ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria III: MATEMATYKA STOSOWANA XVI (1980)

MAGDALENA RUTKOWSKA (Wrocław)

O minimaksowej prognozie dystrybuanty empirycznej*

(*Praca przyjęta do druku 10.9.1977*)

Niech $X = (X_1, X_2, \ldots, X_m)$ i $Y = (Y_1, Y_2, \ldots, Y_n)$ będą niezależnymi próbami prostymi z nieznanej dystrybuanty F. Załóżmy, że tylko wartości X są nam znane. Zajmiemy się problemem minimaksowej prognozy dystrybuanty empirycznej w drugiej próbie na podstawie wartości z pierwszej próby. Niech $\hat{F}(t)$ będzie dystrybuantą empiryczną w drugiej próbie

$$\hat{F}(t) = \frac{1}{n} \sum_{i=1}^{n} \delta_{Y_i}(t), \tag{1}$$

gdzie

$$\delta_{Y_i}(t) = \begin{cases} 1, & \text{gdy} \quad Y_i \leqslant t, \\ 0, & \text{gdy} \quad Y_i > t. \end{cases} \tag{2}$$

Oznaczmy przez $\phi(t)$ predyktor $\hat{F}(t)$. Jako funkcję straty związaną z $\phi(t)$ przyjmijmy wyrażenie postaci

$$L(\hat{F}, \phi) = \int_{-\infty}^{\infty} [\hat{F}(t) - \phi(t)]^2 [F(t)]^{\gamma-1} [1 - F(t)]^{\delta-1} dW(t), \tag{3}$$

gdzie W jest niezerową, skończoną miarą na ciele $\mathscr{B}$ zbiorów borelowskich na prostej R. Ryzyko związane z predyktorem ϕ przyjmie wówczas postać

$$R(F, \phi) = E_F\{L(\hat{F}, \phi)\} = \int_{-\infty}^{\infty} E_F\{[\hat{F}(t) - \phi(t)]^2 [F(t)]^{\gamma-1} [1 - F(t)]^{\delta-1}\} dW(t), \tag{4}$$

gdzie $E_F(\cdot)$ oznacza wartość oczekiwaną ze względu na dystrybuantę F. Przez L_1, L_2, L_3, L_4 oznaczmy cztery szczególne postacie funkcji straty L otrzymane przez podstawienie $\gamma = 0$ lub 1 oraz $\delta = 0$ lub 1 w (3):

$$L_1(\hat{F}, \phi) = \int_{-\infty}^{\infty} \{[\hat{F}(t) - \phi(t)]^2 / F(t) [1 - F(t)]\} dW(t) \qquad (\gamma = \delta = 0),$$

* Praca wykonana w ramach problemu 010.15 „Badania ze statystyki i teorii gier”.

$$
(5)\quad
\begin{aligned}
L_2(\hat{F},\phi) &= \int_{-\infty}^{\infty} \{[\hat{F}(t)-\phi(t)]^2/F(t)\}dW(t) && (\gamma = 0,\ \delta = 1),\\
L_3(\hat{F},\phi) &= \int_{-\infty}^{\infty} \{[\hat{F}(t)-\phi(t)]^2/[1-F(t)]\}dW(t) && (\gamma = 1,\ \delta = 0),\\
L_4(\hat{F},\phi) &= \int_{-\infty}^{\infty} \{[\hat{F}(t)-\phi(t)]^2\}dW(t) && (\gamma = \delta = 1).
\end{aligned}
$$

Będziemy szukali minimaksowego predyktora $\phi(t)$ dystrybuanty empirycznej $\hat{F}(t)$ dla funkcji straty L_1, L_2, L_3. Problem ten dla funkcji straty L_4 rozwiązał S. Trybuła [4]. Udowodnimy następujące twierdzenie.

