ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria II: WIADOMOŚCI MATEMATYCZNE XVIII (1974)

wmptm 1897

Ministerstwo Nauki i Szkolnictwa Wyższego

ROBERT BARTOSZYŃSKI (Warszawa)

Reguły zatrzymywania

1. Wstęp. Celem niniejszego artykułu jest przedstawienie problematyki tzw. reguł zatrzymywania ([1]), stosunkowo nowego, bo liczącego zaledwie kilkanaście lat, działu teorii prawdopodobieństwa.

W największym uproszczeniu, podstawowym problemem jest tu analiza korzyści, jakie może osiągnąć gracz biorący udział w ciągu gier, jeżeli ma on przywilej przerwania gry w dowolnym dogodnym dla siebie momencie.

Nieco bardziej ogólnie, problematykę reguł zatrzymywania można opisać następująco. Obserwujemy pewien proces stochastyczny (dla uproszczenia, w niniejszym artykule omawiane będą jedynie procesy z czasem dyskretnym; obserwowany proces ma więc postać ciągu zmiennych losowych $X_1, X_2, \ldots$). Po dokonaniu kolejnej obserwacji możemy podjąć jedną z dwóch decyzji: „zatrzymać" proces, otrzymując wypłatę zależną od wyników dotychczasowych obserwacji, lub też kontynuować obserwacje dalej. Decyzja o zatrzymaniu lub kontynuacji musi być przy tym oparta tylko na wynikach obserwacji dokonanych do danego momentu czasu i nie może zależeć od przyszłych wartości zmiennych losowych procesu.

Reguła zatrzymywania jest pewną instrukcją, która dla każdego skończonego ciągu wyników obserwacji $x_1, \ldots, x_n$ wartości zmiennych losowych $X_1, \ldots, X_n$ orzeka jednoznacznie, czy proces ma być zatrzymany w chwili n, czy nie; musi być przy tym spełniony warunek, aby reguła ta prowadziła prędzej czy później do zatrzymania procesu.

Przy odpowiednich założeniach o całkowalności funkcji wypłaty, z każdą regułą zatrzymywania można związać oczekiwaną wypłatę. Problem polega na zbadaniu warunków istnienia optymalnej reguły zatrzymywania (tj. reguły o maksymalnej oczekiwanej wypłacie), oraz oczywiście na efektywnym wyznaczeniu takiej reguły, o ile warunki te są spełnione.

2. Przykłady. Zaczniemy od przedstawienia szeregu przykładów, zaczerpniętych głównie z [3] i [2].

([1]) Ang. *stopping rule* lub *stopping time*; franc. *temps d'arrêt*; ros. *случайное время*.

a. Niech $X_n = 1-1/n$ z prawdopodobieństwem 1 (czyli obserwujemy po prostu ciąg liczbowy $\{1-1/n\}$) i niech wypłata za zatrzymanie procesu w chwili n wynosi X_n. Przykład ten nie jest oczywiście specjalnie interesujący; pokazuje on jedynie, że optymalna reguła zatrzymywania może nie istnieć.

b. Jeżeli X_n jest wygraną gracza w n-tej rozgrywce, to $S_n = X_1+\ldots+X_n$ jest jego łączną wygraną, i zarazem wypłatą, jeżeli gracz decyduje się przerwać grę po n-tej rozgrywce.

c. Niech $\{X_n\}$ będą niezależne, i $P(X_n = 1) = P(X_n = 0) = 1/2$ (co odpowiada rzutom symetryczną monetą). Wówczas $S_n = X_1+\ldots+X_n$ jest liczbą orłów w pierwszych n rzutach. Jeżeli przerywamy proces po n-tym rzucie, otrzymujemy wypłatę S_n/n.

Rozważmy regułę; „zatrzymaj się po pierwszym rzucie, jeżeli jego wynikiem jest orzeł; w przeciwnym przypadku, zatrzymaj się po drugim rzucie". Wówczas oczekiwana wypłata równa się

$$\tfrac{1}{2}\cdot 1+\tfrac{1}{4}\cdot\tfrac{1}{2}+\tfrac{1}{4}\cdot 0 = \tfrac{5}{8}.$$

Inną regułą zatrzymywania może być: „zatrzymaj się po pierwszym rzucie, jeżeli jego wynikiem jest orzeł; w przypadku przeciwnym, zatrzymaj się, gdy po raz pierwszy liczba orłów zrówna się z liczbą reszek". Wiadomo (por. np. [5]), że reguła ta doprowadza z prawdopodobieństwem jeden do zatrzymania procesu (tj. z prawdopodobieństwem jeden prędzej czy później nastąpi zrównanie liczby orłów i reszek). Oczekiwana wypłata przy tej regule wynosi

$$\tfrac{1}{2}\cdot 1+\tfrac{1}{2}\cdot\tfrac{1}{2} = \tfrac{3}{4}.$$

Dokładna postać optymalnej reguły zatrzymywania dla tego problemu nie jest znana; wiadomo jedynie, że taka optymalna reguła istnieje (por. [4]).

d. Wyobraźmy sobie włamywacza, który każdej nocy włamuje się do jakiegoś mieszkania. Niech Y_n reprezentuje wartość łupu z n-tego rabunku, pod warunkiem, że włamywacz nie zostanie schwytany. Załóżmy, że Y_n są niezależnymi zmiennymi losowymi o jednakowym rozkładzie o dodatniej wartości oczekiwanej. Ponadto załóżmy, że prawdopodobieństwo, iż włamywacz zostanie schwytany (traci on wówczas cały dotychczasowy łup i zostaje zmuszony do zaprzestania działalności) nie zależy ani od tego, jak długo prowadził on dotąd swój proceder, ani od wartości poszczególnych łupów. Przyjmijmy $d_j = 0$ lub 1 w zależności od tego, czy włamywacz zostaje schwytany j-tej nocy, czy nie. Jego łączny zysk można wówczas wyrazić w postaci

$$d_1 \ldots d_n(Y_1+\ldots+Y_n) \tag{1}$$

i problem polega na wyborze reguły zatrzymywania maksymalizującej oczekiwaną wartość wypłaty (1).

e. Jedziemy samochodem, który chcemy zaparkować możliwie blisko miejsca docelowego. Kolejne stanowiska parkowania mogą być bądź zajęte (w tym przypadku musimy jechać dalej), bądź wolne; wówczas możemy się zatrzymać, lub też

zdecydować się jechać dalej. Do wolnego miejsca, które miniemy nie wolno nam się cofać; ponadto, możemy obserwować jedynie aktualnie mijane miejsce parkowania.

