Journal of the Korean Statistical Society (1998), **27**: *3, pp 279–288*

Multiprocess Dynamic Poisson Models: The Covariates Case†

Joo Yong Shim and Joong Kweon Sohn[1]

ABSTRACT

We propose a multiprocess dynamic Poisson model for the analysis of Poisson process with the covariates. The algorithm for the recursive estimation of the parameter vector modeling time-varying effects of covariates is suggested. Also the algorithm for forecasting of numbers of events at the next time point based on the information gathered until the current time is suggested.

Keywords: Multiprocess dynamic generalized linear model; Covariate; Recursive estimation

1. INTRODUCTION

In the static case the mean of the Poisson distribution is constant over time, but in the dynamic case it is allowed to vary over time. Harvey and Durbin(1986) proposed a modified structural approach to the problem of estimating and forecasting in the dynamic case. West, Harrison and Migon(1985) developed the dynamic generalized linear model which allows the use of an one-dimensional exponential family observation distribution. The guide relationship relates the natural parameter at time i, denoted as λ_i, of the one-dimensional exponential family observation distribution to the parameter vector at time i, denoted as β_i, via $g(\lambda_i) = Z_i\beta_i$, where Z_i is the known (1×p) vector, β_i is the unknown (p×1) parameter vector and $g(\cdot)$ is a specified nonlinear function. The prior distribution is chosen to be conjugate family member having same first two moments as $g^{-1}(Z_i\beta_i)$. The evolution of the parameter vector is determined by the evolution equation, $\beta_i = G_i\beta_{i-1} + w_i$, where G_i is a known (p×p) matrix at time i and w_i is the evolution error vector. The distribution of w_i which is independent of β_i is

†This Study was Supported by the Basic Science Research Institute Program, Ministry of Education, 1997, Project No. BSRI-97-1403.

[1]Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea

specified in terms of the mean vector 0 and the variance-covariance matrix W_i. Harrison and Stevens(1971,1976) introduced the idea of the multiprocess model into the dynamic linear model to consider that there are K different models to be applied with K different probabilities at each time, where K different models differ in the distribution of the parameter vector. To estimate the mean and the variance-covariane matrix of the parameter vector, the mixture of K distributions should be collapsed into a single distribution. The method of collapsing the mixture of K distributions into a single distribution is based on equating first two moments of the mixture to first two moments of the single distribution, where the optimality criterion is minimizing the Kullerbeck-Liebler distance from the mixture to the single distribution. Bolstad(1988) developed the multiprocess dynamic generalized linear model by incorporating the multiprocess approach into the dynamic generalized linear model. Gamerman(1992) proposed the dynamic Poisson model by applying the dynamic generalized linear model to the Poisson process. Bolstad(1995) proposed the multiprocess dynamic Poisson model for the no-covariate case by applying the multiprocess dynamic idea to the Poisson process.

In this article we propose the multiprocess dynamic Poisson model for the covariates case to suggest the recursive estimation of the parameter vector modeling time-varying effects of covariates and the forecasting of the number of the next events. In the single dynamic model, the prior mean of the parameter vector at the next time is equal to the posterior mean of the parameter vector at the current time. But, in the proposed model, the prior mean of the parameter vector at the next time is not always equal to the posterior mean of the parameter vector at the current time, due to the consideration of K different prior distributions of the parameter vector at each time. Thus the proposed model can provide more reasonable prior distribution of the parameter vector than the single dynamic model. The proposed model is described in Section 2. The algorithms for recursive estimations of the parameter vector and forecasting the number of next events are provided in Section 3 and Section 4, respectively. The performance of estimation and forecasting is illustrated via the simulation study in Section 5.

2. MODEL DESCRIPTIONS

The number of events at time i is assumed to follow the Poisson distribution which has a time-varying parameter λ_i for $i = 1, 2, \cdots$. We denote it by

$$(Y_i|\lambda_i) \sim Poisson(\lambda_i),$$

where Y_i is the number of events at time i, λ_i is the mean of the Poisson population of Y_i, which is related to the parameter vector β_i modeling time-varying effects of covariates by the guide relationship. Let D_i be a set of information gathered until time i which can be represented as the set of numbers of events at previous time points including time i.

