Listy Biometryczne – Biometrical Letters
Vol. 32 (1995), No. 1, 15-21

Bayesian estimation of a structural changes model with dependent Bernoulli variable

Krystyna Pruska

Chair of Statistical Methods, University of Łódź,
41, Rewolucji 1905r., 90-214 Łódź

Abstract

In this paper a structural changes model is considered. Two Bernoulli variables are in the model. These variables occur with the probability λ and $1-\lambda$, respectively. The p_1 and p_2 are distribution parameters of the Bernoulli variables.

The Bayesian estimators of this model parameters are derived in two cases:

1) when λ is known and p_1, p_2 are unknown,

2) when λ is unknown and p_1, p_2 are known.

The prior distributions of λ, p_1, p_2 are uniform distributions on the interval (0,1).

1. Introduction

In this paper a special structural changes model with dependent Bernoulli variable and the Bayesian estimation of its parameters are considered[1)]. The formulas for these Bayesian estimators are derived.

This model can be used in biological, medical, economic, social and other research.

Key words: Bayesian estimation, prior and posterior distributions, maximum likelihood method

[1)] The switching regression models are structural changes models too, but their forms are quite different than the model which is proposed in this paper (see for example: Booth and Smith, 1982; Ferreira, 1975; Lubrano, 1985; Pruska, 1994; Tsurumi, 1982).

2. The idea of Bayesian estimation

The idea of Bayesian estimation may be described as follows (see for example Czekała and Dziechciarz, 1990). Let y be an investigated variable in a population. We assume that θ is unknown parameter of y and θ has the prior distribution $g(\theta)$. Let $y_1,\ldots,y_T$ be an independent sample from the population. The Bayesian estimator of parameter θ is the expectation of the posterior distribution of θ for given $y_1,\ldots,y_T$ and for a quadratic loss function. This posterior distribution has the following form:

$$g(\theta|\mathbf{y}) = \frac{g(\theta)L(\mathbf{y}|\theta)}{\int_{\Theta} g(\theta)L(\mathbf{y}|\theta)d\theta}, \tag{2.1}$$

where Θ is the parameter space and $L(\mathbf{y}|\theta)$ is the likelihood for the independent sample $\mathbf{y} = (y_1,\ldots,y_T)$.

3. The form of the statistical model

In this paper we shall deal with the following form of the statistical model:

$$y_t = \begin{cases} y_{1t} & \text{with probability } \lambda \\ y_{2t} & \text{with probability } 1-\lambda \end{cases}, \tag{3.1}$$

where $t = 1,\ldots,T$ and

$$P(y_{1t} = 1) = p_1\,,\ P(y_{1t} = 0) = 1-p_1\ , \tag{3.2}$$

$$P(y_{2t} = 1) = p_2\,,\ P(y_{2t} = 0) = 1-p_2\ , \tag{3.3}$$

$\lambda,\ p_1,\ p_2 \in (0;1),\ p_1 \neq p_2$.

We can notice that

$$P(y_t = 1) = P(y_t = 1\,|y_t = y_{1t})\ P(y_t = y_{1t}) + P(y_t = 1\,|y_t = y_{2t})\ P(y_t = y_{2t}) =$$

$$= p_1\lambda + p_2(1-\lambda) \tag{3.4}$$

Let $p = \lambda p_1 + (1-\lambda)p_2$. The variable y_t has the Bernoulli distribution with parameter p.

4. The estimation of the model

We can estimate parameters of the model (3.1) using different methods. For example we may apply the maximum likelihood method. The maximum likelihood function in this case has the form:

$$L(\mathbf{y}|p_1,p_2,\lambda) = \prod_{t=1}^{T} [y_t p + (1-y_t)(1-p)] = p^{T\bar{y}}(1-p)^{T(1-\bar{y})} =$$

$$= [\lambda p_1 + (1-\lambda)p_2]^{T\bar{y}} [1 - (\lambda p_1 + (1-\lambda)p_2)]^{T(1-\bar{y})} , \tag{4.1}$$

where $\mathbf{y} = (y_1,\dots,y_T)$, $\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$. If we determine the maximum of the function (4.1) we will get the maximum likelihood estimators of parameters p_1, p_2 and λ.

