SEQUENTIAL METHODS IN STATISTICS
BANACH CENTER PUBLICATIONS, VOLUME 16
PWN–POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

SPECIAL SEQUENTIAL ESTIMATION PROBLEMS IN MARKOV PROCESSES

JÜRGEN FRANZ

Section of Mathematics, Technical University of Dresden, Dresden, G.D.R.

1. Introduction

For the last ten years, sequential statistical estimation problems characterized by a random observation time have been considered in a certain closed manner for stochastic processes with stationary independent increments, more exactly: for processes belonging to the exponential class of processes. For a detailed survey one can study Winkler's paper ([14]), published in the same volume of the Banach Center Publications. In sequential considerations of several Markov processes the important fact is observable that basic investigation methods known for exponential-class processes are often applicable to other subclasses of Markov processes too. Thus Do Sun Bai ([4]) used such methods and obtained results on efficient sequential estimators for transition probabilities in finite-state Markov chains. In a paper by Trybuła ([13]) sequential estimation procedures for transition intensities in time-continuous finite-state Markov chains are treated.

In this paper sequential estimators are studied in birth-and-death processes, in an application in a queueing model and, last but not least, in exponential-class processes with minimax and Bayes procedures. The first part deals with a linear birth-and-death process with immigration in the zero state. This is a consideration similar to that in [8]. This special process stands for an example, and in other birth-and-death processes one can find analogous results. So, in the second part we investigate efficient sequential estimators in a queueing model by using the same methods. But these results have a proper importance. Finally, in the case of exponential-class processes minimax and Bayes procedures are given to characterize efficient sequential estimators for the mean parameter θ. This advance seems to be applicable in proving that other efficiently estimable parameter functions $h(\theta)$ have minimax estimators too. And, in view to such investigations in birth-and-death processes, this knowledge will be useful.

2. Sequential estimation procedures in a linear birth-and-death process

We want to study a linear birth-and-death process with immigration in the zero-state as an illustrative example of our investigation methods. Especially, assertions on efficiently estimable functions of the unknown process parameters and on efficient sequential procedures have a more complex form than in other birth-and-death processes (see, for instance, Section 3).

A detailed text concerning sequential estimation problems in birth-and-death processes will appear in [8]. So, in this section we shall omit all proofs and we shall give a survey of the results.

2.1. Likelihood function and random observation time. Let $\{X(t), t \geqslant 0\}$ be a right-continuous Markov process on a probability space $[\Omega, \mathscr{F}, P]$ and with values in $\{0, 1, 2, \ldots\}$. Further, we assume that $P(X(0) = x_0) = 1$ and that the transition probabilities are given by

$$P(X(t+dt) = k \mid X(t) = i) = \begin{cases} \lambda i\, dt + o(dt), & k = i+1,\ i = 1, 2, \ldots, \\ \mu i\, dt + o(dt), & k = i-1,\ i = 1, 2, \ldots, \\ 1-(\lambda+\mu) i\, dt + o(dt), & k = i,\ i = 1, 2, \ldots, \\ \gamma\, dt + o(dt), & k = 1,\ i = 0, \\ 1-\gamma\, dt + o(dt), & k = i = 0, \\ o(dt) & \text{otherwise.} \end{cases} \tag{2.1}$$

The probability measure depends on the unknown parameter vector $\theta \in \Theta := \{(\gamma, \lambda, \mu)^T\colon 0 < \gamma, \lambda, \mu < \infty\}$: $P = P_\theta$. We remark that the state zero is not absorbent and the process jumps upwards or downwards only with jump altitude one.

Let $N_{ij}(t)$ be the number of jumps from i to j up to t $(i, j = 0, 1, 2, \ldots; i \neq j)$,

$$B(t) := \sum_{i=1}^{\infty} N_{i,i+1}(t), \qquad D(t) := \sum_{i=1}^{\infty} N_{i,i-1}(t).$$

The process $\{B(t), t \geqslant 0\}$ is a birth process with intensities $\lambda(i) = \lambda i$ and $\{D(t), t \geqslant 0\}$ is a death process with intensities $\mu(i) = \mu i$. For each $t \geqslant 0$, we have

$$X(t) = x_0 + N_{01}(t) + B(t) - D(t);$$

the process $\{X(t), t \geqslant 0\}$ is called a *linear homogeneous birth-and-death process with immigration into the zero-state.*

We denote by $\mathscr{F}_t$, $t \geqslant 0$, the sub-σ-fields of $\mathscr{F}$ generated by $\{(N_{01}(s), B(s), D(s)), s \leqslant t\}$. Moreover, let P_θ^t denote the restriction of P_θ on $\mathscr{F}_t$. According to Billingsley's book ([2]), for any $\theta_1, \theta_2 \in \Theta$, the two restricted measures $P_{\theta_1}^t$ and $P_{\theta_2}^t$ are mutually absolutely continuous and the likelihood function (Radon–Nikodym derivative) with respect to (w.r.t.) a dominating measure $P^t_{\theta_0}$ has the form

$$L_\theta(t) := \frac{dP^t_\theta}{dP^t_{\theta_0}} = \left(\frac{\gamma}{\gamma_0}\right)^{N_{01}(t)} \exp\{-(\gamma-\gamma_0)\,T_0(t) - (\lambda+\mu-\lambda_0-\mu_0)\sum_{i=1}^{\infty} iT_i(t) + $$
$$+\ln\left(\frac{\lambda}{\lambda_0}\right)\sum_{i=1}^{\infty} N_{i,i+1}(t) + \ln\left(\frac{\mu}{\mu_0}\right)\sum_{i=1}^{\infty} N_{i,i-1}(t)\}$$

where θ_0 is a fixed value in Θ and $T_i(t)$ is the total waiting time in the state i up to $t\,(i = 0, 1, 2, \ldots)$. If we agree that

$$S(t) := \sum_{i=1}^{\infty} iT_i(t) = \int_0^t X(u)\,du \tag{2.2}$$

(we shall call $S(t)$ *integral time*) and

$$W(t) := (T_0(t),\, N_{01}(t),\, S(t),\, B(t),\, D(t))^T,$$

then we write $L_\theta(t)$ as

$$L_\theta(t) = g(W(t))\,\gamma^{N_{01}(t)} \exp\{-\gamma T_0(t) - (\lambda+\mu)\,S(t) + \ln\lambda\, B(t) + \ln\mu\, D(t)\} \tag{2.3}$$

where $g(W(t))$ is the residual factor not depending on θ. From the representation of the likelihood function we can easily see that $W(t)$ is a sufficient statistic w.r.t. P^t_θ.

A sequential estimation procedure is characterized by a stopping time τ w.r.t. $\{\mathscr{F}_t,\, t \geqslant 0\}$, which plays the role of the observation time of the underlying process. Properties of τ essentially determine the sequential procedures.

The pair (τ, φ) will be called a *sequential estimation procedure*, where φ denotes an estimator for any parameter function $h(\theta)$ and φ is based on the observation of the process up to τ. We are interested only in finite procedures (τ, φ), which means that $P(\tau < \infty) = 1$.

