ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XVII, 2 (1982)

S. TRYBUŁA (Wrocław)

SEQUENTIAL ESTIMATION IN FINITE-STATE MARKOV PROCESSES

1. In the paper, the Cramér-Rao inequality for statistical structures is applied, in the sequential case, to some stochastic processes the distribution of which depends on a finite number of parameters. Examples for some processes with independent increments and some Markov processes are presented. The problem of efficient estimation of a function of the intensity matrix of a homogeneous finite-state Markov process is considered and all efficient sequential plans for which the efficiently estimable function depends on one row of the intensity matrix are found. A negative result concerning the existence of efficient plans in the case where the efficiently estimable function depends on two or more rows of the intensity matrix is proved.

The problem of determining efficient sequential plans has been intensively investigated. For binomial samples this problem was solved by De Groot [6]. He formulated important definitions (1) generalized later by other authors. Bhat and Kulkarni [2] solved the problem for multinomial samples, and Do Sun Bai [3] obtained some results for the m-state Markov chain. Efficient sequential estimation for continuous time processes was considered by Trybuła [11]. He investigated the problem for the Poisson process and for some other processes with independent increments. The exponential class of processes was considered by Magiera [7].

In Section 2 of this paper we apply the result of Różański [8] who generalized the important Sudakov lemma [10]. In Sections 3 and 4 we use the results of Do Sun Bai [3] and in Section 5 those of De Groot [6].

2. Let $x(t) = (x_1(t), x_2(t), \ldots, x_r(t))$, $t \geqslant 0$ or $t = 0, 1, 2, \ldots$, be a stochastic process defined on the probability space $(\Omega, \mathscr{F}, P_\theta)$, where Ω is a space of r-dimensional vector-valued right-continuous functions $\omega = x(\cdot)$ for which left-hand limits exist, $\mathscr{F}$ is the least σ-algebra with respect to which all $x(t)$ are measurable, P_θ is a probability measure defined on $\mathscr{F}$ and dependent on a parameter θ. We suppose that $\theta = (\theta_1, \theta_2, \ldots, \theta_m)$ takes its values in an open set $K \subset R^m$.

(1) See also [5].

Let $\mathscr{F}_t$ be the least σ-algebra with respect to which all $x(s)$ are measurable if $s \leqslant t$ and let $P_{\theta,t}$ be the measure P_θ defined on the σ-algebra $\mathscr{F}_t$. Denote by

$$Z(\omega, t) = \big(Z_1(\omega, t), Z_2(\omega, t), \ldots, Z_l(\omega, t)\big)$$

a mapping $\Omega \to R^l$ $\mathscr{F}_t$-measurable and right-continuous with respect to $t \in P_\theta$ almost surely. Let us suppose that the measures $P_{\theta,t}$, $\theta \in K$, are absolutely continuous with respect to the measure $P_{\theta_0,t}$, $\theta_0 \in K$, and that the density function is

$$\frac{dP_{\theta,t}}{dP_{\theta_0,t}} = g\big(t, Z(\omega, t), \theta, \theta_0\big),$$

where g is a continuous function.

Let T be the set of values of $t, T = [0, \infty)$ or $T = \{0, 1, 2, \ldots\}$, and let $\tau(\omega)$ be a random variable mapping $\Omega \to T$ and such that $\{\omega\colon \tau(\omega) \leqslant t\} \in \mathscr{F}_t$ for each $t \in T$. Moreover, let $P_\theta\big(0 < \tau(\omega) < \infty\big) = 1$ for each $\theta \in K$. The random variable $\tau(\omega)$ is called a *sequential plan.*

Let $U = T \times R^l$ and let $\mathscr{B}$ be the σ-field of Borel subsets of U. On $(U, \mathscr{B})$ we define the measure m_θ by

$$m_\theta(A) = P_\theta\Big(\big(\tau(\omega), Z\big(\omega, \tau(\omega)\big)\big) \in A\Big) \quad \text{for each } A \in \mathscr{B}.$$

It follows from the results of Różański [8] that the measures m_θ, $\theta \in K$, are absolutely continuous with respect to m_{θ_0}, and the density function is of the form

$$\frac{dm_\theta}{dm_{\theta_0}} = g(s, z, \theta, \theta_0), \tag{1}$$

where s is a value of $\tau(\omega)$ and z is a value of $Z\big(\omega, \tau(\omega)\big)$.

We observe the process $x(t)$ in the interval $(0, \tau]$ and want to estimate the function $h(\theta)$. An estimator $f(s, z)$ of $h(\theta)$ is a $\mathscr{B}$-measurable function defined on U such that $\mathrm{D}^2\big(f\big(\tau, Z(\tau)\big)\big) < \infty$ and

$$\mathrm{E}(f) = \int_U f(s, z) g(s, z, \theta, \theta_0) m_{\theta_0}(du) = h(\theta). \tag{2}$$

From the assumption that $\mathrm{P}\big(\tau(\omega) < \infty\big) = 1$ we obtain

$$\int_U g(s, z, \theta, \theta_0) m_{\theta_0}(du) = 1. \tag{3}$$

We suppose that we can pass with the derivatives with respect to θ_i $(i = 1, 2, \ldots, m)$ in (2) and (3) under the sign of integration. Let

$$W_i = \frac{\partial \ln g\big(\tau(\omega), Z\big(\omega, \tau(\omega)\big), \theta, \theta_0\big)}{\partial \theta_i}, \quad W = (W_1, W_2, \ldots, W_m).$$

We consider only such sequential plans τ for which

$$\mathrm{E}\Big(\sum_{i=1}^{m} W_i^2\Big) < \infty \quad \text{for each } \theta \in K.$$

From (3) we obtain

$$\mathrm{E}(W_i) = 0 \quad (i = 1, 2, \ldots, m). \tag{4}$$

Moreover, from (2) we get

$$\mathrm{E}(W_i f) = h_i'(\theta) \quad (i = 1, 2, \ldots, m), \tag{5}$$

where $h_i'(\theta) = \partial h / \partial \theta_i$.

Put

$$\Sigma = \mathrm{E}\,\|W_i W_j\|.$$

Let us suppose that for each $\theta \in K$ we have $|\Sigma| \neq 0$. Put

$$\Sigma^{-1} = \|\sigma_{ij}(\theta)\|.$$

Let

$$r_i = \mathrm{E}(W_i f) \quad \text{and} \quad R = (r_1, r_2, \ldots, r_m) \overset{(5)}{=} (h_1', h_2', \ldots, h_m').$$

Then

$$R\Sigma^{-1}R^{\mathrm{T}} = \sum_{i,j=1}^{m} \sigma_{ij} h_i' h_j',$$

where R^{T} is the matrix transposed to R.

Write $W = (W_1, W_2, \ldots, W_m)$.

We are in the conditions of the Cramér-Rao inequality for statistical structures ([1], p. 52, Theorem 5), which implies

$$\mathrm{D}^2(f) \geqslant R\Sigma^{-1}R^{\mathrm{T}} = \sum_{i,j=1}^{m} \sigma_{ij}(\theta) h_i'(\theta) h_j'(\theta). \tag{6}$$

In (6) we have equality for a particular value of θ if and only if for this value the equality

$$f = R\Sigma^{-1}W^{\mathrm{T}} + h(\theta) \tag{7}$$

holds m_{θ_0}-almost surely.

Let us suppose that we can pass in (3) with the second derivatives with respect to θ_i and θ_j $(i, j = 1, 2, \ldots, m)$ under the sign of integration. Then we obtain the second equation of sequential analysis in the form

$$\mathrm{E}(W_i W_j) = -\mathrm{E}\left(\frac{\partial^2 \ln g}{\partial \theta_i \partial \theta_j}\right) \quad (i, j = 1, 2, \ldots, m) \tag{8}$$

with $g = g\big(\tau(\omega), Z(\omega, \tau(\omega)), \theta, \theta_0\big)$, and $\sigma_{ij}(\theta)$ can be easily evaluated by (6).

