The Hierarchical Bayesian Inference Model

Kathrine Kuszon og Sophia Wilson

09/03/2023

When do we need the hierarchical Bayesian inference model?

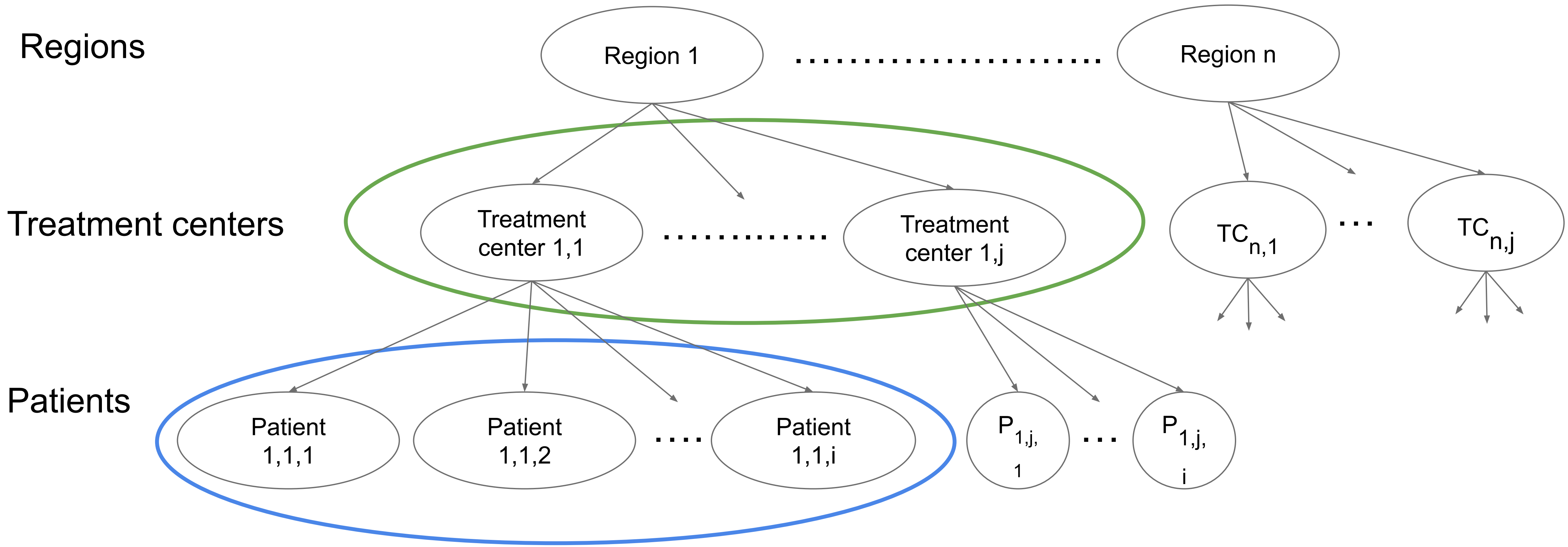
→ Hierarchical data set

When do we need the hierarchical Bayesian inference model?