We define the perturbation of time i as the variation of distributions of the parameter vector β_i which are affected by different variance-covariane matrices of the evolution error vector w_i. Let α_i be the perturbation index variable at time i that determines which distribution of the parameter vector β_i is applied. When $\alpha_i = l$, the distribution of the parameter vector β_i is governed by the evolution equaton $\beta_i = \beta_{i-1} + w_i$, where w_i is the evolution error vector whose distribution is specified in terms of the mean vector 0 and the variance-covariance matrix $W_i^{(l)}$, $l = 1, 2, \cdots, K$. The prior index probabilty at time i which is the prior probability of a perturbation index variable of time i, $\pi_i^{(l)} = P(\alpha_i = l | D_{i-1})$, is assumed to be fixed prior to obtaining any information from observation the time i.

The multiprocess dynamic Poisson model for the covariates case is defined as follows.

i) Observation equation :

$$(Y_i | \lambda_i) \sim Poisson(\lambda_i) \quad \text{for } i = 1, 2, \cdots .$$

ii) Guide relationship:

$$\lambda_i = \exp(Z_i \beta_i) \quad \text{for} \quad i = 1, 2, \cdots ,$$

where Z_i is the known (1×p) covariate vector and β_i is the (p×1) parameter vector.

iii) Evolution equation:

$$\beta_i = \beta_{i-1} + w_i \quad \text{for} \quad i = 1, 2, \cdots ,$$

where w_i is the evolution error vector whose distribution is specified in terms of the mean vector 0 and the variance-covariance matrix which depends on the value of the perturbation index variable of time i such as the variance-covariance matrix of w_i given $\alpha_i = l$ is $W_i^{(l)}$, $l = 1, 2, \cdots, K$, and w_i is independent of the parameter vector β_{i-1}.

3. RECURSIVE ESTIMATIONS OF THE PARAMETER VECTOR

The process is started with the initial distribution of the parameter vector at time 0, β_0, specified in terms of the mean vector $\widehat{\beta}_0$ and the variance-covariance matrix V_0, where $\widehat{\beta}_0$ and V_0 are given prior to time 1.
At each time $i-1$, the posterior distribution of β_{i-1} given $\alpha_{i-1}=k$ is specified in terms of the mean vector $\widehat{\beta}_{i-1}^{(k)}$ and the variance-covariance matrix $V_{i-1}^{(k)}$, which is denoted by

$$(\beta_i|\alpha_{i-1}=k, D_{i-1}) \sim [\widehat{\beta}_{i-1}^{(k)}, V_{i-1}^{(k)}].$$

At time i, each of K posterior distributions of β_{i-1} obtained at time $i-1$ leads to K prior distributions of β_i as

$$(\beta_i|\alpha_{i-1}=k, \alpha_i=l, D_{i-1}) \sim [a_i^{(kl)}, R_i^{(kl)}],$$

where

$$a_i^{(kl)} = \widehat{\beta}_{i-1}^{(k)} \quad \text{and} \quad R_i^{(kl)} = V_{i-1}^{(k)} + W_i^{(l)}.$$

The joint prior distribution of β_i and $\log\lambda_i$ is obtained by the guide relationship,

$$\left(\begin{array}{c} \beta_i \\ \log\lambda_i \end{array} \mid \alpha_{i-1}=k, \alpha_i=l, D_{i-1} \right) \sim \left[\left(\begin{array}{c} a_i^{(kl)} \\ f_i^{(kl)} \end{array} \right), \left(\begin{array}{cc} R_i^{(kl)} & S_i^{(kl)} \\ S_i^{(kl)'} & q_i^{(kl)} \end{array} \right) \right], \tag{3.1}$$

where

$$f_i^{(kl)} = Z_i a_i^{(kl)}, \quad S_i^{(kl)} = Z_i R_i^{(kl)}, \quad q_i^{(kl)} = S_i^{(kl)} Z_i'.$$

Here the prior distribution of λ_i is assumed to be a conjugate gamma distribution $Ga(b_i^{(kl)}, r_i^{(kl)})$, where $b_i^{(kl)}$ and $r_i^{(kl)}$ are obtained in terms of $f_i^{(kl)}$ and $q_i^{(kl)}$ in (3.1) as, respectively, $q_i^{-1(kl)}$ and $q_i^{-1(kl)}\exp(-f_i^{(kl)})$.