TWIERDZENIE 1. *Niech Φ oznacza klasę predyktorów*

$$
(6)\quad \Phi = \Big\{\phi(t):\ \phi(t) = a + \sum_{i=1}^{m} b_i \delta_{X_i}(t),
$$

$$
a \geqslant 0,\ b_i \geqslant 0,\ i = 1, \ldots, m,\ a + \sum_{i=1}^{m} b_i \leqslant 1\Big\}.
$$

Dla funkcji straty L_i $(i = 1, 2, 3)$, określonej wzorami (5), ryzyko $R_i(F, \phi) = E_F\{L_i(\hat{F}, \phi)\}$, $\phi \in \Phi$, jest niezależne od F wtedy i tylko wtedy, gdy stałe $a, b_1, b_2, \ldots, b_m$ spełniają warunek W$_i$, *gdzie*

$$
(7)\quad \mathrm{W}_1:\ a = 0, \quad \sum_{i=1}^{m} b_i = 1,
$$

$$
(8)\quad \mathrm{W}_2:\ a = 0, \quad \Big(1 - \sum_{i=1}^{m} b_i\Big)^2 - \sum_{i=1}^{m} b_i^2 = \frac{1}{n},
$$

$$
(9)\quad \mathrm{W}_3:\ a = 1 - \sum_{i=1}^{m} b_i, \quad \Big(1 - \sum_{i=1}^{m} b_i\Big)^2 - \sum_{i=1}^{m} b_i^2 = \frac{1}{n}.
$$

D o w ó d. Ponieważ

$$
E_F[\delta_{X_i}(t)] = E_F[\delta_{Y_i}(t)] = F(t),
$$
$$
E_F\{[\delta_{X_i}(t)-F(t)]^2\} = E_F\{[\delta_{Y_i}(t)-F(t)]^2\} = F(t)\,[1-F(t)],
$$

więc ryzyko $R(F, \phi)$, dla $\phi \in \Phi$ przyjmie postać

$$
R(F,\phi) = \int_{-\infty}^{\infty} E_F\Big\{\Big[\frac{1}{n}\sum_{i=1}^{n} \delta_{Y_i}(t) - a - \sum_{i=1}^{m} b_i \delta_{X_i}(t)\Big]^2 [F(t)]^{\gamma-1}\,[1-F(t)]^{\delta-1}\Big\} dW(t) =
$$

$$
= \int_{-\infty}^{\infty} \Big\{\frac{1}{n}F(t)\,[1-F(t)] + F(t)\,[1-F(t)] \sum_{i=1}^{m} b_i^2 + a^2 + \Big(1 - \sum_{i=1}^{m} b_i\Big) F^2(t) -
$$

$$-2a\left(1-\sum_{i=1}^{m} b_i\right)F(t)\Big\}[F(t)]^{\gamma-1}[1-F(t)]^{\delta-1}dW(t).$$

Wyznaczmy $R(F,\phi)$ dla funkcji strat określonych wzorami (5). Otrzymamy wtedy:

$$(10)\qquad R_1(F,\phi) = E_F[L_1(\hat{F},\phi)] = \int_{-\infty}^{\infty}\Big\{\frac{1}{n}+\sum_{i=1}^{m} b_i^2+ \\ +\left[a^2+\left(1-\sum_{i=1}^{m} b_i\right)F^2(t)-2a\left(1-\sum_{i=1}^{m} b_i\right)F(t)\right]/F(t)\,[1-F(t)]\Big\}\,dW(t).$$

$$(11)\qquad R_2(F,\phi) = E_F[L_2(\hat{F},\phi)] = \int_{-\infty}^{\infty}\Big\{\frac{1}{n}+\sum_{i=1}^{m} b_i^2-2a\left(1-\sum_{i=1}^{m} b_i\right)+ \\ +\left[\left(1-\sum_{i=1}^{m} b_i\right)^2-\sum_{i=1}^{m} b_i^2-\frac{1}{n}\right]F(t)+a^2F^{-1}(t)\Big\}\,dW(t).$$

$$(12)\qquad R_3(F,\phi) = E_F[L_3(\hat{F},\phi)] = \int_{-\infty}^{\infty}\Big\{\left[\frac{1}{n}+\sum_{i=1}^{m} b_i^2-\left(1-\sum_{i=1}^{m} b_i\right)^2\right]F(t)+ \\ +\left(1-\sum_{i=1}^{m} b_i\right)\left(2a-1+\sum_{i=1}^{m} b_i\right)+\left[a-\left(1-\sum_{i=1}^{m} b_i\right)\right]^2/[1-F(t)]\Big\}\,dW(t).$$

Jeżeli stałe $a, b_1, \ldots, b_m$ spełniają warunek W_i, to ryzyko $R_i(F,\phi)$, $i = 1, 2, 3$, jest stałe, niezależne od F, co kończy dowód twierdzenia.

