1 Notes from the meeting

1. make initial values from FE multinomial
2. number of samples, number of categories, zeros, are all things that might explain lack of convergence
3. if there is something that perfectly explains a category, you get error messages too. Make sure that there are no prefect spits
4. it's very sensitive to starting point, so ignoring random effects and just fitting fixed effects is likely to give you a good estimate
5. when it doesn't converge, it means that the problem is that the model is not fitting for the data
6. simulate data. observed in x axis, simulated in y. simulate data under these parameters. rank the observations for lowest to highest, for each sample, and then compare the two. do this for each sample independently. 95% coverage as area in the line. this is to show that the model is not unrealistic, not that the mpdel os realistic. check the tails
7. the number of degrees of freedom is a function of n and N and are a function of the correlation of the within-patient. if the correlation is very high we have a problem. if the correlation is very low then we are in a better situation
8. fitting something like 22 parameters and 35 samples might be a problem for convergence
9. check with the standard error whether the true values fall in the confidence interval
10. send data to Dom
    - worst case scenario: small sample size, high cat
    - med
    - best: large number of samples

    (ratio of N/p varying)
11. he's got this method of using some particular initial estimator the estimate, esp. with tricky situations. Estimator found by indirect inference (takes forever, but they found a fast solution). could be applied here. it removes asymptotic bias & small sample bias
12. in correlated RE it can be very tricky to find a solution
13. send him the algorithm with good/optimized initial parameters (multinom reg)

14. when you're integrating out the RE, TMB does estimate the random effects, to take them out
15. I use optim, but he doesn't. Instead he uses nlminb instead
(`nlminb(start = obj$par, obj = obj$fn, gr = obj$gr)`)
16. when I do opt I already have the information, and sdreport extracts it
17. Share a script in which he can load and run the TMB c function, and then a script which simulates data under the model, and then run code under simulated dataset, and additionally the raw data of the scenarios
18. simplification might come at a big cost
19. check what they did in the article
20. send him an email about what I thought about the likelihood
21. empirically check unidentifiability due to lack of information: simulation under different number of samples
22. send text about categorical

2 Categorical distribution