



Politechnika Wrocławska

Wydział Matematyki

Kierunek studiów: Matematyka stosowana

Specjalność: –

Praca dyplomowa – inżynierska

## ZASTOSOWANIE STOCHASTYCZNEJ OPTYMALIZACJI DO GIER CZĘŚCIOWO OBSERWOWALNYCH

Aleksander Jakóbczyk

słowa kluczowe:  
tutaj podajemy najważniejsze słowa kluczowe (łącznie nie powinny być dłuższe niż 150 znaków).

krótkie streszczenie:

Celem rozprawy jest wyznaczenie nieoczekiwanych strategii w grach częściowo obserwowalnych za pomocą metod stochastycznej optymalizacji. Praca będzie opierać się na wynikach Cauwet i Teytauda z 2018 roku, w których przedstawili nieoczekiwane strategie dla kilku klasycznych gier oraz kilku metod optymalizacji. Podjęta zostanie próba odtworzenia oraz rozszerzenia wyników na kolejną grę. Przeprowadzona zostanie analiza porównawcza dla różnych metod optymalizacji.

Opiekun pracy dyplomowej	dr inż. Andrzej Giniewicz	.....	.....
	Tytuł/stopień naukowy/imię i nazwisko	ocena	podpis

*Do celów archiwalnych pracę dyplomową zakwalifikowano do:\**

*a) kategorii A (akta wieczyste)*

*b) kategorii BE 50 (po 50 latach podlegające ekspertyzie)*

*\* niepotrzebne skreślić*

pieczęćka wydziałowa

Wrocław, rok 2022





Wrocław University  
of Science and Technology

Faculty of Pure and Applied Mathematics

Field of study: Mathematics

Specialty: Theoretical Mathematics

Engineering Thesis

## TYTUŁ PRACY DYPLOMOWEJ W JĘZYKU ANGIELSKIM

Aleksander Jakóbczyk

keywords:

tutaj podajemy najważniejsze słowa kluczowe w języku angielskim (łącznie nie powinny być dłuższe niż 150 znaków)

short summary:

Tutaj piszemy krótkie streszczenie pracy w języku angielskim (nie powinno być dłuższe niż 530 znaków).

Supervisor	dr inż. Andrzej Giniewicz	.....	.....
	Title/degree/name and surname	grade	signature

*For the purposes of archival thesis qualified to:\**

*a) category A (perpetual files)*

*b) category BE 50 (subject to expertise after 50 years)*

*\* delete as appropriate*

stamp of the faculty

Wrocław, 2022



# Spis treści

<b>Wstęp</b>	<b>3</b>
<b>1 Rozdział pierwszy</b>	<b>5</b>
1.1 Podrozdział pierwszy . . . . .	5
1.2 Podrozdział drugi . . . . .	6
<b>2 Definicje, lematy, twierdzenia, przykłady i wnioski</b>	<b>7</b>
2.1 Typy gier . . . . .	7
2.2 Definicje i Oznaczenia . . . . .	7
2.2.1 Strategie proste . . . . .	8
2.2.2 Strategie mieszane . . . . .	8
2.3 Twierdzenia . . . . .	9
2.4 Problem porównań wielokrotnych . . . . .	10
2.5 Algorytmy wyścigowe . . . . .	11
2.5.1 Limited empirical Bernstein race algorithm (LEBR) . . . . .	11
2.5.2 Improved LEBR (ILEBR) . . . . .	12
2.5.3 Bernstein race without maximum race length . . . . .	17
<b>Podsumowanie</b>	<b>19</b>
<b>Dodatek</b>	<b>21</b>



# Wstęp

We wstępie zapowiadamy, o czym będzie praca. Próbujemy zachęcić czytelnika do dalszej lektury, np. krótko informując, dlaczego wybraliśmy właśnie ten temat i co nas w nim zainteresowało.





# Rozdział 1

## Rozdział pierwszy

Tabela 1.1 przedstawia przykładową tabelę. Do tworzenia tabeli służą m.in. środowiska `tabular` oraz `table`. Istnieje możliwość numeracji dwustopniowej, gdzie pierwsza cyfra oznacza numer rozdziału, a druga – kolejny numer tabeli w tym rozdziale. Tytuł powinien znajdować się centralnie nad tabelą, 12 pkt odstępu od tekstu zasadniczego nad i pod tabelą wraz z tytułem. Jeśli tabela jest cytowana – należy podać centralnie pod tabelą źródło jej pochodzenia, np. opracowanie własne, opracowano na podstawie danych z GUS.

Tabela 1.1: Podstawowa Tabela

Państwo	PKB (w milionach USD )	Stopa bezrobocia
Stany Zjednoczone	75 278 049	4,60%
Chiny	11 218 281	4,10%
Japonia	4 938 644	3,10%
Niemcy	3 466 639	6,00%
Wielka Brytania	2 629 188	4,60%

*Źródło: opracowanie własne*

### 1.1 Podrozdział pierwszy

Tabela 1.2: Podstawowa Tabela

Państwo	PKB (w milionach USD )	Stopa bezrobocia
Stany Zjednoczone	75 278 049	4,60%
Chiny	11 218 281	4,10%
Japonia	4 938 644	3,10%
Niemcy	3 466 639	6,00%
Wielka Brytania	2 629 188	4,60%

*Źródło: opracowanie własne*

## 1.2 Podrozdział drugi

Rysunki do pracy dyplomowej należy wstawiać w sposób podobny do wstawiania tabel, z zasadniczą różnicą polegającą na tym, że podpis powinno umieszczać się centralnie pod rysunkiem, a nie powyżej niego. Numeracja i sposób cytowania pozostają bez zmian, przy czym tabele i rysunki nie mają numeracji wspólnej, np. po Tabeli 1.2 występuje Rysunek 1.1 (o ile jest to pierwszy rysunek rozdziału pierwszego), a nie Rysunek 1.3.