Oznaczmy $X_n = 1$ lub 0 w zależności od tego, czy n-te miejsce parkowania jest wolne czy zajęte i przyjmijmy, że X_n są niezależne o rozkładzie $P(X_n = 1) = p = 1 - P(X_n = 0)$. Niech punkt docelowy znajduje się w miejscu parkowania o numerze 0 i przypuśćmy, że proces zaczyna się od miejsca $-N$, tzn. że obserwujemy wartości $X_{-N}, X_{-N+1}, \ldots, X_{-1}, X_0, X_1, \ldots$ Jeżeli zaobserwowaliśmy wartości $X_{-N}, \ldots, X_n$, przy czym $X_n = 0$, to nie możemy się zatrzymać; jeżeli $X_n = 1$ i decydujemy się na zatrzymanie procesu, to strata równa jest $|n|$ (w tym przypadku wygodniej jest używać terminu „strata" i szukać reguły zatrzymywania minimalizującej oczekiwaną stratę).

W odróżnieniu od poprzednich przykładów, w rozważanym przypadku nie możemy się zatrzymać w dowolnym momencie. Aby sprowadzić tę sytuację do schematu poprzednich, wystarczy przyjąć, że zatrzymanie się możliwe jest zawsze i określić funkcję straty jako równą $+\infty$ w przypadku zatrzymania się przy wartości $X_n = 0$.

f. Na ogłoszenie o wolnym stanowisku sekretarki zgłosiło się N kandydatek. Z każdą z nich kolejno przeprowadza się wywiad oceniając jej przydatność, i natychmiast po skończeniu wywiadu kandydatkę można bądź przyjąć (wówczas proces się kończy), bądź też odrzucić i przeprowadzić wywiad z następną; nie wolno przy tym zmieniać decyzji w stosunku do odrzuconych kandydatek.

Dla opisania tej sytuacji w terminach reguł zatrzymywania, załóżmy że kandydatki pojawiają się w losowej kolejności, tak że wszystkie $N!$ permutacji mają jednakowe prawdopodobieństwa. Każdą z kandydatek można ocenić w stosunku do tych, z którymi przeprowadzono już wywiad, ale nie w stosunku do pozostałych i zadanie polega na znalezieniu reguły, która prowadziłaby do maksymalizacji prawdopodobieństwa wyboru najlepszej kandydatki.

Przypiszmy kandydatkom rangi $1, \ldots, N$, gdzie 1 oznacza kandydatkę najlepszą, a N najgorszą. Niech $Y_1, \ldots, Y_N$ będzie losową permutacją liczb $1, \ldots, N$. Permutacji tej nie możemy obserwować bezpośrednio (ponieważ daną kandydatkę możemy porównywać jedynie z poprzedzającymi ją, ale nie z następującymi po niej). Oznaczmy teraz

$$X_n = \text{liczba tych } Y_j \text{ wśród } Y_1, \ldots, Y_n, \text{ które są} \leqslant Y_n.$$

Tak więc, X_n jest (obserwowalną) rangą ostatniej zbadanej (n-tej) kandydatki w stosunku do zbadanych poprzednio.

Jeżeli zaobserwowaliśmy wartości zmiennych losowych $X_1, \ldots, X_n$ i decydujemy się zatrzymać proces angażując kandydatkę o numerze n, to zysk zależy od tego, czy $Y_n = 1$, czy też $Y_n \neq 1$. Jako funkcję wypłaty można w tym przypadku przyjąć $P(Y_n = 1 | X_1, \ldots, X_n)$ i zadanie polega na konstrukcji reguły zatrzymywania maksymalizującej oczekiwaną funkcję wypłaty, czyli maksymalizującą prawdopodobieństwo wyboru najlepszej kandydatki.

3. Sformułowanie problemu. Zaczniemy od sformułowania problemu reguły zatrzymywania w przypadku, gdy zmienne losowe X_1, X_2, ... obserwowanego procesu przyjmują wartości z pewnego skończonego lub przeliczalnego zbioru S. Wówczas proces $\{X_n\}$ scharakteryzowany jest przez rodzinę prawdopodobieństw

$$\{p(s_1, \ldots, s_n),\ s_i \in S,\ i = 1, \ldots, n,\ n = 1, 2, \ldots\}, \tag{2}$$

przy czym prawdopodobieństwa te spełniają warunki: dla każdego n

$$\sum_{s_1, \ldots, s_n} p(s_1, \ldots, s_n) = 1 \tag{3}$$

oraz dla dowolnych $s_1, \ldots, s_{n-1}$

$$\sum_{s_n \in S} p(s_1, \ldots, s_{n-1}, s_n) = p(s_1, \ldots, s_{n-1}). \tag{4}$$

Jak już powiedziano we wstępie, proces $\{X_n\}$ obserwowany jest sekwencyjnie. Jeżeli po zaobserwowaniu wartości $s_1, \ldots, s_n$ zmiennych losowych $X_1, \ldots, X_n$ zdecydujemy się na zatrzymanie procesu, otrzymujemy wypłatę $f(s_1, \ldots, s_n)$ zależną od dotychczasowych (ale nie od przyszłych) wartości procesu.

