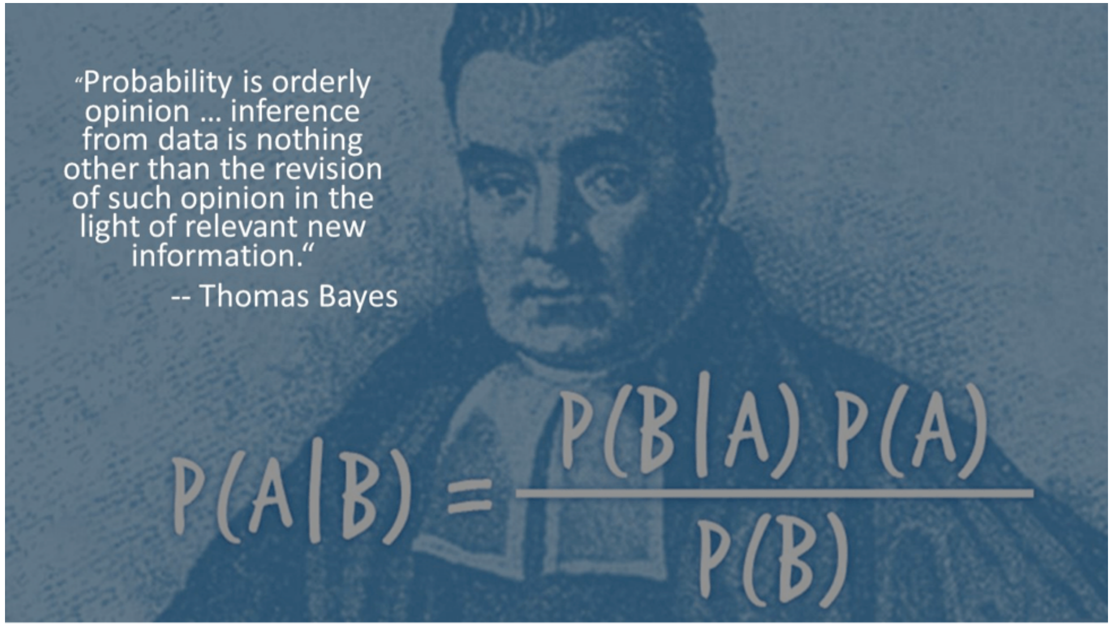
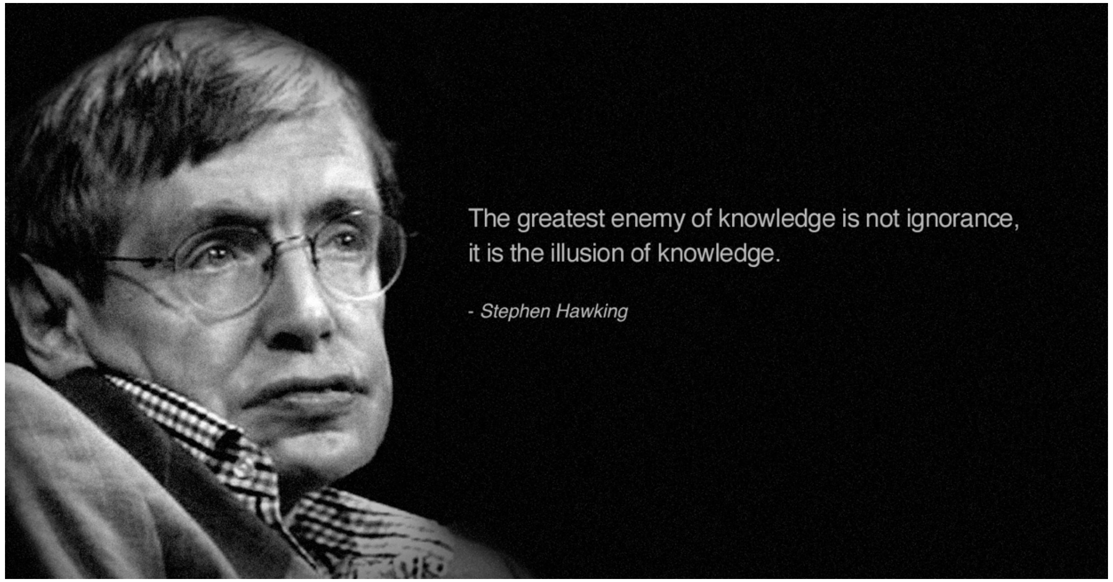
Bayesian Inference