Examples

1. *Multinomial process.* Let $X = (X_1, X_2, \ldots, X_m)$ be a random variable with the multinomial distribution. X takes the values

$$e_1 = (1, 0, \ldots, 0), \quad e_2 = (0, 1, \ldots 0), \quad \ldots \quad e_m = (0, 0, \ldots, 1)$$

with probabilities

$$\text{(9)} \qquad \mathrm{P}(X = e_i) = p_i \quad (i = 1, 2, \ldots, m),$$

where $0 < p_i < 1$ and

$$\sum_{i=1}^{m} p_i = 1.$$

Let $x(1), x(2), \ldots$ be independent and distributed according to (9). Let $x(t) = \big(x_1(t), x_2(t), \ldots, x_m(t)\big), t = 1, 2, \ldots,$ and assume that

$$N_i(t) = \sum_{j=1}^{t} x_i(j) \quad (i = 1, 2, \ldots, m).$$

Let $\theta = (p_1, p_2, \ldots, p_{m-1})$ and $\theta_0 = (p_1^{(0)}, p_2^{(0)}, \ldots, p_{m-1}^{(0)})$. We obtain

$$\frac{dP_{\theta,t}}{dP_{\theta_0,t}} = \left(\frac{p_1}{p_1^{(0)}}\right)^{N_1(t)} \left(\frac{p_2}{p_2^{(0)}}\right)^{N_2(t)} \ldots \left(\frac{p_m}{p_m^{(0)}}\right)^{N_m(t)}, \quad \text{where} \sum_{i=1}^{m} N_i(t) = t.$$

Then we can put

$$Z(t) = \big(N_1(t), N_2(t), \ldots, N_{m-1}(t)\big)$$

and write

$$\frac{dP_{\theta,t}}{dP_{\theta_0,t}} = q\big(t, Z(t), \theta_0\big) p_1^{N_1(t)} p_2^{N_2(t)} \ldots p_m^{N_m(t)}.$$

Hence we obtain the equation

$$\frac{dm_\theta}{dm_{\theta_0}} = q(s, z, \theta_0) p_1^{z_1}, p_2^{z_2} \ldots p_m^{z_m},$$

where $s, z_1, z_2, \ldots, z_m$ are values of the random variables $\tau, N_1(\tau), N_2(\tau), \ldots$ $\ldots, N_m(\tau)$, respectively.

Let $f(s, z)$ be an unbiased estimator of $h(\theta)$. From (6) we obtain

$$\mathrm{D}^2(f) = \frac{1}{\mathrm{E}(\tau)} \Big[\sum_{j=1}^{m-1} p_j (h_j'(\theta))^2 - \Big(\sum_{j=1}^{m-1} p_j h_j'(\theta)\Big)^2\Big],$$

which was proved by Bhat and Kulkarni in [2].

2. *Finite-state Markov chain.* Let $x(0), x(1), x(2), \ldots$ be an m-state Markov chain with initial probabilities $p_1, p_2, \ldots, p_m$ and transition matrix $\|p_{ij}\|$, where

$$0 < p_i < 1, \quad \sum_{i=1}^{m} p_i = 1, \quad 0 < p_{ij} < 1, \quad \sum_{j=1}^{m} p_{ij} = 1$$
$$(i, j = 1, 2, \ldots, m).$$

Let

$$V_k = \begin{cases} 1 & \text{if } x(0) = k, \\ 0 & \text{otherwise,} \end{cases}$$

and let $N_{ij}(t)$ be the number of jumps of the process from the state i to the state j up to the moment t.

Let

$$\theta = (p_1, \ldots, p_{m-1}, p_{11}, \ldots, p_{1,m-1}, \ldots, p_{m1}, \ldots, p_{m,m-1}).$$

Then we obtain

$$\frac{dP_{\theta,t}}{dP_{\theta_0,t}} = \prod_{k=1}^{m} \left(\frac{p_k}{p_k^{(0)}}\right)^{V_k} \prod_{i,j=1}^{m} \left(\frac{p_{ij}}{p_{ij}^{(0)}}\right)^{N_{ij}(t)} \stackrel{\mathrm{df}}{=} q\big(t, Z(t), \theta_0\big) \prod_{k=1}^{m} p_k^{V_k} \prod_{i,j=1}^{m} p_{ij}^{N_{ij}(t)},$$

where

$$Z(t) = \big(V_1, \ldots, V_m, N_{11}(t), \ldots, N_{1m}(t), N_{21}(t), \ldots, N_{2m}(t), \ldots, N_{m1}(t), \ldots, N_{m,m-1}(t)\big).$$

In this case we get

$$\frac{dm_\theta}{dm_{\theta_0}} = q(s, z, \theta_0) \prod_{k=1}^{m} p_k^{z_k} \prod_{i,j=1}^{m} p_{ij}^{z_{ij}},$$

where $s, z_1, \ldots, z_m, z_{11}, \ldots, z_{mm}$ are values of the random variables $\tau, V_1, \ldots, V_m, N_{11}(\tau), \ldots, N_{m,m}(\tau)$, respectively. From (6) we obtain

$$\text{(10)} \quad \mathrm{D}^2(f) \geqslant \sum_{k=1}^{m-1} p_k (h_k')^2 - \Big(\sum_{k=1}^{m-1} p_k h_k'\Big)^2 + \sum_{i=1}^{m} \frac{1}{\mathrm{E}\big(N_i(\tau)\big)} \Big[\sum_{j=1}^{m-1} p_{ij} (h_{ij}')^2 - \Big(\sum_{j=1}^{m-1} p_{ij} h_{ij}'\Big)^2\Big],$$

where

$$N_i(\tau) = \sum_{j=1}^{m} N_{ij}(\tau), \quad h'_k = \frac{\partial h}{\partial p_k}, \quad h_{ij} = \frac{\partial h}{\partial p_{ij}}.$$

Inequality (10) was obtained by Do Sun Bai in [3].

3. *Finite-state Markov process.* Let $x(t)$ be an m-state homogeneous Markov process with initial probabilities $p_1, p_2, \ldots, p_m$, intensity matrix Λ, $0 < p_i < 1$ $(i = 1, 2, \ldots, m)$, $\sum_{i=1}^{m} p_i = 1$, and with intensities $\lambda_{ij} > 0$ $(i, j = 1, 2, \ldots, m, i \neq j)$.

Let

$$V_k = \begin{cases} 1 & \text{if } x(0) = k, \\ 0 & \text{otherwise,} \end{cases}$$

and let $N_{ij}(t)$ be the number of jumps of the process from the state i to the state j in the interval $(0, t]$, $i, j = 1, 2, \ldots, m$, $i \neq j$. Let $T_i(t)$ be the joint time of staying the process in the state i in the interval $(0, t]$.

Let

$$\theta = (p_1, \ldots, p_{m-1}, \lambda_{12}, \ldots, \lambda_{1m}, \ldots, \lambda_{m1}, \ldots, \lambda_{m,m-1})$$

and let

$$\lambda_i = \sum_{\substack{j=1 \\ j \neq i}}^{m} \lambda_{ij}.$$

From [4] we obtain

$$(11) \qquad \frac{dP_{\theta,t}}{dP_{\theta_0,t}} = \prod_{k=1}^{m} \left(\frac{p_k}{p_k^{(0)}}\right)^{V_k} \prod_{\substack{i,j=1 \\ i \neq j}}^{m} \left(\frac{\lambda_{ij}}{\lambda_{ij}^{(0)}}\right)^{N_{ij}(t)} \exp\Big(\sum_{r=1}^{m} (\lambda_r^{(0)} - \lambda_r) T_r(t)\Big)$$

$$\stackrel{\text{df}}{=} q\big(t, Z(t), \theta_0\big) \prod_{k=1}^{m} p_k^{V_k} \prod_{\substack{i,j=1 \\ i \neq j}}^{m} (\lambda_{ij})^{N_{ij}(t)} \exp\Big(-\sum_{t=1}^{m} \lambda_r T_r(t)\Big),$$

where

$$Z(t) = \big(V_1, \ldots, V_m, N_{12}(t), \ldots, N_{1m}(t), \ldots, N_{m1}(t), \ldots, N_{m,m-1}(t), \\ T_1(t), \ldots, T_{m-1}(t)\big).$$

Then

$$\frac{dm_\theta}{dm_{\theta_0}} = q(s, z, \theta_0) \prod_{k=1}^{m} p_k^{v_k} \prod_{\substack{i,j=1 \\ i \neq j}}^{m} \lambda_{ij}^{n_{ij}} \exp\Big(-\sum_{r=1}^{m} \lambda_r t_r\Big),$$

where $v_1, \ldots, v_m, n_{12}, \ldots, n_{m,m-1}, t_1, \ldots, t_m$ are the values of $V_1, \ldots, V_m$, $N_{12}(\tau), \ldots, N_{m,m-1}(\tau), T_1(\tau), \ldots, T_m(\tau)$, respectively.