Reguła zatrzymywania powinna mieć postać instrukcji, która pozwala rozstrzygać, czy proces ma być zatrzymany, czy nie. Instrukcję taką można oczywiście wyrazić na różne sposoby; ogólnie (w rozważanym przypadku procesu o przeliczalnej ilości stanów) można ją przedstawić w postaci listy tych ciągów wyników, dla których proces należy zatrzymać.

Formalnie, oznaczmy przez S^* klasę wszystkich ciągów skończonych utworzonych z elementów zbioru S, tj.

$$S^* = \bigcup_{n=1}^{\infty} S^n,$$

gdzie oczywiście $S^n = S \times \ldots \times S$ (n razy).

DEFINICJA. *Regułą zatrzymania* będziemy nazywali zbiór $L \subset S^*$ spełniający następujące warunki:

(a) jeżeli $(s_1, \ldots, s_n) \in L$, to $(s_1, \ldots, s_j) \notin L$ dla $j = 1, \ldots, n-1$;

(b) $\sum_{(s_1, \ldots, s_n) \in L} p(s_1, \ldots, s_n) = 1$.

Wartością $E(L)$ reguły zatrzymywania L nazywa się oczekiwaną wypłatę przy przyjęciu reguły L, tj.

$$E(L) = \sum_{(s_1, \ldots, s_n) \in L} f(s_1, \ldots, s_n)\, p(s_1, \ldots, s_n), \tag{5}$$

jeżeli szereg ten jest bezwzględnie zbieżny.

Oznaczmy przez $\mathscr{L}$ klasę tych reguł zatrzymywania, dla których (5) ma sens, oraz przyjmijmy

$$V = \sup_{L \in \mathscr{L}} E(L).$$

Jeżeli $E(L^*) = V$, to regułę zatrzymywania L^* nazywa się *optymalną*.

Powyższe definicja reguły zatrzymywania, choć bardzo przejrzysta, ma jednak dwie wady: po pierwsze, dotyczy jedynie dyskretnych przestrzeni realizacji procesu stochastycznego, i po drugie, jest niezbyt wygodna do stosowania w praktyce.

Główna idea ogólnej definicji polega na wyrażeniu zasadniczych pojęć, mianowicie funkcji wypłaty i reguły zatrzymywania, jako funkcji liczbowych określonych na tej samej przestrzeni.

Można to osiągnąć za pomocą następujących konstrukcji. Niech Ω oznacza zbiór wszystkich ciągów nieskończonych $\omega = (x_1, x_2, \ldots)$, gdzie x_j są liczbami rzeczywistymi. Następnie, niech $\mathscr{B}_n$ będzie σ-ciałem podzbiorów Ω generowanym przez wszystkie zbiory postaci

$$\{\omega = (x_1, x_2, \ldots): x_1 \in E_1, \ldots, x_n \in E_n\},$$

gdzie $E_1, \ldots, E_n$ są zbiorami borelowskimi na prostej. Wreszcie, niech $\mathscr{B}$ będzie najmniejszym σ-ciałem zawierającym wszystkie σ-ciała $\mathscr{B}_n$. Z konstrukcji tej wynika, że $\mathscr{B}_1 \subset \mathscr{B}_2 \subset \ldots \subset \mathscr{B}$. Ponadto, σ-ciało $\mathscr{B}_n$ ma następującą własność: jeżeli $A \in \mathscr{B}_n$ i $\omega = (x_1, \ldots, x_n, x_{n+1}, \ldots) \in A$ oraz $\omega' = (x_1, \ldots, x_n, x'_{n+1}, \ldots)$, to $\omega' \in A$ dla dowolnych $x'_{n+1}, x'_{n+2}, \ldots$ Oznacza to, że kryterium przynależności elementu ω do danego zbioru $A \in \mathscr{B}_n$ zależy jedynie od współrzędnych $x_1, \ldots, x_n$.

Zbiór Ω będzie traktowany jako przestrzeń realizacji obserwowanego procesu; proces ten można utożsamić z pewną miarą prawdopodobieństwa P na przestrzeni mierzalnej[2] $(\Omega, \mathscr{B})$.

Dla konstrukcji ogólnej definicji funkcje wypłaty i reguły zatrzymywania traktowane będą jako funkcje określone na zbiorze Ω, a więc formalnie rzecz biorąc, funkcje zależne od nieskończonego ciągu współrzędnych $x_1, x_2, \ldots$ Niezależność danej funkcji od współrzędnych $x_{n+1}, x_{n+2}, \ldots$ będzie przy tym wyrażona przez żądanie mierzalności względem σ-ciała $\mathscr{B}_n$.

W dalszym ciągu, jeżeli v jest funkcją mierzalną względem $\mathscr{B}_n$, jej wartość w punkcie $\omega = (x_1, x_2, \ldots)$ będzie niekiedy zapisywana jako $v(x_1, \ldots, x_n)$. W szczególności będzie to dotyczyć warunkowych wartości oczekiwanych: zamiast niewygodnego symbolu $E(v|\mathscr{B}_n)(x_1, \ldots, x_n)$ będziemy przy tym używali bardziej przejrzystego[3] (choć mniej precyzyjnego) symbolu $E(v|x_1, \ldots, x_n)$.

---

[2] Dokładniej, dla określonej powyżej przestrzeni probabilistycznej $(\Omega, \mathscr{B}, \mathrm{P})$, zmienne losowe X_j procesu określa się jako $X_j(\omega) = X_j((x_1, x_2, \ldots)) = x_j$.