Min Ding

March 21, 2022

Thomas Bayes (1701-1761)





Bayes' Theorem

(1976, by Richard Price)

LII. *An Essay towards solving a Problem in the Doctrine of Chances. By the late Rev. Mr.* Bayes, *communicated by Mr.* Price, *in a letter to* John Canton, *M. A. and F. R. S.*

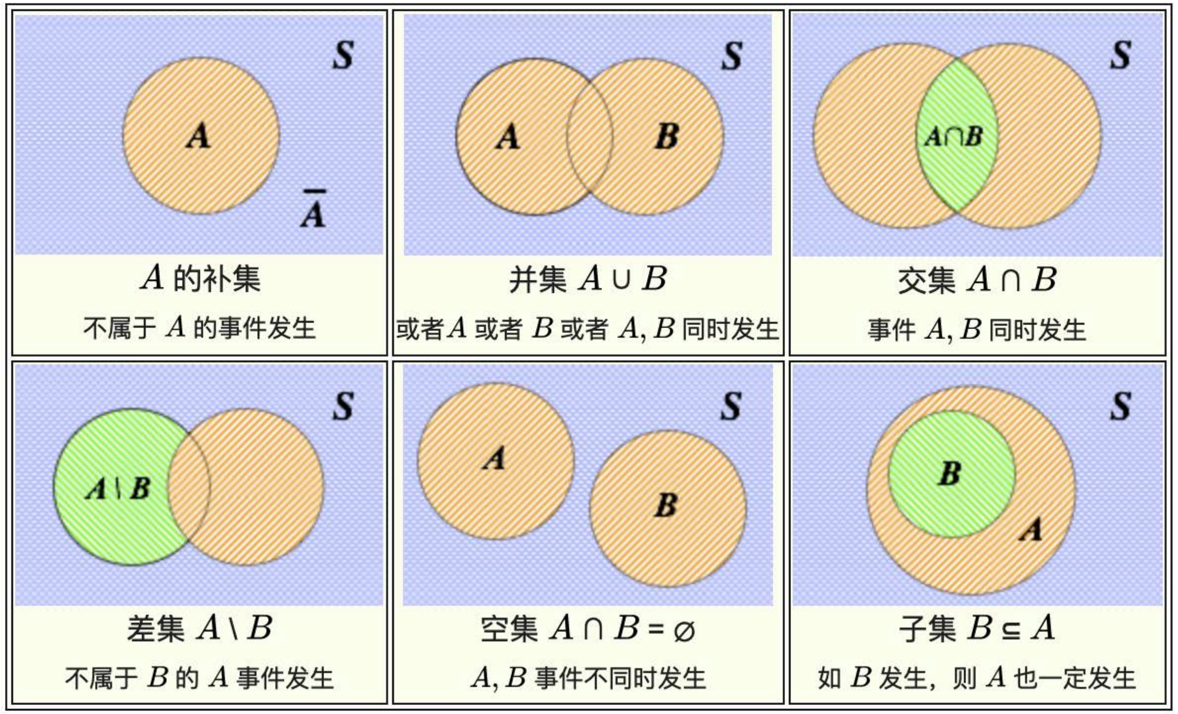
Dear Sir,

Read Dec. 23, 1763. I now send you an essay which I have found among the papers of our deceased friend Mr. Bayes, and which, in my opinion, has great merit, and well deserves to be preserved. Experimental philosophy, you will find, is nearly interested in the subject of it; and on this account there seems to be particular reason for thinking that a communication of it to the Royal Society cannot be improper.

