The Korean Communications
in Statistics Vol. 9, No. 1, 2002
pp. 129-139

Default Bayesian Method for Detecting the Changes in Sequences of Independent Exponential and Poisson Random Variates[1]

Su Youn Jeong[2] and Young Sook Son[3]

Abstract

Default Bayesian method for detecting the changes in sequences of independent exponential random variates and independent Poisson random variates is considered. Noninformative priors are assumed for all the parameters in both of change models. Default Bayes factors, AIBF, MIBF, FBF, to check whether there is any change or not on each sequence and the posterior probability densities of change at each time point are derived. Theoretical results discussed in this paper are applied to some numerical data.

Keywords : noninformative prior, default Bayes factor, arithmetic intrinsic Bayes fator(AIBF), median intrinsic Bayes factor(MIBF), fractional Bayes factor(FBF), posterior probability, change point, exponential sequence, Poisson sequence.

1. Introduction

A sequence $\boldsymbol{X}=\{X_1, X_2, \ldots, X_n\}$ of independent random variates, $X_1, X_2, \ldots, X_n$, is said to have a change at an unknown time γ if it is observed from the change model M_1,

$$M_1 : X_t \sim \begin{cases} f(x \mid \theta_0), & t=1,2,\ldots,\gamma, \\ f(x \mid \theta_1), & t=\gamma+1, \gamma+2, \ldots, n, \end{cases}$$

where $f(x \mid \cdot)$ is a probability density function and unknown parameters vectors, θ_0 and θ_1 $(\theta_0 \neq \theta_1)$. The no-change model M_0 compared with the change model M_1 is defined by

---

1) This study was financially supported by Chonnam National University in the program, 1999.

2) Kwangyang SM Section, POSDATA, Sung Nam, 463-775, Korea.
E-mail : syjeong@mail.posdata.co.kr

3) Dapartment of Statistics, Chonnam National University, Kwangju, 500-757, Korea.
E-mail : ysson@chonnam.chonnam.ac.kr

$$M_0 : X_t \sim f(x|\theta_0), \quad t=1,2,\ldots,n.$$

The Bayes factor can be used as a Bayesian tool to assess whether there is any change in a sequence of random variates or not, that is, which of two models, M_0 or M_1, is correct. Let $\boldsymbol{x}=\{x_1,x_2,\ldots,x_n\}$, $\xi_0=\theta_0$, and $\xi_1=(\theta_0,\theta_1,\gamma)$, where $\xi_j\in\Xi_j, j=0,1$, and Ξ_j is a parameter space under the model $M_j, j=0,1$. Then the Bayes factor B_{10} to compare a change model M_1 to a no-change model M_0 is defined by $B_{10}(\boldsymbol{x}|b=1)$, where

$$B_{10}(\boldsymbol{x}|b)=\frac{m_1(\boldsymbol{x}|b)}{m_0(\boldsymbol{x}|b)}, \tag{1.1}$$

and

$$m_j(\boldsymbol{x}|b) = \int_{\Xi_j}\pi_j(\xi_j)l_j^b(\xi_j|\boldsymbol{x})d\xi_j \tag{1.2}$$

with a probability density function $f_j(x|\xi_j)$, a likelihood function $l_j(\xi_j|\boldsymbol{x})=\Pi_{k=1}^n f_j(x_k|\xi_j)$, a fraction $b(0<b\leq 1)$ of likelihood function, and a prior distribution $\pi_j(\xi_j)$ of parameter ξ_j under the model M_j, $j=0,1$. In equation (1.2), $m_j(\boldsymbol{x}|b=1)$, $j=0,1$, is specially called a marginal or a predictive density of the model M_j, $j=0,1$.

Bayesian approach requires the prior distributions for all the unknown parameters in models.

The objective Bayesian methods based on noninformative priors are called 'default' or 'automatic' Bayesian methods in comparison to the subjective Bayesian method with subjective prior information. Default priors like noninformative priors are objective but most of them are improper, so the caution is needed because of unknown constants incorporated into the Bayes factor. The intrinsic Bayes factor(IBF) of Berger and Pericchi(1996) and the fractional Bayes factor(FBF) of O'Hagan(1995) are classified as 'default' or 'automatic' Bayes factors free from arbitrariness of noninformative improper priors.