With information from the observation, y_i, the posterior distribution of λ_i is obtained as

$$(\lambda_i|\alpha_{i-1}=k, \alpha_i=l, D_i) \sim Ga(b_i^{(kl)}+y_i, r_i^{(kl)}+1). \tag{3.2}$$

From (3.2) the mean and the variance of $log\lambda_i$ given $(\alpha_{i-1}=k, \alpha_i=l, D_i)$ are obtained as, respectively, $\gamma(b_i^{(kl)}+y_i) - \log(r_i^{(kl)})$ and $\dot{\gamma}(b_i^{(kl)}+y_i)$, where $\gamma(\cdot)$ is the digamma function.

Applying the linear Bayes estimation on (3.1) the posterior distribution of β_i given D_i is obtained as

$$(\beta_i|\alpha_{i-1}=k, \alpha_i=l, D_i) \sim [\widehat{\beta}_i^{(kl)}, V_i^{(kl)}],$$

where

$$\widehat{\beta}_i^{(kl)} = a_i^{(kl)} + S_i^{(kl)} q_i^{-1(kl)} \Big(\gamma(q_i^{-1(kl)} + y_i) - \log(1 + q_i^{-1(kl)} \exp(-f_i^{(kl)})) - f_i^{(kl)} \Big)$$

and

$$V_i^{(kl)} = R_i^{(kl)} - S_i^{(kl)} S_i^{(kl)'} q_i^{-2(kl)} \Big(\dot{\gamma}(q_i^{-1(kl)} + y_i) - q_i^{(kl)} \Big).$$

The posterior distribution of $(\beta_i|\alpha_i = l, D_i)$ is represented as the mixture of K posterior distributions of $(\beta_i|\alpha_{i-1} = k, \alpha_i = l, D_i)$ with the posterior index probability $p_i^{(kl)}$. Using that

$$\begin{aligned} & p(y_i|\alpha_{i-1} = k, \alpha_i = l, D_{i-1}) \\ &= \int p(y_i|\lambda_i, \alpha_{i-1} = k, \alpha_i = l, D_{i-1}) p(\lambda_i|\alpha_{i-1} = k, \alpha_i = l, D_{i-1}) d\lambda_i \\ &= \frac{\Gamma(y_i + b_i^{(kl)})}{\Gamma(b_i^{(kl)})\Gamma(y_i + 1)} (\frac{r_i^{(kl)}}{r_i^{(kl)} + 1})^{b_i^{(kl)}} (\frac{1}{r_i^{(kl)} + 1})^{y_i}, \end{aligned}$$

the posterior index probability is obtained as

$$\begin{aligned} p_i^{(kl)} &= P(\alpha_{i-1} = k, \alpha_i = l|D_i) \\ &\propto p(y_i|\alpha_{i-1} = k, \alpha_i = l, D_{i-1}) p_{i-1}^{(k)} \pi_i^{(l)}, \end{aligned}$$

where $p_{i-1}^{(k)} = P(\alpha_{i-1} = k|D_{i-1})$. Thus the posterior distribution of β_i given $\alpha_i = l$ and D_i has the mean vector $\widehat{\beta}_i^{(l)}$ and the variance-covariance matrix $V_i^{(l)}$, where

$$\widehat{\beta}_i^{(l)} = \sum_{k=1}^{K} \widehat{\beta}_i^{(kl)} p_i^{(kl)} / p_i^{(l)}$$

and

$$V_i^{(l)} = \sum_{k=1}^{K} [V_i^{(kl)} + (\widehat{\beta}_i^{(l)} - \widehat{\beta}_i^{(kl)})(\widehat{\beta}_i^{(l)} - \widehat{\beta}_i^{(kl)})'] p_i^{(kl)} / p_i^{(l)}.$$

