Transition Models

$t_{ij}'s$ are assumed to be equally spaced.

Let $H_i = \{y_k, k = 1, \cdots, j-1\}$.

Consider

$$f(y_{ij} \mid H_{ij}, \alpha, \beta) = \exp\left\{\frac{y_{ij} - \psi(\theta_{ij})}{\phi} + c(y_{ij}, \phi)\right\},$$

where $\psi(\theta_{ij})$ and $c(y_{ij}, \phi)$ are known functions.

One has

$$\mu_{ij}^c = E\left[y_{ij} \mid H_{ij}\right] = \psi'\left(\theta_{ij}\right)$$

and

$$v_{ij}^c = V\left[y_{ij} \mid H_{ij}\right] = \psi''\left(\theta_{ij}\right)\phi$$

with

$$h(\mu_{ij}^c) = x_{ij}^T\beta + \sum_{r=1}^{s} f_r(H_{ij}; \alpha) \text{ for suitable functions } f_r(\cdot)'s,$$

and

$$v_{ij}^c = v(\mu_{ij}^c)\phi.$$

Problem 1. Fitting transition models: (A markov model of order q)

By

$$L_i(y_{i1}, \cdots, y_{im_i}) = f(y_{i1}, \cdots, y_{iq}) \prod_{j=q+1}^{m_i} f(y_{ij} \mid y_{i\,j-1}, \cdots, y_{i\,j-q}), i = 1, \cdots, n,$$

one can get the likelihood function

$$L(\alpha, \beta) = \prod_{i=1}^{n} f(y_{i1}, \cdots, y_{iq}) \prod_{j=q+1}^{m_i} f(y_{ij} \mid H_{ij}, \alpha, \beta),$$

where

$$H_{ij} = \{y_{i\,j-1}, \cdots, y_{i\,j-q}\}.$$

Since the term $f\left(y_{i1}, \cdots, y_{iq}\right)$ is always unavailable, the estimators of (α, β) are obtained via maximizing the conditional likelihood

$$\prod_{i=1}^{n} \prod_{j=q+1}^{m_i} f\left(y_{ij} \mid H_{ij}, \alpha, \beta\right).$$

Let $\theta = (\alpha, \beta)$.
Show that the log-conditional likelihood or conditional score function has the form

$$S^c(\theta) = \sum_{i=1}^{n} \sum_{j=(q+1)}^{m_i} \frac{\partial \mu_{ij}^c}{\partial \theta} {v_{ij}^c}^{-1}(y_{ij} - \mu_{ij}^c).$$

$$L^c(\theta) = \prod_{i=1}^{n} \prod_{j=q+1}^{m_i} f\left(y_{ij} \mid H_{ij}, \alpha, \beta\right).$$

$$l^c(\theta) = \ln L^c(\theta) = \frac{\sum_{i=1}^{n} \sum_{j=(q+1)}^{m_i} (y_{ij}\theta_{ij} - \psi(\theta_{ij}))}{\phi} + \sum_{i=1}^{n} \sum_{j=(q+1)}^{m_i} c(y_{ij}, \phi).$$

We have

$$\begin{aligned} S^c(\theta) = \frac{\partial l^c(\theta)}{\partial \theta} &= \frac{\sum_{i=1}^{n} \sum_{j=(q+1)}^{m_i} (y_{ij} - \psi'(\theta_{ij}))}{\phi} \\ &= \sum_{i=1}^{n} \sum_{j=(q+1)}^{m_i} \frac{1}{\phi}(y_{ij} - \mu_{ij}^c) \\ &= \sum_{i=1}^{n} \sum_{j=(q+1)}^{m_i} \frac{\partial \mu_{ij}^c}{\partial \theta} {v_{ij}^c}^{-1}(y_{ij} - \mu_{ij}^c), \end{aligned}$$

where

$$\begin{aligned} \mathbb{E}[y_{ij} \mid H_{ij}] = \psi'(\theta_{ij}) \triangleq \mu_{ij}^c, \\ \mathbf{V}[y_{ij} \mid H_{ij}] = \psi''(\theta_{ij})\phi = \frac{\partial \mu_{ij}^c}{\partial \theta}\phi \triangleq v_{ij}^c \implies \frac{1}{\phi} = \frac{\partial \mu_{ij}^c}{\partial \theta} {v_{ij}^c}^{-1}. \end{aligned}$$

