APPLICATIONES MATHEMATICAE
24,4 (1997), pp. 457–463

M. MĘCZARSKI and R. ZIELIŃSKI (Warszawa)

BAYES OPTIMAL STOPPING OF A HOMOGENEOUS POISSON PROCESS UNDER LINEX LOSS FUNCTION AND VARIATION IN THE PRIOR

Abstract. A homogeneous Poisson process $(N(t), t \geq 0)$ with the intensity function $m(t) = \theta$ is observed on the interval $[0, T]$. The problem consists in estimating θ with balancing the LINEX loss due to an error of estimation and the cost of sampling which depends linearly on T. The optimal T is given when the prior distribution of θ is not uniquely specified.

1. Introduction. The homogeneous Poisson process with the intensity function $m(t) = \theta$, $\theta > 0$, is widely used in many different fields of applied probability and statistics.

Suppose that the process is observed on the interval $[0, T]$ and the problem consists in optimal stopping of observation (optimal choice of T) and optimal estimation of θ under the following circumstances:

1. The loss in estimating θ by an estimator $\widehat{\theta}$ is measured by the LINEX function (see Zellner (1986))

$$L(\widehat{\theta}, \theta) = b[\exp\{a(\widehat{\theta} - \theta)\} - a(\widehat{\theta} - \theta) - 1], \quad a < 0. \tag{1.1}$$

2. Costs of observations are assumed to grow linearly in T, and the total cost of observation and estimation is defined as

$$L_T(\widehat{\theta}, \theta) = kL(\widehat{\theta}, \theta) + (1 - k)(c_1 T + c_2), 0 < k < 1, \ c_1, c_2 > 0. \tag{1.2}$$

3. The prior knowledge about θ is that θ has a gamma $\mathcal{G}(\alpha, \beta)$ distribution

$$g(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha - 1} \exp\{-\beta\theta\}, \quad \theta > 0, \ \alpha, \beta > 0. \tag{1.3}$$

1991 *Mathematics Subject Classification*: 62F15, 62F35, 62L15, 62M09.

Key words and phrases: homogeneous Poisson process, LINEX loss function, prior distribution uncertainty.

Under these conditions Ebrahimi (1992) gave the full solution of the problem in the more general case of the power law process. We agree with the choice of the prior for θ but we believe that a strict specification of the parameters (α, β) of that distribution is not easy. Hence we are interested in solving the problem under the following additional condition.

4. The prior knowledge about θ is that the distribution of θ belongs to the family

$$\Gamma = \{\mathcal{G}(\alpha,\beta) : 0 < \alpha_{\mathrm{L}} \le \alpha \le \alpha_{\mathrm{U}},\ 0 < \beta_{\mathrm{L}} \le \beta \le \beta_{\mathrm{U}}\} \tag{1.4}$$

for some fixed lower bounds $\alpha_{\mathrm{L}}, \beta_{\mathrm{L}}$, and upper bounds $\alpha_{\mathrm{U}}, \beta_{\mathrm{U}}$.

Considerable attention has been paid to the issue of optimal decisions if the prior is not uniquely determined and it is known to belong to a suitable family of priors; see for example Berger (1994) and references therein. Typical Bayesian approach under condition 4 (conditional Γ-minimaxity) is however not applicable to our case because the optimal decision depends on observations while we have to decide on T before the observations are collected. The situation is exactly as that in the case of experimental design. An approach which consists in averaging future observations under a given prior has been presented in DasGupta and Studden (1991). Another approach in the context of Bayesian robustness, consisting in analysing a posterior quantity as a function on the sample space of future observations has been developed in Męczarski and Zieliński (1997). Below we present an approach which is a combination of the principle of conditional Γ-minimaxity and averaging future observations with respect to plausible posterior distributions, i.e. the Γ-minimaxity. The approach enables us to uniquely choose an optimal stopping time T despite the nonuniqueness of the prior distribution. An alternative to our approach presented below is a pure sequential procedure which will be treated elsewhere.

