Methods 3, Week 5:

Towards Bayesian Multi-Level Modelling

The plan

1. Introduction to the data set
2. Step-by-step Regression Modelling on the project data
    - Linear Regression
    - Bayesian Linear Regression
    - Bayesian Multilevel Linear Regression

Introduction to the research project

- Ca 30 children with ASD and 30 typically developing controls matched by gender, Socio-Economic Status (SES), and language production (starting around 2 yo)
- 6 video-recorded visits every 4 months
    - 30 minutes of controlled playful activities with a parent
        - Missing data points (dropouts, unavailabilities…)
- Videos were transcribed at word level
    - Total words, lexicon size, and MLU automatically assessed for both children and parents

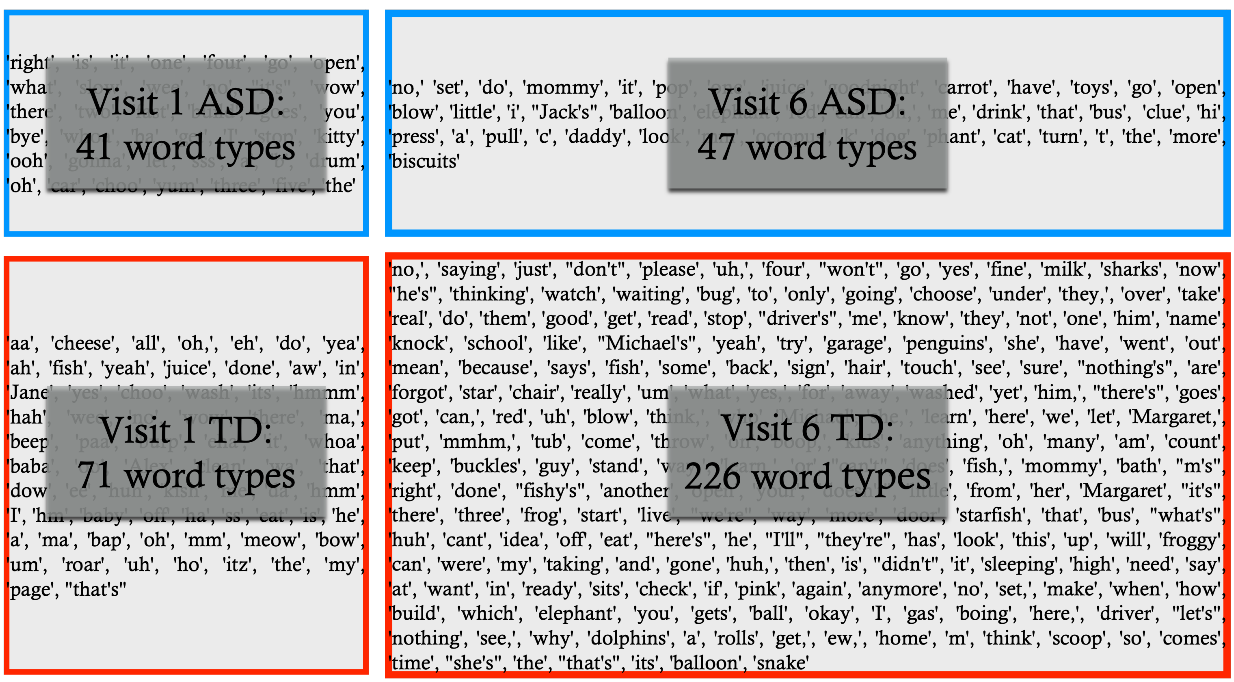
An interaction sample

```
MOT:look at the fun toys bc
MOT:ya, he's already there so
MOT:what, what is that
MOT:where are the kids
MOT:huh, they going to school
CHI:uh huh
MOT:uh huh
CHI:in a bus
MOT:in a bus, good job
MOT:xxx
CHI:uh oh
MOT:that's ok he stays in that, that's a special chair for him
MOT:ok that's a special chair for him
CHI:uh oh
CHI:a bath's broken
MOT:the bath is broken, yeah we'll leave it though
MOT:but you know what, you can still give the baby a bath
MOT:you can pretend
MOT:wanna give her a bath
CHI:eh eh
CHI:how xxx get her dress off
MOT:wanna take her dress off to give her a bath
CHI:uh huh
MOT:can you do it
CHI:uh huh
MOT:ok you try
MOT:ok take her arms out first
CHI:ok
CHI:her arms out first
MOT:now the other arm
MOT:all right give her a bath
CHI:oh
CHI:how to getting this off
CHI:how to getting this off
```

