Categorical: Repeated measures

Table of contents

1 Goals **2**
1.1 Goals 2

2 Review **2**
2.1 Linear mixed models 2

3 Data example **5**
3.1 Data example 5

4 Marginal model **9**
4.1 Marginal model 9
4.2 Example 10
4.3 Marginal wrap-up 12

5 Conditional model **13**
5.1 Conditional model 13
5.2 Example 14
5.3 Conditional wrap-up 16

6 Quick comparison **17**
6.1 Quick comparison 17

7 Summary **18**
7.1 Summary 18

1 Goals

1.1 Goals

1.1.1 Goals of this lecture

- Mixed model for **categorical** outcomes
  - Marginal model: **R** matrix, GEE
  - Conditional model: **G** matrix, GLMM

2 Review

2.1 Linear mixed models

2.1.1 Repeated measures = non-independence

- Repeated measures from the same person are **not independent**
  - An observation from a person provides information about other observations from that person
  - Observations from the same person are more like one another than observations from different people
  - Observations from the same person are **correlated**

2.1.2 Linear mixed models

$$\mathbf{Y}_{ij} = \mathbf{X}\beta + \mathbf{Z} \ + \epsilon$$

- Very general model
- Allows for repeated measures via
  - Random effects: **Z**
  - Correlated residuals: ϵ

2.1.3 Linear mixed model: Marginal approach

$$\mathbf{Y}_{ij} = \mathbf{X}\beta + \mathbf{Z} \ + \epsilon$$

- Fixed effects: $\mathbf{X}\beta$
  - $\mathbf{X}$ is a matrix of the predictors
  - β are regression coefficients

2.1.4 Linear mixed model: Marginal approach

$$\mathbf{Y}_{ij} = \mathbf{X}\beta + \mathbf{Z} \quad + \epsilon$$

- Random effects: $\mathbf{Z}$
  - No random effects
    - No predictors in $\mathbf{Z}$
    - This term drops out

2.1.5 Linear mixed model: Marginal approach

$$\mathbf{Y}_{ij} = \mathbf{X}\beta + \mathbf{Z} \quad + \epsilon$$

- Residuals: ϵ
  - Residuals are correlated / covary across timepoints
    - $t \times t$ matrix

2.1.6 Linear mixed model: Conditional approach

$$\mathbf{Y}_{ij} = \mathbf{X}\beta + \mathbf{Z} \quad + \epsilon$$

- Fixed effects: $\mathbf{X}\beta$
  - $\mathbf{X}$ is a matrix the predictors
  - β are regression coefficients

2.1.7 Linear mixed model: Conditional approach

$$\mathbf{Y}_{ij} = \mathbf{X}\beta + \mathbf{Z} \quad + \epsilon$$

- Random effects: $\mathbf{Z}$
  - $\mathbf{Z}$ is a matrix of random effects predictors (dummy codes)
    - Which observations go with which subject
  - is a variance-covariance matrix of random effects
    - Intercept variance, slope variance, intercept-slope covariance

2.1.8 Linear mixed model: Conditional approach

$$\mathbf{Y}_{ij} = \mathbf{X}\beta + \mathbf{Z} \ + \epsilon$$

- Residuals: ϵ
  - Residuals aren't correlated
  - Single residual variance value
    - Variance of all residuals across all people and timepoints

2.1.9 Marginal vs conditional

- Deal with non-independence in different ways: Different interpretation
  - *Marginal*: Effect based on all observations, then adjust standard errors
  - *Conditional*: Effect for each subject, then average across subjects
- For linear models, marginal and conditional don't differ numerically
  - Average of linear function = linear function of average

2.1.10 Marginal vs conditional

- For non-linear models, marginal and conditional differ numerically
  - Average of a **non-linear** function $\neq$ **non-linear** function of average
  - *Remember from GLiM*: Non-linear function of predicted value $\neq$ predicted value of non-linear function
- Again, that's ok
  - Marginal and conditional models are actually **answering different questions**

2.1.11 Comparison

- Marginal models
  - Cluster (person) is nuisance
  - Population average
  - Usually repeated measures, also cross-sectional
  - GEE
- Conditional models
  - Cluster (person) is of interest

