



Využití elo systému k předpovědi výsledků šachových turnajů

Diplomová práce

Studijní program:

N0114A300106 Učitelství pro střední školy a 2. stupeň základních škol

Studijní obory:

Matematika
Dějepis

Autor práce:

Bc. Filip Zdražil

Vedoucí práce:

Mgr. Martin Schindler, Ph.D.
Katedra aplikované matematiky





Zadání diplomové práce

Využití elo systému k předpovědi výsledků šachových turnajů

Jméno a příjmení: **Bc. Filip Zdražil**
Osobní číslo: P21000941
Studijní program: N0114A300106 Učitelství pro střední školy a 2. stupeň základních škol
Specializace: Matematika
Dějepis
Zadávací katedra: Katedra aplikované matematiky
Akademický rok: **2020/2021**

Zásady pro vypracování:

Student naváže na svou bakalářskou práci, ve které popsal vybrané vlastnosti ratingu elo v šachovém prostředí. V této práci se zaměří na to, jak může vhodná predikce výsledku jednotlivých partií založená na elo ratingu hráčů i dalších faktorech pomoci odhadnout výsledek celého turnaje. K tomu budou použity především kombinatorické nebo simulační (např. bootstrap příp. jiné resampling) metody s využitím vhodného softwaru (zejména R). Důraz bude položen na využitelnost výstupů v konkrétních šachové praxi.

Rozsah grafických prací:
Rozsah pracovní zprávy:
Forma zpracování práce:
Jazyk práce:

tištěná/elektronická
Čeština



Seznam odborné literatury:

ANDĚL, J. Statistické metody. Matfyzpress, 2007.
ANDĚL, J. Základy matematické statistiky. Matfyzpress, 2007.
CARSEY, T. M., HARDEN, J. J. Monte Carlo Simulation and Resampling Methods for Social Science. SAGE Publications, 2013.
EFRON, Bradley, TIBSHIRAMI, Robert J. An introduction to the bootstrap. Chapman & Hall, 1993.
ELO, A. The Rating of Chess Players Past and Present. Arco. 1978.
ZADRAŽIL, F. Elo systém očima matematické statistiky. Technická univerzita v Liberci, 2020.

Vedoucí práce:

Mgr. Martin Schindler, Ph.D.
Katedra aplikované matematiky

Datum zadání práce:

2. března 2021

Předpokládaný termín odevzdání:

30. dubna 2022

prof. RNDr. Jan Pícek, CSc.
děkan

L.S.

doc. RNDr. Miroslav Koucký, CSc.
vedoucí katedry

V Liberci dne 2. března 2021

Prohlášení

Prohlašuji, že svou diplomovou práci jsem vypracoval samostatně jako původní dílo s použitím uvedené literatury a na základě konzultací s vedoucím mé diplomové práce a konzultantem.

Jsem si vědom toho, že na mou diplomovou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci nezasahuje do mých autorských práv užitím mé diplomové práce pro vnitřní potřebu Technické univerzity v Liberci.

Užiji-li diplomovou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti Technickou univerzitu v Liberci; v tomto případě má Technická univerzita v Liberci právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Současně čestně prohlašuji, že text elektronické podoby práce vložený do IS/STAG se shoduje s textem tištěné podoby práce.

Beru na vědomí, že má diplomová práce bude zveřejněna Technickou univerzitou v Liberci v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů.

Jsem si vědom následků, které podle zákona o vysokých školách mohou vyplývat z porušení tohoto prohlášení.

12. července 2022

Bc. Filip Zdražil

Poděkování

Děkuji svému vedoucímu práce panu Martinu Schindlerovi, PhD. Má velkou zásluhu na tom, že jsem se při psaní diplomové práce něco naučil, a hlavně, že mě tato činnost těšila.

Anotace

Využití elo systému k předpovědi výsledků šachových turnajů

klíčová slova: Arpád Elo, elo systém, ratingový systém, švýcarský systém, šachový turnaj, výhoda bílých figur

Tato diplomová práce analyzuje, jaké jsou možnosti předpovězení výsledků šachových turnajů. V návaznosti na bakalářskou práci *Elo systém očima matematické statistiky* jsou zohledňovány nejen ratingy hráčů, ale i další faktory, jakými jsou výhoda bílých figur nebo rozdíly v hráčském chování různě silných hráčů. Na základě těchto informací jsou představeny původní metody, jak odhadovat výsledek šachové partie. Tyto odhady se mohou využít k předpovědi výsledků celých turnajů. Pozornost je věnována výhradně posledním kolům turnajů hraných švýcarským systémem. Důraz je položený na možnosti využití této předpovědi jednotlivými hráči s ohledem na maximalizaci jejich finančních prémie. V příloze je skript funkce v R, která provede takovou analýzu v jakémkoli turnaji daného typu.

Abstract

The use of Elo system for prediction of results of chess tournaments

key words: Arpad Elo, elo system, rating system, swiss tournament, chess tournament, white pieces advantage

This diploma thesis analyzes the possibilities of predicting of the chess tournaments results. Following the bachelor thesis *Elo system through the eye of mathematical statistics*, ratings of players and also other factors such as the white pieces advantage and differences in the player behavior of players of different strengths are taken into account.

Based on this information, original methods for estimating the result of a chess game are presented. These estimates can be used to predict the results of the entire tournament. Focus is placed on the last rounds of tournaments played by the Swiss system.

Emphasis is placed on the possibility of using this prediction by individual players with regard to maximizing maximize their prize money. A script of function in R, that can perform such an analysis in any tournament of given type, is attached.

Obsah

Obsah	9
Seznam tabulek.....	11
Seznam grafů.....	13
Úvod	14
1 Parametry vstupující do šachových turnajů hraných švýcarským systémem.....	16
1.1 Fungování elo systému.....	16
1.2 Švýcarský systém šachových turnajů a Buchholzovo hodnocení.....	18
1.3 Dělení finančních cen v šachových turnajích	21
1.4 Motivační úloha	22
2 Odhad výsledků partií posledního kola turnaje	24
2.1 Kritéria dobrého odhadu výsledků partií.....	26
2.2 Metody založené na předpokladech elo systému	33
2.3 Srovnání metod založených na předpokladech elo systému.....	37
2.4 Metody založené na turnajové praxi	38
2.5 Srovnání metod založených na turnajové praxi.....	45
2.6 Srovnání všech metod jako celku.....	45
3 Vyhledky hráče před posledním kolem turnaje.....	48
3.1 Očekávané pořadí hráče	48
3.2 Podmíněné očekávané pořadí hráče	49
3.3 Dopad na očekávaný finanční zisk hráče	51
3.4 Význam očekávaných výsledků turnaje a možnosti využití.....	54
Závěr	56
Zdroje.....	58
Přílohy.....	59

A. Návod k používání vlastní funkce v R analyzující turnaje	59
B. Další grafy srovnávající metody 1–8 na turnaji Czech Open A.....	62
C. Srovnání metod 1 až 8 na turnajích Czech Open A, B, C, D a Open Praha.....	63
D. Další grafy srovnávající metody 9–20 na turnaji Czech Open A.....	65
E. Srovnání metod 9 až 20 na turnajích Czech Open A, B, C, D a Open Praha	66
F. Srovnání všech 20 metod na turnajích Czech Open A, B, C, D a Open Praha v letech 2015–2019..	67
G. Skript funkce v R analyzující libovolný turnaj	68

Seznam tabulek

Tabulka 1: Příklad počítání pomocného hodnocení s fiktivním soupeřem (virtual opponent).....	20
Tabulka 2: Příklad dělení finančních cen systémem hort.....	22
Tabulka 3: los posledního kola turnaje Czech Open A v roce 2021	22
Tabulka 4: kvalita 1. způsobu odhadování výsledků	33
Tabulka 5: kvalita 2. způsobu odhadování výsledků	33
Tabulka 6: kvalita 3. způsobu odhadování výsledků	34
Tabulka 7: kvalita 4. způsobu odhadování výsledků	34
Tabulka 8: kvalita 5. způsobu odhadování výsledků	35
Tabulka 9: kvalita 6. způsobu odhadování výsledků	35
Tabulka 10: kvalita 7. způsobu odhadování výsledků	36
Tabulka 11: kvalita 8. způsobu odhadování výsledků	37
Tabulka 12: kvalita 9. způsobu odhadování výsledků	38
Tabulka 13: kvalita 10. způsobu odhadování výsledků	39
Tabulka 14: kvalita 11. způsobu odhadování výsledků	39
Tabulka 15: kvalita 12. způsobu odhadování výsledků	40
Tabulka 16: kvalita 13. způsobu odhadování výsledků	41
Tabulka 17: kvalita 14. způsobu odhadování výsledků	41
Tabulka 18: kvalita 15. způsobu odhadování výsledků	42
Tabulka 19: kvalita 16. způsobu odhadování výsledků	43
Tabulka 20: kvalita 17. způsobu odhadování výsledků	43
Tabulka 21: kvalita 18. způsobu odhadování výsledků	44
Tabulka 22: kvalita 19. způsobu odhadování výsledků	44
Tabulka 23: kvalita 20. způsobu odhadování výsledků	44
Tabulka 24: Analýza rozptylu dvojného třídění s interakcemi metod 1 až 20 podle korelace odhadnutých a skutečných umístění elitních hráčů	46

Tabulka 25: Porovnání turnajů podle jejich předvídatelnosti z hlediska korelace odhadnutých a skutečných umístění.....	46
Tabulka 26: Porovnání všech 20 metod z hlediska korelace odhadnutých a skutečných umístění	47
Tabulka 27: los posledního kola turnaje Czech Open A v roce 2021	48
Tabulka 28: Možné zisky dalších hráčů v závislosti na jejich poslední vzájemné partii	53
Tabulka 29: Analýza rozptylu dvojného třídění s interakcemi metod 1 až 8 podle korelace odhadnutých a skutečných umístění	63
Tabulka 30: Srovnání metod 1–8 podle korelací odhadnutých a skutečných umístění	64
Tabulka 31: Analýza rozptylu dvojného třídění s interakcemi metod 1 až 8 podle korelace odhadnutých a skutečných výsledků	64
Tabulka 32: Srovnání metod 1–8 podle korelací odhadnutých a skutečných výsledků.....	64
Tabulka 33: Analýza rozptylu dvojného třídění s interakcemi metod 1 až 8 podle korelace odhadnutých a skutečných umístění elitních hráčů	64
Tabulka 34: Srovnání metod 1–8 podle korelací odhadnutých a skutečných umístění elitních hráčů...	65
Tabulka 35: Analýza rozptylu dvojného třídění s interakcemi metod 9 až 20 podle korelace odhadnutých a skutečných výsledků	66
Tabulka 36: Analýza rozptylu dvojného třídění s interakcemi metod 9 až 20 podle korelace odhadnutých a skutečných umístění	67
Tabulka 37: Analýza rozptylu dvojného třídění s interakcemi metod 9 až 20 podle korelace odhadnutých a skutečných umístění elitních hráčů	67
Tabulka 38: Analýza rozptylu dvojného třídění s interakcemi všech metod podle korelace odhadnutých a skutečných výsledků	67
Tabulka 39: Srovnání všech metod podle korelace odhadnutých a skutečných výsledků	68
Tabulka 40: Analýza rozptylu dvojného třídění s interakcemi všech podle korelace odhadnutých a skutečných umístění.....	68

Seznam grafů

Graf 1: Ela účastníků na zkoumaných turnajích	32
Graf 2: rozložení remíz v závislosti na rozdílu v turnaji Czech Open A 2021 podle metody 7	36
Graf 3: korelace odhadnutých a skutečných umístění hráčů v turnaji Czech Open A 2021 podle metod 1–8	37
Graf 4: počet partií odehraných turnaji Czech Open A v roce 2021 v závislosti na rozdílu el hráčů	40
Graf 5: partie odehrané na turnaji Czech Open A v roce 2021 v závislosti na rozdílu el hráčů a průměru jejich el	42
Graf 6: partie odehrané na turnajích Czech Open A, B, C, D a Open Praha v letech 2015–2019 v závislosti na rozdílu el hráčů a průměru jejich el	43
Graf 7: korelace odhadnutých a skutečných výsledků v turnaji Czech Open A 2021 v metodách 9 až 20	45
Graf 8: umístění hráče Babuly ve 100 nasimulovaných průbězích posledního kola	49
Graf 9: umístění hráče Babuly ve 100 nasimulovaných průbězích posledního kola, když v něm remizuje	50
Graf 10: umístění hráče Babuly ve 100 nasimulovaných průbězích posledního kola, když v něm prohraje	50
Graf 11: zisk hráče Babuly ve 100 nasimulovaných průbězích posledního kola, když v něm remizuje	51
Graf 12: zisk hráče Babuly ve 100 nasimulovaných průbězích posledního kola, když v něm prohraje	52
Graf 13: celkové možné zisky hráče Babuly v závislosti na výsledku jeho poslední partie	52
Graf 14: korelace odhadnutých a skutečných výsledků hráčů v turnaji Czech Open A 2021 podle metod 1–8	62
Graf 15: korelace odhadu umístění elitních hráčů v turnaji Czech Open A 2021 podle metod 1–8	63
Graf 16: korelace odhadnutých a skutečných umístění hráčů v turnaji Czech Open A 2021 podle metod 9–20	65
Graf 17: korelace odhadnutých a skutečných umístění elitních hráčů v turnaji Czech Open A 2021 podle metod 9–20	66

Úvod

Začneme tím, co znamená slovo elo. Jde o pojem ratingového systému vyvinutého maďarsko-americkým fyzikem Arpadem Elem, který v původním významu vztáhnutém k šachovému prostředí číselně vyjadřuje průměrnou sílu hráče.

V bakalářské práci Elo systém očima matematické statistiky obhájené před dvěma lety jsme elo zavedli, nabídli jeho správnou interpretaci a potom zkoumali, jaké faktory se vedle něj projevují na výsledcích šachových partií. Všimli jsme si zejména vlivu bílých figur a rozdílného hráčského chování účastníků turnajů rozličného zaměření. Poznamenali jsme, že tyto informace mohou vést k lepší schopnosti předvídat výsledek partie, ale prakticky jsme nasbírané informace neověřili.

V této diplomové práci se pokusíme předvídat výsledky celých turnajů. K tomu budeme potřebovat co nejlepší odhady výsledků jednotlivých partií a využijeme tak řadu poznatků bakalářské práce. Některým poznatkům dáme velký prostor, jiné odmítneme jako neosvědčené a nahradíme novými myšlenkami.

Abychom dosáhli co nejhmatatelnějších výsledků, zaměříme se na specifickou situaci z šachové praxe, a sice poslední kola turnajů hraných švýcarským systémem. Ta jsou významná tím, že spolu s již známým dřívějším průběhem turnaje rozhodují o celkovém umístění hráčů a, podtrhněme, i o jejich finančním ohodnocení.

Budeme mít dvě měřítka úspěšnosti. Jednak nás bude zajímat, jak dobře by vyvinutý způsob odhadování výsledků předvídal průběh již odehraných turnajů, to znamená, jestli odhadnuté výsledky jednotlivých partií dobře korespondují s reálnými, a jestli tedy dobře odpovídají i konečná umístění jednotlivých hráčů. Jednak nás bude zajímat, jak užitečnou informaci poskytuje vyvinutý způsob jednotlivým účastníkům turnaje a jestli může vést k úpravě jejich strategie do posledního kola například ve smyslu ochoty remizovat s ohledem na zisk co nejlepšího umístění.

V první kapitole si zopakujeme fungování ratingového elo systému a doplníme potřebná fakta vázaná k turnajům hraným švýcarským systémem, zejména používaná pomocná hodnocení a způsob dělení finančních cen. Díky tomu si budeme moci přesně stanovit motivační úlohu, na níž ukážeme, jaké všechny informace mohou být důležité pro turnajové hráče.

V druhé kapitole nastíníme paletu původních způsobů, jak odhadovat výsledek partie. Určíme, podle čeho budeme posuzovat kvalitu jednotlivých způsobů, tyto způsoby mezi sebou vzájemně porovnáme a vybereme ten nejlepší.

Ve třetí kapitole si ukážeme jak odhadovat výsledky turnajů a jak tyto odhady interpretovat. Soustředovat se budeme na pohled jednotlivých hráčů, a na hledání jejich optimální turnajové strategie. Zhodnotíme, zda jim odhadnuté výsledky turnajů mohou pomoci nebo nikoliv.

V příloze A čtenář nalezne návod k používání vlastní funkce v R, která umí předpovídat umístění hráčů v turnaji na základě odhadnutých výsledků posledního kola a analyzovat související věci jako je dopad na finanční prémie jednotlivých hráčů. Příloha G obsahuje více než 2000 řádků dlouhý skript této funkce.

1 Parametry vstupující do šachových turnajů hraných švýcarským systémem

1.1 Fungování elo systému

Na tomto místě připomeneme jen ty základní myšlenky elo systému, které potřebujeme v dalším textu. Pokud by čtenář potřeboval pozvolnější uvedení do problematiky, může nahlédnout do bakalářské práce [1], případně si pustit na jejím základě natočené video [2]. Výborným zdrojem informací je potom kniha Arpada Ela *The rating of Chessplayers: Past and Present* [3].

Mějme jednorozměrnou číselnou osu. Jeden dílek této číselné osy nazveme elo bodem. Arpad Elo, tvůrce elo systému, který mezinárodní šachová federace FIDE používá od roku 1970, předpokládá, že na této ose lze měřit výkonnost hráče. Výkonnost hráče je potom normálně rozdělená náhodná veličina se směrodatnou odchylkou 200 elo bodů. Střední hodnota této náhodné veličiny se nazývá elo hráče.

Pokud má tedy hráč elo například 1750, stává se tedy s příslušnou frekvencí určenou normálním rozdělením, že v jedné partii podá výkon třeba 1650, v druhé 1800 a tak podobně. Ověřit že jeho výkonnost je normálně rozdělená, je prakticky nemožné, protože není jasné, co znamená, že hráč podal v partii ten nebo onen výkon. Dá se tušit, že v praxi dochází i k opačnému jevu, že někteří hráči, například děti, prochází výkonnostním růstem a jejich výkony nejsou normálně rozdělené. V každém případě elo systém s tímto předpokladem pracuje.

Samotné elo hráče nemá žádný hmatatelný obsah, nelze jej ztotožnit například se schopností nalézat v pozici dobré tahy nebo znalostí šachové teorie. Elo hráče se mění v závislosti na výsledcích hráče v partiích s dalšími hráči.

Další předpoklad Arpada Ela byl, že výkonnosti dvou hráčů jsou nezávislé. Obecně není důvod k rozporu, ale v šachové praxi nacházíme příklady, v nichž se vzájemná bilance dvou hráčů prokazatelně liší od předpokladů. Šachový publicista Jiří Veselý psal v této souvislosti o ochočených soupeřích [4]. Může jít o případy, kdy hráči extrémně sedí nebo nesedí soupeřův herní projev, což je pro něj větší výhoda nebo handicap než prostý rozdíl v jejich výkonnostech.

Jinou pochybnost mohou působit faktory psychologické. Může se stát, že slabý hráč se v partii se silným cítí předem poražený a v závislosti na informaci o síle soupeře sníží svou výkonnost. Podobně může silný hráč slabého hráče podcenit a hrát pod svoje možnosti. Stejně jako u předchozího ovšem platí, že elo systém s předpokladem nezávislých výkonností hráčů počítá.

Jestliže máme dvě nezávislé náhodné veličiny, můžeme určit pravděpodobnost, že jedna je větší než druhá. Pravděpodobnosti, že výkonnost prvního hráče je větší než druhého říkáme očekávaný výsledek partie a značíme ho W_e . Očekávaný výsledek partie přirozeně nabývá hodnot od 0 do 1. V závislosti na rozdílu el soupeřů z pohledu bílého D jej můžeme vypočítat podle vzorce

$$W_e = \int_{-\infty}^D \frac{1}{282,84\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2 \cdot 282,84^2}} dt \quad (1)$$

Čtenář se může podívat do [1] pro podrobnější odvození vzorce. My ho zde uvádíme pro snazší zavedení některých způsobů odhadu výsledku turnaje. Pro běžný výpočet se místo vyčíslování vzorce používá jednoduchá tabulka, která shrnuje očekávané výsledky v závislosti na rozdílech el hráčů. Je součástí klasifikačního řádu Mezinárodní šachové federace FIDE [5] i Šachového svazu České republiky [6].

Příklad 1: Hráč hájící bílé figury s elem 1670 se utká se soupeřem s elem 1820. Jaký je očekávaný výsledek partie?

Řešení: Jestliže el hráčů jsou střední hodnoty normálně rozdělených náhodných veličin vyjadřujících výkonnosti hráčů o směrodatné odchylce 200 elo bodů a liší se o $D = -150$, pak pravděpodobnost, že výkon bílého bude v partii vyšší než výkon černého, činí 30 %.

Zamysleme se nad tím, jakou informaci nám W_e dává. Očekávaný výsledek hráče v partii nemůžeme interpretovat jako pravděpodobnost jeho výhry. Označme si b jako pravděpodobnost výhry bílého a r jako pravděpodobnost remízy. Očekávaný výsledek bílého v partii je vlastně výraz $b + 0,5r$, tedy počet bodů, který očekáváme, že bílý uhraje. Ovšem poměr parametrů b a r ale neznáme.

Odpovídá očekávaný výsledek partie šachové praxi? Neuhrávají hráči o 150 elo bodů slabší než jejich soupeři víc než 30 % bodů? Hraje v tom roli barva figur? Nebo ještě něco jiného? Není vyloučené, že po zahrnutí některých dodatečných informací můžeme určit očekávaný výsledek partie lépe, ostatně v druhé kapitole se o to pokusíme. Mějme však na paměti, že v tu chvíli opouštíme elo systém a ocitáme se na poli vlastních odhadů.

Další informací, kterou nám elo systém může nabídnout, a kterou budeme potřebovat, je ohodnocení výkonu hráče v turnaji. Mějme P procentuální úspěšnost hráče v turnaji. Určíme

$$D(P) = 282,84 \cdot u(P) \quad (2)$$

kde přitom $u(P)$ je kvantilová funkce normovaného normálního rozdělení v bodě P . Potom výkon hráče v turnaji R_p (slangově perfo)

$$R_p = R_c + D(P) \quad (3)$$

kde přitom R_c je průměrné elo soupeřů, s nimiž se hráč v turnaji utkal.

1.2 Švýcarský systém šachových turnajů a Buchholzovo hodnocení

Švýcarský systém se používá tehdy, pokud má turnaj řádově desítky až stovky účastníků, ale je čas odehrát jen omezený počet kol, například 9 nebo 11. Hlavní myšlenka je párovat hráče, kterým se v turnaji daří podobně. Zjednodušeně můžeme například říct, že pokud hráč uhrál v prvních pěti kolech turnaje 3 body, bude hrát v šestém kole s jiným hráčem, jež uhrál v prvních pěti kolech turnaje taky 3 body. Detaily způsobu nasazování hráčů ovšem pro naši úlohu nejsou podstatné (případně viz [7]), protože v naší úloze budeme vždy los posledního kola znát. Ostatně los posledního kola zpravidla probíhá prakticky bezprostředně po skončení předposledního kola.

Snadno nahlédneme, že stejný počet bodů často bude mít větší skupina hráčů. Pro rozlišení konečného a vlastně i průběžného pořadí tedy nestačí počet bodů a je potřeba zavést další pomocná hodnocení. Nejčastější jsou různé modifikace Buchholzova hodnocení. Hlavní myšlenkou je přiřknout hráči jako pomocné hodnocení tolik bodů, kolik je součet bodů, jež uhráli všichni jeho dosavadní soupeři. Snahou je zvýhodnit hráče, kteří v turnaji čelili úspěšným soupeřům. Princip si ukážeme na jednoduchém příkladu:

Příklad 2: Aleš hrál pětikolový turnaj. V prvním kole hrál s Bedřichem, který celkově uhrál 4 body, se Ctíradem, který celkově uhrál 2 body, s Drahomilem, který celkově uhrál 3 body, s Evženem, který celkově uhrál 3,5 bodu a s Františkem, který celkově uhrál 1,5 bodu. Jaké bude Alešovo Buchholzovo hodnocení?

Řešení: Alešovo Buchholzovo hodnocení bude 14 bodů, protože soupeři, s nimiž v turnaji hrál, uhráli dohromady 14 bodů.

Častěji než klasické Buchholzovo hodnocení se používá tzv. bucholz cut 1, v němž se od zmíněného součtu odečte nejmenší počet bodů jednoho soupeře. V takovém případě by Buchholzovo hodnocení Aleše v příkladu 1 činilo 12,5 bodu, protože příspěvek Františka by se nepočítal.

Dále existuje střední bucholz (Median Bucholz), v němž se odečítá nejmenší počet bodů jednoho soupeře a nejvyšší počet bodů jednoho soupeře. V takovém případě by Buchholzovo hodnocení Aleše v příkladu 1 činilo 8,5 bodu, protože příspěvek Bedřicha a Františka by se nepočítal.

Existují i další méně používané varianty, ale těm se nebudeme věnovat.

V šachovém prostředí je zvykem udělovat za výhru v partii 1 bod, za remízu 0,5 bodu a za prohru 0 bodů. Pokud partie skončí kontumačně, to se stává například kvůli pozdnímu příchodu nebo zazvonění telefonu, dostanou oba hráči odpovídající počet bodů, jako kdyby vyhráli nebo prohráli.

Pokud se hráč do určitého kola nenechá nalosovat, například kvůli pracovním povinnostem, dostane 0 bodů. Pokud se hráč nalosovat nechá, ale vyjde jako lichý, dostane tzv. volno a zapíše si 1 bod.

Tyto partie, které neskončily přirozeným výsledkem, vnáší nutnost zavedení dalších pravidel do výpočtu Bucholzova hodnocení. Mezinárodní šachová federace FIDE od roku 2012 upřednostňuje výpočet pomocí konceptu fiktivního soupeře (virtual opponent). Lze očekávat, že ve většině turnajů dnes výpočet probíhá právě takto. Problémem ovšem je, že nejde o informaci, která by se zveřejňovala v propozicích turnajů, jde jen o nastavení losovacího programu. Pro případnou analýzu tak musíme vždy otevřít turnajový soubor ve Swiss Manageru a podívat se na nastavení pomocných hodnocení nebo u starších turnajů dokonce kontaktovat pořadatele.

Způsob počítání Bucholzova hodnocení s využitím fiktivních soupeřů je dost komplikovaný. Podrobné vysvětlení najdeme v [8, oddíl C, kapitola 2, článek 13.15.2–3]. Pro lepší pochopitelnost se níže pokusíme o vlastní vyložení a ilustrativní příklad.

Upozorňujeme čtenáře, že důkladným vysvětlením ztrácíme čas proto, aby bylo jasné, jak budeme konstruovat výsledkové listiny v rámci praktické úlohy. Pokud je však čtenář nebude chtít přímo přepočítávat, může s klidem tuto pasáž přeskočit, protože ji z jiných důvodů potřebovat nebude.

Podstatné je zaprvé, že pokud jednotliví soupeři hráče neodehráli všechny partie standardně (a tedy postihla je kontumace, volno nebo nenasazení), nepočítá se Bucholzovo hodnocení hráče z prostého součtu bodů jeho soupeřů, ale z upravených součtů tak, že právě kontumace, volna nebo nenasazení se počítají výhradně jako 0,5 bodu.

Zadruhé, pokud hráče samotného postihla kontumace, volno nebo nenasazení, musí si vypočítat příspěvek do Bucholzova hodnocení pomocí fiktivního soupeře. Uvažuje se, že místo dotčené kontumace, volna nebo nenasazení hrál s fiktivním soupeřem s výsledkem, který odpovídá zisku hráče z kontumace, volna nebo nenasazení. Tento fiktivní soupeř měl před daným kolem stejný počet bodů jako náš hráč, partii prohrál nebo vyhrál tak, jak to odpovídá zisku hráče, a zbytek partií remizoval. Zisk fiktivního soupeře se započte hráči do jeho Bucholzova hodnocení.

Příklad 3: Jak s pomocí konceptu fiktivního soupeře vypočteme Bucholzovo hodnocení Aleše, Bedřicha, Ctirada a Drahomila, jestliže máme k dispozici dané výsledky?

Tabulka 1: Příklad počítání pomocného hodnocení s fiktivním soupeřem (virtual opponent)

výsledky Aleše			výsledky Bedřicha		
kolo	soupeř	výsledek	kolo	soupeř	výsledek
1	Bedřich	1	1	Aleš	0
2	Ctirad	0	2	Bedřich	1 (kontumačně)
3	žádný	0 (nenasazen)	3	žádný	1 (volno)
4	Drahomil	0,5	4	Ctirad	1
výsledky Ctirada			výsledky Drahomila		
kolo	soupeř	výsledek	kolo	soupeř	výsledek
1	žádný	0 (nenasazen)	1	žádný	1 (volno)
2	Aleš	1	2	Bedřich	0 (kontumačně)
3	Drahomil	0,5	3	Ctirad	0,5
4	Bedřich	0	4	Aleš	0,5

Řešení: Bedřich přispěje Alešovi do Alešova Bucholzova hodnocení ne svým bodovým ziskem (3 body), ale jen 2 body, protože jeho „výhry“ z druhého a třetího kola počítáme jako remízy. Podobně Ctirad Alešovi nepřispěje 1,5 bodem, ale 2 body. Drahomil Alešovi přispěje 2 body, totiž 0,5 bodu za každé kolo.

Do Alešova Bucholzova hodnocení tedy počítáme příspěvek 2 body od Bedřicha, 1,5 bodu od Ctirada a 2 body od Drahomila. Příspěvek a Alešovu Bucholzovu hodnocení za třetí kolo, v němž nehrál, protože se nenechal nalosovat, řeší koncept fiktivního soupeře.

Představme si, že Aleš ve 3. kole nastoupil k partii a hrál proti fiktivnímu hráči Evženovi. Když se potkali ve 3. kole, bylo to proto, že oba měli po 2 kolech stejný počet bodů. Evžen měl tedy po 2 kolech 1 bod. Protože Aleš ve 3. kole připsal 0, můžeme usoudit, že Evžen vyhrál, měl tedy na kontě 2 body po 3 kolech. A protože o Evženovi už nic nevíme, budeme se domnívat, že Evžen celý zbytek turnaje (4. kolo) odremizoval. Získal tedy 2,5 bodu.

Takto vypočtené Bucholzovo hodnocení Aleše činí 8 bodů. Stejným způsobem vypočítáme Bucholzovo hodnocení i Bedřichovi, Ctiradovi a Drahomilovi, pomůžeme si přitom fiktivními hráči Františkem, Gustavem, Hugem, Chlodvíkem a Ivanem.

Bedřichovo Bucholzovo hodnocení vypočteme následovně. Příspěvek Aleše činí 2 body. Ve 2. kole Bedřich kontumačně vyhrál s Drahomilem. Kontumační výhra se počítá jako výhra nad fiktivním soupeřem Františkem, který měl po 1. kole stejně jako Bedřich 0 bodů, v 2. kole prohrál s Bedřichem, ale zbylé partie (ve 3. a 4. kole) už odremizoval. Proto František přispěje Bedřichovi do Bucholzova hodnocení 1 bod. Ve 3. kole Bedřich nehrál, protože vyšel jako lichý a dostal volno. To se počítá jako výhra s fiktivním soupeřem Gustavem. Gustav měl stejně jako Bedřich po 2 kolech na kontě 1 bod a partii s Bedřichem prohrál. Zbytek turnaje (4. kolo) odremizoval. Příspěvek Gustava tedy činí 1,5 bodu. Ctirad Bedřichovi do Bucholzova hodnocení přispěje 2 body. Bedřichovo Bucholzovo hodnocení tedy činí 6,5 bodu.

Ctirad se v prvním kole nenechal nalosovat. To se bere jako prohra s fiktivním soupeřem Hugem, který měl před 1. kolem stejně jako Ctirad 0 bodů. Po 1. kole měl Hugo už 1 bod a zbytek turnaje (2., 3. a 4. kolo odremizoval). Hugo Ctiradovi přispěje do Bucholzova hodnocení 2,5 bodu. Aleš Ctiradovi přispěl 2 body. Drahomil Ctiradovi přispěl 2 body. Bedřich Ctiradovi přispěl 2 body. Bucholzovo hodnocení Ctirad činí 8,5 bodu.