Inequality (6) takes now the form

$$\text{(12)} \qquad \mathrm{D}^2(f) \geqslant \sum_{k=1}^{m-1} p_k (h_k')^2 - \Big(\sum_{k=1}^{m-1} p_k h_k'\Big)^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{m} \frac{\lambda_{ij}(h_{ij}')^2}{\mathrm{E}\big(T_i(\tau)\big)},$$

where $h_k' = \partial h/\partial p_k$ and $h_{ij}' = \partial h/\partial \lambda_{ij}$.

4. *Multivariate Poisson process.* Let $x(t) = \big(x_1(t), x_2(t), \ldots, x_r(t)\big)$ be a process with independent increments for which $\mathrm{P}\big(x(0) = (0, 0, \ldots, 0)\big) = 1$ and

$$\mathrm{P}\big(x_1(t) = k_1, x_2(t) = k_2, \ldots, x_r(t) = k_r\big) = \frac{\lambda_1^{k_1}\lambda_2^{k_2}\ldots\lambda_r^{k_r}}{k_1!k_2!\ldots k_r!} t^k e^{-\lambda t},$$

where

$$k = \sum_{i=1}^{r} k_i, \qquad \lambda = \sum_{i=1}^{r} \lambda_i, \qquad \lambda_i > 0 \quad (i = 1, 2, \ldots, r).$$

Let $\theta = (\lambda_1, \lambda_2, \ldots, \lambda_m)$. We have

$$\frac{dP_{\theta,t}}{dP_{\theta_0,t}} = \prod_{i=1}^{r} \left(\frac{\lambda_i}{\lambda_i^{(0)}}\right)^{x_i(t)} \exp\big((\lambda^{(0)} - \lambda)t\big) \overset{\mathrm{df}}{=} q\big(t, Z(t), \theta_0\big) \prod_{i=1}^{r} \lambda_i^{x_i(t)} e^{-\lambda t},$$

where $Z(t) = x(t)$.

Then

$$\frac{dm_\theta}{dm_{\theta_0}} = q(s, z, \theta_0) \prod_{i=1}^{r} \lambda_i^{z_i} e^{-\lambda s},$$

where $z_1, z_2, \ldots, z_r$ are the values of $x_1(\tau), x_2(\tau), \ldots, x_r(\tau)$, and

$$\mathrm{D}^2(f) \geqslant \frac{1}{\mathrm{E}(\tau)} \sum_{i=1}^{r} \lambda_i (h_i')^2.$$

5. *Multivariate Wiener process.* Let $x(t) = \big(x_1(t), x_2(t), \ldots, x_r(t)\big)$ be an r-dimensional process with independent increments for which $\mathrm{P}\big(x(0) = (0, 0, \ldots, 0)\big) = 1$, $x(t)$ has a normal distribution with density

$$p(t, x) = \frac{1}{(2\pi t)^{r/2}} \exp\left(-\frac{\sum\limits_{i=1}^{r} (x_i - \lambda_i t)^2}{2t}\right).$$

Let $\theta = (\lambda_1, \lambda_2, \ldots, \lambda_r)$. We have

$$\frac{dP_{\theta,t}}{dP_{\theta_0,t}} = q\big(t, Z(t), \theta_0\big)\exp\Big(\sum_{i=1}^{r}\lambda_i x_i(t) - \frac{t}{2}\sum_{i=1}^{r}\lambda_i^2\Big),$$

where $Z(t) = x(t)$. Then

$$\frac{dm_\theta}{dm_{\theta_0}} = q(s, z, \theta_0)\exp\Big(\sum_{i=1}^{r}\lambda_i z_i - \frac{s}{2}\sum_{i=1}^{r}\lambda_i^2\Big),$$

where $z_1, z_2, \ldots, z_r$ are the values of $x_1(\tau), x_2(\tau), \ldots, x_r(\tau)$, respectively, and

$$\mathrm{D}^2(f) \geqslant \frac{1}{\mathrm{E}(\tau)}\sum_{i=1}^{r}(h_i')^2.$$

This inequality (in a more general form) was given by Skrjabin in [9].

6. *Finite-state birth and death process.* Let $x(t), t \geqslant 0$, be a birth and death process with values $0, 1, \ldots, m$ and intensities

$$\lambda_i = a_i\lambda \quad (i = 0, 1, \ldots, m-1),$$
$$\mu_i = b_i\mu \quad (i = 1, 2, \ldots, m)$$

for given $a_i > 0$ and $b_i > 0$. Let $\lambda > 0$, $\mu > 0$ and let $p = (p_0, p_1, \ldots, p_m)$, $0 < p_i < 1$, $\sum_{i=0}^{m} p_i = 1$, be the vector of initial probabilities. Then

$$\theta = (p_0, p_1, \ldots, p_{m-1}, \lambda, \mu),$$

and from [4] we obtain

$$\frac{dP_{\theta,t}}{dP_{\theta_0,t}} = \prod_{k=0}^{m}\Big(\frac{p_k}{p_k^{(0)}}\Big)^{V_k}\Big(\frac{\lambda_k}{\lambda^{(0)}}\Big)^{N^+(t)}\Big(\frac{\mu}{\mu^{(0)}}\Big)^{N^-(t)}\exp\big((\lambda^{(0)}-\lambda)T^{(1)}(t) + (\mu^{(0)}-\mu)T^{(2)}(t)\big),$$

where $N^+(t)$ and $N^-(t)$ are the numbers of jumps of the process to the higher and the lower state, respectively, in the interval $(0, t]$,

$$T^{(1)}(t) = \sum_{i=0}^{m-1} a_i T_i(t), \qquad T^{(2)}(t) = \sum_{i=1}^{m} b_i T_i(t),$$

$T_i(t)$ being the joint time of staying the process in the state i in the interval $(0, t]$, and V_k are defined similarly as in the previous examples. Then we have

$$Z(t) = \big(V_0, \ldots, V_m, N^+(t), N^-(t), T^{(1)}(t), T^{(2)}(t)\big)$$

and

$$\frac{dm_\theta}{dm_{\theta_0}} = q(z,\theta_0)\prod_{k=0}^{m} p_k^{z_k}\lambda^{z^+}\mu^{z^-}\exp(-\lambda\zeta_1-\mu\zeta_2),$$

where $z_0,\ldots,z_m,z^+,z^-,\zeta_1,\zeta_2$ are the values of the random variables $V_0,\ldots,V_m,N^+(\tau),N^-(\tau),T^{(1)}(\tau),T^{(2)}(\tau)$, respectively.

The information inequality takes now the form

$$D^2(f)\geqslant\sum_{k=0}^{m-1}p_k(h_k')^2-\Big(\sum_{k=0}^{m-1}p_kh_k'\Big)^2+\frac{\lambda(h_\lambda')^2}{E\big(T^{(1)}(\tau)\big)}+\frac{\mu(h_\mu')^2}{E\big(T^{(2)}(\tau)\big)},$$

where $h_k = \partial h/\partial p_k$, $h_\lambda' = \partial h/\partial\lambda$ and $h_\mu' = \partial h/\partial\mu$.

