Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra matematiky

Diplomová práce

Arbitrážní příležitosti a arbitrážní rozbor při sázení na sportovní výsledky

Plzeň, 2020 Bc. Andrei Sinkonen

ZÁPADOČESKÁ UNIVERZITA V PLZNI
Fakulta aplikovaných věd
Akademický rok: 2019/2020

ZADÁNÍ DIPLOMOVÉ PRÁCE
(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení:	**Bc. Andrei SINKONEN**
Osobní číslo:	**A19N0119P**
Studijní program:	**N0541A170005 Matematika a finanční studia**
Studijní obor:	**Matematika a finanční studia**
Téma práce:	**Arbitrážní příležitosti a arbitrážní rozbor při sázení na sportovní výsledky**
Zadávající katedra:	**Katedra matematiky**

Zásady pro vypracování

1. Stručně a jasně popište realitu sázení na výsledky (na úrovni: vyhraje A, vyhraje B, případně remíza) sportovních zápasů (max. jedna strana textu).
2. Stručně a jasně popište pojem „Arbitrážní příležitost" při takovém sázení (max. jedna strana).
3. Formulujte a dokažte nutné a postačující podmínky pro výskyt „arbitrážní příležitosti" při takovém sázení.
4. Na základě „arbitrážního skóre" otevíracích kurzů navrhněte metodiku vyhodnocení toho, jak si je sázková kancelář (sázkové kanceláře) jistá (jisté) s výsledkem. Exaktně formulujte pojem „jistota sázkové kanceláře s výsledkem". Pokuste se zahrnout vliv konkurence mezi sázkovými kancelářemi.
5. Sestavte pravděpodobnostní model výskytu arbitrážní příležitosti a navrhněte metodiku statistického odhadu jeho parametrů. Případně i zde analyzujte vliv konkurenčního prostředí mezi sázkovými kancelářemi.
6. Pokuste se vysvětlit některá (zdánlivě) paradoxní chování sázkových kanceláří.
7. Navržené postupy ověřte a komentujte na souborech s veřejně dostupnými daty.
8. K užitým datovým zdrojům doložte význam, v nich uvedených, informací. Datové zdroje podrobte kritice a eliminujte „nejvíce pravděpodobné" datové chyby.
9. Presentujete písemnou zprávou (s jedno-jednoznačnými odkazy do případných příloh). Výpočty a podklady doložte v přílohách na připojeném CD.

Rozsah diplomové práce: **cca 45 stran**
Rozsah grafických prací: **dle potřeby**
Forma zpracování diplomové práce: **tištěná**

Seznam doporučené literatury:

- Marek Patrice, Vávra František - Sports Betting Arbitrage in English Football. Proceedings, Aplimat 2019, Slovak University of Technology, Bratislava.
- František Vávra – Sázková (individuální) arbitráž. Pracovní text ZČU Plzeň, FAV/KMA, září 2018.
- František Vávra – Maržový model. ZČU Plzeň, FAV/KMA, srpen 2019.
- Jiří Jansa – Fair Bets in Sports Betting – IES FSV UK, diplomová práce FSV UK 2012.
- Egon Franck, Erwin Verbeek, Stephan Nüesch – Inter – market Arbitrage in Sports Betting.
- Brian Matejek – A Computational Analysis of Arbitrage Opportunities in Sports Gambling.
- Stanley Ong, Chris Kwek - A Sure Win Strategy for Football Betting.

Vedoucí diplomové práce: **Doc. Ing. František Vávra, CSc.**
Katedra matematiky

Datum zadání diplomové práce: **1. října 2019**
Termín odevzdání diplomové práce: **15. května 2020**

Radová

Západočeská univerzita v Plzni
L.S.
Fakulta aplikovaných věd

Doc. Dr. Ing. Vlasta Radová
děkanka

Doc. Ing. Marek Brandner, Ph.D.
vedoucí katedry

V Plzni dne 1. října 2019

Prohlášení

Prohlašuji, že jsem diplomovou práci na téma „Arbitrážní příležitosti a arbitrážní rozbor při sázení na sportovní výsledky“ vypracoval samostatně a výhradně s použitím uvedené literatury a pramenů.

V Plzni dne 5. června 2020

Bc. Andrei Sinkonen

Poděkování

Rád bych poděkoval panu doc. Ing. Františkovi Vávrovi, CSc., za odborné vedení diplomové práce, zájem, cenné rady a čas, který mi věnoval.

Abstract

The work is devoted to the study of the appearance of arbitrage opportunity in betting on sports results, in particular football. This analysis treats of betting reality and conditions of arbitrage opportunities. The work displays conditions for the bookmaker to choose the betting odds. The next step illustrates the methodology to assess how confident the bookmaker is in the outcome of the match. The analysis is followed by the representation of a probabilistic statistical model of the appearance of arbitrage opportunity. Further in the next is discussed the competition between bookmakers, the analyze of possible betting strategies and the description of some seemingly paradoxical actions of bookmakers.

Abstrakt

Práce se zabývá zkoumáním výskytů arbitrážních příležitostí při sázení na sportovní výsledky, konkrétně na fotbal. Rozebereme realitu sázení a podmínky výskytu arbitrážní příležitosti. Během práce ukážeme některé podmínky volby bookmakerem jednotlivých kurzů. V dalším kroku navrhneme metodiku odhadu toho, jak si je sázková kancelář jistá s výsledkem jednotlivého zápasu. Dále sestavíme pravděpodobnostně statistický model výskytu arbitrážní příležitosti. Provedeme analýzu konkurence mezi sázkovými kancelářemi, rozebereme možné strategie sázení a popíšeme některá zdánlivě paradoxní chování bookmakerů.

Obsah

Seznam obrázků

Seznam tabulek

Seznam použitých symbolů

2	k	Počet všech sázek složeného tiketu
3	K_i	Kurz sázení na i-tý výsledek
3	M	Počet všech možných výsledků zápasu
3	m_i	Poměr peněžních prostředků vsazených jednotlivým sázejícím na i-tý výsledek
4	E_0	Liga Premier League
4	E_1	Liga Championship
4	E_2	Liga League 1
4	E_3	Liga League 2
4	E_C	Liga Conference
4	D_1	Liga Bundesliga 1
4	D_2	Liga Bundesliga 2
4	I_1	Liga Serie A
4	I_2	Liga Serie B
4	SP_1	Liga La Liga Primera Division
4	SP_2	Liga La Liga Segunda Division
5	$B365$	Sázková kancelář Bet365
5	BW	Sázková kancelář Bet&Win
5	IW	Sázková kancelář Interwetten
5	LB	Sázková kancelář Ladbrokes
5	PS	Sázková kancelář Pinnacle
5	SJ	Sázková kancelář Stan James
5	VC	Sázková kancelář VC Bet
5	WH	Sázková kancelář William Hill
6	C_i	Peněžní prostředky vsazené na i-tý výsledek
6	r_i	Poměr vsazených prostředků na i-tý výsledek ku všem prostředkům vsazených na všechny možné výsledky daného zápasu
6	α_i	Individuální marže při uskutečnění i-tého výsledku
7	α	Reprezentativní marže
8	min	Nejnižší kurz
8	max	Nejvyšší kurz
8	d	Rozdíl mezi nejvyšším a nejnižším kurzem
9	$mean$	Průměr kurzů
9	a	„Arbitrážní“ skóre