→ Hierarchical data set

CLINICAL TRIAL

Patients clustered within treatment centers nested within regions



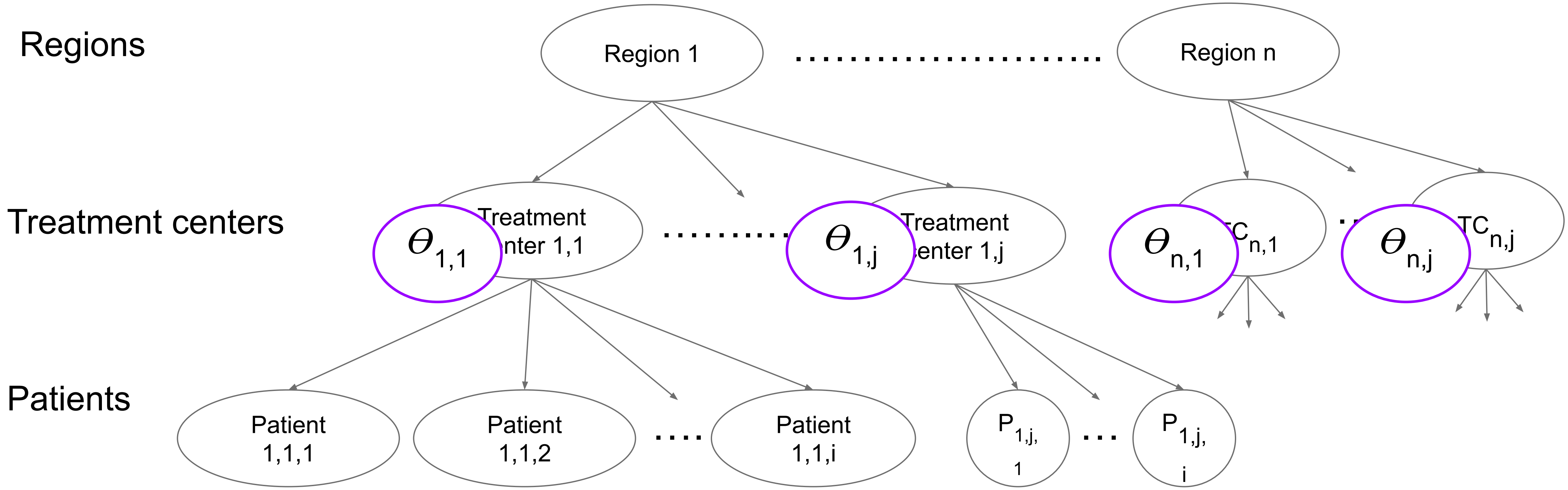
When do we need the hierarchical Bayesian inference model?

→ Hierarchical data set

Parameter of interest:
Survival probability $\theta_{n,j}$

CLINICAL TRIAL

Patients clustered within treatment centers nested within regions



How does the HBI model differ from 'ordinary' Bayesian statistics?

'Ordinary' Bayes

$$\underbrace{P(\theta|y)}_{\text{Posterior}} \propto \underbrace{P(y|\theta)}_{\text{Likelihood}}\underbrace{P(\theta)}_{\text{Prior}}$$

$$y|\theta \sim P(y|\theta)$$

$$\theta \sim P(\theta)$$

How does the HBI model differ from 'ordinary' Bayesian statistics?

'Ordinary' Bayes

$$\underbrace{P(\theta|y)}_{\text{Posterior}} \propto \underbrace{P(y|\theta)}_{\text{Likelihood}}\underbrace{P(\theta)}_{\text{Prior}}$$

$$y|\theta \sim P(y|\theta)$$
$$\theta \sim P(\theta)$$

HBI model

$$y|\theta,\phi \sim P(y|\theta,\phi)$$
$$\theta|\phi \sim P(\theta|\phi)$$
$$\phi \sim P(\phi)$$

How does the HBI model differ from 'ordinary' Bayesian statistics?

'Ordinary' Bayes

$$\underbrace{P(\theta|y)}_{\text{Posterior}} \propto \underbrace{P(y|\theta)}_{\text{Likelihood}}\underbrace{P(\theta)}_{\text{Prior}}$$

$$y|\theta \sim P(y|\theta)$$

$$\theta \sim P(\theta)$$

HBI model

$$\underbrace{P(\phi, \theta|y)}_{\text{Posterior}} \propto \underbrace{P(y|\theta, \phi)}_{\text{Likelihood}}\underbrace{P(\theta|\phi)}_{\text{Prior}}\underbrace{P(\phi)}_{\text{Hyperprior}}$$