The interaction sample, parsed


```
,Interlocutor,Transcription,Token,Lemma,Tagged,TaggedL
0,MOT,look at the fun toys Bc,"[look, at, the, fun, toys, bc]","[look, at, the, fun, toy, bc]","[(look, VB), (at, IN), (the, DT),
(fun, NN), (toys, NNS), (bc, VBP)]","[(look, VB), (at, IN), (the, DT), (fun, NN), (toy, NN), (bc, NN)]"
1,MOT,"ya, he's already there so","[ya, he, 's, already, there, so]","[ya, he, 's, already, there, so]","[(ya, PRP), (he, PRP), ('s,
VBZ), (already, RB), (there, EX), (so, RB)]","[(ya, PRP), (he, PRP), ('s, VBZ), (already, RB), (there, EX), (so, RB)]"
2,MOT,"what, what is that","[what, what, is, that]","[what, what, be, that]","[(what, WP), (what, WP), (is, VBZ), (that,
IN)]","[(what, WP), (what, WP), (be, VB), (that, IN)]"
3,MOT,where are the kids,"[where, are, the, kids]","[where, be, the, kid]","[(where, WRB), (are, VBP), (the, DT), (kids,
NNS)]","[(where, WRB), (be, VB), (the, DT), (kid, NN)]"
4,MOT,"huh, they going to school","[huh, they, going, to, school]","[huh, they, go, to, school]","[(huh, VB), (they, PRP), (going,
VBG), (to, TO), (school, VB)]","[(huh, VB), (they, PRP), (go, VB), (to, TO), (school, NN)]"
5,CHI,uh huh,[huh],[huh],"[(huh, VB)]","[(huh, VB)]"
6,MOT,uh huh,[huh],[huh],"[(huh, VB)]","[(huh, VB)]"
7,CHI,in a bus,"[in, a, bus]","[in, a, bus]","[(in, IN), (a, DT), (bus, NN)]","[(in, IN), (a, DT), (bus, NN)]"
8,MOT,"in a bus, good job","[in, a, bus, good, job]","[in, a, bus, good, job]","[(in, IN), (a, DT), (bus, NN), (good, JJ), (job,
NN)]","[(in, IN), (a, DT), (bus, NN), (good, JJ), (job, NN)]"
9,MOT,xxx,[xxx],[xxx],"[(xxx, NN)]","[(xxx, NN)]"
10,CHI,uh oh,[oh],[oh],"[(oh, UH)]","[(oh, UH)]"
11,MOT,"that's ok he stays in that, that's a special chair for him","[that, 's, ok, he, stays, in, that, that, 's, a, special, chair,
for, him]","[that, 's, ok, he, stay, in, that, that, 's, a, special, chair, for, him]","[(that, DT), ('s, VBZ), (ok, IN), (he, PRP),
(stays, VBZ), (in, IN), (that, IN), (that, DT), ('s, VBZ), (a, DT), (special, JJ), (chair, NN), (for, IN), (him, PRP)]","[(that, DT),
('s, VBZ), (ok, IN), (he, PRP), (stay, VB), (in, RP), (that, IN), (that, DT), ('s, VBZ), (a, DT), (special, JJ), (chair, NN), (for,
IN), (him, PRP)]"
12,MOT,ok that's a special chair for him,"[ok, that, 's, a, special, chair, for, him]","[ok, that, 's, a, special, chair, for,
him]","[(ok, IN), (that, DT), ('s, VBZ), (a, DT), (special, JJ), (chair, NN), (for, IN), (him, PRP)]","[(ok, IN), (that, DT), ('s,
VBZ), (a, DT), (special, JJ), (chair, NN), (for, IN), (him, PRP)]"
13,CHI,uh oh,[oh],[oh],"[(oh, UH)]","[(oh, UH)]"
14,CHI,a bath's broken,"[a, bath, 's, broken]","[a, bath, 's, broken]","[(a, DT), (bath, NN), ('s, POS), (broken, JJ)]","[(a, DT),
(bath, NN), ('s, POS), (broken, JJ)]"
```


The Lexicon



Linear Regression Models

- Basic linear regression model with a single predictor, a slope β, an intercept α and normally-distributed residuals