Drahomil v prvním kole vyšel jako lichý a dostal volno. To se počítá jako výhra s fiktivním soupeřem Chlodvíkem. Chlodvík měl před 1. kolem na kontě 0 bodů, v 1. kole s Drahomilem prohrál, ale od té doby (2., 3. a 4. kolo) všechny partie odremizoval. Drahomilovi do Bucholzova hodnocení přispěje 1,5 bodu. Ve 2. kole Drahomil kontumačně prohrál s Bedřichem. To se bere jako prohra s fiktivním soupeřem Ivanem. Ten měl před 2. kolem na kontě stejně jako Drahomil 1 bod. Po 2. kole měl Ivan díky výhře nad Drahomilem už 2 body. Zbytek turnaje (3. a 4. kolo) Ivan odremizoval, Drahomilovi proto přispěje 3 body. Ctirad Drahomilovi přispěje 2 body. Aleš Drahomilovi přispěje 2 body. Drahomilovo Bucholzovo pomocné hodnocení činí 8,5 bodu.

1.3 Dělení finančních cen v šachových turnajích

Pomocná hodnocení jako je Bucholzovo hodnocení jsou potřeba mimo jiné proto, aby bylo možné spravedlivě rozdělit finanční ceny. S pomocí pomocného hodnocení jednoznačně určíme pořadí hráčů v turnaji a můžeme například říct, že vítěz dostane 10 000 Kč, druhý hráč 6 000 Kč, třetí 4 000 Kč a podobně.

Nabízí se ovšem otázka, jestli je takto striktní rozdělení spravedlivé například tehdy, když první tři hráči mají stejně bodů a liší se jen o pomocná hodnocení. Proto se někdy ceny dělí rovným dílem mezi hráče se stejným počtem bodů.

Přijatelným a často používaným kompromisem je způsob dělení finančních cen, s kterým přišel československý šachista Vlastimil Hort. Polovinu peněz určených pro konkrétní místo vždy dostane hráč, který jej obsadil, zatímco zbylá polovina jde do banku. Bank si rozdělí rovným dílem všichni hráči, kteří měli stejný bodový zisk. Peníze se rozdělují jen mezi hráče, kteří jsou na místech určených pro finanční ohodnocení. Princip si ukážeme na příkladu.

Příklad 4: Pořadatel před turnajem určil, že za 1. místo udělí 10 000 Kč, 2. místo 6 000 Kč, 3. místo 4 000 Kč, 4. místo 2 000 Kč, 5. místo 1 000 Kč a 6. místo 500 Kč. Turnaj skončil následovně:

Tabulka 2: Příklad dělení finančních cen systémem hort

umístění	hráč	počet bodů	pomocné hodnocení	finanční zisk při dělení systémem hort
1. místo	Aleš	7	35	$5000 + (5\ 000 + 3000)/2 = 9\ 000$ Kč
2. místo	Bedřich	7	34,5	$3\ 000 + (5\ 000 + 3000)/2 = 7\ 000$ Kč
3. místo	Ctirad	6,5	33	$2000 + (2000 + 1\ 000 + 500)/3 = 3\ 166$ Kč
4. místo	Drahomil	6,5	32,5	$1000 + (2000 + 1\ 000 + 500)/3 = 2\ 166$ Kč
5. místo	Evžen	6,5	31	$500 + (2000 + 1\ 000 + 500)/3 = 1\ 666$ Kč
6. místo	František	6	28	500 Kč
7. místo	Gustav	6	27	nic

Můžeme nahlédnout, že systém hort vytváří mezery mezi skupinami hráčů s různým bodovým ziskem. Kdyby se peníze přidělovaly prostě za umístění v turnaji, dostal by Bedřich 6 000 Kč a Ctirad 4 000 Kč. Protože je ale mezi hráči bodový rozestup a oba dva patří do různých skupin, přiděluje systém hort Bedřichovi 7 000 Kč a Ctiradovi 3 166 Kč.

1.4 Motivační úloha

Na největším českém šachovém festivalu v Pardubicích v roce 2021 si mohli diváci před posledním kolem přehledně nalosování posledního kola turnaje A a mohli se dohadovat, který z hráčů se stane vítězem. Kritéria určení vítěze byla počet bodů, v případě rovnosti bucholz cut 1. Finanční ceny pro prvních 10 hráčů dělených systémem hort tvořilo 50 000 Kč pro 1. místo, pro 2. místo se počítalo s 30 000 Kč, pro 3. místo s 20 000 Kč, pro 4. místo s 15 000 Kč, pro 5. místo s 10 000 Kč, pro 6. místo s 8 000 Kč, pro 7. místo s 6 000 Kč, pro 8. místo s 5 000 Kč a pro 9. a 10. místo s 3 000 Kč.

Tabulka 3: los posledního kola turnaje Czech Open A v roce 2021

hráč	elo	bucholz cut 1	počet bodů	výsledek	počet bodů	bucholz cut 1	elo	Hráč
Babula	2563	35,5	6,5		6,5	32,5	2306	Czerw
Štoček	2604	35	6		6	32,5	2514	Liviac
Zwardoń	2466	40	6		6	37	2536	Pavlidis
Plát	2525	37	6		6	35,5	2291	Mládek

dále 5 dvojic hráčů majících na kontě 5,5 bodu

dále 11 dvojic hráčů majících na kontě 5 bodů

a dalších 51 dvojic s hráči majících na kontě méně než 5 bodů

Můžeme se vžít do role těchto diváků a pokusit se rozklíčovat možný budoucí vývoj.

Pokud hráči Babula nebo Czerw v partii vyhrají, mají jisté samostatné první místo a nemusí se na nic jiného ohlížet.

Pokud spolu hráči Babula a Czerw remizují, budou mít oba dva 7 bodů. Na 7 bodů by se pak mohli dotáhnout nejvýše tři další hráči. Mezi všemi by potom rozhodovalo pomocné hodnocení. Ze zkušenosti by šlo odhadnout, že hráči Czerwovi zůstane nižší pomocné hodnocení než hráči Babulovi, protože rozdíl tří bodů v bucholz už je značný. Bez podrobné analýzy je ale těžké například určit, s jakou pravděpodobností by měl vyšší bucholz cut 1 Babula po remíze ve své partii než Štoček v případě výhry své partie.

Hráči Babula a hráč Czerw by také měli zvažovat, co by pro ně znamenala prohra v posledním kole. Tam už je úvaha složitější. Hranici 6,5 bodů můžou přeskočit 3 hráči z 2. až 4. šachovnice, dotáhnout se na ni ale může teoreticky až 11 hráčů. Takový případ by nastal tehdy, pokud by partie na 2. až 4. šachovnici skončily remízou a na 5. až 9. šachovnici naopak remízou neskončily. Hráči Babula i Czerw by pravděpodobně měli vyšší bucholz cut 1 než tito dotáhnuvší se hráči, shodný počet bodů s nimi by ovšem mohl výrazně snížit finanční prémii, na níž by dosáhli.

A nyní se ptáme. Co by měl udělat hráč Babula v případě, že by mu hráč Czerw nabídl hned na začátku partie remízu, s ohledem na maximalizaci svého středního finančního zisku? Čekal ho souboj s teoreticky slabším hráčem a případná výhra by mu přinesla zisk 50 000 Kč, což je podstatně víc, než kolik by bral v případě dělení prvního místa například s dalšími třemi hráči. Současně by odmítnutím remízy riskoval pád výsledkovou listinou a klidně by se mohlo stát, že by bral například jen 5 000 Kč nebo ještě méně.

Nebudeme nyní pokračovat v rozvíjení diváckých spekulací o možném průběhu turnaje. V další kapitole se pokusíme vyvinout matematický model, který by do těchto úvah vnesl určitý řád, a teprve potom se vrátíme do předvečera posledního kola tohoto turnaje.

2 Odhad výsledků partií posledního kola turnaje

Předcházející odstavec osvětlil cíl práce. Chceme se naučit odhadovat průběh posledního kola turnaje hraného švýcarským systémem, abychom viděli, jaké jsou vyhlídky jednotlivého hráče s ohledem na výsledky ostatních. Budeme různými způsoby odhadovat pravděpodobnosti výhry bílého, remízy a výhry černého v každé jedné partii posledního kola daného turnaje. Až budeme znát odhady těchto pravděpodobností, zkusíme podle nich turnaj s pomocí výpočetní techniky několikrát dohrát a podíváme se, jak příznivé jsou výsledkové listiny takto dohraných turnajů pro jakéhokoli hráče, kterého si vybereme.

Tuto myšlenku musíme nejprve formalizovat. Každá partie může skončit výhrou bílého, remízou nebo výhrou černého, což znamená, že elementární výsledky náhodného pokusu, kterým rozumíme odehrání šachové partie, jsou $\Omega = \{\text{výhra bílého}, \text{remíza}, \text{výhra černého}\}$.

Všimněme si, že součet pravděpodobností, že nastane výhra bílého, remíza, respektive výhra černého, je roven jedné. To platí za předpokladu, že kontumační výsledky partie způsobené nedostavením se k partii a podobně, označíme za jevy nemožné.

V posledním kole se hraje n různých partií. Můžeme proto hovořit o elementárních výsledcích náhodného pokusu, jímž rozumíme odehrání i -té šachové partie posledního kola, v následující podobě: $\Omega_i = \{\text{výhra bílého v } i\text{-té partii}, \text{remíza v } i\text{-té partii}, \text{výhra černého v } i\text{-té partii}\}$.

Označíme pravděpodobnost b_i pravděpodobnost, že v i -té partii posledního kola bílý vyhraje, r_i pravděpodobnost, že remízuje, a \check{c}_i pravděpodobnost, že vyhraje černý. Dále označíme v_i výsledek i -té partie posledního kola. Náhodná veličina v_i vyjadřující výsledek i -té partie má následující rozdělení pravděpodobnosti:

$$v_i \sim \begin{cases} b_i = P(v_i = \text{výhra bílého v } i\text{-té partii}) \\ r_i = P(v_i = \text{remíza v } i\text{-té partii}) \\ \check{c}_i = P(v_i = \text{výhra černého v } i\text{-té partii}) \end{cases} \quad (4)$$

Skutečné pravděpodobnosti b_i , r_i a \check{c}_i neznáme, budeme je v dalším textu pouze odhadovat. Z typografických důvodů budeme výrazy b_i , r_i a \check{c}_i označovat i odhady pravděpodobností, že partie skončí tím kterým výsledkem.

S pomocí odhadů pravděpodobností jednotlivých výsledků jednotlivých partií se nám otevírají nové možnosti, zejména pokud učiníme předpoklad, že náhodné veličiny vyjadřující odhadnuté výsledky jednotlivých i -tých partií jsou nezávislé. Pak můžeme například snadno vypočítat odhad pravděpodobnosti, že všech n odhadnutých výsledků partií posledního kola bude vítězství bílých

$$P(\text{bílí vyhrají všechny partie posledního kola}) = \prod_{i=1}^n b_i \quad (5)$$

To, že bílí vyhrají všechny partie posledního kola je jen jednou z mnoha možných kombinací výsledků posledního kola. S přihlédnutím k $|\Omega| = 3$ a počtu partií posledního kola n víme, že jich je 3^n . Počítat odhad pravděpodobnosti každé jedné kombinace je prakticky neúnosné. Už při kombinaci výsledků 13 partií bychom překročili milion možností a v případě například $n = 100$, jak tomu v některých turnajích bývá, se dostáváme na $\approx 10^{47}$. Naštěstí nás nezajímají ani tak pravděpodobnosti jednotlivých kombinací výsledků, jako jejich promítnutí do závěrečné výsledkové listiny.

Z každé náhodné veličiny vyjadřující odhad výsledku i -té partie posledního kola provedeme náhodný výběr délky 1. Seřadíme-li n těchto jednorozměrných náhodných výběrů do vektoru, dostáváme vektor v , o němž budeme mluvit jako o vektoru odhadnutých výsledků.

Pro naše potřeby je výhodné sestavit víc vektorů v , můžeme jich sestavit s . Zavedeme označení $v^{(k)}$ jako k -tý vektor odhadnutých výsledků posledního kola. V každém z s vektorů děláme jednorozměrný náhodný výběr pro každou partii znovu.

Pokud všech s takových vektorů $v^{(k)}$ seřadíme svísele vedle sebe do matice a nazveme ji *TIPY*. Pod označením $v_i^{(k)}$ se skrývá k -tý odhad výsledku i -té partie posledního kola.

$$TIPY = \begin{pmatrix} v_1^{(1)} & v_1^{(2)} & \dots & v_1^{(k)} & v_1^{(s)} \\ v_2^{(1)} & v_2^{(2)} & \dots & v_2^{(k)} & v_2^{(s)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_i^{(1)} & v_i^{(2)} & \dots & v_i^{(k)} & v_i^{(s)} \\ v_n^{(1)} & v_n^{(2)} & \dots & v_n^{(k)} & v_n^{(s)} \end{pmatrix} \quad (6)$$

Zavedme dále vektor u , který má délku shodnou s počtem hráčů turnaje h , a v němž j -tý prvek u_j znamená umístění j -tého hráče v turnaji. Protože známe o turnaji všechny podrobnosti z předcházejícího průběhu, můžeme s pomocí vektoru v určit, jak by vypadala výsledková listina v případě, že by hráči hráli v turnaji skutečně tak, jak předpověděl vektor v . V tomto smyslu je vektor u funkcí vektoru v .

Rozebereme matici *TIPY*. Její sloupce představují náhodné nezávislé kombinace výsledků posledního kola. Každý takový k -tý sloupec matice *TIPY* můžeme využít k určení k -té výsledkové listiny $u^{(k)}$ a takových listin může být pochopitelně stejně, jaký je počet sloupečků matice *TIPY* tedy s . Pod označením $u_j^{(k)}$ budeme nově rozumět umístění j -tého hráče v turnaji, při nastání kombinace výsledků $v_i^{(k)}$.

Pokud uspořádáme vektory $u^{(k)}$ do matice svle vedle sebe, dostaneme matici *MÍSTA*. V jejích řádcích budeme číst umístění j -tého hráče v k -tém možném průběhu turnaje.

$$MÍSTA = \begin{pmatrix} u_1^{(1)} & u_1^{(2)} & \dots & u_1^{(k)} & u_1^{(s)} \\ u_2^{(1)} & u_2^{(2)} & \dots & u_2^{(k)} & u_2^{(s)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ u_j^{(1)} & u_j^{(2)} & \dots & u_j^{(k)} & u_j^{(s)} \\ u_h^{(1)} & u_h^{(2)} & \dots & u_h^{(k)} & u_h^{(s)} \end{pmatrix} \quad (7)$$

Jestliže jednotlivé sloupce matice *TIPY* jsou vlastně náhodnými výběry z odhadnutých kombinací výsledků, pak jednotlivé sloupce matice *MÍSTA* jsou náhodné výběry z odhadnutých pořadí hráčů po skončení turnaje. Dále j -tý řádek matice *MÍSTA* představuje náhodný výběr délky s z odhadnutých umístění hráče, a právě tomu budeme věnovat největší díl své pozornosti.

Protože lze ukázat, že střední hodnota výběrových průměrů ze všech myslitelných výběrových průměrů se rovná skutečnému průměru, můžeme říct, že při velkém s , tedy při učinění co největšího počtu různých náhodných výběrů z možných průběhů turnaje, se výběrový průměr blíží populačnímu. Jinými slovy pokud v j -tém řádku matice *MÍSTA* převažuje číslo 1, lze očekávat, že j -tý hráč turnaj vyhraje. Proto budeme-li analyzovat šance j -tého hráče, budeme sledovat rozdělení četnosti jemu přisouzených pořadí v j -tém řádku matice *MÍSTA*.

Způsoby jak odhadovat pravděpodobnosti b_i, r_i a \tilde{c}_i je logické posuzovat podle toho, jestli vytváří sloupce matice *TIPY* respektive sloupce matice *MÍSTA*, které odpovídají průběhům skutečných turnajů. Zavedeme proto vektor v^* znamenající skutečné výsledky daného turnaje v posledním kole a z toho plynoucí vektor u^* znamenající skutečná pořadí hráčů v daném turnaji.

2.1 Kritéria dobrého odhadu výsledků partií

Jak moc odhadnuté výsledky odpovídají skutečným výsledkům a odhadnutá pořadí skutečným pořadím budeme posuzovat pomocí korelačního koeficientu a střední kvadratické chyby. Nyní pojmy příslušně zavedeme a ukážeme, jak s nimi konkrétně budeme pracovat. Opíráme se v této fázi zejména o [9], [10].

Mějme dvě náhodné veličiny X a Y , z nichž obě mají své rozptyly. Výrazem

$$cov(X, Y) = E(X - EX)(Y - EY) \quad (8)$$

rozumíme kovarianci náhodných veličin X a Y .

Speciálním případem kovariance je korelační koeficient ρ vzniklý kovariancí dvou normovaných náhodných veličin:

$$\rho = \frac{\text{cov}(X, Y)}{\sqrt{\text{var } X \text{ var } Y}} \quad (9)$$

Tento koeficient se odhaduje z náhodného výběru pomocí výběrového korelačního koeficientu a značí se r . Označme nejprve výběrový průměr \bar{x} a výběrovou směrodatnou odchylku

$$s_X = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}} \quad (10)$$

(analogicky \bar{y} a s_Y), pak je

$$r = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{n s_X s_Y} \quad (11)$$

Korelační koeficient kvantifikuje vzájemný lineární vztah mezi veličinami X a Y , přičemž nijak nezaručuje příčinnou souvislost této vzájemnosti. Jeho hodnoty nabývají hodnoty od -1 do 1 , přičemž právě blízkost jedné ukazuje na onen vztah.

Nebudeme používat korelační koeficient k určení významnosti vzájemného vztahu, ale k tomu, abychom porovnávali způsoby odhadu pravděpodobností. Jestliže jeden způsob bude lépe korelovat s realitou, tj. bude mít vyšší hodnoty korelačního koeficientu, bude to pro nás jeden ze signálů, že takový způsob je lepší.

Korelačním koeficientem budeme zaprvé porovnávat shodu vektorů v a v^* . Pro $v^{(k)}$ dostáváme:

$$r_{\text{výsledků}}^{(k)} = \frac{\sum_{i=1}^n v_i^{(k)} v_i^* - n \overline{v_i^{(k)}} \overline{v_i^*}}{n s_{v^{(k)}} s_{v^*}} \quad (12)$$

Protože vektorů v budeme mít s , můžeme si dovolit vypočítat průměr z takovýchto korelačních koeficientů. Výsledná hodnota, kterou budeme používat pro analýzu, tedy bude:

$$\bar{r}_{\text{výsledků}} = \frac{\sum_{k=1}^s r_{\text{výsledků}}^{(k)}}{s} \quad (13)$$

Abychom měli možnost si vytvořit představu o variabilitě k -tých korelačních koeficientů, budeme s pomocí vlastní funkce v R počítat i směrodatnou odchylku:

$$s_{r_{\text{výsledků}}} = \sqrt{\frac{\sum_{k=1}^s \left(r_{\text{výsledků}}^{(k)} - \bar{r}_{\text{výsledků}} \right)^2}{s}} \quad (14)$$

Korelační koeficient použijeme zadruhé pro srovnání toho, zda umístění hráčů $u^{(k)}$ po k -tém náhodném výběru výsledků posledního kola $v^{(k)}$ odpovídají reálným výsledkům turnajů a tedy vektoru u^* . Protože jde o speciální případ porovnávání pořadí, mohli bychom se stejným výsledkem použít Spearmanův korelační koeficient, ale pro jednoduchost se přidržíme standardního výpočtu korelačního koeficientu. Vypočteme:

$$r_{\text{umístění}}^{(k)} = \frac{\sum_{j=1}^h u_j^{(k)} u_j^* - n \overline{u_j^{(k)}} \overline{u_j^*}}{n s_{u^{(k)}} s_{u^*}} \quad (15)$$

a stejně jako v minulém případě si dovolíme vypočítat i průměr korelací odhadnutých umístění se skutečnými

$$\bar{r}_{\text{umístění}} = \frac{\sum_{k=1}^s r_{\text{umístění}}^{(k)}}{s} \quad (16)$$

a směrodatnou odchylku těchto s korelačních koeficientů:

$$s_{r_{\text{umístění}}} = \sqrt{\frac{\sum_{k=1}^s \left(r_{\text{umístění}}^{(k)} - \bar{r}_{\text{umístění}} \right)^2}{s}} \quad (17)$$

Až budeme ověřovat, zda umístění $u^{(k)}$ dobře odpovídají realitě u^* , budou nás nejvíce zajímat umístění těch hráčů, kteří bojují o přední příčky. Pro potřeby práce je pojmenujeme elitní hráči. Nebudeme pod tímto pojmem chápat hráče s nejvyšším ratingem, ale ty, kteří mají před posledním kolem buď stejně bodů jako první hráč průběžného pořadí, o 0,5 bodu méně nebo o 1 bod méně. Tyto hráče sledujeme odděleně proto, že jsou to jen oni, kdo mají alespoň teoretickou šanci na vítězství v turnaji. Navíc z jejich řad se zpravidla rekrutují příjemci finančních prémie vyhrazených obvykle pro několik nejúspěšnějších hráčů. Odhad umístění těchto hráčů je proto důležitý. Pokud bychom se chtěli zaměřit jen na problematiku finančních zisků jednotlivých hráčů, mohla by se pro nás stát právě korelace odhadů umístění elitních hráčů s realitou tou nejvýznamnější metrikou.

Označíme e^k vektor umístění f elitních hráčů po k -tém náhodném výběru výsledků posledního kola s prvky $e_g^{(k)}$ vyjadřující umístění g -tého elitního hráče po k -tém náhodném výběru výsledků posledního kola. Dále označíme e^* vektor skutečných umístění f elitních hráčů s prvky $e_g^{(k)}$ vyjadřující umístění g -tého elitního hráče po k -tém náhodném výběru výsledků posledního kola. Třetí způsob využití korelačního koeficientu proto bude mít následující podobu:

$$r_{\text{umístění elity}}^{(k)} = \frac{\sum_{g=1}^f e_g^{(k)} e_g^* - f \overline{e_g^{(k)}} \overline{e_g^*}}{f s_{e^{(k)}} s_{e^*}} \quad (18)$$

Průměr korelací mezi k -tým vektorem odhadnutých výsledků elitních hráčů $e^{(k)}$ a vektorem skutečných výsledků e^* takto:

$$\bar{r}_{\text{umístění elity}} = \frac{\sum_{k=1}^S r_{\text{umístění elity}}^{(k)}}{S} \quad (19)$$

A směrodatnou odchylku těchto korelací takto:

$$s_{r_{\text{umístění elity}}} = \sqrt{\frac{\sum_{k=1}^S \left(r_{\text{umístění elity}}^{(k)} - \bar{r}_{\text{umístění elity}} \right)^2}{S}} \quad (20)$$

Zatímco korelační koeficient pracuje s normovanými náhodnými veličinami, střední kvadratická chyba nabízí průhlednější pohled na to, jak moc se liší dvě náhodné veličiny, respektive dva náhodné výběry.

Mějme veličinu X a její odhad Y . Střední kvadratická chyba odhadu Y představuje střední hodnotu druhých mocnin rozdílů mezi X a Y . Střední kvadratická chyba se někdy nazývá MSE z anglického mean squared error se tedy vypočítá:

$$MSE = E(X - Y)^2 \quad (21)$$

Jestliže porovnáme k -tý odhad výsledků posledního kola $v^{(k)}$ s reálnými výsledky posledního kola v^* , počítáme:

$$MSE_{\text{výsledků}}^{(k)} = \frac{\sum_{i=1}^n \left(v_i^{(k)} - v_i^* \right)^2}{n} \quad (22)$$

Protože budeme porovnávat s takových vektorů, znovu nás bude zajímat průměr s takových $MSE_{\text{výsledků}}$ a jejich směrodatná odchylka. Ty vypočteme následovně:

$$\overline{MSE}_{\text{výsledků}} = \frac{\sum_{k=1}^S MSE_{\text{výsledků}}^{(k)}}{S} \quad (23)$$

a

$$s_{MSE_{\text{výsledků}}} = \sqrt{\frac{\sum_{k=1}^S \left(MSE_{\text{výsledků}}^{(k)} - \overline{MSE}_{\text{výsledků}} \right)^2}{S}} \quad (24)$$

Dále uved'eme vzorce pro střední kvadratickou chybu pro odhad $u^{(k)}$

$$MSE_{umístění}^{(k)} = \frac{\sum_{j=1}^h (u_j^{(k)} - u_j^*)^2}{h} \quad (25)$$

průměr s takových odhadů

$$\overline{MSE}_{umístění} = \frac{\sum_{k=1}^s MSE_{umístění}^{(k)}}{s} \quad (26)$$

a jejich směrodatnou odchylku:

$$S_{MSE_{umístění}} = \sqrt{\frac{\sum_{k=1}^s (MSE_{umístění}^{(k)} - \overline{MSE}_{umístění})^2}{s}} \quad (27)$$

Stojí za povšimnutí, že $MSE_{umístění}^{(k)}$ a $\rho_{umístění}^{(k)}$ jsou nepřímo úměrné. Oboje ukazuje totéž v jiném kontextu.

Nakonec ještě uved'eme vzorce pro střední kvadratickou chybu k -tého odhadu umístění elitních hráčů

$$MSE_{umístění\ elite}^{(k)} = \frac{\sum_{g=1}^f (e_g^{(k)} - u_g^*)^2}{f} \quad (28)$$

průměr s takových odhadů

$$\overline{MSE}_{umístění\ elite} = \frac{\sum_{k=1}^s MSE_{umístění\ elite}^{(k)}}{s} \quad (29)$$

a jejich směrodatnou odchylku

$$S_{MSE_{umístění\ elite}} = \sqrt{\frac{\sum_{k=1}^s (MSE_{umístění\ elite}^{(k)} - \overline{MSE}_{umístění\ elite})^2}{s}} \quad (30)$$

Nyní už máme nastavená základní kritéria, jimiž budeme posuzovat kvalitu způsobů, jimiž budeme odhadovat pravděpodobnosti b_i , r_i a \check{c}_i . Za každou metodou uvedeme šestici $\bar{r}_{výsledků}$, $\bar{r}_{umístění}$, $\bar{r}_{umístění\ elite}$, $\overline{MSE}_{výsledků}$, $\overline{MSE}_{umístění}$, $\overline{MSE}_{umístění\ elite}$, která bude odpovídat shodě odhadnutých výsledků a reality na turnaji, který známe z motivačního příkladu, totiž z turnaje Czech Open A z roku 2021. Pro každou metodu provedeme 1000 simulací, které budeme porovnávat s realitou.

Výše zmíněnou šestici budeme používat k vytvoření představy. Pokud budeme chtít některé metody porovnat odpovědně, půjdeme jinou cestou, a sice aplikací analýzy rozptylu dvojného třídění. Každou metodu navíc prověříme v počtu 100 simulací na pěti ročnících z let 2015–2019 na pěti typech turnajů, a sice na turnajích Czech Open A, Czech Open B, Czech Open C, Czech Open D a Open Praha. Pro každou dvojici typ turnaje a metoda tak budeme mít k dispozici pět měření. Například pro korelace odhadnutých umístění hráčů a reality na turnaji Czech Open A metodou 1 dostaneme vektor, můžeme si ho pro ukázkou označit $COA1$, v následujícím tvaru:

$$COA1 = \begin{bmatrix} \bar{T}_{umístění\text{Czech Open A 2015 metodou 1}} \\ \bar{T}_{umístění\text{Czech Open A 2016 metodou 1}} \\ \bar{T}_{umístění\text{Czech Open A 2017 metodou 1}} \\ \bar{T}_{umístění\text{Czech Open A 2018 metodou 1}} \\ \bar{T}_{umístění\text{Czech Open A 2019 metodou 1}} \end{bmatrix} \quad (31)$$

Vektor $COA1$ chápeme jako náhodný výběr měřítek úspěšnosti metody 1 aplikované na turnaj Czech Open A. Vzhledem k tomu, že budeme prověřovat úspěšnost 20 metod na 5 typech turnajů, dostaneme 100 takových vektorů (například $COA2$, $COB15$ atp.).

Co se týče předpokladů analýzy rozptylu, je logické předpokládat, že náhodné výběry tohoto typu jsou normálně rozdělené. Ještě zřejmější je, že vektory jsou nezávislé, vždyť jde o úplně jiné turnaje s jinými hráči. K domněnce, že určitý vliv mají roky odehrání turnaje, lze zaujmout odmítavé stanovisko, jelikož k tomu není žádný dobrý důvod. Snad by byl, kdybychom zahrnuli covidový rok 2020, ale ten je právě z těchto důvodů vynechán. K ověření toho, že jednotlivé vektory jsou náhodné výběry z rozdělení ze zhruba stejným rozptylem, budeme používat Leveneův test, ale v žádné z prováděných analýz nenarazíme na signifikantní výsledek ukazující na rozdílnost rozptylů.

Analýzu rozptylu dvojného třídění budeme aplikovat buď na všech 100 vektorů, nebo jen na část z nich, pokud si to vyžádá situace. Bude nás přitom zajímat zaujetí postoje k následujícím třem nulovým hypotézám.

Zaprvé úspěšnost odhadu nezávisí na volbě metody. Pokud bychom takovou hypotézu nezamítli, znamenalo by to, že rozdíly mezi úspěšnostmi metod nejsou signifikantní, jinými slovy v souboru vektorů není žádná metoda, která by byla významně lepší než jiná.

Zadruhé úspěšnost odhadu nezávisí na typu turnaje. Pokud bychom takovou hypotézu nezamítli, znamenalo by to, že všechny typy turnajů jdou odhadovat podobně dobře, jinými slovy že jejich specifika nehrají významnou roli pro předpověď výsledků.

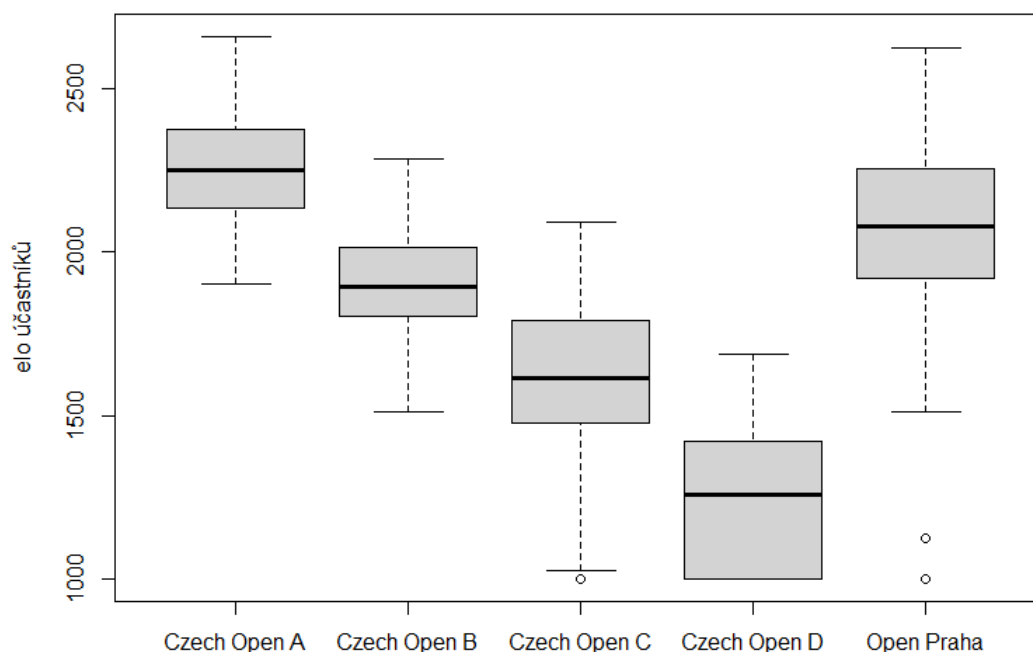
Zatřetí úspěšnost odhadu nezávisí na kombinaci volby metody a typu turnaje. Pokud bychom takovou hypotézu nezamítli, znamenalo by to, že dochází k tomu, že některé metody fungují dobře u některých typů turnajů, zatímco u jiných ne.

Analýzu rozptylu jako takovou budeme vždy celou dělat v programu R, proto zde nebudeme příliš důkladně rozebírat její konstrukci. Zajímat nás bude zejména testová statistika jednotlivých hypotéz a p-hodnota, tedy nejnižší hladina, při které bychom hypotézu zamítli. Pokud bude p-hodnota nižší než 0,05, budeme říkat, že příslušnou hypotézu zamítáme, pokud nižší než 0,01, že hypotézu zamítáme s jistotou, a pokud by byla nižší než 0,001, že hypotézu zamítáme s velkou jistotou.

Pokud by některý z efektů, ať už vliv výběru metody, typu turnaje nebo jejich vzájemného vztahu, vyšel významný, je na místě jednotlivé veličiny porovnat. K tomu budeme v programu R používat Tukeyovu metodu mnohonásobného porovnávání. Ta nám ukáže, mezi kterými veličinami jsou signifikantní rozdíly.

