The Korean Communications
in Statistics Vol. 4, No. 3, 1997
pp. 795 - 805

Bayesian Prediction under Dynamic Generalized Linear Models in Finite Population Sampling

Dal Ho Kim[1] and Sang Gil Kang[2]

Abstract

In this paper, we consider a Bayesian forecasting method for the analysis of repeated surveys. It is assumed that the parameters of the superpopulation model at each time follow a stochastic model. We propose Bayesian prediction procedures for the finite population total under dynamic generalized linear models. Some numerical studies are provided to illustrate the behavior of the proposed predictors.

1. Introduction

Often finite populations are subject to change in time. Usually the time series information is available through the repeated surveys, carried out at regular time intervals by industries and government agencies (e.g. market research survey, crop cutting experiments etc.). As mentioned in Scott and Smith (1974), Indian National Sample Survey, a multipurpose survey carried out annually since 1950 and the Medical Data Index in The U.K., a market research survey of doctors are examples of nonoverlapping surveys with human populations. In such surveys, one has at one's disposal not only the current data, but also data from similar past experiments.

In the past, standard time series methods have been applied to the analysis of repeated survey data from the random samples. For example, Blight and Scott(1973) and Scott and Smith(1974) have derived the estimates for the mean of a time dependent population using a first-order autoregressive model under the assumption that all the parameters of the model are known. Recently, Rodrigues and Bolfarine(1987) considered the prediction of the population total in a finite population using a Bayesian approach based on the Kalman Filter estimating algorithm. Bolfarine(1988) replaced the normality assumption by the more general exponential family of distributions, but he tried to build system equation directly to canonical parameter in

1) Assistant Professor, Department of Statistics, Kyungpook National University, Teagu, 702-701, Korea
2) Lecturer, Department of Statistics, Kyungpook National University, Teagu, 702-701, Korea

the exponential family of distributions so that his dynamic model has no state vector and oversimplified system equation.

In this paper, we consider a Bayesian forecasting under dynamic generalized linear models in sense of Harrison and Stevens(1976) and West, Harrison and Migon(1985) in the context of finite population sampling. Since the analysis under dynamic generalized linear models can not be performed exactly, we have used an approximation based on linear Bayes and conjugate priors. In Section 2, we develop the Bayesian prediction procedures under the dynamic generalized linear model for repeated surveys. Here the model is only partially specified in terms of their first and second moments. In Section 3, we develop the predictor for the normal, binomial and Poisson dynamic superpopulation model. Finally, Section 4 provides simulation studies to illustrate the behavior of the proposed predictors.

2. Bayesian Prediction under Dynamic Models

2.1 Dynamic Models in Repeated Surveys

Consider a finite population U with units labeled $1,\cdots,N$. Let y_i denote the value of a single characteristic attached to the unit i. It is considered in the sequel that y_i follows an univariate exponential family of distributions with superparameter η and a known scale parameter ϕ. That is, $f(y_i \mid \eta,\phi)=\exp[\phi(\eta t-a(\eta))+c(y,\phi)]$. We select a sample S of size n from the finite population to get information about the finite population total $T=\sum_{i=1}^{N} y_i$. Let $D=\{y_i, i\in S\}$ denotes the observed data from the finite population. Then the quantity T can be partitioned as $T=\sum_{i\in S} y_i+\sum_{i\notin S} y_i$. Thus predicting T is equivalent to predicting $\sum_{i\notin S} y_i$. Since

$$E[\sum_{i\notin S} y_i \mid D]=(N-n)E[\dot{a}(\eta) \mid D],$$

where $\dot{a}(\eta)=E[y_i \mid \eta,\phi]$, the Bayesian predictor of T (cf. Ericson(1969) and Bolfarine and Zacks(1992)) is given by

$$\hat{T}_B=E[T \mid D]=n\bar{y}+(N-n)E[\dot{a}(\eta) \mid D], \tag{2.1}$$

where $\bar{y}$ is the sample mean. To generalize the above formulation to dynamic models in repeated surveys, we replace y, T, D, η, n and $\hat{T}_B$ by $y_t, T_t, D_t, \eta_t, n_t$ and $\hat{T}_{D_t}$. Here $D_t=\{y_t, D_{t-1}\}$ represents all the relevant information set available at any time t.