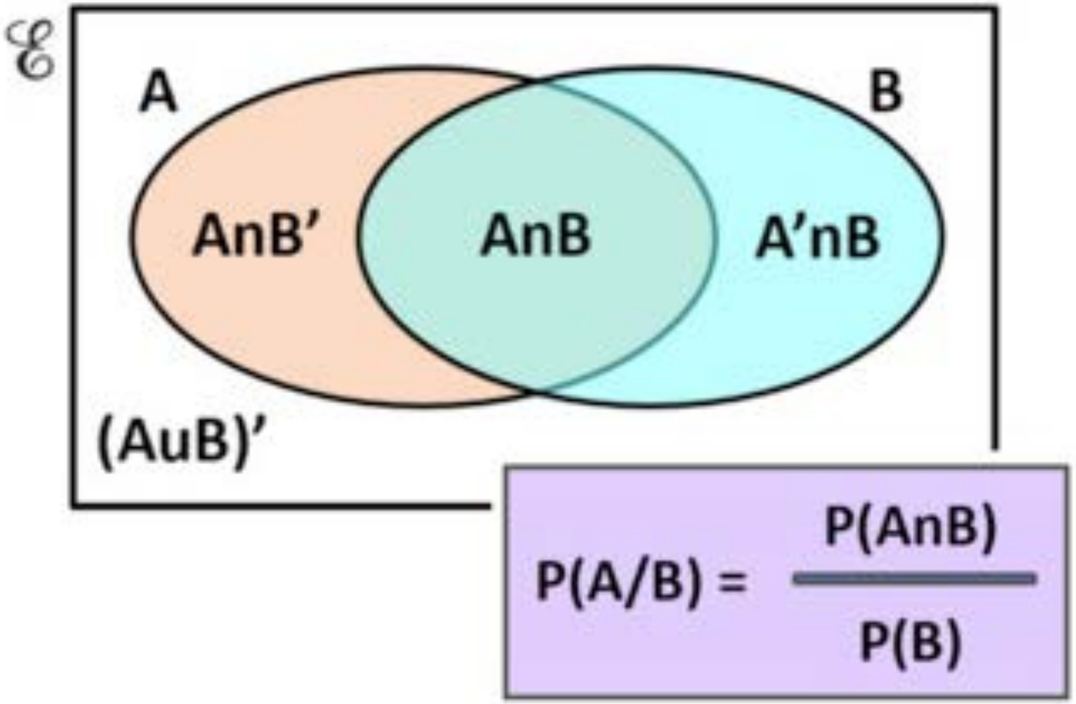
He had, you know, the honour of being a member of that illustrious Society, and was much esteemed by many as a very able mathematician. In an introduction which he has writ to this Essay, he says, that his design at first in thinking on the subject of it was, to find out a method by which we might judge concerning the probability that an event has to happen, in given circumstances, upon supposition that we know nothing concerning it but that, under the same circumstances, it has happened a certain number of times, and failed a certain other number of times. He adds, that he soon perceived that it would not be very difficult to do this, provided some rule could be found, according to which we ought to estimate the chance that the probability for the happening of an event perfectly unknown, should lie between any two named degrees of probability, antecedently to any experiments made about it; and that it appeared to him that the rule must be to suppose the chance the same that it should lie between any two equidifferent degrees; which, if it were allowed, all the rest might be easily calculated in the common method of proceeding in the doctrine of chances. Accordingly, I find among his papers a very ingenious solution of this problem in this way. But he afterwards considered, that the *postulate* on which he had argued might not perhaps be looked upon by all as reasonable; and therefore he chose to lay down in another form the proposition in which he thought the solution of the problem is contained, and in a *Scholium* to subjoin the reasons why he thought it so, rather than to take into his mathematical reasoning any thing that might admit dispute. This, you will observe, is the method which he has pursued in this essay.

Every judicious person will be sensible that the problem now mentioned is by no means merely a curious speculation in the doctrine of chances, but necessay to be solved in order to a sure foundation for all our reasonings concerning past facts, and what is likely to be hereafter. Common sense is indeed sufficient to shew us that, form the observation of what has in former instances been the consequence of a certain cause or action, one may make a judgement what is likely to be the consequence of it another time. and that the larger number of experiments we have to suypport a conclusion, so much more the reason we have to take it for granted. But it is certain that we cannot determine, at least not to

