Table of contents

Overview of models

Overview of models

The following tables provide an overview of the specified models, with model extensions/changes highlighted in **bold**. For every model, the observed data are assumed to be complete (i.e., no missing data).

Model	Dependent variable	Class membership	Constant & trend component[1]
1	continuous, Normal, iid err	time-independent	pooled, linear
2	continuous, Normal, iid err	time-independent	**individual effects**, linear
3	continuous, Normal, iid err	time-independent	individual effects, **quadratic**
4	**count**, **Poisson**, iid err	time-independent	individual effects, quadratic

Table 1: Overview of specified models

[1] for each latent class

General notation

General notation

Individual n, for $n = 1, ..., N$, where N is the number of individuals

Time period t, for $t = 1, ..., T$, where T is the number of time periods

$\boldsymbol{Y}^{obs}$ is a $N \times T$ matrix representing the observed dependent variable, where *obs* refers to simulated or actual data

$\boldsymbol{X}$ is a matrix of size $N \times T$ representing the explanatory variable (e.g., time periods, starting at zero)

Latent class (aka, sub-population) c, for $c = 1, ..., C$, where C is the number of classes

$\boldsymbol{\lambda}$ is a row vector of size C representing the mixture proportions, averaged over individuals and time periods; thus:

$$0 \leq \lambda_c \leq 1 \quad \text{and} \quad \sum_{c=1}^{C} \lambda_c = 1 \tag{1}$$

Model 1

Model 1 — dependent variable

Continuous

Normal (aka, Gaussian)

iid err: independent (i.e., uncorrelated over individuals and time periods) and identically distributed errors

No missing data

Model 1 — constant and trend component

Constant without interindividual differences within-class (i.e., the constant is allowed to vary between classes but not within classes)

Linear (i.e., non-stationary, deterministic) trend component without interindividual differences within-class (i.e., the trend component is allowed to vary between classes but not within classes)

Based on Koop (2003), the assumptions of no interindividual differences within-class regarding the constant and trend component correspond to a pooled model (i.e., a pooled model for each class)

According to Ram and Grimm (2009), latent class growth analysis (or LCGA for short) refers to a subset of GMMs assuming no interindividual differences within-class regarding the constant and trend component

Model 1 — likelihood

$$y^{obs}_{n,t} \sim \sum_{c=1}^{C} \lambda_c \, Normal(\mu_{c,n,t}, \sigma_c) \tag{2}$$

where $\boldsymbol{M}$ is a C-tuple containing $N \times T$ matrices, and $\boldsymbol{\sigma}$ is a row vector of size C.

$$\mu_{c,n,t} = \beta_{0,c} + \beta_{1,c} \, x_{n,t} \tag{3}$$

where $\boldsymbol{\beta}_0$ is a row vector of size C representing the constant, and $\boldsymbol{\beta}_1$ is a row vector of size C with

$$\beta_{1,c} < \beta_{1,c+1} \tag{4}$$

for identification. In combination with $\boldsymbol{X}$, $\boldsymbol{\beta}_1$ represents the linear trend component.

Model 1 — deduction of likelihood

For each $y_{n,t}$, there is a latent discrete parameter $z_{n,t}$ in $\{1, ..., C\}$, indicating that individual n belongs to class c at time period t:

$$z_{n,t} \sim Categorical(\boldsymbol{\lambda}) \tag{5}$$

where $\boldsymbol{Z}$ is a $N \times T$ matrix. Therefore, the likelihood presented in equation 2 is deduced by marginalizing out z_n:

$$\begin{aligned} p(y_{n,t}|z_n, \boldsymbol{\lambda}, \boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \boldsymbol{\sigma}) &= p(z_n|\boldsymbol{\lambda})\, p(y_{n,t}|z_n, \boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \boldsymbol{\sigma}) \\ &= \sum_{c=1}^{C} \lambda_c\, p(y_{n,t}|\beta_{0,c}, \beta_{1,c}, \sigma_c) \\ &= \sum_{c=1}^{C} \lambda_c\, Normal(\mu_{c,n,t}, \sigma_c) \end{aligned} \tag{6}$$

where $\mu_{c,n,t}$ is defined by equation 3.

Model 1 — prior

$$\boldsymbol{\lambda} \sim \mathit{Dirichlet}(\boldsymbol{\alpha}) \tag{7}$$

where $\boldsymbol{\alpha}$ is a row vector of size C with $\alpha_c = 1$ (i.e., $\boldsymbol{\lambda}$ is assigned a proper flat prior).

$$\beta_{0,c} \sim \mathit{Normal}(0, 10) \tag{8}$$

where the hyperparameters are advised to be specified based on the observed dependent variable.

$$\beta_{1,c} \sim \mathit{Normal}(0, 1) \tag{9}$$

where the hyperparameters are advised to be specified based on the observed dependent variable.

