

Ireneusz J. JÓŹWIAK, Łukasz ŁOZIUK  
Wydział Informatyki i Zarządzania  
Politechnika Wroclawska

## ALGORYTMY GRY W ZAKŁADACH BUKMACHERSKICH

**Streszczenie.** Artykuł opisuje usystematyzowane podejścia do gry hazardowej w zakłady bukmacherskie. Przedstawiono systemy gry oraz metody analizy statystycznej rynku bukmacherskiego. Zaproponowano algorytm przewidywania zwiększonego prawdopodobieństwa wystąpienia wyniku remisowego w meczach piłkarskich. Algorytm został sprawdzony w kolekturach bukmacherskich przy użyciu realnych środków finansowych i przyniósł zysk.

**Słowa kluczowe:** zakład bukmacherski, algorytm gry bukmacherskiej, statystyka meczów piłkarskich.

## ALGORITHMS IN BETTING

**Summary.** Paper describes a systematized soccer betting approach and presents methods of betting market analysis. An algorithm of increased probability of draw result anticipation is proposed. The algorithm, tested on the betting market, gave positive results.

**Keywords:** betting, betting algorithm, soccer matches statistics.

### 1. Wprowadzenie

Algorytm gry w zakłady bukmacherskie jest usystematyzowanym podejściem do hazardu, którego celem jest przeciwdziałanie polityce finansowej kasyna i wypracowanie zysku. Skuteczny algorytm powinien zwiększać szanse zwycięstwa w celu wypracowania długoterminowego zysku. Skuteczne systemy bukmacherskie, oparte na analizie statystycznej, wykorzystują rzadkie, sprzyjające graczowi okoliczności. Kurs meczu, jaki bukmacher oferuje graczom, wynika z wyliczonego przez niego prawdopodobieństwa

wystąpienia określonego wyniku meczu. Bukmacher zawsze stara się oferować graczom kurs dający ujemną wartość oczekiwaną zysku z zakładu. Zysk firmy bukmacherskiej wynosi  $I = vW$ , gdzie  $W$  jest sumą kwot wniesionych zakładów,  $v$ , czyli vig [8], jest opisany równaniem:

$$v = 100 \left( \frac{pq}{p+q} \right) \quad (1)$$

gdzie  $p$  i  $q$  to oferowane kursy meczu.

Wielkość  $v$  można nazwać prowizją pobieraną przez bukmachera. Kolektury bukmacherskie silnie konkurują ze sobą, oferując czasem zawyżone kursy meczów. Zawyżony kurs meczu sprawia, że zawarcie zakładu staje się opłacalne.

Pierwsze prace na temat statystycznego modelowania danych pochodzących z lig piłkarskich koncentrują się głównie na rozkładzie liczby zdobytych bramek w meczu. Moroney [6] oraz Reep [7] zgodnie stwierdzili, iż wynik gry jest w znacznym stopniu losowy, a przewidzenie na podstawie danych historycznych rezultatu nie jest możliwe. Hill [5] zastosował prosty test porównawczy końcowych pozycji zespołów w lidze z przewidywaniami ekspertów, stwierdzając ich silną korelację. Znacznie bardziej złożona analiza została przeprowadzona przez Fahrmeira [8], który użył skonstruowanych przez niego technik do analizy zależnych od czasu porównań do danych pochodzących z niemieckiej ligi piłkarskiej. Fahrmeir zauważył następującą dwudzielność: nie jest problemem dość dokładnie przewidzieć, które z drużyn odniosą sukces, ale opracowanie modelu, który miałby wystarczająco wysoką rozdzielczość, by wykorzystać długoterminowe przewidywania, jest znacznie trudniejsze. Fahrmeir [8] zbadał korelacje między rozgrywaniem meczu w roli gospodarza a jego rezultatem. Owen [4] zastosował dynamiczne modele liniowe, uwzględniające zmianę potencjału zespołów, do przewidywania ich wyników, stwierdzając wzrost trafności przewidywań.

