The marketing department at your organization has asked you to analyze their latest brand equity survey. They are particularly interested in responses to how satisfied respondents were to their last interaction with the company – rated Extremely Dissatisfied to Extremely Satisfied. Obviously, a linear model is out of the question since you have ordinal data. As you start digging deeper you realize that the survey was conducted over multiple counties. This is a problem since now you have multilevel data – there may be unobserved things about the different counties or the purchasing behavior of people in different counties that affects their responses. Now you are worried because you have never heard of a multilevel ordered logistic regression – yes that is a mouthful. Not to worry, there is such a thing and it provides a solution to your problem.

The aim of multilevel ordered logistic regression (MOLR) is to estimate the odds that a certain ordered response occurs while at the same time accounting for the fact that you have dependencies in the data. In other words, what are the odds that someone gives a certain rating knowing that individuals belong to certain groups. MOLR can be applied to a number of common situations in market analysis, public policy analysis, and social psychology. For example, MOLR has been used to test the influence of patient satisfaction with health clinics in different counties (O'Connel 2010) and on attitudes to LGBTQ individuals across countries (Mathisen 2018). Below I provide a primer on how to apply these models using the ordinal package in the R statistical programming language.

General Principles

Assume your data has N=1,000 respondents across 4 counties, e.g. you have N participants in level one nested within K clusters (level 2 units). Counties pertain to clusters rather than predictor variables since counties have no intrinsic meaning per se but there may still be variation in responses that occurs because of counties. This is different than other potential grouping variables such as age or socioeconomic status which may indeed have some meaningful effect in and of themselves on the outcome of interest.

Given this data structure you cannot simply run an ordered logistic regression since it violates the assumption of a lack of correlation in the residuals – recall that there may be things about counties that we do not observe but which nonetheless affect our outcomes. This also means that participants nested in one cluster are more likely to behave similarly to those in the same county than they are to those in another county. MOLR aims to disentangle the correlations between the groups in order to provide accurate means of estimating the odds of a given response. This has a couple of implications for interpreting results:

1. The (log) odds that the outcome variable takes a given value is allowed to vary between the clusters. In the marketing example, the chances that someone provides a response Extremely Dissatisfied to Extremely Satisfied varies from county to county. Specifically we differentiate the odds that of a response overall from the odds in a given county.
2. The effects of lower level (level 1) variables on the outcome also varies between clusters. If, for example, you had a variable on the amount of time someone spent in a store as a predictor of how satisfied they were then we would again differentiate the average (log) odds of a response based on time from the odds within each county.

How do we apply this?

Christensen's (2011) ordinal package fortunately has a built-in MOLR function. Below I provide commands for applying the model to a survey of hospital patients on their level of satisfaction (rate) based on four predictors and one grouping variable:

age: The age of the respondent
wt.time: The time spent in the waiting room
nurse: Whether the respondent saw a nurse or a physician
medicaid: Whether the patient uses Medicaid
county: The county in which the hospital's office was located

Note that I just made up this data so it should not be taken as a true analysis.

```
>library(ordinal)
>library(lrtest)
>mydata<-read.csv(file = "~/some.data.csv")
>fm1<-clmm(as.factor(RATE)~AGE+TIME+NURSE+(1|COUNTY), data=mydata)
>fm2<-clm(as.factor(RATE)~AGE+TIME+NURSE+(1|COUNTY), data=mydata) # we also
```
want to runa standard ordered logit to determine if the ordered model is adds any additional leverage. We assume we need to account for county level variation, but it may not actually exist.
```
>lrtest(fm2, fm1)

>summary(fm1)
```

Random effects:

Groups Name	Variance	Std.Dev.
county (Intercept)	0.10271	0.00147

Number of groups: county 4

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)	
age	-0.07100	0.03639	-1.951	0.051	.
wt.time	-0.92626	0.05250	-17.645	<2e-16	***
nurse	-0.07872	0.11431	-0.689	0.491	
medicaid	-0.04076	0.17804	-0.229	0.819	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

	Estimate	Std. Error	z value	+random	-random
1|2	-3.762303	0.208117	-18.078	-3.659	-3.865
2|3	-2.282557	0.183391	-12.446	-2.180	-2.386
3|4	-1.147724	0.171056	-6.710	-1.044	-1.248
4|5	-0.009327	0.170376	-0.055	0.062	-.0510

logLik -1767.21
df 9

>summary(fm2)

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)
age	-0.07192	0.03640	-1.976	0.0482 *
wt.time	-0.92672	0.05254	-17.639	<2e-16 ***
nurse	-0.06675	0.11472	-0.582	0.5606
countyCounty B	-0.03515	0.14925	-0.235	0.8138
countyCounty C	-0.22030	0.17271	-1.276	0.2021
countyCounty D	-0.02836	0.16165	-0.175	0.8608

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

	Estimate	Std. Error	z value
1|2	-3.81638	0.22956	-16.625
2|3	-2.33382	0.20702	-11.273
3|4	-1.19735	0.19647	-6.094
4|5	-0.05871	0.19563	-0.300

logLik -1794.97

df 11

>lrtest(fm2, fm1)

```
  #Df LogLik     Df Chisq   Pr(>Chisq)

1  11  -1794.97

2  9   -1767.21  2  55.400   ***
```

We see that the more complex model explains more variation and is therefore preferred. On to the interpretation. We see that only wait time is a significant predictor of satisfaction, but how do we interpret this information? Like ordered logistic regression, we must convert the estimated coefficients from log odds to odds ratios in order to determine probabilities.

```
>logprobs<-function(logit){
odds<-exp(logit)
prob<-odds/(1+odds)
return(prob)
}

>logprobs(coef(fm1))
1|2     2|3     3|4     4|5      age      wt.time   nurse   medicaid
0.023   0.093   0.241   0.498    0.482    0.284     0.480   0.490


1|2     2|3     3|4     4|5
0.025   0.102   0.260   0.515 (+random)
0.021   0.084   0.223   0.487 (-random)
```

We now have the average probabilities for our cut points (1,2,3,4,5, i.e. Extremely Dissatisfied to Extremely Satisfied) and the average probabilities for our level 1 predictors. As with interpretation of standard multilevel models, the intercepts can vary by as much as the random effects terms. For simplicity's sake I have already shown these above. Note that for random effects terms the interpretation is fairly straightforward however the interpretation becomes more cumbersome and we must address some distributional assumptions that I will not address here. Note that we are saying that the intercepts can vary by this much, not that they always do. An easier means of interpretation therefore is to simply say that the level of satisfaction varies significantly by county.

To get the probabilities for wait time's effect on satisfaction (rate) we can simply add our probabilities:

1|2 + TIME = .023 + .284 = 0.307 [+/- random = 0.305; 0.309]
2|3 + TIME = .093 + .284 = 0.377 [+/- random = 0.368; 0.386]
3|4 + TIME = .241 + .284 = 0.525 [+/- random = 0.507; 0.544]
4|5 + TIME = .498 + .284 = 0.782 [+/- random = 0.771; 0.799]

To address how changes in wait times affect the intercepts, it is necessary to use first differences of wt.time which I will not address here. As with transforms from log odds to probabilities though, the process is essentially the same as with ordered logistic regression.

The interpretation of these models is not as difficult as may initially be assumed. There are some sample size limitations, e.g. they should not be attempted with less than 500 observations and even then, the interpretations are a bit questionable. In any case, MOLR models provide a tool for addressing nesting problems that may come about in our data.