Now, the likelihood function in the sequential case follows from a more general result published by Sudakov in 1969 ([12]) or from the general assertions proved by Döhler ([3]) or the results in [8]. Observing the birth-and-death process up to τ, we then obtain the likelihood function in the sequential case:

$$L_\theta(\tau) = g(W(\tau))\,\gamma^{N_{01}(\tau)} \exp\{-\gamma T_0(\tau) - (\lambda+\mu)\,S(\tau) + \ln\lambda B(\tau) + \ln\mu\, D(\tau)\}. \tag{2.4}$$

Furthermore, we are able to prove that $W(\tau)$ is sufficient w.r.t. P^τ_θ, where $P^\tau_\theta = P_\theta|\mathscr{F}_\tau$ is the measure restricted on the τ-past $\mathscr{F}_\tau$.

2.2. Inequality of the Cramér–Rao type. At first, we consider some moment relations in our birth-and-death process. In connection with this we shall give the sequential maximum likelihood estimator of θ. Finally, we look at the special Cramér–Rao inequality associated with sequential estimators in our process.

Let us deal with the vector of derivatives of $\ln L_\theta(\tau)$ with respect to all components of θ:

$$U_\theta(\tau) := \mathrm{grad}_\theta(\ln L_\theta(\tau)) = \left(\frac{N_{01}(\tau)}{\gamma} - T_0(\tau),\ \frac{B(\tau)}{\lambda} - S(\tau),\ \frac{D(\tau)}{\mu} - S(\tau)\right)^T \tag{2.5}$$

Using the equation $\int L_\theta(\tau)\, dP_\theta = 1$ and the interchange of integration w.r.t. P_{θ_0} and differentiation w.r.t. θ, we find that the expectation of $U_\theta(\tau)$ is finite for all $\theta \in \Theta$ and

$$E_\theta U_\theta(\tau) = 0.$$

Especially, we have

$$E_\theta N_{01}(\tau) = \gamma E_\theta T_0(\tau), \qquad E_\theta B(\tau) = \lambda E_\theta S(\tau), \qquad E_\theta D(\tau) = \mu E_\theta S(\tau), \tag{2.6}$$

where the left-hand sides in (2.6) are finite if $E_\theta T_0(\tau) < \infty$ and $E_\theta S(\tau) < \infty$.

By differentiating once more we get

$$\Sigma := E_\theta U_\theta(\tau) U_\theta^T(\tau) = -E_\theta \mathrm{grad}_\theta U_\theta^T(\tau)$$
$$= E_\theta \begin{vmatrix} \dfrac{N_{01}(\tau)}{\gamma^2} & 0 & 0 \\ 0 & \dfrac{B(\tau)}{\lambda^2} & 0 \\ 0 & 0 & \dfrac{D(\tau)}{\mu^2} \end{vmatrix}. \tag{2.7}$$

And finally, if φ is a measurable function with $E_\theta \varphi = h(\theta)$ and $h(\theta)$ differentiable w.r.t. θ, then

$$\mathrm{grad}_\theta(h(\theta)) = E_\theta[U_\theta(\tau)\varphi]. \tag{2.8}$$

Now, we are interested in sequential maximum likelihood estimators. Basing ourselves on the likelihood function (2.4), we write down the likelihood equation $U_\theta(\tau) = 0$. So, we are able to determine the estimators as the solution of the equation given above:

$$\hat{\theta}(\tau) = \begin{vmatrix} N_{01}(\tau)/T_0(\tau) \\ B(\tau)/S(\tau) \\ D(\tau)/S(\tau) \end{vmatrix} =: \begin{vmatrix} \hat{\gamma}(\tau) \\ \hat{\lambda}(\tau) \\ \hat{\mu}(\tau) \end{vmatrix}. \tag{2.9}$$

Combining (2.9) with the representation (2.5) of $U_\theta(\tau)$, we obtain

$$U_\theta(\tau) = \begin{vmatrix} \frac{1}{\gamma} T_0(\tau) & 0 & 0 \\ 0 & \frac{1}{\lambda} S(\tau) & 0 \\ 0 & 0 & \frac{1}{\mu} S(\tau) \end{vmatrix} \begin{vmatrix} \hat{\gamma}(\tau) - \gamma \\ \hat{\lambda}(\tau) - \lambda \\ \hat{\mu}(\tau) - \mu \end{vmatrix}$$

$$=: I_\theta(\tau)\big(\hat{\theta}(\tau) - \theta\big). \tag{2.10}$$

This relation is similar to the well-known characterization of efficient estimators in the nonsequential case. We can also easily see that $E_\theta I_\theta(\tau) = \Sigma$, i.e., $I_\theta(\tau)$ is a random version of the Fisher information. But, in more general cases, we want to estimate optimally certain given functions $h(\theta)$ of the unknown process parameter θ. So, we have to find finite sequential estimation procedures (τ, φ) such that $E_\theta \varphi = h(\theta)$ and φ has minimal variances.

Since $W(\tau)$ is sufficient w.r.t. P_θ, we can restrict our considerations to estimators $\varphi = \varphi(W(\tau))$. (According to the Rao-Blackwell theorem $\varphi(W(\tau))$ has minimal variance.) Using the moment relations (2.6)-(2.8) given above and considering the variable

$$\Psi = (\operatorname{grad}_\theta h(\theta))^T \Sigma^{-1} U_\theta(\tau) - (\varphi - h(\theta)),$$

we are able to prove the following

THEOREM 2.1. *Let τ be a finite stopping time and let $E_\theta T_0(\tau) < \infty$, $E_\theta S(\tau) < \infty$. Assume that $h(\theta)$ is a differentiable function w.r.t. θ and $\varphi(W(\tau))$ is its unbiased estimator with $E_\theta \varphi^2(W(\tau)) < \infty$. Then, for all $\theta \in \Theta$,*

$$D_\theta^2 \varphi(W(\tau)) \geqslant \frac{\gamma}{E_\theta T_0(\tau)} \left(\frac{\partial h(\theta)}{\partial_\gamma}\right)^2 + \frac{1}{E_\theta S(\tau)} \left[\lambda \left(\frac{\partial h(\theta)}{\partial \lambda}\right)^2 + \mu \left(\frac{\partial h(\theta)}{\partial \mu}\right)^2\right]. \tag{2.11}$$

In (2.11) *the equality sign is valid for $\theta = \theta^*$ if and only if the estimator φ can be represented as*

$$\varphi(W(\tau)) = \frac{1}{E_{\theta^*} T_0(\tau)} \frac{\partial h(\theta^*)}{\partial \gamma} [N_{01}(\tau) - \gamma^* T_0(\tau)] +$$

$$+ \frac{1}{E_{\theta^*} S(\tau)} \left[\frac{\partial h(\theta^*)}{\partial \gamma} (B(\tau) - \lambda^* S(\tau)) + \frac{\partial h(\theta^*)}{\partial \mu} (D(\tau) - \mu^* S(\tau))\right] + h(\theta^*). \tag{2.12}$$

Now, it is usual to take this theorem as a foundation in defining efficiency. We say that a sequential procedure (τ, φ) is *efficient at $\theta = \theta^*$* if, for τ, $h(\theta)$ and φ in (2.11), the equality sign holds at $\theta = \theta^*$. The sequential procedure (τ, φ) is said to be *efficient* if it is efficient for all $\theta \in \Theta$. The estimator φ in an efficient procedure (τ, φ) is also called *efficient*. The parameter function $h(\theta)$ is said to be *efficiently estimable* if there exists an unbiased efficient estimator φ.

