ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XVIII, 4 (1985), p.577-595

M. WILCZYŃSKI (Wrocław)

# MINIMAX SEQUENTIAL ESTIMATION FOR THE MULTINOMIAL AND GAMMA PROCESSES

**Introduction.** In this paper we consider the problem of minimax sequential estimation of the parameters of the multinomial and gamma processes. The loss incurred by the statistician is due not only to the error of estimation but also to the cost of observation. We prove Theorem 1 and next, we give some examples of its application in which the fixed-time plans and the inverse plans are minimax.

**1. Preliminaries.** In this section we introduce the notation which is used in the sequel and recall some basic definitions. Next, we prove Theorem 1 which is a slight modification of that used by Dvoretzky, Kiefer and Wolfowitz [1], Różański [3] and Trybuła [5].

Suppose $X(t) = (X_1(t), \ldots, X_r(t))$, $t \in T$, $T = [0, \infty)$ or $T \subset \{0, 1, \ldots\}$ is a stochastic process defined on the probability space

$$(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\in T}, P_\theta),$$

where $\Omega$ is the space of $r$-dimensional, vector valued, right continuous functions $\omega = x(\cdot)$: $T \to R^r$ for which left side limits exist, $\mathscr{F}$ is the least $\sigma$-algebra with respect to (w.r.t.) which all $x(t)$ are measurable for $t \in T$, $\mathscr{F}_t$ is the least $\sigma$-algebra w.r.t. which all $x(s)$ are measurable for $s \leqslant t$, $P_\theta$ is the probability measure defined on $\mathscr{F}$ and dependent on the parameter $\theta = (\theta_1, \ldots, \theta_m)$ which takes its values in an open set $\Theta \subset R^m$.

DEFINITION 1. The random variable $\tau$: $\Omega \to T$ is called a *stopping time w.r.t.* $\{\mathscr{F}_t\}_{t\in T}$ if

$$\{\omega\colon\ \tau(\omega) \leqslant t\} \in \mathscr{F}_t \qquad \text{for each } t \in T,$$

$$P_\theta\{\omega\colon\ 0 \leqslant \tau(\omega) < \infty\} = 1, \qquad \text{for each } \theta \in \Theta.$$

Let $P_{\theta,t}$ be the restriction of the measure $P_\theta$ to the $\sigma$-algebra $\mathscr{F}_t$. Let us suppose that for each $\theta\in\Theta$ and each $t\in T$ the measure $P_{\theta,t}$ is absolutely continuous w.r.t. the measure $P_{\theta_0,t}$ and that the density function takes the form

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

where $Z(\omega, t)$: $\Omega\times R^l$ is an $\mathscr{F}_t$ measurable and right continuous mapping w.r.t. $t$, $P_\theta$ almost surely and $g(\cdot, \cdot, \theta, \theta_0)$ is a continuous function. Denote by $\mathscr{B}_U$ the $\sigma$-algebra of Borel subsets of $U = T\times R^l$. On $(U, \mathscr{B}_U)$ we define the measure $m_\theta$ in the following way:

$$m_\theta(B) = P_\theta\left\{\omega\colon \big(\tau(\omega), Z(\omega, \tau(\omega))\big)\in B\right\} \quad \text{for all } B\in\mathscr{B}_U.$$

By the modification of Sudakov's lemma (see [3]) the measure $m_\theta$ is absolutely continuous w.r.t. $m_{\theta_0}$ and

$$(1) \qquad \frac{dm_\theta}{dm_{\theta_0}} = g(t, z, \theta, \theta_0),$$

where $t$ and $z$ are the values of $\tau(\omega)$ and $Z(\omega, \tau(\omega))$, respectively.

Suppose $h$ is a function from $\Theta$ into $R^k$.

DEFINITION 2. The $\mathscr{B}_U$ measurable mapping $d$: $T\times R^l\to R^k$ is called an *estimator* of $h$.

DEFINITION 3. By a *sequential plan* we mean any pair $\delta = (\tau, d)$ consisting of a stopping time $\tau$ and an estimator $d$ of $h$.

Denote by $L(d, \theta)$ the loss function determining the loss incurred by the statistician if $\theta$ is the true value of the parameter and $d$ is the chosen estimator. Let $c(t, z)$, $c(\cdot, \cdot)$: $T\times R^l\to[0, \infty)$, be the cost function which represents the cost to the statistician of observing the sample path up to the time $t$.

Denote by $R(\delta, \theta)$ the risk function, i.e.

$$\begin{aligned} R(\delta, \theta) &= \int_U \big[L\big(d(t, z), \theta\big)+c(t, z)\big]\, m_\theta(dt, dz) \\ &= \mathrm{E}_\theta\big[L\big(d(\tau, Z(\tau)), \theta\big)+c\big(\tau, Z(\tau)\big)\big], \end{aligned}$$

where we write $\tau$ and $Z(\tau)$ for the functions $\omega\to Z\big(\omega, \tau(\omega)\big)$ and $\omega\to\tau(\omega)$, respectively. In the sequel we assume that $R(\delta, \theta) < \infty$ for each $\theta\in\Theta$. We denote the class of all sequential plans satisfying this condition by $\mathscr{D}$.

DEFINITION 4. A sequential plan $\delta_0 = (\tau_0, d_0)$ is said to be *minimax* if

$$\sup_{\theta\in\Theta} R(\delta_0, \theta) = \inf_{\delta\in\mathscr{D}} \sup_{\theta\in\Theta} R(\delta, \theta).$$

Let $\pi$ be the prior distribution of the parameter $\theta$ on the space $(\Theta, \mathscr{B}_\Theta)$

($\mathscr{B}_\Theta$ is the $\sigma$-algebra of Borel subsets of $\Theta$). Then assuming that $R(\delta, \theta)$ is $\mathscr{B}_\Theta$ measurable function of the variable $\theta$ we define the Bayes risk by

$$r(\delta, \pi) = \int_\Theta R(\delta, \theta)\pi(d\theta).$$

DEFINITION 5. A sequential plan $\hat{\delta} = (\hat{\tau}, \hat{d})$ is called *Bayes for* $\pi$ if

$$r(\hat{\delta}, \pi) = \inf_{\delta\in\mathscr{D}} r(\delta, \pi).$$

DEFINITION 6. By the *posterior risk* corresponding to $\pi$ and an estimator $d$ we mean

$$r_1\big(\pi(\cdot|\tau = t, Z(\tau) = z), d\big) = \int_\Theta [L(d(t, z), \theta)]\,\pi(d\theta|\tau = t, Z(\tau) = z),$$

where $\pi(\cdot|\tau = t, Z(\tau) = z)$ denotes the posterior distribution of $\theta$ given $\tau = t$, $Z(\tau) = z$.

DEFINITION 7. An estimator $d_\pi^\bullet$ is called $(t, z)$*-Bayes for* $\pi$ if

$$r_1\big(\pi(\cdot|\tau = t, Z(\tau) = z), d_\pi^\bullet\big) = \inf_d r_1\big(\pi(\cdot|\tau = t, Z(\tau) = z), d\big)$$

for all $(t, z)\in T\times R^l$.