Rysunek 1.1: Podstawowy Rysunek

# Rozdział 2

## Definicje, lematy, twierdzenia, przykłady i wnioski

Celem algorytmów, których będziemy wykorzystywać, jest znalezienie optymalnej strategii w grach częściowo obserwowalnych. Zdefiniujmy zatem podstawowe pojęcia potrzebne nam do tego, aby matematycznie opisać czym jest gra i czym jest strategia optymalna. W tym celu wprowadzi kilka podstawowych definicji.

### 2.1 Typy gier

Podstawową kategorią, na jaką możemy podzielić gry, jest podział ze względu na czas, w którym gracze podejmują decyzje:

**Definicja 2.1** (Gra w postaci strategicznej). Jest to typ gry, w której gracze podejmują decyzje w tym samym momencie.

**Definicja 2.2** (Gra w postaci ekstensywnej). Jest to typ gry, w której gracze podejmują decyzje we wcześniej ustalonej kolejności.

Przykładami gier w postaci strategicznej są gry papier-kamień-nożyce, oszust czy też mora. Natomiast przykładami gier w postaci ekstensywnej są szachy, warcaby oraz go.

Gry możemy również dzielić ze względu na posiadaną wiedzę.

**Definicja 2.3** (Gra z kompletną informacją). Jest to typ gry, w której gracze mają informacje o możliwych przyszłych wynikach gry i o zbiorach możliwych strategii.

**Definicja 2.4** (Gra częściowo obserwowalna). Jest to przeciwieństwo gier z kompletną informacją.

Przykładami gier z kompletną informacją są szachy, warcaby oraz go. Natomiast przykładami gier częściowo obserwowanymi są wszelkie gry posiadające w rozgrywce pewien element losowy takie jak rzut kostką czy też dobieranie kart.

Istnieje jeszcze wiele innych podziałów gier ze względu na kategorie takie jak liczba graczy, ze względu na zbiory dostępnych akcji, ze względu na możliwość tworzenia koalicji i wiele innych.

### 2.2 Definicje i Oznaczenia

Wprowadźmy podstawowe oznaczenia potrzebne nam do tego, aby móc zdefiniować czym jest strategia optymalna.

### 2.2.1 Strategie proste

- $N = \{1, 2, \dots, n\}$  – zbiór graczy,
- $A_i$ ,  $i \in N$  – niepusty zbiór strategii czystych gracza  $i$ ,
- $m_i = |A_i|$  – liczba strategii gracza  $i$ ,
- $A = \prod_{i \in N} A_i$  – zbiór wszystkich strategii gry,
- $u_i : A \rightarrow \mathbb{R}$  – funkcja wypłaty gracza  $i$ ,
- $a = (a_1, a_2, \dots, a_n) = (a_i)_{i \in N}$ ,  $a_i \in A_i$  – profil gry w strategiach czystych,
- $u_i(a) = u_i(a, a_{-i})$  – wypłata gracza  $i$  z profilu  $a$ ,
- $a_{-i} = (a_i)_{i \in N \setminus \{i\}}$ , – profil wszystkich strategii poza strategią gracza  $i$ .

**Definicja 2.5** (Gra strategiczna). Grą strategiczną nazywamy trójkę  $GS = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$ .

**Definicja 2.6** (Równowaga Nasha w strategiach czystych gry strategicznej). Równowaga Nasha w strategiach czystych gry strategicznej jest to taki profil gry  $a^* = (a_1^*, a_2^*, \dots, a_n^*) \in A$ , takim, że:

$$\forall_{i \in N} \forall_{a_i \in A_i} u_i(a_i^*, a_{-i}^*) \geq u_i(a_i, a_{-i}^*)$$

Zatem jest to profil gry, w którym istnieje strategia czystej dająca niegorsze wyniki od dowolnej innej strategii czystej. Okazuje się jednak, że taki stan nie zawsze istnieje w strategiach czystych, np. w grze kamień papier nożyce strategia grania tylko kamienia daje gorsze rezultaty przeciwko graniu tylko papieru. Podobnie ze strategią grania tylko nożyc i grania tylko papieru.

### 2.2.2 Strategie mieszane

**Definicja 2.7** (Strategia mieszana). Strategia mieszana  $\sigma_i$  gracza  $i$  w grze strategicznej  $GS = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$ . Nazywamy rozkład prawdopodobieństwa na zbiorze strategii czystych  $A_i$ :

$$\sigma_i = (\sigma_{i1}, \sigma_{i2}, \dots, \sigma_{im_i})$$

gdzie  $\sigma_{ik}$  oznacza prawdopodobieństwo, że gracz  $i$  zagra strategią czystą  $k \in A_i$ .