- Person-specific
- Repeated measures or cross-sectional
- GLMM

3 Data example

3.1 Data example

3.1.1 Schizophrenia over time

- Schizophrenia treatment effects over the course of 7 weeks ($N = 437$), measured by the Inpatient Multidimensional Psychiatric Scale (IMPS)
  - `id`: ID variable
  - `imps79`: Continuous measure of schizophrenia (1 to 7)
  - `imps79b`: Binary measure of schizophrenia (3.5+)
  - `imps79o`: Ordinal measure of schizophrenia (Cuts: 2.5+, 4.5+, 5.5+)
  - `tx`: Placebo (0) or treatment (1)
  - `week`: Week of study (0, 1, 3, 6)

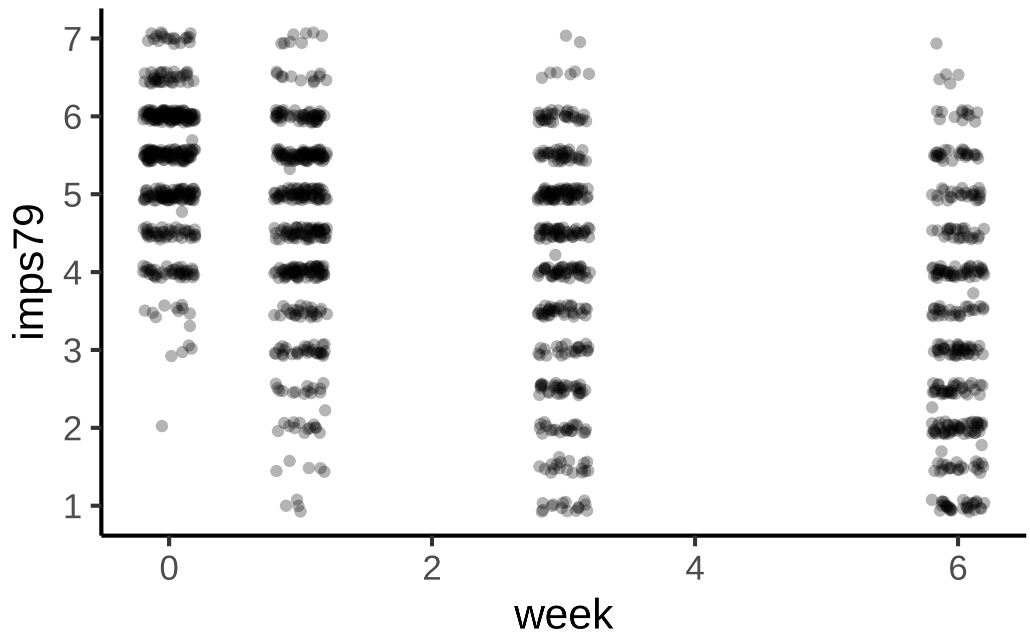
3.1.2 Data

id	imps79	imps79b	imps79o	tx	week
1103	5.5	1	4	1	0
1103	3.0	0	2	1	1
1103	2.5	0	2	1	3
1103	4.0	1	2	1	6
1104	6.0	1	4	1	0
1104	3.0	0	2	1	1
1104	1.5	0	1	1	3
1104	2.5	0	2	1	6
1105	4.0	1	2	1	0
1105	3.0	0	2	1	1
1105	1.0	0	1	1	3
1105	NA	NA	NA	1	6