Followed by the smoothing steps(West and Harrison, 1989), the smoothed distribution of β_i given $(\alpha_i = k, \alpha_{i+1} = l, D_N)$ for $i \leq N = 1, 2, \cdots$ is obtained in terms of the mean vector and the variance-covariance as respectively,

$$\widehat{\beta}_{i:N}^{(kl)} = E(\beta_i|\alpha_i = k, \alpha_{i+1} = l, D_N)$$

and

$$V_{i:N}^{(kl)} = V(\beta_i|\alpha_i = k, \alpha_{i+1} = l, D_N),$$

which lead to the mean vector and the variance-covariance of β_i given D_N, $i \leq N$, as respectively,

$$\widehat{\beta}_{i:N} = E(\beta_i|D_N) = \sum_{k,l}^{K} \widehat{\beta}_{i:N}^{(kl)} p_{i+1}^{(kl)} \tag{3.3}$$

and

$$V_{i:N} = V(\beta_i|D_N) = \sum_{k,l}^{K} [V_{i:N}^{(kl)} + (\widehat{\beta}_{i:N} - \widehat{\beta}_{i:N}^{(kl)})(\widehat{\beta}_{i:N} - \widehat{\beta}_{i:N}^{(kl)})'] p_{i+1}^{(kl)}. \tag{3.4}$$

4. FORECASTING OF THE NUMBER OF EVENTS

In this section we obtain the forecasted number of events in terms of $E[Y_{i+1}|D_i]$ based on information gathered until time i.

K prior distributions of β_{i+1} given D_i are obtained from each of K posterior distributions of β_i given D_i through the evolution equation, which are

$$(\beta_{i+1}|\alpha_i = k, \alpha_{i+1} = l, D_i) \sim [a_{i+1}^{(kl)}, R_{i+1}^{(kl)}], \quad k, l = 1, \cdots, K.$$

By the guide relationship the prior distribution of $\log \lambda_{i+1}$ is specified in terms of the mean and the variance such as, respectively, $f_{i+1}^{(kl)}$ and $q_{i+1}^{(kl)}$, where

$$f_{i+1}^{(kl)} = Z_{i+1} a_{i+1}^{(kl)} \quad \text{and} \quad q_{i+1}^{(kl)} = Z_{i+1} R_{i+1}^{(kl)} Z_{i+1}'. \tag{4.1}$$

Here the prior distribution of λ_{i+1} is assumed to be a conjugate gamma distribution $(b_{i+1}^{(kl)}, r_{i+1}^{(kl)})$. Note that $b_{i+1}^{(kl)}$ and $r_{i+1}^{(kl)}$ are obtained in terms of the mean and the variance of the distribution of $\log \lambda_{i+1}$ given D_i in (4.1) as, respectively, $q_{i+1}^{-1(kl)}$ and $q_{i+1}^{-1(kl)} \exp(-f_{i+1}^{(kl)})$. Then the mean and the variance of $(Y_{i+1}|\alpha_i = k, \alpha_{i+1} = l, D_i)$ is obtained as, respectively, $\widehat{\mu}_{i+1}^{(kl)}$ and $\widehat{v}_{i+1}^{(kl)}$, where

$$\widehat{\mu}_{i+1}^{(kl)} = E(\lambda_{i+1}|\alpha_i = k, \alpha_{i+1} = l, D_i) = \frac{b_{i+1}^{(kl)}}{r_{i+1}^{(kl)}}$$

and

$$\begin{aligned} \widehat{v}_{i+1}^{(kl)} &= E(\lambda_{i+1}|\alpha_i = k, \alpha_{i+1} = l, D_i) + V(\lambda_{i+1}|\alpha_i = k, \alpha_{i+1} = l, D_i) \\ &= \frac{b_{i+1}^{(kl)}}{r_{i+1}^{(kl)}} + \frac{b_{i+1}^{(kl)}}{r_{i+1}^{(kl)2}}. \end{aligned}$$

Thus we obtain the mean and the variance of Y_{i+1} given D_i, respectively, as $\widehat{\mu}_{i+1}$ and $\widehat{v}_{i+1}$, where

$$\widehat{\mu}_{i+1} = \sum_{k,l=1}^{K} \widehat{\mu}_{i+1}^{(kl)} p_i^{(k)} \pi_{i+1}^{(l)} \tag{4.2}$$

and

$$\widehat{v}_{i+1} = \sum_{k,l=1}^{K} [\widehat{v}_{i+1}^{(kl)} + (\widehat{\mu}_{i+1} - \widehat{\mu}_{i+1}^{(kl)})(\widehat{\mu}_{i+1} - \widehat{\mu}_{i+1}^{(kl)})] p_i^{(k)} \pi_{i+1}^{(l)}. \tag{4.3}$$

5. ILLUSTRATIONS

In this section we consider the performance of the estimation proposed in Section 3 and Section 4 via the simulation study. The data set consists of 100 simulated random samples from the Poisson population, whose mean is 4 for the first 25 samples, 8 for the next 25 samples, 4 for the next 25 samples, and 6 for last 25 samples.