A general Bayesian approach to the analysis of dynamic generalized linear models was given by West, Harrison and Migon(1985). Since y_t is assumed to have a sampling distribution in the exponential family, the density of y_t can be expressed as

$$p(y_t \mid \eta_t,\phi_t)=\exp[\phi_t(\eta_t y_t-a(\eta_t))+c(y_t,\phi_t)], \tag{2.2}$$

where η_t is the natural continuous parameter of the distribution, ϕ_t is a known scale parameter. The mean and variance of y_t are given by $\mu_t = E(y_t \mid \eta_t, \phi_t) = \dot{a}(\eta_t)$ and $Var(y_t \mid \eta_t, \phi_t) = \phi_t^{-1} \ddot{a}(\eta_t)$, respectively. Following West, Harrison and Migon(1985), we consider the dynamic generalized linear model for y_t in the context of repeated survey with the following components.

(i) Observation equation :

$$p(y_t \mid \eta_t, \phi_t) = \exp[\phi_t(\eta_t y_t - a(\eta_t)) + c(y_t, \phi_t)\ ],$$

$$g(\eta_t) = \lambda_t = h_t^T \theta_t, \tag{2.3}$$

(ii) Evolution equation :

$$\theta_t = G_t \theta_{t-1} + w_t, \qquad w_t \sim (0, W_t), \tag{2.4}$$

where θ_t is a n-dimensional state vector, h_t is a known n-dimensional regression vector, G_t is a known, $n \times n$ evolution matrix, w_t is a n-vector of evolution errors having zero mean and known variance-covariance matrix W_t, denoted by $w_t \sim (0, W_t)$, $\lambda_t = h_t^T \theta_t$ is a linear function of the state vector, and $g(\eta_t)$ is a known, continuous and monotonic function mapping η_t to the real line.

2.2 Bayesian Prediction Procedures

Given initial prior information D_0 at $t=0$, the information available at time t is simply

$$D_t = \{y_t, D_{t-1}\},$$

where y_t is observed value of the series at t. Here it is assumed that the initial prior for θ_0 is given as

$$(\theta_0 \mid D_0) \sim (m_0, C_0),$$

for some prior moments m_0 and C_0. The observation and evolution error sequences are assumed to be independent and mutually independent, and are independent of $(\theta_0 \mid D_0)$.

2.2.1 Evolution Step

In this step, evolving to time t, we find the prior distributions about θ_t and λ_t at both $t-1$ and t. For the estimation, the initial conditions require that the first and second moments for posterior distribution of θ_{t-1} given D_{t-1} is known at time t. This distribution is

$$(\theta_{t-1} \mid D_{t-1}) \sim (m_{t-1}, C_{t-1}).$$

By using the posterior distribution of θ_{t-1} and evolution equation (2.4), we obtain the moments of θ_t given D_{t-1}. Thus the mean vector and variance-covariance matrix are

$$a_t = E[\theta_t \mid D_{t-1}] = G_t m_{t-1}$$

and

$$R_t = Var[\theta_t \mid D_{t-1}] = G_t C_{t-1} G_t^T + W_t,$$

respectively. Hence the prior distribution for θ_t given D_{t-1} is

$$(\theta_t \mid D_{t-1}) \sim (a_t, R_t). \tag{2.5}$$

By using the prior distribution (2.5) and the observation equation (2.3), we obtain the joint distribution of θ_t and λ_t. Since $E[\theta_t \mid D_{t-1}]$ and $V[\theta_t \mid D_{t-1}]$ can be calculated in prior distribution (2.5), we only need to calculate $E[\lambda_t \mid D_{t-1}]$, $Var[\lambda_t \mid D_{t-1}]$ and $Cov[\theta_t, \lambda_t \mid D_{t-1}]$. Thus the moments are