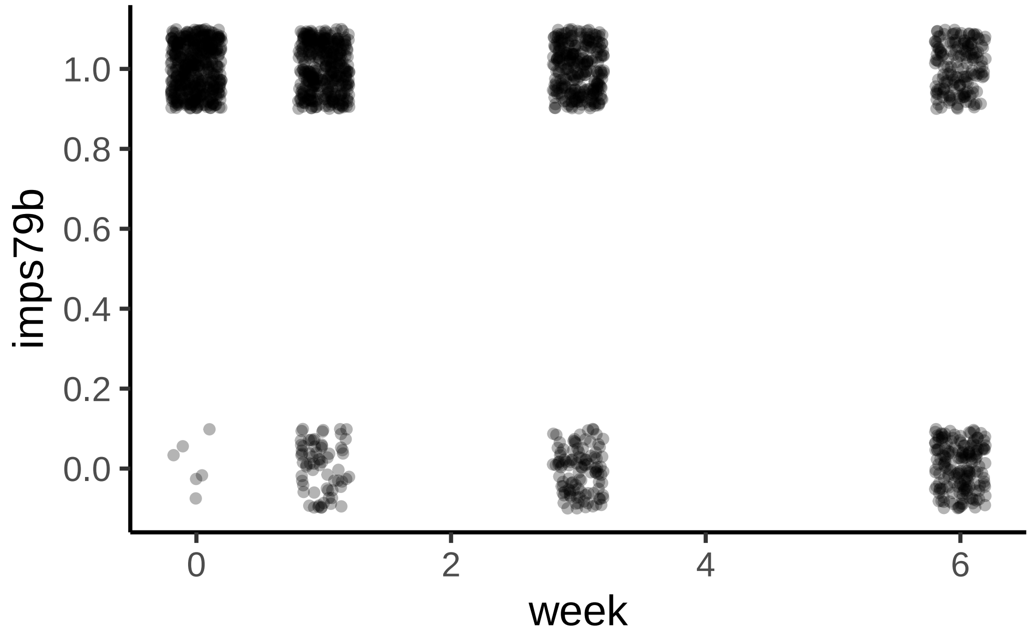
3.1.3 Means and N by week

week	mean_c	mean_b	N
0	5.367	0.986	434
1	4.571	0.843	426
3	4.020	0.711	374
6	3.310	0.484	335