The sampling distribution of the number of events at each time is assumed to follow the Poisson distribution such as,

$$(Y_i|\lambda_i) \sim Poisson(\lambda_i), \text{ for } i = 1, 2, \cdots, 100,$$

where Y_i is the number of events at time i, λ_i is the mean of the Poisson population at time i. In the model, the $\log \lambda_i$ is assumed to be the linear function of the parameter vector β_i such as $\lambda_i = \exp(Z_i \beta_i)$, with Z_i=($1, z_{1i}$) and β_i=$(\beta_{0i}, \beta_{1i})'$, where each value of z_{i1} for $i = 1, 2, \cdots, 100$ is randomly chosen from numbers between 1 and 2. The initial distribution is assumed to be

$$(\beta_0|D_0) \sim [\mathbf{0}, 10\mathbf{I}_2],$$

and the variance-covariance matrix of the evolution error vector is assumed to be

$$W_i^{(1)} = \begin{pmatrix} 0 & 0 \\ 0 & 0.01 \end{pmatrix}, \quad W_i^{(2)} = \begin{pmatrix} 0 & 0 \\ 0 & 1.0 \end{pmatrix},$$

which implies that β_{0i} is a time-constant parameter but β_{1i} is a time-varying parameter. We assume that there are two perturbations for the parameter β_{1i} at each time, steady state and sudden change numbered by 1 and 2, respectively. The model selection probabilities are assumed to be

$$\pi_i^{(1)} = 0.95, \quad \pi_i^{(2)} = 0.05, \quad i = 1, \cdots, 100.$$

Figure 5.1 shows the estimate of the mean of β_{1i} given D_{100}, where the estimate of the mean of β_{0i} given D_{100} is obtained as 1.663 , which are computed from (3.3) with $N = 100$. Figure 5.2 shows the posterior estimate and 95% CI for the mean of each sample from the Poisson population of changing means computed from (3.2). Figure 5.3 shows the observed value and 1-step ahead forecasted value of each sample computed from (4.2). In figures one can see that the estimates react to the changes in the mean quickly.