Przyjmijmy teraz, że

$$(13)\qquad b_1 = b_2 = \ldots = b_m = b.$$

Warunki W_i określone wzorami (7), (8), (9) przybiorą teraz następującą postać

$$(7')\qquad W_1:\quad a = 0,\quad mb = 1,$$

$$(8')\qquad W_2:\quad a = 0,\quad (1-mb)^2-mb^2 = 1/n,$$

$$(9')\qquad W_3:\quad a = 1-mb,\quad (1-mb)^2-mb^2 = 1/n.$$

Z powyższych równań otrzymamy następujące stałe, które równocześnie spełniają warunki zawarte w wyrażeniu (6)

$$(14)\qquad a = 0,\quad b = \frac{1}{m};$$

$$(15)\qquad a = 0,\quad b = \begin{cases} \dfrac{1-\sqrt{\dfrac{1}{m}+\dfrac{1}{n}-\dfrac{1}{mn}}}{m-1}, & \text{gdy}\quad m > 1, \\[2ex] \dfrac{n-1}{2n}, & \text{gdy}\quad m = 1; \end{cases}$$

$$
(16)\qquad
a = \begin{cases} \dfrac{m\sqrt{\dfrac{1}{m}+\dfrac{1}{n}-\dfrac{1}{mn}}-1}{m-1}, & \text{gdy} \quad m > 1, \\[2ex] \dfrac{n+1}{2n}, & \text{gdy} \quad m = 1; \end{cases}
\qquad
b = \begin{cases} \dfrac{1-\sqrt{\dfrac{1}{m}+\dfrac{1}{n}-\dfrac{1}{mn}}}{m-1}, & \text{gdy} \quad m > 1, \\[2ex] \dfrac{n-1}{2n}, & \text{gdy} \quad m = 1. \end{cases}
$$

Oznaczmy przez $\phi_1^0(t)$, $\phi_2^0(t)$, $\phi_3^0(t)$ predyktory liniowe postaci $a+b\sum_{i=1}^{m}\delta_{X_i}(t)$, gdzie stałe a, b są określone odpowiednio wzorami (14), (15), (16). Dla tak określonych predyktorów funkcje ryzyka przybierają następujące wartości

$$
(17)\qquad R_1(F,\phi_1^0) = \left(\frac{1}{n}+\frac{1}{m}\right)\int_{-\infty}^{\infty} dW(t);
$$

$$
(18)\qquad R_2(F,\phi_2^0) = \begin{cases} \left[\dfrac{m\sqrt{\dfrac{1}{m}+\dfrac{1}{n}-\dfrac{1}{mn}}-1}{m-1}\right]^2 \displaystyle\int_{-\infty}^{\infty} dW(t), & \text{dla} \quad m > 1, \\[2ex] \dfrac{1}{4}\left(1+\dfrac{1}{n}\right)^2 \displaystyle\int_{-\infty}^{\infty} dW(t), & \text{dla} \quad m = 1; \end{cases}
$$

$$
(19)\qquad R_3(F,\phi_3^0) = \begin{cases} \left[\dfrac{m\sqrt{\dfrac{1}{m}+\dfrac{1}{n}-\dfrac{1}{mn}}-1}{m-1}\right]^2 \displaystyle\int_{-\infty}^{\infty} dW(t), & \text{dla} \quad m > 1, \\[2ex] \dfrac{1}{4}\left(1+\dfrac{1}{n}\right)^2 \displaystyle\int_{-\infty}^{\infty} dW(t), & \text{dla} \quad m = 1. \end{cases}
$$

