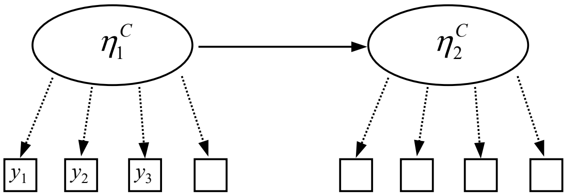
Latent Transition Analysis

Latent transition analysis (LTA) is the extension of latent class analysis to longitudinal data. A variety of model variations are possible to explore specific longitudinal research questions.

Basic Latent Transition Analyses

Latent transition models can be formulated in one of two ways, repeated-measures latent class analysis (RMLCA) or latent transition analysis. In both the latent class model is specified for each time point, based on prior analyses at each time point to determine estimated number of latent classes. The RMCLA model examines transitions among estimated classes across two or more time points without direction prediction, and so is analogous to a matched pairs McNemar or Stuart-Maxwell (marginal homogeneity) test but using estimated latent classes. Latent transitional models are more common and involve a directional prediction of the subsequent time point by the prior time point. This is an autoregressive model in which the predictive portion of the model is a logistic (if there are only two classes) or a multinomial logistic model (for three or more classes). The figure below depicts a latent transition model with four indicators. τ_{jc} as the response probability and $\alpha_{2|1}$ as the intercept/threshold for the multinomial logistic.



Newsom (2015), p. 276

In addition to the response probabilities, transition probabilities are estimated represents the probability that a case will have class membership in one of the classes at the subsequent time point given class membership at the prior time point. This is a conditional probability and is one for each category at the subsequent time point given initial membership in each of the categories at the initial timepoint. For example, if there are three classes of patients based on their attitudes toward annual physical check-ups ("opposed," "interested but infrequent" and "compliant"), then probabilities are computed for predicted membership in each of these categories at Time 2 given which particular class the patient was a membership in at Time 1 (3 × 3 = 9 transitions).

$$\hat{\pi}_{c2|1} = P\left(\eta_{i2}^C = c \mid \eta_{i1}^C = c\right) = \frac{e^{\alpha_{c2|1}}}{1 + e^{\alpha_{c2|1}}}$$

The transition probabilities are often organized into a matrix (Collins & Lanza, 2010) in which the rows are classes at initial time point, *t*, and columns classes at the later time point $t+1$. For example,

Time 1	Time 2		
	Opposed	Interested	Compliant
Opposed	$\hat{\pi}_{c2=1\|c1=1}$	$\hat{\pi}_{c2=2\|c1=1}$	$\hat{\pi}_{c2=3\|c1=1}$
Interested	$\hat{\pi}_{c2=1\|c1=2}$	$\hat{\pi}_{c2=2\|c1=2}$	$\hat{\pi}_{c2=3\|c1=2}$
Compliant	$\hat{\pi}_{c2=1\|c1=3}$	$\hat{\pi}_{c2=2\|c1=3}$	$\hat{\pi}_{c2=3\|c1=3}$

Given the assumption that all class memberships are mutually exclusive and mutually exhaustive, the conditional probabilities for transitions out of a particular class at Time 1 should sum to 1.0 (i.e., sum of row values in the above matrix). It is common to require that the response probabilities be equal over time to ensure that the contribution of one of the indicators (e.g., "I get nervous around doctors") to the probable class membership is equally important at both time points.[1] Predictive models also examine the logistic or multinomial regression regression coefficients from which odds ratios can be computed (e.g., odds or membership in the compliant group if initially I the interested group). As with multinomial regression, if there are more than two classes, a referent group must be chosen for the outcome latent class variable.

LTA with More Time Points

More time points can be investigated in latent transition models (in which the models may be referred to as "Markov chain models"). With more time points, several interesting specific hypotheses may be of interest. The idea of stage sequential modeling (Collins & Wugalter, 1992; Lanza, 2003) investigates certain patterns of class membership (e.g., stages of change in health behavior research, or Piagetian developmental stages). The concept of the stage sequence is that an individual has a specific class membership pattern. For example, in a set of six consecutive binary latent classes, one individual's class membership might be 1, 2, 2, 1,1, 1. This individual moves from Class 1 to Class 2 at the second occasion but then moves back to Class 1 at the fourth occasion. The mover–stayer model is a specific type of the general state sequential model that classifies individuals in terms of their transition patterns over time (van de Pol & Langeheine, 1989). "Movers" are individuals who change from one state to another, whereas "stayers" are individuals who do not change.

