A few Bayesian lectures for the Uninitiated

—

Part 5: a mixed-effect model example

Carsten F. Dormann

Biometry & Environmental System Analysis, University of Freiburg

2. April 2022

Mixed effect model

- Mixed effect model $\in$ hierarchical models
- hierarchy: some parameters describe the response y, others internal, higher-level *unobservables*
- PISA school study: all measurements are at the level of the individual pupil, but there are parameters for schools, states and nations in the model
- mixed effect: subjects are random draws from the population of interest $\longrightarrow$ *subject-level* effects, and *population-level* effects

Mixed effect model and its hierarchies

$y = X\beta + Zu + \varepsilon$ (random intercept model)

Mixed effect model and its hierarchies

$y = X\beta + Zu + \varepsilon$ (random intercept model)
$y_i = X_i\beta + Zu_? + \varepsilon_i$

Mixed effect model and its hierarchies

$y = X\beta + Zu + \varepsilon$ (random intercept model)
$y_i = X_i\beta + Zu_? + \varepsilon_i$
$y_i = X_i\beta + Zu_{j_i} + \varepsilon_i$!!

Mixed effect model and its hierarchies

$y = X\beta + Zu + \varepsilon$ (random intercept model)
$y_i = X_i\beta + Zu_? + \varepsilon_i$
$y_i = X_i\beta + Zu_{j_i} + \varepsilon_i$!!
In words: the ith data point belongs to a specific jth random effect.

Mixed effect model and its hierarchies

$y = X\beta + Zu + \varepsilon$ (random intercept model)
$y_i = X_i\beta + Zu_? + \varepsilon_i$
$y_i = X_i\beta + Zu_{j_i} + \varepsilon_i$!!
In words: the ith data point belongs to a specific jth random effect.

Example: 10 subject were measured 20 times. $j \in \{1, \ldots, 10\}$, while $i \in \{1, \ldots, 200\}$.

Z is in this case a dummy matrix with 10 columns, and each column has one estimate for its subject u_j.

One final step

The likelihood of y_i is based on the estimate of population-level β, but each subject gets an additional "offset" u.

One final step

The likelihood of y_i is based on the estimate of population-level β, but each subject gets an additional "offset" u.
These "offsets" u are computed relative to their mean. That mean is the actual regression intercept β_0.

One final step

The likelihood of y_i is based on the estimate of population-level β, but each subject gets an additional "offset" u.
These "offsets" u are computed relative to their mean. That mean is the actual regression intercept β_0.
That means, we use the individual-level intercepts u_j to estimate the population-level intercept $\bar{u}$.

One final step

The likelihood of y_i is based on the estimate of population-level β, but each subject gets an additional "offset" u.
These "offsets" u are computed relative to their mean. That mean is the actual regression intercept β_0.
That means, we use the individual-level intercepts u_j to estimate the population-level intercept $\bar{u}$.
That is the parameter at the other hierarchical level!

One final step

The likelihood of y_i is based on the estimate of population-level β, but each subject gets an additional "offset" u.
These "offsets" u are computed relative to their mean. That mean is the actual regression intercept β_0.
That means, we use the individual-level intercepts u_j to estimate the population-level intercept $\bar{u}$.
That is the parameter at the other hierarchical level!
Let's write this as equations:

A random intercept mixed effect model

$$y_i \sim N(\mu = \hat{y}_i, \sigma = \hat{s})$$

A random intercept mixed effect model

$$
\begin{aligned}
y_i &\sim N(\mu = \hat{y}_i, \sigma = \hat{s}) \\
\hat{y}_i &= X_i \beta + Z_i u_{j_i}
\end{aligned}
$$

A random intercept mixed effect model

$$
\begin{aligned}
&y_i \sim N(\mu = \hat{y}_i, \sigma = \hat{s}) \\
&\hat{y}_i = X_i\beta + Z_i u_{j_i} \\
&u_j \sim N(\mu = \hat{u}, \sigma = \hat{s_u})
\end{aligned}
$$

A random intercept mixed effect model

$$
\begin{aligned}
&y_i \sim N(\mu = \hat{y}_i, \sigma = \hat{s}) \\
&\hat{y}_i = X_i\beta + Z_i u_{j_i} \\
&u_j \sim N(\mu = \hat{u}, \sigma = \hat{s_u}) \\
&\hat{u} = c
\end{aligned}
$$

A random intercept mixed effect model

$$
\begin{aligned}
&y_i \sim N(\mu = \hat{y}_i, \sigma = \hat{s}) \\
&\hat{y}_i = X_i\beta + Z_i u_{j_i} \\
&u_j \sim N(\mu = \hat{u}, \sigma = \hat{s_u}) \\
&\hat{u} = c
\end{aligned}
$$

Final touch: Bayesian here only means that you have to have a prior for the βs, the σs, the zs and indeed for c.

In JAGS

```
for (i in 1:N){
    Reaction[i] ~ dnorm(muHat[i], tau)
    # random intercept model:
    muHat[i] <- beta0[subjectNr[i]] + beta1 * Days[i]
}
# priors
for (j in 1:k){
    beta0[j] ~ dnorm(beta00, tau.random) # random intercept
}
beta00 ~ dnorm(300, 0.001)          # grand mean
tau ~ dgamma(0.001, 0.001)
tau.random ~ dgamma(0.001, 0.001)
```

Will this never end?

Next time:
State-space models