Niech θ będzie zbiorem wszystkich dystrybuant na R. Ze względu na podane niżej twierdzenia 2 i 3 potrzebny nam będzie ciąg rozkładów prawdopodobieństwa na θ, które w stosowanych w pracy metodach noszą nazwę rozkładów *a priori nieznanej dystrybuanty*. Przyjmiemy mianowicie, że rozkład τ_k jest skupiony na zbiorze

$$
\theta_k = \{F_{k,p};\ 0 < p < 1\}, \quad \text{gdy} \quad k = 1, 2, \ldots
$$

gdzie

$$
(20)\qquad F_{k,p}(t) = \begin{cases} 0, & t < -k, \\ p, & -k \leqslant t < k, \\ 1, & t \geqslant k, \end{cases}
$$

przy czym p ma rozkład beta o gęstości

$$(21) \qquad g(p) = \frac{1}{B(\alpha,\beta)} p^{\alpha-1}(1-p)^{\beta-1}, \quad 0 < p < 1,\ \alpha > 0,\ \beta > 0.$$

Dla tak określonego rozkładu a priori τ_k oczekiwane ryzyko przyjmie postać

$$(22) \qquad r(\tau_k, \phi) \overset{\mathrm{df}}{=} E_{\tau_k}[R(F,\phi)] = E_{\tau_k}\{E_F[L(\hat{F},\phi)]\} =$$
$$= \int_{-\infty}^{\infty} E_{\tau_k}\{E_F[\hat{F}(t)-\phi(t)]\,[F(t)]^{\gamma-1}[1-F(t)]^{\delta-1}\}\,dW(t).$$

Zajmiemy się teraz obliczeniem minimum funkcjonału $r(\tau_k, \phi)$. Oczekiwane ryzyko $r(\tau_k, \phi)$ osiągnie minimum, jeżeli dla każdego t wybierzemy $\phi(t)$ tak, że

$$E_{\tau_k}\{E_F[\hat{F}(t)-\phi(t)]^2\,[F(t)]^{\gamma-1}[1-F(t)]^{\delta-1}\}$$

osiągnie minimum. Ponieważ

$$E_F[\hat{F}(t)-\phi(t)]^2 = E_F[\hat{F}(t)-F(t)]^2 + E_F[\phi(t)-F(t)]^2,$$

zajmiemy się poszukiwaniem minimum wyrażenia

$$(23) \qquad \varrho_t(\tau_k, \phi) = E_{\tau_k}\{E_F[\phi(t)-F(t)]^2\,[F(t)]^{\gamma-1}[1-F(t)]^{\delta-1}\}.$$

Aby wyznaczyć min $\varrho_t(\tau_k, \phi)$ wykorzystamy twierdzenie udowodnione przez F. G. Phadia [3], który rozpatrywał problem minimaksowej estymacji dystrybuanty $F(t)$.

TWIERDZENIE 2 (Phadia). *Niech $X_1, X_2, \ldots, X_m$ będzie próbą losową z nieznanej dystrybuanty F i niech τ_k będzie rozkładem a priori określonym na zbiorze dystrybuant θ i zdefiniowanym powyżej wzorami* (20), (21). *Załóżmy również, że $\alpha+\gamma > 1$, $\beta+\delta > 1$. Wówczas*

$$(24) \qquad \min_{\phi} \varrho_t(\tau_k, \phi) = \frac{1_{[-k,k]}(t)\,B(\alpha+\gamma, \beta+\delta)}{(\alpha+\beta+\gamma+\delta+m-2)\,B(\alpha,\beta)},$$

gdzie $1_A(t)$ oznacza funkcję charakterystyczną zbioru A.

Dla rozkładu τ_k określonego powyżej można łatwo wyznaczyć wartość wyrażenia

$$E_{\tau_k}\{E_F[\hat{F}(t)-F(t)]^2\,[F(t)]^{\gamma-1}[1-F(t)]^{\delta-1}\} = \frac{B(\alpha+\gamma, \beta+\delta)}{nB(\alpha,\beta)} \int_{-\infty}^{\infty} 1_{[-k,k]}(t)\,dW(t).$$

Otrzymamy zatem

$$(25) \qquad \min_{\phi} r(\tau_k, \phi) = \frac{B(\alpha+\gamma, \beta+\delta)}{B(\alpha,\beta)}\left[\frac{1}{\alpha+\beta+\gamma+\delta+m-2} + \frac{1}{n}\right] \int_{-\infty}^{\infty} 1_{[-k,k]}(t)\,dW(t).$$

W dalszej części pracy wykorzystamy twierdzenie analogiczne do twierdzenia 2, § 2.11, [1].