2.3. Efficient sequential estimation procedures. In the sequel we are especially interested in efficient sequential estimation procedures. At first we shall characterize efficient procedures in such a way that the distribution of $W(\tau)$ is accumulated in a 4-dimensional hyperplane of the state space. Then, using such a characterization, we are able to derive the structure of efficiently estimable functions.

Under the same conditions as in Theorem 2.1 the characterization of efficient procedures is induced (analogically to the statement for exponential-class processes in [9]).

THEOREM 2.2. *If the sequential procedure (τ, φ) is efficient, then there exist a vector $c = (c_1, \ldots, c_5)^T$ of real coefficients with $c^T c > 0$ and a constant $d \neq 0$ such that*

$$c^T W(\tau) = d \qquad a.s.\ (P_\theta). \tag{2.13}$$

We remark that the random observation time τ in an efficient sequential procedure (τ, φ) is determined by (2.13). In order to solve the estimation problem mentioned in Section 2.1 we have to carry out two steps:

(i) selection of τ by a specialized form of (2.13)

(ii) search for an optimal estimator φ for a given $h(\theta)$.

Now, we want to answer the question what kind of parameter functions are efficiently estimable:

THEOREM 2.3. *Under the conditions of Theorem 2.1 the only efficiently estimable functions $h(\theta)$ are*

(i)

$$h_1(\gamma) = \frac{\alpha_1 + \alpha_2 \gamma}{c_1 + c_2 \gamma} \qquad (c_1^2 + c_2^2 > 0,\ \lambda < \mu), \tag{2.14}$$

(ii)

$$h_2(\lambda, \mu) = \frac{\alpha_3 + \alpha_4 \lambda + \alpha_5 \mu}{c_3 + c_4 \lambda + c_5 \mu} \qquad (c_3^2 + c_4^2 + c_5^2 > 0), \tag{2.15}$$

where α_i $(i = 1, \ldots, 5)$ are real numbers.

The corresponding efficient procedures (τ, φ) and the corresponding estimators φ are given by

(i)

$$c_1 T_0(\tau) + c_2 N_{01}(\tau) = d, \tag{2.16}$$

$$\varphi_1 = \begin{cases} k_1 + k_2 \dfrac{N_{01}(\tau)}{d} & \text{if} \quad c_1 \neq 0, \\ \tilde{k}_1 \dfrac{T_0(\tau)}{d} + \tilde{k}_2 & \text{if} \quad c_1 = 0 \end{cases} \tag{2.17}$$

$(k_1 = \alpha_1/c_1,\ \tilde{k}_1 = \alpha_1,\ k_2 = \alpha_2 - \alpha_1 c_2/c_1,\ \tilde{k}_2 = a_2/c_2)$;

(ii)

$$c_3 S(\tau) + c_4 B(\tau) + c_5 D(\tau) = d, \tag{2.18}$$

$$\varphi_2 = \begin{cases} k_3 + k_4 \dfrac{B(\tau)}{d} + k_5 \dfrac{D(\tau)}{d} & \text{if} \quad c_3 \neq 0, \\ \tilde{k}_3 \dfrac{S(\tau)}{d} + \tilde{k}_4 \dfrac{B(\tau)}{d} + \tilde{k}_5 & \text{if} \quad c_3 = 0 (c_5 \neq 0) \end{cases} \tag{2.19}$$

$(k_3 = \alpha_3/c_3,\ \tilde{k}_3 = \alpha_3,\ k_4 = \alpha_4 + c_4 \alpha_3/c_3,$
$\tilde{k}_4 = \alpha_4 - c_4 \alpha_5/c_5,\ k_5 = \alpha_5 + c_5 \alpha_3/c_3,\ \tilde{k}_5 = \alpha_5/c_5).$

As an example which is based on the assertion of Theorem 2.3 we want to estimate efficiently the function $h(\lambda, \mu) = \dfrac{\lambda}{\lambda + \mu}$, i.e., the jump probability upwards. According to our investigation we select the procedure (τ, φ) characterized by

$$B(\tau) + D(\tau) = d \qquad \text{a.s. } (P_\theta),$$

where d is any given integer, and then we have to apply the estimator

$$\varphi(W(\tau)) = \frac{1}{d} B(\tau),$$

i.e., we have to register the birth number up to τ.

In Theorem 2.3 the problem of finiteness of an efficient sequential procedure is not treated. In what follows a brief investigation of finiteness is given. But, before we have specified the conditions on the coefficients c_i $(i = 1, \ldots, 5)$, we shall introduce a random time transformation.

Let us denote by

$$\sigma(s) = \inf\{t\colon\ S(t) = s\} \tag{2.20}$$

a new time scale. We define the transformed process $\{\tilde{X}(s),\ s \geqslant 0\}$, where $\tilde{X}(s) = X(\sigma(s))$ (this means an extension or a pressing of the realizations of the underlying process). We use the representation (see (2.2))

$$S(t) = \sum_{i=0}^{N(t)} X(t_i)(t_{i+1} - t_i) \tag{2.21}$$

of the integral time, where $N(t)$ gives the total number of all jumps ($N(t) = N_{01}(t)+B(t)+D(t)$) up to t and t_i ($i = 1, 2, \ldots; t_0 = 0$) are the (random) jump points of $\{X(t), t \geq 0\}$. Now, the jump points of the transformed process are determined by

$$\tilde{t}_i = S(t_i), \qquad i = 0, 1, 2, \ldots$$

and, according to [1], we have the following properties:

(i) $\tilde{t}_{i+1}-\tilde{t}_i = X(t_i)(t_{i+1}-t_i)$ ($i = 0, 1, 2, \ldots$) is exponential distributed with the parameter $(\lambda+\mu)$,

(ii) $\tilde{X}(\tilde{t}_{i+1})-\tilde{X}(\tilde{t}_i) = X(t_{i+1})-X(t_i)$, independent of $\tilde{t}_{i+1}-\tilde{t}_i$,

(iii) the processes $\{\tilde{B}(s), s > 0\}$ and $\{\tilde{D}(s), s \geq 0\}$ corresponding to $\{\tilde{X}(s), s \geq 0\}$ are independent Poisson processes.

As an illustrative case we look at (ii), $c_3 \neq 0$, in Theorem 2.3. If $\tilde{\tau} = S(\tau)$, the characteristic equation (2.18) becomes

$$c_3\tilde{\tau}+c_4\tilde{B}(\tilde{\tau})+c_5\tilde{D}(\tilde{\tau}) = d \tag{2.22}$$

and, if we use the knowledge of exponential-class processes, the procedure is finite if, choosing $d > 0$, we obtain

$$c_3 > 0, \quad c_4 \leq 0, \quad c_5 \leq 0 \qquad \text{and} \qquad c_3+c_4\lambda+c_5\mu > 0.$$

Moreover, we get

$$E_\theta\tilde{\tau} = E_\theta S(\tau) = d(c_3+c_4\lambda+c_5\mu)^{-1}. \tag{2.23}$$

We add that in case (ii), $c_3 = 0$, the same time transformation (2.20) must be applied; in case (i) the time transformation $\sigma(s) = \inf\{t\colon T_0(t) = s\}$ is taken and the reasoning is similar to that applied above.