An arithmetic IBF(AIBF) of Berger and Pericchi(1996) and median IBF(MIBF) of Berger and Pericchi(1998) are defined as follows

$$B_{10}^{AIBF}=B_{10}(\boldsymbol{x}|b=1)\cdot\frac{1}{L}\sum_{l=1}^{L}B_{01}(\boldsymbol{x}(l)|b=1), \tag{1.3}$$

$$B_{10}^{MIBF}=B_{10}(\boldsymbol{x}|b=1)\cdot\underset{1\leq l\leq L}{Median}\{B_{01}(\boldsymbol{x}(l)|b=1)\}, \tag{1.4}$$

where $B_{01}(\boldsymbol{x}(l)|b=1)=m_0(\boldsymbol{x}(l)|b=1)/m_1(\boldsymbol{x}(l)|b=1)$, $m_j(\boldsymbol{x}(l)|b=1)$ is defined in (1.2) by letting b by 1 and a full sample $\boldsymbol{x}$ by the $l-$th minimal training sample, $\boldsymbol{x}(l)$, and L is the number of minimal training samples possible in the sample. The minimal training sample implies the part of full sample with the minimal sample size to guarantee $0<m_j(\boldsymbol{x}\,|b=1)<\infty$ for $j=0,1$ in equation (1.2). Finally the FBF of O'Hagan(1995) is defined by

$$B_{10}^{FBF}=B_{10}(\boldsymbol{x}\,|b=1)\cdot B_{01}(\boldsymbol{x}\,|b), \tag{1.5}$$

where $B_{01}(\boldsymbol{x}\,|b)$ is obtained from the reciprocal of equation (1.1), and the fraction b of likelihood function is usually and simply used as $b=m/n$ with the size m of a minimal training sample.

The Bayes factor B_{10} with the value greater than 1 will support the change model M_1. The posterior probability of one model gives the information of relative support for it to the others. The posterior probability of change model M_1 is defined by

$$P\,(M_1|\boldsymbol{x})=\frac{p_1}{p_1+p_0B^*_{01}}, \tag{1.6}$$

where $*$ denotes *AIBF*, *MIBF*, or *FBF*, p_j is the prior probability of the model $M_j, j=0,\ 1$, and B^*_{01} is the reciprocal of B^*_{10}.

Approachs to previous studies on the change point problems in a sequence of independent random variates are divided by two, non-Bayesian and Bayesian approach. We are interested in its Bayesian approach. Underlying distributions for which the change point problems were discussed are mainly binomial, normal, exponential, and Poisson (Broemeling(1974), Smith(1975),

Lee and Heghinian(1977), Menzefricke(1981), Booth and Smith(1982), Raftery and Akman(1986), Carlin, Gelfand, and Smith(1992), and Chung and Dey(1996)). Generally they used conjugate priors(hierachial conjugate priors in Carlin, Gelfand, and Smith(1992)) or nonimformative improper priors in the work of computing the marginal posterior probability of each change time point and used conjugate priors in the work of computing the Bayes factors. Exceptionally, Booth and Smith(1982) and Raftery and Akman(1986) solved an arbitrary constant problem incorporated into a Bayes factor due to the assumption of the noninformative improper prior by the imaginary constant idea of a thought experiment in Spiegelhalter and Smith(1982). Jeong and Son(2000) obtained output on default Bayesian method for detecting the change in a sequence of independent multivariate normal vectors under the assumption of noninformative improper priors for parameters of multivariate normal distribution.

The default Bayesian method used in this paper for solving change point problems is simpler to use than usual Bayesian method under conjugate priors or the imaginary constant method of Spiegelhalter and Smith(1982). So it can be directly used in the beginning of Bayesian model selection with no subjective prior consideration .

In this paper we discuss default Bayesian method for detecting the change in a sequence of independent exponential random variaties(section 2) and Poisson random variates(section 3). As priors are assumed noninformative improper priors for the parameters of each distribution and the uniform prior for the change point. Default Bayes factors, AIBF, MIBF, and FBF, to check whether there is any change or not on each sequence and the posterior density of change at each time point are derived. A numerical study of section 4 shows the application of theoretical results discussed in section 2 and section 3.