$$y|\theta, \phi \sim P(y|\theta, \phi)$$

$$\theta|\phi \sim P(\theta|\phi)$$

$$\phi \sim P(\phi)$$

Testing the HBI model

→ Creating a hierarchical data set

Testing the HBI model

→ Creating a hierarchical data set

- 75 groups ($j \in \{1,75\}$)
- 5 observations ($i \in \{1,5\}$) in each group

Testing the HBI model

→ Creating a hierarchical data set

Data

$$y_{ij}|\alpha_j, \beta_j \sim N(\alpha_j + \beta_j x_{ij}, \sigma^2)$$

- 75 groups (j ∈ {1,75})
- 5 observations (i ∈ {1,5}) in each group

Testing the HBI model

→ Creating a hierarchical data set

- 75 groups (j ∈ {1,75})
- 5 observations (i ∈ {1,5}) in each group

Data

$$y_{ij}|\alpha_j, \beta_j \sim N(\alpha_j + \beta_j x_{ij}, \sigma^2)$$

Parameters

$$\alpha_j|\mu_\alpha, \sigma_\alpha \sim N(\mu_\alpha, \sigma_\alpha^2)$$

$$\beta_j|\mu_\beta, \sigma_\beta \sim N(\mu_\beta, \sigma_\beta^2)$$

$$\sigma \sim U(0, 10)$$

Testing the HBI model

→ Creating a hierarchical data set

- 75 groups (j ∈ {1,75})
- 5 observations (i ∈ {1,5}) in each group

Data

$$y_{ij}|\alpha_j, \beta_j \sim N(\alpha_j + \beta_j x_{ij}, \sigma^2)$$

