Bayesian Election Prediction: BJP vs Congress

Statistics Module

1 Problem Statement

1.1 What Are We Solving?

We aim to predict the election outcome in a village of 250 voters between two political parties:

- **BJP**: Bharatiya Janata Party

- **Congress**: Indian National Congress

Given:

- Limited survey data (20 voters)

- Historical election patterns

- Village size: 250 voters

Our goal is to compute:

1. Probability of BJP winning

2. Probability of Congress winning

3. Uncertainty in predictions

1.2 Why Bayesian Approach?

Traditional polling has limitations:

- Small samples create large margins of error

- Historical context is ignored

- Uncertainty isn't quantified

Bayesian methods solve these by:

- Incorporating prior knowledge

- Updating beliefs with new evidence

- Providing probabilistic forecasts

2 Solution Approach: Bayes' Theorem

We use the fundamental Bayesian update rule:

$$\underbrace{P(\text{Party}|\text{Survey})}_{\text{Posterior}} = \frac{\overbrace{P(\text{Survey}|\text{Party})}^{\text{Likelihood}} \cdot \overbrace{P(\text{Party})}^{\text{Prior}}}{\underbrace{P(\text{Survey})}_{\text{Evidence}}}$$

3 Step-by-Step Solution

3.1 Step 1: Define Prior Probabilities

- **What**: Initial belief before seeing survey data

- **Why**: Incorporates historical election results

- **How**: Based on past village elections

$$P(\text{BJP}) = 0.6 \quad (\text{BJP won 60\% of similar elections})$$
$$P(\text{Congress}) = 0.4 \quad (\text{Congress won 40\%})$$

3.2 Step 2: Gather New Evidence (Likelihood)

- **What**: Survey of 20 randomly selected voters

- **Results**: 12 support BJP, 8 support Congress

- **Why**: Measures how probable the survey is under each hypothesis

Likelihood calculations:

$$P(\text{Survey}|\text{BJP}) = \binom{20}{12}(0.6)^{12}(0.4)^{8} \quad (\text{If BJP truly has 60\% support})$$

$$P(\text{Survey}|\text{Congress}) = \binom{20}{12}(0.4)^{12}(0.6)^{8} \quad (\text{If Congress truly has 40\% support})$$

3.3 Step 3: Apply Bayes' Theorem

- **What**: Combine prior and likelihood

- **Why**: To update our beliefs with new evidence

- **How**: Compute posterior probabilities

First, calculate normalizing constant:

$$P(\text{Survey}) = (0.18 \times 0.6) + (0.12 \times 0.4) = 0.156$$

Then update beliefs:

$$P(\text{BJP}|\text{Survey}) = \frac{0.18 \times 0.6}{0.156} \approx 0.73$$
$$P(\text{Congress}|\text{Survey}) = \frac{0.12 \times 0.4}{0.156} \approx 0.27$$

3.4 Step 4: Extrapolate to Full Village

- **What**: Predict actual election outcome

- **Why**: Survey sample (20) $\neq$ Village size (250)

- **How**: Model votes as binomial distributions

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

Probability BJP wins (BJP votes > Congress votes):

$$P(X_{\text{BJP}} \geq 126) \quad (\text{since } 126 > 125)$$

Using normal approximation:

$$\mu = 250 \times 0.6 = 150$$
$$\sigma = \sqrt{250 \times 0.6 \times 0.4} \approx 7.75$$
$$P(\text{BJP wins}) \approx P\left(Z \geq \frac{126 - 150}{7.75}\right) = P(Z \geq -3.1) \approx 0.999$$

4 Interpretation of Results

	BJP	Congress
Probability of being preferred	73%	27%
Probability of winning election	¿99%	¡1%

Table 1: Key Differences in Prediction Types

4.1   Why Two Different Probabilities?

- **73% vs ¿99%**: The first measures which party *voters prefer* based on survey, the second measures which party will *win the election* given voting mechanics

- **Crucial insight**: Even a modest preference (60-40) leads to near-certain victory in large populations

5   Limitations and Improvements

5.1   Potential Biases

- Sampling bias: Survey may not represent village

- Prior misspecification: Historical data may not apply

- Non-random voting: Social influences not modeled

5.2   Enhancement Strategies

Method	Benefit
Larger survey sample	Reduces margin of error
Hierarchical modeling	Accounts for village differences
Sensitivity analysis	Tests prior assumption robustness