$$f_t = E[\lambda_t \mid D_{t-1}] = h_t^T a_t$$

and

$$q_t = Var[\lambda_t \mid D_{t-1}] = h_t^T R_t h_t,$$

respectively. The covariance of θ_t and λ_t given D_{t-1} is

$$Cov[\theta_t, \lambda_t \mid D_{t-1}] = h_t^T R_t.$$

Hence the joint distribution of θ_t and λ_t given D_{t-1} is

$$\begin{pmatrix} \theta_t \\ \lambda_t \end{pmatrix} \mid D_{t-1} \Bigg) \sim \left[\begin{pmatrix} a_t \\ f_t \end{pmatrix}, \begin{pmatrix} R_t & h_t^T R_t \\ R_t h_t & q_t \end{pmatrix} \right]. \tag{2.6}$$

By using the method of linear Bayes estimation, the moments of the conditional distribution of θ_t given λ_t are directly obtained in (2.6). Therefore the conditional distribution of θ_t given λ_t is

$$(\theta_t \mid \lambda_t, D_{t-1}) \sim [(a_t + R_t h_t q_t^{-1}(\lambda_t - f_t)), (R_t - R_t h_t q_t^{-1} h_t^T R_t)].$$

2.2.2 Updating Step

In this step we update the prior distribution of the parameter given the observation y_t. Assume that the prior distribution $(\eta_t \mid D_{t-1})$ has the conjugate prior distribution $CP(r_t, s_t)$. That is, $f(\eta_t \mid D_{t-1}) = \exp[c(r_t, s_t) + \eta_t r_t - s_t a(\eta_t)]$. The parameters r_t and s_t are chosen to be consistent with the moments for λ_t in joint distribution (2.6). That is,

$$E[g(\eta_t) \mid D_{t-1}] = f_t \text{ and } Var[g(\eta_t) \mid D_{t-1}] = q_t.$$

The relationship between the moments of η_t and the moments of λ_t is called the guide relationship by West, Harrison and Migon(1985). Now the joint distribution of y_t and η_t is

$$p(y_t, \eta_t \mid D_{t-1}) = \exp[c(y_t, \phi_t) + c(r_t, s_t) + \eta_t(r_t + y_t\phi_t) - (s_t + \phi_t)a(\eta_t)]$$

and the marginal distribution of y_t is

$$p(y_t \mid D_{t-1}) = \exp[c(y_t, \phi_t) + c(r_t, s_t) - c(r_t + \phi_t y_t, s_t + \phi_t)].$$

Thus posterior distribution of η_t given D_t is

$$p(\eta_t \mid D_t) = \exp[c(r_t + \phi_t y_t, s_t + \phi_t) + \eta_t(r_t + \phi_t y_t) - (s_t + \phi_t)a(\eta_t)],$$

that is, the posterior distribution of η_t given D_t is the conjugate posterior $CP(r_t^*, s_t^*)$ where $r_t^* = r_t + \phi_t y_t$ and $s_t^* = s_t + \phi_t$. Now we find the moments of the posterior distribution of θ_t by using the moments of λ_t. Since

$$E[\theta_t \mid D_t] = E[E(\theta_t \mid \lambda_t, D_{t-1}) \mid D_t]$$

and

$$Var[\theta_t \mid D_t] = E[Var(\theta_t \mid \lambda_t, D_{t-1}) \mid D_t] + Var[E(\theta_t \mid \lambda_t, D_{t-1}) \mid D_t].$$

We obtain the mean vector and variance-covariance matrix as

$$m_t = E[\theta_t \mid D_t] = a_t + R_t h_t q_t^{-1}(f_t^* - f_t)$$

and

$$C_t = Var[\theta_t \mid D_t] = R_t - R_t h_t (q_t^{-1} - q_t^{-2} q_t^*) h_t^T R_t,$$

where $f_t^* = E[g(\eta_t) \mid D_t]$ and $q_t^* = Var[g(\eta_t) \mid D_t]$. Therefore when y_t is observed, the posterior distribution of θ_t given D_t is

$$(\theta_t \mid D_t) \sim (m_t, C_t).$$

This completes the determination of the posterior distributions of θ_t.