[3] Warunkowa wartość oczekiwana względem σ-ciała $\mathscr{B}_n$ jest funkcją $\mathscr{B}_n$-mierzalną, a więc zależy właściwie jedynie od współrzędnych $x_1, \ldots, x_n$. Ponieważ pojęcie wartości oczekiwanej względem σ-ciała nie jest zbyt powszechnie znane, przydatne może być następujące wyjaśnienie. Niech g będzie funkcją całkowalną na przestrzeni probabilistycznej $(\Omega, \mathscr{B}, \mathrm{P})$, tj. zmienną losową o skończonej wartości oczekiwanej Eg, i niech $\mathscr{B}'$ będzie σ-ciałem zawartym w $\mathscr{B}$. Jeżeli $\mathscr{B}'$ jest generowane przez skończoną liczbę zdarzeń $A_1, \ldots, A_k$ (tak, że $A_i \cap A_j = \emptyset$ oraz $\bigcup_i A_i = \Omega$, przy czym każdy element $\mathscr{B}'$ jest sumą pewnych zdarzeń A_i lub ich uzupełnień), oraz $\mathrm{P}(A_i) > 0$, dla $i = 1, \ldots, k$, to warunkowa wartość oczekiwana $E(g|\mathscr{B}')$ względem σ-ciała $\mathscr{B}'$ określona jest jako funkcja ω równa $E(g|A_j)$ dla $\omega \in A_j$, gdzie $E(g|A_j)$ jest znaną z klasycznej teorii prawdopodobieństwa warunkową wartością oczekiwaną pod warunkiem zajścia zdarzenia A_j. Dla definicji wartości oczekiwanej względem σ-ciała w przypadku ogólnym por. np. [6].

Możemy obecnie określić funkcję wypłaty. Zgodnie z warunkami sformułowanymi we wstępie, żądającymi aby wypłata w przypadku zatrzymania procesu w chwili n zależała jedynie od $x_1, \ldots, x_n$, założymy, że wypłata ma postać ciągu funkcji $f_1, f_2, \ldots$ takich, że f_n jest mierzalne względem $\mathscr{B}_n$. Funkcja f_n jest tu oczywiście wypłatą, jeżeli proces zostaje zatrzymany w chwili n.

Określony powyżej ciąg $\{f_n, \mathscr{B}_n\}$, gdzie $\{\mathscr{B}_n\}$ jest monotonicznie rosnącym ciągiem σ-ciał zawartych w σ-ciele $\mathscr{B}$ pewnej przestrzeni probabilistycznej $(\Omega, \mathscr{B}, \mathrm{P})$, a f_n jest funkcją $\mathscr{B}_n$-mierzalną, nazywa się *ciągiem stochastycznym*. Jeżeli $E|f_n| = \int |f_n| dP < \infty$ dla $n = 1, 2, \ldots$, to ciąg stochastyczny nazywa się *całkowalny*. W dalszym ciągu milcząco zakładamy, że rozważamy jedynie ciągi całkowalne.

Możemy obecnie określić pojęcie reguły zatrzymywania. W myśl sformułowanego we wstępie intuicyjnego założenia, reguła zatrzymywania powinna każdemu ciągowi $\omega = (x_1, x_2, \ldots)$ przyporządkowywać moment zatrzymania procesu. Tak więc, reguła taka jest funkcją t: $\Omega \to \mathscr{N} = \{1, 2, \ldots\}$. Ponadto, niezależność decyzji o zatrzymaniu od przyszłych wartości procesu wymaga, aby spełniony był następujący warunek: jeżeli $t(\omega) = n$, gdzie $\omega = (x_1, x_2, \ldots)$, to $t(\omega') = n$ dla każdego ω' postaci $(x_1, \ldots, x_n, x'_{n+1}, \ldots)$ przy dowolnych $x'_{n+1}, x'_{n+2}, \ldots$ Wreszcie, reguła zatrzymywania powinna doprowadzić (z prawdopodobieństwem 1) do zatrzymania procesu. Łącząc te wymagania razem, możemy sformułować następującą definicję:

DEFINICJA. *Regułą zatrzymywania* nazywamy funkcję t: $\Omega \to \mathscr{N}$ taką, że

$$\text{(a)} \qquad \{\omega: t(\omega) = n\} \in \mathscr{B}_n, \quad n = 1, 2, \ldots,$$

$$\text{(b)} \qquad \mathrm{P}(t < \infty) = 1.$$

Określmy teraz, dla danej reguły zatrzymywania t, wypłatę f_t jako $f_{t(\omega)}(\omega)$ i przyjmijmy $Ef_t = \int f_{t(\omega)}(\omega)\, \mathrm{P}(d\omega)$, o ile całka ta istnieje.

Oznaczmy przez T klasę wszystkich reguł zatrzymywania, dla których Ef_t istnieje [4]) i niech $V = \sup_{t \in T} Ef_t$. Każda reguła zatrzymywania t^*, dla której $Ef_{t^*} = V$ nazywa się *optymalną*.

4. Programy skończone. Indukcja wsteczna. Rozważmy ciąg stochastyczny $\{f_n, \mathscr{B}_n\}$ i oznaczmy przez T^N klasę wszystkich reguł zatrzymywania t, dla których $t \leqslant N$. Niech $V^N = \sup \{Ef_t: t \in T^N\}$. Można pokazać, że w klasie T^N istnieje reguła optymalna; co więcej, można wskazać metodę jej efektywnego wyznaczania, za pomocą tzw. indukcji wstecznej.

Rozumowanie jest w tym przypadku następujące. Dla $n \leqslant N$ oznaczmy przez T_n^N podklasę T^N składającą się z reguł zatrzymywania t spełniających warunek $n \leqslant t \leqslant N$ i niech $V_n^N = \sup \{Ef_t: t \in T_n^N\}$ (tak, że $T_1^N = T^N$ i $V_1^N = V^N$).