Problem 2. Ordered Categorical data

Y: ordinal response with categories labeled $1, 2, \cdots, k$.
Let

$$F(a \mid x) = P(Y \leq a \mid x),$$

where $a = 1, \cdots, (k-1), x = \left(x_1, \cdots, x_p\right)^T$.
Proportional odds model:

$$\operatorname{logit} F(a \mid x) = \theta_a + x^T \beta, \quad a = 1, \cdots, (k-1).$$

Define $Y^* = \left(Y_1^*, \cdots, Y_{k-1}^*\right)$ with $Y_a^* = 1_{(Y \leq a)}$.
Then,

$$\operatorname{logit} F(a \mid x) = \operatorname{logit} P\left(Y_a^* = 1 \mid x\right).$$

Y	1	2	3	$\cdots$	$k-1$	k
Y_1^*	1	0	0	$\cdots$	0	0
Y_2^*	1	1	0	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$				$\vdots$
Y_{k-1}^*	1	1	1	$\cdots$	1	0

....

Example:
Assume that

$$\operatorname{logit} P\left(Y_j \leq b \mid Y_{i\,j-1} = a\right) = \theta_{ab} + {x_i}^T \beta_a, \quad a, b = 1, \cdots, (k-1).$$

It can be derived that

$$\operatorname{logit} P(Y_{ij} \leq b \mid Y_{i\,j-1}^* = y_{i\,j-1}^*) = \theta_b + \sum_{l=1}^{k-1} \alpha_{lb} y_{i(j-1)l}^* + x_{ij}^T (\beta + \sum_{l=1}^{k-1} r_l y_{i(j-1)l}^*),$$

$$\text{where} \begin{cases} \theta_{kb} = \theta_b, \\ \alpha_{lb} = \theta_{lb} - \theta_{l+1\,b}, \\ \beta_k = \beta, \\ r_l = \beta_l - \beta_{l+1} \end{cases}.$$

$$
\begin{aligned}
&Y_{ij} = a \\
\Rightarrow\ &Y_{ij}^* = (\underbrace{0 \cdots 0}_{a-1}\ \underbrace{1 \cdots 1}_{k-a}).
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{logit} P(Y_{ij} \le b \mid Y_{i\,j-1}^* = y_{i\,j-1}^*) = \theta_b &+ (\theta_{1b} - \theta_{2\,b}) y_{i(j-1)1}^* & +x_{ij}^T\{\beta &+ (\beta_1 - \beta_2) y_{i(j-1)1}^* \\
&+ (\theta_{2b} - \theta_{3\,b}) y_{i(j-1)2}^* & &+ (\beta_2 - \beta_3) y_{i(j-1)2}^* \\
&+ \cdots & &+ \cdots \\
&+ (\theta_{a-1\,b} - \theta_{ab}) y_{i(j-1)(a-1)}^* & &+ (\beta_{a-1} - \beta_a) y_{i(j-1)(a-1)}^* \\
&+ (\theta_{ab} - \theta_{a+1\,b}) y_{i(j-1)a}^* & &+ (\beta_a - \beta_{a+1}) y_{i(j-1)a}^* \\
&+ (\theta_{a+1\,b} - \theta_{a+2\,b}) y_{i(j-1)(a+1)}^* & &+ (\beta_{a+1} - \beta_{a+2}) y_{i(j-1)(a+1)}^* \\
&+ \cdots & &+ \cdots \\
&+ (\theta_{(k-1)b} - \theta_{k\,b}) y_{i(j-1)(k-1)}^* & &+ (\beta_{k-1} - \beta_k) y_{i(j-1)(k-1)}^* \}.
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{logit} P(Y_{ij} \le b \mid Y_{i\,j-1} = a) = \theta_b &+ (\theta_{ab} - \theta_{a+1\,b}) & +x_{ij}^T\{\beta &+ (\beta_a - \beta_{a+1}) \\
&+ (\theta_{a+1\,b} - \theta_{a+2\,b}) & &+ (\beta_{a+1} - \beta_{a+2}) \\
&+ \cdots & &+ \cdots \\
&+ (\theta_{(k-1)b} - \theta_{k\,b}) & &+ (\beta_{k-1} - \beta_k) \}.
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{logit} P(Y_{ij} \le b \mid Y_{i\,j-1} = a) &= \theta_b + \theta_{ab} - \theta_{kb} + x_{ij}^T\{\beta + \beta_a - \beta_k\} \\
&= \theta_{ab} + x_{ij}^T \beta_a.
\end{aligned}
$$

Problem 3. Log-linear transition models for count data

$Y_{ij} \mid \left(H_{ij}, x_{ij}\right) \sim$ Poissom $\left(\mu_{ij}^{c}\right)$.