2.2.3 Prediction Step

In this step, we predict the population total by using the posterior distribution of η_t given D_t. Since $\hat{T}_B = n\bar{y} + (N-n)E[\dot{a}(\eta) \mid D]$ for T, T_t can be predicted by

$$\hat{T}_{D_t} = n_t \bar{y}_t + (N_t - n_t)\widehat{\dot{a}}(\eta_t), \tag{2.7}$$

where $\widehat{\dot{a}}(\eta_t) = E[\dot{a}(\eta_t) \mid D_t]$ calculated from the posterior distribution of η_t and $\bar{y}_t$ is the mean of the sample S_t of size n_t selected at time t. Also at time t the posterior variance of T_t is given by

$$Var[T_t \mid D_t] = (N_t - n_t)^2 Var[\dot{a}(\eta_t) \mid D_t] + (N_t - n_t)E[\ddot{a}(\eta_t)/\phi_t \mid D_t]. \tag{2.8}$$

Note that prediction of T_t depends on the posterior moments of the function $\dot{a}(\eta_t)$ and $\ddot{a}(\eta_t)$. Also these moments are easily and explicitly computed for most exponential families and thus leading to explicit expressions for $\hat{T}_{D_t}$.

At this point, we are in the same position as we were when we started the prediction procedure, so we are ready to repeat the prediction process when the time index from $t-1$ to t.

3. Dynamic Superpopulation Models

3.1 Normal Dynamic Superpopulation Model

We consider that the finite population at each time generated according to a normal superpopulation model with known variances. More specifically, at time t, y_{it} is normally distributed with mean μ_t and variance V_t. That is, $y_{it} \sim N(\mu_t, V_t)$ with $\mu_t = \eta_t$, $i=1,\cdots,N_t$, $t=1,\cdots,T$. So for the observed sample, $\bar{y}_t$ is normally distributed, i.e., $(\bar{y}_t \mid \eta_t) \sim N(\mu_t, V_t/n_t)$. Here $\dot{a}(\eta_t)=\eta_t=\mu_t$, $\phi_t = n_t/V_t$ and $\bar{y}_t = \sum_{i \in S_t} y/n_t$ stands for the sample mean at time t.

The dynamic model is obtained by $g(\eta_t)=\eta_t$ so that $\mu_t=\eta_t=\lambda_t=h_t^T\theta_t$. We work in terms of the μ_t notation. At time $t-1$ the dynamic model is completely the posterior distribution of θ_{t-1} given D_{t-1}, that is,

$$(\theta_{t-1} \mid D_{t-1}) \sim N(m_{t-1}, C_{t-1}).$$

Now the prediction procedures under the normal dynamic superpopulation model is as follows:

(i) The joint distribution of μ_t and θ_t given D_{t-1} is

$$\left(\begin{matrix}\theta_t \\ \mu_t\end{matrix} \mid D_{t-1}\right) \sim \left[\left(\begin{matrix}a_t \\ f_t\end{matrix}\right), \left(\begin{matrix}R_t & h_t^T R_t \\ R_t h_t & q_t\end{matrix}\right)\right], \tag{3.1}$$

where $a_t = G_t m_{t-1}$, $R_t = G_t C_{t-1} G_t^T + W$, $f_t = h_t^T a_t$ and $q_t = h_t^T R_t h_t$.