9	o_i	Koeficienty individuálních marží
9	min_o	Minimální koeficient individuálních marží
9	max_o	Maximální koeficient individuálních marží
9	d_o	Rozdíl mezi maximálním a minimálním koeficientem individuálních marží
9	n	Celkový počet pozorování
9	$\widehat{x}$	Průměr ze všech dat x_i
10	r_{xy}	Výraz shodný s výrazem pro výběrový korelační koeficient
10	s_x	Výraz shodný s výrazem pro výběrovou směrodatnou odchylku
11	r_{xy}^2	Výraz shodný s výrazem pro výběrový koeficient determinace
11	$Q(r_{xy})$	Kvadratická chyba lineární aproximace dat metodou nejmenších čtverců
20	N	Počet všech dostupných sázkových kanceláří
20	$\widetilde{K}_i$	Nejvyšší možný kurz sázení na i-tý výsledek mezi všemi dostupnými sázkovými kancelářemi
20	$\widetilde{a}$	„Arbitrážní“ skóre vypočtený z nejvyšších dostupných kurzů
20	$N(t)$	Počet arbitrážních příležitostí, které se vyskytly od počátku sezóny zvolené ligy až do času t
20	$E(X)$	Střední hodnota náhodné proměnné X
21	$\sigma(X)$	Směrodatná odchylka náhodné proměnné X
21	$\sigma^2(X)$	Rozptyl náhodné proměnné X
21	$F(x)$	Distribuční funkce
21	T	Časy mezi detekováním dvou bezprostředně následujících arbitrážních příležitostí
20	τ	Parametr Poissonové náhodné proměnné $N(t)$
22	$\widetilde{\tau}$	Odhad metodou nejmenších čtverců parametru τ
22	$\widehat{\tau}$	Odhad parametru τ při splněné podmínce nulového součtu odchylek od odhadu aproximační přímky
22	$\widetilde{p}(t)$	Odhad aproximační přímky při $\widetilde{\tau}$
21	t_i	Čas měřený od počátku sezóny zvolené ligy, kdy byl detekován i-tý výskyt arbitrážní příležitosti
21	n	Počet všech výskytů arbitrážních příležitostí během sezóny zvolené ligy
23	l	Počet všech „využitelných“ arbitrážních příležitostí během sezóny zvolené ligy
22	$\widehat{p}(t)$	Odhad aproximační přímky při $\widehat{\tau}$
22	ε	Násobek σ-meze Čebyševovy nerovnosti 2. typu
37	SK	Sázková kancelář
38	H	Vítězství domácích

38	D	Remíza
38	A	Vítězství hostů
38	H_{SK}	Sázková kancelář tipovala vítězství domácích
38	D_{SK}	Sázková kancelář tipovala remízu
38	A_{SK}	Sázková kancelář tipovala vítězství hostů
39	$Q1$	První kvartil
39	$Q2$	Medián
39	$Q3$	Třetí kvartil
39	IQR	Kvartilové rozpětí

Úvod

Cílem této práce je prozkoumat výskyty arbitrážních příležitostí při sázení na sportovní výsledky. Během práce rozebereme realitu takového sázení. Popíšeme pojem „arbitrážní příležitost“ a následně zformulujeme a dokážeme nutné a postačující podmínky pro výskyt arbitrážní příležitosti při takovém sázení. Dále zformulujeme podmínky volby jednotlivých kurzů sázkovou kanceláří a na základě vypsaných kurzů navrhneme odhady bookmakerem zvolených/předpokládaných veličin: poměr vsazených prostředků, reprezentativní marže, individuální marže. Na základě „arbitrážního“ skóre navrhneme metodiku odhadu toho, jak si je bookmaker jistý s výsledkem jednotlivého zápasu. Dále sestavíme pravděpodobnostně statistický model výskytu arbitrážní příležitosti a následně navrhneme metodiku statistického odhadu parametrů takového modelu. V dalším kroku provedeme analýzu konkurence mezi sázkovými kancelářemi. Nakonec rozebereme některé možné strategie sázení a pokusíme se vysvětlit některá zdánlivě paradoxní chování sázkových kanceláří. Všechny výpočty provedeme v jazyce `R` prostřednictvím softwaru `RStudio`, který lze stáhnout např. z [1].

1 Základní informace

1.1 Sázení na výsledky sportovních zápasů

Sportovní sázení je setkání sázkové kanceláře, resp. *bookmakera*, a sázejícího s cílem uzavření sázky na sportovní výsledek. Základní sportovní sázka je sázka typu „$1-0-2$“, která umožňuje sázet na vítězství jednoho ze dvou soupeřů nebo na remízu. V případě, že remíza ve vybrané sportovní disciplině není možná, používá se sázka typu „$1-2$“. Ačkoliv se během této práce budou používat jenom výše uvedené typy sázek, je dobré zmínit o existenci dalších druhů sázení, například na:

- rozdíl bodových jednotek, neboli *asijský handicap*
- počet bodových jednotek, neboli *over/under*
- přesný výsledek
- vítězství na turnaji
- konkretní událost během zápasu
- ...

Dalším příkladem je *kumulované sázení*, které umožňuje zahrnout několik sázek do jednoho tiketu. Tiket s počtem sázek $k \geq 2$ se pak nazývá *složený*. Realizace všech k výsledků znamená pro sázecího výhru ve výši násobku vsazené částky a součinu kurzů složeného tiketu.

Cílem sázkové kanceláře je maximalizace zisku (při respektování poptávky sázejících). Jednou z výhod bookmakera je nekonstantnost kurzů. Jelikož sázková kancelář vypisuje kurzy několik dnů před zápasem, má možnost měnit hodnoty kurzů na základě poměru vsazených prostředků. Výplata se ale provádí v kurzu, ve kterém byla sázka uzavřena. Další výhodou sázkové kanceláře je možnost zvýšení zisku prostřednictvím rozdílu mezi *„spravedlivým“ (ze strany bookmakera) kurzem* ([2], str. 8-10), resp. kurzem odpovídajícím předpokládané bookmakerem pravděpodobnosti, a vypsaným kurzem.

1.2 Arbitrážní příležitost

Výhodou sázejícího je možnost výběru sázkové kanceláře, která nabídne nejvýhodnější kurz na trhu. Důsledkem je možnost využití odlišnosti kurzů jednotlivých bookmakerů pro dosažení garantovaného zisku na straně sázejícího, resp. zisku, který bude dosažen při jakémkoliv výsledku zápasu. Obchodní praktika, která využívá podobné rozdíly, se nazývá *arbitráž* ([3], str. 284).

Dále je vhodné uvést příklad, který demonstruje výskyt arbitrážní příležitosti při sázení na sportovní výsledky. Předpokládáme, že na jednotlivý zápas bez možnosti remízy bookmaker B_1 vypíše kurzy $1.600 - 2.250$ a bookmaker B_2 vypíše na stejný zápas kurzy $1.503 - 2.760$. Tedy nejvýhodnější kurzy pro sázejícího jsou $1.600 - 2.760$. Což znamená, že vsazení prostředků v poměru $63.303\ \% - 36.697\ \%$ přinese sázejícímu garantovaný zisk ve výši 0.0128 Kč ze vsazené 1 Kč.[1]

Rozdílnost kurzů je obvykle vyvolána konkurencí mezi sázkovými kancelářemi. Z toho plyne, že řešením pro sázejícího je porovnání kurzů, které vypisují bookmakeři z prostředí se silnou konkurencí. Dále je třeba zmínit, že k využití uvedené sázecí strategii je potřeba, aby sázející měl otevřené účty u různých sázkových kanceláří a stále převáděl mezi účty peněžní prostředky prostřednictvím vlastního zdrojového účtu.

1.3 Nutné a postačující podmínky pro výskyt arbitrážní příležitosti

Teorie pro popis nutné a postačující podmínky pro výskyt arbitrážní příležitosti je čerpána především z [4], [5].

Cílem sázejícího je rozdělit své peněžní prostředky mezi jednotlivými variantami výsledku zápasu tak, aby neprohrál bez ohledu na to, jak vybraný zápas skončí. Předpokládáme, že ze všech možných výsledků zápasu $M \geq 2$ může nastat právě jeden. Kurz sázení na i-tý výsledek označíme $K_i > 1$. Kurzy se mohou používat od různých sázkových kanceláří, ale platí, že pro každý z možných výsledků bude použit pravě jeden kurz.

Poměr peněžních prostředků vsazených na i-tý výsledek označíme $0 \leq m_i \leq 1$. Předpokládáme, že budou vsázeny všechny peněžní prostředky, což lze zapsat v následujícím tvaru:

$$\sum_{i=1}^{M} m_i = 1 \tag{1.1}$$

[1] Postup pro výpočet poměru vsazených prostředků je uveden v podkapitole 1.3.