Wprowadźmy dodatkowe oznaczenia:

**Fakt 2.8.** *Strategia czysta jest szczególnym przypadkiem strategii mieszanej, w którym prawdopodobieństwo zagrania jednej z dostępnych strategii wynosi 1.*

Wprowadźmy dodatkowe oznaczenia:

- $\Sigma_i = \left\{ \sigma_i : A_i \rightarrow [0, 1], \sum_{k=1}^n \sigma_{ik} = 1, \sigma_{ik} \geq 0 \right\}$  – zbiór strategii mieszanych gracza  $i$ ,
- $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$  – profil gry,

- $u_i(\sigma) = u_i(\sigma_i, \sigma_{-i})$  – wypłata gracza  $i$  z profilu  $\sigma$ ,
- $\sigma_{-i} = (\sigma_j)_{j \in N \setminus \{i\}}$ , – profil wszystkich strategii poza strategią gracza  $i$ .

**Definicja 2.9** (Równowaga Nasha w strategiach mieszanej gry strategicznej). Profil gry strategicznej  $\sigma_i^*$  jest Równowagą Nasha gdy:

$$\forall_{i \in N} \forall_{\sigma_i \in \Sigma_i} u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i, \sigma_{-i}^*)$$

Równowaga Nasha interpretujemy jako taki profil gry, w którym żaden z graczy nie opłaca się zmieniać swojej strategii, ponieważ nie skutkuje to zwiększeniem swoich zysków.

## 2.3 Twierdzenia

Algorytmy wykorzystywane w poniższej pracy oparte są o dwa twierdzenia a dokładniej o szczególne przypadki wynikające z nierówności 2.10 i 2.12:

**Twierdzenie 2.10** (Nierówność Hoffdinga). *Niech  $X_1, X_2, \dots, X_t$  będzie ciągiem niezależnych zmiennych losowych (i.i.d.) takim, że  $a_i \leq X_i \leq b_i$ , wtedy:*

$$S_t = \sum_{i=1}^t X_i, \quad c_i = b_i - a_i,$$

$$P(|S_t - \mathbb{E}(S_t)| \geq \epsilon) \leq 2 \exp \left( -\frac{2\epsilon^2}{\sum_{i=1}^n c_i^2} \right)$$

**Lemat 2.11.** *Niech  $X_1, X_2, \dots, X_t$  będzie ciągiem niezależnych zmiennych losowych (i.i.d.) takim, że  $0 \leq X_i \leq 1$ , wtedy:*

$$\bar{X}_t = \frac{S_t}{t}, \quad \mu = \mathbb{E}(X_i), \quad P(|\bar{X}_t - \mu| \leq \epsilon) = 1 - \delta,$$

$$\epsilon \leq \sqrt{\frac{\ln(2/\delta)}{2t}}$$

**Twierdzenie 2.12** (Empiryczna nierówność Bernsteina). *Niech  $X_1, X_2, \dots, X_t$  będzie ciągiem niezależnych zmiennych losowych (i.i.d.) takim, że  $a \leq X_i \leq b$ , wtedy:*

$$P(|\bar{X}_t - \mu| \geq \epsilon) \leq \delta, \quad \bar{\sigma}_t^2 = \frac{1}{t} \sum_{i=1}^t (X_i - \bar{X}_t)^2,$$

$$|\bar{X}_t - \mu| \leq \bar{\sigma}_t \sqrt{\frac{2 \ln(3/\delta)}{t}} + \frac{3R \ln(3/\delta)}{t}$$

**Lemat 2.13.** *Niech  $X_1, X_2, \dots, X_t$  będzie ciągiem niezależnych zmiennych losowych (i.i.d.) takim, że  $0 \leq X_i \leq 1$ , wtedy:*

$$P(|\bar{X}_t - \mu| \leq \epsilon) = 1 - \delta,$$

$$\epsilon \leq \bar{\sigma}_t \sqrt{\frac{2 \ln(3/\delta)}{t}} + \frac{3 \ln(3/\delta)}{t}$$

## 2.4 Problem porównań wielokrotnych

Założmy, że z prawdopodobieństwem  $1 - \delta$  chcemy wiedzieć który z 2 graczy,  $p_1$  i  $p_2$  jest lepszy. W tym celu będziemy przeprowadzać testy statystyczne, dla których prawdopodobieństwo pomyłki  $k$ -tego testu wynosi  $\delta_k$ , aż do momentu, gdy jeden z graczy wygra przeważającą ilość razy. Wtedy po przeprowadzeniu  $n$  takich testów

$$P(\text{Chociaż jeden z } n \text{ testów się pomylił}) \stackrel{(*)}{\leq} \sum_{k=1}^n \delta_k \implies$$

$$P(\text{Żaden test się nie pomylił}) \leq 1 - \sum_{k=1}^n \delta_k$$

$$(*) \quad P(X + Y) \leq P(X) + P(Y)$$

Oznacza to, że musimy wprowadzić pewną korektę, aby ostateczne prawdopodobieństwo popełnienia błędu było mniejsze niż  $\delta$ . Możemy wprowadzić jedna z dwóch poprawek:

2.4.1 Niech  $n$  będzie maksymalną liczbą testów jaką pozwalamy wykonać, aby wyznaczyć lepszego gracza. Wtedy  $\delta_k = \frac{\delta}{n}$ .