2.4. Asymptotic properties of efficient estimators. We shall study the asymptotic distribution and the convergence behaviour of a sequence of estimators. Applying methods precisely given for exponential-class processes, we are able to prove that a sequence of sequential estimators is strongly consistent and asymptotically normal. We use, as in [6], the following notion of consistency.

Denote by $\{\sigma_d, d\in D\}$ a sequence of increasing random variables, where $D = R^+$ or $D = \{0, 1, 2, \ldots\}$. Let $\{\varphi_d, d\in D\}$ with $\varphi_d = \varphi(W(\sigma_d))$ be a sequence of estimators for $h(\theta)$ and let $\sigma_d/d \to a$ (= const.) in prob. (with prob. 1) as $d \to \infty$. Then, $\{\varphi_d, d\in D\}$ is said to be *weakly* (*strongly*) *consistent* with respect to $\{\sigma_d\}$ if $\varphi_d \to h(\theta)$ in prob. (with prob. 1) as $d \to \infty$.

Now, applying time transformations introduced above and some results for exponential-class processes, we get

THEOREM 2.4. *Let $\{\tau_d, d\in D\}$ be a sequence of finite stopping times and let $h_1(\gamma)$, $h_2(\lambda, \mu)$ and $\varphi_{1,d} = \varphi(T_0(\tau_d), N_{01}(\tau_d))$, $\varphi_{2,d} = \varphi(S(\tau_d), B(\tau_d), D(\tau_d))$*

be, respectively, the only efficiently estimable functions and the corresponding efficient estimators. Moreover, assume that $\lambda < \mu$. *Then we have*

(i) $\{\varphi_{i,d},\ d \in D\}$ *is strongly consistent w.r.t.* $\{\sigma_d\}$ *for* $h_i(\theta)$ $(i = 1, 2)$, *where* $\sigma_d = T_0(\tau_d)$ *or* $\sigma_d = S(\tau_d)$;

(ii) $\varphi_{i,d}$ *is asymptotically* $(d \to \infty)$ *normal with the mean* $h_i(\theta)$ $(i = 1, 2)$.

3. Sequential estimation problems in queueing models

In this section we would like to point out that the results obtained in sequential procedures in birth-and-death processes are applicable in some queueing situations. The underlying investigations are made by Rüger in his diploma thesis ([11]). At first, a problem of machine service is treated. Secondly we comment briefly on a more complicated queueing model.

3.1. Characteristic process of an $M_N/M/s$-queueing model. In this special queueing model we have a situation in which N machines stand in a machine room and each of those machines has an exponentially distributed (with parameter λ) random life time, independent of any other. In a service station s repairmen are engaged in repair, one for each defective machine. The repair time is exponentially distributed with intensity μ.

Now we are interested in the characteristic process $\{X(t),\ t \geqslant 0\}$ which gives the number of defective machines at time t (waiting for repair or being under repair). The state space is $E = \{0, 1, 2, \ldots, N\}$. We assume that $P_\theta(X(0) = x_0) = 1$. The transition intensities are

$$q_{i,i+1} = \lambda_i = \lambda(N-i) \quad (0 \leqslant i \leqslant N),$$

$$q_{i,i-1} = \mu_i = \begin{cases} i\mu & (0 \leqslant i \leqslant s), \\ s\mu & (s < i \leqslant N), \end{cases}$$

$$q_i = -(\lambda_i + \mu_i) \qquad (0 \leqslant i \leqslant N).$$

The waiting time in each state i $(0 \leqslant i \leqslant N)$ is exponentially distributed with parameter $(\lambda_i + \mu_i)$.

In our statistical investigations we look upon the parameter $\theta = (\lambda, \mu)^T$ as the unknown parameter where $\theta \in \{(\lambda, \mu)^T\colon\ 0 < \lambda,\ \mu < +\infty\}$. According to Billingsley's book ([2]) we are able to give the following representation of the likelihood function of probability measure P_θ w.r.t. a dominating measure P_{θ_0} for fixed $\theta_0 \in \Theta$

$$L_\theta(t) = \left(\frac{\lambda}{\lambda_0}\right)^{B(t)} \left(\frac{\mu}{\mu_0}\right)^{D(t)} \exp\{-(\lambda-\lambda_0)Nt + (\lambda-\lambda_0)S(t) - (\mu-\mu_0)T^s(t)\}, \tag{3.1}$$

where $B(t)$ is the number of jumps $i \to i+1$, $D(t)$ is the number of jumps i

$\to i-1$, $S(t) = \int_0^t X(u)\,du$ (integral time) and $T^s(t) = \int_0^t X^s(u)\,du$ (total repair time) with $X^s(u) = \min(X(u), s)$. If we set $T(t) = Nt - S(t)$, then, according to the factorization theorem of densities, we can easily see that

$$W(t) = (T(t),\ T^s(t),\ B(t),\ D(t))^T$$

is a sufficient statistic w.r.t. P_θ^t.

If we want to estimate λ and μ, then we quickly obtain the maximum likelihood estimators:

$$\hat{\lambda}(t) = \frac{B(t)}{T(t)}, \qquad \hat{\mu}(t) = \frac{D(t)}{T^s(t)}. \tag{3.2}$$

Maximum likelihood estimators have good asymptotical properties. But, in general, it is difficult to explore properties in the case of finite observation time or, naturally, in the case of estimating functions of λ and μ and a finite observation time. We hit upon an expedient by changing over to sequential estimation procedures. We formulate our general estimation problem:

For a given function $h(\theta)$ of the unknown parameter $\theta = (\lambda, \mu)^T$ we have to search for a finite random observation time τ and for an estimator φ such that φ is unbiased and with minimal variance.

3.2. Sequential estimators and efficiently estimable parameter functions. Analogically to our consideration in Section 2.1 (for instance, by application of the Sudakov lemma ([12])) we are able to write down the sequential case of the likelihood function:

$$L_\theta(\tau) = g(W(\tau))\,\lambda^{B(\tau)}\mu^{D(\tau)}\exp\{-\lambda T(\tau) - \mu T^s(\tau)\} \tag{3.3}$$

where $g(W(\tau))$ is the residual factor which does not depend on $\theta = (\lambda, \mu)^T$. Moreover, we know that $W(\tau) = (T(\tau),\ T^s(\tau),\ B(\tau),\ D(\tau))^T$ is sufficient w.r.t. the probability measure P_θ^τ restricted on the σ-field of the τ-past. Here, it is also possible to give sequential maximum likelihood estimators for λ and μ:

$$\hat{\lambda}(\tau) = \frac{B(\tau)}{T(\tau)}, \qquad \hat{\mu}(\tau) = \frac{D(\tau)}{T^s(\tau)}.$$

Again we put $U_\theta(\tau) := \mathrm{grad}_\theta(\ln L_\theta(\tau))$. Hence we have

$$U_\theta(\tau) = \left(\frac{B(\tau)}{\lambda} - T(\tau),\ \frac{D(\tau)}{\mu} - T^s(\tau)\right)^T. \tag{3.4}$$

