ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria III: MATEMATYKA STOSOWANA XVII (1981)

WALERY STEFANOW (Sofia)

Plany ukośne dla procesu wielomianowego

(Praca przyjęta do druku 26.06.1978)

1. Wstęp. Rozważmy wektor losowy $Y = (y_1, y_2, \ldots, y_n)$, dla którego $P\{Y = e_j\} = p_j > 0$ $(j = 1, 2, \ldots, n)$, gdzie e_j jest wektorem z jedynką na j-tym miejscu i zerami na miejscach pozostałych oraz gdzie $\sum_{j=1}^{n} p_j = 1$.

Niech

$$X = (X_1, X_2, \ldots, X_n) = \sum_{i=1}^{\tau} Y_i$$

będzie sumą τ niezależnych obserwacji wektora losowego Y. Wtedy $\sum_{j=1}^{n} X_j = \tau$ jest liczbą obserwacji, która w naszym przypadku jest zmienną losową przybierającą tylko wartości całkowite nieujemne i taką, że zdarzenie $\{\tau = k\}$ jest mierzalne względem σ-algebry generowanej przez wektory losowe $Y_1, Y_2, \ldots, Y_k$ dla każdej całkowitej wartości k.

Proces stochastyczny z czasem dyskretnym o przyrostach niezależnych, którego przyrosty mają taki sam rozkład jak wektor Y, startujący z początku układu współrzędnych, będziemy nazywać *procesem wielomianowym*.

Zadaniem naszym będzie oszacowanie parametru $Q = g(p)$, $p = (p_1, p_2, \ldots, p_n)$, w przypadku, gdy p jest nie znane. Znajdziemy pewną klasę optymalnych planów sekwencyjnych tego parametru. Wykażemy mianowicie, że reguła stopu τ, zdefiniowana jako chwila pierwszego wejścia do zbioru

$$A = \{(x_1, x_1, \ldots, x_n)\colon a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b\}, \tag{1.1}$$

gdzie $a_1, a_2, \ldots, a_n, b$ są liczbami całkowitymi o największym wspólnym dzielniku równym 1, $b > 0$, $a_k \leqslant 1$ dla $k = 1, 2, \ldots, n$ i co najmniej dla jednego $k = k_0$, $a_{k_0} = 1$, wyznacza optymalny plan sekwencyjny. Później udowodnimy, że jedynymi planami optymalnymi są plany pierwszego osiągnięcia.

2. Określenia i nierówność informacyjna.

DEFINICJA 1. Niech τ będzie regułą stopu. Funkcję $f = f(X, \tau)$ będziemy nazywać *estymatorem parametru* $Q = g(p)$. Estymator f nazywamy estymatorem *nieobcią-*

żonym, jeżeli $E_p(f) = g(p)$, gdzie $E_p(f)$ jest wartością oczekiwaną estymatora f dla danych p i τ.

DEFINICJA 2. *Planem sekwencyjnym* będziemy nazywać trójkę (τ, g, f) składającą się z reguły stopu τ, funkcji $g(p)$ i jej nieobciążonego estymatora f.

Niech $\mathscr{P}$ będzie jakimś ustalonym zbiorem wartości parametru wektorowego p. Plan sekwencyjny S będziemy nazywać *domkniętym dla* $\mathscr{P}$, jeżeli dla każdego $p \in \mathscr{P}$ $P\{\tau < +\infty\} = 1$.

Wśród wszystkich planów S wyróżniamy klasę planów pierwszego wejścia. Plan pierwszego wejścia określony jest za pomocą zbioru brzegowego A, po wejściu do którego trajektoria zatrzymuje się. Punkty zbioru A nazywamy *brzegowymi*. Punkt nazywamy *przejściowym*, jeśli trajektoria wychodząc z początku układu współrzędnych może go osiągnąć i wyjść z niego z dodatnim prawdopodobieństwem. Punktami *nieosiągalnymi* nazywamy punkty, które nie są ani przejściowymi, ani brzegowymi.

DEFINICJA 3. Plan sekwencyjny S, którego zbiór brzegowy A jest określony wzorem (1.1), będziemy nazywać *stałym*, jeśli $a_k = 1$, $1 \leqslant k \leqslant n$.

DEFINICJA 4. Plan sekwencyjny S, którego zbiór brzegowy A jest określony wzorem (1.1), będziemy nazywać *odwrotnym*, jeśli $a_{i} = a_{i} = \ldots = a_{i_k} = 1$, $a_{i_{k+1}} = a_{i_{k+}} = \ldots = a_{i_n} = 0$, $1 \leqslant k < n$, gdzie $(i_1, i_2, \ldots, i_n)$ jest dowolną ustaloną permutacją zbioru $(1, 2, \ldots, n)$.