$y_i = \alpha + \beta x_i + \epsilon$, with $\epsilon_i \sim Normal(0, \sigma)$

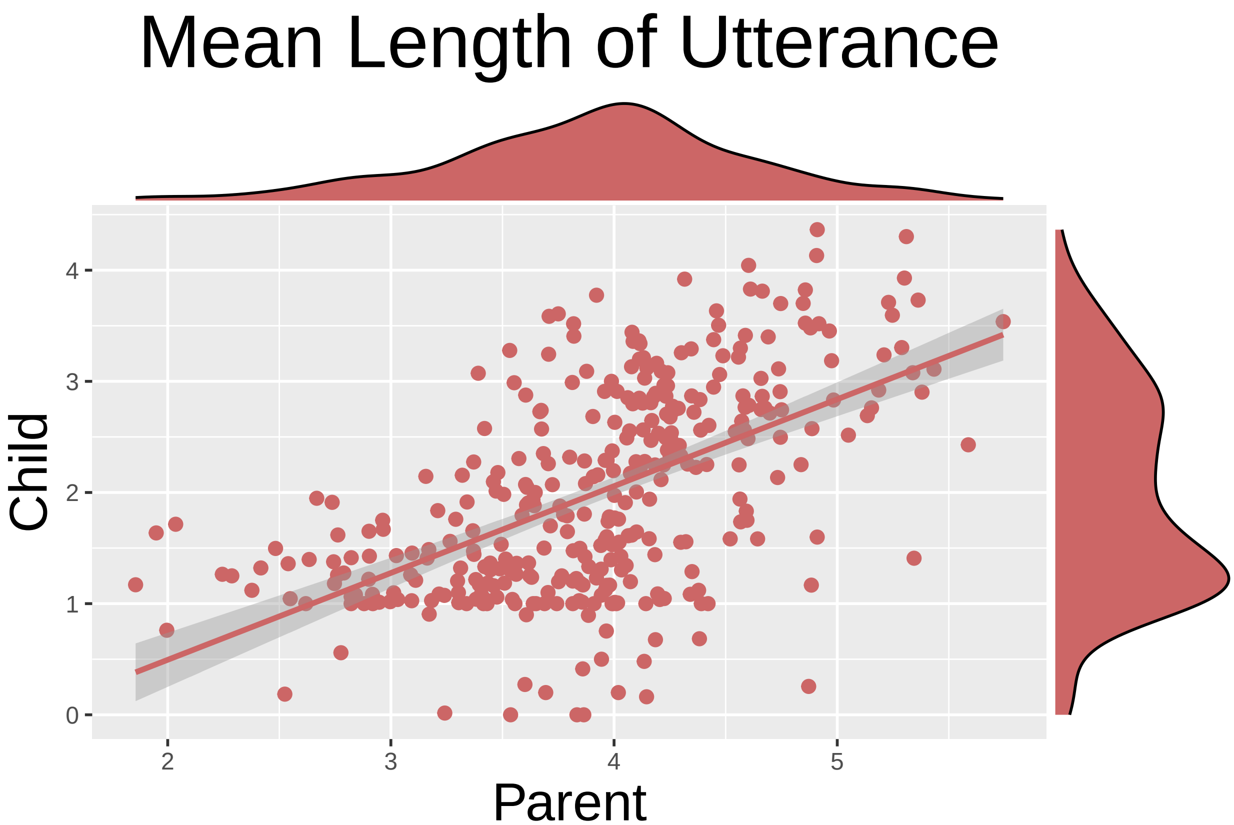
This is equivalent to

$y_i - (\alpha + \beta x_i) \sim Normal(0, \sigma)$

and finally to the more multilevel-friendly form

$y_i \sim Normal(\alpha + \beta x_i, \sigma)$

A Linear Regression Model for MLU



Model Fitting

$MLU(child_i) \sim Normal(\alpha + \beta \cdot MLU(mother_i), \sigma)$

```
mod_utt_td <- lm(formula = CHI_MLU ~ MOT_MLU, data = data)
mod_utt_td$coefficients
```

```
## (Intercept)     MOT_MLU
##  -1.0680205   0.7813712
```

$$MLU(child) = -1.07 + 0.78 \cdot MLU(mother) \quad (1)$$

$$R^2_{adj} = 0.32 \quad (2)$$

- The intercept (-1.1) represents the number of words uttered by the child when the mother says nothing (0)
- The slope (0.78) represents the change in the number of words uttered by the child when the number of words uttered by the mother increases by one