Dále vysvětlíme, proč volíme právě turnaje Czech Open A, B, C, D a Open Praha. Důvodem je jejich etablovanost v českém šachovém prostředí a vyrovnanost jednotlivých ročníků. Pardubické turnaje Czech Open pokrývají úroveň svých účastníků prakticky celé spektrum turnajově aktivních šachistů (připomeňme, že nejmenší možné elo je 1000 a jen zhruba 50 hráčů na světě má elo vyšší než 2700). Turnaj Open Praha je zahrnutý pro srovnání, má trochu širší rozptyl hráčů a nižší podíl cizinců na startovní listině.

Graf 1: Elo účastníků na zkoumaných turnajích



Ještě než přistoupíme k vlastní analýze, uveďme, že veškeré výpočty a sestavování výsledkových listin probíhají v programu R, verze 4.0.2, semínko nastavené na hodnotu 1506, a že veškerá šachová data jsou stažena z webu chess-results.com [11].

2.2 Metody založené na předpokladech elo systému

V této podkapitole rozebereme osm v principu velmi jednoduchých cest, jak odhadovat výsledek partie. Jejich společným jmenovatelem bude využití elo systému, případně jeho korigování v duchu myšlenek, které jsme diskutovali v bakalářské práci. Zohledníme tak zejména výhodu bílých figur a dvěma různými způsoby výskyt remíz.

1. způsob odhadování výsledků: Začneme naivním odhadem a stanovíme, že pravděpodobnost výhry bílého, remízy i výhry černého je vždy stejná, tedy že $b_i = r_i = \check{c}_i = \frac{1}{3}$. Dosáhneme nepřekvapivých výsledků, jež shrnuje tabulka 1.

Tabulka 4: kvalita 1. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{r}_{\text{výsledků}} = -0,0033 & \bar{r}_{\text{umístění}} = 0,9220 & \bar{r}_{\text{umístění elity}} = 0,5378 \\ \overline{MSE}_{\text{výsledků}} = 0,3235 & \overline{MSE}_{\text{umístění}} = 288,4 & \overline{MSE}_{\text{umístění elity}} = 85,76 \end{array}$$

Výsledky partií by patrně neměly korelovat s náhodně volenými odhady výsledků. Pokud se počet simulací $s \rightarrow \infty$, očekáváme $\bar{r}_{\text{výsledků}} = 0$ a tomuto očekávání průměr korelací v podstatě odpovídá.

Jiná je situace s $\bar{r}_{\text{umístění}}$, tam je korelace poměrně dobrá. Souvisí to s tím, že celkové rozložení pořadí hráčů v celém turnaje se neliší dramaticky od jejich už po předposledním kole, hráči, kteří byli v popředí, v popředí skončí, ti, kteří byli na chvostu tabulky, na něm skončí.

Nižší hodnotu korelace nabízí $\bar{r}_{\text{umístění elity}}$. Rozdíl spočívá v tom, že v souboru dat jsou jen hráči z popředí a vyniknou tedy drobné rozdíly v umístění. Náhodný výběr výhry, remízy nebo prohry není dost dobrou oporou k odhadu umístění jednotlivců.

Celkově se ale dá říct, že už tato metoda přináší hmatatelný odhad, který vnáší řád do diváckých intuitivních odhadů a těží ze síly simulačního přístupu.

2. způsob odhadování výsledků: Jak víme, elo systém neumí určit pravděpodobnost všech tří možných výsledků partie. Uměl by to tehdy, pokud by partie nemohla skončit remízou. Zkusíme tuhle myšlenku využít. Položíme $r = 0$, $b = W_e$ a $\check{c} = 1 - b$. To znamená, že žádný odhad partie neskončí remízou, pravděpodobnost b ztotožníme s očekávaným ziskem bílého W_e zavedeným výše a pravděpodobnost výhry černého bude doplňkem b do jedné.

Tabulka 5: kvalita 2. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{\rho}_{\text{výsledků}} = 0,1553 & \bar{\rho}_{\text{umístění}} = 0,9197 & \bar{\rho}_{\text{umístění elity}} = 0,5816 \\ \overline{MSE}_{\text{výsledků}} = 0,3444 & \overline{MSE}_{\text{umístění}} = 297,2 & \overline{MSE}_{\text{umístění elity}} = 97,24 \end{array}$$

Nebudeme rozebírat jednotlivé ukazovatele shody s reálnými výsledky u každého způsobu do detailu zvláště, srovnáme je najednou. Obecně ovšem můžeme říct, že tato metoda nepřinesla výrazně lepší výsledky než první.

3. způsob odhadování výsledků: V bakalářské práci jsme ukázali, že bílé figury přináší hráčům výhodu, která je ekvivalentní řádu desítek elo bodů v závislosti na typu hráče. Je tedy na místě prověřit, jestli dosáhneme lepších odhadů, pokud bílým k jejich elu přičteme určitý počet elo bodů. Jaký?

Podíváme se na podíl bodů, kteří uhráli bílí v daném turnaji v předcházejících kolech a dosadíme je do vzorce pro $D(P)$. Protože bílí uhráli v roce 2021 v turnaji Czech Open A 56,21 % bodů, budeme při odhadování výsledků 9. kola bílým přičítat $D(0,5625) \doteq 44$ elo bodů.

Tabulka 6: kvalita 3. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{r}_{\text{výsledků}} = 0,1771 & \bar{r}_{\text{umístění}} = 0,9217 & \bar{r}_{\text{umístění elity}} = 0,5670 \\ \overline{MSE}_{\text{výsledků}} = 0,3359 & \overline{MSE}_{\text{umístění}} = 289,9 & \overline{MSE}_{\text{umístění elity}} = 101,2 \end{array}$$

4. způsob odhadování výsledků: Informace, kterou nám dává elo hráče o jeho výkonnosti nemusí pružně reagovat na jeho aktuální formu. Tu lépe odráží výkon hráče v turnaji, slangově perfo. V této metodě vyjdeme z druhého způsobu, jen zkusíme nahradit elu hráčů jejich výkony v předcházejících kolech turnaje.

Tabulka 7: kvalita 4. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{r}_{\text{výsledků}} = -0,0218 & \bar{r}_{\text{umístění}} = 0,9063 & \bar{r}_{\text{umístění elity}} = 0,5114 \\ \overline{MSE}_{\text{výsledků}} = 0,4143 & \overline{MSE}_{\text{umístění}} = 346,5 & \overline{MSE}_{\text{umístění elity}} = 109,0 \end{array}$$

Opření se o předcházející výkon hráčů v turnaji není nosnou cestou. Ne úplně špatně odhaduje umístění elitních hráčů, ovšem v ostatních parametrech dává spíše horší výsledky.

5. způsob odhadování výsledků: Zdá se, že největším nedostatkem dosud ukázaných způsobů je ignorování remíz. Určíme r jako podíl remíz v dosavadním průběhu turnaje (v našem konkrétním případě to je 34,6 % a odhadneme

$$b = W_e - \frac{r}{2} \tag{32}$$

a

$$\check{c} = 1 - W_e - \frac{r}{2} \tag{33}$$

Musíme ještě ošetřit případy, kdy $r > 2W_e$ nebo $r > 2(1 - W_e)$. Tehdy by pravděpodobnosti b , respektive $č$ vycházely záporné, což nechceme. V programu tyto situace ošetříme podmínkou, že pro taková r změním jejich hodnotu na $r = 2W_e$ respektive $r = 2(1 - W_e)$. Čeho tím dosáhneme? Pokud bude v partii posledního kola jeden z hráčů, ať už bílý nebo černý, výrazně slabší, řekneme, že je nemožné, aby partii vyhrál, a jeho odpovídající očekávaný zisk realizujeme pouze remízami. Takový přístup není úplně vzdálený realitě.

Tabulka 8: kvalita 5. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{r}_{\text{výsledků}} = 0,1554 & \bar{r}_{\text{umístění}} = 0,9198 & \bar{r}_{\text{umístění elity}} = 0,5812 \\ \overline{MSE}_{\text{výsledků}} = 0,3438 & \overline{MSE}_{\text{umístění}} = 296,7 & \overline{MSE}_{\text{umístění elity}} = 97,18 \end{array}$$

6. způsob odhadování výsledků: Spojíme třetí a pátou metodu, tedy přihlédneme k výhodě bílých figur a zavedeme fixní odhad remízy v partii, obojí na základě předchozích kol turnaje.

Tabulka 9: kvalita 6. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{r}_{\text{výsledků}} = 0,1554 & \bar{r}_{\text{umístění}} = 0,9198 & \bar{r}_{\text{umístění elity}} = 0,5812 \\ \overline{MSE}_{\text{výsledků}} = 0,3438 & \overline{MSE}_{\text{umístění}} = 296,7 & \overline{MSE}_{\text{umístění elity}} = 97,18 \end{array}$$

7. způsob odhadování výsledků: Pevný odhad remízy je dost toporný, je zřejmé, že jiný podíl remíz se bude vyskytovat v partiích dvou obdobně silných hráčů a jiný v partiích dvou rozdílně silných hráčů. Zkusíme proto odhad r pro jednotlivé partie trochu vylepšit.

Není problém vypočítat podíly výher bílých, remíz a výher černých předcházejícím průběhu turnaje. Označme je po řadě \check{b} , \check{r} a $\check{č}$. Bílý za výhru dostane 1 bod, za remízu 0,5 bodu a za prohru 0 bodů. Můžeme tedy říct, že výhry se na jeho celkovém zisku podílejí průměrně podílem z_b , přičemž

$$z_b = \frac{\check{b}}{\check{b} + 0,5\check{r}} \quad (34)$$

Stanovme proto $b = z_b \cdot W_e$. Jak se to projeví na našem konkrétním turnaji? V něm je $\check{b} = 38,93\%$ a $\check{r} = 34,62\%$. Bílé výhry se tedy průměrně podílejí na celkovém zisku bílého podílem $z_b = 69,22\%$. Jestliže spolu například hrají dva hráči, z nichž bílý je silnější o 150 elo bodů, vypočítáme $W_e = 0,7$. Dosadíme do (33) a dostáváme $b = 0,6922 \cdot 0,7 = 0,4845$. Touto metodou přicházíme na to, že pravděpodobnost výhry bílého v této specifické partii je 48,45 %.

Analogickým způsobem vypočítáme $č$, tedy

$$\check{č} = \frac{\check{č}}{\check{č} + 0,5\check{r}} \cdot (1 - W_e) \quad (35)$$

To v našem konkrétním případě při $\check{c} = 26,45 \%$ dává $\check{c} = 18,13 \%$.

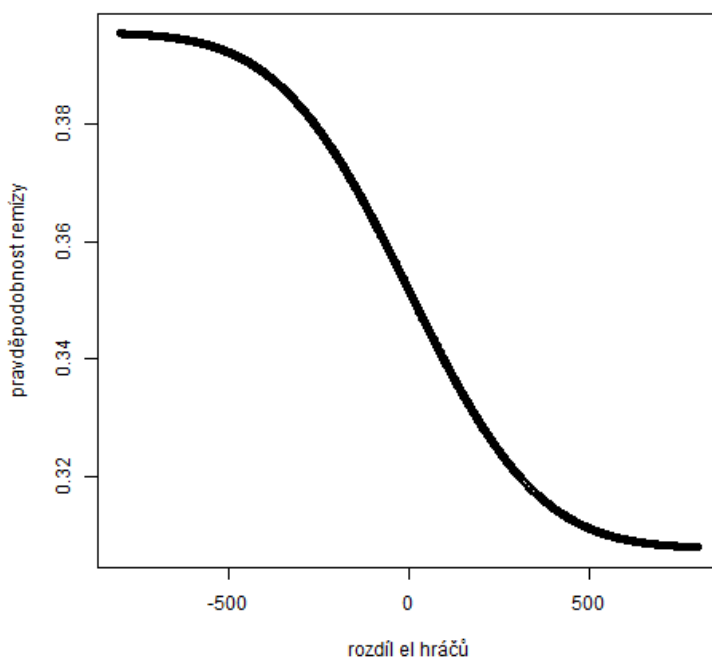
Pravděpodobnost remízy určíme jako

$$r = 1 - b - \check{c} \quad (36)$$

v našem případě tedy dostáváme $r = 33,42 \%$.

Výhodou této metody je, že r se mění v závislosti na rozdílu el hráčů. Ukažme si, jaká r budou vycházet v našem konkrétním příkladu v případě různých rozdílů el hráčů z pohledu bílého.

Graf 2: rozložení remíz v závislosti na rozdílu v turnaji Czech Open A 2021 podle metody 7



Nepřekvapuje nás distribuce hodnot kolem 34,5 %, vzhledem k tomu, že této hodnotě odpovídá podíl remíz v předcházejícím průběhu turnaje. Rozložení hodnot není sudé, což se může zdát zvláštní, ale ve skutečnosti není žádný důvod, proč by tomu tak mělo být. Situace, kdy je v partii silnější bílý a kdy černý, jsou zkrátka různé z hlediska výskytu remíz.

Tabulka 10: kvalita 7. způsobu odhadování výsledků

$\bar{r}_{\text{výsledků}} = 0,1364$	$\bar{r}_{\text{umístění}} = 0,9345$	$\bar{r}_{\text{umístění elity}} = 0,6048$
$\overline{MSE}_{\text{výsledků}} = 0,2742$	$\overline{MSE}_{\text{umístění}} = 242,5$	$\overline{MSE}_{\text{umístění elity}} = 73,87$

8. způsob odhadu výsledků: Protože výsledky předcházejícího způsobu shledáváme povzbudivými, vylepšíme jej ještě o zvýhodnění bílých fixní ratingovou výhodou v duchu třetí metody.

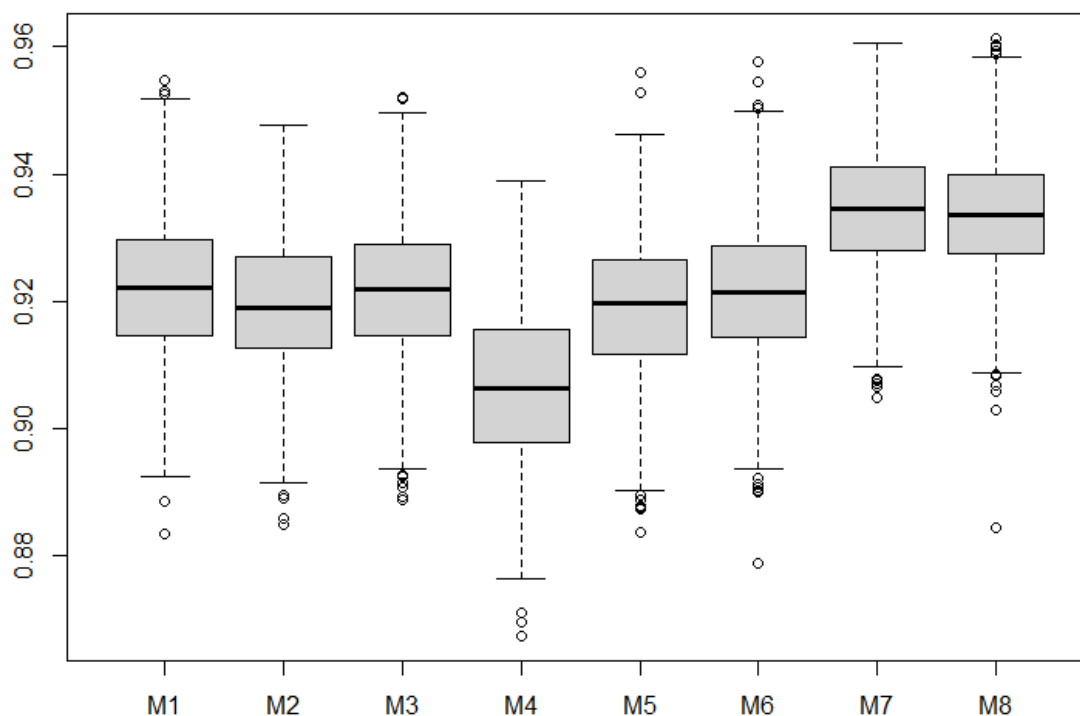
Tabulka 11: kvalita 8. způsobu odhadování výsledků

$$\begin{aligned} \bar{r}_{\text{výsledků}} &= 0,1609 & \bar{r}_{\text{umístění}} &= 0,9340 & \bar{r}_{\text{umístění elity}} &= 0,5876 \\ \overline{MSE}_{\text{výsledků}} &= 0,2767 & \overline{MSE}_{\text{umístění}} &= 244,2 & \overline{MSE}_{\text{umístění elity}} &= 79,49 \end{aligned}$$

2.3 Srovnání metod založených na předpokladech elo systému

Srovnáme nyní osm výše uvedených metod podle korelací odhadovaných a skutečných umístění. Srovnání korelací odhadovaných a skutečných výsledků a korelací odhadovaných a skutečných umístění elitních hráčů uvádíme v příloze B.

Graf 3: korelace odhadnutých a skutečných umístění hráčů v turnaji Czech Open A 2021 podle metod 1–8



Co můžeme vyčíst z grafu? Testování na turnaji Czech Open A v roce 2021 ukazuje na nejhorší výsledky čtvrté metody postavené na výkonu hráče v předcházejícím průběhu turnaje. Za povšimnutí také stojí, že metody, které jsou oproti předcházejícím vylepšené o zahrnutí výhody bílých figur (metoda 3 vůči metodě 2, metoda 6 vůči metodě 5 a metoda 8 vůči metodě 7) pomocí zvýšení ely bílým, se nezdají být výrazně lepší.

Jak jsme slíbili, metody prověřujeme nejenom na turnaji Czech Open A v roce 2021, ale obecně na turnajích Czech Open A, B, C, D a Open Praha v letech 2015–2019. Analýzy rozptylu metod případě jejich porovnání vztahující se k prvních osmi metodám nalezne čtenář v příloze C.

Na tomto místě se omezíme jen na konstatování, že rozdíly v úspěšnostech metod jsou signifikantní, stejně jako rozdíly v předvídatelnosti jednotlivých typů turnajů. Metody 7 a 8 mají obdobné výsledky optikou korelací odhadovaných a skutečných výsledků a odhadovaných a skutečných umístění. Za nejlepší metodu z prvních osmi prohlašujeme metodu číslo 8, vzhledem k tomu, že je významně lepší z hlediska korelací odhadovaných a skutečných umístění elitních hráčů.

2.4 Metody založené na turnajové praxi

V předcházející podkapitole jsme se pokoušeli sestavit odhady pravděpodobností jednotlivých výsledků pomocí zavedení nových parametrů, fakticky jsme vylepšovali elo systém. V této části na něj brát ohled vůbec nebudeme. Vyjdeme z toho, jak dopadly podobné partie hrané v minulosti. Představíme čtyři postupy, které ukážeme vždy ve třech variantách. V první variantě se opřeme o dosavadní průběh daného turnaje, v druhé variantě o partie hrané hned v několika turnajích daného typu a ve třetí variantě o partie hrané v malé databázi, kterou si sestojíme ze všech typů turnajů dohromady.

U všech následujících metod bude platit, že pokud počet podobných partií bude menší než 4, odhad pravděpodobnosti automaticky proběhne přes 8. metodu. Ostatně proto jsme v předcházející podkapitole vybírali nejlepší metodu z prvních osmi.

9. způsob odhadování výsledků: Jedním z nejdůležitějších závěrů bakalářské práce, na kterou navazujeme, bylo, že výhoda bílých figur se nerealizuje stejně silně vzhledem k rozdílu el hráčů. Všimli jsme si například, že pokud je bílý v partii slabší, přináší mu výhoda prvního tahu zpravidla větší posun oproti očekávání, než když je v partii už tak silnější.

Představme si, že chceme předpovědět výsledek partie, v níž se elo hráčů liší o D . V této metodě se podíváme na všechny partie v předcházejícím průběhu turnaje, v nichž rozdíl el hráčů spadl do intervalu $[D - 50; D + 50]$, a podíly výher bílých, remíz a proher bílých v tomto intervalu vezmeme jako odhady pravděpodobností naší partie.

Tabulka 12: kvalita 9. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{r}_{\text{výsledků}} = 0,1563 & \bar{r}_{\text{umístění}} = 0,9371 & \bar{r}_{\text{umístění elity}} = 0,5884 \\ \overline{MSE}_{\text{výsledků}} = 0,2628 & \overline{MSE}_{\text{umístění}} = 232,7 & \overline{MSE}_{\text{umístění elity}} = 76,63 \end{array}$$

10. způsob odhadování výsledků: Použijeme stejnou myšlenku jako v předcházející metodě s tím rozdílem, že se nebudeme dívat na podobné partie z předcházejícího průběhu turnaje, ale odehraných turnajích podobného typu. V případě sledovaného turnaje Czech Open A v roce 2021 se tedy podíváme na turnaje Czech Open A z let 2015–2019.

Tabulka 13: kvalita 10. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{\rho}_{\text{výsledků}} = 0,1775 & \bar{\rho}_{\text{umístění}} = 0,9377 & \bar{\rho}_{\text{umístění elity}} = 0,5904 \\ \overline{MSE}_{\text{výsledků}} = 0,2594 & \overline{MSE}_{\text{umístění}} = 230,1 & \overline{MSE}_{\text{umístění elity}} = 77,60 \end{array}$$

11. způsob odhadování výsledků: Stejný způsob provedeme potřetí. K odhadu pravděpodobností použijeme podíly z podobných partií z hlediska rozdílu el hráčů, které se odehrály v turnajích Czech Open A, B, C, D a Open Praha v letech 2015–2019.

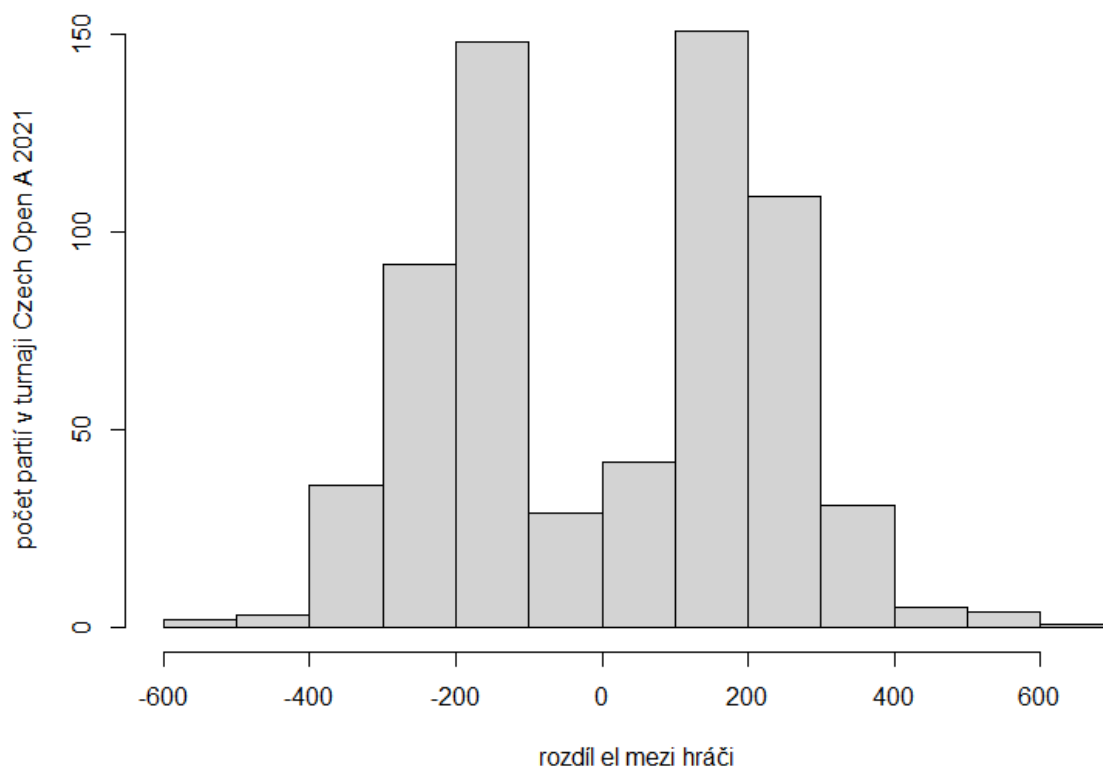
Tabulka 14: kvalita 11. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{\rho}_{\text{výsledků}} = 0,1328 & \bar{\rho}_{\text{umístění}} = 0,9315 & \bar{\rho}_{\text{umístění elity}} = 0,5695 \\ \overline{MSE}_{\text{výsledků}} = 0,2849 & \overline{MSE}_{\text{umístění}} = 253,3 & \overline{MSE}_{\text{umístění elity}} = 83,22 \end{array}$$

Slabinou předchozí trojjediné metody je fakt, že odhad se opírá o partie, jejichž očekávaný výsledek se liší nestejně významným způsobem. Například k předpovědi výsledku partie, v níž byl bílý silnější o 400 bodů a tedy $W_e = 0,92$ je docela rozumné použít i výsledek partie, v níž byl bílý silnější o 350 bodů a měl tedy $W_e = 0,89$. Naproti tomu použít k předpovědi výsledku partie, v níž byl bílý silnější o 25 bodů, výsledek partie, v níž byl bílý o 25 bodů slabší, to se výhodné nezdá, už proto, že v prvním případě je $W_e = 0,53$, zatímco ve druhém je $W_e = 0,47$.

Možným řešením této slabiny by byla úprava šíře intervalu, z něž partie vybíráme, v závislosti na vzdálenosti od 0, protože pokud není rozdíl el hráčů daleko od 0, je ve smyslu proměnlivosti W_e cítit každý elo bod mezi hráči cítit. Jenže vzhledem k tomu, že analyzujeme turnaje hrané švýcarským systémem, právě v této oblasti, kde by bylo záhodno interval zúžit, máme k dispozici zdaleka nejméně dat. Roli hraje způsob nasazování v rámci skupiny hráčů se stejným počtem bodů, podrobněji viz [7].

Graf 4: počet partií odehraných turnaji Czech Open A v roce 2021 v závislosti na rozdílu el hráčů



12. způsob odhadování výsledků: V návaznosti na výše uvedenou výtku zkusíme znovu odhadnout pravděpodobnosti jednotlivých výsledků partie, v níž se el hráčů liší o D . Označme d rozdíel el v již odehrané partii. Pokud $d \in [D - 50; D + 50]$, pak tuto partii zahrneme do odhadu. Vylepšení oproti metodám 9 až 11 bude spočívat v tom, že každé takové partii přiřkneme určitou váhu ω v závislosti na $|D - d|$.

Označme ϕ distribuční funkci normálního rozdělení se střední hodnotou 0 a zvolenou směrodatnou odchylkou $\sigma = 25$. Váhu ω jednotlivých partií určíme jako

$$\omega = \phi(|D - d|) \quad (37)$$

Nyní například odhad pravděpodobnosti výhry bílého vypočteme jako součet všech ω již odehraných partií, v nichž bílí vyhráli, děleno součet všech ω .

V této metodě se budeme dívat výhradně na partie odehrané v předcházejícím průběhu daného turnaje, v našem případě Czech Open A z roku 2021.

Tabulka 15: kvalita 12. způsobu odhadování výsledků

$\bar{\rho}_{\text{výsledků}} = 0,1607$	$\bar{\rho}_{\text{umístění}} = 0,9332$	$\bar{\rho}_{\text{umístění elity}} = 0,5726$
$\overline{MSE}_{\text{výsledků}} = 0,2846$	$\overline{MSE}_{\text{umístění}} = 246,9$	$\overline{MSE}_{\text{umístění elity}} = 83,08$

13. způsob odhadování výsledků: Zopakujeme všechny myšlenky z metody 12 s tím rozdílem, že se nebudeme dívat na předcházející průběh turnaje, ale na partie odehrané obecně v turnajích daného typu, v našem případě tedy na turnajích Czech Open A v letech 2015–2019.

Tabulka 16: kvalita 13. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{\rho}_{\text{výsledků}} = 0,1821 & \bar{\rho}_{\text{umístění}} = 0,9342 & \bar{\rho}_{\text{umístění elity}} = 0,5810 \\ \overline{MSE}_{\text{výsledků}} = 0,2795 & \overline{MSE}_{\text{umístění}} = 243,2 & \overline{MSE}_{\text{umístění elity}} = 83,40 \end{array}$$

14. způsob odhadování výsledků: Zopakujeme všechny myšlenky z metody 12 s tím rozdílem, že se nebudeme dívat na předcházející průběh turnaje, ale na partie odehrané ve všech turnajích, které máme k dispozici, tedy turnaje Czech Open A, B, C, D a Open Praha z let 2015–2019.

Tabulka 11: kvalita 14. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{\rho}_{\text{výsledků}} = 0,1450 & \bar{\rho}_{\text{umístění}} = 0,9316 & \bar{\rho}_{\text{umístění elity}} = 0,5981 \\ \overline{MSE}_{\text{výsledků}} = 0,2866 & \overline{MSE}_{\text{umístění}} = 252,7 & \overline{MSE}_{\text{umístění elity}} = 77,77 \end{array}$$

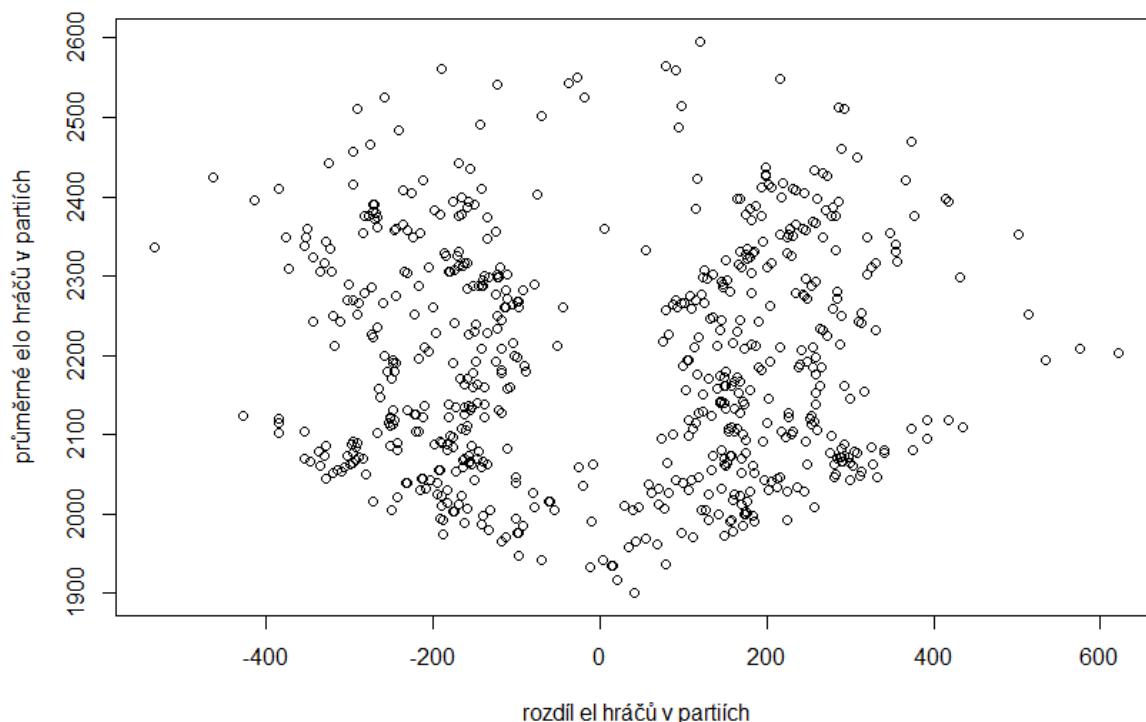
15. způsob odhadování výsledků: V bakalářské práci jsme se mimo jiné zabývali otázkou, zda určitý rozdíl el přináší stejnou výhodu silnějším hráčům na různých typech turnajů, a v některých případech jsme těmto domněnkám dali za pravdu. Nezodpověděli jsme ovšem otázku, jestli je dominantním důvodem jiný typ turnaje (například turnaj žen oproti turnaji mládežníků) nebo rozdílná síla hráčů, kteří se na turnajích potkávají. V další metodě si výběr již odehraných partií omezíme nejenom podobným rozdílem el hráčů, ale i podobným průměrem el obou hráčů.

Chceme předpovědět výsledek partie, v níž se el hráčů liší o D a průměr jejich el je P . Rozdíl el v již odehraných partiích v předcházejícím průběhu turnaje označíme d a průměr el hráčů v těchto partiích p . K tvorbě odhadu použijeme jen partie, které budou svými parametry odpovídat nerovnici:

$$\sqrt{(D - d)^2 + (P - p)^2} < 50 \quad (38)$$

Věnujeme pozornost následujícímu grafu. Jednotlivé body představují partie, které se odehrály na turnaji Czech Open A v roce 2021. Vzorec výše lze interpretovat jako kruh o poloměru 50 se středem v bodě $[D; P]$. Stejně jako v předcházejícím grafu nás může zarazit prostor uprostřed grafu, v němž je nižší hustota dat. I to je způsobeno vlastnostmi švýcarského systému.