Parameters

$$\alpha_j|\mu_\alpha, \sigma_\alpha \sim N(\mu_\alpha, \sigma_\alpha^2)$$

$$\beta_j|\mu_\beta, \sigma_\beta \sim N(\mu_\beta, \sigma_\beta^2)$$

$$\sigma \sim U(0, 10)$$

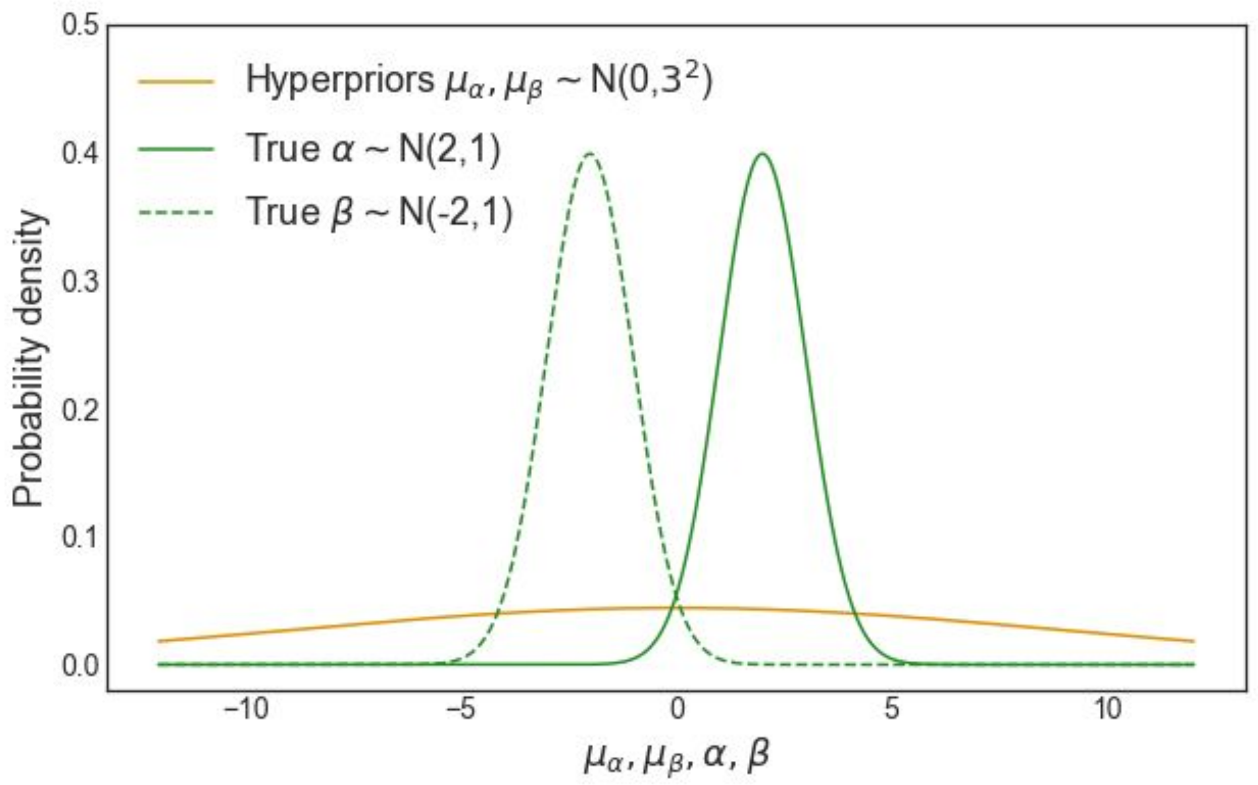
Hyperparameters

$$\mu_\alpha, \mu_\beta \sim N(0, 3^2)$$

$$\sigma_\alpha, \sigma_\beta \sim U(0, 10)$$

Testing the HBI model

→ Creating a hierarchical data set



Data

$$y_{ij}|\alpha_j, \beta_j \sim N(\alpha_j + \beta_j x_{ij}, \sigma^2)$$

Parameters

$$\alpha_j|\mu_\alpha, \sigma_\alpha \sim N(\mu_\alpha, \sigma_\alpha^2)$$

$$\beta_j|\mu_\beta, \sigma_\beta \sim N(\mu_\beta, \sigma_\beta^2)$$

$$\sigma \sim U(0, 10)$$

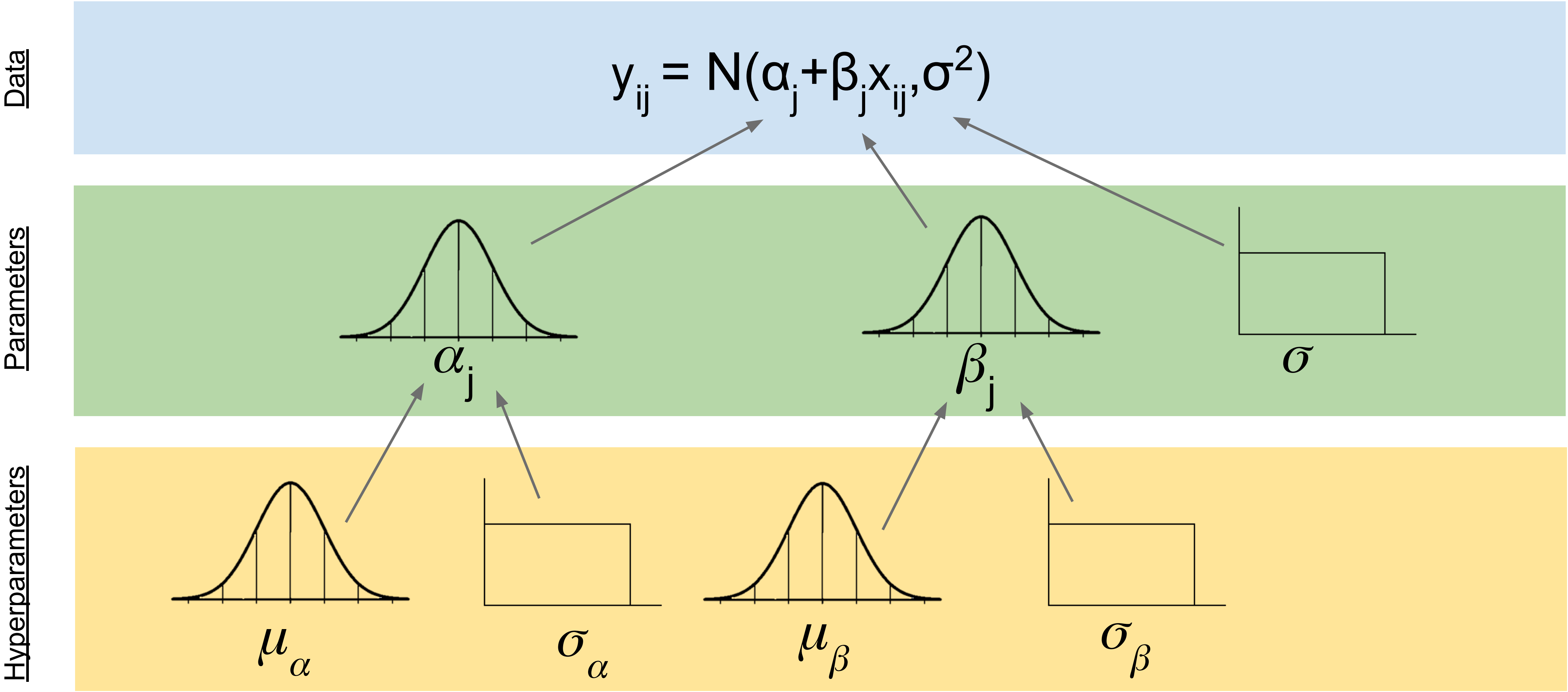
Hyperparameters

$$\mu_\alpha, \mu_\beta \sim N(0, 3^2)$$

$$\sigma_\alpha, \sigma_\beta \sim U(0, 10)$$

Testing the HBI model

→ Creating a hierarchical data set



Comparing the HBI model to other models

→ Overview of alternative models

Comparing the HBI model to other models

→ Overview of alternative models

1) Completely pooling

↓

Assume that α and □ are equal for all groups.

