Beta-Binomial Model Summary

1. Purpose of the Model

The Beta-Binomial model is a Bayesian approach for estimating proportions or probabilities ($\theta \in [0,1]$). It solves problems involving:

- Estimating success probability in Bernoulli trials
- Quantifying uncertainty about proportions
- Making predictions about future outcomes
- Updating beliefs as new data arrives

Example problems:

- Estimating coin fairness after observing flips
- Predicting click-through rates in marketing
- Assessing treatment effectiveness in clinical trials

2. Prior and Likelihood Definition

Prior Distribution

The prior is a Beta distribution: `p(θ) = Beta(θ|a₀, b₀)`

$$p(\theta|a_0, b_0) = \frac{1}{B(a_0,b_0)}\theta^{a_0-1}(1-\theta)^{b_0-1}$$

where:

- `a₀, b₀ > 0` are hyperparameters
- `B(a₀,b₀)` is the Beta function (normalization constant)

Common prior choices:

- Uniform prior: `a₀ = b₀ = 1`
- Jeffreys prior: `a₀ = b₀ = 0.5`
- Weakly informative: `a₀ = b₀ = 2`

Likelihood

The likelihood is a Binomial distribution: `p(y|θ) = Binomial(y|N, θ)`

$p(y|\theta) = \binom{N}{y}\theta^y(1-\theta)^{N-y}$

where:

- `N` = number of trials
- `y` = number of successes
- `θ` = probability of success

3. Posterior Distribution Computation

The posterior is computed analytically using Bayes' theorem:

$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$

For the Beta-Binomial model, the posterior is also a Beta distribution (conjugacy):

$p(\theta|y) = \text{Beta}(\theta|a_0 + y, b_0 + N - y)$

Key property: The Beta distribution is the conjugate prior for the Binomial likelihood.

Posterior Parameters:

- Updated alpha: `a_n = a₀ + y`
- Updated beta: `b_n = b₀ + N - y`

4. Posterior Predictive Distribution

The posterior predictive distribution for a new observation y* given N* new trials:

$p(y^*|y) = \int p(y^*|\theta)p(\theta|y)d\theta$

This yields the Beta-Binomial distribution:

$p(y^*|y) = \binom{N^*}{y^*}\frac{B(y^*+a_n, N^*-y^*+b_n)}{B(a_n, b_n)}$

where `a_n = a₀ + y` and `b_n = b₀ + N - y`.

5. Making Predictions with the Model

Several approaches for point predictions:

Posterior Mean (Expected value)

$\hat{\theta}_{\text{mean}} = \mathbb{E}[\theta|y] = \frac{a_0+y}{a_0+b_0+N}$

Posterior Mode (MAP estimate)

For `a_n, b_n > 1`: $\hat{\theta}_{\text{mode}} = \frac{a_0+y-1}{a_0+b_0+N-2}$

Credible Intervals

For uncertainty quantification, use quantiles of Beta(a_n, b_n).

Model Evidence

The marginal likelihood (model evidence) can be computed analytically:

$$p(y) = \int p(y|\theta)p(\theta)d\theta = \binom{N}{y}\frac{B(y+a_0,N-y+b_0)}{B(a_0,b_0)}$$

Summary

The Beta-Binomial model provides:

1. **Analytically tractable inference** due to conjugacy
2. **Uncertainty quantification** through posterior distribution
3. **Sequential updating** as new data arrives
4. **Flexible prior specification** through a_0, b_0 parameters

This makes it ideal for proportion estimation problems where we want both point estimates and uncertainty measures.