TWIERDZENIE 3. *Niech $\phi_k(t)$ będzie ciągiem bayesowskich predyktorów $\hat{F}(t)$ dla rozkładów a priori τ_k i niech $r(\tau_k, \phi_k)$ będzie odpowiadającym ciągiem ryzyk baye-*

sowskich. Jeżeli $r(\tau_k, \phi_k) \to C$ (const), gdy $k \to \infty$ i ϕ^0 jest dowolnym predyktorem, takim że $R(F, \phi^0) \leqslant C$ dla każdego $F \in \theta$, to ϕ^0 jest predyktorem minimaksowym.

Udowodnimy teraz następujące twierdzenie.

TWIERDZENIE 4. *Niech $X = (X_1, X_2, \ldots, X_m)$ i $Y = (Y_1, Y_2, \ldots, Y_n)$ będą niezależnymi próbami losowymi z nieznanej dystrybuanty F na R i niech $\hat{F}(t)$ będzie dystrybuantą empiryczną w drugiej próbie. Wówczas dla funkcji strat L_1, L_2, L_3 określonych wzorami (5), predyktory minimaksowe dystrybuanty empirycznej dane są odpowiednio wzorami*

$$\phi_1^0(t) = \frac{1}{m}\sum_{i=1}^{m}\delta_{X_i}(t),$$

$$\phi_2^0(t) = \begin{cases} \dfrac{1-\sqrt{\dfrac{1}{m}+\dfrac{1}{n}-\dfrac{1}{mn}}}{m-1}\displaystyle\sum_{i=1}^{m}\delta_{X_i}(t), & \text{gdy } m > 1, \\[2ex] \dfrac{n-1}{2n}\,\delta_X(t), & \text{gdy } m = 1, \end{cases}$$

$$\phi_3^0(t) = \begin{cases} \dfrac{m\sqrt{\dfrac{1}{m}+\dfrac{1}{n}-\dfrac{1}{mn}}-1+\left(1-\sqrt{\dfrac{1}{m}+\dfrac{1}{n}-\dfrac{1}{mn}}\right)\displaystyle\sum_{i=1}^{m}\delta_{X_i}(t)}{m-1}, & \text{gdy } m > 1, \\[2ex] \dfrac{n+1+(n-1)\,\delta_X(t)}{2n}, & \text{gdy } m = 1. \end{cases}$$

D o w ó d. Niech τ_k będzie rozkładem a priori określonym na zbiorze dystrybuant θ i zdefiniowanym powyżej wzorami (20), (21). Oznaczmy przez $r_i(\tau_k, \phi)$ ryzyko bayesowskie dla funkcji straty L_i, $i = 1, 2, 3$.

Aby udowodnić twierdzenie, wystarczy wykazać, że

$$\lim_{k\to\infty}\min_{\phi} r_i(\tau_k, \phi) = R_i(F, \phi_i^0), \qquad i = 1, 2, 3, \tag{26}$$

gdzie $R_i(F, \phi_i^0)$ określone są wzorami (17), (18), (19). Dla funkcji straty L_1 podstawmy dla rozkładu τ_k

$$\alpha = \beta = 1. \tag{27}$$

Otrzymamy wówczas z (25)

$$\min_{\phi} r_1(\tau_k, \phi) = \left(\frac{1}{\alpha+\beta+m-2}+\frac{1}{n}\right)\int_{-\infty}^{\infty} 1_{[-k,k]}(t)\,dW(t) =$$

$$= \left(\frac{1}{m}+\frac{1}{n}\right)\int_{-\infty}^{\infty} 1_{[-k,k]}(t)\,dW(t),$$

$$\lim_{k\to\infty}\min_{\phi} r_1(\tau_k, \phi) = \left(\frac{1}{m}+\frac{1}{n}\right)\int_{-\infty}^{\infty} dW(t) \overset{(17)}{=} R_1(F, \phi_1^0).$$

