Mixed Effects Model - Nepalese Children
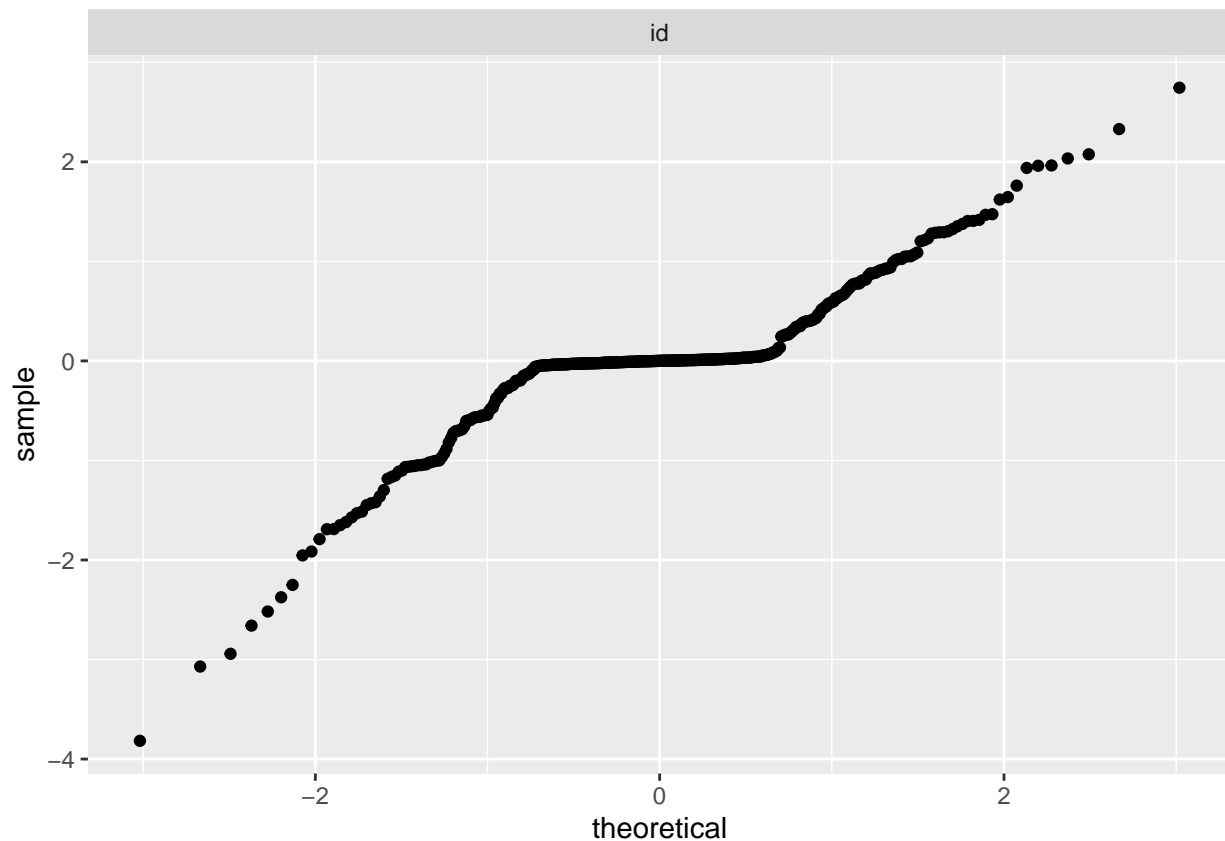
Sivaram Ainkaran

18/09/2021

Question 1

```
lmer1 <- lmer(wt ~ sex + mage + lit + died + alive + age + (age|id),  data=nepali1)
```

Question 2

```
ranef(lmer1) %>%  as_tibble() %>%
  ggplot(aes(sample=condval)) + facet_wrap(~ grpvar) + geom_qq()
```


As we can see in the Q-Q plot above, the random effects on the data peak alot at the centre - shown by the almost straight line around 0. This is quite similar to a normal distribution, just with most of the effects peaking at the center, with long tails.