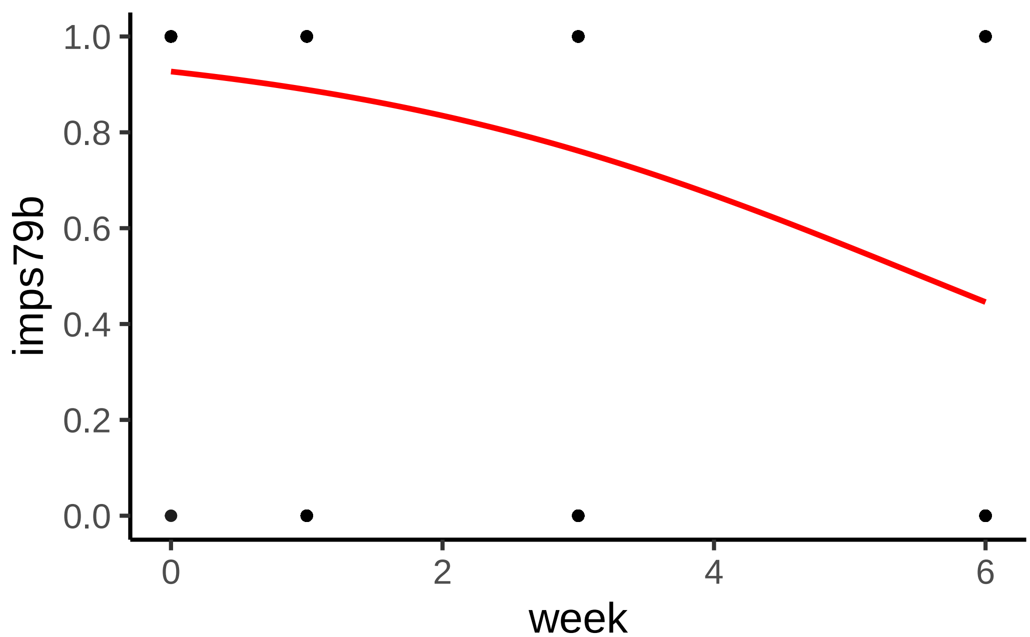
3.1.4 Plot: Continuous measure by week



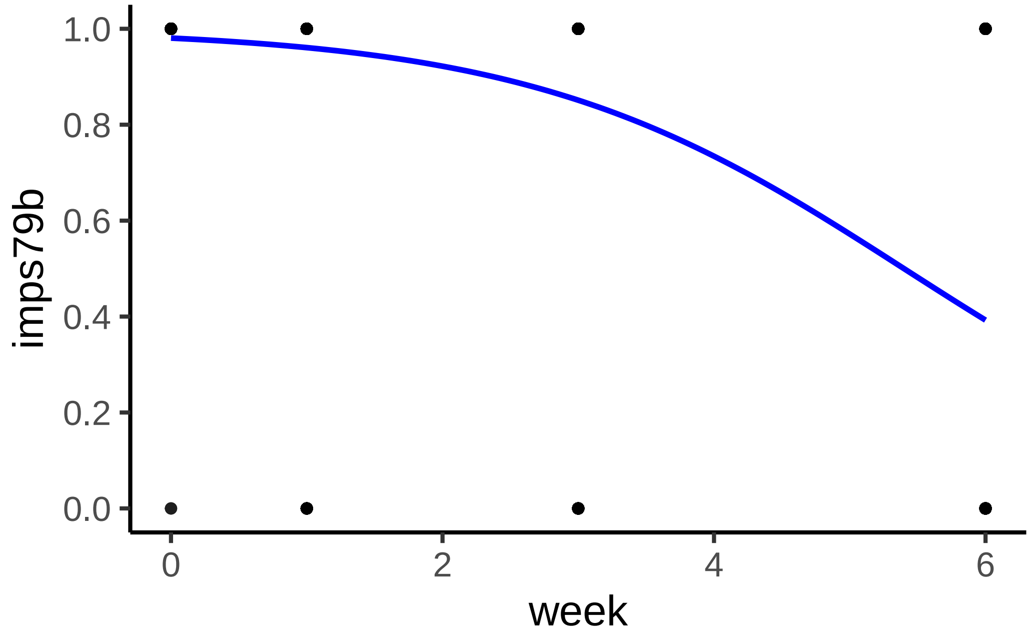
3.1.5 Plot: Binary measure by week



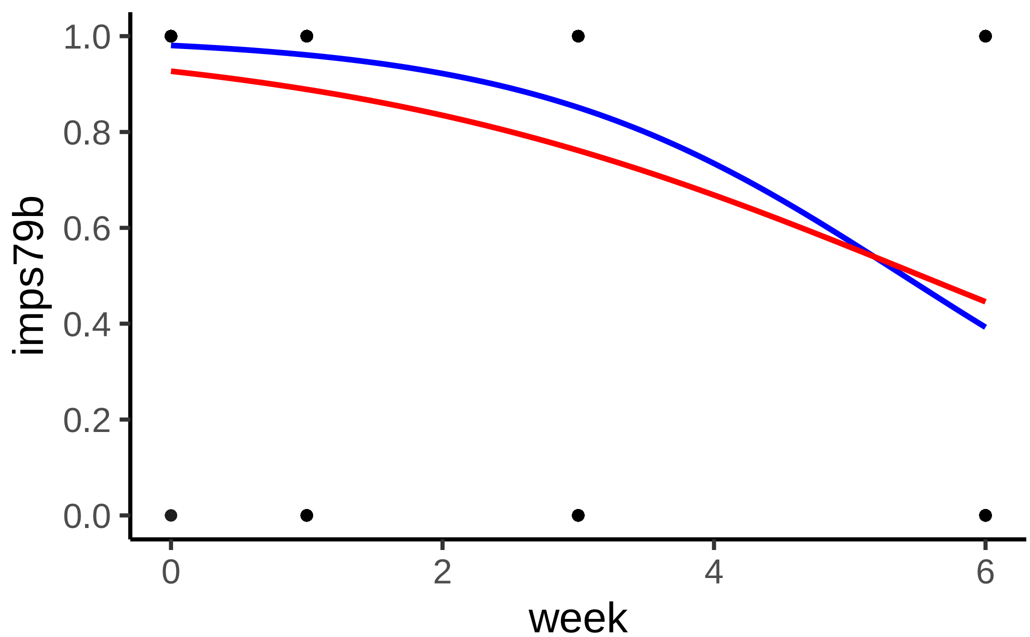
3.1.6 Plot: *Marginal* effect of time



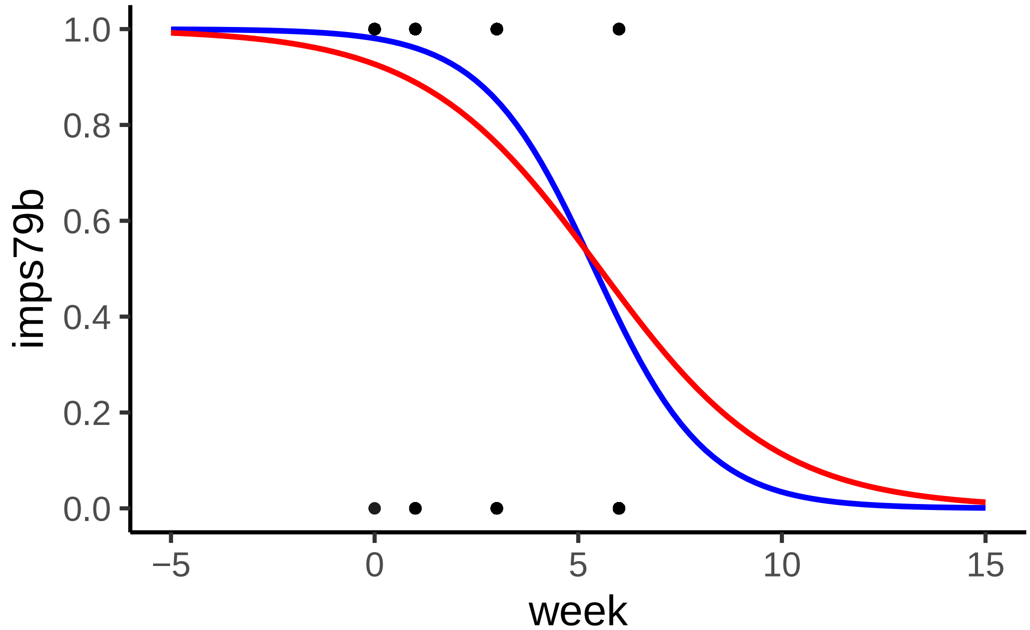
3.1.7 Plot: *Conditional* effect of time



3.1.8 Plot: *Marginal* and *conditional* lines



3.1.9 Plot: $\color{red}{Marginal}$ and $\color{blue}{conditional}$ lines



3.1.10 Overall model

- **Research question: Now**
  1. How does schizophrenia diagnosis **change** over these 7 weeks?
- **Research question: Class**
  1. How does schizophrenia diagnosis **change** over these 7 weeks?
  2. How do the **treatment** groups differ (at baseline)?
  3. Does **change in diagnosis differ** depending on treatment condition?

4 Marginal model

4.1 Marginal model

4.1.1 Marginal model: Generalized estimating equations (GEE)

$$\eta = \mathbf{X}\beta$$

- η: Transformation of predicted value (from GLiM)

- Depends on the specific model (i.e., logistic, Poisson)

- Variance
  - ϵ: Matrix of correlated residuals
    - $t \times t$ matrix: t is number of repeated measures

4.1.2 Marginal model: Generalized estimating equations (GEE)

- Fixed effects (regression coefficients)
  - **Population-averaged effects**
    - Averaging across all observations (ignore people)
  - For "people", not for "a person"
    - Public health application