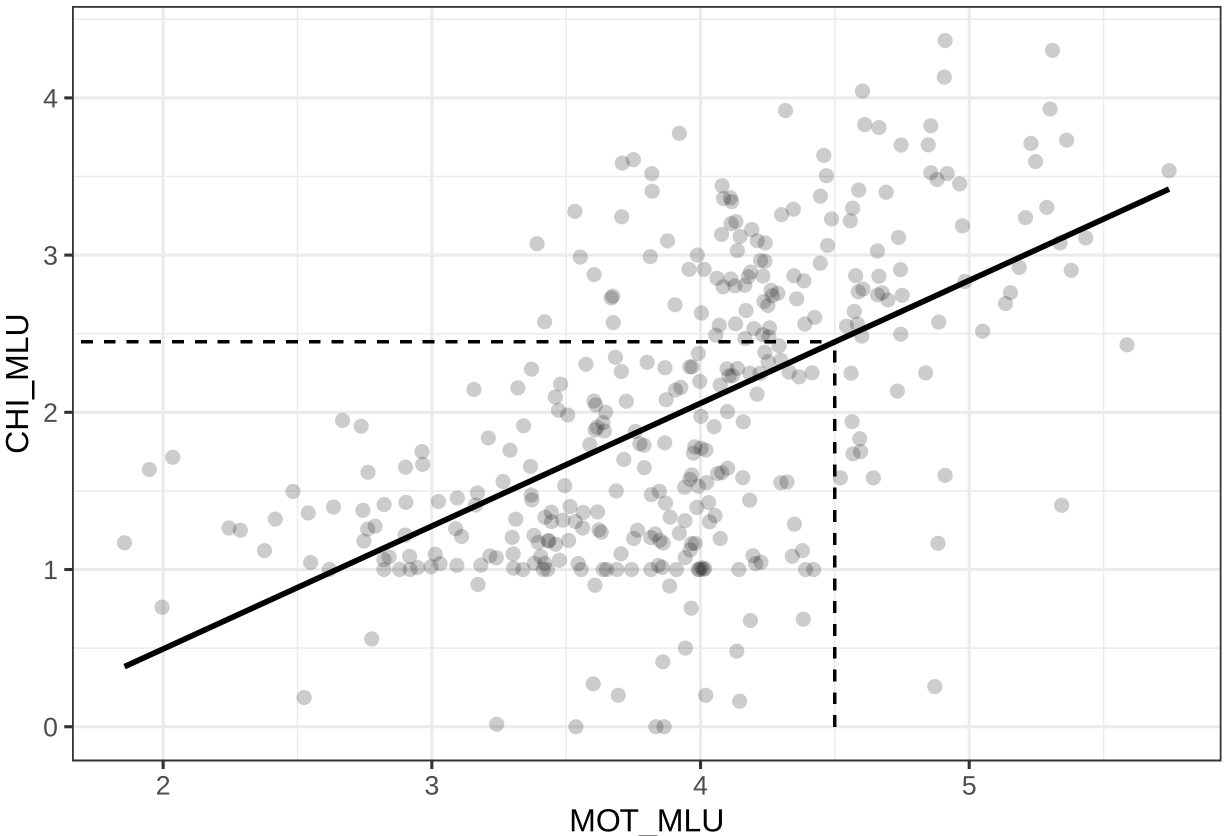
Regression and Prediction

- Linear models allow to make predictions for any given predictor variable value (any given number of words uttered by the parent)
- The function **predict()** retrieves parameters of the fitted model (in our case, the intercept and the slope) to make predictions about the outcome variable, given some values of the predictor(s)

