Multiple nested scales

Multilevel model, with 3 nested levels

Observed value for house i

$$y_i \sim \text{Normal}\left(\mu_{j[i]}, \sigma_h^2\right)$$

House-level stochastic model

Variance among houses within counties

Mean among houses within county j

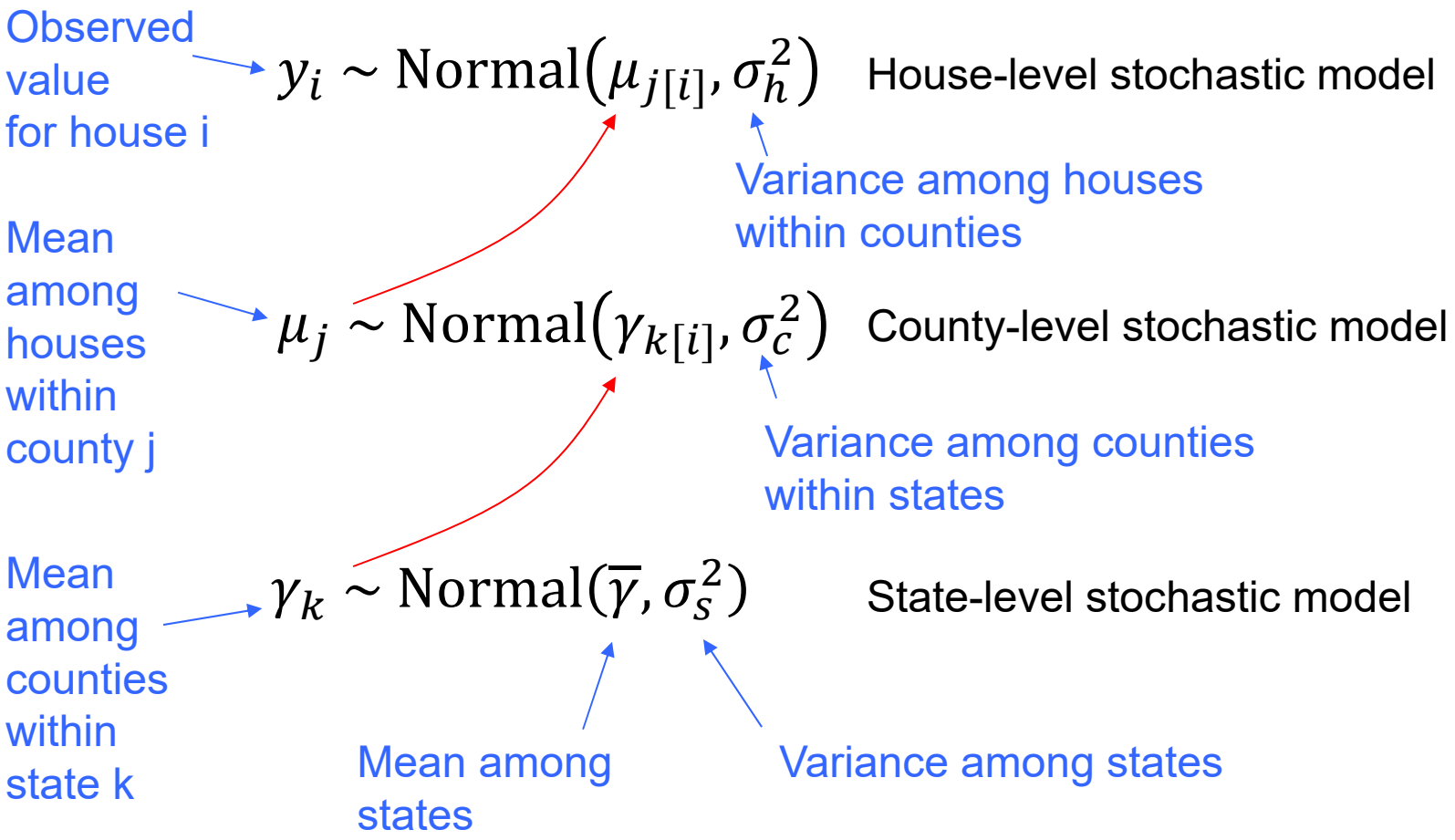
$$\mu_j \sim \text{Normal}(\overline{\mu}, \sigma_c^2)$$

County-level stochastic model

Variance among counties within states

Multiple nested scales

Multilevel model, with 3 nested levels



Multiple nested scales

Additive parameterization

$$y_i = \beta_0 + s_{k[i]} + c_{j[i]} + e_i$$

Residual: deviation of house i from county j mean

Overall mean among states

Deviation of state k mean from overall mean

Deviation of county j mean from state k mean

$$s_k \sim \text{Normal}(0, \sigma_s^2)$$

Variance among states

$$c_j \sim \text{Normal}(0, \sigma_c^2)$$

Variance among counties within states

$$e_i \sim \text{Normal}(0, \sigma_e^2)$$

Variance among houses within counties

Multiple nested scales

Linear model syntax for lmer

If states and counties uniquely identified in data:

```
y ~ 1 + (1|state) + (1|county)
```