Estymacji sekwencyjnej dla procesu wielomianowego poświęcono pracę [1], część pracy [4] i parę uwag w książce [2] (rozdział 12). Na przykład, w pracy [1] rozpatrywano plany stałe i odwrotne.

Niech teraz $\alpha \in A$ i niech $\pi(\alpha)$ oznacza liczbę różnych trajektorii wychodzących z początku układu współrzędnych, które po raz pierwszy osiągają zbiór A w punkcie α. Wtedy prawdopodobieństwo osiągnięcia tego punktu wynosi

$$\pi(\alpha) p_1^{\alpha} p_2^{\alpha} \ldots p_n^{\alpha_n}, \qquad \text{gdzie} \qquad \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n).$$

A zatem plan A jest domknięty dla $p \in \mathscr{P}$, gdy

$$\sum_{\alpha \in A} \pi(\alpha) p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_n^{\alpha_n} = 1. \tag{2.1}$$

Będziemy rozpatrywać takie plany, dla których zachodzi (2.1) i ponadto takie, dla których szereg po lewej stronie równości (2.1) można różniczkować wyraz po wyrazie względem p_j. Niech f będzie nieobciążonym estymatorem funkcji

$$g(p) = E_p(f) = \sum_{\alpha \in A} f(\alpha) \pi(\alpha) p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_n^{\alpha_n}.$$

Dalej będziemy zakładali, że szereg ten jest absolutnie zbieżny i że możemy go różniczkować wyraz po wyrazie względem p_j:

$$\frac{\partial E(f)}{\partial p_j} = \sum_{\alpha \in A} f(\alpha) \pi(\alpha) \left(\frac{\alpha_j}{p_j} - \frac{\alpha_n}{p_n} \right) \prod_{j=1}^{n} p_j^{\alpha_j} = E(f) U_j, \qquad 1 \leqslant j \leqslant n-1,$$

gdzie

$$U_j = \frac{X_j}{p_j} - \frac{X_n}{p_n}, \quad 1 \leqslant j \leqslant n-1. \tag{2.2}$$

$(X_1, X_2, \dots, X_n) = \alpha$ z prawdopodobieństwem $\pi(\alpha)\prod_{j=1}^{n} p_j^{\alpha_j}$. Oczywiście, $(X_1, X_2, \dots$ $\dots, X_n) = \sum_{i=1}^{\tau} Y_i$, jak to było zdefiniowane we wstępie; τ oznacza moment pierwszego osiągnięcia zbioru A.

Różniczkując (2.1) względem p_j, otrzymujemy $EU_j = 0$, a zatem

$$p_n E(X_j) = p_j E(X_n),$$

i z kolei

$$E(X_j) = p_j E(\tau). \tag{2.3}$$

TWIERDZENIE 1. (Nierówność informacyjna). *Dla każdego estymatora f i planu sekwencyjnego S, które spełniają powyższe założenia, zachodzi nierówność*

$$\mathrm{Var}_p f \geqslant \{E(\tau)\}^{-1}\left[\sum_{j=1}^{n-1} p_j (g_j'(p))^2 - \left(\sum_{j=1}^{n-1} p_j g_j'(p)\right)^2\right], \tag{2.4}$$

gdzie $g_j'(p)$ oznacza pochodną funkcji $g(p)$ względem p_j.

Nierówność (2.4) *przechodzi w równość wtedy i tylko wtedy, gdy f jest liniową funkcją zmiennej U_j, $j = 1, 2, \dots, n-1$.*

D o w ó d. Patrz [1].

DEFINICJA 5. Plan sekwencyjny (τ, g, f) jest *efektywny w punkcie* $p = p^0$, jeśli w (2.4) zachodzi równość dla $p = p^0$. Estymator f jest wtedy *efektywny w punkcie* $p = p^0$, a funkcja $g(p)$ jest *efektywnie estymowalna* w tym punkcie.

DEFINICJA 6. Plan sekwencyjny (τ, g, f) nazywamy planem *efektywnym w* $\mathscr{P}$, jeśli jest on efektywny dla każdej wartości $p \in \mathscr{P}$. Estymator f nazywa się wtedy estymatorem *efektywnym w* $\mathscr{P}$, a funkcja $g(p)$ jest *efektywnie estymowalna w* $\mathscr{P}$.