Pokud pro všechna $i = 1, ..., M$ platí nerovnost $K_i m_i \geq 1$, je pak zaručeno, že sázející při vhodném rozdělení prostředků neprohraje při nastání jakéhokoliv z výsledků i. Převedením dané nerovnosti do tvaru $m_i \geq \frac{1}{K_i}$, která platí pro všechna $i = 1, ..., M$, a následným dosazením do již uvedené podmínky (1.1) získáváme *nutnou podmínku* pro zabezpečení neprohry při sázení, kterou zapíšeme následovně:

$$\sum_{i=1}^{M} \frac{1}{K_i} \leq 1 \tag{1.2}$$

Jelikož platí $\left(\sum_{i=1}^{M} \frac{1}{K_i}\right)^{-1} \geq 1$, pak bezrizikový poměr vsázených prostředků na i-tý výsledek pro $i = 1, ..., M$ lze například určit následovně:

$$m_i = \left(K_i \sum_{j=1}^{M} \frac{1}{K_j}\right)^{-1}, \quad i = 1, ..., M \tag{1.3}$$

Potom pro garantovaný výnos z takové sázky platí:

$$\min_{i=1,...,M} K_i m_i \geq 1 \tag{1.4}$$

Odtud plyne, že uvedená podmínka (1.2) je i *podmínkou postačující*.

Výnos při takové volbě m_i pak činí $K_i m_i = \left(\sum_{j=1}^{M} \frac{1}{K_j}\right)^{-1}$, přičemž platí, že zisk ze sázení je vždy stejný bez ohledu na výsledek zápasu. Uvedenou strategii sázení lze pojmenovat jako *maximální garance*. Samozřejmě existují i další strategie, o kterých bude zmíněno v podkapitole 2.5.

1.4 Použitá data

Označení	**Země**	**Název ligy**
E_0	England	Premier League
E_1	England	Championship
E_2	England	League 1
E_3	England	League 2
E_C	England	Conference
D_1	Germany	Bundesliga 1
D_2	Germany	Bundesliga 2
I_1	Italy	Serie A
I_2	Italy	Serie B
SP_1	Spain	La Liga Primera Division
SP_2	Spain	La Liga Segunda Division

Tabulka 1.1: Použité fotbalové ligy

Označení	**Sázková kancelář**
$B365$	Bet365
BW	Bet&Win
IW	Interwetten
LB	Ladbrokes
PS	Pinnacle
SJ	Stan James
VC	VC Bet
WH	William Hill

Tabulka 1.2: Použité sázkové kanceláře

Ke zpracování použijeme historická data otevírajících sázkových kurzů příslušných fotbalovým výsledkům ze sezóny 2014/2015 až do sezóny 2018/2019 [6]. Seznam použitých sázkových kanceláří je k dispozici v tabulce 1.2. Informace o použitých fotbalových ligách je uvedena v tabulce 1.1. Data jsou uložena ve složce „`data/`" podle sezóny a ligy ve formátu `*.csv`. Informace ke staženým datům lze dohledat v příloze A1 [7].

2 Zpracování

2.1 Maržový model

Postup tvorby maržového modelu je inspirován především [8].

Označíme $C_i \geq 0$ peněžní prostředky vsazené na i-tý výsledek, kde $i = 1, ..., M$. Potom $\sum_{i=1}^{M} C_i - C_j K_j$ je výnos sázkové kanceláře při uskutečnění j-tého výsledku, kde $j = 1, ..., M$. Označíme $0 \leq \alpha_j < 1$ požadovanou marži sázkové kanceláře při uskutečnění j-tého výsledku. Potom je pro bookmakera ideální, pokud platí $\sum_{i=1}^{M} C_i - C_j K_j \geq \alpha_j \sum_{i=1}^{M} C_i$. Úpravou dané nerovnosti dostaneme:

$$1 - \alpha_j \geq K_j r_j, \quad j = 1, ..., M \tag{2.1}$$

kde $r_j = C_j (\sum_{i=1}^{M} C_i)^{-1}$ je poměr vsazených prostředků na j-tý výsledek ku všem prostředkům vsazeným na všechny možné výsledky daného zápasu, proto platí:

$$\sum_{j=1}^{M} r_j = 1 \tag{2.2}$$

Předpokládá se, že pro $j = 1, ..., M$ platí $r_j > 0$, neboli na všech M variant výsledku zápasu je něco vsazeno. *Nutný důsledek* uskutečnění nerovnosti (2.1) lze pak zapsat následovně:

$$\sum_{j=1}^{M} \frac{1 - \alpha_j}{K_j} \geq 1 \tag{2.3}$$

Další úpravou nerovnosti (2.1) získáváme podmínku volby kurzů sázkovou kanceláří při známém r_j a daném α_j:

$$K_j \leq \frac{1 - \alpha_j}{r_j}, \quad j = 1, ..., M \tag{2.4}$$

Pokud zahrneme vliv konkurence mezi sázkovými kancelářemi, což odpovídá tomu, že sázková kancelář volí pro ni nejvyšší možný kurz při zvolených α_j a známých r_j, potom volbu kurzů sázkovou kanceláří zapíšeme následovně:

$$K_j = \frac{1 - \alpha_j}{r_j}, \quad j = 1, ..., M \tag{2.5}$$

Jelikož platí vztah (2.2), důsledkem uvedené rovnosti (2.5) je pak následující vztah:

$$\sum_{j=1}^{M} \frac{1 - \alpha_j}{K_j} = 1 \tag{2.6}$$

Jak již bylo uvedeno, pro $j = 1, ..., M$ platí $K_j > 1$. Předpokládá se prozatím, že hodnoty r_j pro $j = 1, ..., M$ jsou známé (nebo „dobře“ odhadnutelné). Odtud plyne omezení na volbu marže:

$$\alpha_j \leq 1 - r_j, \quad j = 1, ..., M \tag{2.7}$$

Předpokládáme, že sázková kancelář neplánuje v dané hře ztrátu, neboli $\alpha_j \geq 0$. Potom podmínku volby kurzů lze zapsat následovně:

$$K_j \leq \frac{1}{r_j}, \quad j = 1, ..., M \tag{2.8}$$

Jelikož platí vztah (2.2), důsledkem nerovnosti (2.8) je pak následující vztah:[1]

$$\sum_{j=1}^{M} \frac{1}{K_j} \geq 1 \tag{2.9}$$

Na základě nerovností (2.8) a předpokladu, že na všech M variant výsledku zápasu je něco vsazeno, získáváme omezení na r_j pro $j = 1, ..., M$:

$$0 < r_j \leq \frac{1}{K_j}, \quad j = 1, ..., M \tag{2.10}$$

Platí tedy, že splnění nerovnosti (2.9) je *nutnou podmínkou* pro existenci přípustného řešení vztahu (2.10). V realitě ale musí být splněna přísnější podmínka (2.3).

Pokud r_j pro $j = 1, ..., M$ volíme následovně:[2]

$$r_j = \left(K_j \sum_{i=1}^{M} \frac{1}{K_i} \right)^{-1}, \quad j = 1, ..., M \tag{2.11}$$

a zároveň platí vztah (2.9), platí i vztahy (2.2) a (2.10), což nám říká, že splnění nerovnosti (2.9) je i *podmínkou postačující* pro existenci přípustného řešení vztahu (2.10).

Předpokládejme, že sázková kancelář, která působí v konkurenčním prostředí, stanoví na všechny výsledky zápasu stejnou marži, neboli pro $j = 1, ..., M$ platí $\alpha_j = \alpha$. Potom ze vztahu (2.6) lze vyjádřit *reprezentativní marži* α:

$$\sum_{j=1}^{M} \frac{1 - \alpha}{K_j} = 1 \tag{2.12}$$

[1] Tedy arbitrážní příležitost (1.2) se nemůže vyskytnout jenom u jedné sázkové kanceláře. Pokud se to někdy v historických datech vyskytne, jedná se o datovou chybu.

[2] Jelikož platí vztah (2.9), potom volba r_j (2.11) splňuje podmínku (2.10).

$$\alpha = 1 - \frac{1}{\sum_{j=1}^{M} \frac{1}{K_j}} \tag{2.13}$$

Porovnání vztahů (2.6) a (2.12) nám říká, že reprezentativní marže α je kombinací *individuálních marží* α_j, kde $j = 1, ..., M$:

$$\begin{aligned}
\sum_{j=1}^{M} \frac{1-\alpha}{K_j} &= \sum_{j=1}^{M} \frac{1-\alpha_j}{K_j} \\
1 - \alpha &= \sum_{j=1}^{M} \left(\frac{\frac{1}{K_j}}{\sum_{i=1}^{M} \frac{1}{K_i}} \right) (1 - \alpha_j) \\
\alpha &= \sum_{j=1}^{M} \left(\frac{\frac{1}{K_j}}{\sum_{i=1}^{M} \frac{1}{K_i}} \right) \alpha_j = (1-\alpha) \sum_{j=1}^{M} \frac{\alpha_j}{K_j} \qquad (2.14)
\end{aligned}$$

2.2 Jistota sázkové kanceláře s výsledkem

Metodika vyhodnocení toho, jak si je sázková kancelář jistá s výsledkem je inspirována především [9].