Now we will derive the formulas for the Bayesian estimators of parameters p_1, p_2 and λ. We will consider two cases:

1) when parameter λ is unknown and p_1, p_2 are known,

2) when parameter λ is known and p_1, p_2 are unknown.

Let $F_{c,d}$ be the cumulative density function of the beta distribution with parameters c, d and let $B(c,d)$ be the value of the beta function for c and d. Let $a = T\bar{y} + 1$, $b = T(1-\bar{y}) + 1$.

Let p_1, p_2 be known and let the prior distribution of λ be the uniform distribution on the interval (0;1) i.e.

$$g(\lambda) = \begin{cases} 1 & \text{for } \lambda \in (0;1) \\ 0 & \text{for } \lambda \notin (0;1) \end{cases} . \tag{4.2}$$

We can calculate that

$$D = \int_0^1 g(\lambda)\, L(\mathbf{y}|\lambda, p_1, p_2)d\lambda = \frac{B(a,b)}{p_1 - p_2}[F_{a,b}(p_1) - F_{a,b}(p_2)] . \tag{4.3}$$

The expression D is the denominator of formula (2.1) for the model (3.1) with known p_1 and p_2.

The Bayesian estimator of parameter λ is a conditional expectation of the posterior distribution of this parameter for the quadratic loss function, i.e.

$$\hat{\lambda} = E(\lambda|\mathbf{y}) = \int_0^1 \lambda g(\lambda|\mathbf{y})d\lambda =$$

$$= \frac{1}{D}\int_0^1 \lambda[\lambda p_1 + (1-\lambda)p_2]^{T\bar{y}} [1 - (\lambda p_1 + (1-\lambda)p_2)]^{T(1-\bar{y})}d\lambda = \tag{4.4}$$

$$= \frac{B(a+1,b)}{B(a,b)(p_1-p_2)} \frac{F_{a+1,b}(p_1) - F_{a+1,b}(p_2)}{F_{a,b}(p_1) - F_{a,b}(p_2)} - \frac{p_2}{p_1 - p_2} .$$

Now let λ be known and the prior distribution of p_1 and p_2 be the uniform distribution with density functions:

$$f_1(p_1) = \begin{cases} 1 & \text{for } p_1 \in S_1 \\ 0 & \text{for } p_1 \notin S_1 \end{cases} \tag{4.5}$$

$$f_2(p_2) = \begin{cases} 1 & \text{for } p_2 \in S_2 \\ 0 & \text{for } p_2 \notin S_2 \end{cases} \tag{4.6}$$

where $S_1 = (0;p_2) \cup (p_2;1)$ and $S_2 = (0;p_1) \cup (p_1;1)$. We can calculate that

$$M = \int_{S_1}\int_{S_2} [\lambda p_1 + (1-\lambda)p_2]^{T\bar{y}} [1 - (\lambda p_1 + (1-\lambda)p_2)]^{T(1-\bar{y})} dp_2 dp_1 =$$

$$= \frac{B(a,b)}{\lambda} [1 - F_{a,b}(1-\lambda)] + \frac{B(a,b)}{1-\lambda} [1 - F_{a,b}(\lambda)] + \tag{4.7}$$

$$+ \frac{B(a+1,b)}{\lambda(1-\lambda)} [F_{a+1,b}(\lambda) + F_{a+1,b}(1-\lambda) - 1] .$$

The M is the denominator of formula (2.1) for the model (3.1) with known λ.