(ii) When we observed $\bar{y}_t$, the posterior distribution of μ_t is

$$(\mu_t \mid D_t) \sim N(f_t^*, q_t^*), \tag{3.2}$$

where $f_t^* = f_t + \frac{q_t}{q_t + V_t/n_t}(\bar{y}_t - f_t)$ and $q_t^* = q_t - \frac{q_t^2}{q_t + V_t/n_t}$ and the posterior distribution of θ_t is

$$(\theta_t \mid D_t) \sim N(m_t, C_t), \tag{3.3}$$

where $m_t = a_t + R_t h_t (f_t^* - f_t)/q_t$ and $R_t = R_t - R_t h_t h_t^T R_t (1 - q_t^*/q_t)/q_t$.

(iii) The predictor of population total is

$$\begin{aligned}\hat{T}_{D_t} &= n_t \bar{y}_t + (N_t - n_t) E[\dot{a}(\eta_t) \mid D_t] \\ &= n_t \bar{y}_t + (N_t - n_t) f_t^*\end{aligned}$$

and the posterior variance of population total is

$$\begin{aligned}Var[T_t \mid D_t] &= (N_t - n_t)^2 Var[\dot{a}(\eta_t) \mid D_t] + (N_t - n_t) E[\ddot{a}(\eta_t)/\phi_t \mid D_t] \\ &= (N_t - n_t)^2 q_t^* + (N_t - n_t) V_t.\end{aligned}$$

Since the usual expansion predictor at time t is $\hat{T}_{E_t}=N_t\,\bar{y}_t$, the variance of the predictor is

$$\begin{aligned} Var[\ \hat{T}_{E_t}-T_t\ ] &= (N_t-n_t)^2 V_t/n_t+(N_t-n_t)V_t \\ &\geq Var[\ T_t \mid D_t\ ]. \end{aligned}$$

Thus there is some improvement of $\hat{T}_{D_t}$ over $\hat{T}_t$.

3.2 Binomial Dynamic Superpopulation Model

In this case, the random quantity y_{it} associated with unit i at time t is such that $P[y_{it}=1\ ]=\pi_t=1-P[y_{it}=0\ ]$. The case $y_{it}=1$ corresponds to a success. So at time t, $\bar{y}_t$, the average number of successes in S_t, follows the exponential sampling model with $\eta_t=\log[\pi_t/(1-\pi_t)\ ]$, $\phi_t=n_t$ and $a(\eta_t)=\log(1+e^{\eta_t})$, $t=1,\cdots,T$. The dynamic model is obtained by $g(\eta_t)=\eta_t$ so that $\eta_t=\lambda_t=h_t^T\theta_t$ is given by $\eta_t=\lambda_t=\log[\pi_t/(1-\pi_t)\ ]=h_t^T\theta_t$. At time $t-1$ the dynamic model is completely the posterior distribution of θ_{t-1} given D_{t-1}, that is,

$$(\theta_{t-1} \mid D_{t-1}) \sim (m_{t-1}, C_{t-1}).$$

The prediction procedures under the binomial dynamic superpopulation model is as follows:

(i) The joint distribution of θ_t and λ_t given D_{t-1} is the same as (3.1).

(ii) The conjugate prior, $CP(r_t, s_t)$, is a beta form for $\pi_t=(1+e^{-\eta_t})$. Here using the mode and curvature of $(\eta_t \mid D_{t-1})$ for f_t and q_t^{-1} leads to $r_t=q_t^{-1}(1+e^{f_t})$ and $s_t=q_t^{-1}(1+e^{f_t})^2 e^{-f_t}$. When we observed $\bar{y}_t$, the posterior distribution of η_t is

$$(\eta_t \mid D_t) \sim (f_t^*, q_t^*),$$

where $f_t^*=\log[r_t^*/(s_t^*-r_t^*)]$, $q_t^*=1/r_t^*+1/(s_t^*-r_t^*)$, $r_t^*=r_t+n_t\,\bar{y}_t$ and $s_t^*=s_t+n_t$. Note that the posterior distribution of θ_t is the same as (3.3).