4.2 Example

4.2.1 Example

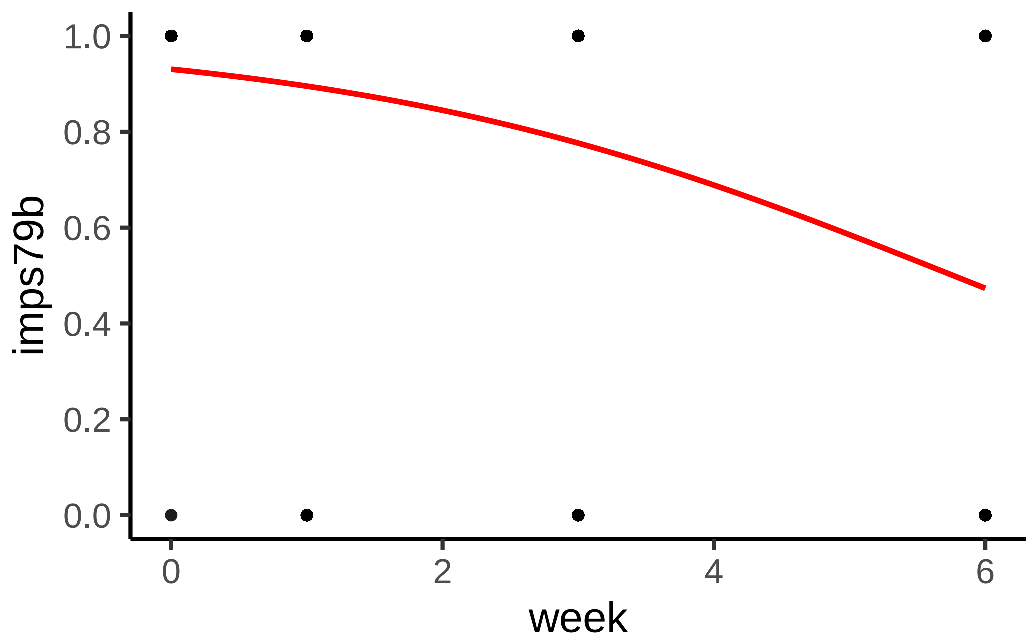
- Time predicts schizophrenia diagnosis
  - `week` as a predictor of `imps79b`
  - `week`: 0, 1, 3, 6
  - `imps79b`: 0 (less than 3.5 on `imps79`), 1 (3.5+ on `imps79`)
- In this example, I'm using **unstructured R** matrix
  - We'll look at others in class
  - We'll also look at treatment effects (`tx`)

4.2.2 Marginal model

```
Call:
geeglm(formula = imps79b ~ 1 + week, family = binomial("logit"),
    data = schizx1, id = schizx1$id, corstr = "unstructured")

 Coefficients:
            Estimate  Std.err  Wald            Pr(>|W|)
(Intercept)  2.59459  0.11876 477.3 <0.0000000000000002 ***
week        -0.45017  0.02767 264.7 <0.0000000000000002 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Correlation structure = unstructured
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)   0.9646  0.1007
  Link = identity

Estimated Correlation Parameters:
          Estimate Std.err
alpha.1:2  0.05390 0.05343
alpha.1:3 -0.02855 0.03265
alpha.1:4 -0.01890 0.03342
alpha.2:3  0.56341 0.10114
alpha.2:4  0.15242 0.06617
alpha.3:4  0.51550 0.07964
Number of clusters:   437  Maximum cluster size: 4
```

4.2.3 Plot: Marginal model



4.3 Marginal wrap-up

4.3.1 Population-averaged effects

$$ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = b_0 + b_1(week)$$

- b_1: Time effect
  - $e^{b_1} = \frac{odds\ of\ event\ at\ week\ t+1}{odds\ of\ event\ at\ week\ t}$
  - Ignoring the repeated measures
    - But standard errors are *adjusted* for non-independence

4.3.2 Population-averaged effects: Pros

- *Robust* to mis-specification of the **R** matrix
  - You use compound symmetry but that's not very close to reality
- Can account for *unobserved or unknown dependence*
  - Things besides repeated measures
- Easier to *estimate* than conditional models
  - Marginal values are readily available

4.3.3 Population-averaged effects: Cons

- Ignores that *individuals* make up these effects
  - Just wants to deal with correlated observations
- Ignores that individuals may have *different patterns* over time
  - Are some individuals *helped a lot* by the treatment? Who knows.
- Documentation for newer R package says it requires complete data
  - Worked fine here so ???
  - You can delete all NA rows, but an additional step

4.3.4 Why is it called "marginal"?

tx	week0	week1	week3	week6
0	0.972	0.880	0.713	0.463
1	0.982	0.802	0.574	0.340

- Estimated using only **marginal** proportions
  - Not joint proportions: Therefore no conditional values either

5 Conditional model

5.1 Conditional model

5.1.1 Conditional model: Generalized linear mixed model

$$\eta = \mathbf{X}\beta$$

- η: Transformation of predicted value (from GLiM)
  - Depends on the specific model (i.e., logistic, Poisson)
- Variance
  - γ: Matrix of random effects
    - Intercept variance, slope variance, intercept-slope covariance
  - ϵ: Residual variance
    - Single number (depends on model: e.g., fixed at $\pi^2/3$ in logistic)

5.1.2 Conditional model as multi-level model

- Two parts ("levels" for multi-level models) of the model
  - Within-person / within-cluster
  - Between-person / between-cluster
- Equations at each level
  - Combine into the full model