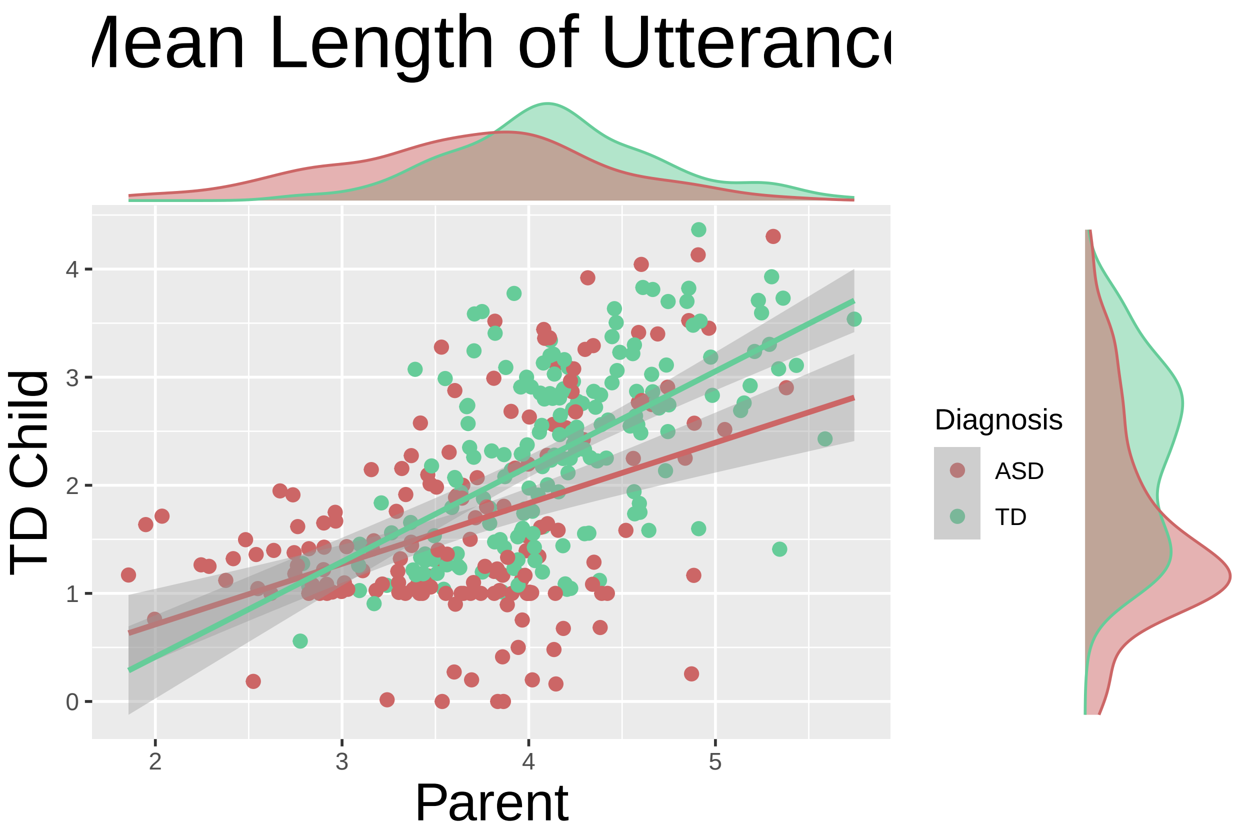
  - Second line of the linear model
- What is MLU_{CHILD}, given that $MLU_{MOT} = 4.5$?

$$
\begin{aligned}
MLU(child) &= -1.07 + 0.78 \cdot MLU(mother) && (3)\\
&= -1.07 + .78 \cdot 4.5 && (4)\\
&= 2.44 && (5)
\end{aligned}
$$

Plotting the prediction



Multiple Regression Models



Multiple Fit and Prediction

```
mod_utt_both <- lm(formula = CHI_MLU ~ MOT_MLU * Diagnosis, data = data)
mod_utt_both$coefficients
```

```
##         (Intercept)             MOT_MLU         DiagnosisTD MOT_MLU:DiagnosisTD
##          -0.4088249           0.5608042          -0.9397872           0.3198611
```

$$MLU(child_{ASD}) = -0.41 + 0.56 \cdot MLU(mother) \tag{6}$$
$$MLU(child_{TD}) = -.41 + .56 \cdot MLU(mother) - .94 + .32 \cdot MLU(mother) \tag{7}$$
$$= -1.35 + .88 \cdot MLU(mother) \tag{8}$$
$$R^2_{adj} = .35 \tag{9}$$

Multiple Regression and Prediction

- What are $MLU(child_{ASD})$ and $MLU(child_{TD})$, given that $MLU(mother) = 4.5$?