Optymalna reguła zatrzymywania w klasie T_n^N stanowi rozwiązanie problemu „częściowego", mianowicie optymalnego zatrzymania procesu nie później niż

([4]) Klasa T jest niepusta (dla całkowalnych ciągów stochastycznych $\{f_n, \mathscr{B}_n\}$): na przykład dla reguły $t = 1$ (zatrzymaj proces po pierwszej obserwacji) jest $Ef_t = Ef_1 < \infty$.

w N-tym kroku, w sytuacji gdy dokonaliśmy już $n \leqslant N$ obserwacji bez zatrzymywania procesu. Załóżmy na chwilę, że w każdej z klas T_n^N istnieje reguła optymalna.

Oznaczmy przez $u_n^N(x_1, \ldots, x_n)$ oczekiwaną wypłatę, przy zastosowaniu optymalnej reguły z klasy T_n^N, jeżeli dotychczasowe obserwacje są $x_1, \ldots, x_n$. Podobnie jak poprzednio, wygodnie będzie traktować funkcję u_n^N jako funkcję od ω mierzalną względem σ-ciała $\mathscr{B}_n$.

Funkcje u_n^N można wyznaczyć rekurencyjnie w sposób następujący. Przede wszystkim $u_N^N = f_N$, gdyż po zaobserwowaniu wartości $x_1, \ldots, x_N$ proces musi zostać zatrzymany i wypłata wynosi $f_N(x_1, \ldots, x_N)$ (ponieważ f_N jest funkcją $\mathscr{B}_N$-mierzalną, jej wartość zależy jedynie od pierwszych N współrzędnych punktu ω).

Przypuśćmy teraz, że dokonaliśmy $N-1$ obserwacji, powiedzmy $x_1, \ldots, x_{N-1}$. Jeżeli zatrzymamy proces, otrzymamy wypłatę $f_{N-1}(x_1, \ldots, x_{N-1})$. Jeżeli będziemy go kontynuować, pojawi się nowa wartość x_N i wypłata wyniesie $f_N(x_1, \ldots, x_N) = u_N^N(x_1, \ldots, x_N)$. Wobec tego oczekiwana wartość wypłaty, jeżeli proces będzie kontynuowany, równa się wartości funkcji $E(f_N|\mathscr{B}_{N-1}) = E(u_N^N|\mathscr{B}_{N-1})$ w punkcie $x_1, \ldots, x_{N-1}$ (przy „uproszczonym" zapisie, wyniesie ona $E(u_N^N|x_1, \ldots, x_{N-1})$). Optymalna reguła zatrzymywania każe więc proces zatrzymać lub kontynuować w zależności od tego, czy $f_{N-1}(x_1, \ldots, x_{N-1})$ jest większe czy mniejsze od $E(u_N^N|x_1, \ldots, x_{N-1})$.

Rozumowanie takie stosuje się do każdego punktu $x_1, \ldots, x_{N-1}$ i wobec tego opisane postępowanie zapewnia (oczekiwaną) wypłatę równą

$$\max [f_{N-1}, E(u_N^N|\mathscr{B}_{N-1})];$$

wyrażenie to równe jest oczywiście u_{N-1}^N.

Mając określone wypłaty u_{N-1}^N możemy cofnąć się o krok wstecz, wyznaczając w podobny sposób u_{N-2}^N, itd. Ogólnie, funkcje u_N^N określone są indukcyjnie przez równości

$$\begin{aligned} &u_N^N = f_N, \\ &u_n^N = \max [f_n, E(u_{n+1}^N|\mathscr{B}_n)], \quad n = N-1, N-2, \ldots, 1. \end{aligned} \tag{6}$$

Możemy teraz opuścić założenie o istnieniu optymalnych reguł zatrzymywania w klasach T_n^N i wyjść bezpośrednio od wzorów (6), określając

$$s_n^N = \textit{pierwsze } j \geqslant n \textit{ takie, że } u_j^N = f_j. \tag{7}$$

Z definicji, $s_n^N \geqslant n$, a ponieważ $u_N^N = f_N$, więc $s_n^N \leqslant N$, wobec czego $s_n^N \in T_n^N$. Można udowodnić (por. [3]), że

$$E(f_{s_n^N}|\mathscr{B}_n) = u_n^N \geqslant E(f_t|\mathscr{B}_n) \quad \text{dla } t \in T_n^N, \tag{8}$$

skąd całkując otrzymujemy

$$Ef_{s_n^N} = Eu_n^N \geqslant Ef_t \quad \text{dla } t \in T_n^N, \tag{9}$$

a więc

$$V_n^N = Eu_n^N, \tag{10}$$

czyli s_n^N jest optymalną regułą zatrzymywania w klasie T_n^N. Dla dowodu tego twierdzenia por. np. [3].

Z powyższego twierdzenia wynika, że w każdym skończonym problemie istnieje optymalna reguła zatrzymywania.

Pokażemy teraz zastosowanie tego twierdzenia, wyznaczając mianowicie optymalną regułę wyboru najlepszej kandydatki na sekretarkę (ostatni przykład z paragrafu 2).

Przede wszystkim łatwo zauważyć, że zmienne losowe X_n (gdzie X_n jest rangą n-tej kandydatki w stosunku do poprzednich) są niezależne i

$$\mathrm{P}(X_n = j) = \frac{1}{n}, \quad j = 1, \ldots, n.$$

Ponadto,

$$f_n = \mathrm{P}(Y_n = 1|\mathscr{B}_n) = \mathrm{P}(Y_n = 1|X_1, \ldots, X_n) =$$

$$= \mathrm{P}(Y_n = 1|X_n) = \begin{cases} 0 & \text{jeżeli } X_n > 1, \\ n/N & \text{jeżeli } X_n = 1. \end{cases}$$

Na mocy twierdzenia, mamy zatem

$$u_N^N = f_N = \begin{cases} 1 & \text{jeżeli } X_N = 1, \\ 0 & \text{jeżeli } X_N > 1, \end{cases}$$

wobec czego, na podstawie (6) i (10), jest $Ef_N = V_N^N = 1/N$.