It is easy to prove that the following moment relations hold:

(i) $$E_\theta B(\tau) = \lambda E_\theta T(\tau), \qquad E_\theta D(\tau) = \mu E_\theta T^s(\tau), \tag{3.5}$$

$$\Sigma := E_\theta U_\theta(\tau) U_\theta^T(\tau) = \begin{vmatrix} \frac{1}{\lambda} E_\theta T(\tau) & 0 \\ 0 & \frac{1}{\mu} E_\theta T^s(\tau) \end{vmatrix}; \tag{3.6}$$

(ii) for φ independent of θ, and $E_\theta \varphi = h(\theta)$ differentiable:

$$\operatorname{grad}_\theta h(\theta) = E_\theta [U_\theta(\tau)\varphi]. \tag{3.7}$$

Now, on the analogy of Theorem 2.1 we formulate an inequality of the Rao–Cramér type for the lower bound of the variance of the estimators φ; here we mention that, according to the Rao–Blackwell theorem, we restrict our interest to estimators depending on $W(\tau)$ only: $\varphi = \varphi(W(\tau))$:

$$D_\theta^2 \varphi(W(\tau)) \geqslant \frac{\lambda}{E_\theta T(\tau)} \left(\frac{\partial h(\theta)}{\partial \lambda}\right)^2 + \frac{\mu}{E_\theta T^s(\tau)} \left(\frac{\partial h(\theta)}{\partial \mu}\right)^2. \tag{3.8}$$

We find the structure of efficient estimators:

$$\varphi(W(\tau)) = \operatorname{grad}_\theta h(\theta) \Sigma^{-1} U_\theta(\tau) + h(\theta) =$$

$$\left(\frac{\partial h(\theta)}{\partial \lambda}, \frac{\partial h(\theta)}{\partial \mu}\right) \begin{vmatrix} \lambda(E_\theta T(\tau))^{-1} & 0 \\ 0 & \mu(E_\theta T^s(\tau))^{-1} \end{vmatrix} \begin{vmatrix} \lambda^{-1} B(\tau) - T(\tau) \\ \mu^{-1} D(\tau) - Ts(\tau) \end{vmatrix} + h(\theta) \tag{3.9}$$

The investigations of efficient estimators are like those of the birth-and-death process in Section 2. We can apply the same methods. So we find the following efficiently estimable functions and the corresponding estimators and the finite sequential procedures:

(i) $$h_1(\theta) = h_1(\lambda) = k_1 - \frac{\tilde{k}_2 \lambda}{c_1 - \lambda c_2}, \qquad \varphi_1 = k_1 + \frac{\tilde{k}_2}{d} B(\tau),$$

$$c_1 T(\tau) + c_2 B(\tau) = d \qquad (c_1 \leqslant \lambda c_2,\ c_2 < 0,\ \tilde{k}_2 \neq 0); \tag{3.10}$$

(ii)

$$h_2(\theta) = h_2(\lambda) = k_2 - \frac{\tilde{k}_1}{c_2 \lambda}, \qquad \varphi_2 = \frac{\tilde{k}_1}{d} T(\tau) + k_2,$$

$$c_2 B(\tau) = d \qquad \left(\tilde{k}_1 \neq 0, \frac{d}{c_2} \text{ integer-valued}\right); \tag{3.11}$$

(iii)

$$h_3(\theta) = h_3(\mu) = k_3 - \frac{\tilde{k}_4 \mu}{c_3 + \mu c_4}, \qquad \varphi_3 = k_3 + \frac{\tilde{k}_4}{d} D(\tau),$$

$$c_3 T^s(\tau) + c_4 D(\tau) = d \qquad (-c_3 < c_4 \mu,\ c_4 < 0,\ \tilde{k}_4 \neq 0); \tag{3.12}$$

(iv)

$$h_4(\theta) = h_4(\mu) = k_4 - \frac{\tilde{k}_3}{c_4 \mu}, \qquad \varphi_4 = -\frac{\tilde{k}_3}{d} T^s(\tau) + k_4$$

$$c_4 D(\tau) = d \qquad \left(\tilde{k}_4 \neq 0, \frac{d}{c_4} \text{ integer-valued}\right) \tag{3.13}$$

where $\tilde{k}_i = dk_i (i = 1, \ldots, 4)$.

3.3. Semi-Markov processes characterizing queueing models. In this section we want to outline an investigation of characteristic processes in a more complicated queueing model. We select an example of a $GI/M/s/\infty$-queueing system in which customers arrive at time points t_i and the differences $T_i = t_{i+1} - t_i$ $(i = 0, 1, \ldots)$ are independent identically distributed random variables in which all service times are independent and identically exponentially distributed and where there exist s service places and ∞ waiting places.

The characteristic process $\{X(t), t \geq 0\}$ of the system is the number of customers found in the system at time t (waiting or in service). We consider the imbedded Markov chain $\{X_k := X(t_k + 0), k = 1, 2, \ldots\}$ in the case $s = 1$. If we choose as input distributions gamma distributions, then we obtain the transition probabilities

$$P_{ij} = P(X_{k+1} = j \mid X_k = i) = \left(\frac{b}{\mu+b}\right)^p \left(\frac{\mu}{\mu+b}\right)^{i-j+1} \frac{\Gamma(i-j+p)}{\Gamma(p)(i-j+1)!} \tag{3.14}$$

and the likelihood function of the underlying process

$$L_\theta(t) = \prod_{k=1}^{N(t)} g(\cdot) P_{X_k X_{k+1}}(\theta) f_\theta(T_k)$$

$$= g(\cdot) b^{pN(t)} e^{-bt} \left(\frac{b}{\mu+b}\right)^{pN(t)} \left(\frac{\mu}{\mu+b}\right)^{M(t)} \tag{3.15}$$

where

$$M(t) = N(t) + \sum_{i=1}^{x_0 + N(t)} \sum_{j=1}^{i+1} (i-j) N_{ij}(t),$$

$N_{ij}(t)$ is the number of jumps from i to j and $N(t)$ is the total number of all jumps up to t. We find that $W(t) := (t, N(t), M(t))^T$ is a sufficient statistic.

Now we are able to carry out considerations of efficient sequential estimators similar to birth-and-death processes, but the mathematical tools are more complicated.

4. Sequential minimax estimation in exponential-class processes

In our considerations of efficient sequential estimation procedures in birth-and-death processes we could see that such procedures are not completely determined. For instance, the real number d in characteristic equation (2.13) is not fixed; each possible value of d belongs to an efficient procedure.

As a possibility to determine efficient estimation procedures completely we choose a risk function consisting of the mean value of a quadratic loss function and of the mean value of a continuous cost function. Here we shall investigate this method in the case of exponential-class processes and we use minimax and Bayes procedures. Since there exist some connections between sequential investigations of exponential-class and birth-and-death processes, we hope that this method is also applicable in birth-and-death processes.