2. A Change in an Exponential Sequence

Consider two models, a no-change model M_0 and a change model M_1, that shows the change in a sequence of independent exponential random variates as follows

$$M_0 : X_t \sim \text{Exponential}(\beta), \quad t=1,2,\ldots,n, \tag{2.1}$$

$$M_1 : X_t \sim \begin{cases} \text{Exponential}(\beta), & t=1,2,\ldots,\gamma, \\ \text{Exponential}(\delta), & t=\gamma+1,\gamma+2,\ldots,n, \end{cases} \tag{2.2}$$

where β and δ are unknown positive parameters, and γ is an integer with $1 \le \gamma \le n-1$.

For a no-change model of (2.1) the likelihood function is given by

$$l_0(\beta \mid \boldsymbol{x}) = \beta^{-n} \exp\left\{-\sum_{j=1}^{n} x_j \,/\, \beta\right\}$$

and a noninformative improper prior, $\pi_0(\beta) = 1/\beta,\ \beta > 0$, is assumed. Thus we can compute

$$m_0(\boldsymbol{x} \mid b) = \int_0^\infty \pi_0(\beta) l_0^b(\beta \mid \boldsymbol{x}) d\beta = \Gamma(bn)\,(b \sum_{j=1}^{n} x_j)^{-bn}. \tag{2.3}$$

For a change model of (2.2) the likelihood function is given by

$$l_1(\beta, \delta, \gamma \mid \boldsymbol{x}) = \beta^{-\gamma}\delta^{-(n-\gamma)} \exp\left\{- \sum_{j=1}^{\gamma} x_j \,/\, \beta \right\} \exp\left\{- \sum_{j=\gamma+1}^{n} x_j \,/\, \delta \right\}$$

and a noninformative improper prior, $\pi_1(\beta, \delta) = 1/(\beta\delta)$, $\beta > 0$, $\delta > 0$, an uniform prior, $\pi_1(\gamma) = 1/(n-1)$, $\gamma = 1, 2, \ldots, n-1$, and $\pi_1(\beta, \delta, \gamma) = \pi_1(\beta, \delta) \cdot \pi_1(\gamma)$ are assumed.

Now, we can compute

$$\begin{aligned} m_1(\gamma, \boldsymbol{x} \mid b) &= \int_0^\infty \int_0^\infty \pi_1(\beta, \delta, \gamma) l_1^b(\beta, \delta, \gamma \mid \boldsymbol{x}) d\beta d\delta \\ &= \frac{\Gamma(b\gamma)\Gamma\{b(n-\gamma)\}}{(n-1)\, b^{bn} (\sum_{j=1}^{\gamma} x_j)^{b\gamma} (\sum_{j=\gamma+1}^{n} x_j)^{b(n-\gamma)}} . \end{aligned}$$

The marginal posterior probability density, $f(\gamma \mid \boldsymbol{x})$, of γ is obtained by

$$f(\gamma \mid \boldsymbol{x}) = \frac{m_1(\gamma, \boldsymbol{x} \mid b=1)}{m_1(\boldsymbol{x} \mid b=1)}, \tag{2.4}$$

where

$$m_1(\boldsymbol{x} \mid b) = \sum_{\gamma=1}^{n-1} m_1(\gamma, \boldsymbol{x} \mid b). \tag{2.5}$$

The size of a minimal training sample is 2 equal to a minimal sample size to guarantee finite marginal densities for both models, M_0 and M_1. Our minimal training sampling plan is to sample in order to preserve the continuity of time. Thus, under the no-change model M_0 of (2.1) minimal training samples are

$$\boldsymbol{x}(l) = \{(x_l, x_{l+1}) \mid x_l, x_{l+1} \sim \text{Exponential}(\beta)\}, \qquad l = 1, 2, \ldots, n-1,$$

and under the change model M_1 of (2.2) minimal training samples are

$$\boldsymbol{x}(l) = \{(x_l, x_{l+1}) \mid x_l \sim \text{Exponential}(\beta),\ x_{l+1} \sim \text{Exponential}(\delta)\}, \qquad l = 1, 2, \ldots, n-1.$$