Existuje několik způsobů měření tzv. *jistoty sázkové kanceláře s výsledkem zápasu*. Některé z možných ukazatelů jistoty:

1. Minimální kurz

   $$min = \min_{i=1,...,M} K_i$$

   Čim je nižší hodnota tohoto ukazatele, tím si je sázková kancelář jistější tím, že takový výsledek nastane.

2. Maximální kurz

   $$max = \max_{i=1,...,M} K_i$$

   Čim je vyšší hodnota tohoto ukazatele, tím si je sázková kancelář jistější tím, že takový výsledek nenastane.

3. Rozdíl mezi maximálním a minimálnim kurzem

   $$d = max - min$$

   Čím je vyšší hodnota tohoto ukazatele, tím si je sázková kancelář jistější s výsledkem zápasu.

Dále prozkoumáme několik dalších ukazatelů souvisejících s již zmíněnými, pomocí kterých bychom mohli odhadnout jistotu sázkové kanceláře s vypisovanými kurzy:

4. Průměr kurzů

$$mean = \frac{1}{M}\sum_{i=1}^{M} K_i$$

5. „Arbitrážní" skóre

$$a = \sum_{i=1}^{M} \frac{1}{K_i}$$

6. Koeficienty individuálních marží

$$o_i = \frac{1}{aK_i}, \quad i = 1, ..., M$$

7. Minimální koeficient individuálních marží

$$min_o = \min_{i=1,...,M} o_i$$

8. Maximální koeficient individuálních marží

$$max_o = \max_{i=1,...,M} o_i$$

9. Rozdíl mezi maximálním a minimálním koeficientem individuálních marží

$$d_o = max_o - min_o$$

Před analýzou souvislostí mezi ukazateli je potřeba vysvětlit význam zde užívaného pojmu *„korelace"* pro danou úlohu. Postup je inspirován především [10].

Koeficienty přímky $y = Ax + B$ prokládající data x_i, y_i, $i = 1, ..., n$ metodou nejmenších čtverců vypočteme následovně ([11], str. 372, vztahy (13.7.2) a (13.7.3)):

$$A = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2.15}$$

$$B = \bar{y} - A\bar{x} \tag{2.16}$$

kde $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ a $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$. Platí následující vztah:

$$\sum_{i=1}^{n}(x_i - \bar{x})\bar{y} = \bar{y}\sum_{i=1}^{n}(x_i - \bar{x}) = 0 \tag{2.17}$$

Potom:

$$A = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2.18}$$

Dále:

$$\begin{aligned}
A &= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \frac{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} \\
&= \frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}} \frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}} \qquad (2.19)
\end{aligned}$$

kde výraz

$$r_{xy} = \frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{2.20}$$

je shodný s výrazem pro výběrový korelační koeficient ([11], str. 182, vztah (10.5.4)), a to nehledě na to, zda jsou x_i, y_i, $i = 1, ..., n$ nezávislými stejně rozdělenými náhodnými výběry, či nejsou. Analogicky výrazy

$$s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2.21}$$

$$s_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2.22}$$

jsou shodné s výrazy pro výběrové směrodatné odchylky ([11], str. 153, vztah (10.2.9)). Potom výrazy pro koeficienty

$$A = r_{xy}\frac{s_y}{s_x} \tag{2.23}$$

$$B = \bar{y} - r_{xy}\frac{s_y}{s_x}\bar{x} \tag{2.24}$$

jsou shodné s výrazy pro koeficienty regresní přímky ([12], str. 684, vztahy (12) a (13)). Kvadratickou chybu lineární aproximace dat metodou nejmenších čtverců lze vyjádřit následovně:

$$\begin{aligned}
Q(r_{xy}) &= \sum_{i=1}^{n}(y_i - ax_i - b)^2 \\
&= \sum_{i=1}^{n}\left(y_i - r_{xy}\frac{s_y}{s_x}x_i - \left(\bar{y} - r_{xy}\frac{s_y}{s_x}\bar{x}\right)\right)^2 \\
&= \sum_{i=1}^{n}\left((y_i - \bar{y}) - r_{xy}\frac{s_y}{s_x}(x_i - \bar{x})\right)^2
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \left((y_i - \bar{y})^2 + r_{xy}^2 \frac{s_y^2}{s_x^2} (x_i - \bar{x})^2 - 2 r_{xy} \frac{s_y}{s_x} (x_i - \bar{x}) (y_i - \bar{y}) \right) \\
&= (n-1)\, s_y^2 + (n-1)\, r_{xy}^2 \frac{s_y^2}{s_x^2} s_x^2 - 2 r_{xy} \frac{s_y}{s_x} s_x s_y \,(n-1)\, r_{xy} \\
&= (n-1)\, s_y^2 \left(1 - r_{xy}^2\right) \qquad (2.25)
\end{aligned}
$$

Potom pro $r_{xy} = \pm 1$ je $Q(r_{xy}) = 0$ a pro $r_{xy} = 0$ je $Q(r_{xy}) = (n-1)\, s_y^2$. Tedy kvadratická chyba lineární aproximace dat má stejné vlastnosti jako kvadratická chyba lineární regrese ([13], str. 161-162). Proto

$$
\begin{aligned}
r_{xy} &= \frac{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2}} \\
&= \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \qquad (2.26)
\end{aligned}
$$

je *mírou „lineární závislosti“* mezi x a y v rámci intervalu $\langle -1, 1 \rangle$, analogicky jako standardní korelační koeficient, resp. jeho odhad [14]. Tedy hodnotu r_{xy} lze interpretovat následovně:

$r_{xy} = 1$	mezi x a y je kladná „lineární závislost“
$r_{xy} = -1$	mezi x a y je záporná „lineární závislost“
$r_{xy} = 0$	x a y jsou „nekorelované“

Obdobně platí, že druhá mocnina „korelačního koeficientu“ r_{xy}

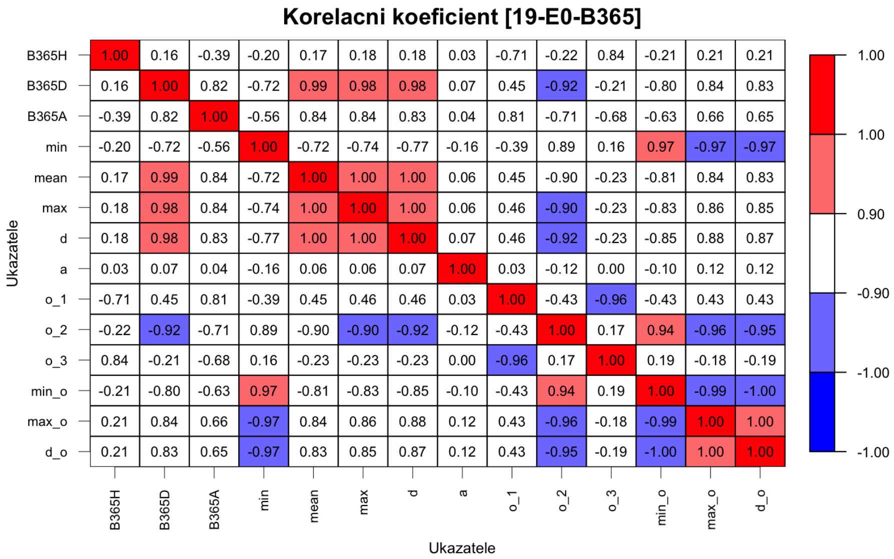
$$
r_{xy}^2 = \left(\frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \right)^2 \qquad (2.27)
$$

je *mírou „lineární závislosti“* mezi x a y v rámci intervalu $\langle 0, 1 \rangle$, analogicky jako koeficient determinace, resp. jeho odhad ([12], str. 682, vztah (5)). Tedy hodnotu r_{xy}^2 můžeme interpretovat následovně:

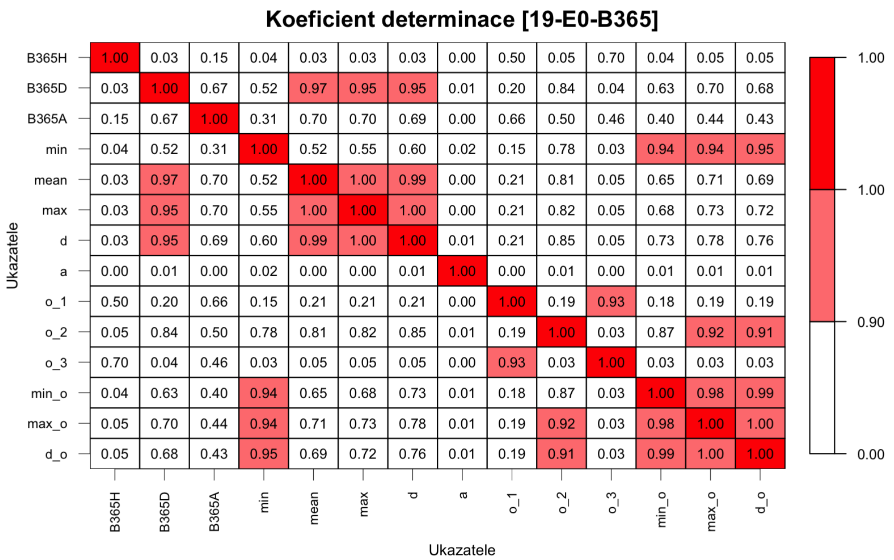
$r_{xy}^2 = 1$	mezi x a y je „lineární závislost“
$r_{xy}^2 = 0$	x a y jsou „nekorelované“

Analýzu souvislostí (lineárně modelovatelných) mezi ukazateli tedy lze provést pomocí „korelačního koeficientu“ r_{xy} a „koeficientu determinace“ r_{xy}^2.

Tabulky bodových odhadů těchto koeficientů pro dvojice zmíněných ukazatelů, vypočtených z dat ligy Premier League, sezóny 2018/2019, bookmakera Bet365, vidíme na obrázcích 2.1 a 2.2. Absolutní hodnoty v tabulkách převyšující 0.9 nám zvýrazňují nejsilnější „lineární závislost“. Některé dvojice ukazatelů s odlišnými koeficienty r_{xy}^2 jsou demonstrovány na obrázcích 2.3a až 2.3j.[3] Všechny obrázky jsou generovány pomocí přílohy A2. Vygenerované obrázky se automaticky uloží do příslušných složek „`img/cor/`“, „`img/r2/`“ a „`img/r2_priklad/`“ ve formátu `*.png`.

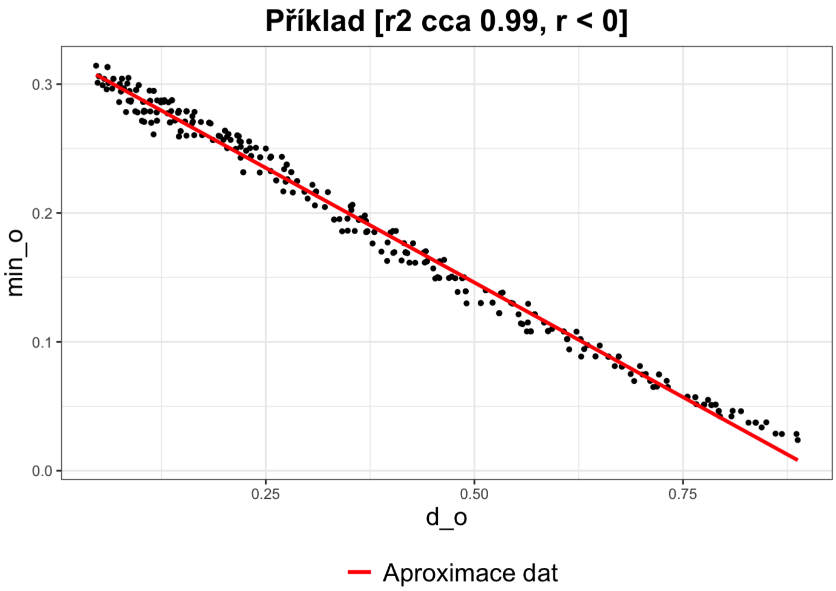


Obrázek 2.1: Tabulka bodových odhadů r_{xy}
[2018/2019 – Premier League – Bet365]

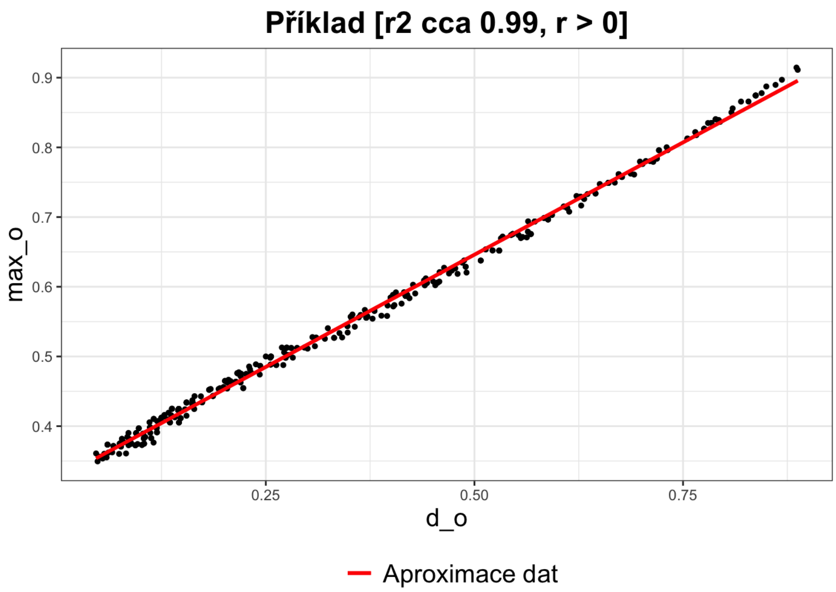


Obrázek 2.2: Tabulka bodových odhadů r_{xy}^2
[2018/2019 – Premier League – Bet365]

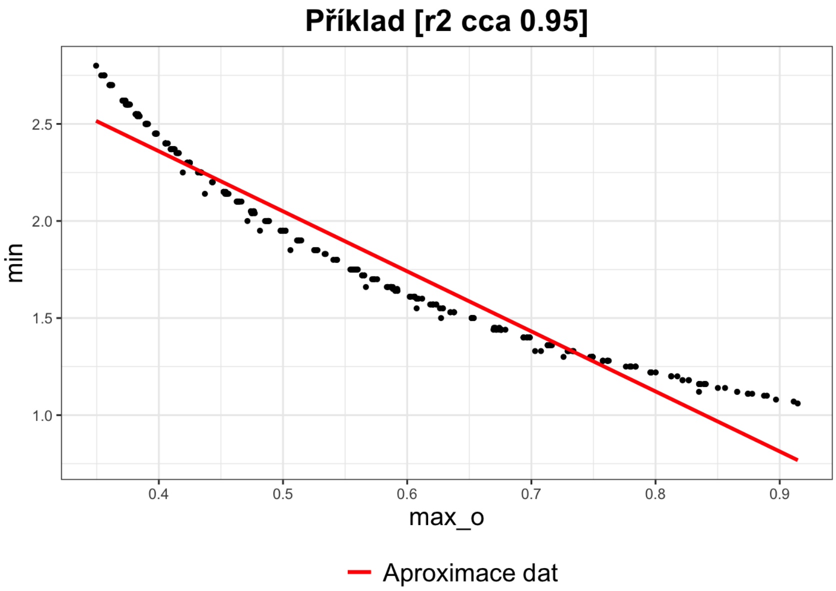
[3] Některé vazby (např. 2.3c, 2.3f, 2.3g, 2.3h, ..., 2.4a, 2.4b, 2.4c) nejsou lineární, ale vypadají jako silné nelineární vztahy, což je námět pro další zkoumání.