Następnie, $u_{N-1}^N = \max\,[f_{N-1},\ E(u_N^N|\mathscr{B}_{N-1})] = \max\,[f_{N-1},\ E(f_N|\mathscr{B}_{N-1})]$.

Ponieważ $f_{N-1} = 0$ lub $(N-1)/N$ w zależności od tego, czy $X_{N-1} > 1$ czy $X_{N-1} = 1$, podczas gdy $E(f_N|X_{N-1} = 1) = E(f_N|X_{N-1} > 1) = 1/N$, więc u_{N-1}^N równa się $(N-1)/N$ lub $1/N$ w zależności od tego, czy $X_{N-1} = 1$ czy $X_{N-1} > 1$. Ponadto, $E(f_N|\mathscr{B}_{N-1}) = Ef_N = V_N^N$, wobec czego

$$u_{N-1}^N = \max\,[f_{N-1},\ V_N^N].$$

Podobnie,

$$u_{N-2}^N = \max\,[f_{N-2},\ E(u_{N-1}^N|\mathscr{B}_{N-2})]$$

i $E(u_{N-1}^N|\mathscr{B}_{N-2}) = E(u_{N-1}^N|X_{N-2}) = Eu_{N-1}^N = V_{N-1}^N$ na podstawie relacji (10), skąd

$$u_{N-2}^N = \max\,[f_{N-2},\ V_{N-1}^N].$$

Ogólnie, mamy

$$u_n^N = \max\,[f_n,\ V_{N+1}^N],$$

skąd wynika, że $V_n^N = Eu_n^N \geqslant V_{N+1}^N$, czyli

$$V_1^N \geqslant V_2^N \geqslant \ldots \geqslant V_N^N = 1/N. \tag{11}$$

Optymalna reguła zatrzymywania $s = s_1^N$ określona przez (7) ma postać

$$s = \textit{pierwsze } n \geqslant 1 \textit{ takie, że } f_n \geqslant V_{n+1}^N;$$

ponieważ $f_n = 0$ jeżeli $X_n > 1$, można tę regułę, na mocy (11), scharakteryzować następująco: istnieje r $(1 \leqslant r \leqslant N)$ takie, że

$$s = \textit{pierwsze } n \geqslant r, \textit{ dla którego } X_n = 1 \tag{12}$$

(innymi słowy, należy przeprowadzić wywiady z $r-1$ kandydatkami nie angażując żadnej z nich, a następnie zaangażować pierwszą, która okaże się lepsza od wszystkich poprzednich).

Tak więc, problem został znacznie zredukowany, przez wskazanie bardzo prostej klasy reguł zatrzymywania postaci (12), wśród których znajduje się optymalna. Pozostaje jedynie wyznaczyć liczbę r. W tym celu należy obliczyć prawdopodobieństwo wyboru najlepszej kandydatki przy opisanym sposobie postępowania jako funkcję r. Łatwo sprawdzić, że prawdopodobieństwo to równa się $1/N$ dla $r = 1$, podczas gdy dla $r > 1$ wynosi ono

$$\frac{r-1}{N} \sum_{k=r}^{N} \frac{1}{k-1}. \tag{13}$$

Istotnie, aby dla permutacji $a_1, \ldots, a_N$ liczb $1, \ldots, N$ reguła zatrzymywania (12) doprowadziła do wyboru najlepszej kandydatki, 1 musi znajdować się na jednym z miejsc $r, r+1, \ldots, N$. Jeżeli 1 znajduje się na miejscu k gdzie $r \leqslant k \leqslant N$ (prawdopodobieństwo tego zdarzenia wynosi $1/N$), to wśród wyrazów $a_1, \ldots, a_{r-1}$, $a_r, \ldots, a_{k-1}$ minimum musi wystąpić na jednym z miejsc $1, \ldots, r-1$ (prawdopodobieństwo $(r-1)/(k-1)$).

Oznaczając wyrażenie (13) przez $A(r)$, mamy $A(r) - A(r+1) = N^{-1}\left(1 - \sum_{k=r}^{N-1} \frac{1}{k}\right)$, i $A(r)$ osiąga maksimum dla

$$r(N) = \inf\{r\colon A(r) - A(r+1) \geqslant 0\} =$$

$$= \inf\left\{r\colon \frac{1}{r} + \frac{1}{r+1} + \ldots + \frac{1}{N-1} \leqslant 1\right\}.$$

Drogą łatwych oszacowań otrzymuje się stąd relację

$$\lim_{N\to\infty} N^{-1} r(N) = e^{-1},$$

jak również

$$\lim_{N\to\infty} V^N = \lim_{N\to\infty} A\big(r(N)\big) = e^{-1}.$$

Tak więc, dla dużych N należy przebadać bez zatrzymywania około 37% kandydatek, a następnie zatrzymać pierwszą, która jest lepsza od wszystkich poprzednich.

Postępowanie takie zapewnia prawdopodobieństwo około 0,37 wyboru najlepszej kandydatki.

5. Najprostsze programy nieskończone. Przypadek monotoniczny. Zaczniemy rozważania od najprostszego przypadku, mianowicie gdy obserwowany proces jest procesem Markowa. Jak można oczekiwać zgodnie z intuicją, jeżeli optymalna reguła zatrzymywania istnieje, to decyzję o zatrzymaniu można uzależnić jedynie od aktualnego (ostatniego) stanu procesu. Co więcej, jeżeli proces jest jednorodny w czasie (tj. prawdopodobieństwa przejścia nie zależą od n), to optymalna reguła zatrzymywania daje się wyrazić w postaci

$$t^* = \textit{pierwsze } n \textit{ takie, że } X_n \in Q,$$

gdzie Q jest pewnym zbiorem stanów.