Graf 5: partie odehrané na turnaji Czech Open A v roce 2021 v závislosti na rozdílu el hráčů a průměru jejich el



Tuto metodu ještě vylepšíme analogicky jako předcházející tři přidělením vah jednotlivým partiím. Jak s nimi budeme pracovat, jsme vysvětlili už v metodě 12. Zde jen uvedme, jak váhu ω vůbec vypočteme. Jestliže ϕ je distribuční funkce normálního rozdělení se střední hodnotou 0 a zvolenou směrodatnou odchylkou 25, pak

$$\omega = \phi \left(\left| \sqrt{(D-d)^2 + (P-p)^2} \right| \right) \quad (39)$$

Zdůrazněme, že v této metodě se budeme opírat jen o partie odehrané v předcházejícím průběhu daného turnaje, v našem případě Czech Open A v roce 2021.

Tabulka 18: kvalita 15. způsobu odhadování výsledků

$\bar{\rho}_{\text{výsledků}} = 0,1409$	$\bar{\rho}_{\text{umístění}} = 0,9364$	$\bar{\rho}_{\text{umístění elity}} = 0,5724$
$\overline{MSE}_{\text{výsledků}} = 0,2682$	$\overline{MSE}_{\text{umístění}} = 235,1$	$\overline{MSE}_{\text{umístění elity}} = 78,26$

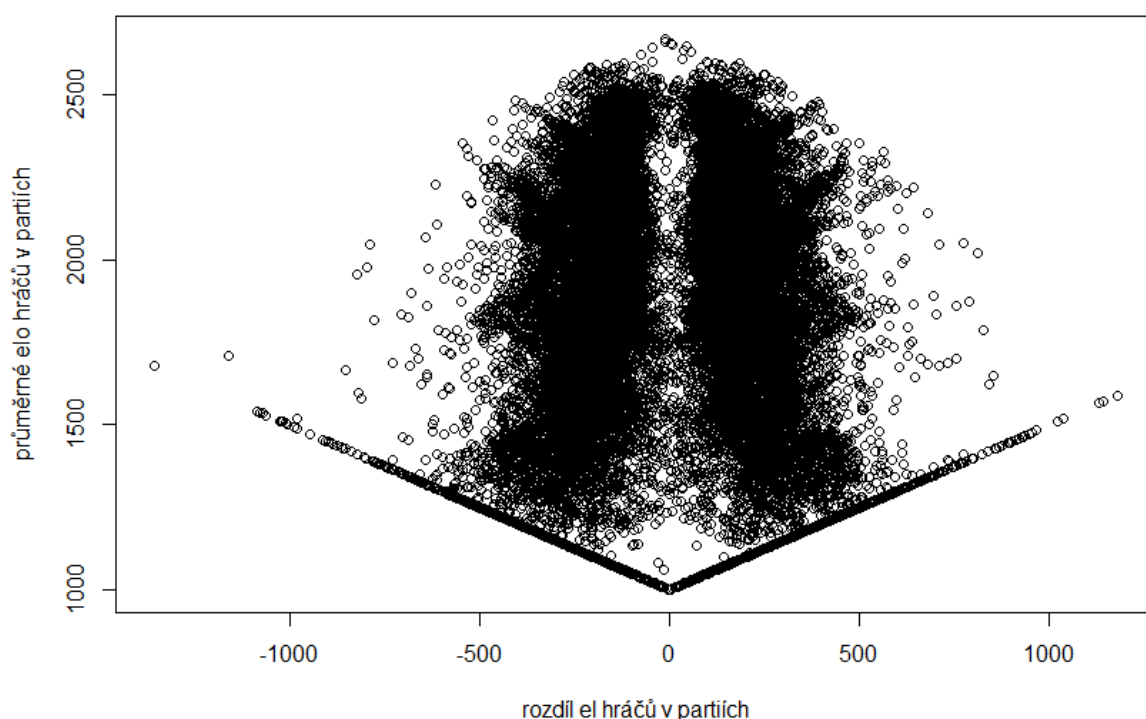
16. způsob odhadování výsledků: V této metodě zopakujeme všechny myšlenky metody 15 s tím rozdílem, že se nebudeme opírat o partie z předcházejícího průběhu daného turnaje, ale o partie odehrané v turnajích daného typu, v našem případě v turnajích Czech Open A v letech 2015–2019.

Tabulka 19: kvalita 16. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{\rho}_{\text{výsledků}} = 0,1739 & \bar{\rho}_{\text{umístění}} = 0,9363 & \bar{\rho}_{\text{umístění elity}} = 0,6294 \\ \overline{MSE}_{\text{výsledků}} = 0,2639 & \overline{MSE}_{\text{umístění}} = 235,5 & \overline{MSE}_{\text{umístění elity}} = 69,99 \end{array}$$

17. způsob odhadování výsledků: V této metodě znovu zopakujeme všechny myšlenky metody 15. Jen zohledníme partie ze všech turnajů, které máme k dispozici, tedy turnaje Czech Open A, B, C, D a Open Praha z let 2015–2019. Výrazná odlišnost od 15. metody spočívá v daleko větším počtu vzorových partií, srovnajme s předchozím grafem. Tam jsme měli k dispozici 653 partií, zde to je 34 398.

Graf 6: partie odehrané na turnajích Czech Open A, B, C, D a Open Praha v letech 2015–2019 v závislosti na rozdílu el hráčů a průměru jejich el



Tvar ve tvaru písmene V v grafu x je důsledkem partií s hráči s elem 1000. Těch je totiž hodně vzhledem k tomu, že elo 1000 dostane každý nový hráč, dokud si nevyhraje vlastní elo.

Tabulka 20: kvalita 17. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{\rho}_{\text{výsledků}} = 0,1574 & \bar{\rho}_{\text{umístění}} = 0,9354 & \bar{\rho}_{\text{umístění elity}} = 0,6207 \\ \overline{MSE}_{\text{výsledků}} = 0,2713 & \overline{MSE}_{\text{umístění}} = 238,7 & \overline{MSE}_{\text{umístění elity}} = 71,12 \end{array}$$

18. způsob odhadování výsledků: Je důvodné se domnívat, že rozdíl el je důležitější údaj než průměrné elo soupeřů. Možná by proto nebylo od věci kruh z metod 15 až 17 protáhnout do tvaru elipsy tak, abychom k předpovědi brali víc partií s podobným rozdílem el i třeba s výraznějším rozdílem v průměru el hráčů. Poradíme si jednodušším způsobem, zvolíme obdélník.

Jestliže si uchováme označení D rozdíl el v odhadované partii a d v již odehraných partiích a P průměr el v odhadované partii a p v již odehraných partiích, pak řekneme, že k odhadu použijeme jen ty odehrané partie, pro které platí, že

$$d \in [D - 30; D + 30] \wedge p \in [P - 100; P + 100] \quad (40)$$

Podíly výher bílých, remíz a proher bílých v tomto obdélníku vezmeme jako odhady pravděpodobností příslušných výsledků.

Podotkněme, že v této metodě se opíráme jen o partie odehrané v předcházejícím průběhu turnaje, v analyzovaném případě tedy turnaji Czech Open A v roce 2021.

Tabulka 21: kvalita 18. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{\rho}_{\text{výsledků}} = 0,1553 & \bar{\rho}_{\text{umístění}} = 0,9367 & \bar{\rho}_{\text{umístění elity}} = 0,5506 \\ \overline{MSE}_{\text{výsledků}} = 0,2642 & \overline{MSE}_{\text{umístění}} = 234,1 & \overline{MSE}_{\text{umístění elity}} = 82,17 \end{array}$$

19. způsob odhadování výsledků: Tato metoda odpovídá metodě 18 jen místo předcházejícího průběhu turnaje využijeme partie ze všech turnajů podobného typu, tedy turnajů Czech Open A z let 2015–2019.

Tabulka 22: kvalita 19. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{\rho}_{\text{výsledků}} = 0,1732 & \bar{\rho}_{\text{umístění}} = 0,9365 & \bar{\rho}_{\text{umístění elity}} = 0,6040 \\ \overline{MSE}_{\text{výsledků}} = 0,2652 & \overline{MSE}_{\text{umístění}} = 234,8 & \overline{MSE}_{\text{umístění elity}} = 74,30 \end{array}$$

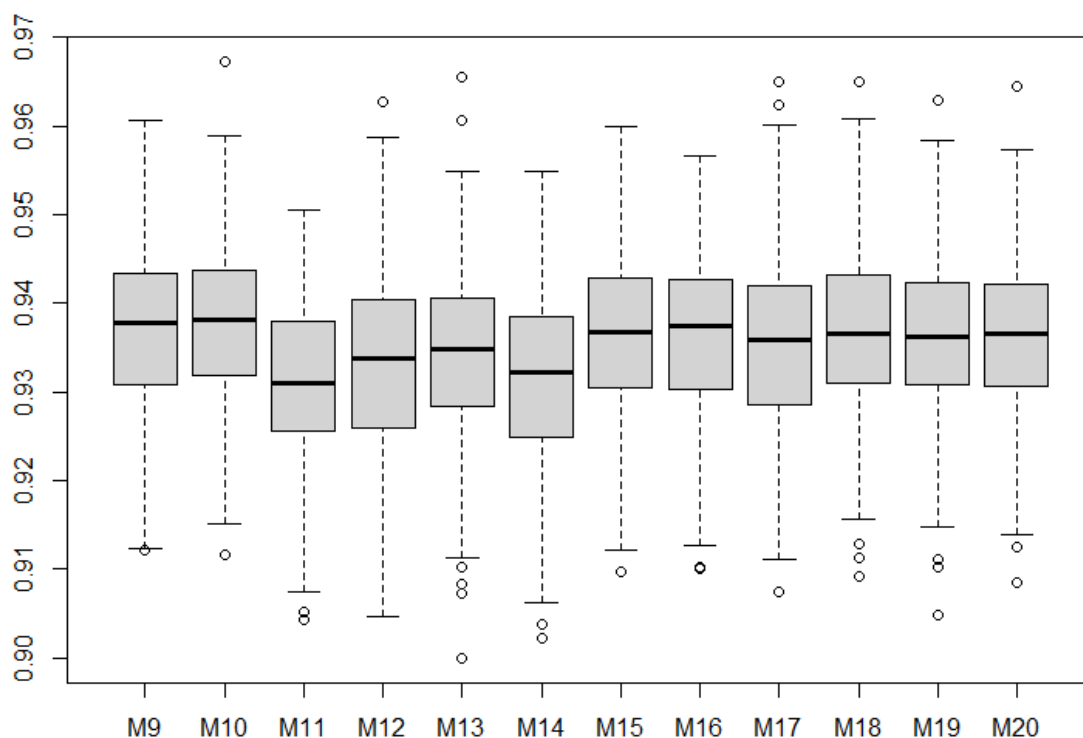
20. způsob odhadování výsledků: Dvacátý způsob odhadování výsledků koresponduje s metodou číslo 18. Místo partií předcházejícího průběhu turnaje jsou zohledněny všechny partie z dostupných turnajů, tedy turnajů Czech Open A, B, C, D a Open Praha z let 2015–2019.

Tabulka 23: kvalita 20. způsobu odhadování výsledků

$$\begin{array}{lll} \bar{\rho}_{\text{výsledků}} = 0,1784 & \bar{\rho}_{\text{umístění}} = 0,9364 & \bar{\rho}_{\text{umístění elity}} = 0,6039 \\ \overline{MSE}_{\text{výsledků}} = 0,2652 & \overline{MSE}_{\text{umístění}} = 235,3 & \overline{MSE}_{\text{umístění elity}} = 75,10 \end{array}$$

2.5 Srovnání metod založených na turnajové praxi

Graf 7: korelace odhadnutých a skutečných výsledků v turnaji Czech Open A 2021 v metodách 9 až 20



Krabicový diagram překvapuje vyrovnanou úspěšností jednotlivých metod. Kromě toho, že 11. metoda se zdá být přes svou podobnost horší než metody 9 a 10, neshledáváme pohledem žádné významné rozdíly. Čtenář si může v příloze D prohlédnout podobné grafy z hlediska korelace odhadových a skutečných umístění a z hlediska odhadovaných a skutečných umístění elitních hráčů.

V příloze E uvádíme výsledky analýz rozptylu jednotlivých metod aplikovaných na turnaje Czech Open A, B, C, D a Open Praha v letech 2015–2019. Zde konstatujeme, že z hlediska všech tří kritérií s velkou jistotou zamítáme hypotézu o tom, že všechny turnaje se předvídají našimi metodami stejně dobře, ovšem nezamítáme hypotézy o tom, že mezi kvalitou metod nejsou žádné významné rozdíly, a že úspěšnost jednotlivých metod nezávisí na volbě turnaje.

2.6 Srovnání všech metod jako celku

Několik stránek jsme představovali metody jak odhadovat výsledky partií posledního kola a průběžně jsme je stručně komentovali. Teď se pokusíme podívat se na všech 20 metod s určitým nadhledem. Začneme tím, že provedeme analýzu rozptylu dvojnásobného třídění s interakcemi podle korelací odhadovaných a skutečných umístění.

Tabulka 24: Analýza rozptylu dvojného třídění s interakcemi metod 1 až 20 podle korelace odhadnutých a skutečných umístění elitních hráčů

zdroj	df	SS	MS	F	p-hodnota
měnlivosti					
typ turnaje	4	0,0259	0,0064	223,4	0,0000
metoda	19	0,0197	0,0010	35,84	0,0000
interakce	76	0,0035	0,0000	1,601	0,0022
reziduální	400	0,0116	0,0000		

Tabulka 24 nám zaprvé říká, že z hlediska všech 20 metod můžeme s velkou jistotou zamítnout hypotézu o tom, že všechny turnaje jsou stejně dobře předvídatelné. Můžeme proto nasadit Tukeyovu metodu mnohonásobného porovnávání, abychom viděli, které turnaje se předvídají lépe než ostatní.

Tabulka 25: Porovnání turnajů podle jejich předvídatelnosti z hlediska korelace odhadnutých a skutečných umístění

Open Praha	Czech Open D	Czech Open A	Czech Open C	Czech Open B
0,9406				
	0,9318			
		0,9287		
			0,9216	
				0,9212

Není překvapivé, že nejlépe předvídatelným turnajem je Open Praha. Jde o poměrně silný turnaj, kde můžeme očekávat účast hráčů se stabilizovanou výkonností a na rozdíl od letního Czech Open v Pardubicích nemá tak vysokou účast cizinců, kteří typicky v turnajích Czech Open B a Czech Open C podávají překvapivé výkony. Není jasné, proč je lépe předvídatelným turnajem Czech Open D než Czech Open A. V obecné rovině však pro nás tyto informace jsou důležité jen pro uvědomění, že ne všechny turnaje jsou pro předpověď výsledků stejně vhodné. Obsahem této práce není postihování detailních rozdílů mezi turnaji.

Daleko důležitější je pro nás druhý řádek tabulky x, který říká, že s velkou jistotou zamítáme hypotézu o tom, že by všechny metody byly stejně efektivní. Jsou mezi nimi rozdíly a Tukeyova metoda mnohonásobného porovnávání nám ukáže, mezi kterými metodami.

Tabulka 26: Porovnání všech 20 metod z hlediska korelace odhadnutých a skutečných umístění

M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
18	19	9	13	12	20	11	10	17	16	15	14	7	8	1	6	3	5	2	4
,93	,93	,93	,93	,93	,93	,93	,93	,93	,93	,93	,93	,93	,93	,93					
								,93	,93	,93	,93	,93	,93	,92					
															,92	,92	,91	,91	
																			,91

Už jsme několikrát zmínili, že metoda 4 postavená na porovnání výkonů hráčů v předcházejícím průběhu turnaje se neosvědčila. Metody 2–3 a 5–6, které se nejvíc drží myšlenek elo systému mají horší výsledky než nahodilé vybírání výsledků metodou 1.

Z jednoduchých metod se nejvíc osvědčily metody 7 a 8 založené na kombinaci elových předpokladů a pohyblivého odhadu remízy v závislosti na předcházejícím průběhu turnaje. Jejich výhodou je, že jsou výpočetně nenáročné.

Spolehlivé výsledky podávají metody 9–20 založené na turnajové praxi. Zajímavé ovšem je, že se mezi nimi neprojevují významné rozdíly. Při dalším zkoumání bychom určitě na některé narazili, relativní vyrovnanost metod nám však umožňuje zvolit pro další práci tu, která nám vyhovuje nejvíc. Pro řešení motivační úlohy si vybereme metodu číslo 18. Čtenář se může přesvědčit prohlédnutím analýz rozptylů a srovnání metod v příloze F, že nemáme k dispozici žádnou významně lepší metodu. Výběr metody číslo 18 zůstává částečně věcí vkusu.

Prakticky užitečná je informace, že se významně neprojevuje rozdíl, jestli čerpáme turnajovou zkušenost jen z předcházejícího průběhu turnaje v řádu stovek partií, z turnajů podobného typu, kde máme k dispozici tisíce partií nebo z databáze různých turnajů, kde máme k dispozici desetitisíce partií. Čerpat jen z předcházejícího průběhu turnaje je totiž výrazně jednodušší na implementaci, není potřeba složitě získávat data z množství podobných turnajů. Dá se ovšem očekávat, že kdybychom vycházeli jen z průběhu malého turnaje s nízkou hráčskou účastí, nedostatek partií by se mohl negativně projevit.

Rozdíly mezi některými skupinami metod jsou sice průkazné, ovšem mějme stále na paměti, že se pohybujeme zhruba v rozmezí dvou až tří setin korelačního koeficientu. Nepodařilo se nám vyvinout metodu, která by byla nějak daleko lepší než náhodné tipování výsledků v metodě číslo 1. Na druhou stranu každé zlepšení má svou cenu vzhledem k tomu, že výsledek šachové partie zkrátka není determinován jen ely obou hráčů a dosáhnout korelace odhadů se skutečností hodnoty například 0,99 je patrně vyloučené.

Poslední informace, která vyplývá z tabulky x je skutečnost, že s jistotou zamítáme hypotézu o tom, že úspěšnost předpovědi nezáleží na vzájemné interakci mezi metodou a typem turnaje. Z toho vyplývá, že pro účely získání co nejvyšší kvality předpovědi může mít smysl ladit metodu na míru tomu kterému turnaji.

3 Vyhlídky hráče před posledním kolem turnaje

V tuto chvíli se můžeme vrátit do odstavce 1.4, v němž jsme si ukázali motivační úlohu. Jsme v předvečer posledního kola turnaje hraného švýcarským systémem, ve kterém rozhoduje počet bodů a pomocným hodnocením je Buchholz cut 1. Cenový fond činí 50 000 Kč pro 1. místo, pro 2. místo 30 000 Kč, pro 3. místo 20 000 Kč, pro 4. místo 15 000 Kč, pro 5. místo 10 000 Kč, pro 6. místo 8 000 Kč, pro 7. místo 6 000 Kč, pro 8. místo 5 000 Kč a pro 9. a 10. místo 3 000 Kč. Ceny se rozdělují systémem hort. Rozlosování posledního kola elitních hráčů je následující:

Tabulka 27: los posledního kola turnaje Czech Open A v roce 2021

hráč	elo	bucholz cut 1	počet bodů	výsledek	počet bodů	bucholz cut 1	elo	hráč
Babula	2563	35,5	6,5		6,5	32,5	2306	Czerw
Štoček	2604	35	6		6	32,5	2514	Liviac
Zwardoń	2466	40	6		6	37	2536	Pavlidis
Plát	2525	37	6		6	35,5	2291	Mládek

dále 5 dvojic hráčů majících na kontě 5,5 bodu

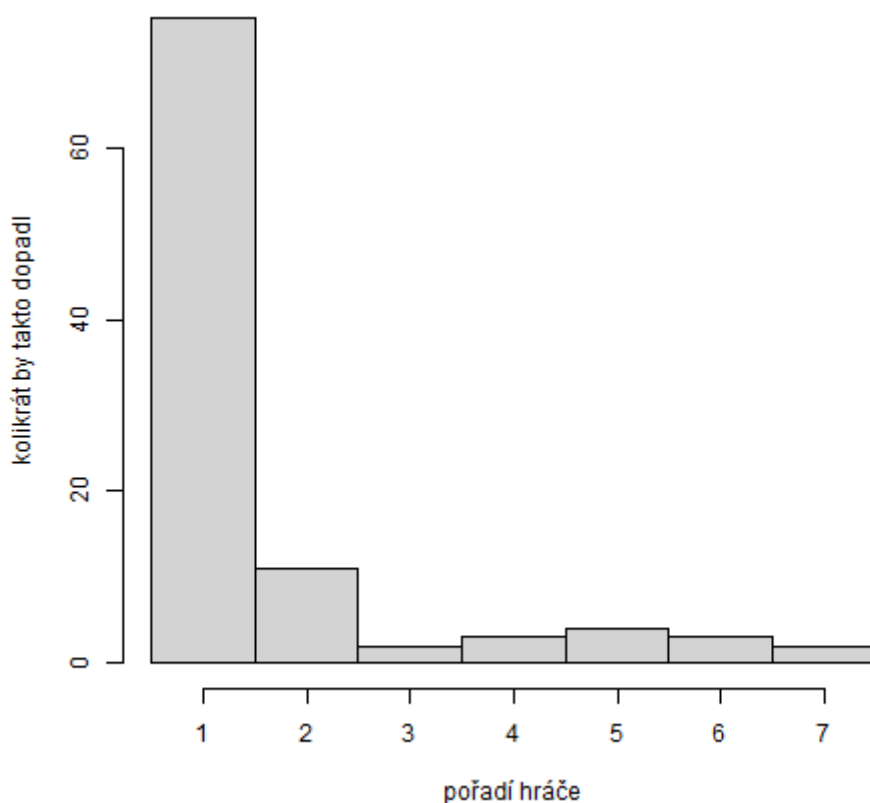
dále 11 dvojic hráčů majících na kontě 5 bodů

a dalších 51 dvojic s hráči majících na kontě méně než 5 bodů

3.1 Očekávané pořadí hráče

Abychom pomohli hráči Babulovi k orientaci v turnajové situaci, odhadneme stokrát výsledek všech 71 partií posledního kola daného turnaje a s pomocí vlastního programu se podíváme, na jaká umístění by sto takových dohrání hráči Babulovi vydalo.

Graf 8: umístění hráče Babuly ve 100 nasimulovaných průbězích posledního kola



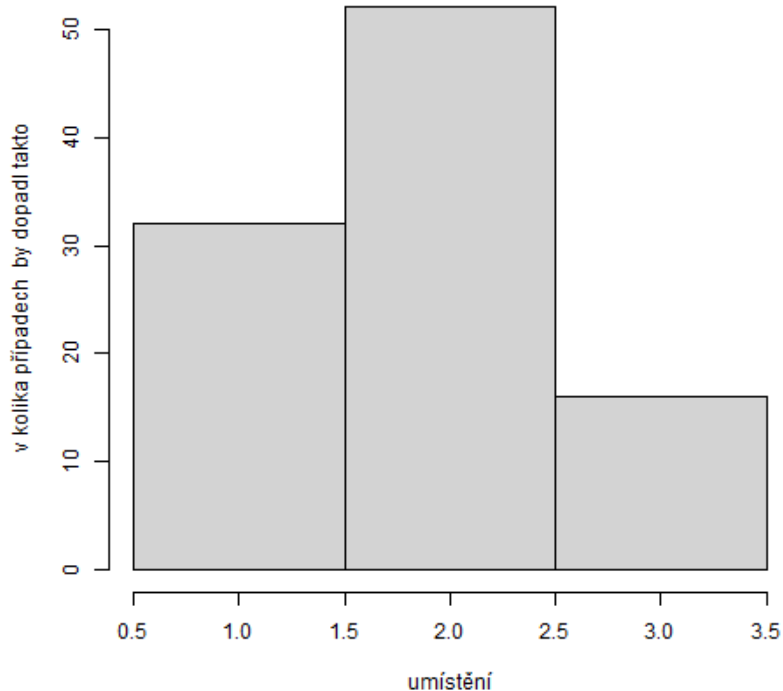
Díky vytvořenému programu máme představu o tom, jak může hráč Babula v turnaji skončit. Očekávané umístění je vítězství, možná jsou ale i umístění mimo první pětku. Můžeme dodat, že průměrné umístění v nasimulovaném konci turnaje hráče Babuly je 1,67.

Jenže k čemu je taková informace dobrá? Hráč přece ví, že před posledním kolem drží první místo a nejuvýhodnější pozici před koncem turnaje, rozhodně nepotřebuje složitou analýzu k tomu, abychom mu řekli, že dost možná vyhraje. Ostatně na to jsme přišli už při povrchní analýze v rámci první kapitoly. Pro hráče by mohlo být daleko zajímavější informací mít představu o tom, jak dopadne, když v posledním kole vyhraje, remizuje nebo dokonce prohraje.

3.2 Podmíněné očekávané pořadí hráče

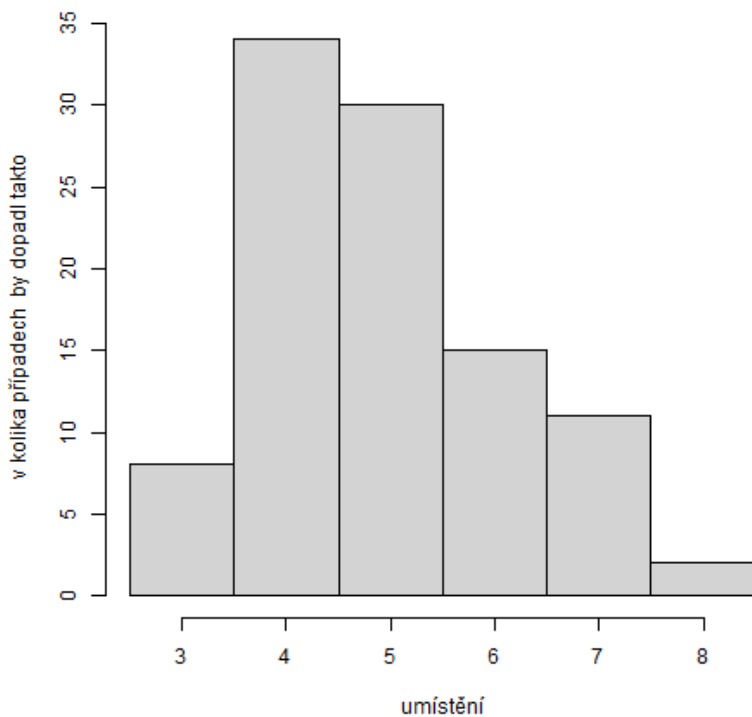
Když hráč Babula ve své partii posledního kola vyhraje, potom zcela jistě vyhraje celý turnaj, protože žádný z hráčů se na něj nemůže bodově dotáhnout, to už víme. Co se stane, když hráč Babula ve své partii posledního kola remizuje?

Graf 9: umístění hráče Babuly ve 100 nasimulovaných průbězích posledního kola, když v něm remizuje



A to už je výrazně zajímavější informace, simulační postup ukázal, že remíza hráče Babulu pravděpodobně podrží na stupních, očekáváme, že v takovém případě skončí druhý, ale možné je i 1. a 3. místo. Průměrné očekávané umístění za předpokladu remízy v posledním kole je 1,84. A konečně co bude pro hráče Babulu znamenat případná prohra v posledním kole?

Graf 10: umístění hráče Babuly ve 100 nasimulovaných průbězích posledního kola, když v něm prohraje



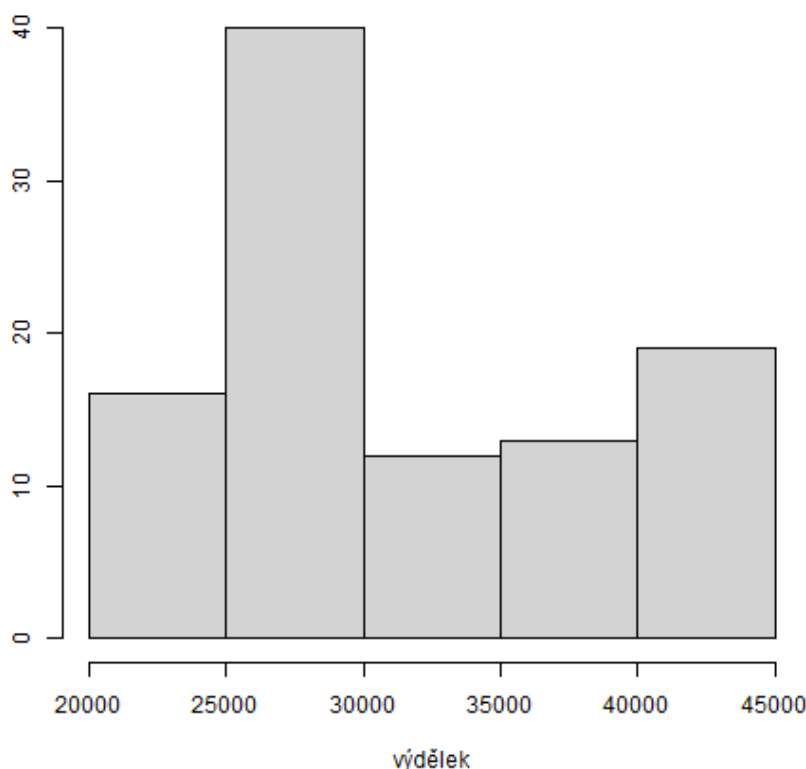
Prohra by pro hráče Babulu znamenala pád z prvního místa, ale jak naznačuje graf, ne zas tak veliký. Jeho průměrné očekávané umístění v takovém průběhu turnaje by bylo 4,93.

Umístění může být pro hráče zajímavou a dobře uchopitelnou informací. Ve výjimečných případech z turnajů hráči umístění do určitého místa postupují na nějaký lepší turnaj, pak jde o zprávu prvořadého významu. Pokud ale hráč po očku sleduje finanční prémie, není mu informace jen o umístění nic platné. Pokud jsou totiž ceny dělené systémem hort, nemají jednotlivá umístění daný finanční ekvivalent, neboť záleží i na tom, kolik hráčů získalo stejný počet bodů.

3.3 Dopad na očekávaný finanční zisk hráče

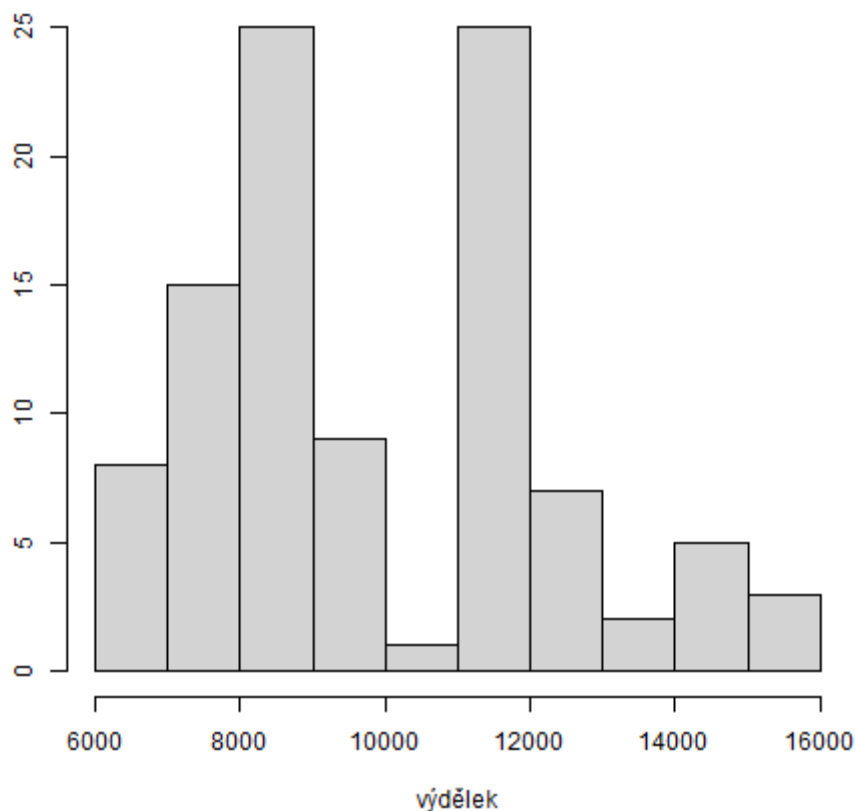
Opět je zcela zřejmé, že pokud hráč Babula v posledním kole vyhraje, stane se vítězem turnaje a získá 50 000 Kč. Očekávané rozdělení finančních prémie v případě remízy ukazuje následující histogram.