(iii) Since

$$E[\dot{a}(\eta_t) \mid D_t\ ]=\frac{r_t^*}{s_t^*} \quad \text{and} \quad Var[\dot{a}(\eta_t) \mid D_t\ ]=\frac{r_t^*(s_t^*-r_t^*)}{s_t^{*2}\,(s_t^*+1)},$$

the predictor of population total is given by

$$\hat{T}_{D_t}=n_t\,\bar{y}_t+(N_t-n_t)\frac{r_t^*}{s_t^*}.$$

Also the posterior variance of population total is given by

$$Var[\ T_t \mid D_t\ ]=(N_t-n_t)^2\frac{r_t^*(s_t^*-r_t^*)}{s_t^{*2}\,(s_t^*+1)}+(N_t-n_t)\frac{r_t^*(s_t^*-r_t^*)}{s_t^*(s_t^*+1)}.$$

Since the usual expansion predictor at time t is $\hat{T}_{E_t} = N_t \bar{y}_t$, the variance of the predictor is

$$Var[\ \hat{T}_{E_t} - T_t\ ] = (N_t - n_t)^2 \pi_t (1-\pi_t)/n_t + (N_t - n_t)\pi_t(1-\pi_t).$$

Note that as $W_t \to \infty, q_t^{-1} \to 0$ and so $\hat{T}_{D_t} = \hat{T}_t$. If $\bar{y}_t$ replaces π_t in the above, then usually $Var[\ \hat{T}_{E_t} - T_t\ ] \geq Var[\ T_t \mid D_t\ ]$.

3.3 Poisson Dynamic Superpopulation Model

In this case, the random quantity y_{it} associated with unit i at time t have the Poisson distribution with superpopulation parameter $\pi_t, i = 1, \cdots, T$. Thus at time t, $\bar{y}_t$, the mean of the y values in the selected sample S_t, follows the exponential sampling model with $\eta_t = \log(\pi_t)$, $\phi_t = n_t$ and $a(\eta_t) = \exp(\eta_t)$, $t = 1, \cdots, T$. The dynamic model is obtained by $g(\eta_t) = \eta_t$ so that $\eta_t = \lambda_t = h_t^T \theta_t$ is given by $\eta_t = \lambda_t = \log(\pi_t) = h_t^T \theta_t$. At time $t-1$ the dynamic model is completely the posterior distribution of θ_{t-1} given D_{t-1}, that is,

$$(\theta_t \mid D_{t-1}) \sim (m_{t-1}, C_{t-1}).$$

Now the prediction procedures under the Poisson dynamic superpopulation model is as follows:

(i) The joint distribution of θ_t and λ_t given D_{t-1} is the same as (3.1).

(ii) The conjugate prior, $CP(r_t, s_t)$, is a log-gamma form for $\eta_t = \log(\pi_t)$. Here using the mode and curvature of $(\eta_t \mid D_{t-1})$ for f_t and q_t^{-1} leads to $r_t = q_t^{-1}$ and $s_t = q_t^{-1} e^{-f_t}$. When we observed $\bar{y}_t$, the posterior distribution of η_t is

$$(\eta_t \mid D_t) \sim (f_t^*, q_t^*),$$

where $f_t^* = \log(r_t^*/s_t^*)$, $q_t^* = 1/r_t^*$, $r_t^* = r_t + n_t \bar{y}_t$ and $s_t^* = s_t + n_t$. Note that the posterior distribution of θ_t is the same as (3.3).

(iii) Since

$$E[\ \dot{a}(\eta_t) \mid D_t\ ] = \frac{r_t^*}{s_t^*} \quad \text{and} \quad Var[\ \dot{a}(\eta_t) \mid D_t\ ] = \frac{r_t^*}{s_t^{*2}},$$

the predictor of population total is given by

$$\hat{T}_{D_t} = n_t \bar{y}_t + (N_t - n_t)\frac{r_t^*}{s_t^*}.$$

Also the posterior variance of population total is given by

$$Var[\ T_t \mid D_t\ ] = (N_t - n_t)^2 \frac{r_t^*}{s_t^{*2}} + (N_t - n_t)\frac{r_t^*}{s_t^*}.$$