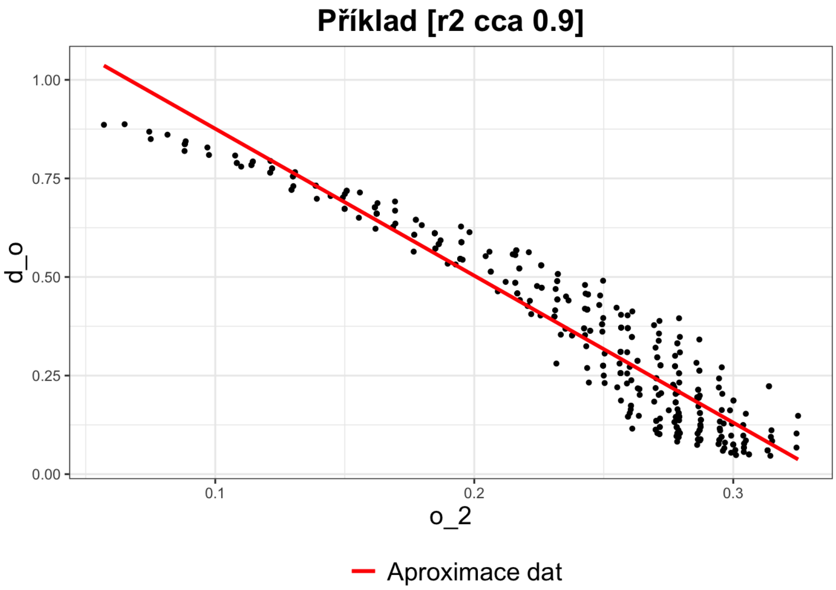
(a) $r_{xy}^2 \approx 0.99$, $r_{xy} < 0$



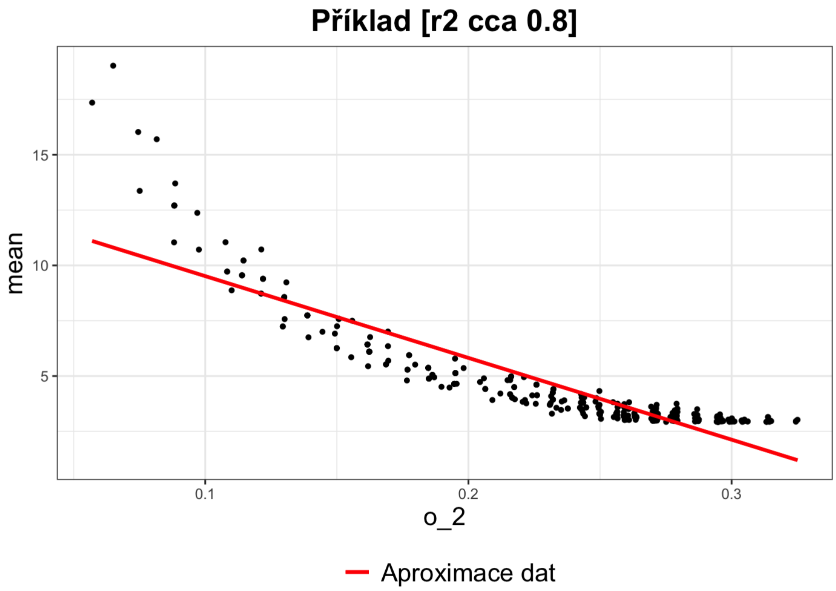
(b) $r_{xy}^2 \approx 0.99$, $r_{xy} > 0$