Bayes' Law - Conditional Probability in Venn Diagram



Proof of Bayes' Law

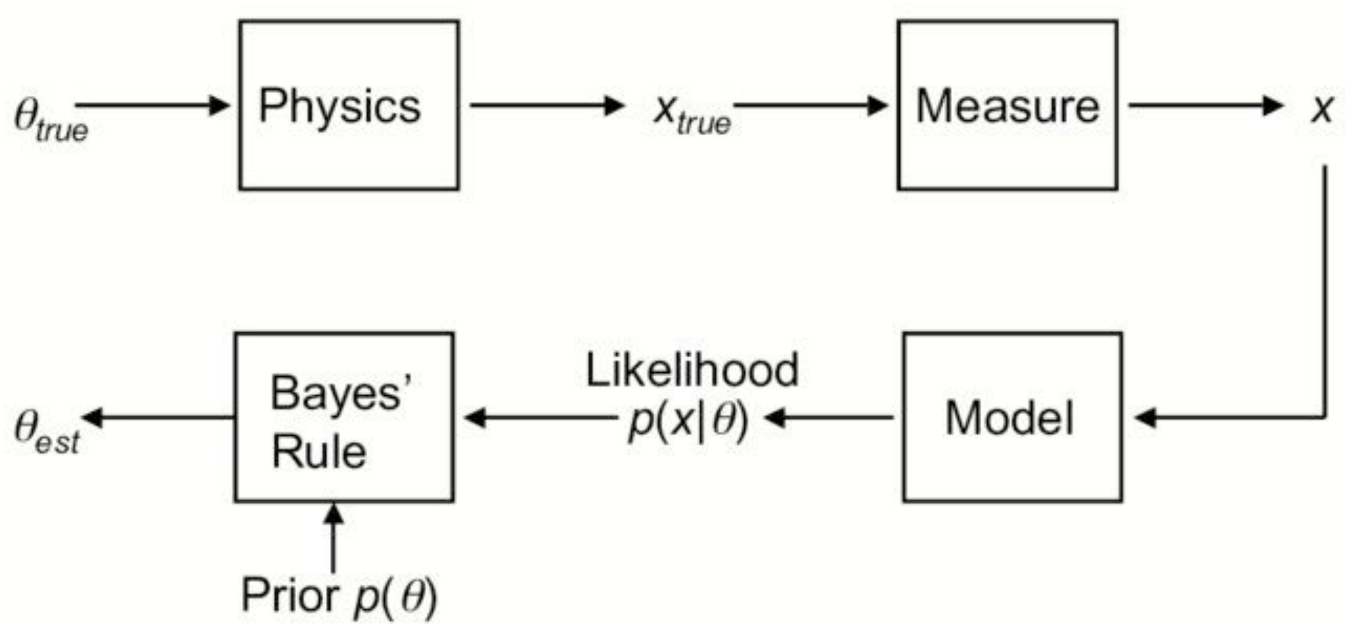


Bayes’ Law - 逆向概率

- 正向概率：假设袋子里面有N个白球，M个黑球，你伸手进去摸一把，摸出黑球的概率是多大？
- 逆概问题：如果我们事先并不知道袋子里面黑白球的比例，而是闭着眼睛摸出一个（或好几个）球，观察这些取出来的球的颜色之后，可以就此对袋子里面的黑白球的比例作出什么样的推测？