Remark 1. Note that for each sequential plan $\delta = (\tau, d)$ and each $\pi$ we have

$$\begin{aligned} r(\delta, \pi) &= \int_\Theta \Big\{\int_U [L(d(t, z), \theta) + c(t, z)]\, m_\theta(dt, dz)\Big\}\pi(d\theta) \\ &= \int_U \Big\{\int_\Theta [L(d(t, z), \theta) + c(t, z)]\,\pi(d\theta|\tau = t, Z(\tau) = z)\Big\} m(dt, dz) \\ &\geqslant \int_U c(t, z)\, m(dt, dz) + \int_U \Big\{\int_\Theta L(d_\pi^\bullet(t, z), \theta)\pi(d\theta|\tau = t, Z(\tau) = z)\Big\} m(dt, dz) \\ &= \int_U \big\{r_1\big(\pi(\cdot|\tau = t, Z(\tau) = z), d_\pi^\bullet\big) + c(t, z)\big\} m(dt, dz), \end{aligned}$$

where

$$m(B) = \int_\Theta m_\theta(B)\pi(d\theta) \quad \text{for each } B\in\mathscr{B}_U.$$

In the sequel we use the following theorem:

THEOREM 1. *Let $f(t, z)$ be a measurable mapping from $U = T\times R^l$ into $R^1$, and let $c(t, z) = c(f(t, z))$. Suppose that there exists a sequence $\{\pi_n\}_{n=1}^\infty$ of the prior distributions of the parameter $\theta\in\Theta$ such that*

$$\liminf_{n\to\infty} r_1\big(\pi_n(\cdot|\tau = t, Z(\tau) = z), d_{\pi_n}^\bullet\big) = K(f(t, z)) \quad \textit{for each } (t, z)\in U$$

*for some real-valued measurable function $K$ defined on $R^1$. Moreover, suppose that $K(y) + c(y)$ attains its minimum over $A$ at a point $y_0\in A$, where $A$ is the* *range of the function* $v\colon \Omega \times T \to R^1$ *defined by* $v(\omega, t) = f\big(t, Z(t, \omega)\big)$. *If the random variable*

$$\tau_{y_0} = \inf\{t \in T\colon f\big(t, Z(t, \omega)\big) = y_0\}$$

*is a stopping time for each* $\theta \in \Theta$, *and if there exists an estimator* $d_{y_0} = d_{y_0}\big(\tau_{y_0}, Z(\tau_{y_0})\big)$ *such that*

$$\sup_{\theta \in \Theta} R(\delta_{y_0}, \theta) \leqslant K(y_0) + c(y_0),$$

*where* $\delta_{y_0} = (\tau_{y_0}, d_{y_0})$, *then the sequential plan* $\delta_{y_0}$ *is minimax in the class of all sequential plans* $\delta = \big(\tau, d\big(\tau, Z(\tau)\big)\big) \in \mathcal{D}$.

Proof. Let $\delta = (\tau, d)$ be a sequential plan. By Remark 1

$$\sup_{\theta \in \Theta} R(\delta, \theta) \geqslant r(\delta, \pi_n)$$
$$\geqslant \int_U \Big\{r_1\big(\pi_n(\cdot | \tau = t, Z(\tau) = z), d^*_{\pi_n}\big) + c\big(f(t, z)\big)\Big\} m(dt, dz)$$

for each $n \geqslant 1$.

But, by Fatou's lemma and by the assumptions of Theorem 1, we have

$$\sup_{\theta \in \Theta} R(\delta, \theta) \geqslant \liminf_{n \to \infty} r(\delta, \pi_n) \geqslant \int_U \{K\big(f(t, z)\big) + c\big(f(t, z)\big)\} m(dt, dz)$$
$$\geqslant K(y_0) + c(y_0) \geqslant \sup_{\theta \in \Theta} R(\delta_{y_0}, \theta)$$

which completes the proof of Theorem 1.

**2.** In this section we consider for the multinomial and gamma processes a few examples of minimax sequential estimation in the case when the cost function does not depend on the state of the process, i.e. if $c(t, z) = c(t)$ for each $(t, z) \in U$, and the obtained minimax plans are the fixed-time ones.

Suppose $Y(t) = \big(Y_1(t), \ldots, Y_r(t)\big)$, $t \in T = \{1, 2, \ldots\}$, is a sequence of independent random variables having the multinomial distribution with parameter $p \in P = \{p = (p_1, \ldots, p_r)\colon p_i > 0, i = 1, \ldots, r, p_1 + \ldots + p_r = 1\}$, i.e.

$$P_p\{Y(t) = e_i\} = p_i, \quad \text{for each } t \in T \text{ and } i = 1, \ldots, r,$$

where $e_1 = (1, 0, \ldots, 0), \ldots, e_r = (0, \ldots, 0, 1)$. Denote $X_i(t) = Y_i(1) + \ldots + Y_i(t)$. The process $X(t) = \big(X_1(t), \ldots, X_r(t)\big)$ is called the multinomial process with parameter $p$. It is well known that for each $p \in P$

$$P_p\{X(t) = x = (x_1, \ldots, x_r)\} = \begin{cases} \dfrac{t!}{x_1! \cdot \ldots \cdot x_r!} p_1^{x_1} \cdot \ldots \cdot p_r^{x_r}, & \text{if } x_1 + \ldots + x_r = t, \\ & x_i \in T \cup \{0\}, \\ 0, & \text{otherwise,} \end{cases}$$

and

(2) $\quad \mathrm{E}_p X_i(t) = tp_i, \quad \mathrm{D}_p^2 X_i(t) = tp_i(1-p_i), \quad \mathrm{Cov}\{X_i(t), X_j(t)\} = -tp_i p_j,$

$$i, j = 1, \ldots, r, \; i \neq j.$$

Moreover, for each $p, p_0 \in P$ holds

$$\frac{dP_{p,t}}{dP_{p_0,t}} = \left(\frac{p_1}{p_r^0}\right)^{X_1(t)} \cdot \ldots \cdot \left(\frac{p_r}{p_r^0}\right)^{X_r(t)} = g\big(t, X(t), p_0\big) p_1^{X_1(t)} \cdot \ldots \cdot p_r^{X_r(t)}$$

(in the case of the multinomial process we have $Z(t) = X(t)$, $\theta = p$) and thus, by (1), for each stopping time $\tau$ is

$$\frac{dm_p}{dm_{p_0}} = g(t, x, p_0) \cdot p_1^{x_1} \cdot \ldots \cdot p_r^{x_r},$$

where $t$ and $x$ are the values of $\tau$ and $X(\tau)$, respectively.

Example 1. Let

$$L(d, p) = \big(d - \sum_{i=1}^{r} a_i p_i\big)^2 \big[\sum_{i=1}^{r} a_i^2 p_i - \big(\sum_{i=1}^{r} a_i p_i\big)^2\big]^{-1} \quad \text{and} \quad c(t, x) = c(t),$$

where $a_1, \ldots, a_r$ are real numbers for which $\sum_{i=1}^{r} (a_i - a_j)^2 > 0$ (since for each $p \in P$ holds

$$\sum_{i=1}^{r} a_i^2 p_i - \big(\sum_{i=1}^{r} a_i p_i\big)^2 = \tfrac{1}{2} \sum_{i,j=1}^{r} (a_i^2 + a_j^2 - 2a_i a_j) p_i p_j = \sum_{i,j=1}^{r} (a_i - a_j)^2 p_i p_j > 0,$$

the loss function is well defined).