There are several potentially important advantages of latent transition models, including the ability to account for measurement error over time, the ability to investigate and ensure measurement invariance over time, and a data reduction function for examine change in qualitative states in which the complexity of many variables is reduced by estimating a smaller number of underlying latent classes (Lanza, Patrick, & Maggs, 2010). There also are a number of challenges with estimating latent class and latent transition analyses (Nylund-Gibson & Choi, 2018). Estimation of latent transition models with several time points becomes increasingly computationally complex. With binary indicators, the number of contingency values increases exponentially (Collins & Lanza, 2010). For example, with just two time points and five indicators there are $2^{2\times5} = 1{,}024$ elements in a crosstabulation table of all of the variables. With three time points, this expands to $2^{3\times5} = 32{,}768$. This poses some potentially very significant computational challenges! In general, more complex models will be more likely to have difficulty with convergence and can produce results from local minima. It is important to use multiple start values (usually randomly generated by the software) and retest models by increasing the number of start values. Many models may require constraints in order to obtain convergence.

Growth Mixture Models

Although they are not latent transition models per se, another approach to longitudinal modeling that utilizes latent class analysis is the growth mixture model (Muthén & Shedden, 1999; Nagin, 1999). Growth mixture models involve a latent class analysis of growth curve model parameters. Growth curve models obtain estimates of initial values and slopes for change over three or more time points. Growth mixture models estimate classes for these parameters. For example, a growth mixture model might find that there are three classes of individual change in social support over time—those who start out with a high amount of support who stay high over time, those who start out with a low amount of support and increase over time, and those who start out with a low about of support and stay low over time. For good didactic illustrations, see Ram and Grimm (2009) and Infurna and Grimm (2018).

---

[1] It is possible to relax this assumption, in which the classes would have different meanings at each time point, and it may even be of interest to have a different number of classes at each time point. Such models would deviate from the usual approach to measurement invariance, but they are potentially useful to examine changes among qualitatively different classes over time.

Examples

Data for this example come from the Health and Retirement Study (HRS). In a recent paper on multimorbidity,[2] we examined latent classes estimated for eight chronic health conditions (e.g., heart disease, stroke, cancer). Prior analyses led us to conclude there were three latent classes (roughly labeled, “minimal disease, ”cardiovascular-musculoskeletal,” “”cardiovascular-musculoskeletal-mental”). We then examined latent transitions across 16 years.

R

There are not many options in R for conducting latent transition analysis. Below, I use the `LMest` package (Bartolucci, Pandolfi, Pennoni, 2017) that estimates “Generalized Latent Markov Models”. Data must be transformed to long format and my initial set of commands does that. On the lmest function, k = is the number of classes extracted, `modBasic = 1` is for “time-homogeneous transition matrices” (for constraining transition probabilities across waves to be equal when there are more than 2), `start = 1` uses random start values, `maxit = 5000` increases the default maximum iterations from 1000 to 5000, and `seed` = sets a random seed for the start value generator for replication purposes. Note that the output does not match the SAS PROC LTA results very closely.

```
> #transform wide data format to long data format
> longd <- reshape(d, idvar="id",
+                   varying=list(c("bp98","bp14"),
+                   c("diab98","diab14"),
+                   c("cancr98","cancr14"),
+                  c("lung98","lung14"),
+                   c("hrt98","hrt14"),
+                    c("strk98","strk14"),
+                    c("arth98","arth14"),
+                     c("depression98","depression14"),
+                    c("time1","time2")),
+                   v.names=c("bp","diab","cancr","lung","hrt","strk",
+                             "arth","depression","time"),direction="long",sep="")

> library(LMest)
> model <- lmest(index = c("id","time"),
+                k = 3,
+                data = longd,
+                modBasic = 1,
+                start = 1,
+                maxit = 5000,
+                seed = 052421)


> model

Basic Latent Markov model
Call:
lmest(data = longd, index = c("id", "time"), k = 3,
    start = 1, modBasic = 1, maxit = 5000, seed = 52421)

Available objects:
 [1] "lk"       "piv"      "Pi"       "Psi"      "np"       "k"        "aic"      "bic"      "lkv"
"V"        "n"        "TT"       "modBasic" "Lk"       "Bic"      "Aic"      "call"     "data"

Convergence info:
     LogLik np k    AIC      BIC     n TT
  -147177.5 56 3 294467 294945.8 38183  2
> summary(model)
Call:
lmest(data = longd, index = c("id", "time"), k = 3,
    start = 1, modBasic = 1, maxit = 5000, seed = 52421)

Coefficients:

Initial probabilities:
     est_piv
[1,]  0.1559
[2,]  0.3121
[3,]  0.5319
```