2. General solution. Given $T > 0$, the number N of events in the interval $[0, T]$ is a Poisson random variable with mean equal to θT so that the likelihood function, given $N = n$, is

$$l_n(\theta) = \theta^n \exp\{-\theta T\}, \qquad n = 0, 1, 2, \ldots \tag{2.1}$$

For a gamma $\mathcal{G}(\alpha, \beta)$ prior for θ, the posterior distribution is gamma $\mathcal{G}(n + \alpha, T + \beta)$. Integrating (1.1) with respect to the posterior distribution (conditioned on the time of observation T and $N = n$) we obtain the posterior risk of the estimator $\widehat{\theta}$,

$$r_{T,n}(\widehat{\theta}, (\alpha,\beta)) = b\left(\frac{T+\beta}{T+\beta+a}\right)^{n+\alpha} \exp\{a\widehat{\theta}\} + ab\frac{n+\alpha}{T+\beta} - b(a\widehat{\theta} - 1) \tag{2.2}$$

and the condition

$$T+\beta+a>0 \tag{2.3}$$

(otherwise the risk is infinite). For the Bayes solution, similarly to Ebrahimi (1992), we obtain

$$\widehat{\theta}^{\mathrm{Bay}}=\widehat{\theta}^{\mathrm{Bay}}(n,T,(\alpha,\beta))=-\frac{n+\alpha}{a}\log\frac{T+\beta}{T+\beta+a} \tag{2.4}$$

which is a simple application of a general formula given in Zellner (1986). The risk function of the Bayes estimator (2.4) is

$$R(\widehat{\theta}^{\mathrm{Bay}}(\cdot,T,(\alpha,\beta)),\theta)=\sum_{n=0}^{\infty}L(\widehat{\theta}^{\mathrm{Bay}},\theta)P_{T,\theta}\{N=n\} \tag{2.5}$$

where

$$P_{T,\theta}\{N=n\}=\frac{(\theta T)^n}{n!}e^{-\theta T},\qquad n=0,1,\ldots, \tag{2.6}$$

gives the distribution of the observation N under a fixed θ. We obtain

$$\begin{aligned}R(\widehat{\theta}^{\mathrm{Bay}}(\cdot,T,(\alpha,\beta)),\theta)=b\biggl(&a\theta+(\theta T+\alpha)\log\frac{T+\beta}{T+\beta+a}\\&+\left(\frac{T+\beta+a}{T+\beta}\right)^{\alpha}\exp\left\{-\frac{a\theta\beta}{T+\beta}\right\}-1\biggr).\end{aligned} \tag{2.7}$$

The Bayes risk, denoted by $r(\widehat{\theta}^{\mathrm{Bay}}(\cdot,T,(\alpha,\beta)))$, of the Bayes estimator under the prior $\Gamma(\alpha,\beta)$ is then obtained as the result of integration of (2.7) with respect to θ:

$$r(\widehat{\theta}^{\mathrm{Bay}}(\cdot,T,(\alpha,\beta)))=b\left\{a\frac{\alpha}{\beta}+\alpha\left(\frac{T+\beta}{\beta}\right)\log\frac{T+\beta}{T+\beta+a}\right\}. \tag{2.8}$$

Now take (α^*,β^*) (i.e. the prior $\Gamma(\alpha^*,\beta^*)$) such that

$$r(\widehat{\theta}^{\mathrm{Bay}}(\cdot,T,(\alpha^*,\beta^*)))=\sup_{(\alpha,\beta)\in[\alpha_{\mathrm{L}},\alpha_{\mathrm{U}}]\times[\beta_{\mathrm{L}},\beta_{\mathrm{U}}]}r(\widehat{\theta}^{\mathrm{Bay}}(\cdot,T,(\alpha,\beta))). \tag{2.9}$$

The supremum on the right-hand side is of course the supremum over the set Γ of plausible prior distributions for θ. The supremum of the total loss (supremum with respect to plausible priors) is equal to

$$\lambda(T)=kr(\widehat{\theta}^{\mathrm{Bay}}(\cdot,T,(\alpha^*,\beta^*)))+(1-k)(c_1T+c_2), \tag{2.10}$$

which depends on T only. The optimal T is T^* for which

$$\lambda(T^*) \le \lambda(T), \qquad T > 0. \tag{2.11}$$

As a conclusion we obtain the following algorithm for finding the Γ-minimax stopping time T^* and the optimal estimator of θ:

1) maximize (2.8) with respect to (α, β) for a fixed T. The minimizer (α^*, β^*) might depend on T;