Define for each $\varepsilon > 0$ the prior distribution $\pi_\varepsilon$ of the parameter $p$ given by the density

$$g_\varepsilon(p) = \begin{cases} C_\varepsilon \big[\sum_{i=1}^{r} a_i^2 p_i - \big(\sum_{i=1}^{r} a_i p_i\big)^2\big] (p_1 \cdot \ldots \cdot p_r)^{\varepsilon - 1}, & \text{if} \quad p \in P, \\ 0, & \text{if} \quad p \notin P, \end{cases}$$

where

$$C_\varepsilon^{-1} = \int \ldots \int_{p \in P} \big[\sum_{i=1}^{r} a_i^2 p_i - \big(\sum_{i=1}^{r} a_i p_i\big)^2\big] (p_1 \cdot \ldots \cdot p_r)^{\varepsilon - 1} dp_1 \cdot \ldots \cdot dp_{r-1}.$$

In what follows we make use of the Liouville identity (see [2], p. 331): for each $y_1 > 0, \ldots, y_{r-1} > 0$ and each $\varphi$: $R^1 \to R^1$ we have

(3) $$\int \ldots \int_{p \in P} \varphi(p_1 + \ldots + p_{r-1}) p_1^{y_1 - 1} \cdot \ldots \cdot p_r^{y_{r-1} - 1} dp_1 \cdot \ldots \cdot dp_{r-1}$$

$$= \frac{\Gamma(y_1) \cdot \ldots \cdot \Gamma(y_{r-1})}{\Gamma(y_1 + \ldots + y_{r-1})} \int_0^1 \varphi(u) u^{y_1 + \ldots + y_{r-1} - 1} du,$$

provided $\int_0^1 |\varphi(u)| u^{y_1 + \ldots + y_{r-1} - 1} du$ exists.

In view of (3) we have

$$C_\varepsilon^{-1} = \frac{(\Gamma(\varepsilon))^r}{\Gamma(r\varepsilon+1)}\left\{\frac{r\varepsilon}{r\varepsilon+1}\sum_{i=1}^{r} a_i^2\varepsilon - \frac{(\sum_{i=1}^{r} a_i\varepsilon)^2}{1+r\varepsilon}\right\}.$$

The density of the posterior distribution $\pi_\varepsilon(\cdot|\tau = t, X(\tau) = x)$ of $p$ takes the form

$$g_\varepsilon(p|\tau = t, X(\tau) = x)$$
$$= \begin{cases} D_\varepsilon\cdot[\sum_{i=1}^{r} a_i^2 p_i - (\sum_{i=1}^{r} a_i p_i)^2]\, p_1^{x_1+\varepsilon-1}\cdot\ldots\cdot p_r^{x_r+\varepsilon-1}, & \text{if} \quad p\in P,\\ 0, & \text{if} \quad p\notin P,\end{cases}$$

where

$$D_\varepsilon^{-1} = \frac{\Gamma(x_1+\varepsilon)\cdot\ldots\cdot\Gamma(x_r+\varepsilon)}{\Gamma(t+r\varepsilon+1)}\left\{\frac{t+r\varepsilon}{t+r\varepsilon+1}\sum_{i=1}^{r} a_i^2(x_i+\varepsilon) - \frac{(\sum_{i=1}^{r} a_i(x_i+\varepsilon))^2}{t+r\varepsilon+1}\right\}.$$

The $(t, x)$-Bayes estimator $d^\bullet_{\pi_\varepsilon} = d^\bullet_\varepsilon$ fulfils the equation

$$\int\ldots\int_{p\in P}(d^\bullet_\varepsilon(t, x) - \sum_{i=1}^{r} a_i p_i)\, p_1^{x_1+\varepsilon-1}\cdot\ldots\cdot p_r^{x_r+\varepsilon-1}\, dp_1\cdot\ldots\cdot dp_{r-1} = 0.$$

Using (3) we obtain

$$d^\bullet_\varepsilon(t, x) = \sum_{i=1}^{r} a_i(x_i+\varepsilon)(t+r\varepsilon)^{-1}.$$

The posterior risk for this estimator is

$$r_1\big(\pi_\varepsilon(\cdot|\tau = t, X(\tau) = x), d^\bullet_\varepsilon\big)$$
$$= D_\varepsilon\cdot\int\ldots\int_{p\in P}(d^\bullet_\varepsilon(t, x) - \sum_{i=1}^{r} a_i p_i)^2\, p_1^{x_1+\varepsilon-1}\cdot\ldots\cdot p_r^{x_r+\varepsilon-1}\, dp_1\cdot\ldots\cdot dp_{r-1}$$
$$= D_\varepsilon\cdot\frac{\Gamma(x_1+\varepsilon)\cdot\ldots\cdot\Gamma(x_r+\varepsilon)}{\Gamma(t+r\varepsilon+1)}\left\{\frac{\sum_{i=1}^{r} a_i(x_i+\varepsilon)}{t+r\varepsilon+1} - \frac{[\sum_{i=1}^{r} a_i(x_i+\varepsilon)]^2}{(t+r\varepsilon)(t+r\varepsilon+1)}\right\}$$
$$= \frac{1}{t+r\varepsilon} \xrightarrow[\varepsilon\to 0]{} \frac{1}{t}.$$

Let us consider the fixed-time plan (i.e. the sequential plan for which the stopping time is equal to some constant with probability 1) $\delta_t = (t, d_t)$ where $d_t = (\sum_{i=1}^{r} a_i X_i(t))t^{-1}$. By (2) we have

$$\mathrm{E}_p(d_t - \sum_{i=1}^{r} a_i p_i)^2 = [\sum_{i=1}^{r} a_i^2 p_i - (\sum_{i=1}^{r} a_i p_i)^2]\, t^{-1}.$$

Hence

$$R(\delta_t, p) = t^{-1} + c(t).$$

Taking $f(t, z) = t$ for each $(t, z) \in U$ and $K(t) = t^{-1}$ for each $t \in T$, we deduce by Theorem 1 that if there exists $t_0 \in T$ such that

$$t_0^{-1} + c(t_0) = \min_{t \in T} \{t^{-1} + c(t)\}, \tag{4}$$

then the fixed-time plan $\delta_{t_0} = (t_0, d_{t_0})$ is minimax in the class $\mathscr{D}$.

EXAMPLE 2. Let

$$L(d, p) = \Big[\sum_{i,j=1}^{r} c_{ij}(d_i - p_i)(d_j - p_j)\Big]\Big[\sum_{i=1}^{r} c_{ii} p_i - \sum_{i,j=1}^{r} c_{ij} p_i p_j\Big]^{-1}, \quad c(t, x) = c(t),$$

where the matrix $C = \|c_{ij}\|_{i,j=1}^{r}$ is positive definite (since $\sum_{i=1}^{r} c_{ii} p_i - \sum_{i,j=1}^{r} c_{ij} p_i p_j = \frac{1}{2} \sum_{i,j=1}^{r} (c_{ii} + c_{jj} - 2c_{ij}) p_i p_j > 0$ for each $p \in P$, the loss function is well defined).

Let for each $\varepsilon > 0$ the prior distribution $\pi_\varepsilon$ of $p$ has the density

$$g_\varepsilon(p) = \begin{cases} C_\varepsilon \cdot \big(\sum_{i=1}^{r} c_{ii} p_i - \sum_{i,j=1}^{r} c_{ij} p_i p_j\big)(p_1 \cdot \ldots \cdot p_r)^{\varepsilon - 1}, & \text{if } \quad p \in P, \\ 0, & \text{if } \quad p \notin P, \end{cases}$$

where

$$C_\varepsilon^{-1} = \int \ldots \int_{p \in P} \big(\sum_{i=1}^{r} c_{ii} p_i - \sum_{i,j=1}^{r} c_{ij} p_i p_j\big)(p_1 \cdot \ldots \cdot p_r)^{\varepsilon - 1}\, dp_1 \cdot \ldots \cdot dp_{r-1}.$$