7. *Finite-state Markov process with related intensities.* Let $x(t)$ be a homogeneous Markov process with values $1,2,\ldots,m$ and intensities

$$\lambda_{ij} = a_{ij}\mu_j,$$

where $a_{ij}>0$ for $i,j=1,2,\ldots,m$, $i\neq j$, and $\mu_j>0$ for $j=1,2,\ldots,m$. We suppose that the constants a_{ij} are known. Assume that the initial probabilities $p_1,p_2,\ldots,p_m$ fulfil the conditions

$$0<p_i<1\quad(i=1,2,\ldots,m),\quad\sum_{i=1}^{m}p_i=1$$

and let V_k $(k=1,2,\ldots,m)$, $N_{ij}(t)$ $(i,j=1,2,\ldots,m,\ i\neq j)$ and $T_i(t)$ $(i=1,2,\ldots,m)$ be the random variables defined as in Example 3. Let us write

$$T_{.j}=\sum_{i=1}^{m}a_{ij}T_i\quad\text{and}\quad N_{.j}=\sum_{\substack{i=1\\ i\neq j}}^{m}N_{ij}.$$

Let $\theta=(p_1,p_2,\ldots,p_{m-1},\mu_1,\mu_2,\ldots,\mu_m)$. By (11) we have

$$\frac{dP_{\theta,t}}{dP_{\theta_0,t}}=q\big(t,Z(t),\theta_0\big)\prod_{k=1}^{m}p_k^{V_k}\prod_{j=1}^{m}\mu_j^{N_{.j}(t)}\exp\Big(-\sum_{r=1}^{m}\mu_rT_r(t)\Big),$$

where

$$Z(t)=\big(V_1,\ldots,V_m,N_{.1}(t),\ldots,N_{.m}(t),\ldots,T_{.1}(t),\ldots,T_{.m}(t)\big),$$

and we obtain

$$D^2(f)\geqslant\sum_{k=1}^{m-1}p_k\Big(\frac{\partial h}{\partial p_k}\Big)^2-\Big(\sum_{k=1}^{m-1}p_k\frac{\partial h}{\partial p_k}\Big)^2+\sum_{j=1}^{m}\frac{\mu_j}{E\big(T_{.j}(\tau)\big)}\Big(\frac{\partial h}{\partial\mu_j}\Big)^2.$$

3. Definition 1. For a fixed sampling plan τ:

(i) the non-constant estimator f_0 is *efficient for* $\mathrm{E}(f_0) = h(\theta)$ at $\theta^{(0)}$ if inequality (6) becomes equality for $f = f_0$ and $\theta = \theta^{(0)}$;

(ii) the estimator f is *efficient* if it is efficient for $\mathrm{E}(f_0) = h(\theta)$ for all values of $\theta \in K$.

(iii) the function $h(\theta)$ is *estimable* if there exists an unbiased estimator of this function;

(iv) the function $h(\theta)$ is *efficiently estimable at* $\theta^{(0)}$ if it is estimable and there exists an unbiased estimator of this function which is efficient at $\theta^{(0)}$;

(v) the function $h(\theta)$ is *efficiently estimable* if there exists an efficient estimator of this function.

Definition 2. The sequential sampling plan τ is *efficient at* $\theta^{(0)}$ if for τ there exists a non-constant function $h(\theta)$ which is efficiently estimable at θ_0. The sequential sampling plan τ is *efficient* if for τ there exists a non-constant function $h(\theta)$ which is efficiently estimable.

Let $x(t)$ be the m-state Markov process considered in Example 3. We write Z, N_{ij} and T_i instead of $Z(\tau), N_{ij}(\tau)$ and $T_i(\tau)$, respectively. Let us put

$$W_k = \frac{\partial \log g(\tau, Z, \theta, \theta_0)}{\partial p_k} = \frac{p_m V_k - p_k V_m}{p_k p_m} \quad (k = 1, 2, \dots, m-1),$$

$$W_{ij} = \frac{\partial \log g(\tau, Z, \theta, \theta_0)}{\partial \lambda_{ij}} = \frac{N_{ij} - \lambda_{ij} T_i}{\lambda_{ij}} \quad (i, j = 1, 2, \dots, m, i \neq j).$$

Then from (4) we obtain $\mathrm{E}(W_k) = \mathrm{E}(W_{ij}) = 0$. Moreover, by (7) we have the following

COROLLARY. *If under the sequential sampling plan τ a non-constant estimator f is efficient for $\mathrm{E}(f) = h(\theta)$ at $\theta^{(0)}$, then there exist constants a_k, b_{ij} $(i, j = 1, 2, \dots, m,\ i \neq j,\ k = 1, 2, \dots, m-1)$ not all equal to zero and a constant c such that*

$$f = \sum_{k=1}^{m-1} a_k (p_m v_k - p_k v_m) + \sum_{\substack{i,j=1 \\ i \neq j}}^{m} b_{ij} (n_{ij} - \lambda_{ij} t_i) + c \tag{13}$$

for almost all (with respect to measure m_0) points $u \in U$.

Let $\Lambda_i = (\lambda_{i1}, \dots, \lambda_{i,i-1}, \lambda_{i,i+1}, \dots, \lambda_{im})$ for $i = 1, 2, \dots, m$. In this section we find all efficient plans in the case where h is a function only of Λ_i (with i fixed). Without loss of generality we may suppose that $i = 1$.

Two distinct values

$$\Lambda_1^{(1)} = (\lambda_{12}^{(1)}, \lambda_{13}^{(1)}, \dots, \lambda_{1m}^{(1)}) \quad \text{and} \quad \Lambda_1^{(2)} = (\lambda_{12}^{(2)}, \lambda_{13}^{(2)}, \dots, \lambda_{1m}^{(2)})$$

are *equivalent* with respect to $h(\Lambda_1)$ if $h(\Lambda_1^{(1)}) = h(\Lambda_1^{(2)})$.

THEOREM 1. *If under the sequential sampling plan τ there exists a non-constant estimator f which is efficient for some function $h(\Lambda_1)$ for two values of Λ_1 which are not equivalent with respect to $h(\Lambda_1)$, then there exist constants $a_2, a_3, \ldots, a_m, \beta$ not all equal to zero and $\gamma \neq 0$ such that*

$$\sum_{j=2}^{m} a_j n_{1j} + \beta t_1 = \gamma$$

for almost all (with respect to measure m_{0_0}) points $u \in U$.

Proof. Inequality (12) in our case takes the form

$$\mathrm{D}^2(f) \geqslant \frac{1}{\mathrm{E}(T_1)} \sum_{j=2}^{m} \lambda_{1j} (h'_{1j})^2 .$$

By the Corollary, if the estimator f is efficient for $\mathrm{E}(f) = h(\Lambda_1)$ at $\Lambda_1^{(1)} = (\lambda_{12}^{(1)}, \lambda_{13}^{(1)}, \ldots, \lambda_{1m}^{(1)})$, then there exist constants $b_2^{(1)}, b_3^{(1)}, \ldots, b_m^{(1)}$ not all equal to zero and a constant $c^{(1)}$ such that

$$f = \sum_{j=2}^{m} b_j^{(1)} (n_{1j} - \lambda_{1j}^{(1)} t_1) + c^{(1)} \tag{14}$$

for m_{0_0}-almost all points $u \in U$.

Similarly, if f is efficient at $\Lambda_1^{(2)} = (\lambda_{12}^{(2)}, \lambda_{13}^{(2)}, \ldots, \lambda_{1m}^{(2)})$, then there exist constants $b_2^{(1)}, b_3^{(2)}, \ldots, b_m^{(2)}$ not all equal to zero and a constant $c^{(2)}$ such that

$$f = \sum_{j=2}^{m} b_j^{(2)} (n_{1j} - \lambda_{1j}^{(2)} t_1) + c^{(2)} \tag{15}$$

for m_{0_0}-almost all points $u \in U$. From (14) and (15) we obtain

$$\sum_{j=2}^{m} (b_j^{(2)} - b_j^{(1)}) n_{1j} - \sum_{j=2}^{m} (b_j^{(2)} \lambda_{1j}^{(2)} - b_j^{(1)} \lambda_{1j}^{(1)}) t_1 + c^{(2)} - c^{(1)} = 0 .$$

But $c^{(i)} = h(\Lambda^{(i)}) = \mathrm{E}_i(f)$ $(i = 1, 2)$, where $\mathrm{E}_i(f)$ is the expectation of f for $\Lambda_1 = \Lambda_1^{(i)}$. Moreover, $\Lambda_1^{(1)}$ and $\Lambda_1^{(2)}$ are not equivalent with respect to $h(\Lambda_1)$. Then $\gamma = c^{(1)} - c^{(2)} \neq 0$ and it follows that the constants

$$a_j = b_j^{(2)} - b_j^{(1)} \quad (j = 2, 3, \ldots, m) \quad \text{and} \quad \beta = -\sum_{j=2}^{m} (b_j^{(2)} \lambda_{1j}^{(2)} - b_j^{(1)} \lambda_{1j}^{(1)})$$

are not all equal to zero.