2.4.2 Niech  $\delta_k$  spełnia nierówność  $\delta \geq \sum_{k=1}^{\infty} \delta_k$ . Wtedy niezależnie od ilości przeprowadzonych testów, ostateczne prawdopodobieństwo pomyłki będzie nie większe niż  $\delta$ .

## 2.5 Algorytmy wyścigowe

Jednymi z algorytmów stosującymi dane korekty są algorytmy wyścigowe (Racing Algorithms). Dwoma najpopularniejszymi typami algorytmów racingowych są "Hoffding race" oraz "Bernstein race". Oparte są one odpowiednio o Twierdzenie 2.10 i Twierdzenie 2.13 oraz korektę 2.4.2. Mają one jednak pewną wadę, mogą zajmować one dużą ilość czasu, a w momencie, gdy poziom umiejętności porównywanych graczy jest sobie równy (prawdopodobieństwo wygranej wynosi 50%), wtedy z prawdopodobieństwem równym  $1 - \delta$  algorytm się nigdy nie zatrzyma.

W celu poradzenia sobie z problemami, jakie wiążą się z klasycznymi algorytmami wyścigowymi, wprowadzimy tzw. Limited Racing algorytm. Założymy zatem dodatkowy warunek, który mówi, że przerywamy działanie algorytmu w momencie, gdy empiryczna wartość oczekiwana z prawdopodobieństwem większym bądź równym  $1 - \delta$  jest znana z dokładnością co do zadanego  $\epsilon$ . W poniższej pracy będziemy przyjmować  $\epsilon = 0.01$  oraz  $\delta = 0.05$ .

### 2.5.1 Limited empirical Bernstein race algorithm (LEBR)

Klasyczny algorytm racingowy opiera się na szeregu  $\epsilon_t$  który posiada właściwość że zdarzenie  $\mathcal{E} = \{|\bar{X}_t - \mu| \leq \epsilon_t, t \in \mathbb{N}^+\}$  występuje z prawdopodobieństwem nie mniejszym niż  $1 - \delta$ . Niech  $\delta_k$  będzie dodatnim szeregiem spełniającym  $\delta \geq \sum_{k=1}^{\infty} \delta_k$  wtedy korzystając z Lematu 2.13

$$\epsilon_t = \bar{\sigma}_t \sqrt{\frac{2 \ln(3/\delta_t)}{t}} + \frac{3 \ln(3/\delta_t)}{t}$$

Ponieważ  $\delta_k$  sumuje się co najmniej do  $\delta$  oraz  $(\bar{X}_t - \epsilon_t, \bar{X}_t + \epsilon_t)$  jest przedziałem ufności dla  $\mu$  o współczynniku ufności  $1 - \delta_t$  otrzymanym z Lematu 2.13, uzyskujemy że zdarzenie  $\mathcal{E}$  występuje z prawdopodobieństwem nie mniejszym niż  $1 - \delta$ . Identyczny rezultat otrzymujemy stosując  $\epsilon_t$  oparte o Lemat 2.11 jednak zmienia się wtedy postać  $\epsilon_t$ . W pracy [1] algorytm opierał się o szereg  $\delta_t = \frac{c\delta_t}{t^2}$ ,  $c = \frac{6}{\pi^2}$ . Pseudo kod jest pokazany jako Algorytm 1 omówmy zarys metody. Rozgrywamy  $t$  gier, i wyznaczamy granice górną  $UB = \min_{1 \leq k \leq t} (\bar{X}_k + c_k)$  oraz granice dolną  $LB = \max(0, \max_{1 \leq k \leq t} (\bar{X}_k - c_k))$ . Algorytm kończy się w momencie gdy  $UB - LB < 2\epsilon$ . Otrzymane w ten sposób  $\bar{X}$  z prawdopodobieństwem nie mniejszym niż  $1 - \delta$  wynosi  $\mu$  z dokładnością co zadanego  $\epsilon$ .

---

**Algorithm 1** LEBR

---

**Ensure:** precision  $\epsilon$ , probabilities  $\delta_k$   
 $LB \leftarrow 0, \quad UB \leftarrow \infty, \quad t \leftarrow 0, \quad n \leftarrow 1$   
**while**  $UB - LB > 2\epsilon$  **do**  
     $t \leftarrow t + 1$   
    Obtain  $X_t$   
     $\epsilon_n \leftarrow \bar{\sigma}_n \sqrt{\frac{2 \ln(3/\delta_n)}{n}} + \frac{3 \ln(3/\delta_n)}{n}$   
     $LB \leftarrow \max(LB, \bar{X}_n - \epsilon_n)$   
     $UB \leftarrow \min(UB, \bar{X}_n + \epsilon_n)$   
     $n \leftarrow n + 1$   
**end while**  
**return**  $\bar{X}_t$