Thus

$$g_\varepsilon(p \mid \tau = t, X(\tau) = x)$$
$$= \begin{cases} D_\varepsilon \cdot \big(\sum_{i=1}^{r} c_{ii} p_i - \sum_{i,j=1}^{r} c_{ij} p_i p_j\big)(p_1 \cdot \ldots \cdot p_r)^{\varepsilon - 1}, & \text{if } \quad p \in P, \\ 0, & \text{if } \quad p \notin P, \end{cases}$$

where by (3)

$$D_\varepsilon^{-1} = \int \ldots \int_{p \in P} \big(\sum_{i=1}^{r} c_{ii} p_i - \sum_{i,j=1}^{r} c_{ij} p_i p_j\big) p_1^{x_1 + \varepsilon - 1} \cdot \ldots \cdot p_r^{x_r + \varepsilon - 1}\, dp_1 \cdot \ldots \cdot dp_{r-1}$$
$$= \frac{\Gamma(x_1 + \varepsilon) \cdot \ldots \cdot \Gamma(x_r + \varepsilon)}{\Gamma(t + r\varepsilon + 2)} \Big\{\sum_{i=1}^{r} c_{ii}(t + r\varepsilon)(x_i + \varepsilon) - \sum_{i,j=1}^{r} c_{ij}(x_i + \varepsilon)(x_j + \varepsilon)\Big\}.$$

Suppose $d_\varepsilon^\bullet = (d_{\varepsilon 1}^\bullet, \ldots, d_{\varepsilon r}^\bullet)$ is a $(t, x)$-Bayes estimator for $\pi_\varepsilon$. Then for each $1 \leqslant i \leqslant r$ is

$$\int_{p\in P}\!\!\ldots\!\int \sum_{i,j=1}^{r} c_{ij}\big(d_{\varepsilon j}^\bullet(t, x)-p_j\big)\, p_1^{x_1+\varepsilon-1}\cdot\ldots\cdot p_r^{x_r+\varepsilon-1}\, dp_1\cdot\ldots\cdot dp_{r-1} = 0.$$

Solving this system of equations we get

$$d_{\varepsilon i}^\bullet(t, x) = (x_i+\varepsilon)(t+r\varepsilon)^{-1} \quad \text{for each } 1 \leqslant i \leqslant r.$$

Further we obtain

$$\int_{p\in P}\!\!\ldots\!\int \big(d_{\varepsilon i}^\bullet(t, x)-p_i\big)\big(d_{\varepsilon j}^\bullet(t, x)-p_j\big)\, p_1^{x_1+\varepsilon-1}\cdot\ldots\cdot p_r^{x_r+\varepsilon-1}\, dp_1\cdot\ldots\cdot dp_{r-1}$$

$$= \frac{\Gamma(x_1+\varepsilon)\cdot\ldots\cdot\Gamma(x_r+\varepsilon)}{(t+r\varepsilon)\,\Gamma(t+r\varepsilon+2)} \times \begin{cases} (x_i+\varepsilon)(t+r\varepsilon)-(x_i+\varepsilon)^2, & \text{if } i=j,\\ -(x_i+\varepsilon)(x_j+\varepsilon), & \text{if } i\neq j. \end{cases}$$

Thus

$$\begin{aligned} & r_1\big(\pi_\varepsilon(\cdot|\tau = t, X(\tau) = x), d_\varepsilon^\bullet\big) \\ &= D_\varepsilon\cdot \int_{p\in P}\!\!\ldots\!\int \sum_{i,j=1}^{r} c_{ij}\big(d_{\varepsilon i}^\bullet(t, x)-p_i\big)\big(d_{\varepsilon j}^\bullet(t, x)-p_j\big)\times \\ &\qquad\qquad \times p_1^{x_1+\varepsilon-1}\cdot\ldots\cdot p_r^{x_r+\varepsilon-1}\, dp_1\cdot\ldots\cdot dp_{r-1} \\ &= \frac{\Gamma(x_1+\varepsilon)\cdot\ldots\cdot\Gamma(x_r+\varepsilon)}{\Gamma(t+r\varepsilon+2)(t+r\varepsilon)}\cdot D_\varepsilon\times \\ &\qquad\qquad \times\Big\{\sum_{i=1}^{r} c_{ii}(x_i+\varepsilon)(t+r\varepsilon)-\sum_{i,j=1}^{r} c_{ij}(x_i+\varepsilon)(x_j+\varepsilon)\Big\} \\ &= \frac{1}{t+r\varepsilon} \xrightarrow[\varepsilon\to 0]{} \frac{1}{t}. \end{aligned}$$

On the other hand for the fixed-time plan $\delta_t = (t, d_t)$, $d_t = (d_{1t}, \ldots, d_{rt})$, with $d_{it} = X_i(t)/t$, $i = 1, \ldots, r$, we have

$$\mathrm{E}_p\{(d_{it}-p_i)(d_{jt}-p_j)\} = \begin{cases} \dfrac{p_i(1-p_i)}{t}, & i=j,\\[2mm] -\dfrac{p_i p_j}{t}, & i\neq j. \end{cases}$$

Thus

$$\begin{aligned} R(\delta_t, p) &= \mathrm{E}_p\Big[\sum_{i,j=1}^{r} c_{ij}(d_{it}-p_i)(d_{jt}-p_j)\Big]\Big[\sum_{i=1}^{r} c_{ii}p_i-\sum_{i,j=1}^{r} c_{ij}p_ip_j\Big]^{-1}+c(t) \\ &= 1/t+c(t). \end{aligned}$$

Therefore, all the assumptions of Theorem 1 are fulfilled and we conclude that if there exists $t_0 \in T$ for which

$$t_0^{-1} + c(t_0) = \min_{t \in T} \{t^{-1} + c(t)\},$$

then the fixed-time plan $\delta_{t_0} = (t_0, d_{t_0})$ is minimax in the class $\mathscr{D}$.

EXAMPLE 3. The problem of minimax sequential estimation of the parameter $p$ of the multinomial process w.r.t. the loss function

$$L(d, p) = \sum_{i=1}^{r} (d_i - p_i)^2/p_i$$

and the cost function $c(t, x) = c(t)$ was considered in [5].

The author limited himself to finding a minimax plan in the class $\mathscr{D}_0 \subset \mathscr{D}$ of all sequential plans $\delta = (\tau, d)$, $d = (d_1, \ldots, d_r)$, for which $d_1 + \ldots + d_r = 1$. The minimax one was the fixed-time plan

$$\delta_{t_0} = (t_0, d_{t_0}), \quad d_{it_0} = \frac{X_i(t_0)}{t_0}, \quad i = 1, \ldots, r,$$

and $t_0$ fulfilling the following equality:

$$c(t_0) + \frac{r-1}{t_0} = \min_{t \in T} \left\{ c(t) + \frac{r-1}{t} \right\}.$$

Now we solve the same problem rejecting the condition $d_1 + \ldots + d_r = 1$. Let us define for each $\varepsilon > 0$ the density of the prior distribution $\pi_\varepsilon$ of $p$, as follows:

$$g_\varepsilon(p) = \begin{cases} \dfrac{\Gamma(r\varepsilon)}{[\Gamma(\varepsilon)]^r} (p_1 \cdot \ldots \cdot p_r)^{\varepsilon - 1}, & \text{if } p \in P, \\ 0, & \text{if } p \notin P. \end{cases} \tag{5}$$

Then

$$g_\varepsilon(p \mid \tau = t, X(\tau) = x) = \begin{cases} \dfrac{\Gamma(t + r\varepsilon)}{\Gamma(x_1 + \varepsilon) \cdot \ldots \cdot \Gamma(x_r + \varepsilon)} p_1^{x_1 + \varepsilon - 1} \cdot \ldots \cdot p_r^{x_r + \varepsilon - 1}, & \text{if } p \in P, \\ 0, & \text{if } p \notin P, \end{cases}$$

and the $(t, x)$-Bayes estimators $d_{\varepsilon i}^{\bullet}$, $i = 1, \ldots, r$, fulfil the equations

$$\int_{p \in P} \cdots \int (d_{\varepsilon i}^{\bullet}(t, x) - p_i) p_i^{-1} p_1^{x_1 + \varepsilon - 1} \cdot \ldots \cdot p_r^{x_r + \varepsilon - 1} \, dp_1 \cdot \ldots \cdot dp_{r-1} = 0, \quad i = 1, \ldots, r.$$