Question 3

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: wt ~ sex + mage + lit + died + alive + age + (age | id)
##    Data: nepali1
##
## REML criterion at convergence: 1708.1
##
## Scaled residuals:
##      Min      1Q    Median      3Q      Max
## -3.01548 -0.44752  0.01788  0.50620  2.68119
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 1.948967 1.39605
##           age         0.001295 0.03598  -0.58
##  Residual             0.140996 0.37549
## Number of obs: 877, groups:  id, 197
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  5.039018   0.538186   9.363
## sexfemale   -0.425426   0.181869  -2.339
## mage         0.045641   0.025360   1.800
## lityes       0.284003   0.451137   0.630
## died        -0.055691   0.116944  -0.476
## alive        0.010976   0.070674   0.155
## age          0.136661   0.003387  40.344
##
## Correlation of Fixed Effects:
##           (Intr) sexfml mage   lityes died   alive
## sexfemale -0.125
## mage      -0.911 -0.036
## lityes    -0.118 -0.097  0.067
## died      -0.112 -0.120  0.203  0.090
## alive      0.545  0.051 -0.786 -0.001 -0.491
## age       -0.122  0.025 -0.041 -0.024 -0.012  0.003
## optimizer (nloptwrap) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 0.0192033 (tol = 0.002, component 1)
```

From this summary we can see that, the average spread of the intercept throughout the group of children differs by 1.948967 while the average spread of age across the group differs by 0.001295. The individual "id" represents an individual child and has different values spaced 4 months apart. From this, we can also see that the spread between different individuals is much greater than the spread between a single individual's own results since the intercept variance is much greater than the residual variance.

Question 4

```
##                    2.5 %       97.5 %
## .sig01        1.16951773   1.63316663
## .sig02       -0.70352844  -0.38289342
## .sig03        0.02950468   0.04188490
## .sigma        0.35029084   0.39790064
## (Intercept)   4.00061591   6.04167941
## sexfemale    -0.76800510  -0.06362264
## mage         -0.00312025   0.09009345
```

```
## lityes      -0.53402656  1.19336180
## died        -0.28640536  0.18078028
## alive       -0.11485790  0.14419421
## age          0.13028153  0.14296107
```

We can use this bootstrapped confidence interval, and the summary from the previous question to determine significance of each term. We can see that the standard deviation for the random intercept does fall within the ".sig01" confidence interval, meaning it is not a significant random effect. The correlation of the random effects of -0.58 does fall within the ".sig02" confidence interval too, meaning it is not significant either. The random slope however, does not have a standard deviation contained within the ".sig03" confidence interval, meaning it is significant.

Question 5

We can use the confidence intervals and summary from the previous 2 questions to interpret the estimates of different fixed effects in the model.

The estimated fixed effects of sex ("sexfemale" here) indicates that female children will, on average be -0.425426kg lighter. The estimated fixed effects of "age" indicates that a child that is 1 month older in age will on average be 0.136661kg heavier. The estimated fixed effects of "mage" indicates that a mother that is 1 year older will have a child that is 0.045641kg heavier, on average. The estimated fixed effects of lit ("lityes" here) indicates that a mother who is literate will have a child that is 0.284003kg heavier on average.

Question 6

```
#null model - model without random effects
lmnull <- lm(wt ~ sex + mage + lit + died + alive + age, data=nepali1)
#original model - with random effects
lmer2 <- lmer(wt ~ sex + mage + lit + died + alive + age + (1|id), data=nepali1)

#estimated Likelihood Ratio Test (LRT) statistic
lrtstat <- as.numeric(2*(logLik(lmer2)-logLik(lmnull)))

#parametric bootstrap of the model with and without random effects (for more accurate LRT statistic)
y <- simulate(lmnull)
lmlst <- numeric(1000)
for(i in 1:1000){
  y <- unlist(simulate(lmnull))
  lm_null <- lm(y ~ sex + mage + lit + died + alive + age, data=nepali1)
  lmer_2 <- lmer(y ~ sex + mage + lit + died + alive + age + (1|id),  data=nepali1, REML=FALSE)
  lmlst[i] <-  as.numeric(2*logLik(lmer_2)-logLik(lm_null))
}

p_value <- mean(lmlst>lrtstat)
std_error <- sqrt((p_value*(1-p_value))/1000)
data.frame(p_value, std_error)
```

```
##   p_value std_error
## 1       0         0
```

Here we use the parametric bootstrapping approach to determine whether the fixed effects in this model are jointly significant. We can see that the p-value produced is 0 so we can reject the null hypothesis that there are no differences between all the fixed effects. This means the fixed effects are not jointly significant. This p-value has a standard error of 0 as well, so we can be quite certain that this is an accurate value.