---

### 2.5.2 Improved LEBR (ILEBR)

Algorytm 1 opiera się na korekcie 2.4.2 która zakłada możliwie nieskończoną ilość testów. Jednak dołożenie ograniczenia na temat żądanej dokładności rzędu  $\epsilon$  pozwala nam znaleźć maksymalną ilość testów jaką należy wykonać aby z prawdopodobieństwem nie mniejszym niż  $1 - \delta$  empiryczna wartość oczekiwana była równa teoretycznej wartości oczekiwanej z dokładnością co do  $\epsilon$ . Niech  $\delta_k = \frac{\delta}{n_{max}}$  gdzie  $n_{max}$  oznacza maksymalną potrzebną ilość testów. Z Lematu 2.11 niezależnie od odchylenia standardowego zmiennej losowej  $X_i$  jesteśmy w stanie oszacować  $n_{max}$  rozwiązując analitycznie poniższe równie 2.1

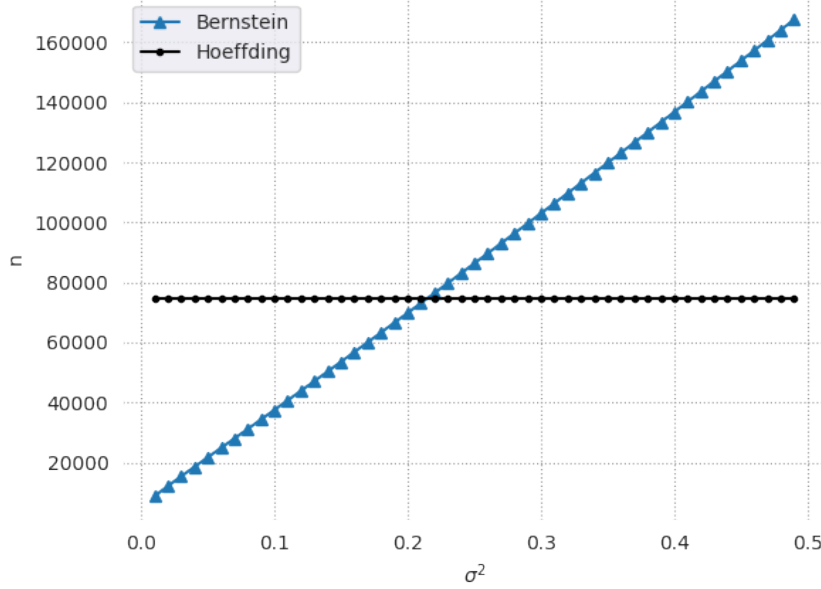
$$\epsilon = \sqrt{\frac{\ln(2n_{max}/\delta)}{2n_{max}}} \quad (2.1)$$

Dla  $\epsilon = 0.01$  i  $\delta = 0.05$  rozwiązaniem numerycznym równania 2.1 jest  $n_{max} = 74539.85$ . Oznacza to że maksymalna ilość gier jaką należy rozegrać między dwoma graczami aby z prawdopodobieństwem nie mniejszym niż  $1 - 0.05$  oszacować prawdopodobieństwo wygrania gry przez pierwszego gracza z dokładniejsi co do 0.01 wynosi  $\lceil n_{max} \rceil = 74540$ .

Podobnie możemy postąpić w przypadku nierówności Bernsteina. Na początku jednak musimy oszacować z góry  $\bar{\sigma}_t$ . Przyjmijmy, że  $X_i$  jest zmienną z rozkładu zero-jedynkowego. Wtedy maksymalna możliwa wariancja dla takiej zmiennej losowej wynosi  $\sigma^2 = 0.5$ . Jest to również maksymalna możliwa wartość dla naszej empirycznej wariancji. Podstawiając wtedy  $\delta_n = \frac{0.05}{n}$ ,  $\epsilon = 0.01$  do Lematu 2.13 otrzymujemy:

$$0.01 \leq \sqrt{\frac{\ln(\frac{3n_{max}}{0.05})}{n_{max}}} + \frac{3 \ln(\frac{3n_{max}}{0.05})}{n_{max}} \quad (2.2)$$

Rozwiązaniem numerycznym równania 2.2 jest  $n_{max} = 170986$ . Jednak granice oparte o nierówność Bernsteina zależą od wariancji. Sprawdźmy zatem jak wygląda  $n_{max}$  w zależności od  $\sigma^2$ .



Rysunek 2.1: Wykres maksymalnej potrzebnej ilości testów w zależności od wariancji zmiennych losowych  $X_i$  w przypadku gdy liczba testów jest równa liczbie gier ( $t = n$ ) dla  $\epsilon = 0.01$  i  $\delta = 0.05$ .

Z Rysunku 2.1 widzimy zatem, że algorytm oparty o nierówność Bernsteina radzi sobie lepiej w przypadku gdy niskiej wariancji. Co więcej niezależnie od ilości wykonywanych testów, dla  $\delta_n = \frac{0.05}{74540}$  maksymalna ilość gier po której z prawdopodobieństwem  $1 - 0.05$  wiemy że  $|\bar{X}_i - \mu| \leq 0.01$  wynosi 74540. Oznacza to że do Algorytmu 1 możemy zastosować funkcje  $\epsilon_k$  opartą o  $\delta_k = \frac{\delta}{n_{max}}$  oraz dołożyć nowy warunek zatrzymania algorytmu gdy ilość wykonanych testów przekroczy  $n_{max}$ . Dotychczas uwzględnione poprawki przedstawione są przy Algorytmu 2.