---

[2] Quiñones, A.R., Newsom, J.T., Elman, M.R., Markwardt, S., Nagel, C.L., Dorr, D.A., Allore, H.G., and Botoseneanu, A. (2021). Racial and ethnic differences in multimorbidity changes over time. *Medical Care*, 59(5), 402-409. doi: 10.1097/MLR.0000000000001527.

```
Transition probabilities:
     state
state      1      2      3
    1 0.9786 0.0000 0.0214
    2 0.2737 0.7263 0.0000
    3 0.1652 0.3099 0.5250

Conditional response probabilities:
, , item = 1

        state
category      1 2      3
       0 0.0000 0 0.0000
       1 0.1554 0 0.9912
       2 0.8446 1 0.0088

, , item = 2

        state
category      1      2      3
       0 0.0000 0.0000 0.0000
       1 0.6002 0.7713 0.9233
       2 0.3998 0.2287 0.0767

, , item = 3
        state
category     1      2      3
       0 0.000 0.0000 0.0000
       1 0.771 0.8796 0.8989
       2 0.229 0.1204 0.1011

, , item = 4
        state
category      1      2      3
       0 0.0000 0.0000 0.0000
       1 0.6905 0.9601 0.9417
       2 0.3095 0.0399 0.0583

, , item = 5
        state
category      1      2     3
       0 0.0000 0.0000 0.000
       1 0.3576 0.8329 0.888
       2 0.6424 0.1671 0.112

, , item = 6
        state
category     1      2      3
       0 0.000 0.0000 0.0000
       1 0.773 0.9488 0.9803
       2 0.227 0.0512 0.0197

, , item = 7
        state
category      1      2     3
       0 0.0000 0.0000 0.000
       1 0.1194 0.4552 0.564
       2 0.8806 0.5448 0.436

, , item = 8
        state
category     1      2      3
       0 0.000 0.0000 0.0000
       1 0.661 0.8851 0.8967
       2 0.339 0.1149 0.1033
```

SAS

The PROC LTA macro developed by Lanza and colleagues[3] to estimate transition probabilities. See notes below in comments section.

```
/*proc lta requires 1, 2 values, seed is a random seed for random starting values,
nstatus is the number of latent classes, assumed to be the same at both time points,
ntimes is the  number of time points, items are the indicator variables and must be
the same number of items at each time point listed in the same order, and categories
is the number of response categories for each variable (assumed to be the same across
```