3. Plany ukośne.

DEFINICJA 7. Plan sekwencyjny nazywa się planem *ukośnym*, jeżeli nie jest planem stałym ani odwrotnym i jeżeli zbiór brzegowy A określony jest następująco:

$$A = \{(x_1, x_2, \dots, x_n) \colon a_1 x_1 + a_2 x_2 + \dots + a_n x_n = b\}, \tag{3.1}$$

gdzie $a_1, a_2, \dots, a_n, b$ są liczbami całkowitymi o największym wspólnym dzielniku równym 1, $b > 0$, $a_k \leqslant 1$ dla $k = 1, 2, \dots, n$ i co najmniej dla jednego $k = k_0$ $a_{k_0} = 1$.

TWIERDZENIE 2 (Małyszew, patrz [2], [4]). *Niech A będzie zbiorem danym wzorem* (3.1). *Jeśli $b > 0$ i a_i, $i = 1, 2, \dots, n$,* są *liczbami całkowitymi o największym wspólnym dzielniku równym* 1, *to warunkiem koniecznym na to, żeby plan ze zbiorem brzegowym A był domknięty jest $a_k \leqslant 1$ dla $k = 1, 2, \dots, n$ i co najmniej dla jednego $k = k_0$, $a_{k_0} = 1$.*

TWIERDZENIE 3. *Jeśli $a_1, a_2, \ldots, a_n, b$ są liczbami całkowitymi o największym wspólnym dzielniku równym 1, $b > 0$, $a_k \leqslant 1$ dla $k = 1, 2, \ldots, n$ i co najmniej dla jednego $k = k_0$, $a_{k_0} = 1$ i jeśli ponadto $a_1p_1+a_2p_2+ \ldots +a_np_n > 0$, to plan, którego zbiór brzegowy A jest określony wzorem (3.1), jest domknięty.*

D o w ó d. W myśl dowodu twierdzenia Małyszewa wystarczy wykazać, że istnieje u_0 takie, że dla każdej liczby naturalnej u, większej od u_0 zachodzi nierówność

$$a_1X_1(u)+a_2X_2(u)+ \ldots +a_nX_n(u) > b \quad \text{p.w.},$$

gdzie $X(u) = (X_1(u), \ldots, X_n(u))$ jest sumą u niezależnych wektorów losowych $Y = (Y_1, Y_2, \ldots, Y_n)$, przy czym $Y = (Y_1, Y_2, \ldots, Y_n)$ ma rozkład podany we wstępie.

Rozpatrzmy zmienną losową $Z = a_1Y_1+a_2Y_2+ \ldots +a_nY_n$. Oczywiście, $a_1X_1(u)+ \ldots +a_nX_n(u)$ można interpretować jako sumę $\sum_{i=1}^{n} Z_i$, gdzie Z_i są niezależnymi zmiennymi losowymi o rozkładzie identycznym z rozkładem zmiennej Z. Wtedy, z prawa iterowanego logarytmu, mamy

$$P\left\{\underline{\lim_u} \frac{Z_1+Z_2+ \ldots +Z_n-u(a_1p_1+a_2p_2+ \ldots +a_np_n)}{\sqrt{\mathrm{Var}Z}\sqrt{2u\ln\ln u}} = -1\right\} = 1,$$

a więc

$$P\left\{\underline{\lim_u} \frac{a_1X_1(u)+ \ldots +a_nX_n(u)-u(a_1p_1+a_2p_2+ \ldots +a_np_n)}{\sqrt{\mathrm{Var}Z}\sqrt{2u\ln\ln u}} = -1\right\} = 1. \tag{3.2}$$

Z (3.2) otrzymujemy, że dla dostatecznie dużych u i dowolnego $\varepsilon > 0$

$$P\{a_1X_1(u)+ \ldots +a_nX_n(u) > \\ > u(a_1p_1+a_2p_2+ \ldots a_np_n)-(1+\varepsilon)\sqrt{\mathrm{Var}Z2u\ln\ln u}\} = 1 \tag{3.3}$$

Ponieważ dla dostatecznie dużych u i dowolnego ustalonego $\varepsilon_1 > 0$ mamy $\sqrt{2u\ln\ln u} < u\varepsilon_1$, więc dobierając odpowiednio małe ε_1 otrzymujemy, że dla dostatecznie dużych u prawa strona nierówności (3.3) jest większa od b, co kończy dowód twierdzenia.