Since the usual expansion predictor at time t is $\hat{T}_{E_t} = N_t \bar{y}_t$, the variance of the predictor is

$$Var[\ \hat{T}_{E_t} - T\ ] = (N_t - n_t)^2 \pi_t / n_t + (N_t - n_t)\pi_t.$$

Note that as $W_t \to \infty, q_t^{-1} \to 0$ and so $\hat{T}_{D_t} = \hat{T}_t$. If $\bar{y}_t$ replaces π_t in the above, then usually $Var[\ \hat{T}_{E_t} - T_t\ ] \geq Var[\ T_t \mid D_t]$.

4. Simulation Studies

In order to illustrate the behavior of the proposed predictors, we consider the following one specific model for our simulation studies.

$$y_t = \mu_t + v_t,$$

$$\mu_t = \mu_{t-1} + \beta_t + \delta\mu_t,$$

$$\beta_t = \beta_{t-1} + \delta\beta_t,$$

where the zero-mean, evolution errors $\delta\mu_t$ and $\delta\beta_t$ uncorrelated. This model can be rewritten in the form

$$y_t = h_t^T \theta_t + v_t,$$

$$\theta_t = G_{t-1} \theta_{t-1} + w_t,$$

where

$$h_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \theta_t = \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix},\ G_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},\ w_t = \begin{bmatrix} \delta\mu_t + \delta\beta_t \\ \delta\beta_t \end{bmatrix}$$

and v_t and w_t are distributed with zero means.

The data of the normal superpopulation model is generated as follow. Starting with $\beta_0 = 0$, the β_t were generated with the errors $\delta\beta_t$ generated according to the normal distribution with zero mean and variance 1. Next starting with $\mu_0 = 0.0$, the μ_t were generated with the errors $\delta\mu_t$ generated according to the normal distribution with zero mean and variance 1. Next for each μ_t a population of size $N_t = 100$ with the errors v_t generated according to the normal distribution with zero mean and variance 1 were generated. From this population, a sample of size $n_t = 10$ was selected at random, without replacement, $t = 1, \cdots, 50$. Figure 4.1 represents the performance of predictor $\hat{T}_{D_t}$, which the sequences $\hat{T}_{D_t}$ follows very closely the sequence T_t.

The binomial superpopulation model is generated as follow. Starting with $\beta_0 = 0.0$, the β_t were generated with the errors $\delta\beta_t$ generated according to the normal distribution with zero mean and variance 1. Next starting with $\mu_0 = 0.0$, the μ_t were generated with the errors $\delta\mu_t$ generated according to the normal distribution with zero mean and variance 1. Next for each $\pi_t = e^{\eta_t}/(1+e^{\eta_t})$ generated, a population of size $N_t = 100$ was generated. From this population, a sample of size $n_t = 10$ was selected at random, without replacement, $t = 1, \cdots, 50$. Figure 4.2 represents the performance of predictor $\hat{T}_{D_t}$, which the sequences $\hat{T}_{D_t}$ follows somewhat closely the sequence T_t.

Similarly, the Poisson superpopulation model is generated as follow. Starting with $\beta_0 = 0.0$, the β_t were generated with the errors $\delta\beta_t$ generated according to the normal distribution with zero mean and variance 1. Next starting with $\mu_0 = 0.0$, the μ_t were generated with the errors $\delta\mu_t$ generated according to the normal distribution with zero mean and variance 1. Next for each $\pi_t = e^{\eta_t}$ generated, a population of size $N_t = 100$ was generated. From this population, a sample of size $n_t = 10$ was selected at random, without replacement, $t = 1, \cdots, 50$. Figure 4.3 represents the performance of predictor $\hat{T}_{D_t}$, which the sequences $\hat{T}_{D_t}$ follows quite closely the sequence T_t.

