

Order Dependence in Bayesian Updating: A Limitation of Convolution-Based Mechanisms

Abstract

Bayesian inference, a powerful method for updating beliefs in the presence of new data, often relies on convolution-like mechanisms to combine prior distributions with likelihoods. However, this mechanism exposes a fundamental limitation: **the result of Bayesian updating is often highly order-dependent**, meaning that the influence of the most recent data can disproportionately affect the posterior distribution. In this paper, we explore this limitation, identifying cases where order dependence arises and highlighting how the **beta distribution** stands as an exception due to its inherent memory of past samples.

1. Introduction

Bayesian inference is widely used to model uncertainty by updating prior beliefs when new data is observed. The updating process typically follows:

$$\text{Posterior} \propto \text{Prior} \times \text{Likelihood}$$

This operation can be interpreted as a form of **convolution** that integrates the prior information with the likelihood of observed data. While effective, this convolution mechanism introduces a critical limitation: **the result of the posterior distribution is highly sensitive to the order in which data is observed**. This order dependence can distort inference, especially when recent data has a stronger likelihood or when the prior distribution lacks sufficient structure to balance the influence of older observations.

2. The Convolution Mechanism and Order Dependence

2.1 Mathematical Basis of Convolution in Bayesian Updating

Consider a sequence of independent observations x_1, x_2, \dots, x_n used to update a prior distribution. The posterior after n observations is:

$$\text{Posterior} = (\text{Prior} * \text{Likelihood}_{x_1}) * \text{Likelihood}_{x_2} * \dots * \text{Likelihood}_{x_n}$$

Each convolution step integrates the current prior with the new data, progressively refining the posterior. However, this refinement often places disproportionate weight on recent data, particularly when the prior does not retain information about earlier samples.

2.2 Sensitivity to Order in Distributions Without Memory

Many commonly used distributions in Bayesian inference, such as Gaussian, Poisson, or exponential distributions, do **not inherently track the number of observations or their sequence**. The updated

posterior reflects the influence of the most recent likelihood function and may be dominated by high-variance or extreme observations. This sensitivity arises because the convolution mechanism continuously shifts the prior without preserving details about earlier updates.

3. The Beta Distribution as an Exception

3.1 Memory of Past Observations

Unlike most distributions, the **beta distribution** inherently encodes the memory of past observations through its parameters α and β :

$$\alpha_{\text{new}} = \alpha_{\text{old}} + \text{number of successes}$$

$$\beta_{\text{new}} = \beta_{\text{old}} + \text{number of failures}$$

These parameters grow proportionally to the number of samples, ensuring that the posterior reflects the entire history of observations. As more data accumulates, the influence of any single observation diminishes, making the result **less order-dependent**.

3.2 Implications for Bayesian Inference

The beta distribution's memory mechanism explains why it is commonly used as a prior in cases involving proportions and binary outcomes (e.g., beta-binomial models). It ensures that **the posterior depends on the**

cumulative evidence, not the order of observations, making it more robust to noisy or extreme data.

4. Potential Solutions to Address Order Dependence

To mitigate the limitations of order dependence in Bayesian inference, several strategies can be considered:

- **Weighted Priors:** Assign weights to past observations, giving older data sufficient influence to balance the effect of new observations.
 - **Hierarchical Bayesian Models:** Use hierarchical structures to model uncertainty across multiple levels, reducing sensitivity to individual updates.
 - **Batch Processing:** Instead of sequentially updating the posterior, process data in batches to reduce the impact of extreme individual observations.
-

5. Conclusion

The convolution mechanism central to Bayesian updating highlights a key limitation: **order dependence in the posterior distribution**. Distributions that do not retain memory of past observations are particularly susceptible to this issue, with later data often dominating the inference process. The beta distribution stands out as an exception due to its parameter structure, which inherently tracks the sample history. By recognizing and addressing this limitation, more robust Bayesian models can be developed, ensuring that inference remains stable and reliable across different data scenarios.