## 2. Dane do badań

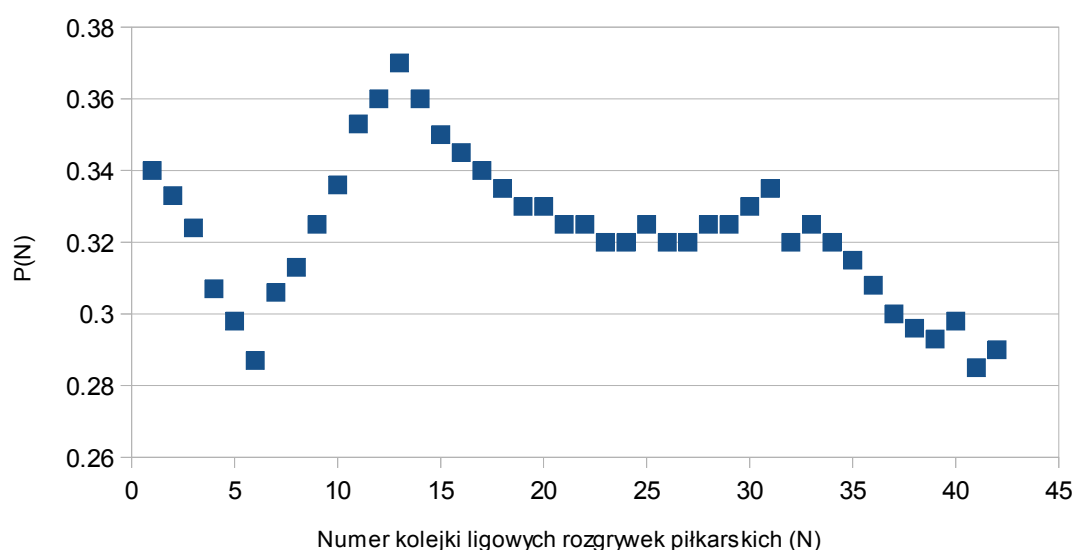
Każdy mecz piłkarski jest bogatym źródłem informacji. Zapisywane są nie tylko wyniki meczu, lecz także czas zdobycia bramek, nazwiska strzelców, pozycja ligowa zespołu, data, godzina gry itd. Indywidualna forma zespołu w wybranym meczu jest również określona przez wiele zewnętrznych czynników: pojawienie się nowych piłkarzy w zespole lub zwolnienie kierownika zespołu. Choć te informacje są również dostępne, są mniej sformalizowane, a ich jakość jest subiektywna. Dane pochodzące z ponad trzydziestu lig i dziesięciu sezonów piłkarskich zostały poddane analizie statystycznej oraz przeglądowi pełnemu. Szukano obciążeń rozkładów prawdopodobieństwa oraz symulowano grę

zadanyм алгоритмом букмacherskim. Dane i алгоритмы, które podczas симуляции przynosiły zysk, były poddawane szczegółowej analizie statystycznej. Istnieją dwa rodzaje obciążen rozkładu prawdopodobieństwa zakończenia meczu z wynikiem remisowym. Te zjawiska są najbardziej widoczne w ligach, w których liczba wyników remisowych jest bardzo duża, np. w drugiej lidze francuskiej i włoskiej.

### 3. Rozkład prawdopodobieństwa wystąpienia remisów

#### 3.1. Zależność czasowa rozkładu prawdopodobieństwa wystąpienia remisu

Analiza statystyczna wykazała, że prawdopodobieństwo zakończenia meczu wynikiem remisowym nie jest jednakowe na przestrzeni całego sezonu piłkarskiego. W środku sezonu piłkarskiego liczba remisów jest większa niż w pozostałych okresach. W przypadku drugiej ligi włoskiej różnica pomiędzy średnim prawdopodobieństwem zakończenia meczu wynikiem remisowym a prawdopodobieństwem wyniku remisowego w meczu wybranym ze środka sezonu piłkarskiego wynosi cztery punkty procentowe, co obrazuje wykres na rysunku 1.



