# Mixture and Group-Based Trajectory Models – part 3.

## Transcript

Video: https://youtu.be/qQby21PSK9Y

So in previous presentations, I have given an introduction to growth mixture models and latent class growth analysis, emphasising that these are person-centred models that allow us to identify finite and smaller numbers of groups that show distinctive trajectories in the variables we observe over time. So if our goal is to identify groups of people that show distinctive patterns of development or developmental trajectories across time, growth mixture models and latent class growth analysis are the best analyses we can use, also because they are methods based on probability rules, so they allow us to give us an indication and a formal assessment of how good is the categorisation we are applying, and how adequate is that categorisation to represent the heterogeneity and the variability in the sample we observe.

In this third presentation, I will talk about how we can add covariates to these models and I will mention how we can introduce time in variant and time varying covariates in the models, and also how these models can help investigate cumulative risk. I will also talk about how we can include distal outcomes in the models, so how we can test whether the different trajectory groups are associated with different outcomes forward in time. And then I will mention some advanced models, in particular dual trajectory models, and multiple trajectory models.

Before I go on, I wanted to outline some key differences between trajectory-based models, and another type of mixture models for longitudinal data, latent transition analysis. And I have also prepared a presentation and material on latent transition analysis for NCRN, so you’re welcome to look at that if you’re interested. The key difference between those two approaches, that trajectory based group models represent groups that differ qualitatively in their behaviour over a period of time. In other words, the latent classes in trajectory-based group models represent distinct patterns of change that happens over a period of time and unfold over a period of time. In the example on the left in this slide, for example, the change unfolds over a period of five years. Latent transition analysis represents qualitative changes in patterns of responses that happen at each specific measurement occasion. So trajectory-based group models are more ideal or adequate for investigating individual differences in developmental patterns that unfold over a significant period of time or over a significant number of measurement occasions, whereas latent transitional analysis is more adequate to investigate stadial developmental models, models where individuals are supposed to move across different stages of development.

[00:03:58]

So in the second presentation I had provided a more formal description of growth mixture models and latent class growth analysis focusing on cases where we deal with ordered categorical variables, however the models I am presenting can be run with any type of variables, from continuous to nominal. And one of the main purposes of these models is to investigate predictors and covariates that might influence the probability of being in one trajectory group or another. For example, does family separation affect the probability of adolescents being in a chronic drug use trajectory group? When we introduce these covariates, because latent classes are nominal variables, and in other words, they are not ordered, we run multinominal logistic regressions of the latent classes on the covariates.

In the exercises, there are some examples of how to introduce covariates and interpret their results, so you are welcome to look at them. However, since growth mixture models also allows individual variability around the growth parameters within classes, the covariates that we introduce in the models can also influence intraclass (inaudible 00:05:19) variation around the intercepts and the slopes. For example, within adolescents that show a chronic trajectory of drug use, the specific trajectory of an adolescent within this group may be affected negatively by parental separation and the increase in drug use may be more dramatic for those adolescents that within the chronic use group were exposed to family separation.

So in this case, models with covariates can be more complex and explore more nuances in mechanism of influence, so when we are using growth mixture models, there are also some more nuances that we can model in this approach.

In the case of latent class growth analysis, the intraclass intercepts and slopes are constant across individuals, so there is no individual variation around the class-specific growth parameters. And when we introduce covariates, therefore we estimate how covariates influence latent class affiliation, or at least the probability of being in one trajectory group or another. Once again, these associations are modelled by multinomial logistic regressions of latent classes on covariates.

So we are often interested in introducing different covariates that can act as predictors of trajectory groups, trajectory group affiliation and we can estimate the influence of these covariates, using multinomial logistic regression of the latent classes on the covariates. But one appealing feature of this method is that we can also calculate probabilities for different configurations of covariates, and there are some exercises on this in the material I have prepared. And this really allows to test hypotheses on the role that the community of exposure to some risk factors may have on the probability of being in a specific trajectory group. So it’s possible to look at the effects, the cumulative effect of risk factors on the probability of displaying some pattern of development trajectories.

And here I display a fictional example where the vertical axis in the graph represents the probability of adolescents being in a chronic antisocial behaviour trajectory group. In this example, the probability of being chronically antisocial is relatively low when adolescents are exposed to only one of these risk factors – low socioeconomic status, parental separation, insecure attachment. But if adolescents are exposed to all three factors over time, the risk of being in the chronically antisocial group is significantly higher. So this is an example of how we could use these approaches to investigate cumulative risk.

Any examples and applications in the covariates that are used to predict trajectory-based groups are invariant, for example, they represent events that happened before the study started. It is, however, possible to include time-variant covariates, and here I refer to a study by Nagin and colleagues, and the QR code links to this publication, if you’re interested. The researchers investigated trajectories of violent behaviour in adolescence, and tested if school failure was associated with changes in expected trajectory. In Quebec, where the study was conducted, school failure led to grade retention. And here I also reported the equation that they also devised to test these hypotheses, where you can see that the latent response Y\* displays a quadratic trajectory across age, and there were additional parameters for the effects of failing school for the first time between age 6 and 10, failing school for the first time between age 11 and 12, or failing school for the first time between age 13 and 15. And note the parameters for school failure indexes for different latent class is k being estimated, or the effects of these covariates varied across classes, or in other words, individual across classes had different trajectories as well as displaying differences in the effects of school failure.