Dla zilustrowania metody znajdowania zbioru Q, załóżmy dla uproszczenia, że wypłata określona jest następująco: jeżeli zatrzymujemy proces w momencie, w którym ostatnim stanem jest s, otrzymujemy $a(s)$, natomiast za prawo kontynuowania płacimy $b(s)$. Wobec tego łączna wypłata, jeżeli zatrzymujemy proces po zaobserwowaniu ciągu $s_1, s_2, \ldots, s_n$ równa się $a(s_n)-b(s_1)-\ldots-b(s_{n-1})$.

Załóżmy ponadto, że zbiór stanów S jest dyskretny, i niech $p(s, s')$ oznacza prawdopodobieństwo przejścia ze stanu s do stanu s'.

Przypuśćmy, że optymalna reguła zatrzymywania istnieje dla każdego stanu początkowego s, i niech $V(s)$ oznacza oczekiwaną wypłatę przy zastosowaniu tej reguły, jeżeli stanem początkowym jest s.

Wyobraźmy sobie, że zaobserwowaliśmy dotychczas ciąg $s_1, s_2, \ldots, s_n$ i niech $s_n = s$. Jasne jest, że możemy zaniedbać dotychczasowe opłaty $b(s_1)+\ldots+b(s_{n-1})$: zostały one wpłacone i nie zwrócą się, niezależnie od tego, czy proces zatrzymamy, czy zdecydujemy się go kontynuować. Tak więc, jeżeli zatrzymamy proces, otrzymamy $a(s)$. Jeżeli zdecydujemy się na jeszcze jeden krok, po którym będziemy postępowali w sposób optymalny, oczekiwana wypłata wyniesie oczywiście

$$-b(s)+\sum_{s'\in S} p(s, s')\ V(s').$$

Wobec tego możemy napisać

$$V(s) = \max\left[a(s), -b(s)+\sum_{s'\in S} p(s, s')\ V(s')\right] \tag{14}$$

dla $s \in S$. Jeżeli wiadomo, że optymalna reguła zatrzymywania istnieje, i znane jest rozwiązanie równań (14), to tę optymalną regułę można wyznaczyć następująco: niech

$$Q = \left\{s \in S:\ a(s) \geqslant -b(s)+\sum_{s'\in S} p(s, s')\ V(s')\right\}.$$

Wówczas $t = \textit{pierwsze } n \geqslant 1 \textit{ takie, że } s_n \in Q$ jest optymalną regułą zatrzymywania.

W ogólnym przypadku, oznaczmy przez T_n klasę wszystkich reguł zatrzymywania t spełniających warunek $t \geqslant n$ i oznaczmy([5])

$$u_n = \sup_{t \in T_n} E(f_t|\mathscr{B}_n).$$

Wówczas odpowiednikiem (14) jest równość

$$u_n = \max [f_n, E(u_{n+1}|\mathscr{B}_n)]$$

powstaje problem znalezienia warunków, przy których

$$t^* = \begin{cases} \text{pierwsze } n \text{ takie, że } u_n = f_n, \\ \infty, \text{ jeżeli } u_n > f_n \text{ dla wszystkich } n, \end{cases}$$

jest optymalną regułą zatrzymywania.

Ponadto, ze względu na to, że wszystkie skończone programy mają rozwiązania optymalne, naturalne wydaje się zbadanie warunków, przy których ciąg rozwiązań problemów skończonych jest zbieżny do rozwiązania problemu ogólnego.

Nie będziemy tu oczywiście przedstawiać teorii ogólnej; ograniczymy się jedynie do wyodrębnienia ważnej klasy problemów, w których można spodziewać się istnienia optymalnych reguł zatrzymywania.

Dla danego ciągu stochastycznego $\{f_n, \mathscr{B}_n\}$ oznaczmy

$$A_n = \{\omega: f_n \geqslant E(f_{n-1}|\mathscr{B}_n)\}.$$

Powiemy, że problem jest *monotoniczny*, jeżeli

(a) $$A_1 \subset A_2 \subset \ldots$$

(b) $$P\Big(\bigcup_n A_n\Big) = 1.$$

Intuicyjna treść pojęcia monotoniczności jest następująca. Jeżeli w jakimś punkci $\omega = (x_1, x_2, \ldots)$ zachodzi nierówność $f_n \geqslant E(f_{n+1}|\mathscr{B}_n)$, czyli w niezbyt precyzyjnym uproszczonym zapisie, nierówność $f_n(x_1, \ldots, x_n) \geqslant E(f_{n+1}|x_1, \ldots, x_n)$, oznacza to, że po zaobserwowaniu $x_1, \ldots, x_n$, z dwóch możliwości: „zatrzymać proces dziś" i „zatrzymać proces jutro", pierwsza nie jest gorsza od drugiej (w sensie oczekiwanej wypłaty). W przypadku monotonicznym, nierówność $f_n \geqslant E(f_{n+1}|\mathscr{B}_n)$ dla pewnego ω pociąga za sobą (warunek (a)) nierówności $f_{n+1} \geqslant E(f_{n+2}|\mathscr{B}_{n+1})$, $f_{n+2} \geqslant E(f_{n+3}|\mathscr{B}_{n+2})$, itd. Mówiąc obrazowo, jeżeli po zaobserwowaniu $x_1, \ldots, x_n$ „zatrzymanie dziś" jest nie gorsze niż „zatrzymanie jutro", to sytuacja ta będzie się powtarzać już zawsze: każda zwłoka o jedną jednostkę czasu może co najwyżej pogorszyć sytuację.

Ponadto warunek (b) orzeka, że taki moment, w którym kontynuacja przestanie się raz na zawsze opłacać, musi na pewno nastąpić prędzej czy później.

([5]) Dla ścisłości należałoby określić u_n jako $\operatorname{ess\,sup}_{t \in T_n} E(f_t|\mathscr{B}_n)$ ze względu na to, że warunkowe wartości oczekiwane wyznaczone są jednoznacznie jedynie prawie wszędzie.