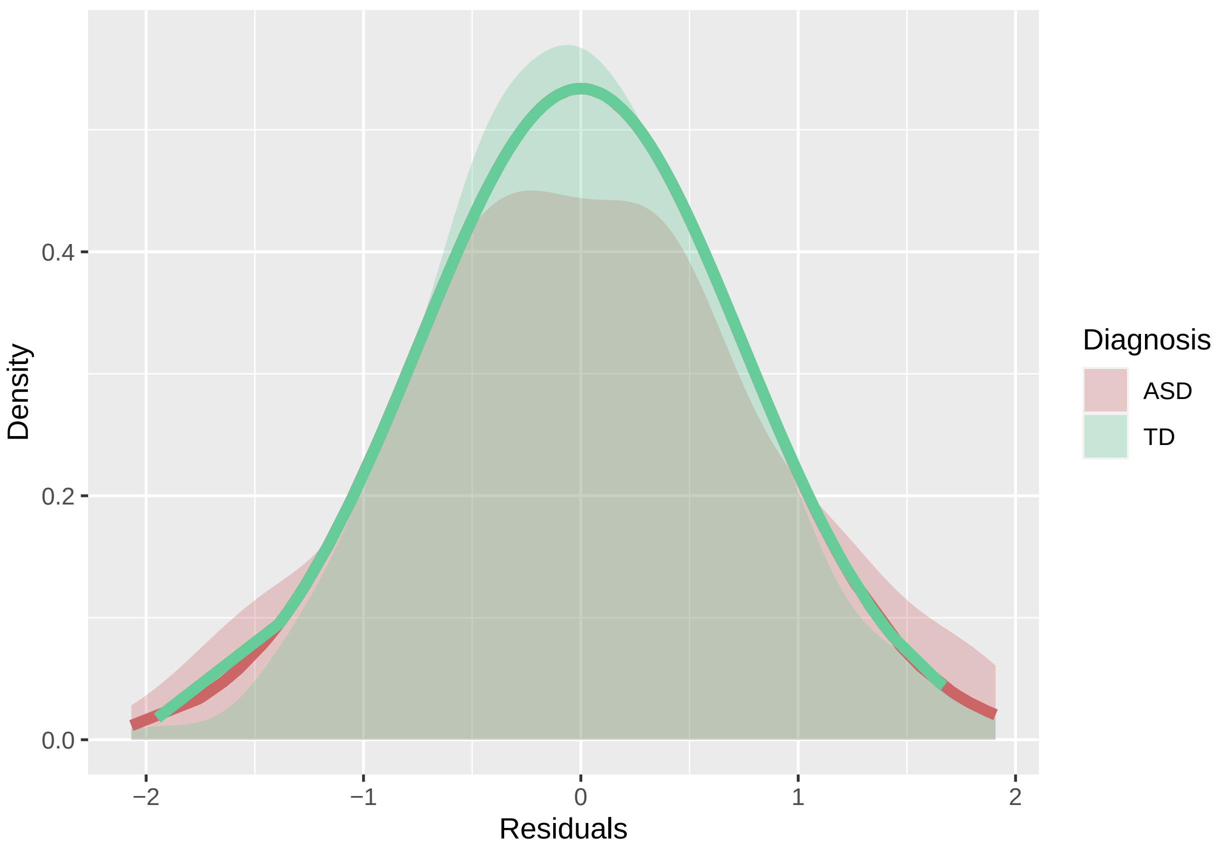
$$MLU(child_{ASD}) = -0.41 + 0.56 \cdot MLU(mother) \tag{10}$$
$$= -0.41 + 0.56 \cdot .45 \tag{11}$$
$$= 2.11 \tag{12}$$

$$
\begin{aligned}
MLU(child_{TD}) &= -1.35 + .88 \cdot MLU(mother) && (13)\\
&= -1.35 + .88 \cdot .45 && (14)\\
&= 2.61 && (15)
\end{aligned}
$$

Residuals

- A common fallacy about the assumptions of the linear (Gaussian) model is that the outcome variable should be normally distributed
  - It is the distribution of the outcome **around its predicted value** that is normally distributed
- Errors are the non-observed differences between the predicted value μ and the observed outcomes.
  - Residuals are an estimate (from the sample) of the errors ϵ_i
- Errors pertain to the data generation process, whereas residuals are the difference between the model's estimation and the observed outcomes