W identyczny sposób dowodzimy, że gdy $a_1p_1+a_2p_2+ \ldots +a_np_n < 0$, wówczas plan, którego zbiór brzegowy jest dany wzorem (3.1), nie jest domknięty.

Teraz zajmiemy się dowodem konieczności i dostateczności spełnienia warunku z twierdzenia 3 do różniczkowania szeregu (2.1) wyraz po wyrazie względem p_j, $j = 1, 2, \ldots, n-1$, w przypadku gdy A jest określony wzorem (3.1).

Rozpatrzmy hiperpłaszczyznę

$$B_b = \{(x_1, x_2, \ldots, x_{m+1}): x_{m+1}-a_1x_1-a_2x_2- \ldots -a_mx_m-b = 0\},$$

gdzie $a_1, a_2, \ldots, a_m$ i b są liczbami całkowitymi i $0 \leqslant a_1 < a_2 < \ldots < a_m$, $b > 0$. Oczywiście możemy ograniczyć się do rozpatrzenia hiperpłaszczyzny B_b, bo wszystkie osie są równouprawnione co do numeracji, a jeżeli w (3.1) $a_i = a_j$ dla $i \neq j$, to domkniętość planu ze zbiorem brzegowym A i istnienie momentów chwili pierwszego zatrzymania w A jest równoważne z sytuacją, gdy zamiast posunięć w kierunku i-tej osi i j-tej osi rozpatrujemy posunięcie tylko w kierunku jednej osi z prawdopodobieństwem p_i+p_j.

Niech teraz $p_{l_1 l_2 \ldots l_m b}^{a_1 a_2 \ldots a_m}$ i $q_{l_1 l_2 \ldots l_m b}^{a_1 a_2 \ldots a_m}$ oznaczają odpowiednio prawdopodobieństwo trafienia do zbioru B_b po raz pierwszy w punkcie $(l_1, l_2, \ldots, l_m, a_1 l_1+a_2 l_2+ \ldots + a_m l_m+b)$, oraz liczbę różnych trajektorii łączących początek układu współrzędnych z punktem $(l_1, l_2, \ldots, l_m, a_1 l_1+a_2 l_2+ \ldots +a_m l_m+b)$ i nie przechodzących przez punkty zbioru B_b. Oczywiście, zachodzą następujące równości:

$$p_{l_1 l_2 \ldots l_m b}^{a_1 a_2 \ldots a_m} = \sum_{i=1}^{m} p_i p_{l_1 \ldots l_{i-1} l_i - 1, l_{i+1} \ldots l_m b + a_i}^{a_1 a_2 \ldots a_m} + p_{m+1} p_{l_1 l_2 \ldots l_m b-1}^{a_1 a_2 \ldots a_m},$$

$$q_{l_1 l_2 \ldots l_m b}^{a_1 a_2 \ldots a_m} = \sum_{i=1}^{m} q_{l_1 l_2 \ldots l_{i-1} l_i - 1, l_{i+1} \ldots l_m b + a_i}^{a_1 a_2 \ldots a_m} + q_{l_1 l_2 \ldots l_m b-1}^{a_1 a_2 \ldots a_m}, \tag{3.4}$$

$$q_{00 \ldots 0b}^{a_1 a_2 \ldots a_m} = 1 \quad (b = 0, 1, 2, \ldots), \qquad q_{0 \ldots 0 l_i 0 \ldots 0}^{a_1 a_2 \ldots a_m} = 0 \quad (l_i = 1, 2, \ldots). \tag{3.5}$$

Równanie (3.4) z warunkami brzegowymi (3.5) ma jedno rozwiązanie i, jak łatwo można się przekonać, jest nim

$$q_{l_1 l_2 \ldots l_m b}^{a_1 a_2 \ldots a_m} = \frac{b[b+l_1(a_1+1)+ \ldots +l_m(a_m+1)-1]!}{l_1! \, l_2! \ldots l_m! [b+l_1 a_1 + \ldots + l_m a_m]!}.$$

W przypadku procesu dwumianowego rozwiązanie to zostało podane w pracy [3].