2) find a minimizer T^* of $\lambda(T)$;

3) given T^*, observe the process $(N(t), t \ge 0)$ on the interval $[0, T^*]$;

4) having observed $N = n$, calculate the Bayes estimator (2.4) under the prior (α^*, β^*):

$$\widehat{\theta}^{\mathrm{Bay}}(\cdot, T^*, (\alpha^*, \beta^*)) = -\frac{n + \alpha^*}{a} \log \frac{T^* + \beta^*}{T^* + \beta^* + a}. \tag{2.12}$$

3. Algorithm. Maximization of (2.8) with respect to (α, β) is easy: it is a linear function of α of the form

$$b \frac{1}{\beta}\left[a + (T + \beta) \log \frac{T + \beta}{T + \beta + a}\right]\alpha \tag{3.1}$$

and hence it is enough to find β^* which maximizes

$$\frac{1}{\beta}\left[a + (T + \beta) \log \frac{T + \beta}{T + \beta + a}\right]. \tag{3.2}$$

Since the function (3.2) is positive and decreasing, we obtain

$$\alpha^* = \alpha_{\mathrm{U}}, \qquad \beta^* = \beta_{\mathrm{L}}. \tag{3.3}$$

Given (α^*, β^*), the optimal T^* is that minimizing (2.10), which now takes on the following form:

$$\lambda(T) = kb \frac{\alpha^*}{\beta^*}\left[a + (T + \beta^*) \log \frac{T + \beta^*}{T + \beta^* + a}\right] + (1 - k)(c_1 T + c_2). \tag{3.4}$$

Numerical minimization of (3.4) is not difficult. It reduces to finding the unique root T^* of the equation

$$-\frac{a}{T + \beta_{\mathrm{L}} + a} - \log\left(1 - \frac{a}{T + \beta_{\mathrm{L}} + a}\right) = \frac{1 - k}{k} \cdot \frac{\beta_{\mathrm{L}}}{\alpha_{\mathrm{U}}} \cdot \frac{c_1}{b}. \tag{3.5}$$

The optimal T^* for some choices of α_{U}, β_{L}, a, b, c_1 and k are presented below. Observe that α_{L}, β_{U} and c_2 have no influence here, so that the optimal solution is absolutely robust to diminishing the shape parameter and/or increasing the scale parameter of the prior distribution.

The oscillation (the range)

$$\sup_{(\alpha,\beta)\in[\alpha_{\mathrm{L}},\alpha_{\mathrm{U}}]\times[\beta_{\mathrm{L}},\beta_{\mathrm{U}}]} r(\widehat{\theta}^{\mathrm{Bay}}(\cdot,T,(\alpha,\beta))) - \inf_{(\alpha,\beta)\in[\alpha_{\mathrm{L}},\alpha_{\mathrm{U}}]\times[\beta_{\mathrm{L}},\beta_{\mathrm{U}}]} r(\widehat{\theta}^{\mathrm{Bay}}(\cdot,T,(\alpha,\beta))) \tag{3.6}$$

of the Bayes risk (2.8) may be considered as a natural indicator of the robustness (see e.g. Sivaganesan and Berger (1989) and Męczarski and Zieliński (1991)). It is interesting to observe that (3.6) is decreasing in T and tends to 0 as $T\to\infty$.

4. Numerical examples. To get a perspicuous intuitive interpretation of numerical examples observe the following:

1) The total loss (1.2) in the case under consideration has the following form:

$$\begin{aligned} L_T(\widehat{\theta},\theta) &= kb[\exp\{a(\widehat{\theta}-\theta)\}-a(\widehat{\theta}-\theta)-1]+(1-k)(c_1T+c_2)\\ &= kb\left[(\exp\{a(\widehat{\theta}-\theta)\}-a(\widehat{\theta}-\theta)-1)+\frac{1-k}{k}\cdot\frac{c_1}{b}T+\frac{1-k}{k}\cdot\frac{c_2}{b}\right]. \end{aligned} \tag{4.1}$$

Now, kb ("scale" of the loss) and $\frac{1-k}{k}\cdot\frac{c_2}{b}$ (an additive constant) are fixed, so that when looking for an optimal stopping time T and an optimal estimator $\widehat{\theta}$ we may confine ourselves to considering the loss of the form

$$(\exp\{a(\widehat{\theta}-\theta)\}-a(\widehat{\theta}-\theta)-1)+\kappa T, \tag{4.2}$$

where $\kappa=\frac{1-k}{k}\cdot\frac{c_1}{b}$.