Replacing n by 2, γ by 1, b by 1, and $\boldsymbol{x}$ by $\boldsymbol{x}(l)$ in (2.3) and (2.5), the predictive densities of minimal training samples under each model are respectively given by for $l = 1, 2, \ldots, n-1$,

$$m_0(\boldsymbol{x}(l) \mid b=1) = (x_l + x_{l+1})^{-2} \tag{2.6}$$

and

$$m_1(\boldsymbol{x}(l)|b=1)=(x_l \cdot x_{l+1})^{-1}. \tag{2.7}$$

Finally, after (1.1)-(1.5) are filled with (2.3), (2.5), (2.6), (2.7), the AIBF, B_{10}^{AIBF}, the MIBF, B_{10}^{MIBF}, and the FBF, B_{10}^{FBF} are straightforwardly obtained.

3. A Change in a Poisson Sequence

Consider two models, a no-change model M_0 and a change model M_1, in a sequence of independent Poisson random variates as follows

$$M_0 : X_t \sim Possion(\mu), \quad t = 1,2,\ldots,n, \tag{3.1}$$

$$M_1 : X_t \sim \begin{cases} Possion(\mu), & t = 1,2,\ldots,\gamma, \\ Possion(\eta), & t=\gamma+1,\gamma+2,\ldots,n, \end{cases} \tag{3.2}$$

where μ and η are unknown positve parameters, and γ is an integer with $1\le \gamma \le n-1$. For a no-change model M_0 of (3.1) the likelihood function is given by

$$l_0(\mu|\boldsymbol{x})=\exp\{-n\mu\}\mu^{\sum_{j=1}^n x_j} / \Pi_{j=1}^n x_j!$$

and a noninformative improper prior , $\pi_0(\mu)=\mu^{-1/2}$, $\mu>0$, is assumed. Now we can obtain

$$m_0(\boldsymbol{x}|b)=\int_0^\infty \pi_0(\mu)\, l_0^b(\mu|\boldsymbol{x})\,d\mu=\frac{\Gamma(b\sum_{j=1}^n x_j+\frac{1}{2})}{(\Pi_{j=1}^n x_j!)^b (bn)^{b\sum_{j=1}^n x_j+\frac{1}{2}}}. \tag{3.3}$$

For change model M_1 of (3.2) the likelihood function is

$$l_1(\mu,\eta,\gamma|\boldsymbol{x})=\exp\{-\gamma\mu-(n-\gamma)\eta\}\mu^{\sum_{j=1}^{\gamma} x_j}\eta^{\sum_{j=\gamma+1}^{n} x_j} / \Pi_{j=1}^n x_j!$$

and a noninformative improper prior, $\pi_1(\mu,\eta)=(\mu\eta)^{-1/2}$, $\mu>0$, $\eta > 0$, an uniform prior

$\pi_1(\gamma)=1/(n-1)$, $\gamma=1,2,\ldots,n-1$, and $\pi_1(\mu,\eta,\gamma)=\pi_1(\mu,\eta)\cdot\pi_1(\gamma)$ are assumed.
Now, we can compute

$$m_1(\gamma,\ \boldsymbol{x}\mid b) = \int_0^\infty\int_0^\infty \pi_1(\mu,\eta,\gamma) l_1^b(\mu,\eta,\gamma\mid \boldsymbol{x})\,d\mu\,d\eta$$

$$= \frac{\Gamma(b\sum_{j=1}^{\gamma} x_j+\frac{1}{2})\Gamma(b\sum_{j=\gamma+1}^{n} x_j+\frac{1}{2})}{(n-1)(\Pi_{j=1}^{n} x_j!)^b (b\gamma)^{b\sum_{j=1}^{\gamma} x_j+\frac{1}{2}}\{b(n-\gamma)\}^{b\sum_{j=\gamma+1}^{n} x_j+\frac{1}{2}}}.$$

The marginal posterior probability density, $f(\gamma|\ \boldsymbol{x})$, of γ given data $\boldsymbol{x}$ is defined as same as equation (2.4) and (2.5).