---

#### Algorithm 2 ILEBR 1

---

**Ensure:** precision  $\epsilon$ , probability  $\delta$   
 $LB \leftarrow 0, \quad UB \leftarrow \infty, \quad t \leftarrow 0, \quad n \leftarrow 1$   
 Find  $n_{max}$  such as  $\epsilon = \sqrt{\frac{\ln(2n_{max}/\delta)}{2n_{max}}}$   
 $\delta_n = \delta/n_{max}$   
**while**  $UB - LB > 2\epsilon$  or  $n < n_{max} + 1$  **do**  
    $t \leftarrow t + 1$   
   Obtain  $X_t$   
    $\epsilon_n \leftarrow \bar{\sigma}_n \sqrt{\frac{2 \ln(3/\delta_n)}{n}} + \frac{3 \ln(3/\delta_n)}{n}$   
    $LB \leftarrow \max(LB, \bar{X}_n - \epsilon_n)$   
    $UB \leftarrow \min(UB, \bar{X}_n + \epsilon_n)$   
    $n \leftarrow n + 1$   
**end while**  
**return**  $\bar{X}_t$

---



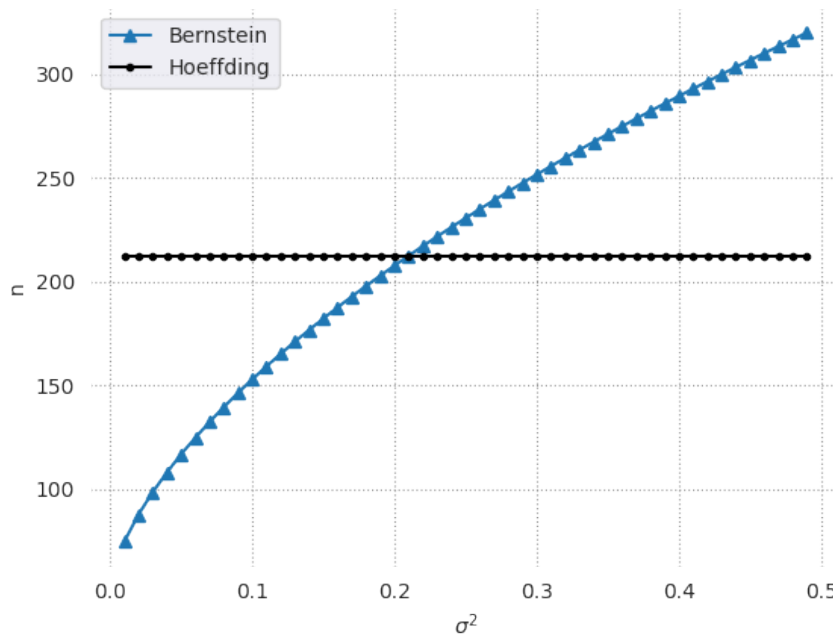
Algorytmy 1 i 2 przeprowadzały testy po każdej rozegranej grze, skutkuje to wytopieniem bardzokiej ilości testów. Zamiast przeprowadzać test po każdej rozegranej grze, niech test odbywa się w momencie gdy liczba rozegranych gier będzie równa pewnej funkcji zależnej od ilości przeprowadzonych testów. Na podstawie artykułu [2] wiemy że funkcja która pozwoli nam na polepszenie wyników jest  $t = n^2$ . Przeprowadźmy identyczną procedurę jak w przypadku równań 2.1 i 2.2. Tym razem jednak zamiast stosować  $t = n$  do Lematów 2.11 i 2.13 zastosujemy  $t = n^2$ .

W przypadku granicy opartej o nierówność Hoffdina otrzymujemy

$$\epsilon = \sqrt{\frac{\ln(2n_{max}/\delta)}{2n_{max}^2}} \implies t_{max} = \lceil n_{max} \rceil^2 = \lceil 213 \rceil^2 = 45369. \quad (2.3)$$

W przypadku granicy opartej o nierówność Bernsteina otrzymujemy

$$0.01 \leq \sqrt{\frac{\ln(\frac{3n_{max}}{0.05})}{n_{max}^2}} + \frac{3 \ln(\frac{3n_{max}}{0.05})^2}{n_{max}} \implies t_{max} = \lceil n_{max} \rceil^2 = \lceil 324 \rceil^2 = 104976. \quad (2.4)$$



Rysunek 2.2: Wykres maksymalnej potrzebnej ilości testów w zależności od wariancji zmiennych losowych  $X_i$  w przypadku gdy liczba rozegranych gier jest równa liczbie przeprowadzonych testów do kwadratu ( $t = n^2$ ) dla  $\epsilon = 0.01$  i  $\delta = 0.05$ .

Porównując ze sobą uzyskane wartości z wzoru 2.1 oraz 2.3 widzimy że udało się zmniejszyć maksymalną ilość gier jakie należy wykonać z 74540 do 45369. Wprowadźmy zatem nowo uzyskane poprawki tworząc Algorytm 3.