↓

Corresponds to running a single regression on the whole data set

Comparing the HBI model to other models

→ Overview of alternative models

1) Completely pooling

↓

Assume that α and are equal for all groups.

↓

Corresponds to running a single regression on the whole data set

2) No pooling

↓

Assign each α and their own wide, flat prior.

↓

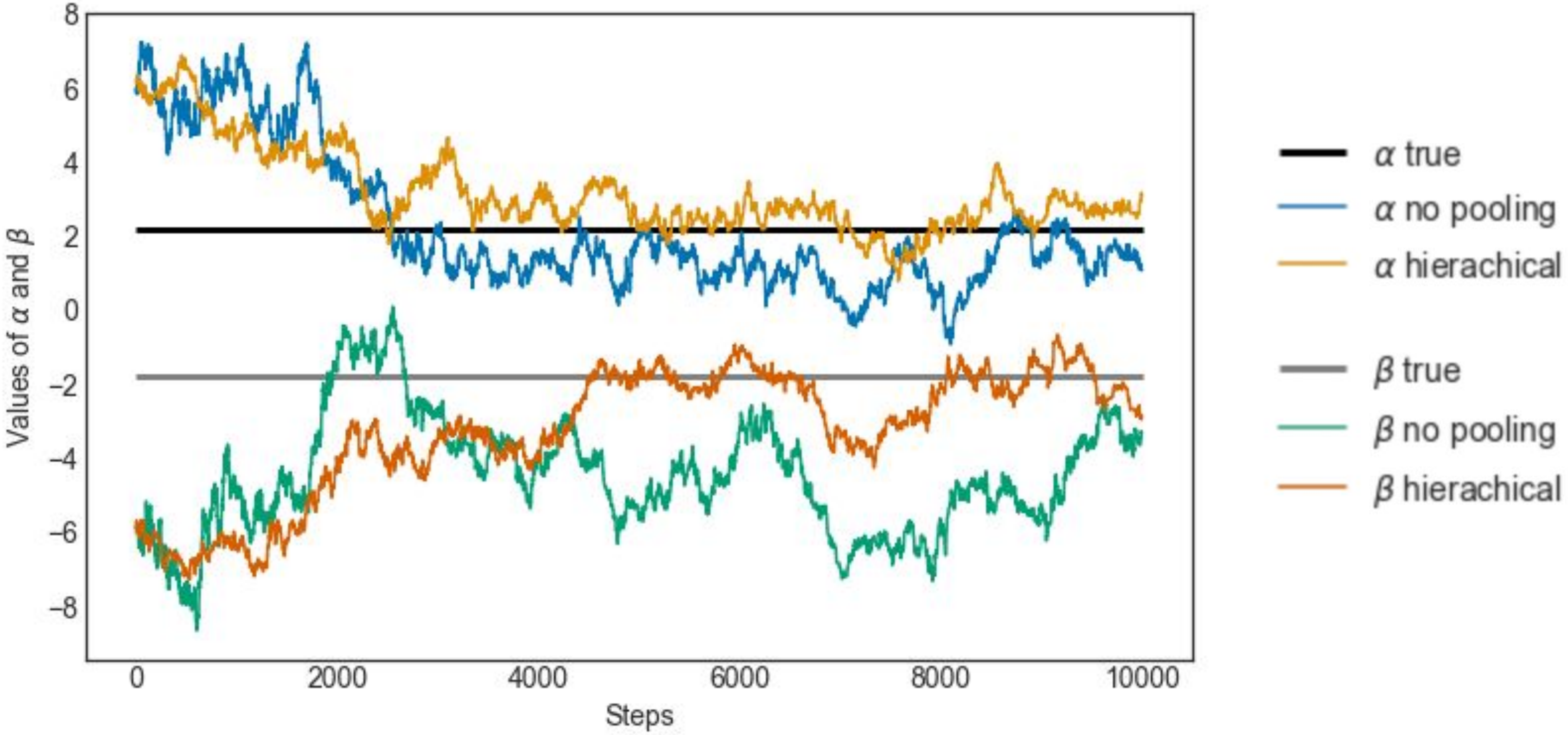
Corresponds to running a single regression on each group.