. . . .

Model 1. Wong (1986) proposed that

$$\mu_{ij}^{c} = \exp\left(x_{ij}^{T}\beta\right)\{1 + \exp\left(-\alpha_0 - \alpha_1 y_{i\,j-1}\right)\},$$

$\alpha_0, \alpha_1 > 0$, where β is the influence of x_{ij} as $y_{i\,j-1} = 0$.

—

Remark. When $y_{i\,j-1} > 0, \mu_{ij}^{c}$ decreases as $y_{i\,j-1}$ increases. A negative association is implied between the prior and current responses.

. . . .

Model 2. $\mu_{ij}^{c} = \exp\left(x_{ij}^{T}\beta + \alpha y_{i\,j-1}\right)$.

—

Properties:

1. μ_{ij}^{c} increases as an exponential function of time as $\alpha > 0$.
2. When $\exp\left(x_{ij}^{T}\beta\right) = \mu$ and $\alpha < 0$, it leads to a stationary process.

$$\mathbb{E}[Y_{ij} \mid \left(H_{ij}, x_{ij}\right)] = Var[Y_{ij} \mid \left(H_{ij}, x_{ij}\right)] = \mu_{ij}^{c}.$$

We can obtain

$$\begin{aligned}\mathbb{E}[Y_{ij}] = \mathbb{E}[\mu_{ij}^{c}] &= \mathbb{E}[\exp\left(x_{ij}^{T}\beta + \alpha y_{i\,j-1}\right)] \\ &= \exp\left(x_{ij}^{T}\beta\right)\mathbb{E}[\exp\left(\alpha y_{i\,j-1}\right)] \\ &= \mu\,\mathbb{E}[\exp\left(\alpha y_{i\,j-1}\right)]\end{aligned}$$

If stationary,

$$\mu = \mu \exp\left(\alpha\mu\right)$$

Problem 4.

Moded 3.

$$\mu_{ij} = \exp\left(x_{ij}{}^T\beta + \alpha\left\{\ln\left(y^*_{i\,j-1}\right) - x^T_{i\,j-1}\beta\right\}\right),$$

where $y^*_{i\,j-1} = \max\left\{y_{i\,j-1}, d\right\}, 0 < d < 1$.

—

$$\text{Property:} \begin{cases} \alpha = 0: \text{ it reduces to an oedinary log-tinear model.} \\ \alpha < 0: \text{ negative correlation between } y_{i\,j-1} \text{ and } y_{ij} \\ \alpha > 0: \text{ positive correlation between } y_{i\,j-1} \text{ and } y_{ij} \end{cases}$$

—

Application to a size-independent branching process:
$\exp(x^T_{ij}\beta) = \mu$

$$\begin{aligned} y_{ij} &: \text{the number of individuals in the } i\text{-th population at generation } j \\ Z_k(y_{i\,j-1}) &: \text{the number of offspring for person } k \text{ in generation } (j-1) \end{aligned}$$

For $y_{i\,j-1} > 0$,

$$y_{ij} = \sum_{k=1}^{y_{i(j-1)}} Z_k(y_{i(j-1)}),$$

where

$$Z_k \overset{iid}{\sim} Poisson\left(\left(\frac{\mu}{y^*_{ij-1}}\right)^{1-\alpha}\right).$$

One can get

$$\mu^c_{ij} = \mu \cdot \left(\frac{y_{i(j-1)}}{\mu}\right)^{\alpha}.$$

Property:

- $\alpha < 0$: the sample paths oscillate back and forth about their long-term average level.
- $\alpha > 0$: the sample paths have sharper peaks and broader valleys.

width=!,height=!,pages=-