[00:10:21]

The graph I produced here represents the potential effect of school failure at an early age, say age 9 on a trajectory, so in this example, the effect of school failure is basically to elevate the expected trajectory to a higher level. And this is an example of how time variant covariates can be included in the model, so I report the QR code so that you can look at this paper if you want to know more.

Another key interest of researchers that use growth mixture models or latent class growth analysis lies in the possibility of testing if trajectory groups differ in the distal outcome or achievements later on. It is easy to include distal outcomes in the models and the exercises I devised with these presentations also provides some examples. However, when we introduce distal outcomes, we introduce further observed variables that vie together with the other variables that are in the model. This means that the latent classes are not just modelling their heterogeneity we observed in the observed variables and their trajectories, but also the heterogeneity of this and the distal outcomes. So consequently, this means that often, the latent class models we had estimated without including distal outcomes may change, and we will have to recheck the latent class models to see how much it has changed and how. And this is sometimes impractical but also requires some caveats in interpreting the results.

And a solution to this problem had been to consider individuals’ latent class affiliation in trajectory groups, as if they were variables that could be used as predictors in other analyses, for example, in a nova. But the problem with this approach where the latent class affiliation is used as if it were similar to an observed variable is that latent class affiliations are uncertain and in the first presentation I gave an example of how one individual could be assigned to one latent class, when the individual could have, say, 50% probability of being in that latent class. So if we fail to take into account these uncertainties, the analyses we run are going to be biased.

An elegant solution to this problem has been provided by the three-step approach. The advantage of this approach lies in separating the step where we estimate the latent class affiliations from the step where we impose associations between latent classes and covariates of distal outcomes, or both. So the first step in this approach is to estimate the model and decide a number of classes, as well as estimate the most likely class of each participant based on the posterior probability. The second step involves calculating parameters that represent the uncertainty in this latent class allocation, and these parameters are (inaudible 00:13:43) of being the assigned latent class, rather than another one. Once we have estimated the most likely class and the parameter that summarises uncertainty around this classification, we can include a latent classification as a variable in the analysis, making sure that we feed the calculated uncertainty parameter into the model, in this way, controlling for this uncertainty. In the online material I provide more details and examples and also I talked in more detail about this in my presentations on latent class analysis and latent transitional analysis.

So this approach can be very helpful when you want to separate the step of developing the model, the latent class model from looking at the associations between the latent classes and distal outcomes or covariates or both. This approach, however, is not adequate if covariates have direct effect on the measurement parameters.

Finally, I’m going to talk about some more advanced models, and for example, we are often interested in investigating the associations between trajectories in different groups, for example, mental health and drug use. Using trajectory-based models, we can model individual differences in the trajectories across the two processes, and we can test then the associations between trajectory groups in one, for example, mental health, and trajectory groups in the other – drug use. In this case, one process, for example, mental health, in the graph you see here, happens slightly before drug use, and this type of model would provide some close stabs, where we can test if those show, for example, chronic mental health problems are significantly more likely to fall in a chronic drug use trajectory group, and we can also test hypotheses about these associations between different trajectory groups by restricting some of the probabilities between…of associations between the latent classes. So it’s possible to create different models where we can investigate the associations between individual differences in trajectories across the two processes and test specific hypotheses about those associations.

[00:16:20]

And finally, it’s possible to extend the models to represent trajectories in different behaviours, and here I provide an example of a study and I have provided a QR code, if you’re interested. And the researchers identified groups of Chinese elderly people that were taking part in a longitudinal study, who differed in their trajectories of physical activities, smoking and social participation. You can see in the pictures a representation of these different trajectory groups we identified, and for example, a large group was labelled isolated, active, non-smokers, because they were consistently physically active across time, but showed consistent low levels of smoking, but also low levels of social participation. The results indicated that trajectory groups characterised by low levels of social participation displayed significantly higher depression scores.

So to finalise, I have talked about how we can add different types of covariates to growth mixture models or latent class growth analysis models, and I’ve talked about how we can include distal outcomes and I mentioned some advanced models. So thank you very much for your attention, and if you want to know more, you can look at the exercises and some of the references I’ve put together with these presentations. Thank you very much. Bye now.

National Centre for Research Methods (NCRM)  
Social Sciences  
Murray Building (Bldg 58)  
University of Southampton  
Southampton SO17 1BJ  
United Kingdom

**Web** www.ncrm.ac.uk   
**Email** info@ncrm.ac.uk  
**Tel** +44 23 8059 4539  
**Twitter** @NCRMUK