Graf 11: zisk hráče Babuly ve 100 nasimulovaných průbězích posledního kola, když v něm remizuje



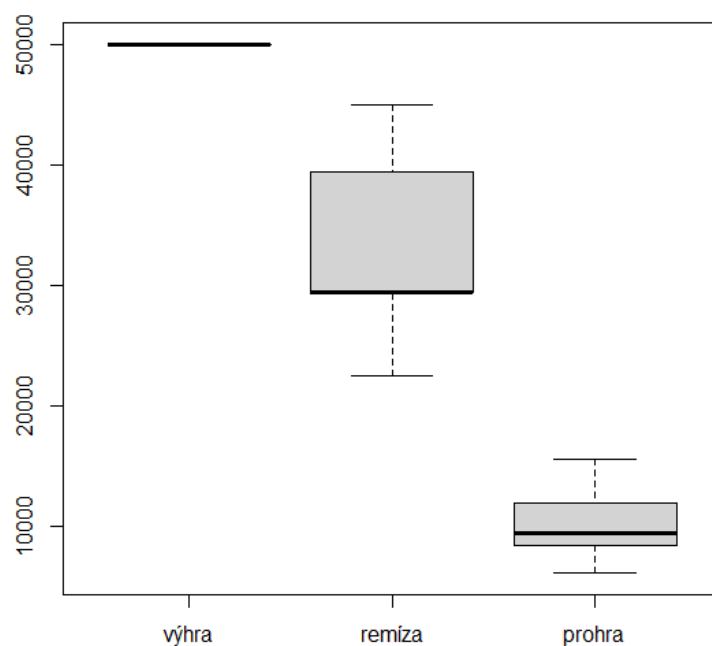
Pokud bychom chtěli grafickou interpretaci zkondenzovat, v případě remízy očekáváme zisk hráče Babuly průměrně 32 314 Kč a to se směrodatnou odchylkou 6560 Kč.

Graf 12: zisk hráče Babuly ve 100 nasimulovaných průbězích posledního kola, když v něm prohraje



Zdá se, že v případě prohry už systém hort výrazně sníží finanční odměnu pro hráče Babulu. Jeho průměrný očekávaný zisk je 10 020 Kč se směrodatnou odchylkou 2345 Kč. K čemu můžou vést různé výsledky hráče Babuly v posledním kole turnaje, ještě shrneme krabicovým diagramem.

Graf 13: celkové možné zisky hráče Babuly v závislosti na výsledku jeho poslední partie



Vidíme, že úspěch v partii posledního kola se v našem případě projevuje finančním zisku docela dramatickým způsobem. Vraťme se nyní ale k původní otázce? Co by měl dělat hráč Babula, kdyby mu jeho soupeř nabídnul hned na začátku partie remízu? Bylo by pro něj rozumné ji přijmout?

Otázku si můžeme zjednodušit. Je očekávaný zisk hráče v partii, tedy při zahrnutí možné výhry, remízy i prohry vyšší, rovný nebo menší než očekávaný zisk v případě remízy? Ponechme označení b (protože má bílé) jako výhra hráče Babuly v posledním kole, r jako remíza hráče Babuly v posledním kole a $č$ jako prohra hráče Babuly v posledním kole. Které ze znamének $=$, $>$ nebo $<$ patří mezi dva následující výrazy na místě otazníku.

$$b \cdot 50\,000 + r \cdot 32\,314 + č \cdot 10\,020 \text{ ? } 32\,314 \quad (41)$$

Určíme pravděpodobnosti b , r a $č$ podle metody číslo 8. Elo hráče Babuly je 2563, a protože má v partii bílé, zvýšíme mu jeho rating o 44 bodů (viz metoda 3) na 2607. Elo hráče Czerwa je 2306, čili rozdíl D činí 301 elo bodů. Umíme určit, že $W_e(301) = 0,85$ a při popisu metody 7 jsme zjistili, že $\check{b} = 38,9\%$ a $\check{č} = 26,4\%$. S pomocí těchto čísel a vzorců (34) a (35) vypočteme $b = 58,8\%$, $č = 9,0\%$ a $r = 32,2\%$. Dosazením do vzorce (41) dostáváme na levé straně 40 555 Kč. To znamená, že očekávaný zisk při hraní partie i při zahrnutí rizika prohry je vyšší než očekávaná výše prémie za okamžitě přijatou remízu. Hráči Babulovi bychom proto na základě těchto úvah přijetí hypotetické nabídky remízy nedoporučili. Ve skutečnosti partie skončila remízou už ve 28. tahu a to po nabídce hráče Babuly.

Parametry tohoto příkladu byly ovlivněny skutečností, že hráč Babula byl v partii podle ratingových předpokladů jasným favoritem navíc s bílými figurami, proto jsme mu nemohli doporučit přijmout hypotetickou nabídku remízy hned po začátku partie. Zaměřme ještě krátce pozornost na třetí šachovnici, kde se utkali hráči Zwardoň a Pavlidis. Následující tabulka shrnuje stejným způsobem vypočtené očekávané zisky v případě jednotlivých výsledků obou dvou hráčů.

Tabulka 28: Možné zisky dalších hráčů v závislosti na jejich poslední vzájemné partii

Zwardoň	očekávaný zisk	Pavlidis	očekávaný zisk
vyhraje ($b = 0,305$)	30 216 Kč	vyhraje ($č = 0,338$)	28 341 Kč
remizuje ($r = 0,357$)	13 490 Kč	remizuje ($r = 0,357$)	10 550 Kč
prohraje ($č = 0,338$)	2 318 Kč	prohraje ($b = 0,305$)	1 938 Kč
odmítne remízu	14 815 Kč	odmítne remízu	13 936 Kč

Oba hráči, Zwardoň i Pavlidis by podle našeho modelu měli odmítnout hypotetickou nabídku v úvodu partie odmítnout, protože výše odměny z případné výhry mírně převažuje nad hrozbou hlubokého propadu v případě prohry. Jejich příklad si detailně ukazujeme proto, že v jejich případě jsou

obě varianty poměrně vyrovnané a volba partiové strategie je proto do značné míry věcí vkusu, jde skoro o rozhodnutí mezi pomyslným vrabcem v hrsti a holubem na střeše. Na taková rozhodnutí je náš model poměrně krátký a je limitovaný četnými omezeními, jak si ještě řekneme v podkapitole 3.4.

Na tomto místě jen doplníme, jak vlastně celý námi zkoumaný turnaj dopadl. Na prvních dvou šachovnicích hráči remizovali, na třetí hráč Zwardoń porazil svého soupeře a na čtvrté hráč Mládek svého soupeře. Na 7 bodů proto dosáhla čtveřice Babula, Czerw, Zwardoń a Mládek, kteří se podle pomocného hodnocení srovnali do pořadí Zwardoń, Babula, Mládek a Czerw. Z toho důvodu brali po řadě 39 375 Kč, 29 375 Kč, 24 375 Kč a 21 875 Kč.

3.4 Význam očekávaných výsledků turnaje a možnosti využití

Na závěr projdeme, co všechno náš systém opomíjí, a jak moc to brání úspěšnému využívání informací, kterých už umíme s pomocí vlastní funkce v R dosáhnout.

Od začátku mlčky předpokládáme, že jednotlivé partie posledního kola jsou nezávislé, a že se tedy šachista může soustředit výhradně na svoji hru a výsledky ostatních hráčů chápat jako náhodné řídicí se nějakými předem stanovenými pravděpodobnostmi. To je ale zásadní zjednodušení. Všechny šachovnice aspirantů na finanční prémii se zpravidla nachází v jedné hrací místnosti a hráči tak v reálném čase a snad i přímo ze své židle vidí, jak se vyvíjí situace na ostatních šachovnicích a s nimi jejich vyhlídky na konkrétní umístění v turnaji. Pokud by například v analyzovaném turnaji skončily remízou nebo k ní nevyhnutelně šly partie na 2. až 4. šachovnici, mohl by si být hráč Babula prakticky jistý, že k vítězství v turnaji mu v tu chvíli stačí remíza, protože potom by měl 7 bodů jen on a jeho soupeř Czerw, který měl ale před posledním kolem výrazně horší pomocné hodnocení.

Z podobných důvodů jsme od začátku upřesňovali otázku o přijetí nebo odmítnutí hypotetické nabídky remízy tím, že by padla bezprostředně po zahájení partie, tedy tak, aby nešlo brát zřetel na vývoj na ostatních šachovnicích. Přitom už první tahy partií mohou napovědět náchylnost hráčů k přijetí remízy, když jsou například známi tím, že některá zahájení volí jen tehdy, jestliže potřebují hrát nekompromisně na výhru. A to nemluvíme o tom, že zejména na lokálních turnajích se špička turnaje může vzájemně znát. Vedle dosavadních výsledků a losu posledního kola se pak může stát i hodiny před začátkem kola obecně známou informací i to, že dva soupeři plánují brzy z místa turnaje odcestovat a dají svou poslední partii brzy za remízu.

Na některých turnajích (ne na námi analyzovaných) se používají různé mutace tzv. sofijského pravidla, jež zapovídá nabídky remízy před určitým předem stanoveným tahem. Taková omezení mohou vnést do turnaje ještě větší chaos. Když si představíme dvojici soupeřů, pro něž je pro oba nejvýhodnější přijmout remízu, je těžké jim bránit, aby svou partii remizovali. Kolikrát může být potom jakési předstírání hry podstatně nedůstojnější než remíza podepsaná svobodně po prvním tahu partie.

Na reálném turnaji se utkávají konkrétní šachisté. Šachisté, kteří mohou mít z minulosti specifickou vzájemnou bilanci, kteří si nemusí vyhovovat svými herními styly a podobně. Ostatní účastníci turnajů mohou jen na základě své šachové odbornosti určit pravděpodobnosti jednotlivých výsledků daleko lépe, než sebelepší metody těžící co nejvíce informací z el hráčů. Zkrátka výsledek šachové partie není determinován jen ely hráčů.

Má smysl se ptát, jestli přes to všechno odhadování výsledků partií posledního kola může přinést hráčům dostatek užitečného. Jestli odhadnuté výsledkové listiny a jejich interpretace jsou lepší než intuitivní úvahy nepodložené pravděpodobnostní teorií. Odpovědná odpověď by vyžadovala provést řádný průzkum mezi šachisty, kteří se do podobných situací mohou vzhledem ke své výkonnosti dostat. Taková cesta by ovšem překročila rozsah práce.

Na druhou stranu, co hráč potřebuje? Základní orientaci před partií umožňující specifickou přípravu s ohledem na míru nutnosti výhry a míru tolerance k případné remíze. Prověření několika základních podmíněných pravděpodobností s ohledem na možné výsledky svých i ostatních partií simulačními metodami je rychlejší než ruční dopočítávání možných kombinací výsledků. Vhled hráče do zákonitostí posledního kola konkrétního turnaje může být díky všemu, co v práci ukazujeme, lepší.

Jakmile se potom poslední kolo rozeběhne, hráč se může soustředit výhradně na svoji partii. A pokud se po několika hodinách hry bude na některé z šachovnic rodit například nějaký překvapivý výsledek, může s ohledem na něj očekávat, že původní informace z modelu se pohnou některým směrem.

Závěr

Diplomovou prací *Využití elo systému k předpovědi výsledků šachových turnajů* jsem navázal na bakalářskou práci *Elo systém očima matematické statistiky*. V textu jsme se věnovali turnajům hraným švýcarským systémem a pokoušeli jsme se zjistit, jaké jsou možnosti předpovídat konečné pořadí turnaje, známe-li všechny výsledky od prvního do předposledního kola a los posledního kola.

Představili jsme postupně 20 původních metod, jak odhadovat na základě el hráčů výsledky posledního kola. V závislosti na ratingovém rozdílu hráčů dává elo systém poměrně dobrou informaci o tom, jaký podíl bodů očekáváme, že uhraje silnější a jaký slabší hráč. Chceme-li však předpovídat výsledek partie, je žádoucí vymyslet způsob, jak určovat pravděpodobnost remízy v partii. Neosvědčilo se nám používání jednoho fixního odhadu remízy, lepší výsledky jsme získávali, když jsme remízu stanovili pomocí průměrného podílu remíz na očekávaném zisku obou hráčů. Shledáváme překvapivým, že k lepším odhadům nevedlo zvýšení elu bílých o hodnotu ekvivalentní výhodě bílých figur.

O trochu lepší byly metody, které se v předpovědích opíraly o zkušenost z podobných již odehraných partií. Na příkladu velkých turnajů, které jsme analyzovali, se zdálo, že opření se o výsledky partií z předcházejícího průběhu daného turnaje je podobně účinné jako opření se o výsledky partií z řádově větší databáze podobných turnajů.

Čtyři typy výběru podobných partií, které jsme vyvinuli, se v úspěšnosti předpovědi významně nelišily. Přiznáváme, že jsme ponechali poměrně velký prostor k vylepšování těchto způsobů pomocí například změny šíře intervalů jednotlivých parametrů podobnosti nebo lepšího systému přiřazování vah různě podobným partiím. Vzhledem k tomu, že jsme prokázali, že se různé turnaje liší v předvídatelnosti svých výsledků, se zdá být vhodné pro praktické použití metody ladit pro zvolené turnaje.

Domníváme se však, že není snadné dosáhnout významně lepších odhadů vzhledem k tomu, že výsledek šachové partie není zdaleka determinován jen rozdílem el soupeřů, ale mnoha dalšími faktory a ještě více v případě rozhodujících turnajových klání. Celkově jsme se u dobrých metod dostali zhruba na mediánovou korelaci odhadnutých a skutečných výsledků 0,2, odhadnutých a skutečných umístění všech hráčů 0,9 a odhadnutých a skutečných umístění hráčů bojujících o čelo turnaje 0,6.

Největší síla i těch nejhorších vyvinutých metod spočívá v implementaci do R a možnosti simulovat poslední kolo turnaje s větším počtem opakování. I při značném zjednodušení, s nímž pracujeme, dosahujeme prakticky uchopitelných výsledků, které značně překračují možnosti intuice jednotlivých hráčů. S pomocí programu můžeme nabídnout hráči kvantitativní představu o tom, jaké umístění a ve svém důsledku jak vysokou finanční prémie v turnaji může očekávat v závislosti na třech možných výsledcích své partie.

Neumíme odpovědně říct, za jak užitečné považuje získané informace šachová obec. Vnímáme míru zjednodušení, kdy se například na výsledky jednotlivých partií turnaje díváme jako na nezávislé, i když víme, že tomu tak v praxi není, protože hráči mohou upravovat svou turnajovou strategii v závislosti na vývoji ostatních partií. Věříme, že hráč, který podobné informace popisující situaci v jeho turnaji interpretuje při vědomí zmíněných nedostatků, získá výhodu oproti ostatním.

V příloze A čtenář nalezne návod k funkci v R, která umí dělat analýzu libovolného turnaje a v příloze G skript dané funkce v R. Funkce není řádně otestovaná pro neautorské použití, chápu to z hlediska své práce jako bonus. Výhledově bych chtěl funkci odladit a učinit ji dostupnou i pro uživatele nepracující s R. Čtenář, který by s funkcí chtěl pracovat a měl jakékoli otázky nebo podněty se na mě může obrátit na e-mail filipzadrazil.jbc@gmail.com.

Zdroje

- [1] ZADRAŽIL, Filip. Elo systém očima matematické statistiky. Liberec, 2020. Bakalářská práce. Technická univerzita v Liberci.
- [2] Robert a Petr šachy, 2022, Fungování elo systému 1, výpočet a matematické pozadí – F. Zadražil, YouTube video. [cit. 14.07.2022]. Dostupné z: URL:
<https://www.youtube.com/watch?v=op4UXCvSB0w>
- [3] ELO, Arpad. The Rating of Chessplayers: Past and Present. 2. New York: Ishi Press International, 2008. ISBN 0-923891-27-7.
- [4] VESELÝ, Jiří. Psychologický průvodce šachovou partií. 1. Praha: Olympia, 1981.
- [5] FIDE. FIDE Rating Regulations effective from 1 July 2017 [online]. [cit. 09.03.2020]. Dostupné z:
<https://handbook.fide.com/chapter/B022017>
- [6] NOVOTNÝ, Viktor. Klasifikační řád ŠSČR [online]. [cit. 09.03.2020]. Dostupné z:
<https://www.chess.cz/sachovy-svaz-cr/legislativa/klasifikacni-rad-sscr/>
- [7] NAGY, Peter. Švýcarský systém šachových turnajů. Brno, 2016. Bakalářská práce. Masarykova univerzita, Přírodovědecká fakulta, Ústav matematiky a statistiky.
- [8] FIDE. General Rules and Technical Recommendations for Tournaments [online]. [cit. 14.07.2022]. Dostupné z: <https://handbook.fide.com/files/handbook/C02Standards.pdf>
- [9] ANDĚL, Jiří. Statistické metody. Čtvrté upravené vydání. Praha: Matfyzpress, 2007. ISBN 80-7378-003-8.
- [10] HOLČÍK, Jiří, KOMENDA, Martin (eds.) a kol. Matematická biologie: e-learningová učebnice [online]. 1. vydání. Brno: Masarykova univerzita, 2015. ISBN 978-80-210-8095-9.
- [11] HERZOG, Heinz. Chess-Results Server [online]. [cit. 14.07.2022]. Dostupné z:
<http://chess-results.com/>

Přílohy

A. Návod k používání vlastní funkce v R analyzující turnaje

- Funkce umí analyzovat jen vývoj, dopad na výsledkovou listinu a na finanční prémie posledních kol turnajů jednotlivců hraných švýcarským systémem na 9 kol.
- Najdu si svůj turnaj na chess-results.com.
- Klepnu na Starting rank crosstable.
- Pokud má turnaj víc než 150 účastníků, sroluji na konec stránky a klepnu na Show complete list.
- Klepnu na Export to Excel (.xlsx).
- Uložím soubor do pracovního adresáře v R.
- Otevřu si soubor v Excelu a smažu podrobnosti o turnaji (ne hlavičku tabulky) pomocí smazání celých řádků. Stejně tak smažu informace pod tabulkou.
- Dám Ctrl H a u najít a nahradit nahradím všechny výrazy $\frac{1}{2}$ číslem 5.
- Funkce je odladěná na analýzu již proběhlých turnajů. Pokud chci analyzovat právě probíhající turnaj (musí mít už nalosované poslední kolo), musím v Excelu ručně upravit celý sloupeček 9. kola tak, že u všech řádků, kde není 0 nebo -1 dopíšu za výrazy jako 5b a podobně ještě číslici 2.
- Uložím si do pracovního adresáře skript z přílohy G jako funkci v R a pomocí příkazu source ji aktivuji.
- Nyní už můžu s funkcí pracovat. Zadáám do příkazového okna v R zadržím() a do závorek píšu jednotlivé argumenty:
 - Argument nazev. Píšu do uvozovek, jak jsem si pojmenoval excelovský soubor s turnajovou tabulkou. Nepíšu konec s tečkou a koncovkou. Zadáám například nazev="turnaj"
 - Argument Nsim. Kolik simulací chci provést. V případě Nsim=100 můžu očekávat výsledek za zhruba 2 minuty, v případě Nsim=1000 za zhruba 20 minut atp.
 - Argument vyrobít. Defaultně je nastavený na 1. Pokud dělám opakovaně víc výpočtů se stejným turnajem, můžu změnit na vyrobít=0. Program si dělá některé pomocné

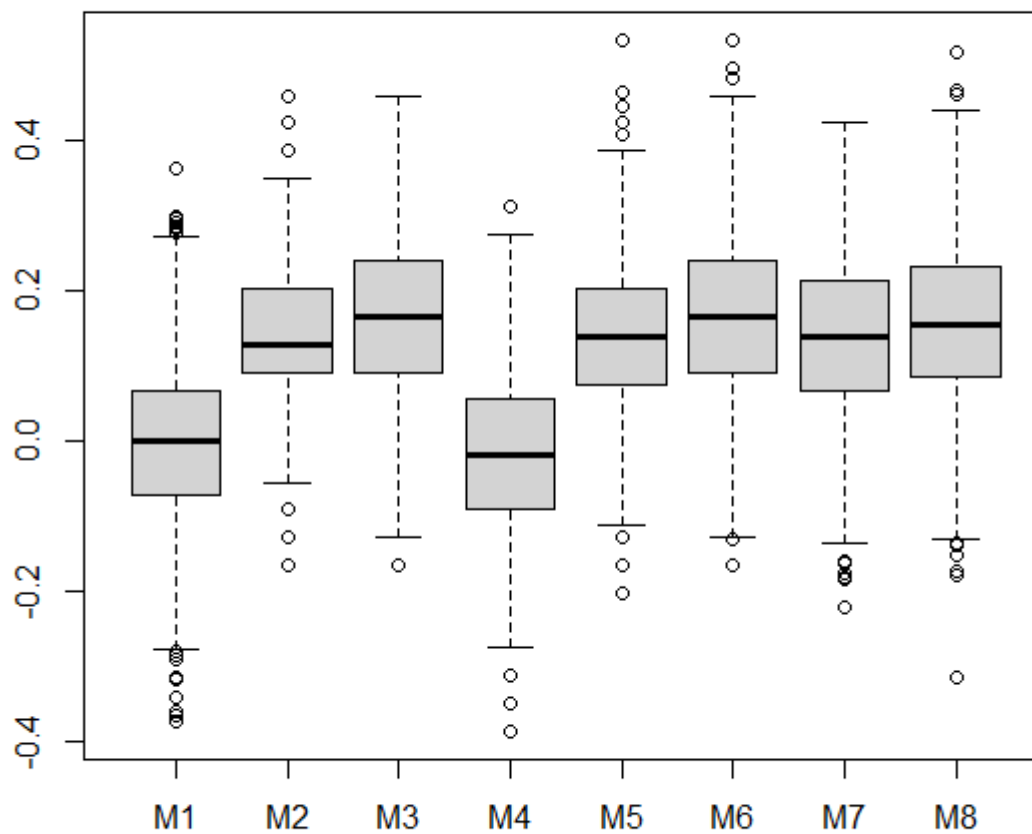
soubory, které ale už z předcházejících výpočtů můžu mít nahrané v pracovním adresáři. S vyrobít=0 bude výpočet o něco rychlejší.

- Argument `ph_1`. Určuje první pomocné hodnocení v případě rovnosti bodů. Funkce umí pracovat s `bucholz cut 1` (3), střední `bucholz` (7), `bucholz` (6), progresivním skóre (9) a průměrným elem soupeřů (8). `Bucholz cut 1` je nastavený defaultně.
- Argument `ph_2`. Určuje druhé pomocné hodnocení v případě rovnosti bodů. Funkce umí pracovat s `bucholz cut 1` (3), střední `bucholz` (7), `bucholz` (6), progresivním skóre (9) a průměrným elem soupeřů (8). Střední `bucholz` je nastavený defaultně.
- Argument `cisla`. Argument je defaultně prázdný vektor. Pokud chceme simulovat průběh turnaje za podmínky, že určitá partie skončí nějak, píšou například `cisla=c(5,2,81,50)`, což znamená, že si určíme výsledek hráčům se startovními čísly 5, 2, 81 a 50. Vždy musím zahrnout oba hráče z takové partie.
- Argument `předpoklady`. Argument je defaultně prázdný vektor, ale pokud vyplním argument `cisla`, musím vyplnit i argument `předpoklady`. Vyplním ho jako vektor, jehož prvky budou odpovídat výsledkům startovních čísel z vektoru `cisla`. Takže například `předpoklady=c(1,0,0.5,0.5)` znamená, že chci zjišťovat, jak to dopadne, když hráč se startovním číslem 5 porazí hráče se startovním číslem 2 a hráč se startovním číslem 81 remizuje s hráčem se startovním číslem 50.
- Argument `hrac`. Argument je defaultně nastavený na 0. Pokud ale chceme získat představu, jak dopadne turnaj z pohledu konkrétního hráče, vyplníme tento argument jeho startovním číslem.
- Argument `ceny`. Argument je defaultně vyplněný na 0. Pokud chceme analyzovat i finanční prémie, dáme `ceny=1`.
- Argument `prize`. Pokud jsme nechali `ceny=0`, argumentu si nemusím všimnout. Argument je defaultně nastavený na „pevne“, což využijeme u turnajů, kde se finanční ceny nedělí. Pokud chceme přepnout na systém dělení hrot, dáme `prize="hort"`.
- Argument `money`. Argument je defaultně nastavený jako prázdný vektor. Pokud chceme analyzovat finanční prémie, vyplníme argument `money` vektorem o takové délce, kolik finančních cen rozdělujeme. Prvním prvkem vektoru je výše ceny pro prvního atd. Tento argument můžeme použít i tehdy, když nám třeba velmi záleží na tom skončit na pódiu. Vyplníme si první tři ceny vysokými částkami a hned budeme vědět, jak moc je to reálné.

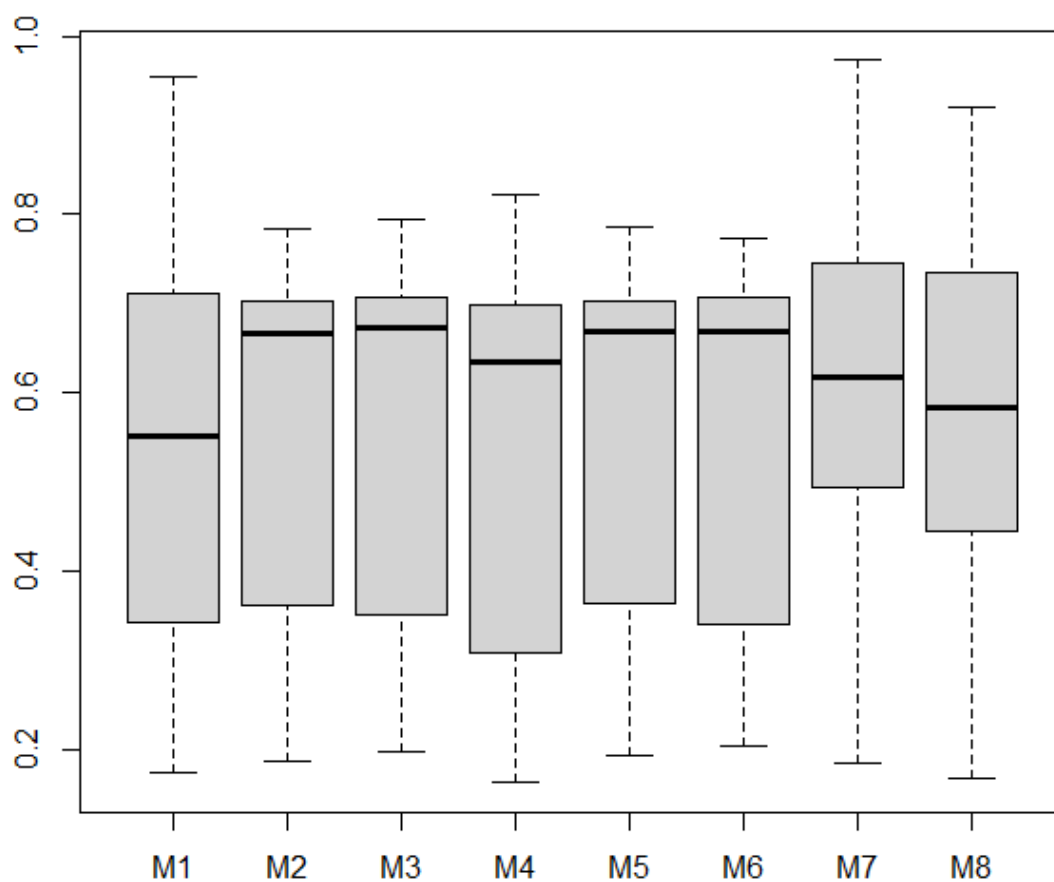
- Když mám příkaz ve formě například
`zadrazil(nazev=„turnaj“,Nsim=10,vyrobit=1,cisla=c(5,2),predpoklady=c(1,0),hrac=5,ceny=1,
prize=„hort“,money=c(300,200,100))`
vyplněný, můžu ho spustit.
- R průběžně píše po 10 % počet hotových simulací.
- Když je proces hotový, mám v pracovním adresáři uložené 2-5 souborů.
 - Určitě tam naleznu excel začínající na `sosnuto`, ten je nezajímavý.
 - Určitě tam naleznu excel začínající `vysledky_simulace`. V něm čtu tak, že na řádcích čtu pořadí hráčů v jednotlivých simulacích. Řádky odpovídají startovním číslům, respektive na řádce 2 jsou výsledky hráče se startovním číslem 1, na řádce 3 jsou výsledky hráče se startovním číslem 2 atd.
 - Pokud jsem vyplnil argument `hrac`, najdu ve složce i graf s histogramem jeho možných umístění. V nadpise grafu je uloženo průměrné očekávané umístění.
 - Pokud jsem vyplnil argument `ceny`, najdu ve složce i excel začínající `finance`. V něm čtu tak, že na řádcích čtu finanční zisk hráčů v jednotlivých simulacích. Řádky odpovídají startovním číslům, respektive na řádce 2 jsou výsledky hráče se startovním číslem 1, na řádce 3 jsou výsledky hráče se startovním číslem 2 atd.
 - Pokud jsem vyplnil argument `ceny` i `hrac`, najdu ve složce i graf s histogramem jeho možných finančních zisků. V nadpise grafu je uložený průměrný očekávaný zisk.
- Pokud budu dělat výpočtů víc, je výhodné si výsledné excely a grafy přesunout do jiných složek mimo pracovní adresář, protože by se mi mohly přemazávat.
- Funkce pracuje výhradně s metodou předpovědi číslo 18.
- Funkce z hlediska veřejného sdílení ve zkušební verzi. Za případné nedostatky se omlouvám.