THEOREM 2. *Let τ be a sequential sampling plan for which* $\mathrm{P}(T_1 > 0) = 1$ *and there exist constants $a_2, a_3, \ldots, a_m$ not all equal to zero and $\gamma \neq 0$*

such that

$$\sum_{j=2}^{m} a_j n_{ij} + \beta t_1 = \gamma$$

for m_{θ_0}-almost all $u \in U$.

Then for m_{θ_0}-almost all $u \in U$ either

$$n_{1\sigma(2)} + n_{1\sigma(3)} + \ldots + n_{1\sigma(k)} = l \tag{16}$$

for some positive integer l, where $(\sigma(2), \sigma(3), \ldots, \sigma(m))$ is a permutation of $(2, 3, \ldots, m)$ and k is an integer, $2 \leqslant k \leqslant m$, or

$$t_1 = a$$

for some $a > 0$.

Proof. Without loss of generality we may assume that $\gamma > 0$. Let us consider all cases:

(i) Let $a_{j_0} < 0$ for some j_0. Since $\mathrm{E}(N_{ij}) = \lambda_{ij}\mathrm{E}(T_i)$, we obtain

$$(a_2\lambda_{12} + a_3\lambda_{13} + \ldots + a_m\lambda_{1m} + \beta)\mathrm{E}(T_1) = \gamma. \tag{17}$$

For sufficiently large λ_{1j_0} and sufficiently small λ_{1j} ($j \neq j_0$) the left-hand side of (17) is negative and the right-hand side is positive, which is impossible.

(ii) Let $a_j \geqslant 0$ ($j = 2, 3, \ldots, m$) and $\beta < 0$. For sufficiently small λ_{1j} ($j = 2, 3, \ldots, m$) the left-hand side of (17) is negative, which is impossible.

(iii) Let $a_j \geqslant 0$ ($j = 2, 3, \ldots, m$), let $a_j > 0$ for at least one $j = j_0$, and let $\beta > 0$. Put

$$R(t) = a_2 N_{12}(t) + \ldots + a_m N_{1m}(t) + \beta T_1(t).$$

Since $R(t)$ is a non-decreasing function of t, for sufficiently small $\varepsilon > 0$ we obtain

$$\mathrm{P}(R(t) \neq \gamma, t \geqslant 0) = \mathrm{P}\left(R\left(\frac{\gamma-\varepsilon}{\beta}\right) = \gamma - \varepsilon, R\left(\frac{\gamma-\varepsilon/2}{\beta}\right) \geqslant \gamma + \varepsilon,\right.$$

$$\left. R(t) \neq \gamma, \frac{\gamma-\varepsilon}{\beta} < t < \frac{\gamma-\varepsilon/2}{\beta}\right) \geqslant \mathrm{P}(A) > 0,$$

where $A = \{\omega : x(t) = 1$ for $0 \leqslant t < (\gamma-\varepsilon)/\beta$, $x(t)$ has in the interval $((\gamma-\varepsilon)/\beta, (\gamma-\varepsilon/2)/\beta]$ the only jump from the state 1 to $j_0\}$. Then

$$\mathrm{P}(R(\tau) \neq \gamma) \geqslant \mathrm{P}(R(t) \neq \gamma, t \geqslant 0) > 0,$$

which contradicts the supposition that $\mathrm{P}(\tau < \infty) = 1$.

(iv) Let $a_j \geqslant 0$ and assume that at least two non-zero a_j are unequal, say $a_3 > a_2 > 0$, and $\beta = 0$. Suppose that there exist $n_{12}^{(0)} \geqslant 1, n_{13}^{(0)}, \ldots, n_{1m}^{(0)}$ such that

$$a_2 n_{12}^{(0)} + a_3 n_{13}^{(0)} + \ldots + a_m n_{1m}^{(0)} = \gamma. \tag{18}$$

Then

$$a \overset{\mathrm{df}}{=} a_2(n_{12}^{(0)} - 1) + a_3 n_{13}^{(0)} + \ldots + a_m n_{1m}^{(0)} < \gamma,$$

$$a_2(n_{12}^{(0)} - 1) + a_3(n_{13}^{(0)} + 1) + a_4 n_{14}^{(0)} + \ldots + a_m n_{1m}^{(0)} > \gamma.$$

Let

$$R_0(t) = a_2 N_{12}(t) + \ldots + a_m N_{1m}(t).$$

Then for any $t^{(0)}$ $(0 < t^{(0)} < t^{(1)})$ we get

$$\mathrm{P}\big(R_0(t) \neq \gamma, t \geqslant 0\big) \geqslant \mathrm{P}\big(R_0(t^{(0)}) = a, R_0(t^{(1)}) = a + a_3 n_{13}^{(0)},$$
$$R_0(t) \neq \gamma, t^{(0)} < t < t^{(1)}\big) \geqslant \mathrm{P}(A_0) > 0,$$

where

$$A_0 = \{\omega : N_{12}(t^{(0)}) = n_{12}^{(0)} - 1, N_{1j}(t^{(0)}) = n_{ij}^{(0)} \ (j = 3, 4, \ldots, m),$$
$$N_{12}(t^{(1)}) = n_{12}^{(0)} - 1, N_{13}(t^{(1)}) = n_{13}^{(0)} + 1, N_{1j}(t^{(1)}) = n_{1j}^{(0)}$$
$$(j = 4, 5, \ldots, m)\}.$$

Thus

$$\mathrm{P}\big(R_0(\tau) \neq \gamma\big) \geqslant \mathrm{P}\big(R_0(t) \neq \gamma, t \geqslant 0\big) > 0,$$

which contradicts the condition that $\mathrm{P}(\tau < \infty) = 1$.

If (18) holds only for $n_{12}^{(0)} = 0$, then $n_{12} = 0$ m_{θ_0}-almost everywhere, and $\mathrm{E}(N_{12}) = 0$, which contradicts the equation $\mathrm{E}(N_{12}) = \lambda_{12}\mathrm{E}(T_1) > 0$.

(v) Let $a_{j_2} = a_{j_3} = \ldots = a_{j_k} = a = 0$ for some k $(2 \leqslant k \leqslant m)$, where $j_s < j_{s+1}$ $(s = 2, 3, \ldots, k-1)$, let the remaining a_j be equal to zero, and let $\beta = 0$. Then we obtain

$$a(n_{1j_2} + n_{1j_3} + \ldots + n_{1j_k}) = \gamma$$

m_{θ_0}-almost everywhere, whence

$$n_{1j_2} + n_{1j_3} + \ldots + n_{1j_k} = \frac{\gamma}{a} \overset{\mathrm{df}}{=} l.$$

Obviously, l must be a positive integer.

(vi) Let $a_j = 0$ $(j = 2, 3, \ldots, m)$ and let $\beta \neq 0$. We obtain

$$\beta t_1 = \gamma \quad \text{and} \quad t_1 = \frac{\gamma}{\beta} \overset{\mathrm{df}}{=} a.$$

Obviously, $a > 0$, which completes the proof of the theorem.

Let $\tau(\omega)$ be the time of the first attaining of the line (16). It can be proved that $\tau(\omega)$ is a sequential sampling plan fulfilling all our regularity suppositions. Then we have

$$E_\tau(N_{1\sigma(2)}+N_{1\sigma(3)}+\ldots+N_{1\sigma(k)}) = l.$$

Since $E_\tau(N_{1j}) = \lambda_{1j}E_\tau(T_1)$ $(j = 2, 3, \ldots, m)$, we obtain

$$\sum_{j=2}^{k} \lambda_{1\sigma(j)}E_\tau(T_1) = l. \tag{19}$$

Let $\tau'(\omega)$ be another sequential sampling plan fulfilling our conditions for which equation (16) holds. For given $\tau'(\omega)$ let $N'_{12}, \ldots, N'_{1m}, T'_1$ be random variables defined analogously as $N_{12}, \ldots, N_{1m}, T_1$ for the sampling plan $\tau(\omega)$. Then

$$E_{\tau'}(N'_{1\sigma(l)}+N'_{1\sigma(3)}+\ldots+N'_{1\sigma(k)}) = l,$$

and since $E_{\tau'}(N'_{1j}) = \lambda_{1j}E_{\tau'}(T_1)$, we get

$$\sum_{j=2}^{k} \lambda_{1\sigma(j)}E_{\tau'}(T'_1) = l. \tag{20}$$

From (19) and (20) we obtain $E_{\tau'}(T'_1) = E_\tau(T_1)$.