Also, the size of a minimal training sample is the same as the exponential case. After following the minimal training sampling scheme as exponential case, the predictive densities of minimal samples under the model M_0 and M_1 are respectively derived by

$$m_0(\ \boldsymbol{x}(l)|b=1)=\Gamma(x_l+x_{l+1}+\frac{1}{2})/(2^{x_l+x_{l+1}+\frac{1}{2}}\cdot x_l!\cdot x_{l+1}!) \tag{3.4}$$

and

$$m_1(\ \boldsymbol{x}(l)|b=1)=\Gamma(x_l+\frac{1}{2})\Gamma(x_{l+1}+\frac{1}{2})/(x_l!\cdot x_{l+1}!)\ . \tag{3.5}$$

Finally, after (1.1)-(1.5) are filled with (3.3), (2.5), (3.4), (3.5) the AIBF, B_{10}^{AIBF}, the MIBF, B_{10}^{MIBF}, and the FBF, B_{10}^{FBF}, are directly obtained.

4. Numerical Study

We perform a simulation study to support theoretical results developed in previous sections. With 10,000 replications five exponential(Poisson) data sets of sample size 30 with $\beta(\mu)=1.0$ and $\delta(\eta)=0.2,\ 0.5,1.0,2.0,5.0$ are simulated from $X_t\sim$ Exponential(1.0)(Poisson(1.0)), $t=1,2,\ \ldots,\gamma$ and $X_t\sim$ Exponential(δ)(Poisson(η)), $t=\gamma+1,\gamma+2,\ldots,n$. The change point is given by $\gamma=20(10)$ in case of Exponential(Piosson).

Table 4.1 shows results of the posterior probability of a change model M_1 computed from (1.6) with the equal prior probability for each model through the AIBF, the MIBF, the FBF computed from (1.3), (1.4.), (1.5). The values of the first line and the second line in each cell of Table are the mean and the standard deviation of 10,000 replications, respectively. The value in the parenthesis implies the proportion of replications with B_{10} greater than 1.

As the size of change is larger, the posterior probabilities to support the change model and the proportions of replications with B_{10} greater than 1 are close to 1. For an exponential(Poisson) data with $\delta(\eta)=1.0$ to explain the no-change model 10~13% (13~14%) of 10,000 replications lead to the wrong model selection.

Proportion histograms of the mode γ_0 that maximize the posterior probability density of the change point γ in (2.4) given exponential data and Poisson data are shown in Figure 4.1 and Figure 4.2 , respectively. The histogram of U-type for a no-change data with $\delta(\eta)=1.0$ implies that the data set is generated from the same distribution with no change. This U-type histogram is changed to the histogram degenerated at $\gamma=20(10)$ in exponential(Poisson) case as the degree of change is larger.

In our research a number of simulation experiments were performed with various $\delta(\eta)$'s, sample sizes, change points γ's. Though their results are not contained in this paper because of space limited we can see that simulation results are more coincident with our theoretical expectation as the size of sample is larger and the value of γ is close to $n/2$.

Table 4.1: Results of posterior probability for the change model M_1.

$\delta(\eta)$	Exponential Model			Poisson Model		
	AIBF	MIBF	FBF	AIBF	MIBF	FBF
0.2	0.893 0.156 (0.96)	0.899 0.149 (0.97)	0.881 0.165 (0.95)	0.791 0.214 (0.85)	0.795 0.216 (0.85)	0.764 0.228 (0.81)
0.5	0.447 0.226 (0.38)	0.496 0.224 (0.41)	0.453 0.224 (0.34)	0.513 0.217 (0.41)	0.492 0.225 (0.39)	0.493 0.213 (0.38)
1.0	0.327 0.148 (0.12)	0.347 0.150 (0.13)	0.307 0.143 (0.10)	0.368 0.140 (0.14)	0.360 0.140 (0.13)	0.371 0.133 (0.13)
2.0	0.500 0.248 (0.42)	0.518 0.245 (0.44)	0.477 0.247 (0.38)	0.588 0.242 (0.55)	0.593 0.237 (0.57)	0.590 0.233 (0.57)
5.0	0.931 0.146 (0.96)	0.935 0.140 (0.97)	0.924 0.155 (0.96)	0.999 0.008 (1.00)	0.999 0.007 (1.00)	0.999 0.008 (1.00)

