Bayes and Price: Reasoning from Effects to Causes (Inverse Probability)

- **The Background Problem - Epistemology and Induction:** The core intellectual challenge during the Enlightenment, when Bayes and Price lived, was understanding how we gain knowledge and justify beliefs based on observations. Philosophers like David Hume had famously raised deep questions about the problem of induction: just because the sun has risen every day *so far* (observed effect), how can we logically justify our strong belief that it will rise tomorrow (inferring a stable underlying cause or regularity)? How do we rationally update our beliefs in the face of new evidence? There wasn't a rigorous mathematical framework for this kind of reasoning.
- **Bayes' Specific Problem (The Billiard Table Thought Experiment):** Bayes' essay, "An Essay towards solving a Problem in the Doctrine of Chances," frames the problem abstractly, often explained with a billiard table analogy (though the essay uses a square frame):
    1. Imagine throwing a ball (Ball W) onto a square table completely at random. Its landing position along one axis defines a hidden probability p (e.g., the probability that subsequent balls land to its right). We don't know p. This p represents the unknown "cause" or "true state" or "parameter."
    2. Now, throw n subsequent balls (Balls O) onto the table, again randomly. Observe how many (k) land to the right of the first ball (Ball W). This observation (k successes in n trials) is the "effect" or the "data."
    3. **The Question:** Given the observed effect (k out of n), what can we rationally conclude about the *location* of the first ball (the value of the unknown probability p, the "cause")? How strongly should we believe that p lies within a certain range?
- **Motivation - Quantifying Belief Update:** Bayes sought a mathematical way to solve this *inverse problem*. Standard probability (forward probability) calculates P(Effect | Cause) – e.g., if we *know* p, what's the probability of getting k successes in n trials (which is given by the Binomial distribution). Bayes wanted to calculate P(Cause | Effect) – given the data k and n, what's the probability that p lies in a specific interval? He was trying to formalize how empirical observations should influence our beliefs about the underlying process generating them.
- **Price's Motivation:** Richard Price found Bayes' unpublished manuscript after his death. Price was a moral philosopher and theologian. He immediately saw the profound implications of Bayes' work for reasoning not just about physical phenomena but also about philosophical and theological arguments. He believed Bayes' method provided a rational basis for:
    - **Inductive Reasoning:** Justifying beliefs about general laws based on specific instances.
    - **Arguments from Design:** Assessing the evidence for a creator or intelligent design (the "cause") based on the observed order and complexity of the universe (the "effect").
    - **Evaluating Testimony:** Assessing the credibility of historical accounts or miracles based on evidence.
      Price wanted a tool to combat skepticism (like Hume's) by showing a mathematical way to strengthen belief in a hypothesis through accumulating evidence.

- **Role of LoTP:** In deriving his famous theorem, Bayes (and later Laplace, who generalized it) needed to calculate the overall probability of observing the effect (k out of n), irrespective of the true underlying p. This is the denominator in Bayes' Theorem, often denoted P(Data). To calculate this, one must consider *all possible values* of the unknown cause p, calculate the probability of the data *given* each specific p, weight it by the prior probability of that p, and sum (or integrate) over all possibilities. This summation/integration step *is* the Law of Total Probability (or its continuous analogue). $P(Data) = \int P(Data \mid p) * P(p)\, dp$
- It's the necessary step of accounting for all the mutually exclusive ways the observed effect could have arisen.

Laplace: Using the Law of Total Probability in service of a Hypothesis

Laplace wielded probability theory as a tool across an astonishing range of scientific problems. While he used the logic of partitioning sample spaces (i.e., LoTP) constantly, a famous example where it contributed to a significant scientific insight is his analysis of the **inclinations of planetary orbits**.

- **The Observation (Effect):** Astronomers observed that the orbits of the known planets in the solar system were not randomly oriented. They were all roughly coplanar, lying close to the plane of the Sun's rotation and Earth's orbit (the ecliptic). Furthermore, they all orbited in the same direction.
- **The Question (Competing Causes):** Was this observed regularity a result of pure chance, or did it point towards a common physical cause during the formation of the solar system?
    - *Hypothesis 1 (Chance):* The orbital planes and directions are random and independent.
    - *Hypothesis 2 (Common Cause):* The planets formed from a single rotating disk of gas and dust (the Nebular Hypothesis, which Laplace strongly supported and developed).
- **Laplace's Application of Probability (Using LoTP Logic):** Laplace set out to calculate the probability of the observed orbital arrangement occurring *purely by chance*.
    1. **Partitioning the Possibilities:** He conceptualized the "event" as the observed low inclinations and common direction of orbits. He needed to calculate P(Observation). Using the LoTP framework, this probability could be thought of as:
       P(Obs) =

       P(Obs | Chance) * P(Chance) + P(Obs | Common Cause) * P(Common Cause)

       (Though he focused mainly on calculating the first term to show how small it was).

    2. **Calculating P(Observation | Chance):** This required considering all possible random orientations of planetary orbits. He treated the orbital inclinations and directions as random variables. He calculated the probability that *if* orientations were truly random, they would *by chance alone* cluster as closely as observed. This

involves calculating probabilities over geometric configurations – a complex counting/integration problem over the space of possible orientations.

3. **The Insight:** Laplace demonstrated that the probability P(Observation | Chance) was extraordinarily small – astronomically improbable. The observed configuration was vastly unlikely to be a coincidence.
4. **Conclusion:** By showing how improbable the effect (observed orbits) was under the "Chance" cause, he provided powerful quantitative evidence against it. This strongly supported the alternative hypothesis: a common physical origin (the "Common Cause"). The observed regularity wasn't luck; it was evidence for the Nebular Hypothesis.

- **Role of LoTP:** The core of the argument relies on the LoTP principle: evaluating the likelihood of an observation by considering the different possible underlying processes (causes) that could have produced it. By calculating the probability of the observation conditional on the "chance" hypothesis and finding it negligible, Laplace made a compelling case for the "common cause" hypothesis, even without directly calculating P(Obs | Common Cause). The logic is inherently comparative, facilitated by partitioning the causal possibilities.

In both cases (Bayes/Price and Laplace), the drive was to move beyond qualitative arguments to quantitative reasoning about uncertainty, causes, and effects. The Law of Total Probability emerged not as an end in itself, but as a fundamental building block necessary to construct these more complex arguments about inference and scientific evidence.