Transition Models

- Extension to GLM for describing the conditional distribution of Y_{ij} conditional to $Y_{i1}, \dots, Y_{ij-1}$ and covariates x_{ij}
- We consider the case where observation t_{ij} are equally spaced

Markov Models

Model conditional distribution of each response given past responses and covariates

$$g(E(Y_{ij}) \mid y_{ij-1}, \dots, y_{i1}, x_{ij}) = \text{function of } y_{ij-1}, \dots, y_{i1}, x_{ij}$$

treat past responses like additional predictors

Markov Models: Examples

- Linear Regression

$$\begin{aligned} E(Y_{ij} \mid \text{Past}_{ij}) &= x_{ij}\beta + \alpha_j(y_{ij-1} - x_{ij-1}\beta) \\ \alpha_j &= \exp(-\phi \mid t_{ij} - t_{ij-1} \mid) \end{aligned}$$

- Logistic Regression

$$\text{logit} Pr(Y_{ij} = 1 \mid \text{Past}_{ij}) = x_{ij}\beta + \alpha y_{ij-1}$$

- Log-Linear Regression

$$\begin{aligned} \log E(Y_{ij} \mid \text{Past}_{ij}) &= x_{ij}\beta + \alpha(\log y^*_{ij-1} - x_{ij-1}\beta) \\ y^*_{ij-1} &= \max(c, y_{ij-1}),\ c > 0 \end{aligned}$$

GLM Markov Models: Inference

- Maximum Likelihood
- GEE

Potential Applications

1. $y_{ij} = \begin{cases} 0 & \text{Pre-clinical stage} \\ 1 & \text{Tumor grade 1} \\ 2 & \text{Tumor grade 2} \\ 3 & \text{Tumor grade 3} \end{cases}$

   Study the efficacy of a new treatment protocol for slowing rate of progression

   Model $Pr(Y_{ij} \mid y_{ij-1}, x_{ij})$

2. y_{ij} = \$s earned by HMO in year j

   Wall street bankers want to predict earnings for the next year

   Model $E(Y_{ij} \mid y_{ij-1}, \dots, y_{i1}, x_{ij})$

Transition Model for Binary Responses

- logit $P(Y_{ij} \mid \text{Past}_{ij}) = \boldsymbol{x}_{ij}\boldsymbol{\beta} + \alpha y_{ij-1}$

  the change of respiratory infection at time t_{ij} depends on explanatory variables but also on whether or not the child had infection 3 months earlier

  $\boldsymbol{\beta}$ = *change* $\times$ *unit change in* x *in the log odds of infection, among children whit outcome* y_{ij-1} *at the prior visit*
- $e^{\alpha} = \frac{\exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta}+\alpha)}{\exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta})}$ = ratio of the odds of infection among children who did and did not have infection at the prior visit

Transition Models for Binary Data: equally spaced data

First Order Markov Chain

Transition Matrix: records the probabilities of making each of the possible transitions from one visit to the next

		y_{ij} 0	y_{ij} 1
y_{ij-1}	0	π_{00}	π_{01}
	1	π_{10}	π_{11}

- $\pi_{11} = Pr(Y_{ij} = 1 \mid y_{ij-1} = 1)$
- $\pi_{10} = Pr(Y_{ij} = 1 \mid y_{ij-1} = 0)$ trans. prob.
- $\pi_{ab} = Pr(Y_{ij} = a \mid y_{ij-1} = b)$

Regression Models for transition Probabilities

$$\boxed{\text{logit}Pr(Y_{ij} = 1 \mid y_{ij-1}) = x_{ij}\beta + y_{ij-1}x_{ij}\delta}$$

y_{ij-1}	Coef
0	β
1	$\beta + \delta$

Second Order Markov Models

y_{ij-2}	y_{ij-1}	y_{ij} 0	y_{ij} 1
0	0	π_{000}	π_{001}
0	1	π_{010}	π_{011}
1	0	π_{100}	π_{101}
1	1	π_{110}	π_{111}

As one Regression

$$\text{logit}Pr(Y_{ij} = 1 \mid y_{ij-1}, y_{ij-2}, x_{ij}) =$$
$$x_{ij}\beta + \delta_1 y_{ij-1}x_{ij} + \delta_2 y_{ij-2}x_{ij} + \delta_3 y_{ij-1}y_{ij-2}x_{ij}$$

y_{ij-2}	y_{ij-1}	Coefs
0	0	β
0	1	$\beta+\delta_1$
1	0	$\beta+\delta_2$
1	1	$\beta+\delta_1+\delta_2+\delta_3$

Indonesian children's study example

- Table 10.1 summarizes the number of transitions from respiratory disease status Y_{ij-1} to disease status Y_{ij}. These rates estimate the transition probabilities $P(Y_{ij} \mid Y_{ij-1})$
- Table 10.2 shows a cross tabulation of respiratory disease Y_{ij} against xerophthalmia status x_{ij}

  frequency of respiratory infection is $1.49 = .119/0.080$ times as high among children who are vitamin A deficient.
- There is correlation among repeated measurement on the same child. We can control for this by examining the effect of vitamin A deficiency separately for transitions starting with $Y_{ij-1} = 0$ or $Y_{ij-1} = 1$ (table 10.3)
- Among children free of infection at the prior visit, the frequency of respiratory disease is $1.44 = 0.108/0.075$ times higher if the child has xerophthalmia.
- Among children who suffered of infection at the prior visit, the xerophthalmia relative risk is $1.54 = 0.200/0.130$

- Therefore xerophthalmia effect is similar for $Y_{ij-1} = 0$ and $Y_{ij-1} = 1$ even though Y_{ij-1} is a strong predictor.

Statistical Model

$$\text{logit}Pr(Y_{ij} = 1 \mid Y_{ij-1} = y_{ij-1}) = \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta} + \alpha y_{ij-1}$$

- Results are in table 10.4 (GEE)
- a comparison between table 10.2 and table 10.3 indicates that the association of xerophthalmia and respiratory infection is similar for children who did and did not have respiratory infection at the previous visit. This is confirmed by the xerophthalmia-by-previous-infection interaction term in model 2 ($.11$ and 95% CI $(-2.1, 2.3)$
- Having controlled for age, season and respiratory infection at the prior visit, there is a mild evidence in these data for an association between xerophthalmia and respiratory infection; the xerophthalmia coeff is $0.78(0.53)$.
- Finally, with transition models, you must check whether the regression inference about $\boldsymbol{\beta}$ change with the model for time dependence, so we add Y_{ij-2} as predictor

- the inclusion of Y_{ij-2} reduce s the influence of season and Y_{ij-1} and increases the xerophthalmia coefficient to 1.73
- If the first order Markov assumption is valid, then the standard errors are valid
- Robust standard errors have a valid coverage in large sample size even when the Markov assumption is incorrect
- A simple check of the Markov assumption is to compare model based standard erros versus robust standard errors

Conclusion

- In transition models, explanatory variables and previous responses are treated symmetrically as predictors of the current response.
- Hence, as the time dependence model changes, so might inferences about explanatory variables
- sensitivity of inferences with respect to time dependence assumptions.