Model 1 — prior — continued

$$\sigma_c \sim \mathit{Normal}(0, 1)\,\mathbf{1}(\sigma_c > 0) \tag{10}$$

where the Normal distribution is truncated to the left at zero, modeled via the indicator function $\mathbf{1}(\sigma_c > 0)$. Furthermore, the standard deviation hyperparameter (here set to one) is advised to be specified based on the observed dependent variable.

Model 2

Model 2 — dependent variable

Continuous

Normal (aka, Gaussian)

iid err: independent (i.e., uncorrelated over individuals and time periods) and identically distributed errors

No missing data

Model 2 — constant and trend component

Constant with interindividual differences within-class (i.e., the constant is allowed to vary between classes and within classes)

Linear (i.e., non-stationary, deterministic) trend component without interindividual differences within-class (i.e., the trend component is allowed to vary between classes but not within classes)

Based on Koop (2003), the assumptions of interindividual differences within-class regarding the constant and no interindividual differences within-class regarding the trend component correspond to an individual effects model (i.e., an individual effects model for each class)

Model 2 — likelihood

$$y_{n,t}^{obs} \sim \sum_{c=1}^{C} \lambda_c \, Normal(\mu_{c,n,t}, \sigma_c) \tag{11}$$

where $\boldsymbol{M}$ is a C-tuple containing $N \times T$ matrices, and $\boldsymbol{\sigma}$ is a row vector of size C.

$$\mu_{c,n,t} = \beta_{0,n,c} + \beta_{1,c} \, x_{n,t} \tag{12}$$

$$\beta_{0,n,c} \sim Normal(\mu_{\beta_{0,c}}, \sigma_{\beta_{0,c}}) \tag{13}$$

where $\boldsymbol{B}_0$ is a $N \times C$ matrix representing the constant, and $\boldsymbol{\beta}_1$ is a row vector of size C with

$$\beta_{1,c} < \beta_{1,c+1} \tag{4}$$

for identification. In combination with $\boldsymbol{X}$, $\boldsymbol{\beta}_1$ represents the linear trend component.

Model 2 — deduction of likelihood

For each $y_{n,t}$, there is a latent discrete parameter $z_{n,t}$ in $\{1, ..., C\}$, indicating that individual n belongs to class c at time period t:

$$z_{n,t} \sim \mathit{Categorical}(\boldsymbol{\lambda}) \tag{5}$$

where $\boldsymbol{Z}$ is a $N \times T$ matrix. Therefore, the likelihood presented in equation 11 is deduced by marginalizing out z_n:

$$\begin{aligned} p(y_{n,t}|z_n, \boldsymbol{\lambda}, \boldsymbol{\beta}_{0,n}, \boldsymbol{\beta}_1, \boldsymbol{\sigma}) &= p(z_n|\boldsymbol{\lambda})\, p(y_{n,t}|z_n, \boldsymbol{\beta}_{0,n}, \boldsymbol{\beta}_1, \boldsymbol{\sigma}) \\ &= \sum_{c=1}^{C} \lambda_c\, p(y_{n,t}|\beta_{0,n,c}, \beta_{1,c}, \sigma_c) \\ &= \sum_{c=1}^{C} \lambda_c\, \mathit{Normal}(\mu_{c,n,t}, \sigma_c) \end{aligned} \tag{14}$$

where $\mu_{c,n,t}$ is defined by equation 12.

Model 2 — prior

$$\boldsymbol{\lambda}_n \sim \mathit{Dirichlet}(\boldsymbol{\alpha}) \tag{7}$$

where $\boldsymbol{\alpha}$ is a row vector of size C with $\alpha_c = 1$ (i.e., $\boldsymbol{\lambda}_n$ is assigned a proper flat prior).

$$\mu_{\beta_{0,c}} \sim \mathit{Normal}(0, 10) \tag{15}$$

where the hyperparameters are advised to be specified based on the observed dependent variable.

$$\sigma_{\beta_{0,c}} \sim \mathit{Normal}(0, 1)\,\mathbf{1}(\sigma_{\beta_{0,c}} > 0) \tag{16}$$

where the Normal distribution is truncated to the left at zero, modeled via the indicator function $\mathbf{1}(\sigma_{\beta_{0,c}} > 0)$. Furthermore, the standard deviation hyperparameter (here set to one) is advised to be specified based on the observed dependent variable.

Model 2 — prior — continued

$$\beta_{1,c} \sim Normal(0, 1) \tag{9}$$

where the hyperparameters are advised to be specified based on the observed dependent variable.

$$\sigma_c \sim Normal(0, 1)\, \mathbf{1}(\sigma_c > 0) \tag{10}$$

where the Normal distribution is truncated to the left at zero, modeled via the indicator function $\mathbf{1}(\sigma_c > 0)$. Furthermore, the standard deviation hyperparameter (here set to one) is advised to be specified based on the observed dependent variable.

Model 3

References

Koop, G. (2003). *Bayesian Econometrics*. Wiley.

Ram, N. & Grimm, K. J. (2009). Growth Mixture Modeling: A Method for Identifying Differences in Longitudinal Change Among Unobserved Groups. *International Journal of Behavioral Development*, *33*(6), 565-576.