Now, we are interested in optimal estimators. For this purpose we have to minimize the risk function. It turns out that, with a proper weight function and for the estimation of the mean parameter in exponential-class processes, the sequential minimax procedures reduce to fixed-time procedures. In the case of one-dimensional exponential-class processes such investigations are carried out for Poisson, negative-binomial, gamma and Wiener processes by Dvoretzky, Kiefer and Wolfowitz ([5]) and in a one-dimensional generalization by Magiera [10].

4.1. Notions and a basic result. Let $\{X(t),\ t\in T\}$ be an m-dimensional exponential-class process as defined in the article given by Winkler ([14]) in the same volume. Such a stochastic process is characterized by the Radon–Nikodym derivatives of the distribution of $X(t)$ (w.r.t. any dominating measure ν):

$$f(x,t,\theta)=\frac{dP_\theta^{(t)}}{d\nu}(x)=g(x,t)\exp\{a^T(\theta)x+b_0(\theta)t\}\quad (t\in T,\ x\in R^m) \tag{4.1}$$

where $\theta=(\theta_1,\ldots,\theta_m)^T\in\Theta\subset R^m$ is the unknown parameter vector, $a(\theta) = (a_1(\theta),\ldots,a_m(\theta))^T$ and $a_i(\theta)$ $(i=1,\ldots,m)$, $b_0(\theta)$ are nonnegative differentiable parameter functions.

If we put

$$A=\operatorname{grad}_\theta a^T(\theta)\qquad\text{and}\qquad b=\operatorname{grad}_\theta b_0(\theta)$$

and if A^{-1} exists, then for the expected value and for the covariance matrix we have

$$E_\theta X(t)=-A^{-1}bt,\qquad K_t=-A^{-1}\operatorname{grad}_\theta(A^{-1}b)^Tt.$$

Now we are interested only in cases in which $A^{-1}b=-\theta$, i.e.,

$$E_\theta X(t)=\theta t,\qquad K_t=A^{-1}t. \tag{4.2}$$

For a given parameter function $h(\theta)$ we denote by φ an unbiased estimator and by $\delta = (\tau, \varphi)$ the sequential procedure consisting of a finite Markov stopping time τ and of the estimator φ. Let Δ be the set of all such sequential procedures δ. The loss involved in using the estimator φ when $h(\theta)$ is the correct value is denoted by $L(h, \varphi)$. We assume that $c(t)$ containing the costs of the inspection up to time t is nonnegative, continuous and increasing in t. Then

$$R(h, \delta) = E_\theta\big(L(h, \varphi) + c(\tau)\big) \tag{4.3}$$

is the risk function, where the existence of $R(h, \delta)$ is assumed for all $\theta \in \Theta$.

A sequential estimation procedure $\delta^* = (\tau^*, \varphi^*)$ is called *minimax* if

$$\sup_{\theta\in\Theta} R\big(h(\theta), \delta^*\big) = \inf_{\delta\in\Delta}\sup_{\theta\in\Theta} R\big(h(\theta), \delta\big). \tag{4.4}$$

Let $\mathfrak{G}$ be a σ-field of Borel subsets of Θ. We define by $F(\theta)$ a probability distribution function on $(\Theta, \mathfrak{G})$ (a priori distribution). Assume that $R\big(h(\theta), \delta\big)$ is a $\mathfrak{G}$-measurable function of θ. Then, for given $\delta \in \Delta$, the expected risk w.r.t. $F(\theta)$ is defined by

$$r(F, \delta) = \int_\Theta R\big(h(\theta), \delta\big)\, dF(\theta).$$

A sequential estimation procedure $\delta^* = (\tau^*, \varphi^*)$ is called a *Bayes procedure for F* if

$$r(F, \delta^*) = \inf_{\delta\in\Delta} r(F, \delta). \tag{4.5}$$

If $\tau = t$ ($t > 0$, fixed) with probability 1 then the procedure $\delta^0 = (t, \varphi^0)$ is said to be a *fixed-time procedure*. In this case we can restrict our considerations to

$$\tilde{R}(h, \varphi^0) = E_\theta L\big(h(\theta), \varphi^0\big) \tag{4.6}$$

and

$$\tilde{r}(F, \varphi^0) = \int_\Theta \tilde{R}\big(h(\theta), \varphi^0\big)\, dF(\theta). \tag{4.7}$$

An estimator $\hat{\varphi}^0$ is said to be a *t-Bayes estimator for* $h(\theta)$ if (4.7) assumes its minimum for $\varphi^0 = \hat{\varphi}^0$.

We assume that the a posteriori distribution function $F(\theta \mid x)$, given $X(t) = x$, is defined for a fixed-time procedure $\delta^0 = (t, \varphi^0)$ and for the a priori distribution $F(\theta)$. Then, the a posteriori t-risk corresponding to F and φ^0 is defined by

$$\tilde{r}(F, \varphi^0 \mid x) = \int_\Theta L\big(h(\theta), \varphi^0\big)\, dF(\theta \mid x). \tag{4.8}$$

We remark that a t-Bayes estimator minimizes the a posteriori t-risk corresponding to F and φ^0.

Now we want to give a well-known method of solving for minimax procedures in the case $h(\theta) = \theta$:

LEMMA (Dvoretzky, Kiefer, Wolfowitz ([5])). *Suppose that, for every $t > 0$, there exists a sequence of distribution functions F_n $(n = 1, 2, \ldots)$ for which there exist corresponding t-Bayes estimators $\hat{\varphi}_n^0$ with the property that the a posteriori t-risk associated with F_n and $\hat{\varphi}_n^0$ is independent of the value of the sufficient statistic $X(t)$. Further, assume that there exists an estimator φ^0 for which*

$$\tilde{R}(t) := \sup_{\theta \in \Theta} \tilde{R}(\theta, \varphi^0) = \lim_{n \to \infty} \tilde{r}(F_n, \hat{\varphi}_n^0). \tag{4.9}$$

If there exists a t_0 $(0 < t_0 < \infty)$ for which

$$c(t_0) + \tilde{R}(t_0) = \min_{t > 0} (c(t) + \tilde{R}(t)) \tag{4.10}$$

holds, then the fixed-time procedure $\delta^0 = (t_0, \varphi^0)$, for $t = t_0$, is minimax.

4.2. Bayes and minimax sequential estimation for the multidimensional exponential-class processes. The following considerations represent the multidimensional generalization of results obtained by Magiera ([10]). We shall derive a sequence of t-Bayes estimators for the parameter vector θ. For this investigation we apply suitable a priori parameter densities determined on the principle "closed under sampling". Choosing a quadratic weighted loss function, we can see that the corresponding a posteriori t-risk is independent of the value $X(t) = x$. If certain technical conditions are fulfilled, then according to the lemma of Dvoretzky et al. we get a minimax estimator φ^0 for θ.