Rys. 1. Prawdopodobieństwo zakończenia meczu remisem w Serie B w latach 2002-2012

Fig. 1. Probability distribution of draw outcomes in Serie B 2002-2012

Źródło: opracowanie własne.

Czteroprocentowa różnica prawdopodobieństwa wystąpienia remisu informuje o zwiększonej wartości oczekiwanej zysku z zakładu w środku sezonu piłkarskiego względem zakładów wniesionych w innych okresach.

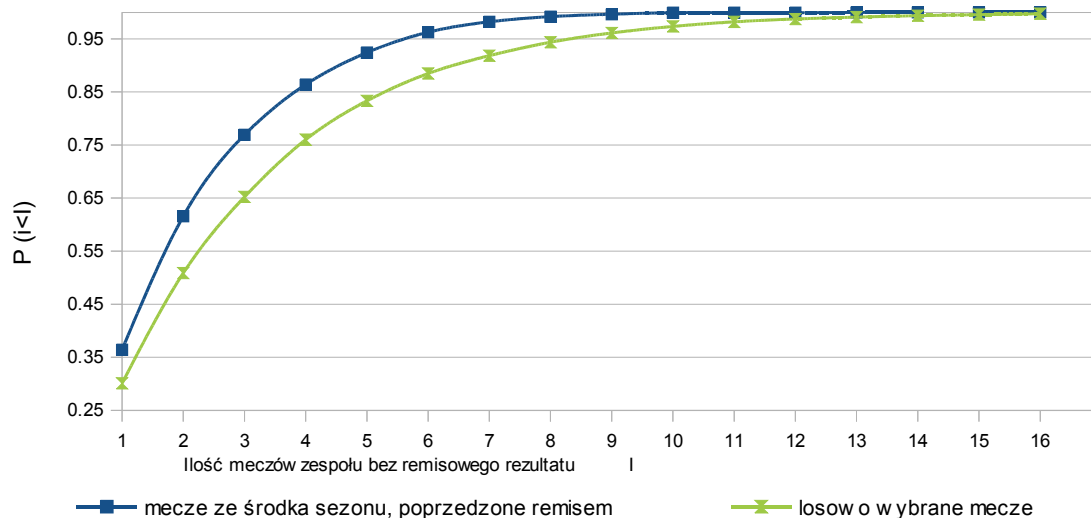
### 3.2. Wpływ formy zespołu na rozkład prawdopodobieństwa wystąpienia remisu

Analiza statystyczna wykazała, iż istnieje korelacja pomiędzy wynikiem meczu a wynikami osiągniętymi przez zespół w poprzednich meczach. Korelacja jest związana ze sportową formą zespołu, którą można określić na podstawie kilku wcześniejszych meczów. Badania wykazały, iż istnieje zwiększone o około dwa punkty procentowe prawdopodobieństwo zakończenia meczu z wynikiem remisowym, jeśli rezultat poprzedniego meczu rozegranego przez analizowany zespół również był remisowy.

### 3.3. Wybór meczów o wysokim prawdopodobieństwie zakończenia remisem

W celu optymalizacji algorytmu typowania meczów należy określić mecze  $M$ , na które ma wpływ zarówno obciążenie związane z formą zespołu  $M_F$ , jak i obciążenie związane z zależnością czasową występowania remisowych wyników  $M_T$ . Mecze  $M$  są częścią wspólną zbiorów  $M_F$  i  $M_T$ , czyli ich iloczynem:

$$M = M_F \cap M_T \quad (2)$$



Rys. 2. Rozkład prawdopodobieństwa wystąpienia meczów zespołu bez wyniku remisowego

Fig. 2. Probability distributions of non draw outcomes of a football team

Źródło: Opracowanie własne.

Prawdopodobieństwo zakończenia meczu ze zbioru  $M$  rezultatem remisowym jest dla drugiej ligi francuskiej i włoskiej o około sześć punktów procentowych wyższe niż w meczu nienależącym do zbioru  $M$ , co obrazuje wykres na rysunku 2.