B. Další grafy srovnávající metody 1–8 na turnaji Czech Open A

Graf 14: korelace odhadnutých a skutečných výsledků hráčů v turnaji Czech Open A 2021 podle metod 1–8



Graf 15: korelace odhadu umístění elitních hráčů v turnaji Czech Open A 2021 podle metod 1–8



C. Srovnání metod 1 až 8 na turnajích Czech Open A, B, C, D a Open Praha

Tabulka 29: Analýza rozptylu dvojného třídění s interakcemi metod 1 až 8 podle korelace odhadnutých a skutečných umístění

zdroj	df	SS	MS	F	p-hodnota
měnlivosti					
typ turnaje	4	0,0101	0,0025	84,08	0,0000
metoda	7	0,0077	0,0011	36,50	0,0000
interakce	28	0,0307	0,0000	0,907	0,6040
reziduální	160	0,0048	0,0000		

Tabulka 30: Srovnání metod 1–8 podle korelací odhadnutých a skutečných umístění

M7	M8	M1	M6	M3	M5	M2	M4
0,9320	0,9312						
	0,9312	0,9272					
			0,9203	0,9201	0,9193	0,9192	
							0,9129

Tabulka 31: Analýza rozptylu dvojného třídění s interakcemi metod 1 až 8 podle korelace odhadnutých a skutečných výsledků

zdroj	df	SS	MS	F	p-hodnota
měnlivosti					
typ turnaje	4	0,0924	0,0231	25,924	0,0000
metoda	7	0,5039	0,0719	80,759	0,0000
interakce	28	0,0307	0,0011	0,211	0,6040
reziduální	160	0,1426	0,0008		

Tabulka 32: Srovnání metod 1–8 podle korelací odhadnutých a skutečných výsledků

M6	M3	M2	M5	M8	M7	M4	M1
0,1377	0,1365	0,1267	0,1259	0,1168			
		0,1267	0,1259	0,1168	0,1096		
						0,0211	0,0026

Tabulka 33: Analýza rozptylu dvojného třídění s interakcemi metod 1 až 8 podle korelace odhadnutých a skutečných umístění elitních hráčů

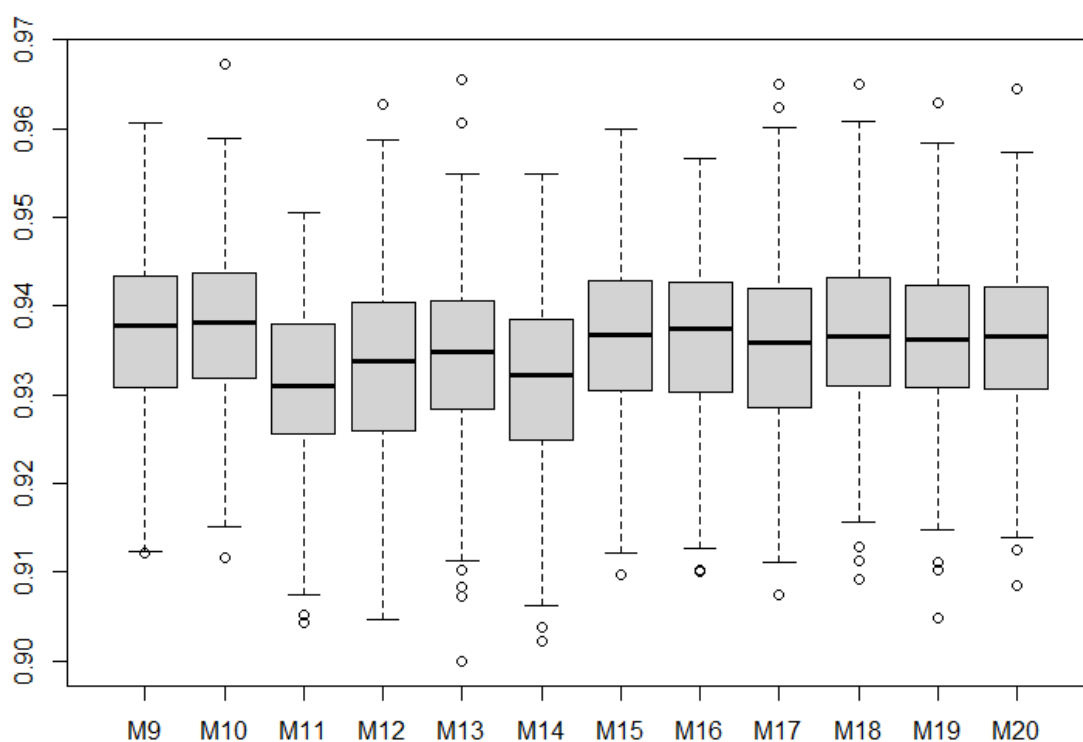
zdroj	df	SS	MS	F	p-hodnota
měnlivosti					
typ turnaje	4	3,841	0,9062	29,176	0,0000
metoda	7	0,031	0,0044	0,134	0,996
interakce	28	0,234	0,0084	0,254	1,000
reziduální	160	5,266	0,0329		

Tabulka 34: Srovnání metod 1–8 podle korelací odhadnutých a skutečných umístění elitních hráčů

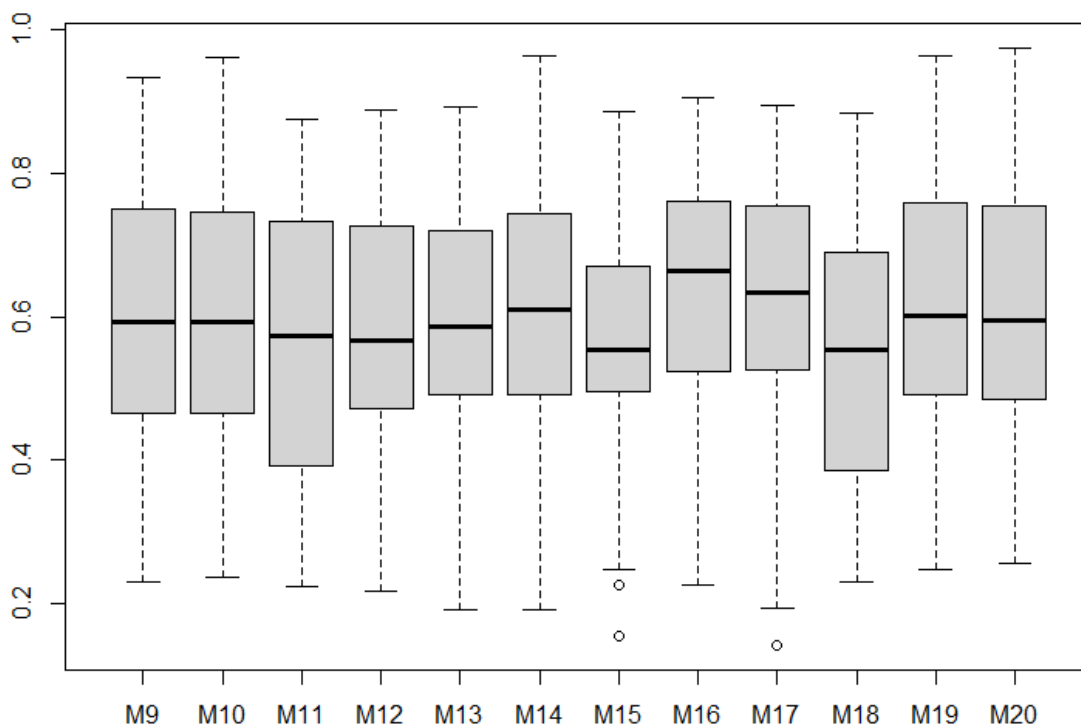
M8	M7	M4	M1	M6	M3	M2	M5
0,4907	0,4896	0,4826	0,4767	0,4691	0,4680	0,4640	0,4524

D. Další grafy srovnávající metody 9–20 na turnaji Czech Open A

Graf 16: korelace odhadnutých a skutečných umístění hráčů v turnaji Czech Open A 2021 podle metod 9–20



Graf 17: korelace odhadnutých a skutečných umístění elitních hráčů v turnaji Czech Open A 2021 podle metod 9–20



E. Srovnání metod 9 až 20 na turnajích Czech Open A, B, C, D a Open Praha

Tabulka 35: Analýza rozptylu dvojného třídění s interakcemi metod 9 až 20 podle korelace odhadnutých a skutečných výsledků

zdroj	df	SS	MS	F	p-hodnota
měnlivosti					
typ turnaje	4	0,6882	0,1720	129,4	0,0000
metoda	11	0,0074	0,0006	0,508	0,897
interakce	44	0,0380	0,0008	0,650	0,956
reziduální	240	0,3191	0,0013		

Tabulka 36: Analýza rozptylu dvojného třídění s interakcemi metod 9 až 20 podle korelace odhadnutých a skutečných umístění

zdroj	df	SS	MS	F	p-hodnota
měnlivosti					
typ turnaje	4	0,0175	0,004375	155,214	0,0000
metoda	11	0,000061	0,000006	0,196	0,998
interakce	44	0,001017	0,000023	0,820	0,782
reziduální	240	0,006765	0,000028		

Tabulka 37: Analýza rozptylu dvojného třídění s interakcemi metod 9 až 20 podle korelace odhadnutých a skutečných umístění elitních hráčů

zdroj	df	SS	MS	F	p-hodnota
měnlivosti					
typ turnaje	4	6,152	1,5381	57,508	0,0000
metoda	11	0,106	0,0097	0,361	0,97
interakce	44	0,160	0,0036	0,136	1,00
reziduální	240	6,419	0,0267		

F. Srovnání všech 20 metod na turnajích Czech Open A, B, C, D a Open Praha v letech 2015–2019

Tabulka 38: Analýza rozptylu dvojného třídění s interakcemi všech metod podle korelace odhadnutých a skutečných výsledků

zdroj	df	SS	MS	F	p-hodnota
měnlivosti					
typ turnaje	4	0,6711	0,1677	145,3	0,0000
metoda	19	0,5544	0,0291	25,28	0,0000
interakce	76	0,1783	0,0023	2,033	0,0000
reziduální	400	0,4617	0,0011		


```

else if (nchar(start_praha22[j,i])==4)
{ bili_praha22[j,i]<-as.numeric(substring(start_praha22[j,i],1,2))}
else if (nchar(start_praha22[j,i])==3)
{ bili_praha22[j,i]<-as.numeric(substring(start_praha22[j,i],1,1))}
else { bili_praha22[j,i]<-0}
}
else { bili_praha22[j,i]<-"nez"}
}
}
vysledky_praha22<-start_praha22
#následující cyklus vypíše výsledky, ale zase jen do buněk z pohledu bílý
for (j in 1:nrow(start_praha22))
{
for (i in 7:14)
{
if (grepl("w",start_praha22[j,i])==TRUE)
{
if(nchar(start_praha22[j,i])==5)
{ vysledky_praha22[j,i]<-as.numeric(substring(start_praha22[j,i],5,5))}
else if (nchar(start_praha22[j,i])==4)
{ vysledky_praha22[j,i]<-as.numeric(substring(start_praha22[j,i],4,4))}
else if (nchar(start_praha22[j,i])==3)
{ vysledky_praha22[j,i]<-as.numeric(substring(start_praha22[j,i],3,3))}
else { vysledky_praha22[j,i]<-0}
}
else { vysledky_praha22[j,i]<-"nez"}
}
}
}
#následující cyklus opraví ty 5, co tam mám místo remíz na 0,5
#jinak tam zůstane, co tam je
for (j in 1:nrow(start_praha22))
{
for (i in 7:14)
{
if(is.na(vysledky_praha22[j,i])==TRUE) { vysledky_praha22[j,i]<-"nez"}
else { vysledky_praha22[j,i]<-vysledky_praha22[j,i]}
}
}
#mám v tom foru dva cykly, ten první opravuje NA (v praxi kontumace) na "nez"
#ten druhý dělá ty remízy
if(vysledky_praha22[j,i]==5) { vysledky_praha22[j,i]<-0.5}
else { vysledky_praha22[j,i]<-vysledky_praha22[j,i]}
}
}
}
#partie_praha22 je první pokus o vytvoření matice všech partií
#zatím má, ale tolik řádků, kolik je výsledků z pohledu obou hráčů
#je to dělané na prvních 8 kol
partie_praha22<-matrix(nrow=8*nrow(start_praha22),ncol=7)
#do toho jenhry_praha22 jsem nasypal jen výsledky, žádná jména, ela nic
jenhry_praha22<-vysledky_praha22[,c(7:14)]
#do jednoho sloupečku sesypám všechny výsledky, co se odehrály, co jsou
#v matici jenhry_praha22
partie_praha22[,5]<-c(t(jenhry_praha22))
jensouperi_praha22<-bili_praha22[,c(7:14)]
#do matice jensouperi_praha22 zase cpu čísla všech možných soupeřů
#i to potom přendávám v podobě vektoru do matice partie_praha22
partie_praha22[,3]<-c(t(jensouperi_praha22))
for (i in 1:nrow(partie_praha22))
{ partie_praha22[i,4]<-start_praha22[c(t(jensouperi_praha22))[i],5]}
#dál potřebuji do 1. sloupečku matice partie_praha22 startovní čísla hráčů
#z jejichž pohledu se to hraje (bílí i černí v jedné kupě)
#na to používám 1+floor(i/8), protože každý soupeř přidává do hry 8 partií
for (i in 1:nrow(partie_praha22))
{ as.numeric(partie_praha22[i,1]<-1+floor((i-1)/8))}
for (i in 1:nrow(partie_praha22))
{ partie_praha22[i,2]<-start_praha22[1+floor((i-1)/8),5]}
partie_praha22[is.na(partie_praha22)]= "nez"
#následuje pomocný vektor, který zjišťuje, kde všude mi do
pomocny_1<-c(grepl("nez",partie_praha22[,5]))
pomocny_2<-c()
#následuje zjištění, v kterých číslech to tak není, tj. které řádky mě
for (i in 1:nrow(partie_praha22)) {
if (pomocny_1[i]==FALSE) {pomocny_2[i]<-i}
else {pomocny_2[i]<-0}
}
pomocny_3<-pomocny_2[pomocny_2>0]
#no a do matice cistepartie_praha22 už mi padají jenom řádky, co mě
#tj. brané z pohledu bílých
cistepartie_praha22<-partie_praha22[pomocny_3,]

```

```

#tady potřebuji vyřešit problém, že čísla jsou kdoví proč v nečíselném
#proto je hodím do vektoru, předělám na čísla a vrátím do matice
#ale pod novým názvem osm_praha22
osm_praha23<-matrix(as.numeric(cistepartie_praha22),
                    nrow=nrow(cistepartie_praha22),ncol=6)
for (i in 1:nrow(osm_praha23))
{ osm_praha23[i,6]<-osm_praha23[i,2]-osm_praha23[i,4]}
osm_praha22<-cbind(osm_praha23,round((osm_praha23[,2]+osm_praha23[,4])/2))
cil<-paste(misto,"sosnuto_",nazev, ".xlsx",sep="")
jakodataframe<-as.data.frame(osm_praha22)
write_xlsx(jakodataframe,path=cil)
}

#ted' jdu pripravovat studnicka, v ni jsou vsechny vysledky 1.-8. kola tur
#který je specifikován v argumentu nazev
{kod_3<-paste(misto,"sosnuto_",nazev, ".xlsx",sep="")
studnicka<-as.data.frame(read_xlsx(kod_3))

mega_informace<-matrix(nrow=6*length(postup),ncol=Nsim+2)
for(q in 1:length(postup)){
metoda<-postup[q]

#ted' přichází největší blok, tj. bývalá funkce info_verze_4
#to je ta, co mi dělá všechno, vypořádává se s výsledkovými listinami
#včetně kontumací, neodehraných her a podobně,
#počítá bucholz, simuluje výsledky
#přijde tolikrát, jaká je délka vektoru postup
{
#PRVNÍ CHCI VYROBIT TABULKU ČÍSEL SOUPEŘŮ, DÁM TO PAK I DO MATICE
#VÝSTUPEM JE SOUPEROVE
{
kod<-paste(misto,nazev, ".xlsx",sep="")
start_praha22<-as.data.frame(read_xlsx(kod))
souperove<-start_praha22
for (i in 1:nrow(start_praha22)){
for (j in 7:15){
if ((start_praha22[i,j]=="-1")==TRUE) {souperove[i,j]<-1000}
if ((nchar(start_praha22[i,j])==1)==TRUE) {souperove[i,j]<-2000}
}
}
}

for (i in 1:nrow(start_praha22)){
for (j in 7:15){
if(nchar(start_praha22[i,j])==5)
{souperove[i,j]<-(substring(start_praha22[i,j],1,3))}
else if (nchar(start_praha22[i,j])==4)
{souperove[i,j]<-(substring(start_praha22[i,j],1,2))}
else if (nchar(start_praha22[i,j])==3)
{souperove[i,j]<-(substring(start_praha22[i,j],1,1))}
}
}
#a v tuhle chvíli mám matici souperove, kde 7:15 jsou startovní čísla
#soupeřů a mám tam 1000 pro volno, 2000 pro nenasazen, kontumace
#nepostřižena
}

#DRUHÝ CHCI VYROBIT TABULKU VÝSLEDKŮ, KDE NENASAZEN JE 0, VOLNO JE 1 A
#KONTUMACE ODPOVÍDÁ VÝSLEDKU V TABULCE
#VÝSTUPEM JE VYSLEDKY_CISLA
{
vysledkove<-start_praha22
for (i in 1:nrow(start_praha22)){
for (j in 7:15){
if(nchar(start_praha22[i,j])==5)
{vysledkove[i,j]<-(substring(start_praha22[i,j],5,5))}
else if (nchar(start_praha22[i,j])==4)
{vysledkove[i,j]<-(substring(start_praha22[i,j],4,4))}
else if (nchar(start_praha22[i,j])==3)

```

```

    {vysledkove[i,j]<-(substring(start_praha22[i,j],3,3))}
  }
}

for (i in 1:nrow(start_praha22)){
  for (j in 7:15){
    if ((vysledkove[i,j]==5)==TRUE) {vysledkove[i,j]<-0.5}
    if ((vysledkove[i,j]=="+")==TRUE) {vysledkove[i,j]<-1}
    if ((vysledkove[i,j]=="-")==TRUE) {vysledkove[i,j]<-0}
    if ((vysledkove[i,j]=="-1")==TRUE) {vysledkove[i,j]<-1}
  }
}
vysledky_cisla<-vysledkove
#a v tuhle chvíli mám matici soupeře, kde 7:15 jsou výsledky ve
#tvaru 0/0,5/1
}

#TŘETÍ CHCI MATICI, KDE VÍM O VŠECH KONTUMOVANÝCH VÝSLEDČÍCH, VOLNECH
#A NENASAZEN
#VÝSTUPEM JE KONTUMACE
{
  kontumace<-start_praha22
  for (i in 1:nrow(start_praha22)){
    for (j in 7:15){
      if(nchar(start_praha22[i,j])==5)
        {kontumace[i,j]<-(substring(start_praha22[i,j],5,5))}
      else if (nchar(start_praha22[i,j])==4)
        {kontumace[i,j]<-(substring(start_praha22[i,j],4,4))}
      else if (nchar(start_praha22[i,j])==3)
        {kontumace[i,j]<-(substring(start_praha22[i,j],3,3))}
    }
  }
}

for (i in 1:nrow(start_praha22)){
  for (j in 7:15){
    if ((kontumace[i,j]==5)==TRUE)
      {kontumace[i,j]<-0.5}
    if ((kontumace[i,j]=="-1")==TRUE)
      {kontumace[i,j]<-"volno"}
    if ((nchar(start_praha22[i,j])==1)==TRUE)
      {kontumace[i,j]<-"nenasazen"}
  }
}

#a v tuhle chvíli mám matici, která 7:15 počítá s partii, co se
#nehraly
#ve tvaru "volno", "nenasazen", "+" nebo "-"

vektor_poctu_zdravych_her<-c()
matice_poctu_bodu<-matrix(nrow=nrow(start_praha22),ncol=9)
matice_el_souperu<-matrix(nrow=nrow(start_praha22),ncol=9)
vektor_perf<-c()
vektor_prumernych_el<-c()
for (i in 1:nrow(start_praha22)){
  vektor_poctu_zdravych_her[i]<-sum(kontumace[i,7:14]==1 |
    kontumace[i,7:14]==0.5 |
    kontumace[i,7:14]==0)}
for (i in 1:nrow(start_praha22)){
  for (j in 1:8){
    if(kontumace[i,6+j]==1 | kontumace[i,6+j]==0.5 | kontumace[i,6+j]==0)
      {matice_poctu_bodu[i,j]<-as.numeric(kontumace[i,6+j])}
    else {matice_poctu_bodu[i,j]<-0}
  }
}
for(i in 1:nrow(start_praha22)){
  matice_poctu_bodu[i,9]<-matice_poctu_bodu[i,1]+matice_poctu_bodu[i,2]+
  matice_poctu_bodu[i,3]+matice_poctu_bodu[i,4]+matice_poctu_bodu[i,5]+
  matice_poctu_bodu[i,6]+matice_poctu_bodu[i,7]+matice_poctu_bodu[i,8]
}

for (i in 1:nrow(start_praha22)){
  for (j in 1:8){
    if(kontumace[i,6+j]==1 | kontumace[i,6+j]==0.5 | kontumace[i,6+j]==0)
      {matice_el_souperu[i,j]<-souperove[as.numeric(souperove[i,6+j]),5]}
    else {matice_el_souperu[i,j]<-0}
  }
}

```

```

    }
  }

  for(i in 1:nrow(start_praha22)){
    matice_el_souperu[i,9]<-matice_el_souperu[i,1]+matice_el_souperu[i,2]+
    matice_el_souperu[i,3]+matice_el_souperu[i,4]+matice_el_souperu[i,5]+
    matice_el_souperu[i,6]+matice_el_souperu[i,7]+matice_el_souperu[i,8]
  }

  for(i in 1:nrow(start_praha22)){
    vektor_prumernych_el[i]<-matice_el_souperu[i,9]/vektor_poctu_zdravych_her[i]
    vektor_perf[i]<-matice_el_souperu[i,9]/vektor_poctu_zdravych_her[i]+
    qnorm(matice_poctu_bodu[i,9]/vektor_poctu_zdravych_her[i],0,282.84)
    if (is.infinite(vektor_perf[i])==TRUE|is.na((vektor_perf[i])==TRUE))
    { vektor_perf[i]<-as.numeric(souperove[i,5]) }
  }

#ČTVRTÁ CHCI MATICI (TEĎ OPRAVDU MATICI, NE KUS TABULKY), KDE BUDU MÍT
#NAPÁROVANÉ DVOJICE, CO SE UTKAJÍ V POSLEDNÍM KOLE
#NEŠLO MI TO ROZRHNOU, PROTO I ROVNOU NEHRALI_POSLEDNI
#VÝSTUPEM JE listina a nehrali_posledni a nehrali_posledni_navic
{
  bilivefinale1<-start_praha22
  cernivefinale1<-start_praha22
  bilivefinale_hrali<-bilivefinale1[grepl("w",bilivefinale1[,15])==TRUE,]
  #to jsou klasicky bílí z partií, které byly nalošovány (tj. včetně
  #kontumací)
  cernivefinale_hrali<-cernivefinale1[grepl("b",cernivefinale1[,15])==TRUE,]
  #to jsou klasicky černí z partií, které byly nalošovány (tj. včetně
  #kontumací)
  vsichnivefinale_nehrali_a<-bilivefinale1[nchar(bilivefinale1[,15])==1,]
  vsichnivefinale_nehrali_b<-bilivefinale1[nchar(bilivefinale1[,15])==2,]
  vsichnivefinale_nehrali<-
  rbind(vsichnivefinale_nehrali_a,vsichnivefinale_nehrali_b)
  nehrali<-matrix(nrow=nrow(vsichnivefinale_nehrali),ncol=3)
  for (i in 1:nrow(vsichnivefinale_nehrali)) {
    nehrali[i,1]<-vsichnivefinale_nehrali[i,1]
    nehrali[i,2]<-vsichnivefinale_nehrali[i,5]
    nehrali[i,3]<-vsichnivefinale_nehrali[i,15] }
  #to jsou hráči, kteří v posledním kole nehrali (tj. buď 1, pokud
  #neměli soupeře
  #nebo 0 pokud se nenechali nalošovat)
  los_hrali_bili<-bilivefinale_hrali
  for (i in 1:nrow(bilivefinale_hrali)){
    if(nchar(bilivefinale_hrali[i,15])==5)
  { los_hrali_bili[i,15]<-as.numeric(substring(los_hrali_bili[i,15],1,3))}
    else if (nchar(bilivefinale_hrali[i,15])==4)
  { los_hrali_bili[i,15]<-as.numeric(substring(los_hrali_bili[i,15],1,2))}
    else if (nchar(bilivefinale_hrali[i,15])==3)
  { los_hrali_bili[i,15]<-as.numeric(substring(los_hrali_bili[i,15],1,1))}
    else {print("něco se pokazilo")}
  }

  listina<-matrix(nrow=nrow(los_hrali_bili),ncol=8)
  listina[,1]<-los_hrali_bili[,1]
  listina[,2]<-los_hrali_bili[,5]
  listina[,3]<-los_hrali_bili[,15]
  for (i in 1:nrow(los_hrali_bili))
  { listina[i,4]<-start_praha22[listina[i,3],5]}
  listina<-matrix(c(as.numeric(listina)),nrow(los_hrali_bili),ncol=8)
  listina[,5]<-listina[,2]-listina[,4]
  listina[,6]<-round((listina[,2]+listina[,4])/2)

  nehrali_posledni<-matrix(c(as.numeric(nehrali)),nrow(nehrali),ncol=3)
  nehrali_posledni_navic<-rep("tvorim",nrow(nehrali_posledni))
  for (i in 1:nrow(nehrali_posledni)){
    if ((nehrali_posledni[i,3]==0)==TRUE)
    { nehrali_posledni_navic[i]<- "nenasazen" }
    if ((nehrali_posledni[i,3]==1)==TRUE)
    { nehrali_posledni_navic[i]<- "volno" }
  }
}

```

#PÁTÁ CHCI ZKUSIT NASIMULOVAT VÝSLEDKY POSLEDNÍHO KOLA
 #to jsou ty moje metody simulací!

```
{
  tipy<-matrix(nrow=nrow(listina),ncol=Nsim)
  if((metoda==1)==TRUE){
    print("Dělám metodu 1.")
    for (i in 1:Nsim){
      for (j in 1:nrow(listina)){
        tipy[j,i]<-sample(x=c(1,0.5,0),size=1,prob=c(1/3,1/3,1/3))
      }
    }
  }
  #všechno stojí na 1/3, výhra, prohra i remíza

  if((metoda==2)==TRUE){
    print("Dělám metodu 2.")
    for (i in 1:Nsim){
      for (j in 1:nrow(listina)){
        tipy[j,i]<-sample(x=c(1,0),1,prob=c
          (pnorm((listina[j,2]-listina[j,4])/282.84),
            pnorm((listina[j,4]-listina[j,2])/282.84)))
      }
    }
  }
  #jedničky a nuly odpovídající očekávaným výsledkům

  if((metoda==3)==TRUE){
    print("Dělám metodu 3.")
    for (i in 1:Nsim){
      for (j in 1:nrow(listina)){
        tipy[j,i]<-sample(x=c(1,0),1,prob=c
          (pnorm((listina[j,2]-listina[j,4]+
            qnorm(mean(studnicka[,5]),0,282.84))/282.84),
            pnorm((listina[j,4]-listina[j,2]-
            qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
      }
    }
  }
  #jedničky a nuly odpovídající očekávaným výsledkům s tím
  #že bílý má fixní výhodu, která vychází z předcházejících výsledků

  if((metoda==4)==TRUE){
    print("Dělám metodu 4.")
    for (i in 1:Nsim){
      for (j in 1:nrow(listina)){
        tipy[j,i]<-sample(x=c(1,0),1,prob=c
          (pnorm((vektor_perf[listina[j,1]]-vektor_perf[listina[j,3]])/282.84),
            1-pnorm((vektor_perf[listina[j,1]]-vektor_perf[listina[j,3]])/282.84)))
      }
    }
  }
  #jedničky a nuly odpovídající očekávaným výsledkům, ale ela tvoří per

  if((metoda==5)==TRUE){
    print("Dělám metodu 5.")
    remiz_6<-rep(sum(studnicka[,5]==0.5)/
      length(studnicka[,5]),nrow(listina))
    for (i in 1:Nsim){
      for (j in 1:nrow(listina)){
        if (((pnorm((listina[j,2]-listina[j,4])/282.84))<remiz_6[j]/2)
          ((pnorm(listina[j,4]-listina[j,2])/282.84))<remiz_6[j]/2)
          ==TRUE)
          {
            remiz_6[j]<-2*min(pnorm((listina[j,2]-listina[j,4])/282.84),
              (pnorm(listina[j,4]-listina[j,2])/282.84))
          }

        tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
          (pnorm((listina[j,2]-listina[j,4])/282.84)-remiz_6[j]/2,
            remiz_6[j],
            pnorm((listina[j,4]-listina[j,2])/282.84)-remiz_6[j]/2))
      }
    }
  }
}
```

```

}

#vezmu remízovost z turnaje a tím podílem poupravím očekávaný výsledek
#pokud je očekávaná remízovost větší než dvojnásobek pravděpodobnosti
#výhry nebo prohry, pak ji snížím právě na tolik
#tady vycházím z remízovosti jen v rámci daného turnaje

if((metoda==6)==TRUE){
  print("Dělám metodu 6.")
  remiz_6<-rep(sum(studnicka[,5]==0.5)/
    length(studnicka[,5]),nrow(listina))
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      if(((pnorm((listina[j,2]+
qnorm(mean(studnicka[,5]),0,282.84)-listina[j,4])/282.84))<remiz_6[j]/2)
      ((pnorm(listina[j,4]-listina[j,2]-
qnorm(mean(studnicka[,5]),0,282.84))/282.84))<remiz_6[j]/2)
      ==TRUE)
      {
        remiz_6[j]<-2*min(pnorm((listina[j,2]+
qnorm(mean(studnicka[,5]),0,282.84)-listina[j,4])/282.84),
          (pnorm(listina[j,4]-listina[j,2]-
qnorm(mean(studnicka[,5]),0,282.84))/282.84))
      }

      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        (pnorm((listina[j,2]+
qnorm(mean(studnicka[,5]),0,282.84)-listina[j,4])/282.84)-remiz_6[j]/2,
          remiz_6[j],
          pnorm((listina[j,4]-listina[j,2]-
qnorm(mean(studnicka[,5]),0,282.84))/282.84)-remiz_6[j]/2))
    }
  }
}

#tohle je v podstatě vylepšená 5. metoda, volím fixní podíl remíz
#z průběhu daného turnaje, dále vylepším očekávaný zisk o fixní
#elovou výhodu a potom to dám dohromady

if((metoda==7)==TRUE){
  print("Dělám metodu 7.")
  remiz_6<-rep(sum(studnicka[,5]==0.5)/
    (2*sum(studnicka[,5]==1)+sum(studnicka[,5]==0.5)),nrow(listina))
  remiz_7<-rep(sum(studnicka[,5]==0.5)/
    (2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5)),nrow(listina))
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        ((1-remiz_6[j])*pnorm((listina[j,2]-listina[j,4])/282.84),
          1-(1-remiz_6[j])*pnorm((listina[j,2]-listina[j,4])/282.84)-
          (1-remiz_7[j])*pnorm((listina[j,4]-listina[j,2])/282.84),
          (1-remiz_7[j])*pnorm((listina[j,4]-listina[j,2])/282.84)))
    }
  }
}

#vycházím z turnaje samotného
#tahle metoda zohledňuje jaký podíl tvoří remízy na zisku hráčů
#v podstatě to je vylepšená 5. metoda

if((metoda==8)==TRUE){
  print("Dělám metodu 8.")
  remiz_6<-rep(sum(studnicka[,5]==0.5)/
    (2*sum(studnicka[,5]==1)+sum(studnicka[,5]==0.5)),
    nrow(listina))
  remiz_7<-rep(sum(studnicka[,5]==0.5)/
    (2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5)),
    nrow(listina))
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        ((1-remiz_6[j])*pnorm((listina[j,2]+
qnorm(mean(studnicka[,5]),0,282.84)
          -listina[j,4])/282.84),
          1-(1-remiz_6[j])*pnorm((listina[j,2]+

```



```

        qnorm(mean(studnicka[,5]),0,282.84)
        -listina[j,4])/282.84)-
    (1-remiz_7[j])*pnorm((listina[j,4]+
    qnorm(mean(studnicka[,5]),0,282.84)
    -listina[j,2])/282.84),
    (1-remiz_7[j])*pnorm((listina[j,4]-
    listina[j,2]-qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
    }
}
}

if((metoda==9)==TRUE){
  print("Dělám metodu 9.")
  pul_rozpeti<-50
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocny_150<-c()
      pomocny_150<-
        which(studnicka[,6]>listina[j,5]-pul_rozpeti & studnicka[,6]<
          listina[j,5]+pul_rozpeti)
      vyber<-studnicka[pomocny_150,5]
      if(length(vyber)<4){
        remiz_6<-sum(studnicka[,5]==0.5)/
          (2*sum(studnicka[,5]==1)+sum(studnicka[,5]==0.5))
        remiz_7<-sum(studnicka[,5]==0.5)/
          (2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5))

        tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
          ((1-remiz_6)*pnorm((listina[j,2]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,4])/282.84),
          1-(1-remiz_6)*pnorm((listina[j,2]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,4])/282.84)-
          (1-remiz_7)*pnorm((listina[j,4]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,2])/282.84),
          (1-remiz_7)*pnorm((listina[j,4]-
            listina[j,2]-qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
        }
      } else
      {
        prst_vyhry<-length(vyber[vyber==1])/length(vyber)
        prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
        prst_prohry<-length(vyber[vyber==0])/length(vyber)
      }
      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c(prst_vyhry,prst_remizy,prst_prohry))
    }
  }
}

#vycházím z turnaje
#odhaduji výsledky na základě četnosti v intervalu
#vycházím z turnaje samotného

if((metoda==10)==TRUE){
  print("Dělám metodu 10.")
  pul_rozpeti<-50
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocny_150<-c()
      pomocny_150<-
        which(stredni_studna[,6]>listina[j,5]-pul_rozpeti & stredni_studna[,6]<
          listina[j,5]+pul_rozpeti)
      vyber<-stredni_studna[pomocny_150,5]
      prst_vyhry<-length(vyber[vyber==1])/length(vyber)
      prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
      prst_prohry<-length(vyber[vyber==0])/length(vyber)

      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c(prst_vyhry,prst_remizy,prst_prohry))
    }
  }
}