We now look at the density $f(x, t, \theta)$ in (4.1). Only the exponential factor

$$l_\theta(t) := \exp\{a^T(\theta)x + b_0(\theta)t\}$$

depends on θ. We assume that $b_0(\theta)$ can be split so that

$$b_0(\theta) = \sum_{i=1}^{m} b_i(\theta).$$

Putting $b(\theta) := (b_1(\theta), \ldots, b_m(\theta))^T$, $t := (t, \ldots, t)^T$ with m components, we write

$$l_\theta(t) = \exp\{a^T(\theta)x + b^T(\theta)t\}. \tag{4.11}$$

Further, we assume that Θ is an m-dimensional interval, i.e.,

$$\Theta = (\underline{\theta}_1, \bar{\theta}_1) \times \ldots \times (\underline{\theta}_m, \bar{\theta}_m).$$

In order to give the a priori parameter densities we make the following assumption:

A1. There exist a factorization of $g(x, t)$ in (4.1), real numbers $r_i^* \in [-1, 1]$ and measurable functions H_i $(i = 1, \ldots, m)$ such that, for all $x \in R^m$,

(i) $g(x, t) = \prod_{i=1}^{m} g_i(x, t)$,

(ii) $\int_{\underline{\theta}_i}^{\bar{\theta}_i} l_\theta(t)\, d\theta_i = \frac{1}{(t - r_i^*) g(x, t)} H_i(\theta, x, t)$,

(iii) $\int_{\Theta} l_\theta(t)\, d\theta = \left(\prod_{i=1}^{m} (t - r_i^*) g_i(x, t)\right)^{-1}$.

Under A1 the term $\prod_{i=1}^{m} (t - r_i^*) g_i(x, t) \exp\{a^T(\theta)x + b^T(\theta)t\}$ yields a density w.r.t. θ. According to the principle "closed under sampling" we choose "small" values for x and t. If we put

$$r_i := \max(0, r_i^*), \qquad r := (r_1, \ldots, r_m)^T, \qquad u := \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)^T,$$

we get a suitable sequence of a priori parameter densities $(n = 1, 2, \ldots)$:

$$f_n(\theta) = \prod_{i=1}^{m} \left(\frac{1}{n} + r_i - r_i^*\right) g_i\left(r, \frac{1}{n} + r_i\right) \exp\{a^T(\theta) r + b^T(\theta)(u + r)\}. \quad (4.12)$$

Let the process $\{X(t), t \in T\}$ be observed up to t and $X(t) = x$. Then the a posteriori densities are

$$f_n(\theta \mid x) = \frac{l_\theta(t) f_n(\theta)}{\int_{\Theta} l_\theta(t) f_n(\theta)\, d\theta} = \prod_{i=1}^{m} \left(t + \frac{1}{n} + r_i - r_i^*\right) \times$$

$$\times g_i\left(r + x, t + \frac{1}{n} + r_i\right) \exp\{a^T(\theta)(r + x) + b^T(\theta)(u + r + t)\}. \quad (4.13)$$

In the next step we choose the loss function representing the loss of the statistician making a false decision

$$L(\theta, \varphi) = (\varphi - \theta)^T \bar{A} (\varphi - \theta) \quad (4.14)$$

where $\bar{A} = (A^T A)^{1/2}$ and A is given in (4.2).

Using the following technical assumptions, we are able to determine the minimal a posteriori t-risk and the corresponding t-Bayes estimators.

A2. There exists an inverse matrix M^{-1} where $M = \int_{\Theta} \bar{A} f_n(\theta \mid x)\, d\theta$.

A3. For each $x \in R^m$ and $t > r$, we have $(i = 1, \ldots, m)$

(i) $\lim_{\theta_i \to \underline{\theta}_i} \exp\{a^T(\theta)x + b^T(\theta)t\} = \lim_{\theta_i \to \bar{\theta}_i} \exp\{a^T(\theta)x + b^T(\theta)t\}$,

(ii) $\lim_{\theta_i \to \underline{\theta}_i} \theta_i \exp\{a^T(\theta)x + b^T(\theta)t\} = \lim_{\theta_i \to \bar{\theta}_i} \theta_i \exp\{a^T(\theta)x + b^T(\theta)t\}$.

A4. With

$$B = \operatorname{grad}_\theta(b^T(\theta)) \qquad \text{and} \qquad V_n = \begin{vmatrix} t+1/n+r_1 & & 0 \\ & \ddots & \\ 0 & & t+1/n+r_m \end{vmatrix}$$

the following relations are valid:

(i) $\int_\Theta (B(t+u+r)-V_n b) f_n(\theta \mid x)\, d\theta = 0$,

(ii) $\int_\Theta \theta^T V_n^{-1}(B(t+u+r)-V_n b) f_n(\theta \mid x)\, d\theta = 0$,

(iii) $\int_\Theta (D\bar{A}-\bar{A}D) f_n(\theta \mid x)\, d\theta = 0$ for a diagonal matrix D independent of θ,

(iv) $\int_\Theta \theta^T (V_n^{-1}\bar{A}-\bar{A}V_n^{-1}) f_n(\theta \mid x)\, d\theta = 0$.

A.5. For all $r_i^* \in [-1, 1]$ $(i = 1, \ldots, m)$:

$$\sup_{\theta\in\Theta}\left[\sum_{i=1}^m \sum_{j=1}^m a_{ij}\frac{1}{t+r_i}\frac{1}{t+r_j}a^{ij}t+\theta^T\bar{A}(\theta-V^{-1}\theta t)- \right.$$
$$\left. -\theta^T V^{-1}\bar{A}(\theta t-V^{-1}\theta t^2)\right] = \sum_{i=1}^m \frac{1}{t+r_i},$$

where

$$\bar{A} = (a_{ij}), \qquad \bar{A}^{-1} = (a^{ij}) \qquad \text{and} \qquad V = \begin{vmatrix} t+r_1 & & 0 \\ & \ddots & \\ 0 & & t+r_m \end{vmatrix}.$$

Now, assuming that integration w.r.t θ and differentiation w.r.t. φ_n^0 are interchangeable, we minimize the a posteriori t-risks $(n = 1, 2, \ldots)$

$$\tilde{r}(F_n, \varphi_n^0 \mid x) = \int_\Theta L(\theta, \varphi_n^0) f_n(\theta \mid x)\, d\theta \tag{4.15}$$

and we get

$$2\int_\Theta \bar{A}(\varphi_n^0-\theta) f_n(\theta \mid x)\, d\theta = 0. \tag{4.16}$$

Finally under the conditions. A2–A4, we obtain the solutions

$$\hat{\varphi}_n^0 = V_n^{-1}(x+r) = \begin{vmatrix} (t+1/n+r_1)^{-1} & & 0 \\ & \ddots & \\ 0 & & (t+1/n+r_m)^{-1} \end{vmatrix} \begin{vmatrix} x_1+r_1 \\ \vdots \\ x_m+r_m \end{vmatrix}$$

and therefore the minimal t-risks

$$\tilde{r}(F_n, \hat{\varphi}_n^0 \mid x) = \sum_{i=1}^m \left(t+\frac{1}{n}+r_i\right)^{-1},$$

which are independent of $X(t) = x$.