Wobec tego, w przypadku monotonicznym naturalnym kandydatem na optymalną regułę zatrzymywania jest

$$t^* = \textit{pierwsze } n \geqslant 1 \textit{ takie, że } f_n \geqslant E(f_{n+1}|\mathscr{B}_n). \tag{15}$$

Zanim sformułujemy warunki, przy których t^* jest optymalną regułą zatrzymywania w problemach monotonicznych, pokażemy przykład (zaczerpnięty z [3]), pokazujący, że t^* bynajmniej nie musi być optymalne.

Rozważmy mianowicie zmodyfikowaną nieco grę „zrezygnuj lub podwój". Niech $X_1, X_2, \ldots$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie $\mathrm{P}(X_j = 1) = \mathrm{P}(X_j = -1) = 1/2$ i przyjmijmy, że wypłata przy zatrzymaniu procesu po zaobserwowaniu wartości $x_1, \ldots, x_n$ zmiennych $X_1, \ldots, X_n$ wynosi

$$f_n = \frac{2n}{n+1} \prod_{j=1}^{n} (x_j+1).$$

Interpretując $+1$ jako orła i -1 jako reszkę, wypłata równa się 0, jeżeli zdarza się choć jedna reszka, natomiast w przypadku przerwania procesu dla ciągu n kolejnych orłów wynosi ona $\frac{2n}{n+1} \cdot 2^n$. Wówczas $E(f_{n+1}|\mathscr{B}_n) = E(f_{n+1}|X_1, \ldots, X_n)$ równa się 0 jeżeli w ciągu $X_1, \ldots, X_n$ zdarzył się składnik -1, oraz

$$E(f_{n+1}|X_1 = \ldots = X_n = 1) = \frac{1}{2} \cdot \frac{2(n+1)}{n+2} 2^{n+1} > \frac{2n}{n+1} \cdot 2^n.$$

Wobec tego, jeżeli zaobserwowaliśmy do danego momentu n orłów, nie opłaca się gry przerywać. Mimo to jednak, postępowanie stale w myśl tej zasady prowadzi z prawdopodobieństwem 1 do osiągnięcia wygranej 0.

Łatwo się przekonać, że problem jest monotoniczny. Istotnie, jeżeli $f_n \geqslant E(f_{n+1}|\mathscr{B}_n)$, to musi być $f_n = 0$, a więc również $f_{n+1} = E(f_{n+2}|\mathscr{B}_{n+1}) = 0$, czyli $f_{n+1} \geqslant E(f_{n+2}|\mathscr{B}_{n+1})$. Ponadto, warunek (b) jest spełniony, ponieważ zdarzenie $X_n = -1$ nastąpi dla pewnego n z prawdopodobieństwem jeden. Reguła (15) ma więc tu postać

$$t^* = \textit{pierwsze } n \geqslant 1, \textit{ dla którego } X_n = -1$$

i jest najgorszą z możliwych.

Powyższy przykład jest oczywiście dość patologiczny i wskazuje na konieczność przyjęcia dodatkowych założeń przy dowodzie optymalności reguły zatrzymywania (15). Prawdziwe jest mianowicie następujące

TWIERDZENIE. *W przypadku monotonicznym, jeżeli*

$$\liminf_{n\to\infty} \int_{\{t^*>n\}} f_n^+ \, d\mathrm{P} = 0,$$

to $Ef_{t^*} \geqslant Ef_t$ *dla każdej reguły zatrzymywania* t, *dla której*

$$\liminf_{n\to\infty} \int_{\{t>n\}} f_n^- \, dP = 0$$

(tutaj f_n^+ i f_n^- oznaczają część dodatnią i część ujemną funkcji f_n).

6. Uwagi końcowe. Ze względu na to, że teoria reguł zatrzymywania powstała bardzo niedawno i w chwili obecnej jest jednym z najbardziej dynamicznie rozwijających się działów rachunku prawdopodobieństwa, jest jeszcze o wiele za wcześnie na próbę syntezy i ogólniejszej oceny tej teorii. Wiąże się ona zresztą ściśle z szeregiem innych teorii, przede wszystkim ze statystyką matematyczną (testy sekwencyjne), z programowaniem dynamicznym, teorią sterowania, itp., a ponadto rozwijana jest również dla przypadku zatrzymywania procesów o czasie ciągłym.

W krótkim artykule nie sposób również omówić dokładniej zastosowań tej teorii. Warto jedynie może wspomnieć o pewnym dość nieoczekiwanym „produkcie ubocznym" rozważań nad regułami zatrzymywania, mianowicie konstrukcją testów statystycznych o mocy 1 (por. [1], [7]).

Ogólnie, wydaje się, że teorię tę dobrze scharakteryzowali jej twórcy, wybierając tytuł swojej książki: *Wielkie nadzieje*: *teoria optymalnego zatrzymywania* (6).

Literatura

[1] G. A. Barnard, *Practical applications of tests with power one*, Proc. of the 37-th Session of the ISI, Londyn 1969.

[2] L. Breiman, *Stopping rule problems*, w *Applied Combinatorial Mathematics*, wyd. E. F. Beckenbach. Wiley, New York 1964.

[3] Y.S. Chow, H. Robbins and D. Siegmund, *Great expectations*: *The theory of optimal stopping*, Houghton Mifflin Co., Boston 1971.

[4] A. Dvoretzky, *Existence and properties of certain optimal stopping rules*, Proc. Fifth Berkeley Symposium on Probability Theory and Mathematical Statistics, Berkeley i Los Angeles. Univ. of California Press, 1967.

[5] W. Feller, *Wstęp do teorii prawdopodobieństwa*, Tom. I, PWN, Warszawa 1966.

[6] M. Loève, *Probability theory*, Van Nostrand, Princeton 1960.

[7] H. Robbins and D. Siegmund, *Confidence intervals and interminable tests*, Proc. of the 37-th Session of the ISI, Londyn 1969.

(6) *Great Expectations*: *The theory of optimal stopping.*