Remark 2. By (3) we have

$$\int_{p\in P}\!\!\cdots\!\int p_i^{-1}\cdot p_1^{x_1+\varepsilon-1}\cdot\ldots\cdot p_r^{x_r+\varepsilon-1}\,dp_1\cdot\ldots\cdot dp_{r-1}$$

$$=\frac{\Gamma(x_1+\varepsilon)\cdot\ldots\cdot\Gamma(x_r+\varepsilon)}{\Gamma(x_i+\varepsilon)\,\Gamma\big(t-x_i+(r-1)\varepsilon\big)}\int_0^1\frac{(1-u)^{x_i+\varepsilon-1}}{1-u}u^{t-x_i+(r-1)\varepsilon-1}\,du,\qquad i=1,\ldots,r,$$

provided $\int_0^1(1-u)^{x_i+\varepsilon-2}u^{t-x_i+(r-1)\varepsilon-1}\,du$ exists and is finite. The last assumption is fulfilled if $x_i+\varepsilon-2>-1$ and $t-x_i+(r-1)\varepsilon-1>-1$, so we have to put $\varepsilon>1$.

Now for each $\varepsilon>1$ we obtain

$$d_{\varepsilon i}^{\bullet}(t,x)=\frac{x_i+\varepsilon-1}{t+r\varepsilon-1}$$

and

$$r_1\big(\pi_\varepsilon(\cdot|\tau=t,\,X(\tau)=x),\,d_\varepsilon^{\bullet}\big)$$
$$=\int_{p\in P}\!\!\cdots\!\int\sum_{i=1}^{r}\big(d_{\varepsilon i}^{\bullet}(t,x)-p_i\big)^2 p_i\cdot g_\varepsilon\big(p|\tau=t,\,X(\tau)=x\big)\,dp_1\cdot\ldots\cdot dp_{r-1}$$
$$=\sum_{i=1}^{r}\left(\frac{x_i+\varepsilon}{t+r\varepsilon}-\frac{x_i+\varepsilon-1}{t+r\varepsilon-1}\right)=\frac{r-1}{t+r\varepsilon-1}\xrightarrow[\varepsilon\to 1]{}\frac{r-1}{t+r-1}.$$

Let $\delta_t=(t,d_t)$, $d_t=(d_{1t},\ldots,d_{rt})$, $d_{it}=\dfrac{X_i(t)}{t+r-1}$, $i=1,\ldots,r$, be the fixed-time plan. In view of (2) we have

$$R(\delta_t,p)=\mathrm{E}_p\left\{\sum_{i=1}^{r}\frac{(d_{it}-p_i)^2}{p_i}+c(t)\right\}$$
$$=\sum_{i=1}^{r}\frac{tp_i(1-p_i)+(r-1)^2p_i}{p_i(t+r-1)^2}+c(t)=\frac{r-1}{t+r-1}+c(t)$$

and it is easy to deduce that if there exists $t_0\in T$ for which

$$\frac{r-1}{t_0+r-1}+c(t_0)=\min_{t\in T}\left\{\frac{r-1}{t+r-1}+c(t)\right\}$$

then the fixed-time plan $\delta_{t_0}=(t_0,d_{t_0})$ is minimax in $\mathcal{D}$.

EXAMPLE 4. Suppose that

$$L(d, p) = \sum_{i=1}^{r} \frac{(d_i - p_i)^2}{1 - p_i} \quad \text{and} \quad c(t, x) = c(t).$$

First we find a minimax plan in the class $\mathscr{D}_0$ defined in Example 3. The $(t, x)$-Bayes estimator $d_\varepsilon^* = (d_{\varepsilon 1}^*, \ldots, d_{\varepsilon r}^*)$ w.r.t. the prior distribution $\pi_\varepsilon$ of $p$ defined by (5) is a solution of the equations

$$\int \ldots \int_{p \in P} \left[ \frac{d_{\varepsilon i}^*(t, x)}{1 - p_i} - \frac{d_{\varepsilon r}^*(t, x)}{1 - p_r} \right] g_\varepsilon(p | \tau = t, X(\tau) = x) dp_1 \cdot \ldots \cdot dp_{r-1} = 0,$$

$$i = 1, \ldots, r.$$

Remark 3. Note that

$$\int \ldots \int_{p \in P} (1 - p_i)^{-1} p_1^{x_1 + \varepsilon - 1} \cdot \ldots \cdot p_r^{x_r + \varepsilon - 1} dp_1 \cdot \ldots \cdot dp_{r-1}$$

$$= \frac{\Gamma(x_1 + \varepsilon) \cdot \ldots \cdot \Gamma(x_r + \varepsilon)}{\Gamma(x_i + \varepsilon) \Gamma(t - x_i + (r-1)\varepsilon)} \int_0^1 (1 - u)^{x_i + \varepsilon - 1} u^{-1} u^{t - x_i + (r-1)\varepsilon - 1} du,$$

$$i = 1, \ldots, r,$$

provided $x_i + \varepsilon - 1 > -1$ and $t - x_i + (r-1)\varepsilon - 2 > -1$, so we have to put $\varepsilon > 1/(r-1)$. Since $d_{\varepsilon 1} + \ldots + d_{\varepsilon r} = 1$ we obtain for each $\varepsilon > 1/(r-1)$ that

$$d_{\varepsilon i}^*(t, x) - 1 = \frac{(1 - r)[t - x_i + (r-1)\varepsilon - 1]}{(r-1)t + r(r-1)\varepsilon - r}, \quad i = 1, \ldots, r,$$

and

$$r_1\big(\pi_\varepsilon(\cdot | \tau = t, X(\tau) = x), d_\varepsilon^*\big)$$

$$= \sum_{i=1}^{r} \int \ldots \int_{p \in P} \left[ \frac{(d_{\varepsilon i}^*(t, x) - 1)^2}{1 - p_i} + 2(d_{\varepsilon i}^*(t, x) - 1) + (1 - p_i) \right] \times$$

$$\times g_\varepsilon(p | \tau = t, X(\tau) = x) dp_1 \cdot \ldots \cdot dp_{r-1}$$

$$= \sum_{i=1}^{r} \left[ (d_{\varepsilon i}^*(t, x) - 1)^2 \frac{t + r\varepsilon - 1}{t - x_i + (r-1)\varepsilon - 1} + 2(d_{\varepsilon i}^*(t, x) - 1) + \frac{t - x_i + (r-1)\varepsilon}{t + r\varepsilon} \right]$$

$$= \sum_{i=1}^{r} (d_{\varepsilon i}^*(t, x) - 1) \frac{(1 - r)(t + r\varepsilon - 1)}{(r-1)t + r(r-1)\varepsilon - r} + 2(1 - r) + (r - 1)$$

$$= \frac{r - 1}{(r-1)t + r(r-1)\varepsilon - r} \xrightarrow[\varepsilon \to (r-1)^{-1}]{} \frac{1}{t}.$$

From the other side for the fixed-time plan $\delta_t = (t, d_t)$, $d_t = (d_{1t}, \ldots, d_{rt})$, with $d_{it} = X_i(t)/t$, $i = 1, \ldots, r$, we obtain

$$R(\delta_t, p) = \mathrm{E}_p\left\{\sum_{i=1}^{r} \frac{(d_{it}-p_i)^2}{1-p_i} + c(t)\right\} = \sum_{i=1}^{r} \frac{p_i(1-p_i)}{t(1-p_i)} + c(t) = \frac{1}{t} + c(t)$$

and if there exists $t_0 \in T$ such that

$$\frac{1}{t_0} + c(t_0) = \min_{t \in T}\left\{\frac{1}{t} + c(t)\right\}$$

then $\delta_{t_0} = (t_0, d_{t_0})$ is minimax in the class $\mathcal{D}_0$.

