Bayesian Election Prediction: Intuition and Application

Dr. Ratnesh Prasad Srivastava, GGV, CSIT, Bilaspur, C.G.

date

1 The Bayesian Mindset: Learning from Evidence

Figure 1: The Bayesian learning cycle: Prior belief + New evidence → Updated belief

1.1 Core Intuition

Bayesian reasoning answers: "How should we *update our beliefs* when new evidence arrives?" For elections:

- **Prior belief**: What we know *before* the survey (historical data)
- **Likelihood**: How *probable* is our survey result under different scenarios
- **Posterior**: Our *updated* belief after seeing evidence

2 Step 1: Setting Prior Beliefs (The Foundation)

2.1 Conceptual Basis

- **Analogy**: Your initial forecast before polling day
- **Key insight**: Quantifies historical knowledge (e.g., BJP won 6/10 similar villages)
- **Mathematical representation**:

$$P(\text{BJP}) = 0.6, \quad P(\text{Congress}) = 0.4$$

2.2 Why This Matters

With Historical Data	Without Historical Data
Strong prior (60% BJP)	Neutral prior (50-50)

3 Step 2: Likelihood - Measuring Evidence Strength

3.1 The Likelihood Question

"If BJP were truly leading (60% support), how *surprising* is our survey result of 12 BJP voters out of 20?"

3.2 Binomial Probability Calculation

$$\begin{aligned} P(12 \text{ BJP}|\text{BJP win}) &= \binom{20}{12}(0.6)^{12}(0.4)^{8} \\ &\approx 18\% \\ P(12 \text{ BJP}|\text{Congress win}) &= \binom{20}{12}(0.4)^{12}(0.6)^{8} \\ &\approx 12\% \end{aligned}$$

3.3 Intuitive Interpretation

Figure 2: The survey result (12 BJP) is more likely under BJP-win scenario

4 Step 3: Bayesian Update - The Magic Step

4.1 The Normalization Process

Posterior probability combines prior beliefs and evidence:

$$P(\text{BJP}|\text{Survey}) = \frac{\text{Prior} \times \text{Likelihood}}{\text{Total Evidence}}$$

4.2 Why Normalize?

- **Analogy**: Scaling survey results to account for prior knowledge
- **Key insight**: Ensures probabilities sum to 100%
- **Calculation**:

$$\begin{aligned} \text{Total Evidence} &= (0.6 \times 0.18) + (0.4 \times 0.12) = 0.156 \\ P(\text{BJP}|\text{Survey}) &= \frac{0.6 \times 0.18}{0.156} \approx 73\% \\ P(\text{Congress}|\text{Survey}) &= \frac{0.4 \times 0.12}{0.156} \approx 27\% \end{aligned}$$

5 Step 4: From Survey to Village Prediction

5.1 The Scaling Challenge

"How do we translate survey results (20 voters) to village predictions (250 voters)?"

5.2 Probability Distribution Approach

Model uncertainty using binomial distributions:

$$\begin{aligned} \text{BJP votes} &\sim \text{Binomial}(250, 0.6) \\ \text{Congress votes} &\sim \text{Binomial}(250, 0.4) \end{aligned}$$

5.3 Winning Condition Visualization

Figure 3: Distribution overlap: BJP needs ≥ 126 votes to win

5.4 Normal Approximation

$$\begin{aligned} \mu_{\text{BJP}} &= 250 \times 0.6 = 150 \\ \sigma_{\text{BJP}} &= \sqrt{250 \times 0.6 \times 0.4} \approx 7.75 \\ P(\text{BJP wins}) &= P(X \geq 126) \\ &\approx P\left(Z \geq \frac{126 - 150}{7.75}\right) \\ &= P(Z \geq -3.1) > 99.9\% \end{aligned}$$

6 Reality Check: Assumptions and Limitations

Assumption	Real-World Consideration
Representative sample	Was survey random? Rural voters may be underrepresented
Static preferences	Voters might change minds before election
Two-party system	Third parties could split votes (AAP in our case)
Binomial model	Assumes independent voters (may cluster)

7 Takeaway: The Bayesian Advantage

- **Adaptive learning**: Updates beliefs with new evidence
- **Uncertainty quantification**: Provides probabilistic forecasts
- **Decision-making tool**: "73% chance BJP wins" better than "BJP leads survey"
- **Iterative refinement**: New polls can be incorporated easily