5.1.3 Conditional model as multi-level model

- Level 1: Within-person equation
  - $\eta = \pi_{0i} + \pi_{1i}(week) + e_{ij}$
- Level 2: Between-person equation
  - $\pi_{0i} = \beta_{00} + r_{0i}$
  - $\pi_{1i} = \beta_{10} + r_{1i}$
- Combined equation
  - $\eta = \beta_{00} + \beta_{10}(week) + r_{0i} + r_{1i}(week) + e_{ij}$

5.2 Example

5.2.1 Example

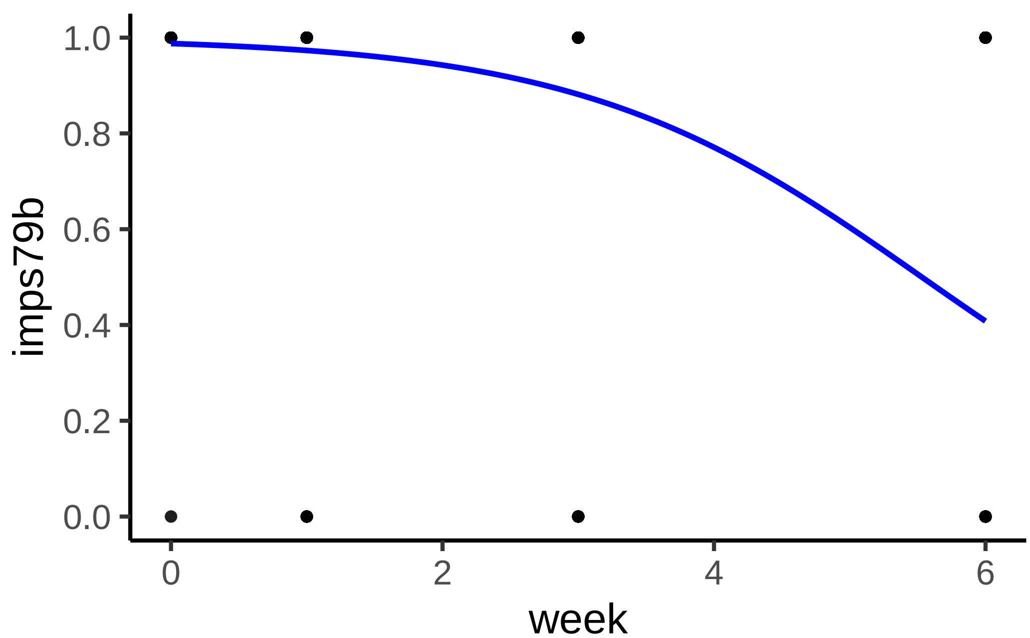
- Time predicts schizophrenia diagnosis
  - `week` as a predictor of `imps79b`
  - `week`: 0, 1, 3, 6
  - `imps79b`: 0 (less than 3.5 on `imps79`), 1 (3.5+ on `imps79`)
- In this example, we have **random intercepts** and **random slopes**
  - We'll also look at treatment effects (`tx`) in class
  - For that model, we'll only be able to use **random intercepts**

5.2.2 Conditional model

```
Generalized linear mixed model fit by maximum likelihood (Laplace
  Approximation) [glmerMod]
 Family: binomial  ( logit )
Formula: imps79b ~ 1 + week + (1 + week | id)
   Data: schizx1

     AIC      BIC   logLik deviance df.resid
  1291.6   1318.4   -640.8   1281.6     1564

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.8446  0.0898  0.1139  0.2646  1.0822
```

```
Random effects:
 Groups Name        Variance Std.Dev. Corr
 id     (Intercept) 4.413    2.101
        week        0.711    0.843    -0.13
Number of obs: 1569, groups:  id, 437

Fixed effects:
            Estimate Std. Error z value            Pr(>|z|)
(Intercept)    4.386      0.539    8.14 0.00000000000000041 ***
week          -0.793      0.118   -6.71 0.00000000001954283 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
     (Intr)
week -0.817
```

5.2.3 Plot: Conditional model



5.3 Conditional wrap-up

5.3.1 Person-specific effects

$$ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = b_0 + b_1(week)$$

- b_1: Time effect
  - $e^{b_1} = \frac{odds\ of\ event\ at\ week\ t+1}{odds\ of\ event\ at\ week\ t}$
  - Estimated **separately** for each person
    * Average individual effects together to get the average effect

5.3.2 Person-specific effects: Pros

- *Individual trajectories* are estimated
  - Not just correlations between repeated measures
- More *flexibility* in individual *variability*
  - Random intercepts and random slopes with respect to time
- Conceptually, fits better with how psychologists think
  - Individuals, trajectories, etc.