```

```

    }
  }
  #vycházím z typu turnaje
  #odhaduji výsledky na základě četnosti v intervalu

  if((metoda==11)==TRUE){
    print("Dělám metodu 11.")
    pul_rozpeti<-50
    for (i in 1:Nsim){
      for (j in 1:nrow(listina)){
        pomocny_150<-c()
        pomocny_150<-
          which(studna[,6]>listina[j,5]-pul_rozpeti & studna[,6]<
            listina[j,5]+pul_rozpeti)
        vyber<-studna[pomocny_150,5]
        prst_vyhry<-length(vyber[vyber==1])/length(vyber)
        prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
        prst_prohry<-length(vyber[vyber==0])/length(vyber)

        tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c(prst_vyhry,prst_remizy,prst_prohry))
      }
    }
  }

  if((metoda==12)==TRUE){
    print("Dělám metodu 12.")
    pul_rozpeti<-50
    for (i in 1:Nsim){
      for (j in 1:nrow(listina)){
        pomocna_matrice<-matrix(nrow=length
          (which(studnicka[,6]>listina[j,5]-pul_rozpeti &
            studnicka[,6]<listina[j,5]+pul_rozpeti)),ncol=5)
        pomocna_matrice[,1]<-
          which(studnicka[,6]>listina[j,5]-pul_rozpeti &
            studnicka[,6]<listina[j,5]+pul_rozpeti)
        pomocna_matrice[,2]<-studnicka[pomocna_matrice[,1],5]
        pomocna_matrice[,3]<-pomocna_matrice[,1]-listina[j,5]
        pomocna_matrice[,4]<-abs(0.5-pnorm(pomocna_matrice[,3],listina[j,5],25))
        pomocna_matrice[,5]<-pomocna_matrice[,2]*pomocna_matrice[,4]
        prst_vyhry<-sum(pomocna_matrice[which(pomocna_matrice[,2]==1),4])/
          sum(pomocna_matrice[,4])
        prst_remizy<-sum(pomocna_matrice[which(pomocna_matrice[,2]==0.5),4])/
          sum(pomocna_matrice[,4])
        prst_prohry<-prst_remizy<-sum(pomocna_matrice[which(pomocna_matrice[,2]==0),4])/
          sum(pomocna_matrice[,4])
        if(length(pomocna_matrice[,1])<4)
        {
          remiz_6<-sum(studnicka[,5]==0.5)/
            (2*sum(studnicka[,5]==1)+sum(studnicka[,5]==0.5))
          remiz_7<-sum(studnicka[,5]==0.5)/
            (2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5))

          tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
            ((1-remiz_6)*pnorm((listina[j,2]+
              qnorm(mean(studnicka[,5]),0,282.84)
              -listina[j,4])/282.84),
              1-(1-remiz_6)*pnorm((listina[j,2]+
              qnorm(mean(studnicka[,5]),0,282.84)
              -listina[j,4])/282.84)-
              (1-remiz_7)*pnorm((listina[j,4]+
              qnorm(mean(studnicka[,5]),0,282.84)
              -listina[j,2])/282.84),
              (1-remiz_7)*pnorm((listina[j,4]-
              listina[j,2]-qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
        }
        else
        {
          tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c(prst_vyhry,prst_remizy,prst_prohry))
        }
      }
    }
  }
}

```

```

if((metoda==13)==TRUE){
  print("Dělám metodu 13.")
  pul_rozpeti<-50
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocna_matice<-matrix(nrow=length
        (which(stredni_studna[,6]>listina[j,5]-pul_rozpeti &
          stredni_studna[,6]<listina[j,5]+pul_rozpeti)),ncol=5)
      pomocna_matice[,1]<-
        which(stredni_studna[,6]>listina[j,5]-pul_rozpeti &
          stredni_studna[,6]<listina[j,5]+pul_rozpeti)
      pomocna_matice[,2]<-stredni_studna[pomocna_matice[,1],5]
      pomocna_matice[,3]<-pomocna_matice[,1]-listina[j,5]
      pomocna_matice[,4]<-abs(0.5-pnorm(pomocna_matice[,3],listina[j,5],25))
      pomocna_matice[,5]<-pomocna_matice[,2]*pomocna_matice[,4]
      prst_vyhry<-sum(pomocna_matice[which(pomocna_matice[,2]==1),4])/
        sum(pomocna_matice[,4])
      prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0.5),4])/
        sum(pomocna_matice[,4])
      prst_prohry<-prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0),4])/
        sum(pomocna_matice[,4])
      if(length(pomocna_matice[,1])<4)
      {
        remiz_6<-sum(studnicka[,5]==0.5)/
          (2*sum(studnicka[,5]==1)+sum(studnicka[,5]==0.5))
        remiz_7<-sum(studnicka[,5]==0.5)/
          (2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5))

        tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
          ((1-remiz_6)*pnorm((listina[j,2]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,4])/282.84),
            1-(1-remiz_6)*pnorm((listina[j,2]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,4])/282.84)-
            (1-remiz_7)*pnorm((listina[j,4]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,2])/282.84),
            (1-remiz_7)*pnorm((listina[j,4]-
            listina[j,2]-qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
        }
      else
      {
        tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c(prst_vyhry,prst_remizy,prst_prohry))
      }
    }
  }
}

```

```

#vycházím z typu turnaje
#odhaduji výsledky na základě četnosti v intervalu
#je tam zakomponované normální rozdělení
#vycházím z turnaje samotného

```

```

if((metoda==14)==TRUE){
  print("Dělám metodu 14.")
  pul_rozpeti<-50
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocna_matice<-matrix(nrow=length
        (which(studna[,6]>listina[j,5]-pul_rozpeti &
          studna[,6]<listina[j,5]+pul_rozpeti)),ncol=5)
      pomocna_matice[,1]<-
        which(studna[,6]>listina[j,5]-pul_rozpeti &
          studna[,6]<listina[j,5]+pul_rozpeti)
      pomocna_matice[,2]<-studna[pomocna_matice[,1],5]
      pomocna_matice[,3]<-pomocna_matice[,1]-listina[j,5]
      pomocna_matice[,4]<-abs(0.5-pnorm(pomocna_matice[,3],listina[j,5],25))
      pomocna_matice[,5]<-pomocna_matice[,2]*pomocna_matice[,4]
      prst_vyhry<-sum(pomocna_matice[which(pomocna_matice[,2]==1),4])/

```

```

sum(pomocna_matice[,4])
prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0.5),4])/
sum(pomocna_matice[,4])
prst_prohry<-prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0),4])/
sum(pomocna_matice[,4])
if(length(pomocna_matice[,1])<4)
{
remiz_6<-sum(studnicka[,5]==0.5)/
(2*sum(studnicka[,5]==1)+sum(studnicka[,5]==0.5))
remiz_7<-sum(studnicka[,5]==0.5)/
(2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5))

tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
((1-remiz_6)*pnorm((listina[j,2]+
qnorm(mean(studnicka[,5]),0,282.84)
-listina[j,4])/282.84),
1-(1-remiz_6)*pnorm((listina[j,2]+
qnorm(mean(studnicka[,5]),0,282.84)
-listina[j,4])/282.84)-
(1-remiz_7)*pnorm((listina[j,4]+
qnorm(mean(studnicka[,5]),0,282.84)
-listina[j,2])/282.84),
(1-remiz_7)*pnorm((listina[j,4]-
listina[j,2]-qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
}
else
{
tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c(prst_vyhry,prst_remizy,prst_prohry))
}
}
}
}

```

```

if((metoda==15)==TRUE){
print("Dělám metodu 15.")
pul_rozpeti<-80
for (i in 1:Nsim){
for (j in 1:nrow(listina)){studnicka[,6]
pomocna_matice<-matrix(nrow=length(which(((studnicka[,6]-listina[j,5])^2)+
((studnicka[,7]-listina[j,6])^2)<2500)),ncol=5)
pomocna_matice[,1]<-which(((studnicka[,6]-listina[j,5])^2)+
((studnicka[,7]-listina[j,6])^2)<2500)
pomocna_matice[,2]<-studnicka[pomocna_matice[,1],5]
pomocna_matice[,3]<-pomocna_matice[,1]-listina[j,5]
pomocna_matice[,4]<-abs(0.5-pnorm(pomocna_matice[,3],listina[j,5],25))
pomocna_matice[,5]<-pomocna_matice[,2]*pomocna_matice[,4]
prst_vyhry<-sum(pomocna_matice[which(pomocna_matice[,2]==1),4])/
sum(pomocna_matice[,4])
prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0.5),4])/
sum(pomocna_matice[,4])
prst_prohry<-prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0),4])/
sum(pomocna_matice[,4])
vyber<-studnicka[pomocna_matice[,1],5]
if(length(pomocna_matice[,1])<4){
remiz_6<-sum(studnicka[,5]==0.5)/
(2*sum(studnicka[,5]==1)+sum(studnicka[,5]==0.5))
remiz_7<-sum(studnicka[,5]==0.5)/
(2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5))

tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
((1-remiz_6)*pnorm((listina[j,2]+
qnorm(mean(studnicka[,5]),0,282.84)
-listina[j,4])/282.84),
1-(1-remiz_6)*pnorm((listina[j,2]+
qnorm(mean(studnicka[,5]),0,282.84)
-listina[j,4])/282.84)-
(1-remiz_7)*pnorm((listina[j,4]+
qnorm(mean(studnicka[,5]),0,282.84)
-listina[j,2])/282.84),
(1-remiz_7)*pnorm((listina[j,4]-
listina[j,2]-qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
}
else
{

```

```

prst_vyhry<-length(vyber[vyber==1])/length(vyber)
prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
prst_prohry<-length(vyber[vyber==0])/length(vyber)
if(length(vyber[vyber==1])==0){prst_vyhry<-
pnorm((listina[j,2]-listina[j,4])/282.84)}
if(length(vyber[vyber==0.5])==0)
{prst_remizy<-0}
if(length(vyber[vyber==0])==0)
{prst_prohry<-pnorm((listina[j,4]-listina[j,2])/282.84)}
tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
(prst_vyhry,prst_remizy,prst_prohry))
}
}
}
}
#vycházím z turnaje
#odhaduji výsledky na základě v kruhu s normálními váhami

```

```

if((metoda==16)==TRUE){
print("Dělám metodu 16.")
pul_rozpeti<-80
for (i in 1:Nsim){
for (j in 1:nrow(listina)){stredni_studna[,6]
pomocna_matrice<-
matrix(nrow=length(which(((stredni_studna[,6]-listina[j,5])^2)+
((stredni_studna[,7]-listina[j,6])^2)<2500)),ncol=5)
pomocna_matrice[,1]<-which(((stredni_studna[,6]-listina[j,5])^2)+
((stredni_studna[,7]-listina[j,6])^2)<2500)
pomocna_matrice[,2]<-stredni_studna[pomocna_matrice[,1],5]
pomocna_matrice[,3]<-pomocna_matrice[,1]-listina[j,5]
pomocna_matrice[,4]<-abs(0.5-pnorm(pomocna_matrice[,3],listina[j,5],25))
pomocna_matrice[,5]<-pomocna_matrice[,2]*pomocna_matrice[,4]
prst_vyhry<-sum(pomocna_matrice[which(pomocna_matrice[,2]==1),4])/
sum(pomocna_matrice[,4])
prst_remizy<-sum(pomocna_matrice[which(pomocna_matrice[,2]==0.5),4])/
sum(pomocna_matrice[,4])
prst_prohry<-prst_remizy<-sum(pomocna_matrice[which(pomocna_matrice[,2]==0),4])/
sum(pomocna_matrice[,4])
vyber<-stredni_studna[pomocna_matrice[,1],5]
if(length(pomocna_matrice[,1])<4){
remiz_6<-sum(studnicka[,5]==0.5)/
(2*sum(studnicka[,5]==1)+sum(studnicka[,5]==0.5))
remiz_7<-sum(studnicka[,5]==0.5)/
(2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5))

tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
((1-remiz_6)*pnorm((listina[j,2]+
qnorm(mean(studnicka[,5]),0,282.84)
-listina[j,4])/282.84),
1-(1-remiz_6)*pnorm((listina[j,2]+
qnorm(mean(studnicka[,5]),0,282.84)
-listina[j,4])/282.84)-
(1-remiz_7)*pnorm((listina[j,4]+
qnorm(mean(studnicka[,5]),0,282.84)
-listina[j,2])/282.84),
(1-remiz_7)*pnorm((listina[j,4]-
listina[j,2]-qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
}
else
{
prst_vyhry<-length(vyber[vyber==1])/length(vyber)
prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
prst_prohry<-length(vyber[vyber==0])/length(vyber)
if(length(vyber[vyber==1])==0){prst_vyhry<-
pnorm((listina[j,2]-listina[j,4])/282.84)}
if(length(vyber[vyber==0.5])==0){prst_remizy<-0}
if(length(vyber[vyber==0])==0){prst_prohry<-
pnorm((listina[j,4]-listina[j,2])/282.84)}
tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
(prst_vyhry,prst_remizy,prst_prohry))
}
}
}
}

```

```

}
}
#vycházím z typu turnaje
#odhaduji výsledky na základě v kruhu s normálními váhami

if((metoda==17)==TRUE){
  print("Dělám metodu 17.")
  pul_rozpeti<-80
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){ studna[,6]
      pomocna_matice<-
        matrix(nrow=length(which(((studna[,6]-listina[j,5])^2)+
          ((studna[,7]-listina[j,6])^2)<2500)),ncol=5)
      pomocna_matice[,1]<-which(((studna[,6]-listina[j,5])^2)+
        ((studna[,7]-listina[j,6])^2)<2500)
      pomocna_matice[,2]<-studna[pomocna_matice[,1],5]
      pomocna_matice[,3]<-pomocna_matice[,1]-listina[j,5]
      pomocna_matice[,4]<-abs(0.5-pnorm(pomocna_matice[,3],listina[j,5],2.5))
      pomocna_matice[,5]<-pomocna_matice[,2]*pomocna_matice[,4]
      prst_vyhry<-sum(pomocna_matice[which(pomocna_matice[,2]==1),4])/
        sum(pomocna_matice[,4])
      prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0.5),4])/
        sum(pomocna_matice[,4])
      prst_prohry<-prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0),4])/
        sum(pomocna_matice[,4])
      vyber<-studna[pomocna_matice[,1],5]
      if(length(pomocna_matice[,1])<4){
        remiz_6<-sum(studna[,5]==0.5)/
          (2*sum(studna[,5]==1)+sum(studnicka[,5]==0.5))
        remiz_7<-sum(studna[,5]==0.5)/
          (2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5))

        tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
          (((1-remiz_6)*pnorm((listina[j,2]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,4])/282.84),
            1-(1-remiz_6)*pnorm((listina[j,2]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,4])/282.84)-
            (1-remiz_7)*pnorm((listina[j,4]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,2])/282.84),
            (1-remiz_7)*pnorm((listina[j,4]-
            listina[j,2]-qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
        }
      else
      {
        prst_vyhry<-length(vyber[vyber==1])/length(vyber)
        prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
        prst_prohry<-length(vyber[vyber==0])/length(vyber)
      }
    if(length(vyber[vyber==1])>0){ prst_vyhry<-
      pnorm((listina[j,2]-listina[j,4])/282.84) }
    if(length(vyber[vyber==0.5])>0){ prst_remizy<-0 }
    if(length(vyber[vyber==0])>0){ prst_prohry<-
      pnorm((listina[j,4]-listina[j,2])/282.84) }
    tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
      (prst_vyhry,prst_remizy,prst_prohry))
  }
}
}
}
#vycházím z turnaje
#odhaduji výsledky na základě v kruhu s normálními váhami

if((metoda==18)==TRUE){
  print("Dělám metodu 18.")
  rozpeti_rozdilu<-30
  rozpeti_sil<-100
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocny_150<-c()
      pomocny_150<-intersect(which(studnicka[,6]>

```

```

        listina[j,5]-rozpeti_rozdil &
        studnicka[,6]<
        listina[j,5]+rozpeti_rozdil),
    which(studnicka[,7]>
        listina[j,6]-rozpeti_sil &
        studnicka[,7]<listina[j,6]
        +rozpeti_sil))
vyber<-studnicka[pomocny_150,5]
prst_vyhry<-length(vyber[vyber==1])/length(vyber)
prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
prst_prohry<-length(vyber[vyber==0])/length(vyber)
if(length(pomocny_150)<4){
    remiz_6<-sum(studnicka[,5]==0.5)/
    (2*sum(studnicka[,5]==1)+sum(studnicka[,5]==0.5))
    remiz_7<-sum(studnicka[,5]==0.5)/
    (2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5))

    tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        ((1-remiz_6)*pnorm((listina[j,2]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,4])/282.84),
        1-(1-remiz_6)*pnorm((listina[j,2]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,4])/282.84)-
        (1-remiz_7)*pnorm((listina[j,4]+
            qnorm(mean(studnicka[,5]),0,282.84)
            -listina[j,2])/282.84),
        (1-remiz_7)*pnorm((listina[j,4]-
            listina[j,2]-qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
    }
else
{
    tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        (prst_vyhry,prst_remizy,prst_prohry))
}
}
}
}
#XXX
#vycházím z turnaje
#odhaduji výsledky na základě četnosti v obdélníčku

if((metoda==19)==TRUE){
    print("Dělám metodu 19.")
    rozpeti_rozdil<-30
    rozpeti_sil<-100
    for (i in 1:Nsim){
        for (j in 1:nrow(listina)){
            pomocny_150<-c()
            pomocny_150<-
                intersect(which(stredni_studna[,6]>
                    listina[j,5]-rozpeti_rozdil &
                    stredni_studna[,6]<listina[j,5]+rozpeti_rozdil),
                    which(stredni_studna[,7]>listina[j,6]-rozpeti_sil &
                    stredni_studna[,7]<listina[j,6]+rozpeti_sil))
            vyber<-stredni_studna[pomocny_150,5]
            prst_vyhry<-length(vyber[vyber==1])/length(vyber)
            prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
            prst_prohry<-length(vyber[vyber==0])/length(vyber)
            if(length(pomocny_150)<4){
                remiz_6<-sum(studnicka[,5]==0.5)/
                (2*sum(studnicka[,5]==1)+sum(studnicka[,5]==0.5))
                remiz_7<-sum(studnicka[,5]==0.5)/
                (2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5))

                tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
                    ((1-remiz_6)*pnorm((listina[j,2]+
                        qnorm(mean(studnicka[,5]),0,282.84)
                        -listina[j,4])/282.84),
                    1-(1-remiz_6)*pnorm((listina[j,2]+
                        qnorm(mean(studnicka[,5]),0,282.84)
                        -listina[j,4])/282.84)-
                    (1-remiz_7)*pnorm((listina[j,4]+
                        qnorm(mean(studnicka[,5]),0,282.84)
                        -listina[j,2])/282.84),
                    (1-remiz_7)*pnorm((listina[j,4]-
                        qnorm(mean(studnicka[,5]),0,282.84)
                        -listina[j,2])/282.84)))
            }
        }
    }
}

```

```

        (1-remiz_7)*pnorm((listina[j,4]-
listina[j,2]-qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
    }
else
{
    tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        (prst_vyhry,prst_remizy,prst_prohry))
}
}
}
}
#XXX
#vycházím z typu turnaje
#odhaduji výsledky na základě četnosti v obdélníčku

if((metoda==20)==TRUE){
    print("Dělám metodu 20.")
    rozpeti_rozdil<-30
    rozpeti_sil<-100
    for (i in 1:Nsim){
        for (j in 1:nrow(listina)){
            pomocny_150<-c()
            pomocny_150<-intersect(which(studna[,6]>
                listina[j,5]-rozpeti_rozdil &
                studna[,6]<listina[j,5]+
                rozpeti_rozdil),
                which(studna[,7]>listina[j,6]-
                    rozpeti_sil &
                    studna[,7]<listina[j,6]+rozpeti_sil))
            vyber<-studna[pomocny_150,5]
            prst_vyhry<-length(vyber[vyber==1])/length(vyber)
            prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
            prst_prohry<-length(vyber[vyber==0])/length(vyber)
            if(length(pomocny_150)<4){
                remiz_6<-sum(studnicka[,5]==0.5)/
                    (2*sum(studnicka[,5]==1)+sum(studnicka[,5]==0.5))
                remiz_7<-sum(studnicka[,5]==0.5)/
                    (2*sum(studnicka[,5]==0)+sum(studnicka[,5]==0.5))

                tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
                    ((1-remiz_6)*pnorm((listina[j,2]+
                        qnorm(mean(studnicka[,5]),0,282.84)
                        -listina[j,4])/282.84),
                    1-(1-remiz_6)*pnorm((listina[j,2]+
                        qnorm(mean(studnicka[,5]),0,282.84)
                        -listina[j,4])/282.84)-
                    (1-remiz_7)*pnorm((listina[j,4]+
                        qnorm(mean(studnicka[,5]),0,282.84)
                        -listina[j,2])/282.84),
                    (1-remiz_7)*pnorm((listina[j,4]-
                        listina[j,2]-qnorm(mean(studnicka[,5]),0,282.84))/282.84)))
            }
            else
            {
                tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
                    (prst_vyhry,prst_remizy,prst_prohry))
            }
        }
    }
}
#XXX
#vycházím z databáze
#odhaduji výsledky na základě četnosti v obdélníčku

#vycházím z databáze
#odhaduji výsledky na základě četnosti v intervalu
#je tam zakomponované normální rozdělení
#vycházím z turnaje samotného

if((metoda==112)==TRUE){
    print("Dělám metodu 12.")

```



```

pul_rozpeti<-50
for (i in 1:Nsim){
  for (j in 1:nrow(listina)){
    pomocny_150<-c()
    pomocny_150<-
      which(studnicka[,6]>listina[j,5]-pul_rozpeti & studnicka[,6]<
        listina[j,5]+pul_rozpeti)
    vyber<-studnicka[pomocny_150,5]
    prst_vyhry<-length(vyber[vyber==1])/length(vyber)
    prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
    prst_prohry<-length(vyber[vyber==0])/length(vyber)

    tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
      (prst_vyhry,prst_remizy,prst_prohry))
  }
}
#vycházím z turnaje
#odhaduji výsledky na základě četnosti v intervalu
#vycházím z turnaje samotného

#vycházím z databáze
#odhaduji výsledky na základě četnosti v intervalu

if((metoda==444)==TRUE){
  print("Dělám metodu 444.")
  remiz_6<-rep(sum(studna[,5]==0.5)/length(studna[,5]),nrow(listina))
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      if (((pnorm((listina[j,2]-listina[j,4])/282.84))<remiz_6[j]/2|
        ((pnorm(listina[j,4]-listina[j,2])/282.84))<remiz_6[j]/2)==TRUE)
      {
        remiz_6[j]<-2*min(pnorm((listina[j,2]-listina[j,4])/282.84),
          (pnorm(listina[j,4]-listina[j,2])/282.84))
      }

      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        (pnorm((listina[j,2]-listina[j,4])/282.84)-remiz_6[j]/2,
          remiz_6[j],
          pnorm((listina[j,4]-listina[j,2])/282.84)-remiz_6[j]/2))
    }
  }
}
#vezmu remizovost z turnaje a tím podílem upravím očekávaný výsledek
#pokud je očekávaná remizovost větší než dvojnásobek pravděpodobnosti
#výhry nebo prohry, pak ji snížím právě na tolik
#tady vycházím z remizovosti jen v rámci daného turnaje

if((metoda==55)==TRUE){
  print("Dělám metodu 5.")
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      tipy[j,i]<-sample(x=c(1,0),1,prob=c
        (pnorm((listina[j,2]-listina[j,4]+
          qnorm(mean(stredni_studna[,5]),0,282.84))/282.84),
          pnorm((listina[j,4]-listina[j,2]-
          qnorm(mean(stredni_studna[,5]),0,282.84))/282.84)))
    }
  }
}
#jedničky a nuly odpovídající očekávaným výsledkům s tím
#že bílý má fixní výhodu, která vychází z turnajů podobného typu

if((metoda==65)==TRUE){
  print("Dělám metodu 6.")
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      tipy[j,i]<-sample(x=c(1,0),1,prob=c
        (pnorm((listina[j,2]-listina[j,4]+qnorm(mean(studna[,5]),0,282.84))/282.84),
          pnorm((listina[j,4]-listina[j,2]-qnorm(mean(studna[,5]),0,282.84))/282.84)))
    }
  }
}

```

```

    }
  }
}
#jedničky a nuly odpovídající očekávaným výsledkům s tím
#že bílý má fixní výhodu, která vychází z celkové databáze

if((metoda==77)==TRUE){
  print("Dělám metodu 7.")
remiz_6<-rep(sum(stredni_studna[,5]==0.5)/
  length(stredni_studna[,5]),nrow(listina))
for (i in 1:Nsim){
  for (j in 1:nrow(listina)){
    if(((pnorm((listina[j,2]-listina[j,4])/282.84))<remiz_6[j]/2|
      ((pnorm(listina[j,4]-listina[j,2])/282.84))<remiz_6[j]/2)==TRUE)
      {
        remiz_6[j]<-2*min(pnorm((listina[j,2]-listina[j,4])/282.84),
          (pnorm(listina[j,4]-listina[j,2])/282.84))
      }

    tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
      (pnorm((listina[j,2]-listina[j,4])/282.84)-remiz_6[j]/2,
        remiz_6[j],
        pnorm((listina[j,4]-listina[j,2])/282.84)-remiz_6[j]/2))
  }
}
}
#vezmu remízovost z turnaje a tím podílem poupravím očekávaný výsledek
#pokud je očekávaná remízovost větší než dvojnásobek pravděpodobnosti
#výhry nebo prohry, pak ji snížím právě na tolik
#tady vycházím z remízovosti jen v rámci daného turnaje

if((metoda==30)==TRUE){
  print("Dělám metodu 10.")
remiz_6<-rep(sum(stredni_studna[,5]==0.5)/
  (2*sum(stredni_studna[,5]==1)+
  sum(stredni_studna[,5]==0.5)),nrow(listina))
remiz_7<-rep(sum(stredni_studna[,5]==0.5)/
  (2*sum(stredni_studna[,5]==0)+
  sum(stredni_studna[,5]==0.5)),nrow(listina))
for (i in 1:Nsim){
  for (j in 1:nrow(listina)){
    tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
      ((1-remiz_6[j])*pnorm((listina[j,2]-listina[j,4])/282.84),
        1-(1-remiz_6[j])*pnorm((listina[j,2]-listina[j,4])/282.84)-
        (1-remiz_7[j])*pnorm((listina[j,4]-listina[j,2])/282.84),
        (1-remiz_7[j])*pnorm((listina[j,4]-listina[j,2])/282.84)))
  }
}
}
#vycházím z typu turnaje
#tahle metoda zohledňuje jaký podíl tvoří remízy na zisku hráčů
#vpodstatě to je vylepšená předchozí metoda

if((metoda==31)==TRUE){
  print("Dělám metodu 11.")
remiz_6<-rep(sum(studna[,5]==0.5)/
  (2*sum(studna[,5]==1)+sum(studna[,5]==0.5)),nrow(listina))
remiz_7<-rep(sum(studna[,5]==0.5)/
  (2*sum(studna[,5]==0)+sum(studna[,5]==0.5)),nrow(listina))
for (i in 1:Nsim){
  for (j in 1:nrow(listina)){
    tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
      ((1-remiz_6[j])*pnorm((listina[j,2]-listina[j,4])/282.84),
        1-(1-remiz_6[j])*pnorm((listina[j,2]-listina[j,4])/282.84)-
        (1-remiz_7[j])*pnorm((listina[j,4]-listina[j,2])/282.84),
        (1-remiz_7[j])*pnorm((listina[j,4]-listina[j,2])/282.84)))
  }
}
}

```

```

}
#vycházím z úplné databáze
#tahle metoda zohledňuje jaký podíl tvoří remízy na zisku hráčů
#vpodstatě to je vylepšená předchozí metoda

if((metoda==112)==TRUE){
  remiz_6<-c()
  remiz_7<-c()
  pul_rozpeti<-45
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocny_150<-c()
      pomocny_150<-
        which(studnicka[,6]>listina[j,5]-pul_rozpeti & studnicka[,6]<
              listina[j,5]+pul_rozpeti)
      vyber<-studnicka[pomocny_150,5]
      remiz_6[j]<-length(vyber[vyber==0.5])/length(vyber)
      remiz_7[j]<-length(vyber[vyber==0.5])/length(vyber)
      if (pnorm((listina[j,2]-listina[j,4])/282.84)<remiz_6[j]/2)
      (pnorm((listina[j,4]-listina[j,2])/282.84)<remiz_6[j]/2)==TRUE){
        remiz_6[j]<-2*min(mean(studnicka[,5])/
                          0.5*pnorm((listina[j,2]-listina[j,4])/282.84),
                          (pnorm(listina[j,4]-listina[j,2])/282.84))

        tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
                          ((1-remiz_6[j])*pnorm((listina[j,2]-listina[j,4])/282.84),
                           1-(1-remiz_6[j])*pnorm((listina[j,2]-listina[j,4])/282.84)-
                           (1-remiz_7[j])*pnorm((listina[j,4]-listina[j,2])/282.84),
                           (1-remiz_7[j])*pnorm((listina[j,4]-listina[j,2])/282.84)))
      }
    }
  }
}
#vycházím z turnaje
#dívám se na partie v obdélníku +-45 a na výskyt remíz odtamtud, le

if((metoda==113)==TRUE){
  remiz_6<-c()
  pul_rozpeti<-45
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocny_150<-c()
      pomocny_150<-
        which(stredni_studna[,6]>listina[j,5]
              -pul_rozpeti & stredni_studna[,6]<
              listina[j,5]+pul_rozpeti)
      vyber<-stredni_studna[pomocny_150,5]
      remiz_6[j]<-length(vyber[vyber==0.5])/length(vyber)
      if (pnorm((listina[j,2]-listina[j,4])/282.84)<remiz_6[j]/2)
      (pnorm((listina[j,4]-listina[j,2])/
              282.84)<remiz_6[j]/2)==TRUE){
        remiz_6[j]<-2*min(mean(stredni_studna[,5])/
                          0.5*pnorm((listina[j,2]-
                                      listina[j,4])/282.84),
                          (pnorm(listina[j,4]-listina[j,2])/282.84))

        tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
                          ((1-remiz_6[j])*pnorm((listina[j,2]
                                                  -listina[j,4])/282.84),
                           1-(1-remiz_6[j])*pnorm((listina[j,2]
                                                  -listina[j,4])/282.84)-
                           (1-remiz_7[j])*pnorm((listina[j,4]
                                                  -listina[j,2])/282.84),
                           (1-remiz_7[j])*pnorm((listina[j,4]-listina[j,2])/282.84)))
      }
    }
  }
}
#vycházím z typu turnaje
#dívám se na partie v obdélníku +-45 a na výskyt remíz odtamtud, lep

if((metoda==114)==TRUE){
  remiz_6<-c()
  pul_rozpeti<-45
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){

```

```

pomocny_150<-c()
pomocny_150<-
  which(studna[,6]>listina[j,5]-pul_rozpeti & studna[,6]<
    listina[j,5]+pul_rozpeti)
vyber<-studna[pomocny_150,5]
remiz_6[j]<-length(vyber[vyber==0.5])/length(vyber)
if (pnorm((listina[j,2]-listina[j,4])/282.84)<remiz_6[j]/2|
(pnorm((listina[j,4]-listina[j,2])/282.84)<remiz_6[j]/2)==TRUE){
  remiz_6[j]<-2*min(mean(studna[,5])/
    0.5*pnorm((listina[j,2]-listina[j,4])/282.84),
    (pnorm(listina[j,4]-listina[j,2])/282.84))

  tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
    ((1-remiz_6[j])*pnorm((listina[j,2]-listina[j,4])/282.84),
    1-(1-remiz_6[j])*pnorm((listina[j,2]-listina[j,4])/282.84)-
    (1-remiz_7[j])*pnorm((listina[j,4]-listina[j,2])/282.84),
    (1-remiz_7[j])*pnorm((listina[j,4]-listina[j,2])/282.84)))
}
}
}
#vycházím z databáze
#dívám se na partie v obdélníku +-45 a na výskyt remíz odtamtud,

if((metoda==188)==TRUE){
  pul_rozpeti<-65
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocny_150<-c()
      pomocny_150<-
        intersect(which(studnicka[,6]>listina[j,5]-pul_rozpeti &
          studnicka[,6]
            <listina[j,5]+pul_rozpeti),
          which(studnicka[,7]>
            listina[j,6]-pul_rozpeti &
            studnicka[,7]<
            listina[j,6]+pul_rozpeti))
      vyber<-studnicka[pomocny_150,5]
      prst_vyhry<-length(vyber[vyber==1])/length(vyber)
      prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
      prst_prohry<-length(vyber[vyber==0])/length(vyber)
      if(length(vyber[vyber==1])==0){prst_vyhry<-
        pnorm((listina[j,2]-listina[j,4])/282.84)}
      if(length(vyber[vyber==0.5])==0){prst_remizy<-0}
      if(length(vyber[vyber==0])==0){prst_prohry<-
        pnorm((listina[j,4]-listina[j,2])/282.84)}
      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        (prst_vyhry,prst_remizy,prst_prohry))
    }
  }
}
#XXX
#vycházím z turnaje
#odhaduji výsledky na základě četnosti v obdélníku

if((metoda==159)==TRUE){
  pul_rozpeti<-45
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocny_150<-c()
      pomocny_150<-intersect(which(stredni_studna[,6]>
        listina[j,5]-pul_rozpeti &
        stredni_studna[,6]<
        listina[j,5]+pul_rozpeti),
        which(stredni_studna[,7]>
          listina[j,6]-pul_rozpeti &
          stredni_studna[,7]<
          listina[j,6]+pul_rozpeti))
      vyber<-stredni_studna[pomocny_150,5]
      prst_vyhry<-length(vyber[vyber==1])/length(vyber)
      prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
      prst_prohry<-length(vyber[vyber==0])/length(vyber)

      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        (prst_vyhry,prst_remizy,prst_prohry))
    }
  }
}