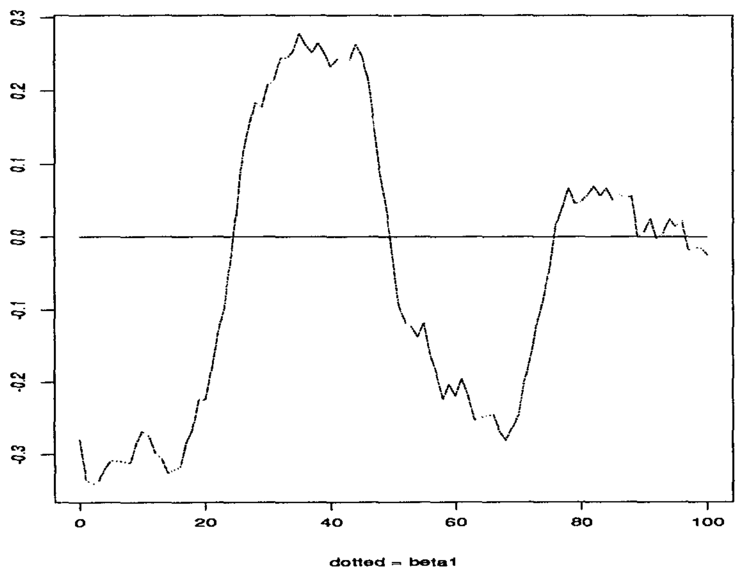

Figure 5.1: The Estimated Mean of The Parameter β_{1i} given D_{100}

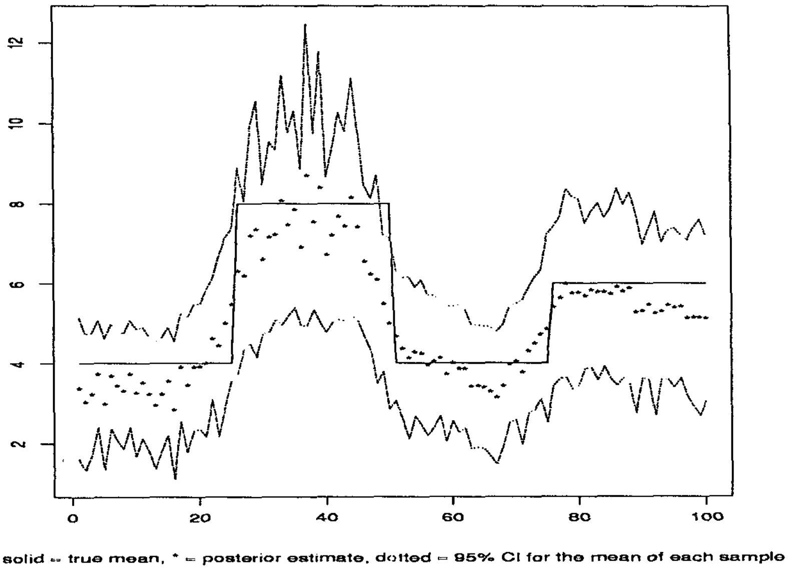

Figure 5.2: Posterior Estimate and 95% CI for Each Mean

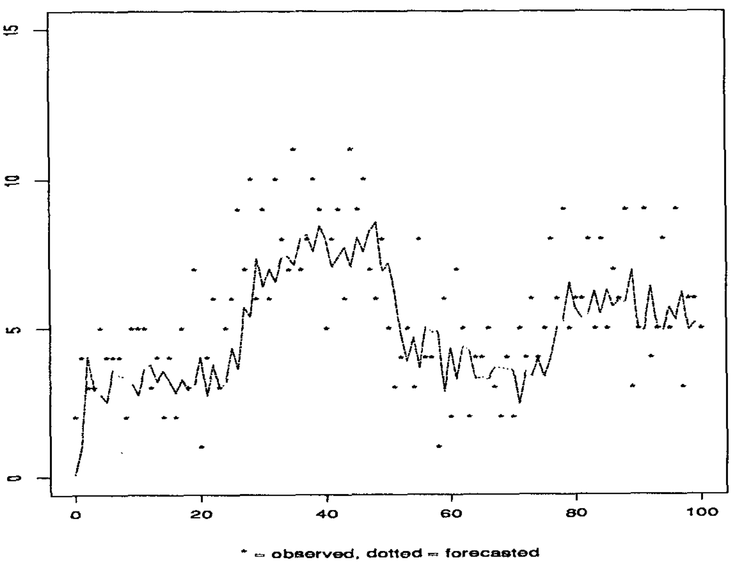

Figure 5.3: 1-Step Ahead Forecasted Value of Each Sample

References

Bolstad, W. M. (1988). "Estimation in the Multiprocess Dynamic Generalized Linear Model," *Communication in Statistics A: Theory and Methods*, 17, 4179-4204.

Bolstad, W. M. (1995). "The Multiprocess Dynamic Poisson Model," *Journal of the American Statistical Association*, 90, 227-232.

Gamerman, D. (1992). "A Dynamic Approach to the Statistical Analysis of Point Processes," *Biometrika* 79, 1, 39-50.

Harrison, P. J. and Stevens, C. F. (1971). "A Bayesian Approach to Short-Term Forecasting," *Operations Research Quarterly*, 22, 341-362.

Harrison, P. J. and Stevens, C. F. (1976). "Bayesian Forecasting(with discussions)," *Journal of the Royal Statistical Society* B, 38, 205-247.

Harvey, A. C. and Durbin, J. (1986). "The Effects of Seat Belt Legislation of Road Casualties: A Case Study in Time Series Modelling," *Journal of the Royal Statistical Society* A, 149, 187-227.

West, M., Harrison, P. J. (1989). *Bayesian Forecasting and Dynamic Models*, Springer-Verlag.

West, M., Harrison, P. J. and Migon, H. S. (1985). "Dynamic Generalized Linear Models and Bayesian Forecasting," *Journal of the American Statistical Association*, 80, 73-97.