---

**Algorithm 3** ILEBR 2

---

**Ensure:** precision  $\epsilon$ , probability  $\delta$   
 $LB \leftarrow 0, \quad UB \leftarrow \infty, \quad t \leftarrow 0, \quad n \leftarrow 1$   
Find  $n_{max}$  such as  $\epsilon = \sqrt{\frac{\ln(2n_{max}/\delta)}{2n_{max}^2}}$   
 $\delta_n = \delta/n_{max}$   
**while**  $UB - LB > 2\epsilon$  or  $n < n_{max} + 1$  **do**  
    **repeat**  
         $t \leftarrow t + 1$   
        Obtain  $X_t$   
    **until**  $t = n^2$   
     $\epsilon_n \leftarrow \bar{\sigma}_n \sqrt{\frac{2\ln(3/\delta_n)}{n^2}} + \frac{3\ln(3/\delta_n)}{n^2}$   
     $LB \leftarrow \max(LB, \bar{X}_n - \epsilon_n)$   
     $UB \leftarrow \min(UB, \bar{X}_n + \epsilon_n)$   
     $n \leftarrow n + 1$   
**end while**  
**return**  $\bar{X}_t$

---

Do tej pory Wszystkie stosowane algorytmy wyznaczały nam prawdopodobieństwo wygranej pierwszego gracza. Jednak nam nie zależny na tym dokładnie znać prawdopodobieństwo wygranej graczy. Głównym celem algorytmów jest odnalezienie lepszego gracza. Pozwala nam to modyfikacje Algorytmu 3 o nowe warunki zatrzymania.

---

**Algorithm 4:** ILEBR\* 1

---

**Ensure:** precision  $\epsilon$ , probability  $\delta$   
 $LB \leftarrow 0, \quad UB \leftarrow \infty, \quad t \leftarrow 1, \quad n \leftarrow 0$   
Find  $n_{max}$  such as  $\epsilon = \sqrt{\frac{\ln(2n_{max}/\delta)}{2n_{max}^2}}$   
 $\delta_n = \delta/n_{max}$   
**while**  $(UB - LB > 2\epsilon$  or  $n < n_{max} + 1)$  and  $(BU > 0.5$  or  $LB < 0.5)$  **do**  
    **repeat**  
         $t \leftarrow t + 1$   
        Obtain  $X_t$   
    **until**  $t = n^2$   
     $\epsilon_n \leftarrow \bar{\sigma}_n \sqrt{\frac{2\ln(3/\delta_n)}{n^2}} + \frac{3\ln(3/\delta_n)}{n^2}$   
     $LB \leftarrow \max(LB, \bar{X}_n - \epsilon_n)$   
     $UB \leftarrow \min(UB, \bar{X}_n + \epsilon_n)$   
     $n \leftarrow n + 1$   
**end while**  
**if**  $LB > 0.5$  **then**  
    **return**  $p_1$  win  
**else if**  $UB < 0.5$  **then**  
    **return**  $p_2$  win  
**else if**  $\bar{X}_n > 0.5$  **then**  
    **return**  $p_1$  win  
**else**  
    **return**  $p_2$  win  
**end if**

---

Do tej pory staraliśmy się ograniczyć z góry maksymalną ilość rozgrywanych gier. Jednak oczywistym faktem jest że nie ma sensu testowania który z graczy jest lepszy w monecie gdy ilość rozegranych gier jest niewystarczająco mały. Jednak jak wyznaczyć nasze minimalne  $t$  po którym zaczniemy testować? Algorytm oparte o nierówność Bernsteina najszybciej wyznaczają wynik w przypadku gdy wariancja naszej zmiennej losowej wynosi 0 (jeden z graczy zawsze wygrywa). Dla Algorytmu ?? interesuje nas monet od którego będziemy w stanie rozróżnić kiedy dolna bać górna granica przekroczy  $\frac{1}{2}$ . Zatem oszacujmy  $n_{min}$  korzystając z Lematu 2.13 dla  $\bar{\sigma}_t = 0$  i  $t = n^2$ .

$$0.5 \leq \frac{3 \ln(\frac{3n_{min}}{0.05})}{n_{min}^2} \implies t_{min} = \lceil n_{min}^2 \rceil = \lceil 5.93749 \rceil^2 = 36.$$

Oznacza to, że przy  $\epsilon = 0.01$  i  $\delta = 0.05$  w przypadku gdy jeden graczy wygrywa zawsze, będziemy w stanie to stwierdzić nie prędzej niż po 36 grach. Na tej podstawie wyznaczmy nowe  $n_{max}$  uwzględniając pomijanie pierwsze niepotrzebne testy.

$$0.01 = \sqrt{\frac{\ln(2n_{max}/0.05)}{2(n+6)^2_{max}}} \implies t_{max} = \lceil (n_{max} + 6) \rceil^2 = \lceil (207 + 6) \rceil^2 = 45369. \quad (2.5)$$

Chociaż równanie 2.5 nie zmniejszyło nam maksymalnej liczby testów jakie należy wykonać to pozwoliło nam ono na zmniejszenie  $\epsilon_n$ . Oznacza to że stosując odroczenie pierwszych testów, pozwala nam na dokładniejsze oszacowanie naszej górnej i dolnej granicy w stosowanych algorytmach.