From the equations

$$E_\tau(N_{1j}) = \lambda_{1j}E_\tau(T_1), \qquad E_{\tau'}(N'_{1j}) = \lambda_{1j}E_{\tau'}(T'_1)$$

and from the result above we obtain

$$E_{\tau'}(N'_{1j}) = E_\tau(N_{1j}) \quad \text{for } j = 2, 3, \ldots, m.$$

Now, let $\tau(\omega)$ be the time of the first attaining of the line $t_1 = a$ for some $a > 0$ with probability 1. Let $\tau'(\omega)$ be the sequential sampling plan for which $t_1 = a$ m_{θ_0}-almost everywhere. It is easy to see that in this case $\tau'(\omega) = \tau(\omega)$ with probability 1.

Definition 3. Let $\tau(\omega)$ be the time of the first attaining of the line (16). The sequential sampling plan $\tau'(\omega)$ is *delimited* by $\tau(\omega)$ if it fulfils our regularity conditions and condition (16) $m^{\tau'}_{\theta_0}$-almost everywhere.

Example. Let $\tau(\omega)$ be the time of the first attaining of the line (16) and let $\tau'(\omega)$ be the time of the first jump of the process $x(t)$ after $\tau(\omega)$. Then it is easily seen that $T'_1 = T_1$, $N'_{1j} = N_{1j}$ $(j = 2, 3, \ldots, m)$ m_{θ_0}-almost everywhere. Then $\tau'(\omega)$ is delimited by $\tau(\omega)$.

Definition 4. If $\tau(\omega)$ is the time of the first attaining of the line (16) for some $k = 2, 3, \ldots, m$ and for some positive integer l, then $\tau(\omega)$ is said to be an *inverse plan*.

If $\tau(\omega)$ is the plan for which $T_1 = a$ holds for some $a = 0$ with probability 1, then $\tau(\omega)$ is called a *simple plan.*

From the Corollary and Theorems 1 and 2 it follows that the only plans which could be efficient when $h = h(\Lambda_1)$ are inverse plans, plans delimited by them, and simple plans.

THEOREM 3. *The following functions $h(\Lambda_1)$ are efficiently estimable:*
(a) *for an inverse plan τ with some k, l and σ or for plans delimited by τ,*

$$h(\Lambda_1) = \frac{a_2\lambda_{12} + \ldots + a_m\lambda_{1m} + \beta}{\lambda_{1\sigma(2)} + \ldots + \lambda_{1\sigma(k)}}; \tag{21}$$

(b) *for a simple plan τ,*

$$h(\Lambda_1) = a_2\lambda_{12} + \ldots + a_m\lambda_{1m} + \beta. \tag{22}$$

Proof. Let us consider an inverse plan τ with some k $(2 \leqslant k \leqslant m)$, l and σ for which

$$f = a_2 N_{1\sigma(2)} + \ldots + a_k N_{1\sigma(m)} + aT_1 + b.$$

For the plan τ we have $n_{1\sigma(2)} + \ldots + n_{1\sigma(k)} = l$ m_{0_0}-almost everywhere and, consequently,

$$\mathrm{E}(N_{1\sigma(2)} + \ldots + N_{1\sigma(k)}) = (\lambda_{1\sigma(2)} + \ldots + \lambda_{1\sigma(k)})\mathrm{E}(T_1) = l,$$

whence

$$\mathrm{E}(T_1) = \frac{l}{\lambda_{1\sigma(2)} + \ldots + \lambda_{1\sigma(k)}},$$

and

$$\mathrm{E}(f) = \frac{l(a_2\lambda_{1\sigma(2)} + \ldots + a_k\lambda_{1\sigma(m)} + a)}{\lambda_{1\sigma(2)} + \ldots + \lambda_{1\sigma(k)}} + b \tag{23}$$

$$\stackrel{\mathrm{df}}{=} \frac{a_2\lambda_{12} + \ldots + a_m\lambda_{1m} + \beta}{\lambda_{1\sigma(2)} + \ldots + \lambda_{1\sigma(k)}}.$$

Since

$$N_{1i} = N_{1i} - \lambda_{1i}T_1 + \lambda_{1i}T_1,$$

we have

$$f = a_2(N_{1\sigma(2)} - \lambda_{1\sigma(2)}T_1) + \ldots + a_m(N_{1\sigma(m)} - \lambda_{1\sigma(m)}T_1) + \tag{24}$$
$$+ (a_2\lambda_{1\sigma(2)} + \ldots + a_m\lambda_{1\sigma(m)} + a)T_1 + b.$$

But

$$(\lambda_{1\sigma(2)}+\ldots+\lambda_{1\sigma(k)})T_1 = N_{1\sigma(2)}+\ldots+N_{1\sigma(k)}-$$

$$-[(N_{1\sigma(2)}-\lambda_{1\sigma(2)}T_1)+\ldots+(N_{1\sigma(k)}-\lambda_{1\sigma(k)}T_1)] = l-\sum_{j=2}^{k}(N_{1\sigma(j)}-\lambda_{1\sigma(j)}T_1)$$

and

$$T_1 = \frac{l}{\lambda_{1\sigma(2)}+\ldots+\lambda_{1\sigma(k)}} - \frac{\sum_{j=2}^{k}(N_{1\sigma(j)}-\lambda_{1\sigma(j)}T_1)}{\lambda_{1\sigma(2)}+\ldots+\lambda_{1\sigma(k)}}. \tag{25}$$

From equations (24) and (25) it follows that

$$f = \sum_{j=2}^{m} a_j(N_{1\sigma(j)}-\lambda_{1\sigma(j)}T_1) - \frac{a_2\lambda_{1\sigma(2)}+\ldots+a_m\lambda_{1\sigma(m)}+a}{\lambda_{1\sigma(1)}+\ldots+\lambda_{1\sigma(m)}}\times$$

$$\times\sum_{j=1}^{k}(N_{1\sigma(j)}-\lambda_{1\sigma(j)}T_1) + \frac{l(a_2\lambda_{1\sigma(2)}+\ldots+a_m\lambda_{1\sigma(m)}+a)}{\lambda_{1\sigma(1)}+\ldots+\lambda_{1\sigma(k)}}+b$$

$$= \sum_{j=2}^{m}\frac{h'_{1j}(\Lambda_1)}{\mathrm{E}(T_1)}(N_{1j}-\lambda_{1j}T_j)+h(\Lambda_1).$$

Then from (27) below we infer that the function $h(\Lambda_1)$ defined by (21) is efficiently estimable.

For the simple plan $t_1 = a$ let $f = a_2N_{12}+\ldots+a_mN_{1m}+b$ be an unbiased estimator of the function (22). We obtain

$$\mathrm{E}(f) = (a_2\lambda_{12}+\ldots+a_m\lambda_{1m})\mathrm{E}(T_1)+b \tag{26}$$

$$= (a_2\lambda_{12}+\ldots+a_m\lambda_{1m})a+b \overset{\mathrm{df}}{=} a_2\lambda_{12}+\ldots+a_m\lambda_{1m}+\beta,$$

$$f = a_2N_{12}+\ldots+a_mN_{1m}+b$$

$$= a_2(N_{12}-\lambda_{12}T_1)+\ldots+a_m(N_{1m}-\lambda_{1m}T_1)+(a_2\lambda_{12}+\ldots+a_m\lambda_{1m})a+\beta$$

$$= \sum_{j=2}^{m}\frac{h'_{1j}(\Lambda_1)}{\mathrm{E}(T_1)}(N_{1j}-\lambda_{1j}T_j)+h(\Lambda_1).$$

Then from (27) below it follows that for a simple plan the function (22) is efficiently estimable.