## 4. Ewaluacja

Na podstawie znalezionych obciążeń rozkładów prawdopodobieństwa sformułowano algorytm zwiększający wartość oczekiwaną zysku z gry w zakładach bukmacherskich. Algorytm (3) zawiera zakład o wynik remisowy meczu ze środka sezonu piłkarskiego granego przez zespół, który poprzedni mecz zakończył remisem:

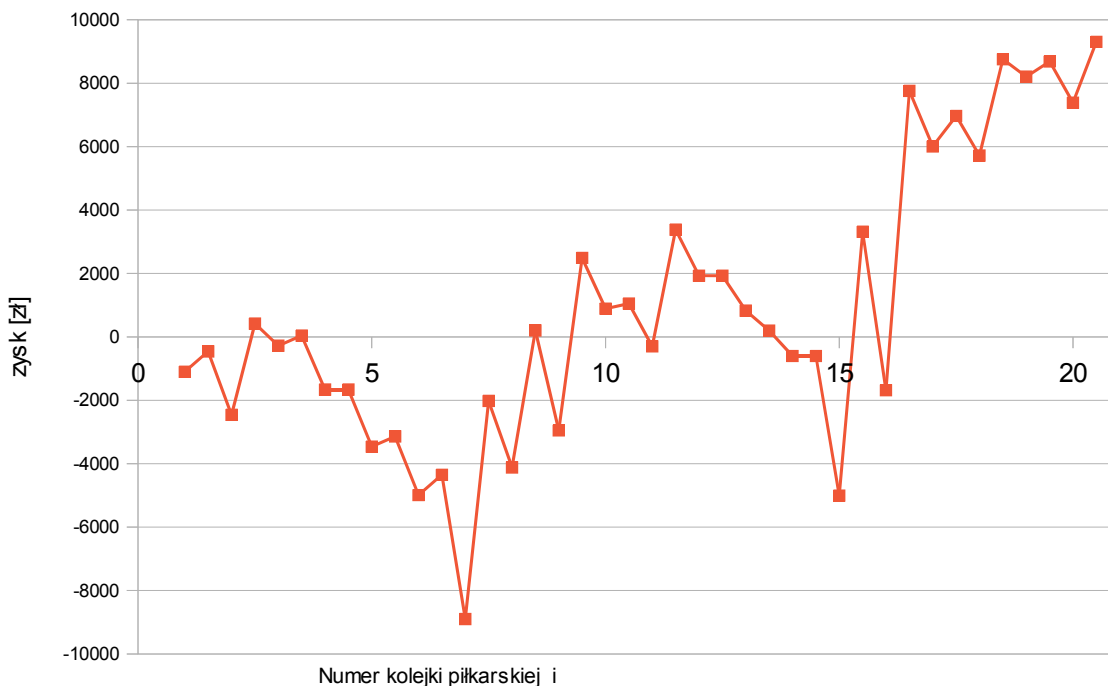
```
l = (J-1)/2
for(i=l/2; i<l; i++):
    for(j=0; j<J; j++):
        if (Wynik(M[j][i-1]==REMIS):
            wnieś_zakład(M[j][i], Kwota)

function wnieś_zakład(Mecz, Kwota):
    bukmacher = wybierzBukmachera(Mecz)
    zakład(Mecz, Kwota, Bukmacher)      (3)
```

Zmienna  $J$  jest liczbą zespołów w lidze piłkarskiej. Funkcja `wybierzBukmachera(Mecz)` przeszukuje oferty pochodzące od różnych kolektur bukmacherskich i wybiera ofertę zakładu o najwyższym kursie, a funkcja `zakład(Mecz, Kwota, Bukmacher)` powoduje zawarcie zakładu u bukmachera oferującego najwyższy kurs meczu. Przedstawiony algorytm został sprawdzony na rzeczywistych środkach finansowych w internetowej kolekturze bukmacherskiej. Po sześciu miesiącach gry wykazał sześćdziesiąt procent rentowności.