Comparing the HBI model to other models

→ Markov Chain Metropolis-Hastings sampler: Obtaining estimates of α and □

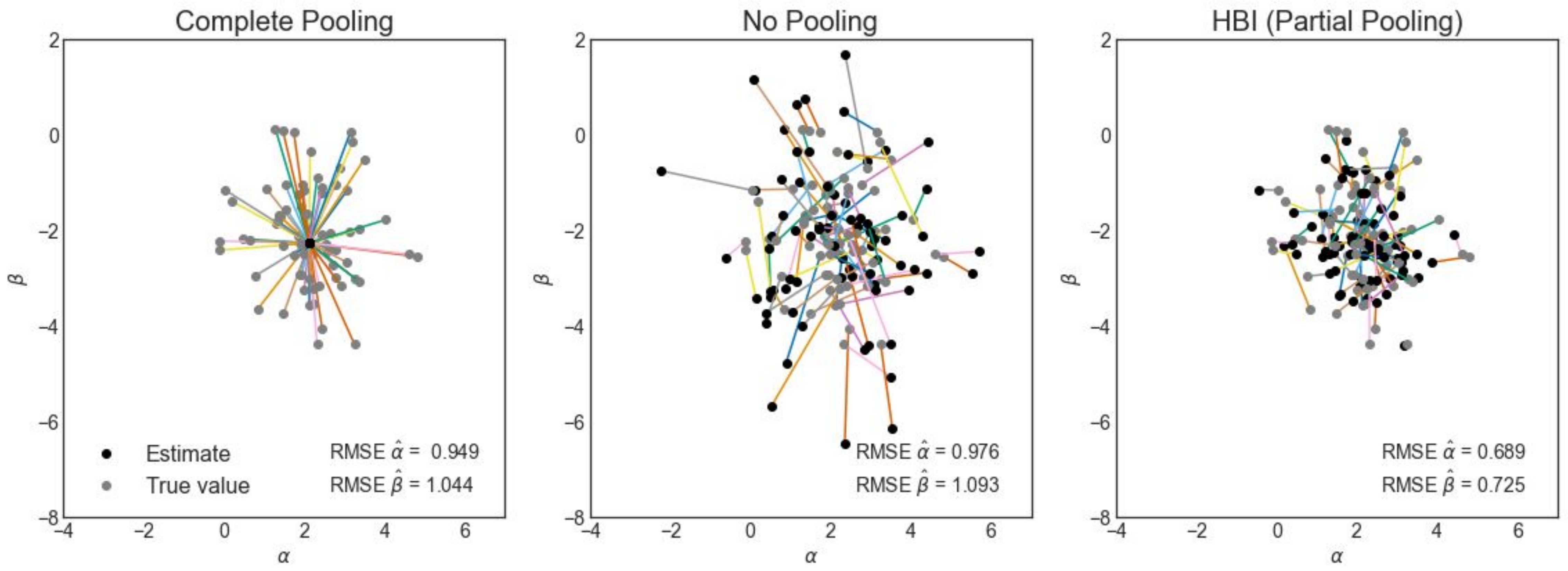
Comparing the HBI model to other models

→ Markov Chain Metropolis-Hastings sampler: Obtaining estimates of α and



Comparing the HBI model to other models

→ Comparing estimates of α and □ to the true values



Summarizing: Hierarchical Bayesian Inference Model

*The HBI model provides a flexible framework for statistical modeling that can capture **variability across groups** and **improve the accuracy and precision of the parameter estimates***

Summarizing: Hierarchical Bayesian Inference Model

*The HBI model provides a flexible framework for statistical modeling that can capture **variability across groups** and **improve the accuracy and precision of the parameter estimates***

→ Allows sharing of statistical strengths between the groups of data by assuming parameters come from common distributions

Summarizing: Hierarchical Bayesian Inference Model

*The HBI model provides a flexible framework for statistical modeling that can capture **variability across groups** and **improve the accuracy and precision of the parameter estimates***

→ Allows sharing of statistical strengths between the groups of data by assuming parameters come from common distributions

→ Lets the data shape the prior itself by introducing the hyperprior

Summarizing: Hierarchical Bayesian Inference Model

*The HBI model provides a flexible framework for statistical modeling that can capture **variability across groups** and **improve the accuracy and precision of the parameter estimates***

- → Allows sharing of statistical strengths between the groups of data by assuming parameters come from common distributions
- → The prior is affected by the data itself by introducing the hyperprior
- → Estimates are less sensitive to noise as the prior structure pulls the estimates towards the population distribution (shrinkage)

Thank you for you attention :)