5.3.3 Person-specific effects: Cons

- Often harder to **estimate**
  - Much more complex model than marginal model
- Deciding on *random effects* can be difficult
  - Both choosing and estimation
- Accounts for *specific sources of non-independence*
  - Cannot account for e.g., multiple members of same family

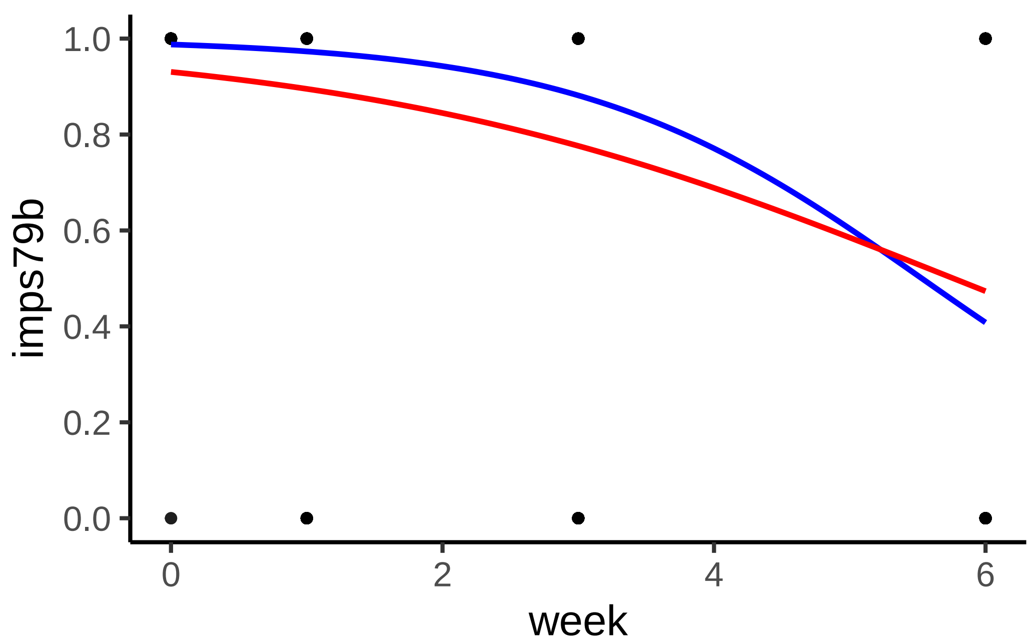
6 Quick comparison

6.1 Quick comparison

6.1.1 Example: Conditional vs marginal effects

- Marginal effect
  - **For people**, one week of time has a certain effect
  - Population level: Interest is the *group of people*
- Conditional effect (GLMM):
  - **For a person**, one week of time has a *different* effect
  - Individual level: Interest is *individuals*

6.1.2 Plot: Comparison of *marginal* and *conditional*



7 Summary

7.1 Summary

7.1.1 Summary of this week

- Extended mixed models to categorical outcomes
  - Marginal: **R** matrix, population averaged, GEE, cluster robust
  - Conditional: **G** matrix, generalized linear mixed models (GLMM)
- Different interpretations, different numbers
  - Population-averaged: Adjusts for non-independence, doesn't care about repeated measures
  - Conditional: Analyzes each person separately and averages

7.1.2 In class

- Look at models including `week`, `tx`, and their product
  - Think about how that more complex model works
  - See some errors you might get

7.1.3 Next week

- ~~All~~ Some of the additional crucial details
  - Estimation issues
    - Maximum likelihood?
    - Tips and tricks to get models to run?
  - Model comparisons
  - Multi-level issues: Adding and centering predictors, contextual effects
  - More on predicted values