W przypadku, gdy zbiorem brzegowym jest B_b, szereg (2.1) ma postać

$$\sum_{l_1, l_2, \ldots, l_m = 0} p_{l_1 l_2 \ldots l_m b}^{a_1 a_2 \ldots a_m} \tag{3.6}$$

Wprowadźmy następujące oznaczenie:

$$c_{l_1 l_2 \ldots l_m} \stackrel{\mathrm{df}}{=} \{p_{l_1 l_2 \ldots l_m b}^{a_1 a_2 \ldots a_m}\}^{1/(l_1+l_2+\ldots+l_m)}.$$

Aby szereg (3.6) można było różniczkować wyraz po wyrazie dowolną ilość razy, wystarczy sprawdzić, czy

$$\limsup_{l_1+l_2+\ldots+l_m} c_{l_1 l_2 \ldots l_m} < 1. \tag{3.7}$$

Wykażemy, jak wyglądają wszystkie punkty skupienia ciągu $\{c_{l_1 l_2 \ldots l_m}\}$. Stosując wzór Stirlinga, otrzymujemy

$$q_{l_1 l_2 \ldots l_m b}^{a_1 a_2 \ldots a_m} = C \frac{[b+(a_1+1)l_1+ \ldots +(a_m+1)l_m-1]^{b+(a_1+1)l_1+\ldots+(a_m+1)l_m}}{l_1^{l_1} l_2^{l_2} \ldots l_m^{l_m} [b+a_1 l_1+ \ldots +a_m l_m]^{b+a_1 l_1+\ldots+a_m l_m}},$$

gdzie $C^{1/(l_1+l_2+\ldots+l_m)} \xrightarrow[l_1+\ldots+l_m \to \infty]{} 1$.

Niech $l_1+l_2+ \ldots +l_m$ dąży do nieskończoności w ten sposób, że

$$\frac{l_i}{l_1+l_2+ \ldots +l_m} \to r_i, \qquad i = 1, 2, \ldots, m, \quad r_i \in [0, 1].$$

Nazwijmy ten sposób dążenia $l_1+l_2+\ldots+l_m$ do nieskończoności sposobem $(r_1, r_2, \ldots, r_m)$ (oczywiście, $\sum_{i=1}^{m} r_i = 1$). Gdy $l_1+l_2+\ldots+l_m$ dąży do nieskończoności w sposób $(r_1, r_2, \ldots, r_m)$, wtedy

$$c_{l_1l_2\ldots l_m} \xrightarrow[l_1+\ldots+l_m+\infty]{} \frac{\Big[\sum_{i=1}^{m}(a_i+1)r_i\Big]^{\sum_{i=1}^{m}(a_i+1)r_i}}{\prod_{i=1}^{m} r_i^{r_i}\Big[\sum_{i=1}^{m} a_ir_i\Big]^{\sum_{i=1}^{m}a_ir_i}} \prod_{i=1}^{m} p_i^{r_i} p_{m+1}^{\sum_{i=1}^{m}a_ir_i} \overset{\mathrm{df}}{=} : \varphi(r_1, r_2, \ldots, r_{m-1}).$$

Oczywiście zbiór wszystkich punktów skupienia ciągu $\{c_{l_1l_2\ldots l_m}\}$ jest zbiorem wartości funkcji $\varphi(r_1, r_2, \ldots, r_{m-1})$ zdefiniowanej na podzbiorze z $[0, 1]^{m-1}$. Niech bowiem s będzie punktem skupienia ciągu $\{c_{l_1l_2l_m}\}$. Wówczas istnieje podciąg ciągu $\{c_{l_1l_2\ldots l_m}\}$ zbieżny do s, gdy $l_1+l_2+\ldots+l_m \to \infty$. Niech to będzie podciąg $\{c_{k_1k_2\ldots k_m}\}$. Ale istnieją $0 \leqslant r_1 \leqslant 1$ oraz taki podciąg $\{c_{t_1t_2\ldots t_m}\}$ ciągu $\{c_{k_1k_2\ldots k_m}\}$, że $\frac{t_1}{t_1+\ldots+t_m} \xrightarrow[t_1+\ldots+t_m\to\infty]{} r_1$. Z tego podciągu można wybrać podciąg $\{c_{m'_1m'_2\ldots m_m}\}$ taki, że $\frac{m'_2}{m'_1+\ldots+m'_m} \xrightarrow[m'_1+\ldots+m'_m\to\infty]{} r_2$ dla pewnego $0 \leqslant r_2 \leqslant 1$, itd. Istnieją więc $r_1, r_2, \ldots, r_{m-1}$ takie, że $0 \leqslant r_i \leqslant 1$, $\sum_{i=1}^{m-1} r_i \leqslant 1$, i podciąg ciągu $\{c_{l_1l_2\ldots l_m}\}$ taki, że suma $l_1+l_2+\ldots+l_m$ dąży do nieskończoności w sposób $(r_1, r_2, \ldots, r_m)$, gdzie $r_m = 1-\sum_{i=1}^{m-1} r_i$, co oznacza, że $s = \varphi(r_1, r_2, \ldots, r_{m-1})$.