4. Let $h(\Lambda)$ depend only on the matrix Λ. From (7) it follows that the estimator f is efficient for $\mathrm{E}(f) = h(\Lambda)$ if and only if

$$f = \sum_{i,j=1}^{m}\frac{h'_{ij}(\Lambda)}{\mathrm{E}(T_i)}(N_{ij}-\lambda_{ij}T_i)+h(\Lambda). \tag{27}$$

Now, let $h(\Lambda)$ depend only on two rows of the matrix Λ, say, $h(\Lambda) = h(\Lambda_1, \Lambda_2)$, where Λ_i denotes the i-th row. Then f is an efficient estimator of $\mathrm{E}(f) = h(\Lambda_1, \Lambda_2)$ if and only if

$$(28) \qquad f = \sum_{j=2}^{m} \frac{h'_{1j}(\Lambda_1, \Lambda_2)}{\mathrm{E}(T_1)} (N_{1j} - \lambda_{1j} T_1) + \sum_{\substack{j=1 \\ j\neq 3}}^{m} \frac{h'_{2j}(\Lambda_1, \Lambda_2)}{\mathrm{E}(T_2)} (N_{2j} - \lambda_{2j} T_2) + h(\Lambda_1, \Lambda_2).$$

Assume that $(\Lambda_1^{(1)}, \Lambda_2^{(1)})$ and $(\Lambda_1^{(2)}, \Lambda_2^{(2)})$ are values of (Λ_1, Λ_2) such that $h(\Lambda_1^{(1)}, \Lambda_2^{(1)}) \neq h(\Lambda_1^{(2)}, \Lambda_2^{(2)})$. Then, similarly as in Theorem 1, we can prove that there exist constants $a_{1j}, a_{2j'}$ $(j = 2, 3, \ldots, m$ and $j' = 1, 3, \ldots, m)$ and β_1, β_2 not all equal to zero such that

$$(29) \qquad \sum_{j=2}^{m} a_{1j} N_{ij} + \beta_1 T_1 + \sum_{\substack{j=1 \\ j\neq 2}}^{m} a_{2j} N_{2j} + \beta_2 T_2 = 1$$

m_{0_0}-almost everywhere.

Let us suppose that there exists no hyperplane

$$(30) \qquad \sum_{j=2}^{m} a'_{1j} N_{1j} + \beta'_1 T_1 + \sum_{\substack{j=1 \\ j\neq 2}}^{m} a'_{2j} N_{2j} + \beta'_2 T_2 = 1,$$

different from (29), such that (30) holds simultaneously m_{0_0}-almost everywhere.

Moreover, assume that all second derivatives $\partial^2 h / \partial\lambda_{ij} \partial\lambda_{i'j'}$ $(i, i' = 1, 2)$ exist and are continuous for all $a_{1j} > 0$, $a_{2j'} > 0$, and $\mathrm{E}(T_1)$, $\mathrm{E}(T_2)$ have continuous derivatives with respect to all parameters λ_{ij} in the whole considered domain.

Differentiating (28) with respect to λ_{1k} we obtain

$$(31) \qquad \sum_{j=2}^{m} \left(\frac{h'_{1j}}{\mathrm{E}(T_1)}\right)'_{\lambda_{1k}} N_{1j} - \sum_{j=2}^{m} \left(\frac{\lambda_{1j} h'_{1j}}{\mathrm{E}(T_1)}\right)'_{\lambda_{1k}} T_1 + \sum_{\substack{j=1 \\ j\neq 3}}^{m} \left(\frac{h'_{2j}}{\mathrm{E}(T_2)}\right)'_{\lambda_{1k}} N_{2j} - \sum_{\substack{j=1 \\ j\neq 3}}^{m} \left(\frac{\lambda_{2j} h'_{2j}}{\mathrm{E}(T_2)}\right)'_{\lambda_{1k}} T_2 = -h'_{1k} \quad (k = 2, 3, \ldots, m).$$

Similarly,

$$(32)\qquad \sum_{j=2}^{m}\left(\frac{h'_{1j}}{\mathrm{E}(T_1)}\right)'_{\lambda_{2k'}} N_{1j}-\sum_{j=2}^{m}\left(\frac{\lambda_{1j}h'_{1j}}{\mathrm{E}(T_1)}\right)'_{\lambda_{2k'}} T_1+$$

$$+\sum_{\substack{j=1\\ j\neq 3}}^{m}\left(\frac{h'_{2j}}{\mathrm{E}(T_2)}\right)'_{\lambda_{2k'}} N_{2j}-\sum_{\substack{j=1\\ j\neq 3}}^{m}\left(\frac{\lambda_{2j}h'_{2j}}{\mathrm{E}(T_2)}\right)'_{\lambda_{2k'}} T_2 = -h'_{2k'}\quad (k' = 1, 3, \ldots, m).$$

Then the following equalities must hold:

$$(33)\qquad \left(\frac{h'_{1j}}{\mathrm{E}(T_1)}\right)'_{\lambda_{1k}} = -a_{1j}h'_{1k}\quad (j, k = 2, 3, \ldots, m),$$

$$(34)\qquad \left(\frac{h'_{1j}}{\mathrm{E}(T_1)}\right)'_{\lambda_{2k'}} = -a_{1j}h'_{2k'}\quad (j = 2, 3, \ldots, m,\ k = 1, 3, \ldots, m).$$

Comparing equalities (33) and (34) for fixed j and $k = 2, 3, \ldots, m$, $k' = 1, 3, \ldots, m$, we obtain

$$(35)\qquad \frac{h'_{1j}}{\mathrm{E}(T_1)} = -a_{1j}h+c_{1j}\quad (j = 2, 3, \ldots, m),$$

where c_{1j} does not depend on Λ_1 and Λ_2 (it may depend on $\Lambda_3, \ldots, \Lambda_m, p$, where p is the vector of initial probabilities).

Similarly, comparing expressions at T_1, N_{2j} $(j = 1, 3, \ldots, m)$ and T_2 in (31), (32) and (29) we obtain

$$(36)\qquad \sum_{j=2}^{m}\frac{\lambda_{1j}h'_{1j}}{\mathrm{E}(T_1)} = \beta_1 h+d_1,$$

$$\frac{h'_{2j}}{\mathrm{E}(T_2)} = -a_{2j}h+c_{2j}\quad (j = 1, 3, \ldots, m),$$

$$\sum_{\substack{j=1\\ j\neq 3}}^{m}\frac{\lambda_{2j}h'_{2j}}{\mathrm{E}(T_2)} = \beta_2 h+d_2,$$

where d_1, c_{2j} and d_2 do not depend on Λ_1 and Λ_2.

From (35) and (36) we get

$$-\sum_{j=2}^{m}a_{1j}\lambda_{1j}h+\sum_{j=2}^{m}c_{1j}\lambda_{1j} = \beta_1 h+d_1.$$

Suppose that not all $a_{12}, a_{13}, \ldots, a_{1m}, \beta_1$ are equal to zero. Then

$$h = \frac{\sum_{j=2}^{m}c_{1j}\lambda_{1j}-d_1}{\sum_{j=2}^{m}a_{1j}\lambda_{1j}+\beta_1}.$$

Suppose that not all $a_{21}, a_{23}, \ldots, a_{2m}, \beta_2$ are equal to zero. Then we have simultaneously

$$h = \frac{\sum_{\substack{j=1\\ j\neq 2}}^{m} c_{2j}\lambda_{2j} - d_2}{\sum_{\substack{j=1\\ j\neq 2}}^{m} a_{2j}\lambda_{2j} + \beta_2}.$$

But this is impossible. Thus h is a function only of Λ_1 or Λ_2.

The generalization for h depending on more than two rows of the matrix Λ is obvious.

It should be noticed that for $m = 2$ the supposition that for a given efficient sampling plan τ there exists only one hyperplane of the form (29) is not always true.

Let τ be the time of the first entry to $n_{21} = l$ when $x(0) = 1$ and let τ be also the time of the first entry to $n_{12} = l$ when $x(0) = 2$. In other words, τ is the time of the first attaining of the line $n_{12}+n_{21} = 2l$. This sampling plan is said to be a *mixed inverse plan*. For this plan we have $N_{12} = N_{21} = l$ m_{0_0}-almost everywhere.

Let τ' be the inverse plan defined by $N_{12} = l$ and let τ'' be the inverse plan defined by $N_{21} = l$. Then τ is simultaneously delimited by τ' and by τ''.

