# Complex Beta-Binomial Models: Summary

## 1. Hierarchical Beta-Binomial Model

### Purpose

The hierarchical Beta-Binomial model addresses situations where we have multiple groups, and we believe these groups share common underlying parameters. It's particularly useful when:

- Data comes from multiple sources (schools, hospitals, regions)
- We want to "borrow strength" across groups
- Individual group sample sizes are small

### Model Structure

```
Level 1 (Hyperpriors): a ~ Gamma(α_a, β_a), b ~ Gamma(α_b, β_b)
Level 2 (Group parameters): θ_i ~ Beta(a, b) for each group i
Level 3 (Observations): y_i ~ Binomial(n_i, θ_i)
```

### Key Benefits

1. **Shrinkage**: Pulls extreme group estimates toward the population mean
2. **Uncertainty quantification**: Accounts for both within-group and between-group variation
3. **Small sample robustness**: Helps stabilize estimates for groups with few observations

### When to Use

- Multiple similar units (schools testing same curriculum)
- Suspected common underlying distribution
- Want to make population-level inferences

## 2. Beta-Binomial Regression

### Purpose

Beta-Binomial regression extends the basic model to include covariates that explain variation in success probabilities. It handles overdispersion better than standard logistic regression.

### Model Structure

```
Link function: logit(μ_i) = X_i @ β
Mean parameterization: θ_i ~ Beta(μ_i × ϕ, (1 - μ_i) × ϕ)
Observations: y_i ~ Binomial(n_i, θ_i)
```

Where:

- `X_i` are covariates for observation i
- `β` are regression coefficients
- `ϕ` is a dispersion parameter
- `μ_i` is the expected probability for observation i

### Key Benefits

1. **Covariate effects**: Quantifies how predictors affect success probability
2. **Overdispersion handling**: Better than logistic regression when variance exceeds binomial
3. **Interaction modeling**: Can include interaction terms

### When to Use

- Success probability depends on measured factors
- Data shows more variation than standard binomial
- Need to control for confounders

## 3. Mixture of Beta-Binomials

### Purpose

Mixture models handle populations with distinct subgroups that have different success probabilities. They can identify latent clusters in the data.

### Model Structure

```
Mixture weights: π_k (sum to 1)
Component parameters: θ ~ Σ_k π_k × Beta(a_k, b_k)
Observations: y ~ Binomial(n, θ)
```

### Estimation via EM Algorithm

1. **E-step**: Compute probability each observation belongs to each component
2. **M-step**: Update component parameters and mixture weights

3. **Iterate**: Until convergence

### Key Benefits

1. **Heterogeneity modeling**: Captures distinct subpopulations
2. **Automatic clustering**: Identifies groups without labels
3. **Flexible distributions**: Can approximate complex shapes

### When to Use

- Suspect multiple distinct groups in population
- Histogram shows multimodal distribution
- Standard models show poor fit

## Model Selection Guidelines

### Choosing Between Models

1. **Use Simple Beta-Binomial when**:
   - Single homogeneous population
   - No covariates of interest
   - Adequate fit to data
2. **Use Hierarchical when**:
   - Multiple related groups
   - Want population-level inference
   - Groups have small sample sizes
3. **Use Regression when**:
   - Have explanatory variables
   - Need to quantify covariate effects
   - Data shows overdispersion
4. **Use Mixture when**:
   - Multiple distinct subpopulations suspected
   - Poor fit with single distribution
   - Interest in identifying clusters

### Model Comparison Tools

- **Log-likelihood**: Higher is better

- **AIC**: 2k - 2ln(L), lower is better
- **BIC**: k×ln(n) - 2ln(L), lower is better
- **Cross-validation**: Out-of-sample prediction

## Implementation Tips

### Numerical Stability

- Work in log-space when possible
- Use bounded optimization for parameters
- Initialize parameters sensibly

### Convergence Diagnostics

- Monitor log-likelihood trajectory
- Check parameter stability
- Validate results with different initializations

### Practical Considerations

1. **Sample size**: Hierarchical models need multiple groups
2. **Identifiability**: Mixture models may have label switching
3. **Interpretability**: Simpler models often more interpretable
4. **Computational cost**: Mixtures most expensive, simple least

## Example Applications

1. **Hierarchical**: School performance across districts, clinical trials across hospitals
2. **Regression**: Marketing campaign effectiveness, dose-response studies
3. **Mixture**: Customer segmentation, disease subtypes

## Summary

These complex Beta-Binomial models extend the basic framework to handle:

- Multiple groups (hierarchical)
- Covariate effects (regression)
- Population heterogeneity (mixture)

Choose based on your data structure and research questions. Start simple and add complexity as needed.