---

**Algorithm 5:** ILEBR\* 2

---

**Ensure:** precision  $\epsilon$ , probability  $\delta$

$LB \leftarrow 0, \quad UB \leftarrow \infty, \quad t \leftarrow 1, \quad n \leftarrow 0$

Find  $n_{max}$  such as  $\epsilon = \sqrt{\frac{\ln(2n_{max}/\delta)}{2(n+6)^2_{max}}}$

$\delta_n = \delta/n_{max}$

**while** ( $UB - LB > 2\epsilon$  or  $n < n_{max} + 1$ ) and ( $BU > 0.5$  or  $LB < 0.5$ ) **do**

**repeat**

$t \leftarrow t + 1$

        Obtain  $X_t$

**until**  $t = n^2$

$\epsilon_n \leftarrow \bar{\sigma}_n \sqrt{\frac{2 \ln(3/\delta_n)}{(n+6)^2} + \frac{3 \ln(3/\delta_n)}{(n+6)^2}}$

$LB \leftarrow \max(LB, \bar{X}_n - \epsilon_n)$

$UB \leftarrow \min(UB, \bar{X}_n + \epsilon_n)$

$n \leftarrow n + 1$

**end while**

**if**  $LB > 0.5$  **then**

**return**  $p_1$  win

**else if**  $UB < 0.5$  **then**

**return**  $p_2$  win

**else if**  $\bar{X}_n > 0.5$  **then**

**return**  $p_1$  win

**else**

**return**  $p_2$  win

**end if**

---

### 2.5.3 Bernstein race without maximum race length

Wykorzystując Lemat 2.13 oraz korektę 2.4.2 z  $\delta_k = \frac{c\delta}{k^2}$ ,  $c = \frac{6}{\pi^2}$  otrzymujemy, że dla ciągu  $X_1, X_2, \dots, X_t$  i.i.d. zmiennych losowych takim, że  $0 \leq X_i \leq 1$

$$\epsilon_{t,k} \leq \bar{\sigma}_t \sqrt{\frac{2 \ln(3/\delta_k)}{t}} + \frac{3 \ln(3/\delta_k)}{t} = \bar{\sigma}_t \sqrt{\frac{2 \ln(\frac{k^2 \pi^2}{2\delta})}{t}} + \frac{3 \ln(\frac{k^2 \pi^2}{2\delta})}{t}$$

Wtedy  $e_{t,k}$  interpretujemy jako maksymalną różnicę między empiryczną a teoretyczną wartością oczekiwaną po przeprowadzeniu  $k$  testów i rozegraniu  $t$  gier, z prawdopodobieństwem pomyłki równym

**Lemat 2.14.** *Niech liczba rozegranych gier będzie funkcją zależną od  $k$  ( $f(k) = t$ ) oraz niech  $\delta_k = \frac{\delta}{g(k)}$  gdzie  $\delta \geq \sum_{k=1}^{\infty} \frac{\delta}{g(k)}$  i  $\ln(g(k)) \in o(f(k))$ . Wtedy  $\lim_{k \rightarrow \infty} e_{f(k),k} = 0$*

*Dowód Faktu 2.15.*

$$0 \leq X_i \leq 1 \implies \bar{\sigma}_t^2 \leq \frac{1}{2}$$

$$\epsilon_{f(k),k} \leq \sqrt{\frac{\ln(\frac{3g(k)}{\delta})}{f(k)}} + \frac{3 \ln(\frac{3g(k)}{\delta})}{f(k)}$$

Z założeń wiemy że  $\ln(g(k)) \in o(f(k))$ , zatem

$$\lim_{k \rightarrow \infty} \frac{\ln(\frac{3f(k)}{\delta})}{f(k)} = 0$$

Co ostatecznie z tw. o trzech ciągach daje nam

$$0 \leq \lim_{k \rightarrow \infty} e_{f(k),k} \leq 0 \implies \lim_{t \rightarrow \infty} e_{f(k),k} = 0$$

□

Lemat 2.15 pozwala nam na to aby ograniczyć ilość przeprowadzanych testów. Ponieważ podejście oparte o testowanie który z graczy jest lepszy po każdej rozegranej grze prowadzi do nadmierowej ilości wykonywanych testów. Przykładem funkcjami jakimi użyć do Lematu 2.15 są  $f(k) = k^2$ ,  $g(k) = \frac{6/\pi^2}{k^2}$ . Teaki dobór funkcji oznacza że  $k$ -ty test odbywa się gdy liczba przeprowadzonych gier wynosi  $k^2$ .

# Podsumowanie

Podsumowanie w pracach matematycznych nie jest obligatoryjne. Warto jednak na zakończenie krótko napisać, co udało nam się zrobić w pracy, a czasem także o tym, czego nie udało się zrobić.



# Dodatek

Dodatek w pracach matematycznych również nie jest wymagany. Można w nim przedstawić np. jakiś dłuższy dowód, który z pewnych przyczyn pominęliśmy we właściwej części pracy lub (np. w przypadku prac statystycznych) umieścić dane, które analizowaliśmy.





# Bibliografia

- [1] CAUWET, M.-L., TEYTAUD, O. Surprising strategies obtained by stochastic optimization in partially observable games. In *2018 IEEE Congress on Evolutionary Computation (CEC)* (2018), IEEE, pp. 1–8.
- [2] HEIDRICH-MEISNER, V., IGEL, C. Non-linearly increasing resampling in racing algorithms. In *European Symposium on Artificial Neural Networks* (2011), Evere, Belgium: d-side publications, pp. 465–470.