Oczywiście, $\varphi(r_1, r_2, \ldots, r_{m-1})$ zależy jeszcze od $p_1, p_2, \ldots, p_m$.

Wykażemy, że dla ustalonych $r_1, r_2, \ldots, r_{m-1}$ funkcja $\varphi(r_1, r_2, \ldots, r_{m-1})$ osiąga maksimum tylko w punkcie $(p_1, p_2, \ldots, p_{m+1})$ spełniającym warunek

$$p_{m+1}-a_1p_1-a_2p_2-\ldots-a_mp_m = 0 \tag{3.8}$$

i że maksimum to jest równe 1.

Dla ustalonych $p_2, p_3, \ldots, p_m$ funkcja $\varphi(r_1, r_2, \ldots, r_{m-1})$ (oczywiście, dalej zakładamy, że $r_1, r_2, \ldots, r_{m-1}$ są ustalone) osiąga maksimum tylko w punkcie

$$p_1 = \frac{(1-p_2-p_3-\ldots-p_m)r_1}{[(a_1+1)r_1+\sum_{i=2}^{m} a_ir_1]}, \tag{3.9}$$

co wynika z równania $\partial\varphi/\partial p_1 = 0$. Zastępując p_1 w $\varphi(r_1, r_2, \ldots, r_{m-1})$ wartością z (3.9), otrzymujemy, że φ zależy tylko od $p_2, p_3, \ldots, p_m$. Analogicznie, ustalając $p_3, p_4, \ldots, p_m$ dochodzimy do wniosku, że φ osiąga maksimum tylko w punkcie

$$p_2 = \frac{(1-p_3-p_4-\ldots-p_m)r_2}{[(a_1+1)r_1+(a_2+1)r_2+\sum_{=3}^{m} a_ir_i]},$$

itd.

Tak więc $\varphi(r_1, r_2, \ldots, r_{m-1})$ osiąga maksimum tylko w punkcie $(p_1, p_2, \ldots, p_{m+1})$ danym układem równań

$$p_i = \frac{(1-p_{i+1}-\ldots-p_m)r_i}{[\sum_{j=1}^{i}(a_j+1)r_j+\sum_{j=i+1}^{m}a_jr_j]}, \quad i = 1, 2, \ldots, m. \tag{3.10}$$

Rozwiązując układ (3.10) znajdujemy

$$p_i = \frac{r_i}{[\sum_{j=1}^{m}(a_j+1)r_j]}, \quad i = 1, 2, \ldots, m. \tag{3.11}$$

Oczywiście z (3.11) otrzymujemy

$$p_{m+1} = \frac{\sum_{j=1}^{m} a_jr_j}{\sum_{j=1}^{m}(a_j+1)r_j},$$

czyli $\varphi(r_1, r_2, \ldots, r_{m-1})$ osiąga maksimum tylko w punkcie $(p_1, p_2, \ldots, p_{m+1})$ spełniającym warunek (3.8) dla ustalonych $r_1, r_2, \ldots, r_{m-1}$. Zastępując we wzorze na funkcję $\varphi(r_1, r_2, \ldots, r_{m-1})$ wartości dla $p_1, p_2, \ldots, p_m, p_{m+1}$ dane w (3.11), otrzymujemy, że $\varphi(r_1, r_2, \ldots, r_{m-1}) = 1$. Tak więc, gdy $(p_1, p_2, \ldots, p_{m+1})$ nie spełnia warunku (3.8), nie istnieje $(r_1, r_2, \ldots, r_m)$ takie, że jednocześnie $\sum_{j=1}^{m} r_i = 1$ i $\varphi(r_1, r_2, \ldots, r_{m-1}) = 1$. Ponieważ jednak $\varphi(r_1, r_2, \ldots, r_{m-1})$ jest funkcją ciągłą ze względu na $(r_1, r_2, \ldots, r_{m-1})$ przy ustalonym $(p_1, p_2, \ldots, p_{m+1})$, określoną na zbiorze zwartym, więc stąd wynika, że (3.7) zachodzi, gdy $(p_1, p_2, \ldots, p_{m+1})$ nie spełnia warunku (3.8).

Teraz możemy łatwo udowodnić następujące twierdzenie:

TWIERDZENIE 4. *Następujące dwa warunki są równoważne*:

(i) *Plan, którego zbiór brzegowy jest dany wzorem* (3.1), *jest domknięty. Można różniczkować* (2.1) *wyraz po wyrazie względem* p_j, $j = 1, 2, \ldots, n-1$. *Istnieje pierwszy moment chwili* τ *pierwszego osiągnięcia zbioru danego wzorem* (3.1).