We summarize this in

THEOREM 4.1. *Assume that* A1–A4 *are fulfilled. Then, for processes belonging to the exponential class and for the loss function* (4.14), *the t-Bayes estimators w.r.t.* $F_n(\theta)$ $(n = 1, 2, \ldots)$ *are*

$$\hat{\varphi}_n^0 = \left(\frac{X_1(t)+r_1}{t+\frac{1}{n}+r_1}, \ldots, \frac{X_m(t)+r_m}{t+\frac{1}{n}+r_m}\right)^T \tag{4.17}$$

and the corresponding a posteriori t-risks are

$$\tilde{r}(F_n, \hat{\varphi}_n^0 \mid x) = \sum_{i=1}^{m} \left(t+\frac{1}{n}+r_i\right)^{-1}. \tag{4.18}$$

Theorem 4.1 and the lemma of Dvoretzky et al. allow us to assert what follows:

THEOREM 4.2. *Suppose that* A1–A5 *are valid. Then, for processes belonging to the exponential class and for the loss function* (4.14), *the fixed-time procedure* $\delta^0 = (t_0, \varphi^0)$ *is minimax, where* t_0 *is the minimal t-value in*

$$c(t) + \sum_{i=1}^{m} (t+r_i)^{-1} \tag{4.19}$$

and φ^0 *is the estimator*

$$\varphi^0 = V^{-1} X(t_0) = \left(\frac{X_1(t_0)}{t_0+r_1}, \ldots, \frac{X_m(t_0)}{t_0+r_m}\right)^T. \tag{4.20}$$

We remark that the assumptions A1–A5 are fulfilled in a large number of known processes which are also important in practical applications. Particularly, one-dimensional exponential-class processes, for instance Poisson processes, Wiener processes, negative-binomial processes and gamma processes, fulfil the conditions in question. Here, it should be mentioned that in the case $m = 1$ condition A5 turns into

$$\sup_{\theta \in \Theta} (r^2 \theta^2 A) = r,$$

which is equal to Magiera's condition (see [10]).

4.3. Examples and a view of applications. At first, let $\{X(t), t \in T\}$ be an m-dimensional Wiener process. Here, we have $a^T(\theta) = \theta^T K^{-1}$, $b_0(\theta) = -\frac{1}{2}\theta^T K^{-1}\theta$, $A = K^{-1}$ and $b^T = -\theta^T K^{-1}$, where K is the covariance matrix at $t = 1$, not depending on θ.

We obtain $r_1 = \ldots = r_m = 0$. Conditions A1–A5 are fulfilled. The t-Bayes estimators are (see (4.17))

$$\hat{\varphi}_n^0 = \left(\frac{X_1(t)}{t+\frac{1}{n}}, \ldots, \frac{X_m(t)}{t+\frac{1}{n}}\right)^T, \tag{4.21}$$

the a posteriori t-risks (see (4.18)) are

$$\tilde{r}(F_n, \hat{\varphi}_n^0 \mid x) = m\left(t+\frac{1}{n}\right)^{-1}, \tag{4.22}$$

and the minimax estimator (see (4.20)) are

$$\varphi^0 = \frac{1}{t_0} X(t_0). \tag{4.23}$$

Second, let $\{X(t), t \in T\}$ be the process with the two independent components $X_1(t)$ and $X_2(t)$ where $X_1(t)$ is Poisson distributed and $X_2(t)$ is gamma-distributed. The density (4.1) takes on the following form

$$f(x_1, x_2, \theta_1, \theta_2) = \frac{t^{x_1} x_2^{t-1}}{x_1! \, \Gamma(t)} \exp\left\{\ln \theta_i x_1 - \frac{x_2}{\theta_2} - (\theta_1 + \ln \theta_2) t\right\}$$

$$(x_1 = 0, 1, \ldots, x_2 \geqslant 0, \ 0 < \theta_1, \ \theta_2 < \infty).$$

According to A1 we obtain $r_1 = 0$, $r_2 = 1$ and the a priori densities

$$f_n(\theta_1, \theta_2) = \frac{1}{n\Gamma(1/n)} \theta_2^{-1/n-1} \exp\left\{-\frac{1}{\theta_2} - \frac{\theta_1}{n}\right\}.$$

Finally, conditions A2–A5 are fulfilled and we get the t-Bayes estimators

$$\hat{\varphi}_n^0 = \left(\frac{X_1(t)}{t+\frac{1}{n}}, \frac{X_2(t)+1}{t+\frac{1}{n}+1}\right)^T, \tag{4.24}$$

the a posteriori t-risks $(n = 1, 2, \ldots)$

$$\tilde{r}(F_n, \hat{\varphi}_n^0 \mid x) = \frac{1}{t+\frac{1}{n}} + \frac{1}{t+\frac{1}{n}+1} \tag{4.25}$$

and the minimax estimator

$$\varphi^0 = \left(\frac{X_1(t_0)}{t_0}, \frac{X_2(t_0)}{t_0+1}\right)^T. \tag{4.26}$$

We remark that the second component does not represent an unbiased estimator for θ_2.

An important application of the above results consists in pointing out that other efficiently estimable parameter functions $h(\theta)$ also have minimax estimators. Considering a stopped exponential-class process, we can prove that, for a sequence of stopping times τ_d $(d \in D)$, the sequence of states corresponding to the random times τ_d also represents an exponential-class process. Therefore, modifying our investigations of this section, we find minimax procedures for these parameter functions $h(\theta)$.

Surely, such a knowledge is useful in investigations of birth-and-death processes.

References

[1] K. B. Athreya and P. E. Ney, *Branching Processes*, Springer-Verlag, Berlin–Heidelberg–New York 1972.

[2] P. Billingsley, *Statistical Inference for Markov Processes*, Univ. of Chicago Press 1961.

[3] R. Döhler, *Dominierbarkeit und Suffizienz in der Sequentialanalyse*, Math. Operationsforsch. Statist. Ser. Statist. **12** (1981), 101–134.

[4] Do Sun Bai, *Efficient estimation of transition probabilities in a Markov chain*, Ann. Statist. **3** (1975), 1305–1317.

[5] A. Dvoretzky, J. Kiefer, and J. Wolfowitz, *Sequential decision problems for processes with continuous time parameter. Problems of estimation*, Ann. Math, Statist. **24** (1953), 403–415.

[6] J. Franz, *Consistency and completeness in sequential estimation problems*, Math. Operationsforsch. Statist. Ser. Statist. **10** (1979), 127–139.

[7] —, *A note on sequential minimax estimation for exponential class processes*, Preprint 07–13–81, Technische Universität Dresden, Dresden 1981.

[8] —, *Sequential estimation and asymptotic properties in birth-and-death processes*, Math. Operationsforsch. Statist. Ser. Statist. **13** (1982), 231–244.

[9] J. Franz and W. Winkler, *Sequential estimation problems for the exponential class of processes with independent increments*, Scand. J. Statist. **6** (1979), 129–139.

[10] R. Magiera, *On sequential minimax estimation for the exponential class of processes*, Zastos. Mat. **15** (1977), 434–445.

[11] W. Rüger, *Sequentialle Schätzverfahren in der Bedienungstheorie*, Diplomarbeit Technische Universität Dresden, Sektion Mathematik, Dresden 1981.

[12] V. N. Sudakov, *On measures defined by Markov stopping times*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **12** (1969), 157–164, in Russian.

[13] S. Trybuła, *Sequential estimation in finite state Markov processes*, Instytut Matematyki, Komunikat Politech. Wrocław no. 120 (1977).

[14] W. Winkler, *A survey of sequential estimation in processes with independent increments*, this volume, 545–554.

Presented to the semester
Sequential Methods in Statistics
September 7–December 11, 1981