Reading summary:

Hidden Markov Models for Zero-Inflated Poisson Counts with an Application to Substance Use

Weijia Xiong

6/3/2020

Data and Objective

- Data

From an experiment performed on 45 cocaine-dependent participants that consist of daily diaries of cocaine use collected using the method of Time-Line Follow-Back (TLFB). 12 weeks prior and 4 weeks after completion of the 2-day stress and cue-reactivity study.

- Objective

Whether participation in a non-treatment outpatient clinical study is protective or detrimental(study participation alters drug seeking behavior or not) in terms of frequency of follow-up drug use among cocaine-dependent participants and to determine relevant demographic predictors of use parameters.

ZIP Hidden Markov Model for Poisson Counts

ZIP-HMM:

$$(Z_{it}|S_{it}) = \begin{cases} 0 & \text{w.p. } \pi_{it} \\ \text{Poisson } (\theta_{it}) & \text{w.p.} 1 - \pi_{it} \end{cases}$$

$$log(\theta_{it}) = \lambda_{0i} + \lambda_{1i} S_{it-1}$$

Poisson distribution:

$$f\left(Z_{it}|\theta_{it}, S_{it} = 1, X_{it}, X_{TRT_{it}}\right) = \frac{1}{z_{it}!} e^{-\theta_{it}} \theta_{it}^{z_{it}}$$

Transition probabilities:

$$\begin{aligned} p_{it}^{00} &= P\left(S_{it} = 0|S_{i,t-1} = 0, X_{it}, X_{TRT_{it}}\right) \\ p_{it}^{10} &= P\left(S_{it} = 1|S_{i,t-1} = 0, X_{it}, X_{TRT_{it}}\right) \\ p_{it}^{01} &= P\left(S_{it} = 0|S_{i,t-1} = 1, X_{it}, X_{TRT_{it}}\right) \\ p_{it}^{11} &= P\left(S_{it} = 1|S_{i,t-1} = 1, X_{it}, X_{TRT_{it}}\right) \end{aligned}$$

Transitioning or remaining high:

$$p_{it}^{1S_{(i,t-1)}} = \log\mathrm{it}^{-1}\left(\beta_0 + \beta_1 S_{i,t-1} + \beta_2 X_{it} + \beta_3 X_{TRT_{it}}\right)$$

Low-risk group:

$$\text{logit}\left(\pi_{it}|Y_{it}\right) = \alpha_0 + \alpha_1 Y_{it1} + \cdots + \alpha_j Y_{it}$$

Bayesian Inference

Estimation of exact posterior distribution:

- Gibbs sampler
- MCMC algorithms

Implement tool:

- WinBUGS

Comparison with other models

- conditional predictive ordinate statistic: $-log(CPO)_i$ (Smaller)
- log pseudo-marginal likelihood: LPML (Larger)

1. Nested 1-state models

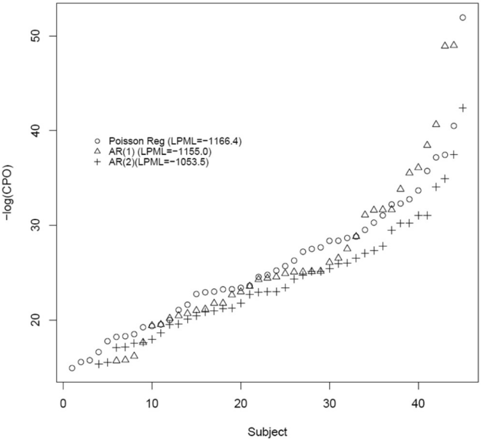


Figure 4.
Relative model fits of nested models based on -log(CPO_i). The value in parenthesis is the LPML (larger is better).

Comparison with other models

2. Second order HMM

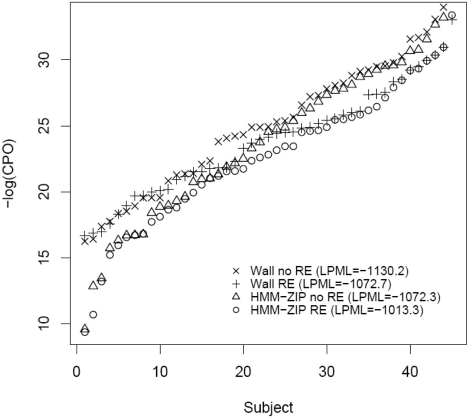
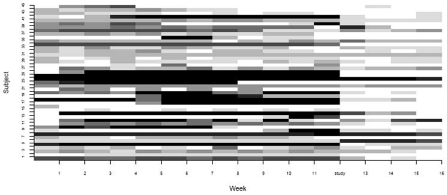
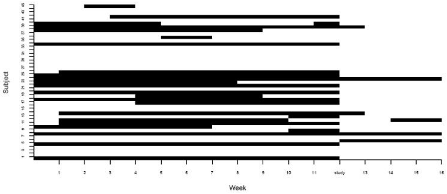


Figure 5.
Relative model fits of the 2-state hidden Markov models based on -log(CPO_i). The value in parenthesis is the LPML (larger is better).

Results

Top: the matrix of weekly drug use counts for each participant in the study.
Bottom: membership in the hidden state derived from posterior probabilities.





- $OR_{TRT_O} = 0.06[0.01, 0.19]$
- Highly preventative effect: study-benefit (seen after week 12) in reducing drug-seeking behavior

Question

1. ZIP vs ZINB
2. The number of state? '1-state','2-state','3-state'...
3. Bayesian approach vs likelihood-based frequentist approach
4. Deviance information criteria (DIC) is more commonly used for for Bayesian model selection? (Mentioned by the author for future work)