(ii) $a_1p_1+a_2p_2+\ldots+a_np_n > 0$.

D o w ó d. (ii) $\Rightarrow$ (i). To wynika z poprzednich rozważań i twierdzenia 3. Różniczkując (2.1) otrzymujemy, że

$$E(\tau) = \frac{b}{a_1p_1+a_2p_2+\ldots+a_np_n}. \tag{3.12}$$

(i) $\Rightarrow$ (ii). Wykażemy, że $\neg$(ii) $\Rightarrow$ $\neg$(i). Niech $a_1p_1+a_2p_2+\ldots+a_np_n < 0$. Wówczas plan nie jest domknięty. Niech $a_1p_1+a_2p_2+\ldots+a_np_n = 0$ i przypuśćmy, że możemy różniczkować (2.1) wyraz po wyrazie względem p_j. Wtedy otrzymujemy, że $E(\tau) = \infty$. To kończy dowód twierdzenia.

WNIOSEK 3.1. *Warunkiem koniecznym dla efektywności planu, którego zbiór brzegowy jest dany wzorem* (3.1), *w punkcie* $(p_1, p_2, \ldots, p_n)$ *jest warunek* (ii) *twierdzenia* 4.

Dowód wynika natychmiast z twierdzenia 4.

4. Efektywność planów ukośnych. Niech $\mathscr{P}$ oznacza zbiór tych p, dla których zachodzi nierówność

$$a_1 p_1 + a_2 p_2 + \ldots + a_n p_n > 0,$$

gdzie $a_1, a_2, \ldots, a_n$ są takie jak we wzorze (3.1).

TWIERDZENIE 5. *Plan, którego zbiór brzegowy jest dany wzorem* (3.1), *jest efektywny w* $\mathscr{P}$. *Następujące funkcje, i tylko one, są efektywnie estymowalne*:

$$g(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{n} \frac{\lambda_i p_i b}{a_1 p_1 + a_2 p_2 + \ldots + a_n p_n} + c, \tag{4.1}$$

gdzie $\lambda_1, \lambda_2, \ldots, \lambda_n$, c *dowolne stałe. Ich jedyne estymatory efektywne są dane wzorem*

$$f = \sum_{i=1}^{n} \lambda_i X_i + c, \tag{4.2}$$

gdzie $(X_1, X_2, \ldots, X_n)$ *zostało zdefiniowane we wstępie, gdy* τ *jest chwilą pierwszego osiągnięcia zbioru danego wzorem* (3.1).

D o w ó d. Niech $f = \sum_{i=1}^{n} \lambda_i X_i + c$. Aby dowieść, że f jest efektywny w punkcie $(p_1, p_2, \ldots, p_n)$, wystarczy wykazać, że każda X_i jest liniową funkcją zmiennej U_j, $j = 1, 2, \ldots, n-1$. Oczywiście, $(X_1, X_2, \ldots, X_n)$ generuje miarę w N^n, gdzie N oznacza zbiór liczb całkowitych nieujemnych, skupioną na hiperpłaszczyźnie danej wzorem (3.1). Ponieważ każdy punkt tej hiperpłaszczyzny ma miarę dodatnią, więc pomiędzy $X_1, X_2, \ldots, X_n$ zachodzi jedynie związek

$$a_1 X_1 + a_2 X_2 + \ldots + a_n X_n = b \quad \text{p.w.} \tag{4.3}$$

Z (4.3) wyrażamy X_n (zakładając, że $a_n \neq 0$; można to założyć, bo indeksacja jest dowolna, a co najmniej jedno $a_i \neq 0$):

$$X_n = \frac{1}{a_n} [b - a_1 X_1 - a_2 X_2 - \ldots - a_{n-1} X_{n-1}]. \tag{4.4}$$

Z (2.2) i (4.4) otrzymujemy

$$U_j = \frac{X_j}{p_j} - \frac{1}{p_n a_n} [b - a_1 X_1 - a_2 X_2 - \ldots - a_{n-1} X_{n-1}], \quad j = 1, 2, \ldots, n-1,$$

czyli mnożąc przez $p_j p_n a_n$ otrzymujemy, że

$$a_n p_j p_n U_j + b p_j = p_j \sum_{\substack{i=1 \\ i \neq j}}^{n-1} a_i X_i + (a_j p_j + a_n p_n) X_j, \quad j = 1, 2, \ldots, n-1. \tag{4.5}$$

