to the way we think about the brain.

BAYES' RULE BASICS

The mathematical foundation of Bayesian concepts stems from the so-called Bayes' rule, named after one of its contributors, the 18th century British Reverend Thomas Bayes. Let's consider a practical example of how Bayes' rule works. A medical doctor faced with the following data D , a patient with a cough, contemplates three hypothetical diseases: a lung cancer (H_1), a cold (H_2) or gastroenteritis (H_3). The relative merit of each hypothesis can be deconstructed as follows according to Bayes' rule. Patients usually cough when afflicted by lung cancer or a cold but rarely in the case of gastroenteritis. Therefore, the likelihood of the potential cause for the cough is high under H_1 and H_2 and low under H_3 . Second, a cold and gastroenteritis are much more prevalent diseases than lung cancer in the general population. The *a priori* likelihood of H_2 and H_3 is much higher than that of H_1 . Given that only H_2 scores high both in *a priori* and current evidence, the most likely disease given the symptoms is a cold.



Florent Meyniel, Ph.D.

Stated more generally, Bayes' rule says that our degree of belief in a hypothesis H given some current data D depends on the *a priori* likelihood of this hypothesis (what we know about it, independent of the current data), and the likelihood of the current data given this hypothesis. Formally, degrees of belief and

likelihoods correspond to probabilities [1] and Bayes' rule reads:

$$p(H|D) = p(D|H) \cdot p(H) / p(D)$$

Bayes' rule distinguishes between our belief *a priori* in the hypothesis $p(H)$ and our belief in this hypothesis *a posteriori*, $p(H|D)$, once particular data are considered to evaluate it. The notation $p(D|H)$ is a shorthand for the probability of D given that we know H (the so-called likelihood of the data) and $p(H|D)$ for the probability of H given that we know D .

Several aspects of Bayes' rule are noteworthy. First, it is extremely general – H and D may be any sort of variables as long as they can be assigned a probability. Second, Bayes' rule is *quantitative*: the posterior probability on the left hand side accepts only one value that depends on the terms in the right hand side. This means that Bayes' rule offers a unique way to combine uncertain quantities such as current evidence and prior knowledge in order to estimate the likelihood of a conclusion. In that sense, Bayes' rule is normative: any other estimate is an over- or under-estimation of the likelihood of the conclusion. This normative nature of Bayes' rule can be seen as an extension of classical logic. With classical logic, one can derive the validity of a conclusion, which is either true or false from premises that are known for sure. With Bayes' rule, one can derive the likelihood of a conclusion, which varies on a continuum, from premises that suffer from uncertainty. Another key aspect of Bayes' rule is its symmetry: $p(H|D)$ and $p(D|H)$ appear on opposite sides of the equation which allows going from one to the other. The likelihood of current data given a particular hypothesis – $p(D|H)$ – corresponds to solving a direct or “forward” problem: estimating what should be observed given a known cause. Bayes' rule allows reversing the logic to infer what might be the unknown cause of particular observations – $p(H|D)$.

HOW THE BRAIN IS BAYESIAN

With these mathematical foundations in mind, the brain can be said to be Bayesian in at least three ways. A first key idea is that the brain computes and represents quantities that are probabilistic [2]. In the perceptual domain, this means that every feature of a visual scene is represented by probabilities. For instance, the orientation of a line is not encoded as a single tilt value, but as a distribution of tilt values across several neurons in the visual cortex. Indeed, each of these neurons is tuned for a particular orientation and it responds more intensely when the input data conform to its preferred orientation. Such a neuron therefore acts as a “likelihood detector”: its activity signals the probability of the line having its preferred orientation. Because different neurons are tuned to different orientations, their activity collectively encodes the likelihood of the tilt [3]. This probabilistic view may contrast with the apparent “oneness” of perception. When viewing a scene, we access only one percept at a time, and not distinct hypothetical percepts associated with probabilities. However, recent theories show that this all-or-none processing is the exception rather than the rule in the brain. This “oneness” results from conscious processes that select and amplify one possible interpretation among many [4]. By contrast, most brain processes operate without consciousness and rely on distributions of values and probabilistic computations.