W przypadku funkcji straty L_2

$$(28)\qquad \min_{\phi} r_2(\tau_k, \phi) = \frac{B(\alpha+1, \beta)}{B(\alpha, \beta)}\left[\frac{1}{\alpha+\beta+m-1}+\frac{1}{n}\right]\int_{-\infty}^{\infty} 1_{[-k,k]}(t)dW(t) =$$

$$= \frac{\beta(\alpha+\beta+m+n-1)}{n(\alpha+\beta)(\alpha+\beta+m-1)}\int_{-\infty}^{\infty} 1_{[-k,k]}(t)dW(t).$$

Załóżmy teraz, że $n > 1$ i podstawmy

$$(29)\qquad \alpha = 1, \quad \beta = \frac{m+\sqrt{mnx}}{n-1}, \quad \text{gdzie} \quad x = m+n-1.$$

Współczynnik występujący przed całką we wzorze (28) przyjmuje wówczas postać

$$(30)\qquad \frac{\beta(\alpha+\beta+m+n-1)}{n(\alpha+\beta)(\alpha+\beta+m-1)} = \frac{(m+\sqrt{mnx})(nx+\sqrt{mnx})}{n(mn+\sqrt{mnx})(x+\sqrt{mnx})} =$$

$$= \left(\frac{m+\sqrt{mnx}}{mn+\sqrt{mnx}}\right)^2 = \begin{cases} \left[\dfrac{\sqrt{mnx}-n}{n(m-1)}\right]^2, & \text{dla } m > 1, \\ \left(\dfrac{n+1}{2n}\right)^2, & \text{dla } m = 1. \end{cases}$$

Zatem

$$\lim_{k\to\infty}\min_{\phi} r_2(\tau_k, \phi) = \begin{cases} \left[\dfrac{\sqrt{mnx}-n}{n(m-1)}\right]^2 \displaystyle\int_{-\infty}^{\infty} dW(t), & \text{dla} \quad m > 1, \\ \left(\dfrac{n+1}{2n}\right)^2 \displaystyle\int_{-\infty}^{\infty} dW(t), & \text{dla} \quad m = 1. \end{cases}$$

Porównując otrzymane wyniki z wyrażeniem określonym wzorem (18) widzimy, że dla $n > 1$, dla funkcji straty L_2 warunek (26) jest spełniony. Rozważmy teraz przypadek $n = 1$. Predyktor $\phi_2^0(t)$ nie zależy wówczas od X. Rozkład a priori τ_k zdefiniujemy następująco. Z prawdopodobieństwem 1 wybieramy dystrybuantę

$$(31)\qquad F(t) = \begin{cases} 0, & \text{gdy} \quad t < -k, \\ 1/k, & \text{gdy} \quad -k \leqslant t < k, \\ 1, & \text{gdy} \quad t \geqslant k. \end{cases}$$

W tym przypadku predyktor bayesowski $\bar{\phi}(t) = F(t)$, a bayesowskie ryzyko wyrazi się wzorem

$$r_2(\tau_k, \bar{\phi}) = \int_{-\infty}^{\infty} E_F[\hat{F}(t)-F(t)]^2\,[F(t)]^{-1}dW(t) = \left(1-\frac{1}{k}\right)\int_{-\infty}^{\infty} 1_{[-k,k]}(t)dW(t)$$

i

$$\lim_{k\to\infty} r_2(\tau_k, \bar{\phi}) = \int_{-\infty}^{\infty} dW(t) = R_2(F, \phi_2^0).$$

Natomiast dla funkcji straty L_3

$$\min_{\phi} r_3(\tau_k, \phi) = \frac{\alpha(\alpha+\beta+m+n-1)}{n(\alpha+\beta)(\alpha+\beta+m-1)} \int_{-\infty}^{\infty} 1_{[-k,k]}(t)\,dW(t).$$