Układ (4.5) ma rozwiązanie, gdy wyznacznik macierzy

$$\text{(4.6)} \qquad \begin{bmatrix} a_1p_1+a_np_n & a_2p_1 & \dots & & a_{n-1}p_1 \\ a_1p_2 & a_2p_2+a_np_n & a_3p_2 & \dots & a_{n-1}p_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1p_{n-1} & a_2p_{n-1} & \dots & a_{n-2}p_{n-1} & a_{n-1}p_{n-1}+a_np_n \end{bmatrix}$$

jest różny od zera. Wyznacznik ten jest równy $p_1p_2 \dots p_{n-1}$ razy wyznacznik macierzy

$$\text{(4.7)} \qquad \begin{bmatrix} a_1+\dfrac{a_np_n}{p_1} & a_2 & \dots & & a_{n-1} \\ a_1 & a_2+\dfrac{a_np_n}{p_2} & a_3 & \dots & a_{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1 & a_2 & \dots & a_{n-2} & a_{n-1}+\dfrac{a_np_n}{p_{n-1}} \end{bmatrix}.$$

Wyznacznik macierzy (4.7) już można łatwo znaleźć za pomocą prostych przekształceń. W końcu otrzymujemy, że wyznacznik macierzy (4.6) jest równy $(a_np_n)^{n-2}(a_1p_1+ \dots +a_np_n)$, czyli w naszym przypadku wyznacznik ten jest różny od zera. Układ (4.5) ma więc rozwiązanie, a zatem każda funkcja X_i, $i = 1, 2, \dots, n-1$, jest liniową funkcją zmiennej U_j, $j = 1, 2, \dots, n-1$. Zatem, estymator f dany wzorem (4.2) jest efektywny w $\mathscr{P}$, a na mocy (2.3) i (3.12) jego funkcja efektywnie estymowalna jest dana wzorem (4.1).

Fakt, że estymatory (4.2) i funkcje (4.1) są odpowiednio jedynymi estymatorami i funkcjami efektywnymi, wynika z pracy [4].

To kończy dowód twierdzenia.

Dotychczas rozpatrywaliśmy tylko plany pierwszego osiągnięcia. Zauważmy, że wszystko, co powiedzieliśmy w ustępie 2, byłoby aktualne dla innych planów, gdybyśmy założyli, że τ jest dowolną chwilą markowską.

TWIERDZENIE 6. *Jedynymi planami efektywnymi w co najmniej dwóch punktach są plany pierwszego osiągnięcia.*

Dowód. Przypuśćmy, że plan z chwilą markowską τ jest efektywny w co najmniej dwóch punktach. Wtedy z prawdopodobieństwem 1 zachodzi (patrz [1]) równość:

$$\text{(4.8)} \qquad a_1X_1+a_2X_2+ \dots +a_nX_n = b,$$

gdzie $a_1, a_2, \dots, a_n, b$ oznaczają pewne stałe. Z (2.3) wynika, że

$$\text{(4.9)} \qquad E(\tau) = \frac{b}{a_1p_1+a_2p_2+\dots+a_np_n}.$$

Niech τ_0 będzie momentem pierwszego osiągnięcia hiperpłaszczyzny $a_1x_1+a_2x_2+ \ldots +a_nx_n = b$, gdzie $a_1, a_2, \ldots, a_n, b$ z (4.8). Oczywiście, zachodzi następująca nierówność:

$$\tau_0 \leqslant \tau \quad \text{p.w.} \tag{4.10}$$

Tak samo, z (2.3), otrzymujemy, że

$$E(\tau_0) = \frac{b}{a_1p_1+a_2p_2+\ldots+a_np_n}. \tag{4.11}$$

Z (4.9), (4.10) i (4.11) wynika, że $\tau = \tau_0$ p.w.

To kończy dowód twierdzenia.

Prace cytowane

[1] B. R. Bhat, N. V. Kulharni, *On efficient multinomial estimation*, J. Roy. Statist. Soc., Ser. B 28, 1965.

[2] А. М. Каган, Ю. В. Линник, С. Р. Рао, *Характеризационные задачи математической статистики*, Москва 1972.

[3] R. Magiera, S. Trybuła, *Plany ukośne dla procesu dwumianowego*, Matematyka Stosowana 6 (1976), str. 41–47.

[4] Д. Д. Скрябин, *Об эффективном последовательном оценивании*, Вестник Ленинградского Университета 7, 1973.