The joint posterior distribution of p_1 and p_2 has the form:

$$h(p_1,p_2|\mathbf{y}) = \frac{1}{M} [\lambda p_1 + (1-\lambda)p_2]^{T\bar{y}} [1 - (\lambda p_1 + (1-\lambda)p_2)]^{T(1-\bar{y})} . \tag{4.8}$$

Now we can determine the Bayesian estimators of p_1 and p_2 for the quadratic loss function. We obtain:

$$\hat{p}_1 = E(p_1|\mathbf{y}) = \int_0^1 p_1 [\int_0^1 h(p_1,p_2|y)dp_2]dp_1 =$$

$$= \frac{B(a,b)}{2M(1-\lambda)} [(1 - F_{a,b}(\lambda)] - \frac{1}{2M\lambda^2(1-\lambda)} \{ B(a+2,b) [1 - F_{a+2,b}(\lambda) \tag{4.9}$$

$$- F_{a+2,b}(1-\lambda)] - 2(1-\lambda)B(a+1,b) [1 - F_{a+1,b}(1-\lambda)] + (1-\lambda)^2 B(a,b) [(1-F_{a,b}(1-\lambda)]\}$$

and

$$\hat{p}_2 = E(p_2|\mathbf{y}) = \int_0^1 p_2 [\int_0^1 h(p_1,p_2|\mathbf{y})dp_1]dp_2 =$$

$$= \frac{B(a,b)}{2M\lambda} [1 - F_{a,b}(1-\lambda)]) - \frac{1}{2M\lambda(1-\lambda)^2} \{ B(a+2,b) [1 - F_{a+2,b}(\lambda) \tag{4.10}$$

$$- F_{a+2,b}(1-\lambda)] - 2\lambda B(a+1,b) [1 - F_{a+1,b}(\lambda)] + \lambda^2 B(a,b)[(1-F_{a,b}(\lambda)]\} .$$

Because

$$B(a+1,b) = aB(a,b) \ / \ (a+b), \tag{4.11}$$

$$B(a+2,b) = a(a+1)\,B(a,b) \ / \ (a+b)(a+b+1) \tag{4.12}$$

we can write the formulas (4.4), (4.9) and (4.10) in the following way:

$$\hat{\lambda} = \frac{a}{(a+b)(p_1-p_2)} \, \frac{F_{a+1,b}(p_1) - F_{a+1,b}(p_2)}{F_{a,b}(p_1) - F_{a,b}(p_2)} - \frac{p_2}{p_1 - p_2} \ , \tag{4.13}$$

$$\hat{p}_1 = \frac{1}{K}\Big[\frac{1-F_{a,b}(\lambda)}{2(1-\lambda)} - \frac{a(a+1)}{(a+b)(a+b+1)} \, \frac{1-F_{a+2,b}(\lambda) - F_{a+2,b}(1-\lambda)}{2\lambda^2(1-\lambda)} +$$

$$+ \frac{a}{a+b} \, \frac{1-F_{a+1,b}(1-\lambda)}{\lambda^2} - \frac{(1-\lambda)(1-F_{a,b}(1-\lambda))}{2\lambda^2}\Big] \ , \tag{4.14}$$

$$\hat{p}_2 = \frac{1}{K}\Big[\frac{1-F_{a,b}(1-\lambda)}{2\lambda} - \frac{a(a+1)}{(a+b)(a+b+1)} \, \frac{1-F_{a+2,b}(\lambda) - F_{a+2,b}(1-\lambda)}{2\lambda(1-\lambda)^2} +$$

$$+ \frac{a}{a+b} \, \frac{1-F_{a+1,b}(\lambda)}{(1-\lambda)^2} - \frac{\lambda(1-F_{a,b}(\lambda))}{2(1-\lambda)^2}\Big] \ , \tag{4.15}$$

where

$$K = \frac{1-F_{a,b}(1-\lambda)}{\lambda} + \frac{1-F_{a,b}(\lambda)}{1-\lambda} + \frac{a}{a+b} \, \frac{F_{a+1,b}(\lambda) + F_{a+1,b}(1-\lambda) - 1}{\lambda(1-\lambda)} \ . \tag{4.16}$$

If p_1, p_2 and λ are unknown and are uniformly distributed we obtain the integral which is not convergent. In this case we cannot determine Bayesian estimators.

5. The numerical example

The model (3.1) can have different applications. We will consider the following example. Statistical observations were generated from suitable distributions.