Now we find a solution of our problem in the class $\mathcal{D}$.

For the prior distribution $\pi_\varepsilon$ of $p$ given by the density (5) we obtain

$$\int_{p \in P} \cdots \int \frac{d^{\bullet}_{\varepsilon i}(t, x) - p_i}{1-p_i} g_\varepsilon(p \mid \tau = t, X(\tau) = x)\, dp_1 \cdot \ldots \cdot dp_{r-1} = 0, \qquad i = 1, \ldots, r,$$

$$d^{\bullet}_{\varepsilon i}(t, x) = \frac{x_i + \varepsilon}{t + r\varepsilon - 1}, \qquad i = 1, \ldots, r,$$

$$\begin{aligned} &r_1\big(\pi_\varepsilon(\cdot \mid \tau = t, X(\tau) = x), d^{\bullet}_\varepsilon\big) \\ &= \int_{p \in P} \cdots \int \sum_{i=1}^{r} \frac{(d^{\bullet}_{\varepsilon i}(t, x) - p_i)^2}{1-p_i} g_\varepsilon(p \mid \tau = t, X(\tau) = x)\, dp_1 \cdot \ldots \cdot dp_{r-1} \\ &= \sum_{i=1}^{r} \left[\frac{-(x_i+\varepsilon)}{t + r\varepsilon - x_i - \varepsilon - 1} \cdot \frac{x_i + \varepsilon}{t + r\varepsilon - 1} + \frac{(x_i+\varepsilon)(x_i+\varepsilon+1)}{(t + r\varepsilon - x_i - \varepsilon - 1)(t + r\varepsilon)}\right] \\ &= \sum_{i=1}^{r} \frac{x_i + \varepsilon}{(t + r\varepsilon - 1)(t + r\varepsilon)} = \frac{1}{t + r\varepsilon - 1} \xrightarrow[\varepsilon \to (r-1)^{-1}]{} \frac{1}{t + (r-1)^{-1}}. \end{aligned}$$

For the fixed-time plan $\delta_t = (t, d_t)$, $d_t = (d_{1t}, \ldots, d_{rt})$, with

$$d_{it} = \frac{X_i(t) + (r-1)^{-1}}{t + (r-1)^{-1}}, \qquad i = 1, \ldots, r,$$

we have

$$R(\delta_t, p) = \mathrm{E}_p\left\{\sum_{i=1}^{r} \frac{(d_{it}-p_i)^2}{1-p_i} + c(t)\right\} = \frac{1}{t + (r-1)^{-1}} + c(t).$$

Thus, if there exists $t_0 \in T$ for which

$$\frac{1}{t_0 + (r-1)^{-1}} + c(t_0) = \min_{t \in T}\left\{\frac{1}{t + (r-1)^{-1}} + c(t)\right\},$$

then $\delta_{t_0} = (t_0, d_{t_0})$ is minimax in the class $\mathcal{D}$.

Remark 4. Let us note that in Examples 3 and 4 the plans which are minimax in the class $\mathscr{D}$ are better (i.e. have smaller risk functions) than the plans which are minimax in the class $\mathscr{D}_0$, although $p_1 + \ldots + p_r = 1$.

Let us now consider a problem of minimax sequential estimation for the gamma process.

EXAMPLE 5. Let $X(t)$, $t \geqslant 0$ be the gamma process, i.e. the homogeneous stochastic process with independent increments for which for each $\theta \in \Theta = (0, \infty)$ holds

$$P_\theta\{X(0) = 0\} = 1, \quad \mathrm{E}_\theta X^2(t) < \infty, \quad t \geqslant 0,$$

$$\frac{dP_{\theta,t}}{d\lambda} = \begin{cases} \dfrac{x^{t-1}}{\Gamma(t)} \cdot \exp\{-t \ln \theta - x \cdot \theta^{-1}\}, & \text{if } x > 0, t > 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $\lambda$ is the Lebesgue measure.

It is easy to deduce that for each $\theta, \theta_0 \in \Theta$ holds

$$\frac{dP_{\theta,t}}{dP_{\theta_0,t}} = \begin{cases} \exp\{-t(\ln \theta - \ln \theta_0) - x(\theta^{-1} - \theta_0^{-1})\}, & \text{if } x > 0, t > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Then, by (1)

$$\frac{dm_{\theta,t}}{dm_{\theta_0,t}} = \begin{cases} g(t, x, \theta_0) \cdot \exp\{-t \ln \theta - x\theta^{-1}\}, & \text{if } x > 0, t > 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $t$ and $x$ are the values of $\tau$ and $X(\tau)$, $g(t, x, \theta_0) = \exp\{t \ln \theta_0 + x\theta_0^{-1}\}$.

Let

$$(6) \qquad L(d, \theta) = (d - \theta^\alpha)^2 \cdot \theta^{-2\alpha} \quad \text{and} \quad c(t, x) = c(t).$$

The problem of minimax sequential estimation for the gamma process and the loss function (6) for $\alpha = 1$ was considered in [1].

For each $\varepsilon > 0$ and $x_0 > 0$ we define the prior distribution $\pi_\varepsilon$ of $\theta$ with the density

$$g_\varepsilon(\theta) = \begin{cases} \dfrac{x_0^\varepsilon}{\Gamma(\varepsilon)} \cdot \theta^{-(\varepsilon+1)} \cdot \exp\{-x_0 \theta^{-1}\}, & \theta > 0, \\ 0, & \theta \leqslant 0. \end{cases}$$

Since for each $\varepsilon > 0$ and $x_0 > 0$ is

$$(7) \quad \int_0^\infty \theta^{-(\varepsilon+1)} \cdot \exp\{-x_0 \theta^{-1}\} \, d\theta = \int_0^\infty u^{\varepsilon+1} \cdot \exp\{-x_0 u\} u^{-2} \, du = \Gamma(\varepsilon_0)/x_0^\varepsilon,$$

the density of the posterior distribution of $\theta$ takes the form

$$g_\varepsilon(\theta|\tau = t,\ X(\tau) = x) = \begin{cases} \dfrac{x_1^T}{\Gamma(T)} \cdot \theta^{-(T+1)} \cdot \exp\{-x_1 \theta^{-1}\}, & \theta > 0, \\ 0, & \theta \leqslant 0, \end{cases}$$

where $x_1 = x_0 + x$, $T = t + \varepsilon$.

The $(t, x)$-Bayes estimator $d_\varepsilon^\bullet$ is a solution of the equation

$$\int_0^\infty (d_\varepsilon^\bullet(t, x) - \theta^\alpha) \cdot \theta^{-(T+1+2\alpha)} \cdot \exp\{-x_1 \cdot \theta^{-1}\}\, d\theta = 0.$$

Using (7) we calculate

$$d_\varepsilon^\bullet(t, x) = \frac{\Gamma(T+\alpha)}{\Gamma(T+2\alpha)} \cdot x_1^\alpha,$$

$$r_1\big(\pi_\varepsilon(\cdot|\tau = t,\ X(\tau) = x),\ d_\varepsilon^\bullet\big) = 1 - \frac{\Gamma^2(t+\alpha+\varepsilon)}{\Gamma(t+2\alpha+\varepsilon) \cdot \Gamma(t+\varepsilon)} \underset{\varepsilon \to 0}{\longrightarrow} 1 - \frac{\Gamma^2(t+\alpha)}{\Gamma(t+2\alpha)\Gamma(t)}.$$

Let us consider the fixed-time plan

$$\delta_t = (t, d_t), \qquad d_t = \frac{\Gamma(t+\alpha)}{\Gamma(t+2\alpha)} \cdot X^\alpha(t).$$

Since for each $\alpha > 0$ holds

$$\mathrm{E}_\theta [X(t)]^\alpha = \frac{\Gamma(t+\alpha)}{\Gamma(t)} \cdot \theta^\alpha,$$

we obtain

$$R(\delta_t, \theta) = \mathrm{E}_\theta \{(d_t - \theta^\alpha)^2 \cdot \theta^{-2\alpha} + c(t)\} = 1 - \frac{\Gamma^2(t+\alpha)}{\Gamma(t)\Gamma(t+2\alpha)} + c(t).$$

Finally, let us note that $\inf_{d \in R^1} \sup_{\theta \in \Theta} (d - \theta^\alpha)^2 \theta^{-2\alpha} \geqslant 1$.