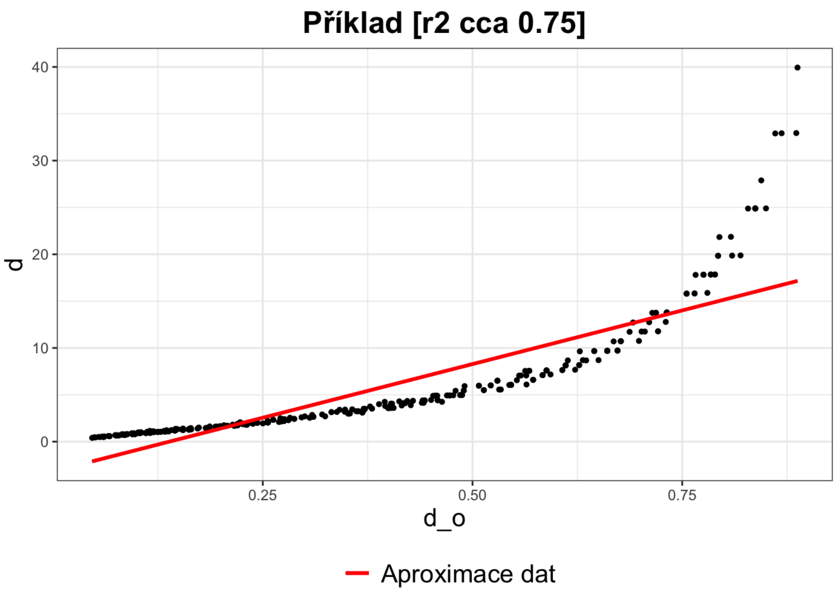
(c) $r_{xy}^2 \approx 0.95$



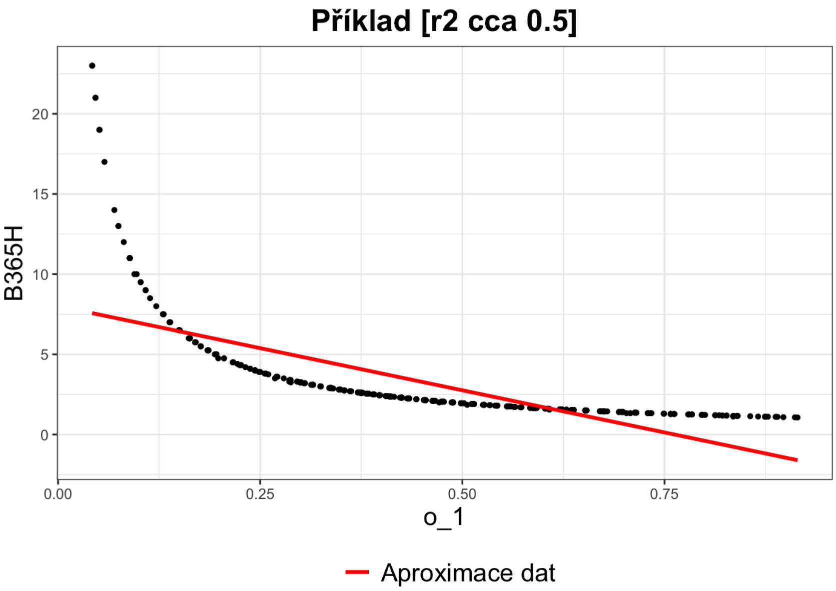
(d) $r_{xy}^2 \approx 0.9$



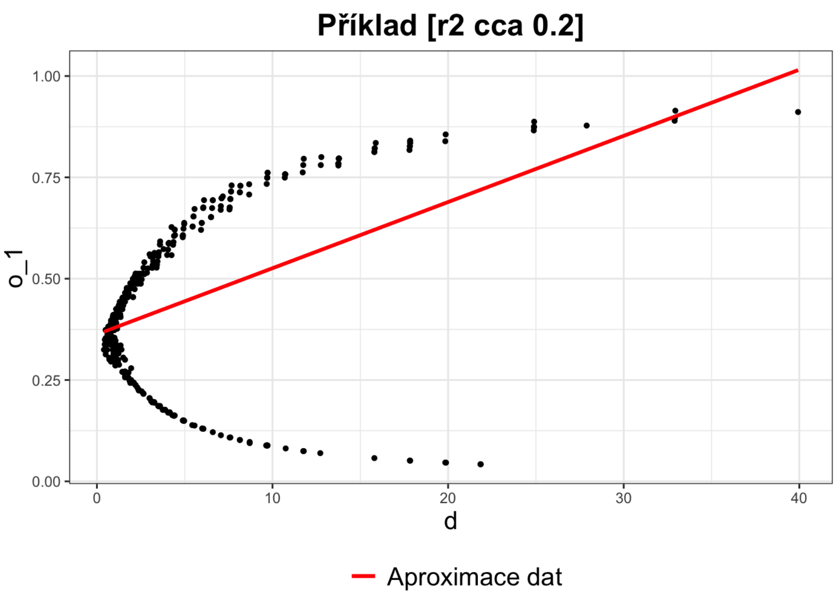
(e) $r_{xy}^2 \approx 0.8$



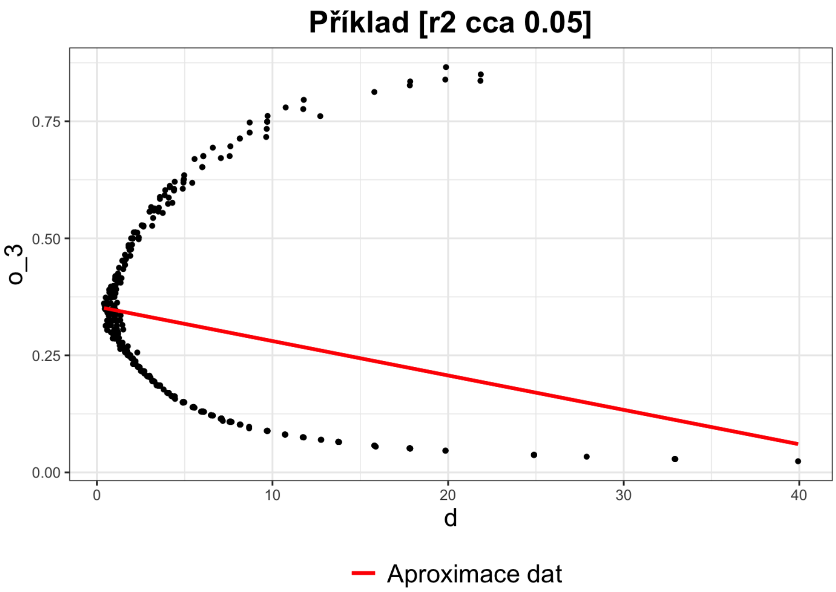
(f) $r_{xy}^2 \approx 0.75$



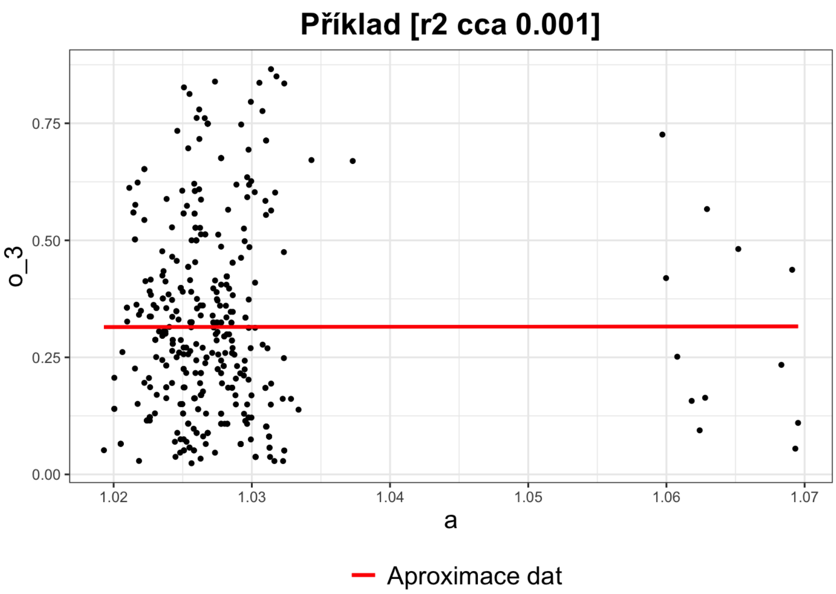
(g) $r_{xy}^2 \approx 0.5$



(h) $r_{xy}^2 \approx 0.2$



(i) $r_{xy}^2 \approx 0.05$

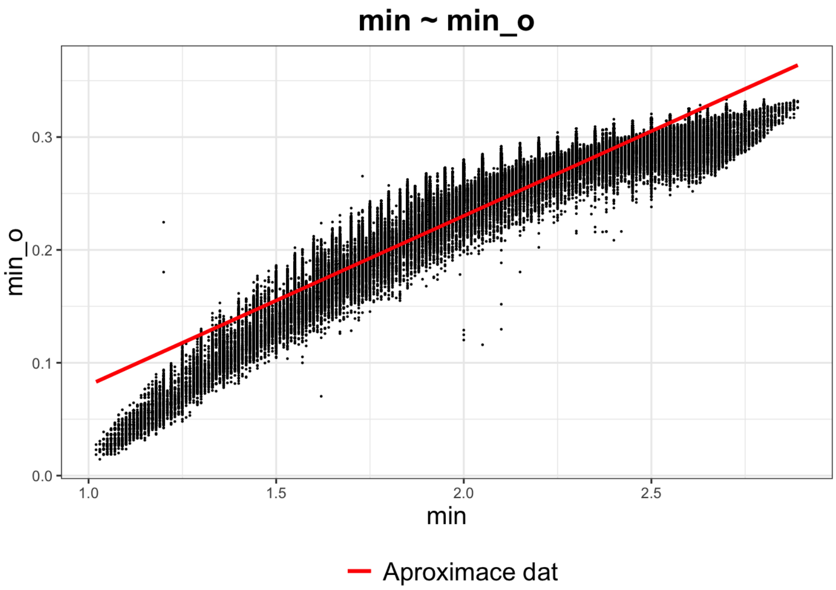


(j) $r_{xy}^2 \approx 0.001$

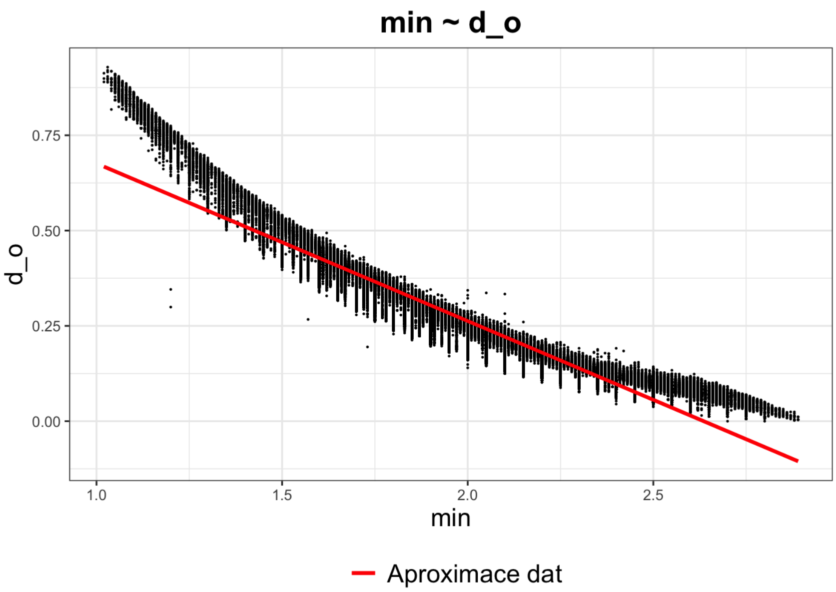
Obrázek 2.3: Některé dvojice ukazatelů s odlišnými koeficienty r_{xy}^2 [2018/2019 – Premier League – Bet365]

Z tabulek bodových odhadů „korelačních koeficientů" r_{xy} a „koeficientů determinace" r_{xy}^2 vidíme několik normovaných ukazatelů, které máji silnou „lineární závislost" s ukazatelem min:

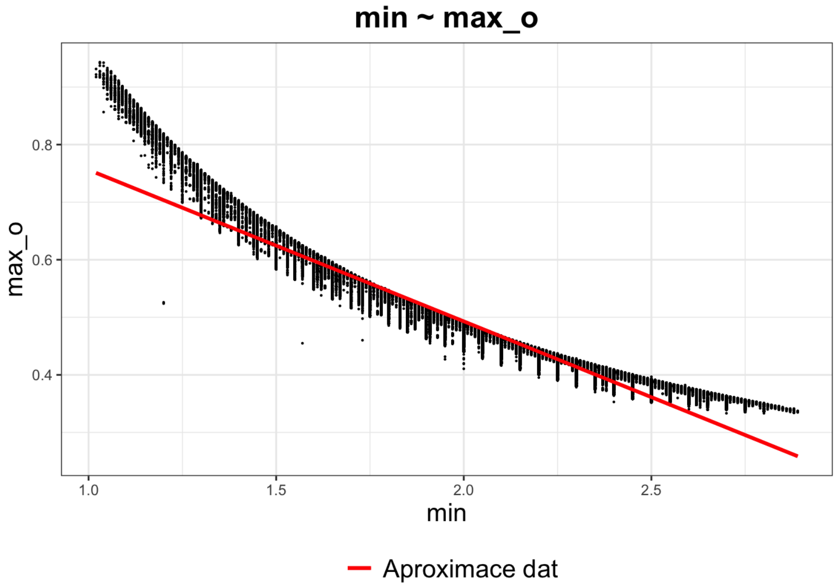
$min \sim min_o$	kladná „lineární závislost"
$min \sim max_o$	záporná „lineární závislost"
$min \sim d_o$	záporná „lineární závislost"



