ALTA model appendix

Brian P. Flaherty
Department of Psychology
University of Washington

Abstract

This appendix accompanies the article *Testing the degree of cross-sectional and longitudinal dependence between two discrete dynamic processes*. It provides a detailed presentation of the ALTA model and the specific parameter restrictions for testing general patterns of cross-sectional and longitudinal association.

ALTA model

The ALTA model may be defined as the following. Let $\boldsymbol{W}$ denote a vector of categorical random variables, qT long, where T is the number of measurement occasions and q is the number of variables measuring both latent variables on a single occasion. For example, if four variables measure X and three measure Y, then $q = 7$. Therefore, W_{tj} denotes the jth random variable at Time t. Let $\boldsymbol{r}$ denote a vector of length qT of the number of response categories of the q items at each of the T times. One can form a contingency table of the variables in $\boldsymbol{W}$, with a cell in this contingency table denoted as $\boldsymbol{w}$, corresponding to a response pattern comprised of all items over all times. For example, w_{32} denotes the response to the second variable at the third occasion in response pattern $\boldsymbol{w}$.

In the ALTA model, the levels of discrete latent variables are the latent classes. Let C_t and D_t denote the number of latent classes in X and Y, respectively, at Time t. Often $C_1 = C_2 = \ldots = C_T = C$ and $D_1 = D_2 = \ldots = D_T = D$, but this is not necessary. Given this notation, the ALTA model is

$$P(\boldsymbol{W} = \boldsymbol{w}) = \sum_{c_1=1}^{C_1} \sum_{d_1=1}^{D_1} \cdots \sum_{c_T=1}^{C_T} \sum_{d_T=1}^{D_T} \alpha_{c_1} \beta_{d_1|c_1} \left(\prod_{j=1}^{q} \prod_{k=1}^{r_j} \rho_{1jk|c_1 d_1}^{I(w_{1j}=k)} \right) \times \prod_{t=2}^{T} \left(\epsilon^{(t-1)}_{c_t|c_{t-1}d_{t-1}} \eta^{(t-1)}_{d_t|c_{t-1}c_t d_{t-1}} \prod_{j=1}^{q} \prod_{k=1}^{r_j} \rho_{tjk|c_t d_t}^{I(w_{tj}=k)} \right), \quad (1)$$

where

- α_{c_1} in the unconditional probability of membership in latent class c_1 of X at Time 1;
- $\beta_{d_1|c_1}$ is the probability that a person is in latent class d_1 in Y conditional on membership in the c_1 latent class in X at Time 1;

- $\epsilon^{(t-1)}_{c_t|c_{t-1}d_{t-1}}$ is the probability of being a member of X latent class c_t at Time t conditional on being in latent classes c_{t-1} and d_{t-1} of X and Y, respectively, at Time $t-1$;
- $\eta^{(t-1)}_{d_t|c_{t-1}c_t d_{t-1}}$ is the probability of being in latent class d_t of Y at Time t conditional on membership in X latent classes c_{t-1} and c_t at Times $t-1$ and t, respectively, and Y latent class d_{t-1} at Time $t-1$; and
- $\rho^{I(w_{tj}=k)}_{tjk|c_t d_t}$ denotes the probability of response category k for item j at Time t conditional on membership in latent classes c_t and d_t in X and Y, respectively.
- $I(\cdot)$ is the indicator function that equals one when the argument is true and zero otherwise;

With this notation, we can precisely specify the broad patterns of constraints for testing associated change described in the body of the article. Independence of X and Y is specified by the following three sets of parameter restrictions:

$$\beta_{d_1|c_1} = \beta_{d_1}, \tag{2}$$

$$\epsilon^{(t-1)}_{c_t|c_{t-1}d_{t-1}} = \epsilon^{(t)}_{c_t|c_{t-1}}, \tag{3}$$

$$\eta^{(t-1)}_{d_t|c_{t-1}c_t d_{t-1}} = \eta^{(t)}_{d_t|d_{t-1}}. \tag{4}$$

Equation 2 implies that the probability of being in a particular latent class of Y is the same across all levels of X. Similarly, Equations 3 and 4 imply that all the $\epsilon^{(t-1)}_{c_t|c_{t-1}d_{t-1}}$ are constrained to be equal across levels of Y (d_{t-1}), and all $\eta^{(t-1)}_{d_t|c_{t-1}c_t d_{t-1}}$ are constrained to be equal across all pairs of X latent classes (c_{t-1} and c_t), respectively. The constraints on the $\epsilon^{(t-1)}_{c_t|c_{t-1}d_{t-1}}$ parameters stipulate that X latent class membership at Time t depends only on X latent class membership at Time $t-1$, and not on Y latent class membership at Time $t-1$. Similarly, the constraints on the $\eta^{(t-1)}_{d_t|c_{t-1}c_t d_{t-1}}$ parameters stipulate that Time t Y latent class membership depends only on Y membership at Time $t-1$, and not on X membership at either time.

The cross-sectional and longitudinal models each allow some dependence between the latent variables. The cross-sectional model allows dependence between X and Y at the first measurement occasion, but maintains statistical independence longitudinally. In this model, the $\beta_{d_1|c_1}$'s are freely estimated, but the restrictions on the $\epsilon^{(t-1)}_{c_t|c_{t-1}d_{t-1}}$'s and $\eta^{(t-1)}_{d_t|c_{t-1}c_t d_{t-1}}$ from the independence model are maintained (i.e., Equations 3 and 4). In contrast, the longitudinal model allows for longitudinal dependence between the latent variables, but no cross-sectional association at Time 1. In this case, the $\epsilon^{(t-1)}_{c_t|c_{t-1}d_{t-1}}$'s and $\eta^{(t-1)}_{d_t|c_{t-1}c_t d_{t-1}}$ are estimated freely, but the restrictions on the $\beta_{d_1|c_1}$'s (i.e., Equation 2) are retained.

The final model, full association, corresponds to the full ALTA parameterization in Equation 1, with no parameter restrictions that limit dependence between the latent class variables. Table 1 summarizes these four models of association.

Table 1: Patterns of parameter restrictions for testing different patterns of dependence between two discrete latent variables.

Pattern of Association	Parameters $\beta_{d_1\|c_1}$	$\epsilon^{(t-1)}_{c_t\|c_{t-1}d_{t-1}}$	$\eta^{(t-1)}_{d_t\|c_{t-1}c_td_{t-1}}$
Independence	Equal across c_1	Equal across d_{t-1}	Equal across c_{t-1} and c_t
Cross-sectional	Unrestricted	Equal across d_{t-1}	Equal across c_{t-1} and c_t
Longitudinal	Equal across c_1	Unrestricted	Unrestricted
Full	Unrestricted	Unrestricted	Unrestricted