Assessing normality of residuals



Going Bayesian with BRMS

- In classical linear regression, we compute point estimates of our parameters and use these estimates to make predictions
  - **lm**: ordinary least squares (OLS) algorithm
- Bayesian linear regression estimates distributions over the parameters and predictions
  - Allows to model uncertainty in predictions

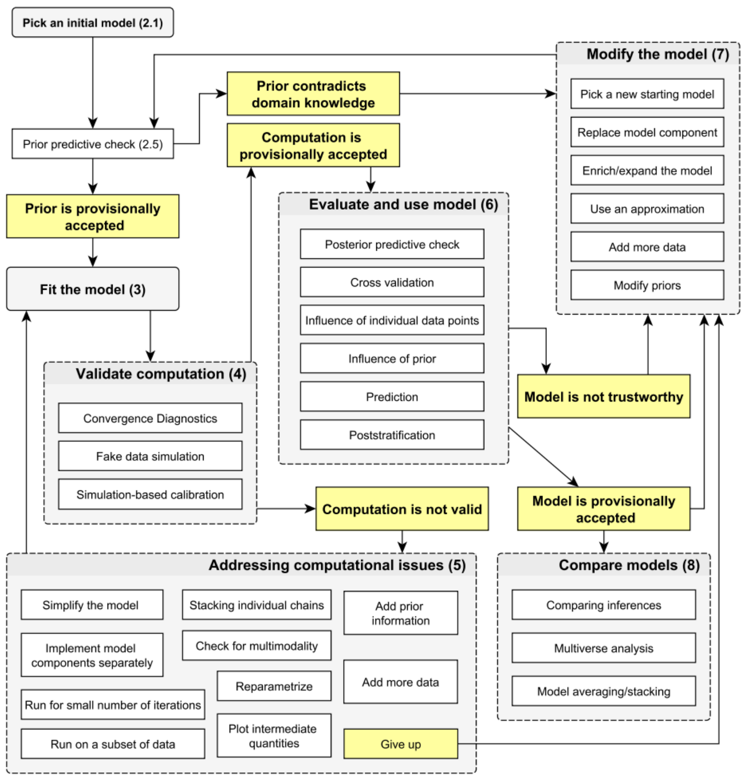
- **brms**: Bayesian Regression Models using STAN (BRMS)
  * Bayesian multilevel linear and non-linear models
  * Attempt at bridging frequentist and Bayesian practices by extending **lme4**'s equation syntax

The general bayesian workflow (Gelman, 2013)

For each statistical problem, we follow three general steps:

1. Model Building: likelihood + priors
2. Model Updating using information contained in data and Bayes theorem (aka approximating the posterior probability)
3. Model Evaluation: fit, assumptions, results summarization, readjusting the model

The bayesian workflow in detail (Gelman, 2020)



Our first model

```r
formula1 <- brms::bf(CHI_MLU ~ MOT_MLU * Diagnosis)
model1 <- brms::brm(formula = formula1,
                    data = data,
                    file = 'data/w5/example')
model1
```

```
##  Family: gaussian
##   Links: mu = identity; sigma = identity
## Formula: CHI_MLU ~ MOT_MLU * Diagnosis
##    Data: data (Number of observations: 352)
##   Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup draws = 4000
```

```
##
## Population-Level Effects:
##                     Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept              -0.40      0.31    -1.02     0.20 1.00     1814     2230
## MOT_MLU                 0.56      0.08     0.40     0.73 1.00     1784     2338
## DiagnosisTD            -0.96      0.53    -2.01     0.08 1.00     1373     1770
## MOT_MLU:DiagnosisTD     0.32      0.13     0.06     0.58 1.00     1352     1812
##
## Family Specific Parameters:
##       Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma     0.75      0.03     0.70     0.81 1.00     2975     2481
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Priors

- By default **brms** uses an uninformative prior based on the average value of the measured variable

```
# This function shows which priors a model fitted with brms has (implicitly) assumed
brms::prior_summary(model1)
```

```
##                   prior     class                coef group resp dpar nlpar lb
##                  (flat)         b
##                  (flat)         b         DiagnosisTD
##                  (flat)         b             MOT_MLU
##                  (flat)         b MOT_MLU:DiagnosisTD
##  student_t(3, 1.9, 2.5) Intercept
##    student_t(3, 0, 2.5)     sigma                                             0
##  ub       source
##          default
##     (vectorized)
##     (vectorized)
##     (vectorized)
##          default
##          default
```