2) The parameter $a<0$ in the LINEX term of (4.2) has a fixed value, hence we may confine ourselves to finding an optimal $T^*/(-a)$ instead of T^* itself.

3) If the prior distribution belongs to Γ (cf. (1.4)), then the maximum value of the prior expectation of θ, denoted by $E_{\max}\theta$, is equal to $\alpha_{\mathrm{U}}/\beta_{\mathrm{L}}$.

Taking all this into account, the equation (3.5) may be written in the form

$$x-\log(1+x)=\frac{\kappa}{E_{\max}\theta}, \tag{4.3}$$

where

$$x=\frac{1}{\frac{T}{-a}+\frac{\beta_{\mathrm{L}}}{-a}-1}. \tag{4.4}$$

The values of $T/(-a)$ are presented in Table 1 (together with the values of x in the last row).

TABLE 1

	$\kappa/E_{\max}\theta$										
$\beta_L/(-a)$	0	0.001	0.006	0.018	0.095	0.307	0.901	7.602	95.385	993.09	$+\infty$
1	$+\infty$	19	9	5	2	1	0.5	0.1	0.01	0.001	0
1.001	$+\infty$	18.999	8.999	4.999	1.999	0.999	0.499	0.099	0.009	0	
1.01	$+\infty$	18.99	8.99	4.99	1.99	0.99	0.49	0.09	0		
1.1	$+\infty$	18.9	8.9	4.9	1.9	0.9	0.4	0			
1.5	$+\infty$	18.5	8.5	4.5	1.5	0.5	0				
2	$+\infty$	18	8	4	1	0					
3	$+\infty$	17	7	3	0						
6	$+\infty$	14	4	0							
10	$+\infty$	10	0								
20	$+\infty$	0									
x	0	0.053	0.111	0.2	0.5	1	2	10	100	1000	$+\infty$

The strange looking values of $\kappa/E_{\max}\theta$ are the values for which, under the values of $\beta_L/(-a)$ in the first column of the table, the optimal stopping time T^* is equal to 0. Of course, the optimal T^* under any greater value of $\kappa/E_{\max}\theta$ is also 0. Optimal $T^* = 0$ means that θ should be estimated on the base of the prior distribution only.

5. Conclusions. The influence of the uncertainty in the prior has a particular form: only α_U and β_L influence T^*. Consequently, only the uncertainty in the upper bound for α and in the lower bound for β is essential.

Acknowledgments. We feel indebted to Bruno Betrò and Fabrizio Ruggeri for inspiring discussions.

References

J. O. Berger (1994), *An overview of robust Bayesian analysis*, Test 3, 5–124.

A. DasGupta and W. J. Studden (1991), *Robust Bayesian experimental design in normal linear models*, Ann. Statist. 19, 1244–1256.

N. Ebrahimi, (1992), *An optimal stopping time for a power law process*, Sequential Anal. 11, 213–227.

M. Męczarski and R. Zieliński (1991), *Stability of the Bayesian estimator of the Poisson mean under the inexactly specified gamma prior*, Statist. Probab. Lett. 12, 329–333.

M. Męczarski and R. Zieliński (1997), *Stability of the posterior mean in linear models: an admissibility property of D-optimum and E-optimum designs*, ibid., to appear.

S. Sivaganesan and J. O. Berger (1989), *Ranges of posterior measures for priors with unimodal contaminations*, Ann. Statist. 17, 868–889.

A. Zellner (1986), *Bayesian estimation and prediction using asymmetric loss functions*, J. Amer. Statist. Assoc. 81, 446–451.

Marek Męczarski
Institute of Econometrics
Warsaw School of Economics
Al. Niepodległości 162
02-554 Warszawa, Poland
E-mail: mecz@sgh.waw.pl

Ryszard Zieliński
Institute of Mathematics
Polish Academy of Sciences
Śniadeckich 8
00-950 Warszawa, Poland
E-mail: rziel@impan.gov.pl

Received on 21.2.1997