---

[3] PROC LCA & PROC LTA (Version 1.3.2) [Software]. (2015). University Park: The Methodology Center, Penn State. Retrieved from http://methodology.psu.edu

```
time),  meaurement times constraints response probabilities to be invariant over
time*/

PROC LTA;
NSTATUS 3;
NTIMES 2;
ITEMS bp98 diab98 cancr98 lung98 hrt98 strk98 arth98 depression98
 bp14 diab14 cancr14 lung14 hrt14 strk14 arth14 depression14;
CATEGORIES 2 2 2 2 2 2 2 2;
seed 052421;
measurement times;
run;
```

```
              Data Summary, Model Information, and Fit Statistics (EM Algorithm)



Number of subjects in dataset:       38183
Number of subjects in analysis:      38183

Number of measurement items per time:    8
Response categories per item:            2 2 2 2 2 2 2 2
Number of occasions (times):             2
Number of groups in the data:            1
Number of latent statuses:               3

Rho starting values were randomly generated (seed = 52421).

Parameter restrictions: Rho (measurement) parameters were constrained to be equal across time.

The model converged in 899 iterations.

Maximum number of iterations: 5000
Convergence method: maximum absolute deviation (MAD)
Convergence criterion:  0.000001000

=============================================
Fit statistics:
=============================================

Log-likelihood:   -147520.88
G-squared:          18571.46
AIC:                18635.46
BIC:                18909.07
Degrees of freedom:    65503

Test for MCAR
      Log-likelihood:   -138235.15
      G-squared:           9934.60
      Degrees of freedom:   578738


                            Parameter Estimates

Delta estimates (status membership probabilities):
Status:                   1          2          3
  Time  1     :       0.4548     0.3344     0.2109
```

```
  Time  2     :       0.2694     0.2711     0.4595

Tau estimates (transition probabilities):
Time   1 latent status (rows) by
    Time   2 latent status (columns)
                         1          2          3
        1     :       0.5129     0.2281     0.2590
        2     :       0.0000     0.5006     0.4994
        3     :       0.1717     0.0000     0.8283


Rho estimates (item-response probabilities):
(All times)
  Response category:  1:
Status:                  1          2          3
  bp98        :       0.6393     0.5837     0.0806
  diab98      :       0.9030     0.9210     0.5446
  cancr98     :       0.9148     0.8701     0.7962
  lung98      :       0.9590     0.9130     0.7910
  hrt98       :       0.8949     0.8393     0.4851
  strk98      :       0.9763     0.9748     0.8057
  arth98      :       0.9980     0.0001     0.2022
  depression98:       0.9149     0.8601     0.7318

  Response category:  2:
Status:                  1          2          3
  bp98        :       0.3607     0.4163     0.9194
  diab98      :       0.0970     0.0790     0.4554
  cancr98     :       0.0852     0.1299     0.2038
  lung98      :       0.0410     0.0870     0.2090
  hrt98       :       0.1051     0.1607     0.5149
  strk98      :       0.0237     0.0252     0.1943
  arth98      :       0.0020     0.9999     0.7978
  depression98:       0.0851     0.1399     0.2682
```

I have not illustrated any here, but a few structural equation modeling packages, such as Mplus or Mx, can estimate latent transition models as well.

References and Further Reading

Collins and Lanza. Ch. 7 Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences* (Vol. 718). Chapter 7 “Repeated Measures Latent Class Analysis and Latent Transition Analysis”.

Collins, L. M., & Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. *Multivariate Behavioral Research, 27,* 131–157.

Infurna, F. J., & Grimm, K. J. (2018). The use of growth mixture modeling for studying resilience to major life stressors in adulthood and old age: Lessons for class size and identification and model selection. *The Journals of Gerontology: Series B, 73*(1), 148-159.

Lanza, S. T., Patrick, M. E., & Maggs, J. L. (2010). Latent transition analysis: Benefits of a latent variable approach to modeling transitions in substance use. Journal of Drug Issues, 40(1), 93-120.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics, 55*, 463–469.

Nagin, D. S. (1999). Analyzing developmental trajectories: a semiparametric, group-based approach. *Psychological Methods, 4*, 139–157.

Newsom, J.T. (2015). *Longitudinal Structural Equation Modeling: A Comprehensive Introduction.* New York: Routledge. Chapter 10 “Latent Transition and Growth Mixture Models”.

Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science, 4*(4), 440.

Ram, N., & Grimm, K. J. (2009). Methods and measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development, 33*(6), 565-576.

van de Pol, F., & Langeheine, R. (1989, October). Mixed Markov Models, mover-stayer models and the EM algorithm. In R. Coppi & S. Bolasco (Eds.), *Multiway data analysis* (pp. 485–495).Amsterdam: North-Holland.