Otherwise you can form unique identifiers with the nested operator:

```
y ~ 1 + (1|state/county)
```

which expands to (you can also use this):

```
y ~ 1 + (1|state) + (1|state:county)
```

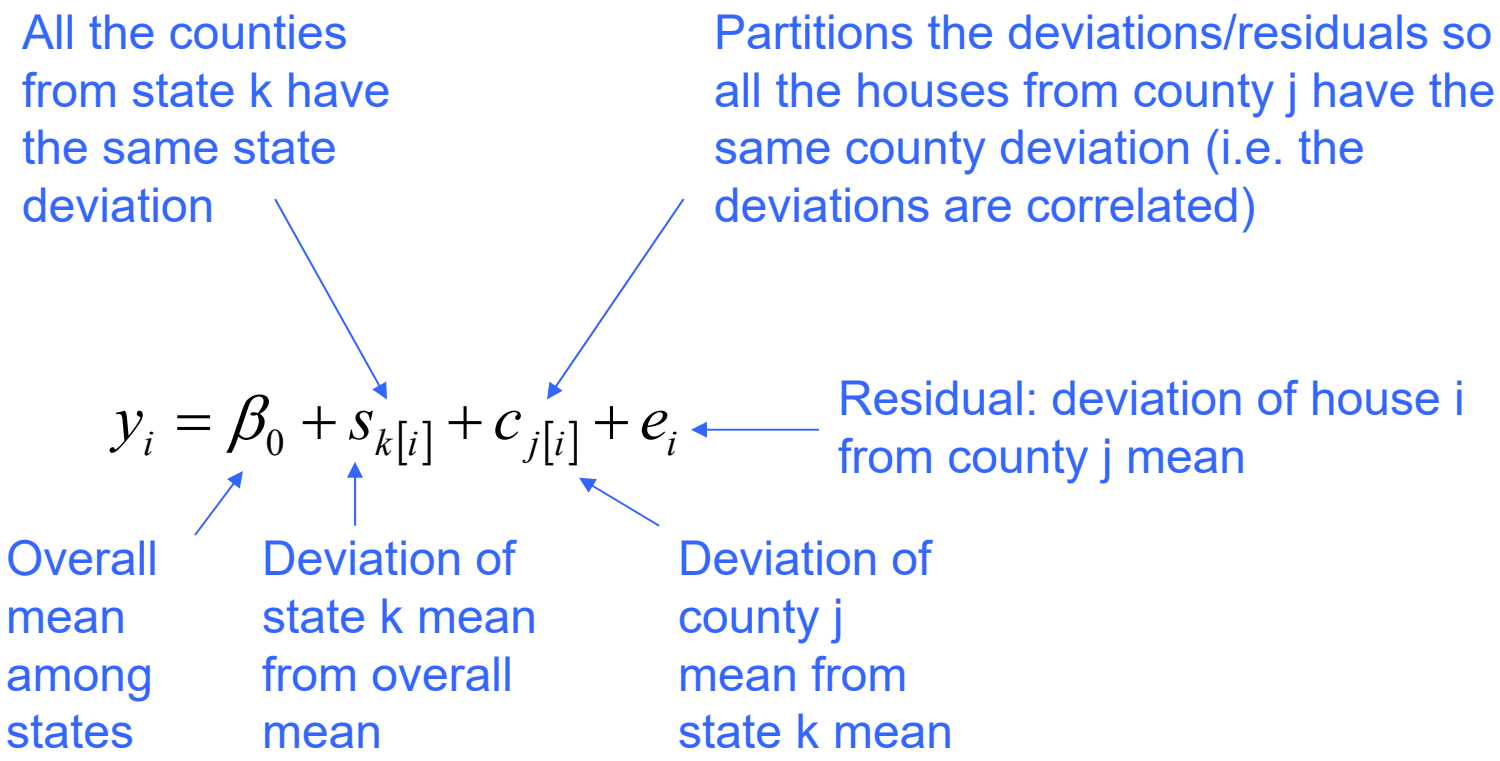
thus, the interaction `state:county` uniquely identifies all counties in the dataset

Autocorrelation

- When present: overoptimistic about the amount of information in the data
- The grouping terms account for autocorrelation (units within groups are more correlated than between groups)
- Specifically, their deviations are correlated. The estimated group deviations (i.e. "random effects") properly partition the correlation into scales of variation

Autocorrelation

How the multilevel model accounts for autocorrelation



LMMs

- Linear mixed models
  - aka linear mixed effects models
- Class of multilevel models
- Mixed = fixed + random effects
  - aka deterministic + stochastic effects
  - aka no pooling + partial pooling
- Fixed effects are linear
- Random effects are all Normal

GLMMs

- Generalized linear mixed models
- Fixed effects have linear predictor but link function can be nonlinear (e.g. log, logit)
- Data level: various distributions
  - e.g. Poisson, Binomial, Exponential, Normal
  - traditionally: exponential family but others work just fine
- Upper levels: Normal distribution