```

```

    }
  }
}
#vycházím z typu turnaje
#odhaduji výsledky na základě četnosti v obdélníčku

if((metoda==200)==TRUE){
  remiz_6<-c()
  pul_rozpeti<-15
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocny_150<-c()
      pomocny_150<-intersect(which(studna[,6]>listina[j,5]-pul_rozpeti &
        studna[,6]<listina[j,5]+pul_rozpeti),
        which(studna[,7]>listina[j,6]-pul_rozpeti &
          studna[,7]<listina[j,6]+pul_rozpeti))
      vyber<-studna[pomocny_150,5]
      prst_vyhry<-length(vyber[vyber==1])/length(vyber)
      prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
      prst_prohry<-length(vyber[vyber==0])/length(vyber)

      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        (prst_vyhry,prst_remizy,prst_prohry))
    }
  }
}
#vycházím z databáze
#odhaduji výsledky na základě četnosti v obdélníčku

if((metoda==31)==TRUE){
  pul_rozpeti<-45
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocna_matice<-matrix(nrow=length
        (which(studnicka[,6]>listina[j,5]-pul_rozpeti &
          studnicka[,6]<listina[j,5]+pul_rozpeti)),ncol=5)
      pomocna_matice[,1]<-
        which(studnicka[,6]>listina[j,5]-pul_rozpeti &
          studnicka[,6]<listina[j,5]+pul_rozpeti)
      pomocna_matice[,2]<-studnicka[pomocna_matice[,1],5]
      pomocna_matice[,3]<-pomocna_matice[,1]-listina[j,5]
      pomocna_matice[,4]<-pnorm(pomocna_matice[,3],listina[j,5],15)
      pomocna_matice[,5]<-pomocna_matice[,2]*pomocna_matice[,4]
      prst_vyhry<-sum(pomocna_matice[which(pomocna_matice[,2]==1),4])/
        sum(pomocna_matice[,4])
      prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0.5),4])/
        sum(pomocna_matice[,4])
      prst_prohry<-prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0),4])/
        sum(pomocna_matice[,4])

      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c(prst_vyhry,prst_remizy,prst_prohry))
    }
  }
}

if((metoda==32)==TRUE){
  pul_rozpeti<-45
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocna_matice<-matrix(nrow=length
        (which(studnicka[,6]>listina[j,5]-pul_rozpeti &
          studnicka[,6]<listina[j,5]+pul_rozpeti)),ncol=5)
      pomocna_matice[,1]<-
        which(studnicka[,6]>listina[j,5]-pul_rozpeti &
          studnicka[,6]<listina[j,5]+pul_rozpeti)
      pomocna_matice[,2]<-studnicka[pomocna_matice[,1],5]
      pomocna_matice[,3]<-pomocna_matice[,1]-listina[j,5]
      pomocna_matice[,4]<-pnorm(pomocna_matice[,3],listina[j,5],15)
      pomocna_matice[,5]<-pomocna_matice[,2]*pomocna_matice[,4]
      prst_vyhry<-sum(pomocna_matice[which(pomocna_matice[,2]==1),4])/
        sum(pomocna_matice[,4])
      prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0.5),4])/
        sum(pomocna_matice[,4])
    }
  }
}

```

```

prst_prohry<-prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0),4])/
sum(pomocna_matice[,4])

tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c(prst_vyhry,prst_remizy,prst_prohry))
}
}

if((metoda==33)==TRUE){
pul_rozpeti<-45
for (i in 1:Nsim){
for (j in 1:nrow(listina)){
pomocna_matice<-matrix(nrow=length
(which(studnicka[,6]>listina[j,5]-pul_rozpeti &
studnicka[,6]<listina[j,5]+pul_rozpeti)),ncol=5)
pomocna_matice[,1]<-
which(studnicka[,6]>listina[j,5]-pul_rozpeti &
studnicka[,6]<listina[j,5]+pul_rozpeti)
pomocna_matice[,2]<-studnicka[pomocna_matice[,1],5]
pomocna_matice[,3]<-pomocna_matice[,1]-listina[j,5]
pomocna_matice[,4]<-pnorm(pomocna_matice[,3],listina[j,5],15)
pomocna_matice[,5]<-pomocna_matice[,2]*pomocna_matice[,4]
prst_vyhry<-sum(pomocna_matice[which(pomocna_matice[,2]==1),4])/
sum(pomocna_matice[,4])
prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0.5),4])/
sum(pomocna_matice[,4])
prst_prohry<-prst_remizy<-sum(pomocna_matice[which(pomocna_matice[,2]==0),4])/
sum(pomocna_matice[,4])

tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c(prst_vyhry,prst_remizy,prst_prohry))
}
}
}

if((metoda==201)==TRUE){
remiz_6<-c()
pul_rozpeti<-45
for (i in 1:Nsim){
for (j in 1:nrow(listina)){
pomocna_matice<-
matrix(nrow=length
(which(stredni_studna[,6]>
listina[j,5]-pul_rozpeti &
stredni_studna[,6]<
listina[j,5]+pul_rozpeti))
,ncol=5)
pomocna_matice[,1]<-
which(studna[,6]>listina[j,5]-pul_rozpeti &
studna[,6]<listina[j,5]+pul_rozpeti)
pomocna_matice[,2]<-stredni_studna[pomocna_matice[,1],5]
pomocna_matice[,3]<-pomocna_matice[,1]-listina[j,5]
pomocna_matice[,4]<-pnorm(pomocna_matice[,3],listina[j,5],15)
pomocna_matice[,5]<-pomocna_matice[,2]*pomocna_matice[,4]
remiz_6[j]<-sum(pomocna_matice[which(pomocna_matice[,2]==0.5),4])/
sum(pomocna_matice[,4])

if (pnorm((listina[j,2]-listina[j,4])/282.84)<remiz_6[j]/2|
(pnorm((listina[j,4]-listina[j,2])/282.84)<remiz_6[j]/2)==TRUE){
remiz_6[j]<-2*min(mean(studna[,5])/0.5*
pnorm((listina[j,2]-listina[j,4])/282.84),
(pnorm(listina[j,4]-listina[j,2])/282.84))}

tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
(pnorm((listina[j,2]-listina[j,4])/282.84)-remiz_6[j]/2,
remiz_6[j],
pnorm((listina[j,4]-listina[j,2])/282.84)-remiz_6[j]/2))
}
}
}

if((metoda==201)==TRUE){
remiz_6<-c()
pul_rozpeti<-45
for (i in 1:Nsim){

```

```

for (j in 1:nrow(listina)){
  pomocna_matice<-matrix(nrow=length
    (which(stredni_studna[,6]>listina[j,5]-pul_rozpeti &
      stredni_studna[,6]<listina[j,5]+pul_rozpeti)),ncol=5)
  pomocna_matice[,1]<-
    which(studna[,6]>listina[j,5]-pul_rozpeti &
      studna[,6]<listina[j,5]+pul_rozpeti)
  pomocna_matice[,2]<-studna[pomocna_matice[,1],5]
  pomocna_matice[,3]<-pomocna_matice[,1]-listina[j,5]
  pomocna_matice[,4]<-pnorm(pomocna_matice[,3],listina[j,5],15)
  pomocna_matice[,5]<-pomocna_matice[,2]*pomocna_matice[,4]
  remiz_6[j]<-sum(pomocna_matice[which(pomocna_matice[,2]==0.5),4])/
    sum(pomocna_matice[,4])

  if (pnorm((listina[j,2]-listina[j,4])/282.84)<remiz_6[j]/2)
  (pnorm((listina[j,4]-listina[j,2])/282.84)<remiz_6[j]/2)==TRUE){
    remiz_6[j]<-2*min(mean(studna[,5])/0.5*
      pnorm((listina[j,2]-listina[j,4])/282.84),
      (pnorm(listina[j,4]-listina[j,2])/282.84))

    tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
      (pnorm((listina[j,2]-listina[j,4])/282.84)-remiz_6[j]/2,
        remiz_6[j],
        pnorm((listina[j,4]-listina[j,2])/282.84)-remiz_6[j]/2))
  }
}

if((metoda==201)==TRUE){
  remiz_6<-c()
  pul_rozpeti<-45
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocny_150<-c()
      pomocny_150<-
        which(stredni_studna[,6]>listina[j,5]-pul_rozpeti & stredni_studna[,6]<
          listina[j,5]+pul_rozpeti)
      vyber<-stredni_studna[pomocny_150,5]
      remiz_6[j]<-length(vyber[vyber==0.5])/length(vyber)
      if (pnorm((listina[j,2]-listina[j,4])/282.84)<remiz_6[j]/2)
      (pnorm((listina[j,4]-listina[j,2])/282.84)<remiz_6[j]/2)==TRUE){
        remiz_6[j]<-2*min(mean(stredni_studna[,5])/
          0.5*pnorm((listina[j,2]-listina[j,4])/282.84),
          (pnorm(listina[j,4]-listina[j,2])/282.84))

        tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
          (pnorm((listina[j,2]-listina[j,4])/282.84)-remiz_6[j]/2,
            remiz_6[j],
            pnorm((listina[j,4]-listina[j,2])/282.84)-remiz_6[j]/2))
      }
    }
  }

if((metoda==171)==TRUE){
  remiz_6<-c()
  pul_rozpeti<-45
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocny_150<-c()
      pomocny_150<-
        which(studna[,6]>listina[j,5]-pul_rozpeti & studna[,6]<
          listina[j,5]+pul_rozpeti)
      vyber<-studna[pomocny_150,5]
      remiz_6[j]<-length(vyber[vyber==0.5])/length(vyber)
      if (pnorm((listina[j,2]-listina[j,4])/282.84)<remiz_6[j]/2)
      (pnorm((listina[j,4]-listina[j,2])/282.84)<remiz_6[j]/2)==TRUE){
        remiz_6[j]<-2*min(mean(studnicka[,5])/
          0.5*pnorm((listina[j,2]-listina[j,4])/282.84),
          (pnorm(listina[j,4]-listina[j,2])/282.84))

        tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
          (pnorm((listina[j,2]-listina[j,4])/282.84)-remiz_6[j]/2,
            remiz_6[j],
            pnorm((listina[j,4]-listina[j,2])/282.84)-remiz_6[j]/2))
      }
    }
  }
}

```

```

}

if((metoda==250)==TRUE){
  remiz_6<-c()
  pul_rozpeti<-45
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocna_matrice<-matrix(nrow=length
        (which(stredni_studna[,6]>listina[j,5]-pul_rozpeti &
          stredni_studna[,6]<listina[j,5]+pul_rozpeti)),ncol=5)
      pomocna_matrice[,1]<-
        which(studna[,6]>listina[j,5]-pul_rozpeti &
          studna[,6]<listina[j,5]+pul_rozpeti)
      pomocna_matrice[,2]<-studna[pomocna_matrice[,1],5]
      pomocna_matrice[,3]<-pomocna_matrice[,1]-listina[j,5]
      pomocna_matrice[,4]<-pnorm(pomocna_matrice[,3],listina[j,5],15)
      pomocna_matrice[,5]<-pomocna_matrice[,2]*pomocna_matrice[,4]
      remiz_6[j]<-sum(pomocna_matrice[which(pomocna_matrice[,2]==0.5),4])/
        sum(pomocna_matrice[,4])

      if (pnorm((listina[j,2]-listina[j,4])/282.84)<remiz_6[j]/2|
        (pnorm((listina[j,4]-listina[j,2])/282.84)<remiz_6[j]/2)==TRUE){
        remiz_6[j]<-2*min(mean(studna[,5])/0.5*
          pnorm((listina[j,2]-listina[j,4])/282.84),
          (pnorm(listina[j,4]-listina[j,2])/282.84))}

      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        (pnorm((listina[j,2]-listina[j,4])/282.84)-remiz_6[j]/2,
          remiz_6[j],
          pnorm((listina[j,4]-listina[j,2])/282.84)-remiz_6[j]/2))
    }
  }
}

if((metoda==30)==TRUE){
  remiz_6<-c()
  pul_rozpeti<-15
  for (i in 1:Nsim){
    for (j in 1:nrow(listina)){
      pomocny_150<-c()
      pomocny_150<-intersect(which(studna[,6]>listina[j,5]-pul_rozpeti &
        studna[,6]<listina[j,5]+pul_rozpeti),
        which(studna[,7]>listina[j,6]-pul_rozpeti &
          studna[,7]<listina[j,6]+pul_rozpeti))
      vyber<-studna[pomocny_150,5]
      prst_vyhry<-length(vyber[vyber==1])/length(vyber)
      prst_remizy<-length(vyber[vyber==0.5])/length(vyber)
      prst_prohry<-length(vyber[vyber==0])/length(vyber)

      tipy[j,i]<-sample(x=c(1,0.5,0),1,prob=c
        (prst_vyhry,prst_remizy,prst_prohry))
    }
  }
}

#a už mám matici tipů, jak dopadnou partie posledního kola, ovšem jen
#pro bílé
}

#ŠESTÁ, PODOBNÁ MATICE TIPŮ PRO ČERNÉ
#VÝSTUPEM JE tipy_cerne
{tipy_cerne<-matrix(nrow=nrow(listina),ncol=Nsim)
for (i in 1:Nsim){
  for (j in 1:nrow(listina)){
    tipy_cerne[j,i]<-1-tipy[j,i]
  }
}
#a už mám i matici tipů pro černé
}

#SEDMÁ PŘÍPRAVNÝ KROK, VYROBÍM SI ZE SOUPEROVE A VYSLEDKY_CISLA MATICE

```



```

{ for (i in 1:nrow(souperove)) {
  for (j in 7:15) {
    souperove[i,j]<-as.numeric(souperove[i,j])
    vysledky_cisla[i,j]<-as.numeric(vysledky_cisla[i,j])
  }
}
}

#OSMÁ DOOPRAVDY VÝSLEDKY
{
  analyza<-matrix(ncol=Nsim,nrow=nrow(souperove))
  bodiky<-matrix(ncol=Nsim,nrow=nrow(souperove))
  matice_tipu_posledniho_kola<-matrix(nrow=2*nrow(listina),ncol=Nsim)
  pomocna_matice<-matrix(nrow=2*nrow(listina),ncol=Nsim)
  for (k in 1:Nsim)
  {
    listina[,7]<-tipy[,k]
    listina[,8]<-tipy_cerne[,k]

    posledni_kolo_verze1<-
      matrix(rbind(matrix(cbind(listina[,1],listina[,7]),ncol=2),
        matrix(cbind(listina[,3],listina[,8]),ncol=2),
        matrix(cbind(nehrali_posledni[,1],nehrali_posledni[,3]),
          ncol=2)),ncol=2)
    posledni_kolo_verze2<-
      posledni_kolo_verze1[order(posledni_kolo_verze1[,1],decreasing=FALSE),]

    posledni_kolo_korelace<-rbind(matrix(cbind(listina[,1],listina[,7]),ncol=2),
      matrix(cbind(listina[,3],listina[,8]),ncol=2))
    posledni_kolo_korelace_2<-posledni_kolo_korelace[order
      (posledni_kolo_korelace[,1],decreasing=FALSE),]

    for (i in 1:length(cisla)) {
      posledni_kolo_verze2[cisla[i],2]<-predpoklady[i]
    }

    vysledky_cisla[,15]<-posledni_kolo_verze2[,2]

    kontumace[nehrali_posledni[,1],15]<-nehrali_posledni_navic

    matice_bucholzu<-matrix(nrow=nrow(start_praha22),ncol=10)
    souperove_v_matici<-matrix(nrow=nrow(start_praha22),ncol=10)
    vysledky_cisla_v_matici_pro_nekontumovane<-
      matrix(nrow=nrow(start_praha22),ncol=9)

    for (i in 1:nrow(souperove)) {
      for (j in 1:9) {
        souperove_v_matici[i,j]<-as.numeric(souperove[i,j+6])
        vysledky_cisla_v_matici_pro_nekontumovane[i,j]<-
          as.numeric(vysledky_cisla[i,j+6])
      }
    }

    for (i in 1:nrow(souperove)) {
      for (j in 1:9){
        if((souperove_v_matici[i,j]<999)==TRUE) {
          matice_bucholzu[i,j]<-(sum(vysledky_cisla_v_matici_pro_nekontumovane
            [souperove_v_matici[i,j],1:9])
            -0.5*sum(kontumace[souperove_v_matici[i,j],
              7:14] == "volno" |
              kontumace[souperove_v_matici[i,j],
              7:14] == "+")
            -0.5*sum(kontumace[souperove_v_matici[i,j],
              15] == "volno")
            +0.5*sum(kontumace[souperove_v_matici[i,j],
              7:14] == "-" |
              kontumace[souperove_v_matici[i,j],
              7:14] == "nenasazen")
            +0.5*sum(kontumace[souperove_v_matici[i,j],
              15] == "nenasazen"))}
        }
      }
    }
  }
}

```

```

}

kumulace_pro_virtualni_nekontumovane<-matrix(nrow=nrow(souperove),ncol=10)
kumulace_pro_virtualni_nekontumovane[,1]<-
  rep(0,nrow(kumulace_pro_virtualni_nekontumovane))
for (i in 1:nrow(souperove)){
  for (j in 1:9){
    kumulace_pro_virtualni_nekontumovane[i,j+1]<-
      sum(vysledky_cisla_v_matici_pro_nekontumovane[i,1:j])
  }

  for (i in 1:nrow(souperove)) {
    for (j in 1:9){
      if (is.na(matice_bucholzu[i,j])==TRUE) {
        matice_bucholzu[i,j]<-kumulace_pro_virtualni_nekontumovane[i,j]+
          (1-vysledky_cisla_v_matici_pro_nekontumovane[i,j])+
          0.5*(9-j)
      }
    }
  }
  for (i in 1:nrow(souperove))
  {
    matice_bucholzu[i,10]<-sum(matice_bucholzu[i,1:9])-
      min(matice_bucholzu[i,1:9])
  }
  matice_poradi<-matrix(nrow=nrow(matice_bucholzu),ncol=9)
  for (i in 1:nrow(matice_bucholzu)){
    matice_poradi[i,1]<-souperove[i,1]
  }
  matice_poradi[i,2]<-sum(vysledky_cisla_v_matici_pro_nekontumovane[i,1:9])
  matice_poradi[i,3]<-sum(matice_bucholzu[i,1:9])-
    min(matice_bucholzu[i,1:9])
  matice_poradi[i,5]<-sum(matice_bucholzu[i,1:9])-
    max(matice_bucholzu[i,1:9])
  matice_poradi[i,6]<-sum(matice_bucholzu[i,1:9])
  matice_poradi[i,7]<-sum(matice_bucholzu[i,1:9])-
    min(matice_bucholzu[i,1:9])-max(matice_bucholzu[i,1:9])
  matice_poradi[i,8]<-vektor_prumernych_el[i]
  matice_poradi[i,9]<-sum(kumulace_pro_virtualni_nekontumovane[i,])
  matice_poradi[,4]<-
    rank(-matice_poradi[,2]-matice_poradi[,ph_1]/1000
      -matice_poradi[,ph_2]/1000000,ties.method = "random")
}
analiza[,k]<-matice_poradi[,4]
bodiky[,k]<-matice_poradi[,2]
matice_tipu_posledniho_kola[,k]<-posledni_kolo_korelace_2[,2]
pomocna_matice[,k]<-posledni_kolo_korelace_2[,1]
if (round(k/(Nsim/10))=k/(Nsim/10))
  {print(paste("Mám už vyzkoušeno",k,"simulací."))}
}
bodiky<<-bodiky
jakodataframe_2<-as.data.frame(analyza)
cil_3<-paste(misto,"vysledky_simulace_",nazev,"_",metoda,".xlsx",sep="")
write_xlsx(jakodataframe_2,path=cil_3)

```

```

#tady jsem skončil, jak udělat pořadí podle počtu bodů a současně
#bucholzu
#to by měla být ta matice_bucholzu[,4]
}

```

```

#myslím, že další příprava na korelace
{
  kod<-paste(misto,nazev,".xlsx",sep="")

```

```

  start_praha22<-as.data.frame(read_xlsx(kod))

```

```

  vysledkove<-start_praha22
  for (i in 1:nrow(start_praha22)){
    for (j in 7:14){
      if(nchar(start_praha22[i,j])==5)
        {vysledkove[i,j]<-substring(start_praha22[i,j],5,5)}
      else if (nchar(start_praha22[i,j])==4)
        {vysledkove[i,j]<-substring(start_praha22[i,j],4,4)}
      else if (nchar(start_praha22[i,j])==3)

```

```

    {vysledkove[i,j]<-(substring(start_praha22[i,j],3,3))}
  }
}
#tímhle jsem si vyrobil výsledky z jednotlivých kol
for (i in 1:nrow(start_praha22)){
  for (j in 7:14){
    if ((vysledkove[i,j]==5)==TRUE) {vysledkove[i,j]<-0.5}
    if ((vysledkove[i,j]=="+")==TRUE) {vysledkove[i,j]<-1}
    if ((vysledkove[i,j]=="-")==TRUE) {vysledkove[i,j]<-0}
    if ((vysledkove[i,j]=="-1")==TRUE) {vysledkove[i,j]<-1}
  }
}
#pětky (ty musím opravit ještě v excelu, protože R mi dělá problém s 1/2)
#opravuji na 0.5
#+ je partie vyhraná kontumačně, tu přepisuji na 1
#- je partie prohraná kontumačně, tu přepisuji na 0
#-1 je výhra za porážku "ducha", přepisuji na 1
cil<-paste(misto,"vysledky_cisla_",navez,".xlsx",sep="")
jakodataframe<-as.data.frame(vysledkove)
#tady jsem vymazal udělení nového excelu

konecne_poradi<-matrix()
konecne_poradi<-cbind(vysledkove[,1],vysledkove[,17])

#další kus je přípravný k tomu, abych těžil informace
{ kod<-paste(misto,navez,".xlsx",sep="")
start_praha22<-as.data.frame(read_xlsx(kod))
hura<-matrix(nrow=nrow(vysledkove),ncol=2)
vysledkove<-start_praha22
for (i in 1:nrow(start_praha22)){
  if ((nchar(vysledkove[i,15])==1)==TRUE) {vysledkove[i,15]<-""}
  if(nchar(start_praha22[i,15])==5)
  {vysledkove[i,15]<-(substring(start_praha22[i,15],5,5))}
  else if (nchar(start_praha22[i,15])==4)
  {vysledkove[i,15]<-(substring(start_praha22[i,15],4,4))}
  else if (nchar(start_praha22[i,15])==3)
  {vysledkove[i,15]<-(substring(start_praha22[i,15],3,3))}
  if ((vysledkove[i,15]==5)==TRUE) {vysledkove[i,15]<-0.5}
  if ((vysledkove[i,15]==1)==TRUE | (vysledkove[i,15]==0.5)==TRUE |
      (vysledkove[i,15]==0)==TRUE)
  {hura[i,1]<-vysledkove[i,1]
  hura[i,2]<-vysledkove[i,15]}
}

odehrano<-as.numeric(na.omit(hura)[,1])
vysledky_posledni_kolo<-matrix(as.numeric(hura[odehrano,]),
  nrow=length(odehrano),ncol=2)

tipy_odehranych_her_posledniho_kola<-
  matrix(nrow=nrow(vysledky_posledni_kolo),ncol=Nsim)
g<-match(setdiff(pomocna_matice[,1],vysledky_posledni_kolo[,1]),
  pomocna_matice[,1])
vysledky_elita_1<-matrix(nrow=nrow(souperove),ncol=2)

for (t in 1:nrow(souperove)){
  vysledky_elita_1[,1]<-souperove[,1]
vysledky_elita_1[t,2]<-sum(vysledky_cisla_v_matici_pro_nekontumovane[t,1:8])
}
elitni<-which(vysledky_elita_1[,2]==max(vysledky_elita_1[,2])|
  vysledky_elita_1[,2]==max(vysledky_elita_1[,2])-0.5|
  vysledky_elita_1[,2]==max(vysledky_elita_1[,2])-1)

if(length(g)>0){for (r in 1:Nsim) {
  tipy_odehranych_her_posledniho_kola[,r]<-
  matice_tipu_posledniho_kola[-g,r]}
else
{tipy_odehranych_her_posledniho_kola[,r]<-
  matice_tipu_posledniho_kola[,r]}
}

#další už to dává do jednotlivých vektorů jako pro_korelace a podobně
{
  pro_korelace<-c()

```

```

pro_korelace_tipu<-c()
pro_korelace_elitni<-c()
pro_MSE<-c()
pro_MSE_tipu<-c()
pro_MSE_elitni<-c()

for (i in 1:Nsim)
{pro_korelace[i]<-cor(analyza[,i],konecne_poradi[,2])}

for (i in 1:Nsim)
{pro_korelace_tipu[i]<-
cor(vysledky_posledni_kolo[,2],tipy_odehranych_her_posledniho_kola[,i])}

for (i in 1:Nsim)
{pro_korelace_elitni[i]<-cor(analyza[elitni,i],konecne_poradi[elitni,2])}

for (i in 1:Nsim)
{pro_MSE[i]<-sum((konecne_poradi[,2]-analyza[,i])^2)/
length(konecne_poradi[,2])}

for (i in 1:Nsim)
{pro_MSE_tipu[i]<-sum((vysledky_posledni_kolo[,2]-
tipy_odehranych_her_posledniho_kola[,i])^2)/
length(vysledky_posledni_kolo[,2])}

for (i in 1:Nsim)
{pro_MSE_elitni[i]<-sum((konecne_poradi[elitni,2]-analyza[elitni,i])^2)/
length(elitni)}
}

#no a poslední krok, zařizují, abych měl všechny informace po kupě ze všech
#metod a všech typů posouzení, + navíc je tam nějaký ten export ven
{
informace<-rbind(pro_korelace,pro_korelace_tipu,pro_korelace_elitni,
pro_MSE,pro_MSE_tipu,pro_MSE_elitni)
informace_2<-cbind(c(mean(pro_korelace),mean(pro_korelace_tipu),
mean(pro_korelace_elitni),mean(pro_MSE),
mean(pro_MSE_tipu),mean(pro_MSE_elitni)),
c(sd(pro_korelace)*(length(pro_korelace)-1)/
length(pro_korelace),
sd(pro_korelace_tipu)*(length(pro_korelace_tipu)-1)/
length(pro_korelace_tipu),
sd(pro_korelace_elitni)*(length(pro_korelace_elitni)-1)/
length(pro_korelace_elitni),
sd(pro_MSE)*(length(pro_MSE)-1)/length(pro_MSE),
sd(pro_MSE_tipu)*(length(pro_MSE_tipu)-1)/
length(pro_MSE_tipu),
sd(pro_MSE_elitni)*(length(pro_MSE_elitni)-1)/
length(pro_MSE_elitni)))
informace_3<-cbind(informace_2,informace)
colnames(informace_3)<-c("průměr", "směrodatná odchylka",c(1:Nsim))
balik<-informace_3
cil_2<-paste(misto,"simulace_porovnani_",navez,"_",metoda,".xlsx",sep="")
jakodataframe_3<-as.data.frame(informace_3)
}
}

#a tohle už sdružuje výsledky do jedné ohromné matice, co se jmenuje
#mega_informace, jsem pořád uvnitř for cyklu od postupu
{ mega_informace[6*q-5,]<-balik[1,]
mega_informace[6*q-4,]<-balik[2,]
mega_informace[6*q-3,]<-balik[3,]
mega_informace[6*q-2,]<-balik[4,]
mega_informace[6*q-1,]<-balik[5,]
mega_informace[6*q,]<-balik[6,]}

#tady upravuji hotovou matici mega_informace
{ colnames(mega_informace)<-c("průměr", "směrodatná odchylka",c(1:Nsim))
radky<-c()
postupove<-c("MSE pořadí elitní", "korelace pořadí", "korelace jedn. výsledků",
"Korelace elitních pořadí", "MSE pořadí", "MSE jednotlivých tipů")
for(f in 1:nrow(mega_informace)){
radky[f]<-c(paste("metoda číslo",postup[1+floor((f-1)/6)],postupove[f%6+1]))
}
rownames(mega_informace)<-radky}

```

```

#kód, která mě připravuje na přítomnost grafů (porovnávání)
{nasobky<-matrix(nrow=6,ncol=length(postup))
  for(k in 1:length(postup)){
    for(t in 1:6){
      nasobky[t,k]<-t+(k-1)*6
    }
  }
}

#pokud dám srovnání==1 (defaultně 0 a tohle se nedělá), pak mi to vyplive
#do složky srovnání všech metod
#pokud dám srovnání==1 (defaultně 0 a tohle se nedělá), pak mi to vyplive
#do složky srovnání všech metod

#tady se zavedu série odstavců, který se použije jen s vylněnými předpoklady
{
  {if(length(predpoklady)>0) {if(predpoklady[1]==1)
  {napsat<-"vyhraje"} } else{napsat<-""} }
  {if(length(predpoklady)>0) {if(predpoklady[1]==0.5)
  {napsat<-"remizuje"} } else{napsat<-""} }
  {if(length(predpoklady)>0) {if(predpoklady[1]==0)
  {napsat<-"prohraje"} } else{napsat<-""} }
  {if(length(predpoklady)>2){dodat<-paste("\n a že bude splněno ",
    "dalších ",(length(predpoklady)-2)/2,
    " podmínek",sep="");}
  else{dodat<-""} }
  {if(length(cisla)>0){
  png(filename="hracovy_sance.png")
  hist(analyza[hrac,],breaks=(seq(min(analyza[hrac,])-1,
    max(analyza[hrac,])+0.5),
    main=paste("průměrné očekávané umístění je",
    sum(analyza[hrac,])/Nsim,xlab="umístění",ylab=paste
    ("v kolika případech"," by dopadl takto"))
  print(sum(analyza[hrac,])/Nsim)
  dev.off()
  }}
  {if(hrac>0&&length(predpoklady)==0){
  png(filename="hracovy_sance.png")
  hist(analyza[hrac,],breaks=(seq(min(analyza[hrac,])-1,
    max(analyza[hrac,])+0.5),main=NULL,
    xlab="pořadí hráče",ylab="kolikrát by takto dopadl")
  print(sum(analyza[hrac,])/Nsim)

  dev.off()
  }}
}

#jdu si hrát s prize money
if(ceny==1){
  {zisk<-matrix(ncol=Nsim,nrow=nrow(souperove))

  if(prize=="pevne"){
    for(j in 1:Nsim){
      for(i in 1:nrow(souperove)){
        if(analyza[i,j]>length(money)){zisk[i,j]<-0}
        else{zisk[i,j]<-money[analyza[i,j]]}
      }}
    }

  if(prize=="hort"){
    for(j in 1:Nsim){
      for(i in 1:nrow(souperove)){
        deleni<-c()
        for(x in 1:19)
          { deleni[x]<-length(bodiky[,j][bodiky[,j]==
            (max(bodiky[,j])-0.5*(x-1)))}

        if(analyza[i,j]>length(money)){zisk[i,j]<-0}
        else{

```

```

if(sum(deleni[1:2]*(max(bodiky[,j])-bodiky[i,j]+0.5))<=
length(money))
{a1<-money[analiza[i,j]]/2
a2<-2*(max(bodiky[,j])-bodiky[i,j]+0.5)
a3<-sum(deleni[1:a2-1])+1
a4<-sum(deleni[1:a2])
a5<-(deleni[a2])
zisk[i,j]<-a1+0.5*sum(money[a3:a4])/a5}
else{zisk[i,j]<-0}
}}
for(j in 1:Nsim){
for(i in 1:nrow(souperove)){
if(is.na(zisk[i,j])==TRUE){zisk[i,j]<-0}
}
}
palma<-zisk
for(j in 1:Nsim){
if(sum(zisk[,j]!=length(money)))

{banka<-sum(money)-sum(zisk[,j])

rrr<-length(money)-sum(zisk[,j]!=0)
h<-1
while(h<=rrr){ {
palma[which(analyza[,j]==
analyza[which(analyza[,j]==
sum(zisk[,j]!=0)+h),j]),j]<-
money[sum(zisk[,j]!=0)+h]/2+0.5*banka/(length(money)-
sum(zisk[,j]!=0))
h<-h+1}
}
}
}

}}

#jdu si hrát s prize money - grafy

if(hrac>0 & ceny==1)
{png(filename=paste("hracuv_vydelek",predpoklady[1],".png"))
hist(palma[hrac,],
main=paste("průměrný očekávaný výsledek je ",sum(palma[hrac,])/Nsim),
xlab="výdělek",ylab=NULL)
cilk<-paste(misto,"finance_",navez,".xlsx",sep="")
gol<-as.data.frame(palma)
write_xlsx(gol,path=cilk)

dev.off()}

}

```