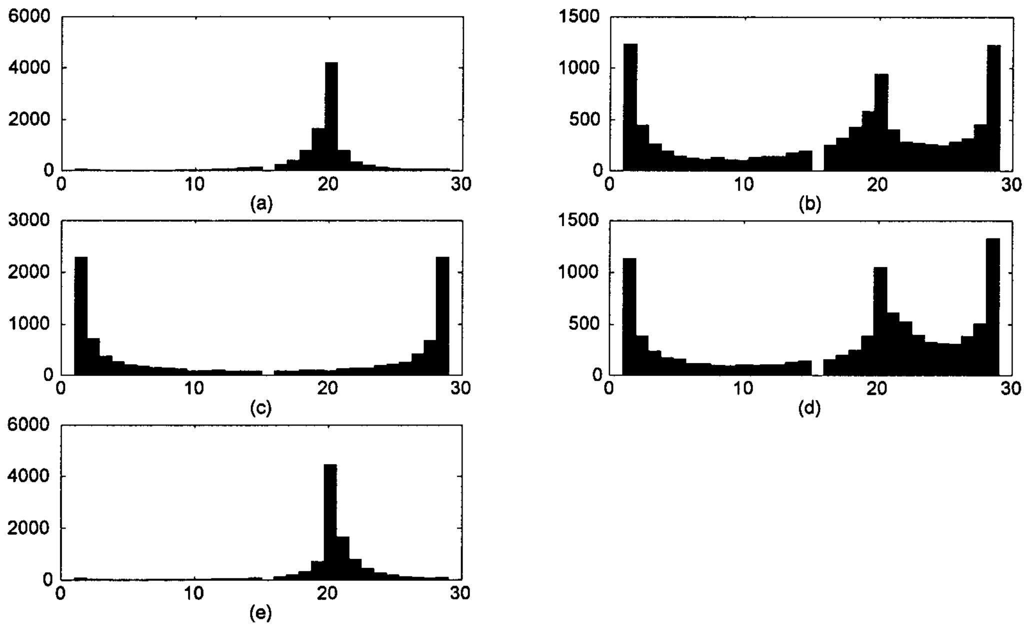

Figure 4.1: Histograms of replications of the mode γ_0 of change point given the exponential data with (a) $\delta=0.2$, (b) $\delta=0.5$, (c) $\delta=1.0$, (d) $\delta=2.0$, (e) $\delta=5.0$.

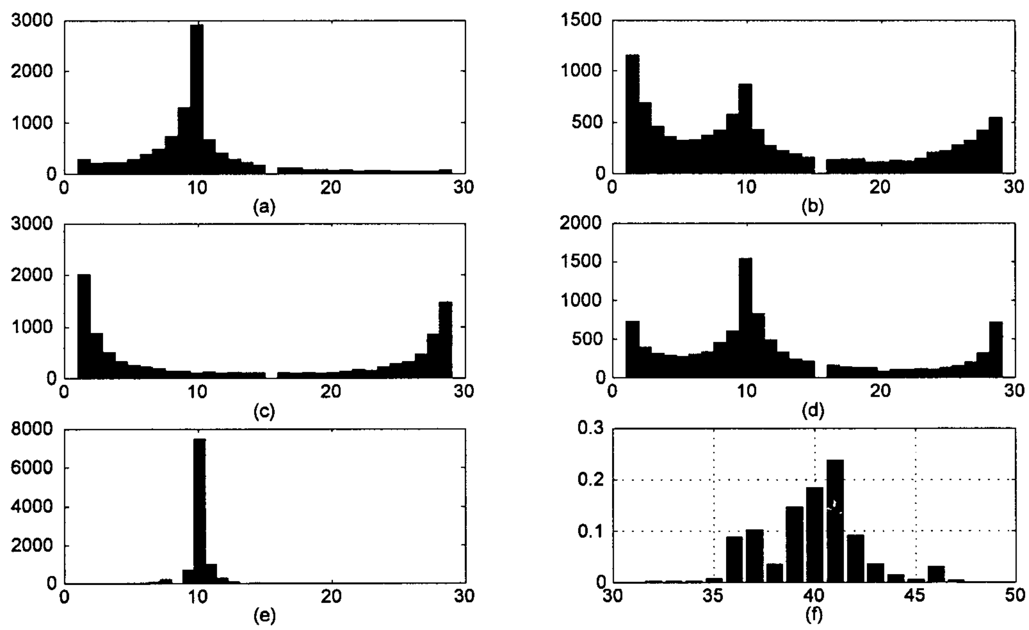

Figure 4.2 : Histograms of replications of the mode γ_0 of change point given the Poisson data with (a) $\eta=0.2$, (b) $\eta=0.5$, (c) $\eta=1.0$, (d) $\eta=2.0$, (e) $\eta=5.0$, and (f) the posterior probability of change point in British coal mining disaster data.