A second Bayesian view of the brain is that the internal knowledge and percepts represented by neurons are constructed following Bayes’ rule. This internal knowledge therefore constitutes a posterior belief about the causes of the inputs received by the brain [5,6]. This inference is usually fraught with uncertainty as the brain must make sense of the world based on inputs that are limited and ambiguous. For instance, different three-dimensional shapes in the world may result in the same image once they are projected onto our eyes. There is therefore a real challenge for the brain to perceive the world despite the paucity and the ambiguity of its inputs. This is an old idea in psychology, identified by the 19th-century German scientist von Helmholtz. The Bayesian framework is made to handle inference from uncertain data, and it even offers a principled remedy: combining the uncertain evidence provided by sensory inputs with prior knowledge.

There is ample experimental evidence that perception relies on prior information to compensate for the poverty of the inputs received. Many biases and visual illusions reveal this automatic reliance on prior

information. For instance, when observers are asked to evaluate the tilt of a line, they tend to perceive lines that are nearly vertical as purely vertical, and nearly horizontal as purely horizontal. These orientations are indeed much more frequent in our world. The perceived orientation of a line that weakly departs from these frequent orientations is therefore dominated by our prior expectations [7]. Studies in non-human animals showed that these priors are learned during development from experience. As a result, priors become part of our cortical networks in such a way that they shape their spontaneous activity [8]. When there is no stimulus to drive neuronal activity, the spontaneous activity is dominated by prior expectations. This is because in the absence of input data, the posterior probability in Bayes’ rule boils down to the prior probability.

Lastly, Bayes’ rule allows for inferring the causes of current observations. By building on this knowledge of the causes, one can in turn predict future observations [5]. This predictive nature of Bayes’ rule is the third pillar of the Bayesian view of the brain. Brain imaging and recordings of neurons show that the brain constantly uses previous observations to form expectations about the upcoming events. Such expectations can build up rapidly even in very simple contexts. For instance, upon hearing the four tones “bip”, “bip”, “bip”, “bip” in a row, you may expect that the fifth sound will be another “bip”. Several brain regions increase their activity if the fifth sound is “bop” instead of “bip” [9–11]. This increased activity signals that there is an error: the current expectation appears violated. Interestingly, this error signal is much larger when the expectation was high. An even larger response is recorded if the deviant sound occurs after ten repetitions of “bip” as compared to only four such repetitions. These error signals are actually quantitative: in this simple experiment, they match the expected frequency of sounds that can be inferred using Bayes’ rule and the sounds already presented. It is noteworthy that individuals with schizophrenia exhibit significantly lower error signals on electroencephalograms than do healthy individuals in this kind of paradigm, suggesting that statistical inference might be impaired in this pathology. [12,13]. Other experiments used carefully designed sequences of stimuli to show that the brain is capable of learning more complex statistics and even abstract rules [14,15]. Remarkably, experiments in infants and babies showed that this Bayesian machinery operates early in life. Young babies are already capable of quantitative predictions based only on a few observations [16,17].

A major strength of the Bayesian view of the brain is its unifying power. The few examples reported here show that many brain processes can be accounted for by Bayesian principles. It is true across species (in humans and other animals), spatial scales (from single neurons to neuronal networks to brain-scale circuits), cognitive domains (perception, learning, decision making) and stages of development (in neonates, infants and adults). It may even be true of evolution. This is because Bayes' rule is normative: if a particular process deviates from it, then other processes, closer to Bayes' rule, will do better. By selection, processes should gradually approach Bayes' rule, as we see in well-tuned systems such as the human visual cortex.

FUTURE CHALLENGES

This Bayesian view has proved quite successful in neuroscience, although controversy should be acknowledged [18,19]. Challenges nonetheless remain for the future. The most critical one is that Bayesian principles constrain what computations should be, but they leave their implementation entirely open. Indeed, there are often many different ways to solve the same computation. Future works will aim at identifying the specific algorithms that the brain uses for Bayesian computations.

Another challenge is that Bayesian views have been applied so far mostly to perception because this is the domain in which neuroscience is the most advanced. However, future works will probe Bayesian computations in other domains, such as decision making [20–23]. They should also probe the extent to which Bayesian computations and their associated uncertainty levels are accessible to introspection. Recent studies showed that the “sense of confidence” – the degree of belief that we attach to our percepts, memories and decisions – is actually much more sophisticated in humans than previously envisaged [24,25].