Thus, if there exists $t_0 > 0$ such that

$$\min_{t>0} \left\{1 - \frac{\Gamma^2(t+\alpha)}{\Gamma(t)\Gamma(t+2\alpha)} + c(t)\right\} = 1 - \frac{\Gamma^2(t_0+\alpha)}{\Gamma(t_0)\Gamma(t_0+2\alpha)} + c(t_0) \leqslant 1 + c(0),$$

then the fixed-time plan $\delta_{t_0} = (t_0, d_{t_0})$ is minimax in the class $\mathscr{D}$.

3. Using Theorem 1 we solve in this section one problem of minimax sequential estimation in the case when $c(t, z) = c(z)$ for each $(t, z) \in U$.

Let $X(t)$, $t \in T = \{1, 2, \ldots\}$, be the multinomial process with the parameter $p$ and let

$$(8) \qquad L(d, p) = \left( d - \frac{\sum_{i=1}^{r} a_i p_i + b}{\sum_{i=1}^{r} b_i p_i} \right)^2 \frac{(\sum_{i=1}^{r} b_i p_i)^2}{z(p)}, \qquad c(t, x) = c\left(\sum_{i=1}^{r} b_i x_i\right),$$

where

$$z(p) = \{(\sum_{i=1}^{r} b_i p_i)(\sum_{i=1}^{r} a_i^2 p_i) - 2(\sum_{i=1}^{r} a_i b_i p_i)(\sum_{i=1}^{r} a_i p_i + b) + \\ + (\sum_{i=1}^{r} a_i p_i + b)^2 - b^2 \cdot (\sum_{i=1}^{r} b_i p_i)\}.$$

We assume throughout the rest of this section that $b_1, \ldots, b_r, a_1, \ldots, a_r, b$ are the real numbers such that $b_i$ is equal to either 0 or 1, $i = 1, \ldots, r$, and one of the following condition is fulfilled:

$$(9) \qquad a_i b_i = 0 \quad \text{for each } 1 \leqslant i \leqslant r,\ b \text{ arbitrary},$$

$$(10) \qquad \text{if } b_i = 0 \text{ then } a_i = 0 \text{ for each } 1 \leqslant i \leqslant r,\ b = 0.$$

The case $b_1 = b_2 = \ldots = b_r = 1$ has been considered in the previous section $\left(\sum_{i=1}^{r} x_i = t\right)$. Hence we assume that there exists $i_0$, $1 \leqslant i_0 \leqslant r$, for which $b_{i_0} = 0$. Without loss of generality we put $b_r = 0$. Moreover, since $p_1 + \ldots + p_r = 1$, we may suppose that $a_r = 0$.

Let us define for each $y \in T$ the random variable

$$\tau_y = \inf \{t \in T: \sum_{i=1}^{r-1} b_i X_i(t) = y\}.$$

By the result obtained in [4] it follows that $\tau_y$ is a stopping time for each $y \in T$ and $p \in P_0 = \{p \in P: \sum_{i=1}^{r-1} b_i p_i > 0\}$ and

$$(11) \qquad \mathrm{E}_p \tau_y = y \cdot \left(\sum_{i=1}^{r-1} b_i p_i\right)^{-1}, \quad \mathrm{E}_p X_i(\tau_y) = p_i \mathrm{E}_p \tau_y,$$

$$(12) \qquad \mathrm{D}_p^2\left(\sum_{i=1}^{r-1} a_i X_i(\tau_y) + b\tau_y\right) = yz(p)\left(\sum_{i=1}^{r-1} b_i p_i\right)^{-2}.$$

Now let us define the sequential plans (the inverse plans)

$$\delta_y^1 = (\tau_y, d_y^1) \quad \text{where} \quad d_y^1 = \Big(\sum_{i=1}^{r-1} a_i X_i(\tau_y) + b(\tau_y+1)\Big)\cdot(y+1)^{-1},$$

$$\delta_y^2 = (\tau_y, d_y^2) \quad \text{where} \quad d_y^2 = \Big(\sum_{i=1}^{r-1} a_i X_i(\tau_y)\Big)\cdot y^{-1}.$$

Making use of (11) and (12) we get

$$R(\delta_y^1, p) = (y+1)^{-2}\,\mathrm{E}_p\left\{\sum_{i=1}^{r-1} a_i X_i(\tau_y) + b(\tau_y+1) - (y+1)\frac{\sum_{i=1}^{r-1} a_i p_i + b}{\sum_{i=1}^{r-1} b_i p_i}\right\}^2 \times$$

$$\times \frac{(\sum_{i=1}^{r-1} b_i p_i)^2}{z(p)} + \mathrm{E}_p\{c(\sum_{i=1}^{r-1} b_i X_i(\tau_y))\}$$

$$= \frac{(\sum_{i=1}^{r-1} b_i p_i)^2}{(y+1)^2\cdot z(p)}\cdot \mathrm{D}_p^2\Big(\sum_{i=1}^{r-1} a_i X_i(\tau_y) + b\tau_y\Big) + c(y) +$$

$$+ \frac{[b(\sum_{i=1}^{r-1} b_i p_i) - (\sum_{i=1}^{r-1} a_i p_i + b)]^2}{z(p)(y+1)^2}$$

$$= \frac{y}{(y+1)^2} + c(y) + \frac{[b(\sum_{i=1}^{r-1} b_i p_i)^2 - (\sum_{i=1}^{r-1} a_i p_i + b)]^2}{z(p)(y+1)^2}.$$

Note that if (9) holds then

$$[b(\sum_{i=1}^{r-1} b_i p_i) - (\sum_{i=1}^{r-1} a_i p_i + b)]^2 - z(p)$$

$$= b^2(\sum_{i=1}^{r-1} b_i p_i)^2 - 2b(\sum_{i=1}^{r-1} b_i p_i)(\sum_{i=1}^{r-1} a_i p_i + b) - (\sum_{i=1}^{r-1} a_i^2 p_i)(\sum_{i=1}^{r-1} b_i p_i) + b^2(\sum_{i=1}^{r-1} b_i p_i)$$

$$= (\sum_{i=1}^{r-1} b_i p_i)\{\sum_{i=1}^{r-1} (b^2 b_i - b^2 - 2ba_i - a_i^2)p_i - b^2(1 - \sum_{i=1}^{r-1} p_i)\}$$

$$\leqslant (\sum_{i=1}^{r-1} b_i p_i)\{\sum_{i=1}^{r-1} (b^2 b_i - b^2 - 2ba_i - a_i^2)p_i\} \leqslant 0$$

(if $b_i = 1$ then $a_i = 0$ and $b^2 b_i - b^2 - 2ba_i - a_i^2 = 0$, if $a_i = 0$ then $b_i = 0$ and $b^2 b_i - b^2 - 2ba_i - a_i^2 = 0 - (b+a_i)^2 \leqslant 0$).