Now, we apply default Bayesian method for the Poisson change model to British coal mining disaster data during 112 year period, 1851-1962 given by Jarrett(1979). The computation gives $B_{10}^{AIBF}=6.7E+12$, $B_{10}^{MIBF}=6.5E+12$, $B_{10}^{FBF}=4.9E+12$, and equally $P(M_1|\mathbf{x})=1.0$ through three default Bayes factors which strongly support the change model. Also, the marginal posterior probability of each change point is shown in (f) of Figure 4.2. The largest three posterior probabilities are 0.238, 0.185, and 0.146 at $\gamma=41, 40$, and 39, respectively. All these results coincide with those of Raftery and Akman(1986), Carlin, Gelfand, and Smith(1992), and Chung and Dey(1996).

References

[1] Berger, J. O. and Pericchi, L.R.(1996), The Intrinsic Bayes Factor for Model Selection and Prediction, *Journal of the American Statistical Asociation*, Vol. 91, No.433, 109-122.

[2] Berger, J. O. and Pericchi, L. R.(1998), Accurate and Stable Bayesian Model Selection: The Median Intrinsic Bayes Factor, Sankhya, B, Vol. 60, 1-18.

[3] Booth, N. B. and Smith, A. F. M.(1982), A Bayesian Approach to Retrospective Identification of Change-Points, *Journal of Economerics*, Vol. 19, 7-22.

[4] Broemeling, L. D.(1974), Bayesian Inferences about a Change Sequence of Random Variables, *Communication in Statistics*, 243-255.

[5] Carlin, B. P., Gelfand, A. E. and Smith, A. F. M.(1992), Hierachial Bayesian Analysis of Change point Problems, *Applied Statistics,* Vol. 41, No. 2 , 389-405.

[6] Chung, Y. S. and Dey, D. K.(1996), Bayes Factor for Change-point Problems with Conjugate Prior, *Journal of the Korean Statistical Society*, Vol. 25, No. 4, 577-588.

[7] Jarrett, R. G.(1979), A note on the intervals between coal-mining disasters, *Biometrika,* 66, 191-193.

[8] Jeong, S. and Son, Y. S.(2000), Default Bayesian Method for Detecting the Changes in a Sequence of Independent Multivariate Normal Vectors, *Proceedings of the Spring Conference, Korean Statistical Society,* 95-100.

[9] Lee, A. F. S. and Heghinian, S. M.(1977), A Shift of the Mean Level in a Sequence of Independent Normal Random Variables-A Bayesian Approach, *Tecnometrics*, Vol. 19, No. 4, 503-506.

[10] Menzefricke, U.(1981), A Bayesian Analysis of a Change in the Precision of a Sequence of Independent Normal Random Variables at an Unknown Time Point, *Applied Statistics,* Vol. 30, No. 2, 141-146.

[11] O'Hagan, A.(1995), Fractional Bayes Factors for Model Comparsion, *Journal of the Royal Statistical Society*, B, Vol. 57, No. 1, 99-138.

[12] Raftery, A. E. and Akman, V. E.(1986), Bayesian analysis of a Poisson process with a change-point, *Biometrika*, Vol. 73, No. 1, 85-89.

[13] Smith, A. F. M.(1975), A Bayesian Approach to Inference about a Change Point in a Sequence of Random Variables, *Biometrika,* Vol. 62, 407-416.
[14] Spiegelhalter, D. J. and Smith, A. F. M.(1982), Bayes Factors for Linear and Log-Linear Models with Vague Prior Information, *Journal of the Royal Statistical Society*, B, Vol. 44, 377-387.

[2002년 1월 접수, 2002년 3월 채택]