Further readings:

https://en.wikipedia.org/wiki/Bayes'_theorem

<http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004305>

<https://elifesciences.org/content/5/e11476>

Supporting references

1. Jaynes ET. Probability Theory: The Logic of Science. Cambridge University Press; 2003.
2. Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 2004;27: 712–719. doi:10.1016/j.tins.2004.10.007
3. Deneve S, Latham PE, Pouget A. Reading population codes: a neural implementation of ideal observers. *Nat Neurosci.* 1999;2: 740–745.
4. Dehaene S. *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts.* New York: Viking; 2014.
5. Friston K. The free-energy principle: a unified brain theory? *Nat Rev Neurosci.* 2010;11: 127–138. doi:10.1038/nrn2787
6. Rao RP. An optimal estimation approach to visual perception and learning. *Vision Res.* 1999;39: 1963–1989.
7. Girshick AR, Landy MS, Simoncelli EP. Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nat Neurosci.* 2011;14: 926–932. doi:10.1038/nn.2831
8. Berkes P, Orbán G, Lengyel M, Fiser J. Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science.* 2011;331: 83–87. doi:10.1126/science.1195870
9. Huettel SA. Decisions under Uncertainty: Probabilistic Context Influences Activation of Prefrontal and Parietal Cortices. *J Neurosci.* 2005;25: 3304–3311. doi:10.1523/JNEUROSCI.5070-04.2005
10. Karoui IE, King J-R, Sitt J, Meyniel F, Gaal SV, Hasboun D, et al. Event-Related Potential, Time-frequency, and Functional Connectivity Facets of Local and Global Auditory Novelty Processing: An Intracranial Study in Humans. *Cereb Cortex.* 2014; bhu143. doi:10.1093/cercor/bhu143
11. Squires KC, Wickens C, Squires NK, Donchin E. The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science.* 1976;193: 1142–1146. doi:10.1126/science.959831
12. Fletcher PC, Frith CD. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci.* 2009;10: 48–58. doi:10.1038/nrn2536
13. Michie PT, Malmierca MS, Harms L, Todd J. The neurobiology of MMN and implications for schizophrenia. *Biol Psychol.* 2016;116: 90–97. doi:10.1016/j.biopsycho.2016.01.011
14. Wacongne C, Changeux J-P, Dehaene S. A Neuronal Model of Predictive Coding Accounting for the Mismatch Negativity. *J Neurosci.* 2012;32: 3665–3678. doi:10.1523/JNEUROSCI.5003-11.2012

15. Wang L, Uhrig L, Jarraya B, Dehaene S. Representation of Numerical and Sequential Patterns in Macaque and Human Brains. *Curr Biol CB*. 2015;25: 1966–1974. doi:10.1016/j.cub.2015.06.035
16. Frank MC, Tenenbaum JB. Three ideal observer models for rule learning in simple languages. *Cognition*. 2011;120: 360–371. doi:10.1016/j.cognition.2010.10.005
17. Téglás E, Vul E, Girotto V, Gonzalez M, Tenenbaum JB, Bonatti LL. Pure Reasoning in 12-Month-Old Infants as Probabilistic Inference. *Science*. 2011;332: 1054–1059. doi:10.1126/science.1196404
18. Bowers JS, Davis CJ. Bayesian just-so stories in psychology and neuroscience. *Psychol Bull*. 2012;138: 389–414. doi:10.1037/a0026450
19. Griffiths TL, Chater N, Norris D, Pouget A. How the Bayesians got their beliefs (and what those beliefs actually are): Comment on Bowers and Davis (2012). *Psychol Bull*. 2012;138: 415–422. doi:10.1037/a0026884
20. Beck JM, Ma WJ, Kiani R, Hanks T, Churchland AK, Roitman J, et al. Probabilistic Population Codes for Bayesian Decision Making. *Neuron*. 2008;60: 1142–1152. doi:10.1016/j.neuron.2008.09.021
21. Chater N, Tenenbaum JB, Yuille A. Probabilistic models of cognition: Conceptual foundations. *Trends Cogn Sci*. 2006;10: 287–291.
22. Pouget A, Beck JM, Ma WJ, Latham PE. Probabilistic brains: knowns and unknowns. *Nat Neurosci*. 2013;16: 1170–1178. doi:10.1038/nn.3495
23. Solway A, Botvinick MM. Goal-directed decision making as probabilistic inference: A computational framework and potential neural correlates. *Psychol Rev*. 2012;119: 120–154. doi:10.1037/a0026435
24. Meyniel F, Schlunegger D, Dehaene S. The Sense of Confidence during Probabilistic Learning: A Normative Account. *PLoS Comput Biol*. 2015;11: e1004305. doi:10.1371/journal.pcbi.1004305
25. Meyniel F, Sigman M, Mainen ZF. Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron*. 2015;88: 78–92. doi:10.1016/j.neuron.2015.09.039