Thus

(13) $$R(\delta_y^1, p) \leqslant 1/(y+1)+c(y), \quad \text{for each } p \in P_0.$$

In the same way we can check that under (10) we have

$$R(\delta_y^2, p) \leqslant 1/y+c(y), \quad \text{for each } p \in P_0.$$

We prove the following

LEMMA 1. *Suppose that the loss function and the cost function are given by* (8).

*Under* (9) *if there exists* $y_1 \in T$ *such that*

(14) $$1/(y_1+1)+c(y_1) = \min_{y \in T} \{1/(y+1)+c(y)\},$$

*then the sequential plan* $\delta_{y_1}^1 = (\tau_{y_1}, d_{y_1}^1)$ *is minimax in* $\mathscr{D}$.

*Under* (10) *if there exists* $y_2 \in T$ *such that*

$$1/y_2+c(y_2) = \min_{y \in T} \{1/y+c(y)\},$$

*then the sequential plan* $\delta_{y_2}^2 = (\tau_{y_2}, d_{y_2})$ *is minimax in* $\mathscr{D}$.

Proof. To prove the first part of Lemma 1 we define for each $\varepsilon > 0$ the prior distribution $\pi_\varepsilon$ of $p$ given by the density

$$g_\varepsilon(p) = \begin{cases} C_\varepsilon z(p)(p_1 \cdot \ldots \cdot p_r)^{\varepsilon-1}, & \text{if} \quad p \in P, \\ 0, & \text{if} \quad p \notin P, \end{cases}$$

where

$$C_\varepsilon^{-1} = \int_{p \in P} \ldots \int z(p)(p_1 \cdot \ldots \cdot p_r)^{\varepsilon-1} \, dp_1 \cdot \ldots \cdot dp_{r-1}.$$

Thus

$$g_\varepsilon\big(p \mid \tau = t, X(\tau) = x\big) = \begin{cases} D_\varepsilon \cdot z(p) p_1^{x_1+\varepsilon-1} \cdot \ldots \cdot p_r^{x_r+\varepsilon-1}, & \text{if} \quad p \in P, \\ 0, & \text{if} \quad p \notin P, \end{cases}$$

where

$$D_\varepsilon^{-1} = \int_{p \in P} \ldots \int z(p) p_1^{x_1+\varepsilon-1} \cdot \ldots \cdot p_r^{x_r+\varepsilon-1} \, dp_1 \cdot \ldots \cdot dp_{r-1}$$

$$= \frac{\Gamma(x_1+\varepsilon) \cdot \ldots \cdot \Gamma(x_r+\varepsilon)}{\Gamma(t+r\varepsilon+2)} \{[\sum_{i=1}^{r} a_i(x_i+\varepsilon)+b(t+r\varepsilon+1)]^2 +$$

$$+[\sum_{i=1}^{r-1} a_i^2(x_i+\varepsilon)-b^2(t+r\varepsilon+1)][\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)+1]\}.$$

Recalling Definition 7 we see that the $(t, x)$-Bayes estimator $d^{\bullet}_{\pi_\varepsilon}(t, x) = d^{\bullet}_{\varepsilon}(t, x)$ fulfils the equation

$$\int_{p\in P}\cdots\int\Big(d^{\bullet}_{\varepsilon}(t, x)-\frac{\sum_{i=1}^{r-1} a_i p_i+b}{\sum_{i=1}^{r-1} b_i p_i}\Big)\Big(\sum_{i=1}^{r-1} b_i p_i\Big)^2 p_1^{x_1+\varepsilon-1}\cdot\ldots\cdot p_r^{x_r+\varepsilon-1}\,dp_1\cdot\ldots\cdot dp_{r-1}=0.$$

Solving this equation we obtain

$$d^{\bullet}_{\varepsilon}(t, x)\Big\{\Big(\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)\Big)\Big(\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)+1\Big)\Big\}$$
$$=\sum_{i=1}^{r-1} a_i b_i(x_i+\varepsilon)+\Big(\sum_{i=1}^{r-1} a_i(x_i+\varepsilon)\Big)\Big(\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)\Big)+$$
$$+b(t+r\varepsilon+1)\Big(\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)\Big).$$

Hence

$$d^{\bullet}_{\varepsilon}(t, x)=\frac{\sum_{i=1}^{r-1} a_i(x_i+\varepsilon)+b(t+r\varepsilon+1)}{\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)+1},$$

provided (9) holds.

The posterior risk is

$$r_1\big(\pi_\varepsilon(\cdot|\tau=t,\, X(\tau)=x),\, d^{\bullet}_{\varepsilon}\big)$$
$$=D_\varepsilon\cdot\Big\{\int_{p\in P}\cdots\int\Big(d^{\bullet}_{\varepsilon}(t, x)-\frac{\sum_{i=1}^{r-1} a_i p_i+b}{\sum_{i=1}^{r-1} b_i p_i}\Big)^2\frac{\big(\sum_{i=1}^{r-1} b_i p_i\big)^2}{z(p)}z(p)\times$$
$$\times p_1^{x_1+\varepsilon-1}\cdot\ldots\cdot p_r^{x_r+\varepsilon-1}\,dp_1\cdot\ldots\cdot dp_{r-1}\Big\}$$
$$=D_\varepsilon\cdot\Big\{\frac{\big(\sum_{i=1}^{r-1} a_i^2(x_i+\varepsilon)-b^2(t+r\varepsilon+1)\big)\big(\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)+1\big)}{\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)+1}+$$
$$+\frac{-\big[\sum_{i=1}^{r-1} a_i(x_i+\varepsilon)+b(t+r\varepsilon+1)\big]^2\cdot\big[\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)\big]}{\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)+1}+$$

$$+\frac{\left[\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)+1\right]\left[\sum_{i=1}^{r-1} a_i(x_i+\varepsilon)+b(t+r\varepsilon+1)\right]^2}{\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)+1}\Bigg\}\frac{\Gamma(x_1+\varepsilon)\cdot\ldots\cdot\Gamma(x_r+\varepsilon)}{\Gamma(t+r\varepsilon+2)}$$

$$=\frac{1}{\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)+1}\xrightarrow[\varepsilon\to 0]{}\frac{1}{\sum_{i=1}^{r-1} b_i x_i+1}.$$

By Theorem 1, (13) and (14) the rest of the proof is now straightforward.

Using the same sequence $\pi_\varepsilon$, $\varepsilon > 0$, of the prior distributions of $p$ we obtain that under (10)

$$d_\varepsilon^{\bullet}(t,\, x) = \frac{\sum_{i=1}^{r-1} a_i(x_i+\varepsilon)}{\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)},$$

$$r_1\big(\pi_\varepsilon(\cdot|\tau = t,\, X(\tau) = x),\, d_\varepsilon^{\bullet}\big) = \frac{1}{\sum_{i=1}^{r-1} b_i(x_i+\varepsilon)}\xrightarrow[\varepsilon\to 0]{}\frac{1}{\sum_{i=1}^{r-1} b_i x_i}.$$

By the arguments used to prove the first part of Lemma 1 we obtain the required results concerning the plan $\delta_{y_2} = (\tau_{y_2},\, d_{y_2}^2)$.

## References


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

[2] G. M. Fichtenholz, *Rachunek różniczkowy i całkowy*, PWN, Warszawa 1975, vol. 3, p. 330.

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

[4] W. Stefanow, *Plany ukośne dla procesu wielomianowego*, Matematyka Stosowana 17 (1981), p. 95-104.

[5] S. Trybuła, *Some investigations in minimax estimation theory*, Dissertationes Math. 240, Warszawa 1985, 44 pp.



INSTITUTE OF MATHEMATICS
TECHNICAL UNIVERSITY WROCŁAW
50-370 WROCŁAW



*Received on 21. 11. 1983;*

*revised version on 30. 05. 1984*