Dla $n > 1$ podstawienie

$$\beta = 1, \quad \alpha = \frac{m+\sqrt{mnx}}{n-1}, \quad x = m+n-1 \tag{32}$$

daje

$$\min_{\phi} r_3(\tau_k, \phi) = \frac{(m+\sqrt{mnx})(nx+\sqrt{mnx})}{n(mn+\sqrt{mnx})(x+\sqrt{mnx})} \int_{-\infty}^{\infty} 1_{[-k,k]}(t)\,dW(t) =$$

$$= \begin{cases} \left[\dfrac{\sqrt{mnx}-n}{n(m-1)}\right]^2 \displaystyle\int_{-\infty}^{\infty} 1_{[-k,k]}(t)\,dW(t), & \text{gdy } m > 1, \\ \left[\dfrac{n+1}{2n}\right]^2 \displaystyle\int_{-\infty}^{\infty} 1_{[-k,k]}(t)\,dW(t), & \text{gdy } m = 1. \end{cases}$$

Zatem w przypadku funkcji straty L_3, dla $n > 1$ spełniona jest równość (26).

Niech $n = 1$. Predyktor $\phi_3^0(t)$, podobnie jak $\phi_2^0(t)$, nie zależy od X. Zdefiniujemy teraz rozkład a priori τ_k następująco. Z prawdopodobieństwem jeden wybieramy dystrybuantę

$$F(t) = \begin{cases} 0, & \text{gdy} \quad t < -k, \\ 1-1/k, & \text{gdy} \quad -k \leqslant t < k, \\ 1, & \text{gdy} \quad t \geqslant k. \end{cases}$$

Wówczas, bayesowski predyktor $\bar{\phi}(t) = F(t)$ i bayesowskie ryzyko przyjmie postać

$$\min_{\phi} r_3(\tau_k, \phi) = \int_{-\infty}^{\infty} \{E_F[\hat{F}(t)-F(t)]^2/[1-F(t)]\}\,dW(t) =$$

$$= \left(1-\frac{1}{k}\right) \int_{-\infty}^{\infty} 1_{[-k,k]}(t)\,dW(t),$$

i

$$\lim_{k\to\infty} \min_{\phi} r_3(\tau_k, \phi) \overset{(19)}{=} R_3(F, \phi_3^0).$$

Wykazaliśmy więc, że również dla $n = 1$ zachodzi równość (26), co kończy dowód.

Jeżeli miara $W(t)$ skupiona jest w punkcie t_0, rozważane zagadnienie sprowadza się do problemu minimaksowej prognozy zmiennej losowej Y/n na podstawie obserwacji zmiennej losowej X/m, gdzie X i Y mają rozkłady dwumianowe z parametrami odpowiednio (m, p), (n, p), gdzie $p = F(t_0)$. Problem ten został rozwiązany przez Hodgesa i Lehmanna w pracy [2]. W przypadku funkcji straty L_1 predyktor $\phi_1^0(t)$

nie zależy od n i jest identyczny z minimaksowym estymatorem dystrybuanty $F(t)$ dla funkcji straty

$$L(F, \phi) = \int_{-\infty}^{\infty} \{[\phi(t) - F(t)]^2 / F(t)\,[1 - F(t)]\} dW(t).$$

Gdy $n \to \infty$

$$\phi_2^0(t) \to \frac{1}{m + \sqrt{m}} \sum_{i=1}^{m} \delta_{X_i}(t),$$

$$\phi_3^0(t) \to \frac{\sqrt{m} + \sum_{i=1}^{m} \delta_{X_i}(t)}{m + \sqrt{m}}$$

i predyktory te pokrywają się w granicy z minimaksowymi estymatorami dystrybuanty $F(t)$ dla funkcji strat odpowiednio

$$L_2(\hat{F}, \phi) = \int_{-\infty}^{\infty} \{[\phi(t) - F(t)]^2 / F(t)\} dW(t),$$

$$L_3(\hat{F}, \phi) = \int_{-\infty}^{\infty} \{[\phi(t) - F(t)]^2 / [1 - F(t)]\} dW(t).$$

Estymatory te znalazł E. S. Phadia [3].

Literatura

[1] T. S. Ferguson, *Mathematical statistic — a decision theoretic approach*, Academic Press, New York 1967.

[2] J. L. Hodges, E. L. Lehmann, *Some problems in minimax point estimation*, AMS 21 (1950), str. 182–197.

[3] E. G. Phadia, *Minimax estimation of a cumulative distribution function*, AS 1 (1973), str. 1149–1157.

[4] S. Trybuła, *O minimaksowej prognozie dystrybuanty empirycznej*, w druku.