Forward probability

Inverse probability



Forward and Inverse Problems

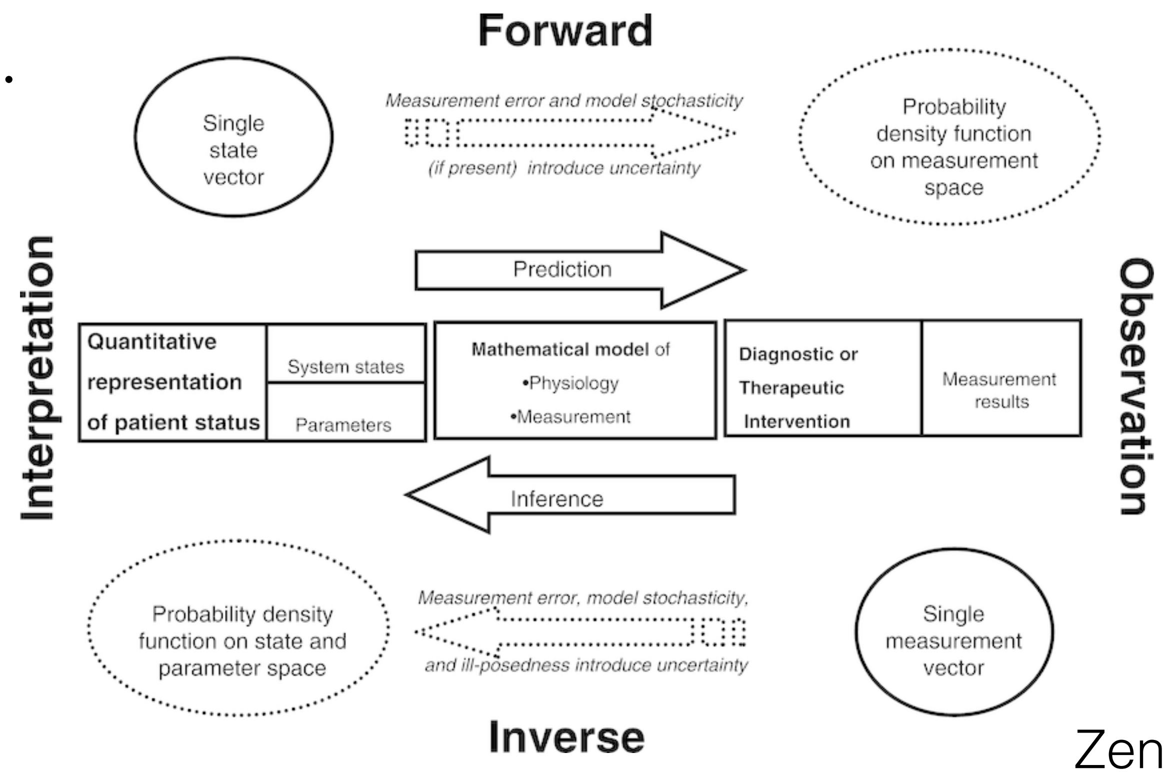
Direct/Forward

Based on Theory d = g(m)

Models *m* ⇄ **Data *d***

Reverse/Inverse

e.g.



Zenker+2007

Probability Density Function

In the case of a continuous random variable, the probability taken by X on some given value x is always 0. In this case, if we find P(X = x), it does not work. Instead of this, we must calculate the probability of X lying in an interval (a, b). Now, we have to figure it for P(a< X< b), and we can calculate this using the formula of PDF. The Probability density function formula is given as,

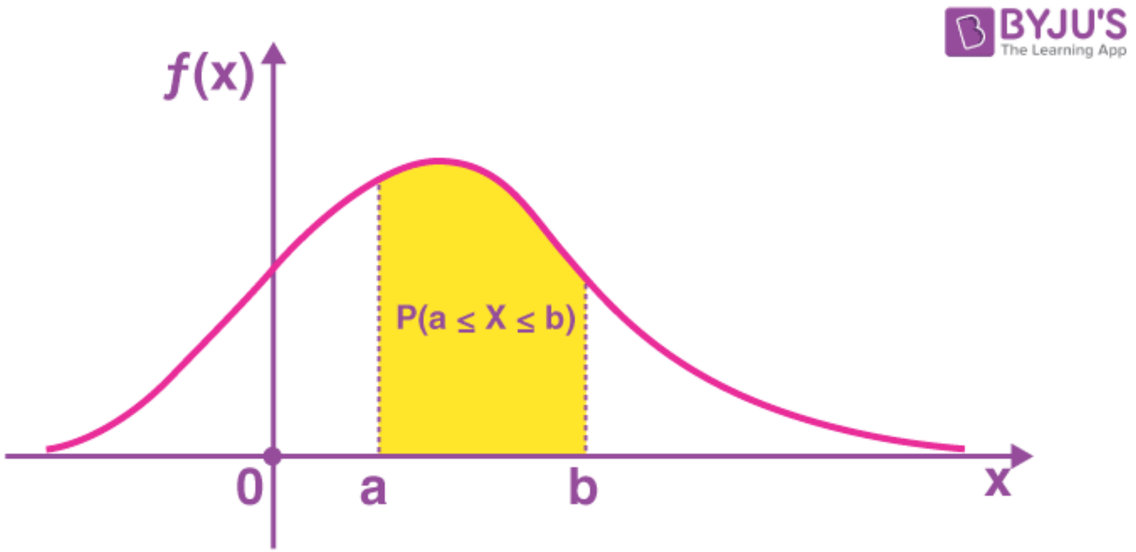
$$\mathbf{P(a < X < b) = \int_a^b f(x)\, dx}$$

Or

$$\mathbf{P(a \le X \le b) = \int_a^b f(x)\, dx}$$

This is because, when X is continuous, we can ignore the endpoints of intervals while finding probabilities of continuous random variables. That means, for any constants a and b,

$P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$.



https://byjus.com/maths/probability-density-function/

Bayesian vs Frequentist (频率学派) - Prior

Bayesian	Frequentist
• Prior knowledge • Updatable with new information • Data are fixed • Parameters are described probabilistically • Complex to compute	• No prior information to model • Data is a repeatable random sample • Parameters are fixed • Easy to compute