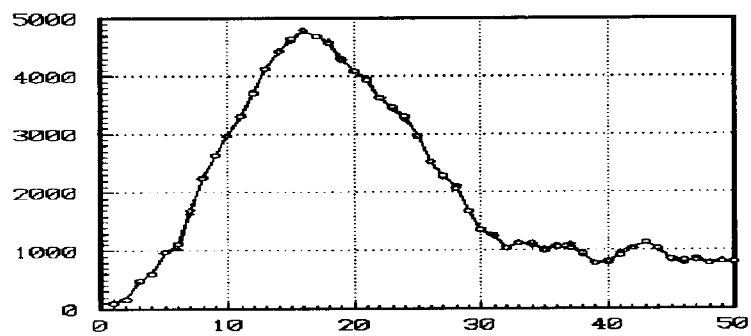

Figure 4.1 Population Total(◦) and Predictor Value(•) : Normal Superpopulation Model

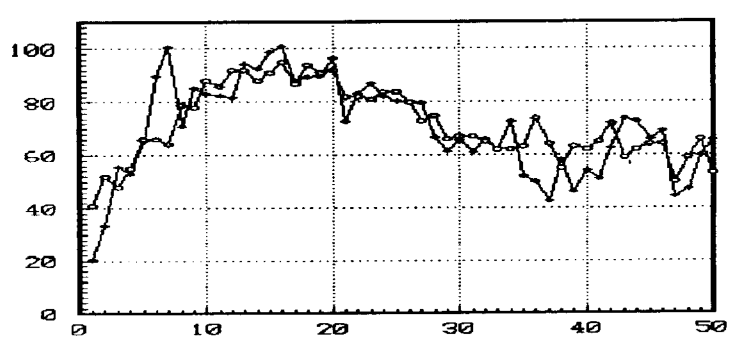

Figure 4.2 Population Total(◦) and Predictor Value(•) : Binomial Superpopulation Model

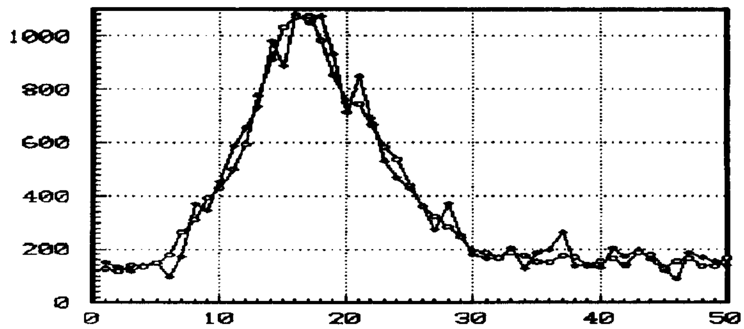

Figure 4.3 Population Total(◦) and Predictor Value(•) : Poisson Superpopulation Model

References

[1] Blight, B.J.N. and Scott, A.J. (1973). A Stochastic Model for Repeated Surveys. *Journal of the Royal Statistical Society, Series B,* 35, 61-66.

[2] Bolfarine, H. (1988). Finite Population Prediction under Dynamic Generalized Linear Models. *Communication in Statistics-Computation and Simulation,* 17, 187-208.

[3] Bolfarine, H. and Zacks, S. (1992). *Prediction Theory for Finite Populations.* Springer Verlag, New York.

[4] Ericson, W.A. (1969). Subjective Bayesian Models in Sampling Finite Populations. *Journal of the Royal Statistical Society, Series B,* 31, 195-233.

[5] Harrison, P.J. and Stevens, C.F. (1976). Bayesian Forecasting. *Journal of the Royal Statistical Society, Series B,* 38, 205-247.

[6] Rodrigues, J. and Bolfarine, H. (1987). A Kalman Filter Model for Single and Two-stage Repeated Surveys. *Statistics and Probability Letters,* 5, 299-303.

[7] Scott, A.J. and Smith, A.F.M. (1974). Analysis of Repeated Surveys Using Time Series Methods. *Journal of the American Statistical Association,* 69, 674-678.

[8] 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.