Zgodnie z symulacjami przeprowadzonymi na danych historycznych z drugiej ligi włoskiej średnia stopa zwrotu z gry algorytmem wynosi około 25%.

Na rysunku 3 przedstawiono wykres skumulowanego zysku z gry uzyskany z zastosowaniem algorytmu (3) w lidze włoskiej w sezonie 2011-2012.



Rys. 3. Wykres skumulowanego zysku z gry algorytmem [3] w lidze włoskiej w sezonie 2011-2012

Fig. 3. Plot of profit form betting according to algorithm [3] in Serie B in season 2011-2012

Źródło: opracowanie własne.

## 5. Wnioski

Rozkład występowania remisów w meczach piłkarskich jest zbliżony do rozkładu Bernoullego. Prawdopodobieństwo wystąpienia remisu zależy między innymi od fazy, w której znajduje się sezon piłkarski, oraz od wyników zespołu w poprzednich meczach. Zaprezentowany algorytm selekcjonuje mecze, w których występuje zwiększone prawdopodobieństwo remisu. Jego struktura jest bardzo prosta i wykazuje wiele słabości. Możliwa jest modyfikacja algorytmu przez określenie dodatkowych warunków wyboru meczów do wniesienia zakładów bukmacherskich. Zgodnie z symulacjami działania algorytmu istnieje ryzyko, że przez dwa sezony algorytm będzie przynosił stratę, by dopiero w trzecim przynieść zysk. Kolejną wadą algorytmu wynika z giełdowego charakteru zakładów bukmacherskich. Zwiększając kwoty naszych zakładów, obniżamy kurs meczu, obniżając tym samym wartość oczekiwaną zysku z zakładu. Najważniejszym wnioskiem płynącym z wykonanych badań jest wykazanie, że istnieje możliwość wypracowania długoterminowego zysku w zakładach bukmacherskich, gdyż średnia wartość oczekiwana zysku z gry przedstawionym algorytmem jest dodatnia.

## Bibliografia

1. Moore D., McCabe D.: Introduction to the Practice of Statistics 6<sup>th</sup>. W. H. Freeman and Company, New York 2007, p. 509-523.
2. Back D., Bielecki A., Peng, H.: Stochastic methods in finance. Fondazione C.I.M.E., Firenze 2003, p. 105-109.
3. William C., Ingenoso M.: Risk formulas for proportional betting. BJMath Publ. Co., Chicago 2005, p. 8-9.
4. Owen A.: Dynamic bayesian forecasting models of football match outcomes with estimation of the evolution variance parameter. „IMA Journal of Management Mathematics”, No. 22, 2011, p. 99-113.
5. Hill I.: Association Football and Statistical Inference. „Journal of the Royal Statistical Society”, No. 23, 1974, p. 203-208.
6. Moroney M.: Facts from Figures. Penguin Books, London 1956, s. 147.
7. Reep C., Benjamin B.: Skill and chance in association football. „Journal of the Royal Statistical Society”, No. 17, 1968, p. 24-26.
8. Fahrmeir L., Tutz G.: Dynamic-stochastic models for time-dependent ordered paired comparison systems. „Journal of the American Statistical Association”, 1994, p. 34-36.

## Abstract

The paper presents the idea and concept of soccer betting market analysis and solutions for soccer betting. Two types of probability distribution bias were found. Statistical analysis reveals an interconnection between the final result of a match and results achieved in previous matches. The correlation is connected with the form of the team, which can be determined from previous matches. Moreover it was found that the probability of a draw result in a match is unequal during the whole football season. In the middle of the season the quantity of draws is higher than in other periods. Discovered dependencies increase the probability of a draw result in a match. An algorithm of increased probability of draw result anticipation was proposed. The algorithm, tested on the betting market, gave positive results. The research shows that it is possible to achieve and preserve a long-term profit in soccer betting.