Let us assume that in the human organism some pathological changes can appear as a result of some factors. Let us assume that the probability of morbidity is equal to 0.6 in the case. If these factors are not influenced then the probability of morbidity is equal to 0.2. We do not know which part of population was influenced by unwholesome factors, but we want to estimate this fraction. We can apply the formula (4.4) in this case.

Let $y_t = 1$ if a man is ill and $y_t = 0$ if a man is not ill. We have the following observations for the variable y_t: 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0. The sample y_t for $t = 1,...,25$ is obtained by the simulation. We know that $p_1 = 0.6$ and $p_2 = 0.2$. We generate values of the variable y_{1t} from the Bernoulli distribution with parameter 0.6 and y_{2t} from the Bernoulli distribution with parameter 0.2. Next we generate values of the variable z_t from the Bernoulli distribution with parameter 0.2 and we take $y_t = z_t y_{1t} + (1-z_t)y_{2t}$ (which means that the true value of parameter λ is equal to 0.2). Now we can estimate parameter λ by the formula (4.4) After calculations we have the estimate $\hat{\lambda} = 0.23$.

If we know which part of the population was influenced by unwholesome factors and the probabilities of morbidity p_1 and p_2 are unknown then we can estimate these probabilities by the formulas (4.9) and (4.10).

Let $\lambda = 0.2$. We use the same sample y_t, $t = 1,...,25$. After calculations we get the following estimates: $\hat{p}_1 = 0.48$ and $\hat{p}_2 = 0.21$ where $\hat{p}_1$ is the estimate of morbidity probability for this part of our population which was influenced by unwholesome factors and $\hat{p}_2$ is the estimate of morbidity probability for other people.

6. Final remarks

In this paper a proposition of the Bayesian estimators for a special model with dependent Bernoulli variable is presented. We can apply this estimation method for small samples. We may notice that the values of these estimators are calculated in a simple way. We only use the cumulative density of the beta distribution. The maximum likelihood method, for example, is more complicated, because we have to determine the maximum of a multivariable function and we need a large sample.

The considered statistical model may be applied in the cases in which observed variable is qualitative and we can assume that it has the Bernoulli distribution with a changing parameter.

REFERENCES

Booth, N.B., Smith, A.F.M. (1982). A Bayesian approach to retrospective identification of change-points. *Journal of Econometrics* **19**, 7-22.

Czekała, M., Dziechciarz, J. (1990). Bayesian estimation. In: Econometric models estimation. S. Bartosiewicz (Ed.), PWN, Warszawa, 260-286 (in Polish).

Ferreira, P.E. (1975). A Bayesian analysis of a switching regression model: known number of regimes. *Journal of the American Statistical Association* **70**, 370-374.

Lubrano, M. (1985). Bayesian analysis of switching regression models. *Journal of Econometrics* **29**, 69-95.

Pruska, K. (1994). Estimation of switching regression models with dependent Bernoulli variable. 12th Prague Conference on "Information Theory, Statistical Decision Function and Random Processes".

Tsurumi, H. (1982). Bayesian and maximum likelihood analysis of a gradual switching regression in a simultaneous equation framework. *Journal of Econometrics* **26**, 165-182.

Received 10 October 1994;
revised 3 April 1995

Bayesowska estymacja modelu zmian strukturalnych z zero-jedynkową zmienną zależną

Streszczenie

W pracy rozpatrywany jest statystyczny model zmian strukturalnych, w którym występują dwie zmienne o rozkładach zero-jedynkowych z parametrami, odpowiednio, p_1 i p_2. Zmienne te nie realizują się jednocześnie. Prawdopodobieństwo zrealizowania się jednej zmiennej wynosi λ, a drugiej zmiennej $1-\lambda$.

Dla powyższego modelu wyznaczone zostały estymatory bayesowskie jego parametrów w przypadku, gdy

1) parametr λ jest znany, a parametry p_1, p_2 nie są znane,

2) parametr λ nie jest znany, a parametry p_1, p_2 są znane,

przy czym przyjmuje się, że rozkłady a priori parametrów λ, p_1, p_2 są jednostajne na przedziale (0;1).

Słowa kluczowe: estymacja bayesowska, rozkłady a priori i a posteriori, metoda największej wiarogodności