We have proved that for the inverse plan $N_{12} = l$ the function $h = a/\lambda_{12}+\beta$ is efficiently estimable. Similarly, for the inverse plan $N_{21} = l$ the function $h = a/\lambda_{21}+\beta$ is efficiently estimable. Let us consider the mixed inverse plan. Let

$$f = a_1T_1 + a_2T_2 + b.$$

But $\mathrm{E}(N_{12}) = \lambda_{12}\mathrm{E}(T_1) = l$ and $\mathrm{E}(N_{21}) = \lambda_{21}\mathrm{E}(T_2) = l$, thus

$$\mathrm{E}(f) = l\left(\frac{a_1}{\lambda_{12}} + \frac{a_2}{\lambda_{21}}\right) + b.$$

On the other hand, since $N_{12} = N_{21} = l$ m_{0_0}-almost everywhere, we have

$$f = -\frac{a_1}{\lambda_{12}}(N_{12}-\lambda_{12}T_1) - \frac{a_2}{\lambda_{21}}(N_{21}-\lambda_{21}T_2) + l\left(\frac{a_1}{\lambda_{12}} + \frac{a_2}{\lambda_{21}}\right) + b.$$

Now also the function

$$h = l\left(\frac{a_1}{\lambda_{12}} + \frac{a_2}{\lambda_{21}}\right) + b$$

is efficiently estimable.

5. Sometimes we have the information that the value of Λ is near a given value $\Lambda^{(0)}$. In this case we can use an estimator efficient at $\Lambda^{(0)}$ hoping that for Λ near $\Lambda^{(0)}$ the variance of the estimator will be near the right-hand side of inequality (12).

Let f be an estimator efficient at $\Lambda^{(0)}$. We know that

$$f = \sum_{\substack{i,j=1\\ i\neq j}}^{m} b_{ij}(N_{ij}-\lambda_{ij}^{(0)}T_1)+c,$$

and from (27) it follows that

$$c = h(\Lambda^{(0)}) \quad \text{and} \quad b_{ij} = \frac{h'_{ij}(\Lambda^{(0)})}{\mathrm{E}^{(0)}(T_i)},$$

where $\mathrm{E}^{(0)}(T_i)$ is the value of $\mathrm{E}(T_i)$ for $\Lambda = \Lambda^{(0)}$. Since

$$h(\Lambda) = \mathrm{E}(f) = \sum_{\substack{i,j=1\\ i\neq j}}^{m} b_{ij}(\lambda_{ij}-\lambda_{ij}^{(0)})\mathrm{E}(T_i)+c,$$

we obtain

$$h(\Lambda) = \sum_{\substack{i,j=1\\ i\neq j}}^{m} h'_{ij}(\Lambda^{(0)})(\lambda_{ij}-\lambda_{ij}^{(0)})\frac{\mathrm{E}(T_i)}{\mathrm{E}^{(0)}(T_i)}+h(\Lambda^{(0)}). \tag{37}$$

Then, if we want to estimate efficiently a given function $h(\Lambda)$ at $\Lambda^{(0)}$, we should find a plan with $\mathrm{E}(T_i)$ $(i = 1, 2, \ldots, m)$ such that equation (37) holds.

Furthermore, we have

$$\begin{aligned}
\mathrm{D}^2(f) &= \mathrm{D}^2\Big(\sum_{\substack{i,j=1\\ i\neq j}}^{m} b_{ij}(N_{ij}-\lambda_{ij}T_i)+b_{ij}(\lambda_{ij}-\lambda_{ij}^{(0)})T_i\Big)\\
&= \sum_{\substack{i,j=1\\ i\neq j}}^{m} b_{ij}^2\mathrm{E}(N_{ij}-\lambda_{ij}T_i)^2+\mathrm{D}^2\Big(\sum_{\substack{i,j=1\\ i\neq j}}^{m} b_{ij}(\lambda_{ij}-\lambda_{ij}^{(0)})T_i\Big)+\\
&\quad+\sum_{\substack{i,j,k,l=1\\ i\neq j,\ k\neq l}}^{m} b_{ij}b_{kl}(\lambda_{kl}-\lambda_{kl}^{(0)})\mathrm{E}\big((N_{ij}-\lambda_{ij}T_i)T_k\big)\\
&\overset{(5),(8)}{=} \sum_{\substack{i,j=1\\ i\neq j}}^{m} b_{ij}^2\lambda_{ij}\mathrm{E}(T_i)+\mathrm{D}^2\Big(\sum_{i,j=1}^{m} b_{ij}(\lambda_{ij}-\lambda_{ij}^{(0)})T_i\Big)+\\
&\quad+\sum_{\substack{i,j,k,l=1\\ i\neq j,\ k\neq l}}^{m} b_{ij}b_{kl}(\lambda_{kl}-\lambda_{kl}^{(0)})\lambda_{ij}\frac{\partial\mathrm{E}(T_k)}{\partial_i}.
\end{aligned}$$

Let $\tau_1(\omega)$ and $\tau_2(\omega)$ be two plans efficient at $\Lambda^{(0)}$ with the same $\mathrm{E}(T_i)$ $(i = 1, 2, \ldots, m)$. Since

$$b_{ij} = \frac{h'_{ij}(\Lambda^{(0)})}{\mathrm{E}^{(0)}(T_i)},$$

for given Λ this plan is better for which the variance

$$\mathrm{D}^2\Bigl(\sum_{\substack{i,j=1\\ i\neq j}}^{m} \frac{h'_{ij}(\Lambda^{(0)})}{\mathrm{E}^{(0)}(T_i)}(\lambda_{ij}-\lambda_{ij}^{(0)})T_i\Bigr)$$

is smaller. Particularly, when $h(\Lambda) = h(\Lambda_i)$, for given Λ_i this plan is better for which $\mathrm{D}^2(T_i)$ is smaller.

References

[1] J.-R. Barra (Ж.-Р. Барра), *Основные понятия математической статистики*, Издательство „Мир", Москва 1974 (translation from French).

[2] B. R. Bhat and N. V. Kulkarni, *On efficient multinomial estimation*, J. Roy. Statist. Soc. Ser. B 28 (1966), p. 45-52.

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

[4] I. I. Gihman and A. V. Skorohod (И. И. Гихман и А. В. Скороход), *О плотностях вероятностных мер в функциональных пространствах*, Успехи мат. наук 21 (1966), p. 83-152.

[5] M. A. Girshick, F. Mosteller and L. J. Savage, *Unbiased estimates for certain binomial sampling with applications*, Ann. Math. Statist. 17 (1946), p. 13-23.

[6] M. H. De Groot, *Unbiased sequential estimation for binomial populations*, ibidem 30 (1959), p. 80-101.

[7] R. Magiera, *On the inequality of Cramér-Rao type in sequential estimation theory*, Zastos. Mat. 14 (1974), p. 227-235.

[8] R. Różański, *A modification of Sudakov's lemma and efficient sequential plans for the Ornstein-Uhlenbeck process*, ibidem 17 (1980), p. 73-86.

[9] D. D. Skrjabin (Д. Д. Скрябин), *Об эффективном последовательном оценивании*, Вестник Ленинград. унив., Мат. мех. астроном., 7 (1973), p. 50-56.

[10] V. N. Sudakov (В. Н. Судаков), *О мерах определяемых марковскими моментами*, Зап. науч. сем. Ленинград. отдел. мат. инст. Стеклов. (ЛОМИ), Т. 12, *Исследования по теории случайных процессов*, Наука 1969, p. 157-164.

[11] S. Trybuła, *Sequential estimation in processes with independent increments*, Dissertationes Math. (Rozprawy Mat.) 60 (1968), p. 1-49.

INSTITUTE OF MATHEMATICS
TECHNICAL UNIVERSITY OF WROCŁAW
50-370 WROCŁAW

Received on 2. 2. 1978;
revised version on 2. 12. 1978

S. TRYBUŁA (Wrocław)

SEKWENCYJNA ESTYMACJA DLA PROCESÓW MARKOWA O SKOŃCZONEJ LICZBIE STANÓW

STRESZCZENIE

W pracy zastosowano nierówność Craméra-Rao dla statystycznych struktur, w sekwencyjnym przypadku, do niektórych procesów stochastycznych o rozkładzie zależnym od skończonej liczby parametrów. Podano przykłady dla niektórych procesów o przyrostach niezależnych i niektórych procesów Markowa. Omówiono problem efektywnej estymacji funkcji macierzy intensywności jednorodnego procesu Markowa o skończonej liczbie stanów i znaleziono wszystkie efektywne plany, dla których efektywnie estymowalna funkcja zależy od jednego wiersza macierzy intensywności. Otrzymano także pewien negatywny rezultat dla efektywnie estymowalnej funkcji, która zależy od dwóch lub więcej wierszy macierzy intensywności.