(a) $min \sim min_o$



(b) $min \sim d_o$



(c) $min \sim max_o$

Obrázek 2.4: Všechna pozorování vybraných dvojic ukazatelů

Tedy čím je vyšší hodnota ukazatelů max_o a d_0, resp. čím je nižší hodnota ukazatele min_o, tím je nižší hodnota min, což odpovídá vyšší jistotě sázkové kanceláře s výsledkem zápasu. Tak bychom pomocí ukazatelů min_o, max_o a d_o mohli odhadnout jistotu sázkové kanceláře v rámci intervalu $(0, 1)$. V dalším kroku prozkoumáme „lineární závislost" mezi těmito dvojicemi ukazatelů pro všechny možné kombinace lig, sezón a bookmakerů, což realizujeme pomocí přílohy A2. Celkový počet pozorování ze všech lig a všech jejích sezón činí $165\,317$. Výsledky jsou k dispozici v tabulce 2.1. Na obrázcích 2.4a až 2.4c vidíme všechna pozorování těchto dvojic ukazatelů a přímku prokládající data.[3]

	$r(min, Y)$			$r^2(min, Y)$		
Y	min_o	max_o	d_o	min_o	max_o	d_o
Minimum	0.8735	−0.9910	−0.9873	0.7630	0.9299	0.9168
1. kvartil	0.9245	−0.9828	−0.9747	0.8547	0.9480	0.9360
Medián	0.9466	−0.9802	−0.9704	0.8961	0.9608	0.9416
Průměr	0.9447	−0.9787	−0.9715	0.8931	0.9578	0.9438
3. kvartil	0.9670	−0.9737	−0.9675	0.9350	0.9659	0.9499
Maximum	0.9880	−0.9643	−0.9575	0.9761	0.9820	0.9747
Kvartilové rozpětí	0.0425	0.0091	0.0072	0.0803	0.0178	0.0139
Variační rozpětí	0.1145	0.0267	0.0297	0.2131	0.0522	0.0579
Celkem	0.9372	−0.9600	−0.9562	0.8783	0.9216	0.9144

Tabulka 2.1: Bodové odhady koeficientů r_{xy} a r_{xy}^2

Z tabulky 2.1 vidíme, že nejvyšší míra „lineární závislosti" je mezi ukazateli min

a max_o. Tedy max_o lze považovat za jeden z možných *normovaných ukazatelů míry jistoty sázkové kanceláře s výsledkem zápasu*. Čím je bližší tento ukazatel k 1, tím si je jistější bookmaker s vypsanými kurzy.

2.3 Pravděpodobnostní a statistický model výskytu arbitrážní příležitosti

Postup sestávení pravděpodobnostně statistického modelu výskytu arbitrážní příležitosti je inspirován především [16].

Nejvyšší možný kurz na výsledek $j = 1, ..., M$ mezi všemi dostupnými $N \geq 1$ sázkovými kancelářemi označíme následovně:

$$\widetilde{K}_j = \max_{i=1,...,N} K_{ij} \tag{2.28}$$

Potom „arbitrážní" skóre na daný zápas vypočtené z nejvyšších dostupných kurzů označíme $\widetilde{a} = \sum_{j=1}^{M} \frac{1}{\widetilde{K}_j}$. Víme, že pro zabezpečení neprohry sázejícího při sázení musí být splněna nutná a postačující podmínka (1.2), kterou lze zapsat následovně:

$$\widetilde{a} \leq 1 \tag{2.29}$$

Splněná podmínka (2.29) poskytuje sázejícímu možnost při vhodném rozdělení sázek na všechny možné výsledky jednotlivého zápasu vsadit tak, aby hráč s jistotou neprodělal. Takovému jevu říkáme *arbitrážní příležitost*. Víme, že sázková kancelář volí kurzy tak, aby byla splněna podmínka (2.3). Potom důsledkem splnění podmínky (2.29) je to, že všechny nejvyšší kurzy $\widetilde{K}_1, ..., \widetilde{K}_M$ nemohou být vypsány jednou sázkovou kanceláří najednou. V takovém případě by sázková kancelář s jistotou prodělala.

Označíme $N(t)$ počet arbitrážních příležitostí, které se vyskytly od počátku sezóny zvolené ligy až do času t. Potom lze pravděpodobnostní rozdělení $N(t)$ aproximovat „Poissonovým čítacím procesem" ([15], str. 59-60):

$$\begin{aligned} P(N(t) = k) &= e^{-\lambda t} \frac{(\lambda t)^k}{k!} \\ &= e^{-\frac{t}{\tau}} \frac{\left(\frac{t}{\tau}\right)^k}{k!}, \quad k = 0, 1, ... \end{aligned} \tag{2.30}$$

Tedy $N(t)$ je Poissonovou náhodnou proměnnou s parametrem $t\lambda \equiv \frac{t}{\tau}$. Potom střední hodnota, rozptyl a směrodatná odchylka $N(t)$ jsou ve tvaru:

$$E(N(t)) = \frac{t}{\tau} \tag{2.31}$$

$$\sigma^2(N(t)) = \frac{t}{\tau} \tag{2.32}$$

$$\sigma(N(t)) = \sqrt{\frac{t}{\tau}} \tag{2.33}$$

Několik vlastností plynoucích z vlastností „Poissonového čítacího procesu“:

- Počet současně detekovaných arbitrážních příležitostí nemůže být více než jedna. Doložíme to tím, že sázkové kanceláře nevyhlašují své kurzy na jednotlivé kolo v jediný okamžik.

- Časy T mezi detekováním dvou bezprostředně následujících arbitrážních příležitostí:

  – se řídí exponenciálním rozdělením pravděpodobnosti ([11], str. 130, vztah (8.4.4)). Tedy distribuční funkce je ve tvaru:

$$F_T(x) = P(T \leq x) = 1 - e^{-\frac{x}{\tau}} \tag{2.34}$$

  – jsou nezávislými náhodnými proměnnými.

Dále naznačíme způsob odhadu a ověření parametru τ. Předpokládáme, že máme k dispozici soubor dat v následujícím tvaru:

t	t_1	...	t_i	...	t_m
$N(t)$	1	...	i	...	m

kde

t_i je (resp. by měl být) čas (měřený od počátku sezóny zvolené ligy), kdy byl detekován i-tý výskyt arbitrážní příležitosti a

m je počet všech výskytů arbitrážních příležitostí během sezóny dané ligy.

Z toho plyne, že časy mezi výskyty arbitrážních příležitostí jsou kombinací dvou náhodných proměnných:

- Časy mezi vyhlášením kurzů a ukončením sázení na zápasy daného kola
- Časy mezi detekováním arbitrážních příležitostí mezi jednotlivými koly

Z veřejně dostupných historických souborů dat jsou k dipozici ale jenom datumy konání zápasů a otevírací kurzy[4]. Vyhlášení kurzů obvykle probíhá nějaký fixní čas před datem konání zápasu, proto by z daného pohledu vypsané datumy neměly být problémem. Ale z pohledu toho, že se často hraje více zápasů v jeden

[4]Někdy i zavírací, ale bez vývoje uvnitř intervalu mezi vyhlášením kurzů a ukončením sázení.

den, počet detekovaných arbitrážních příležitostí ve stejný čas může být víc než jedna, což by již mohlo činit problémem.

K alespoň částečné eliminaci uvedených datových „nedostatků“ využijeme metodu nejmenších čtverců ([11], str. 364-365):

$$\sum_{i=1}^{m}\left(i-\frac{t_i}{\widetilde{\tau}}\right)^2 \underset{\widetilde{\tau}}{\to} \min \tag{2.35}$$

Řešením dané minimalizační úlohy metodou nejmenších čtverců:

$$\widetilde{\tau}=\frac{\sum_{i=1}^{m} t_i^2}{\sum_{i=1}^{m} i t_i} \tag{2.36}$$

Požadujeme-li odhad $\widehat{\tau}$ parametru τ jako odhad „trendové“ přímky (2.31), vzniká problém s nestranností odhadu ([11], str. 187-188) přímky $\widetilde{p}$:

$$\widetilde{p}(t)=E(\widetilde{N}(t))=\frac{t}{\widetilde{\tau}} \tag{2.37}$$

Součet odchylek od takového odhadu aproximační přímky nemusí být nulový, proto se jeví jako nejlepší odhad vypočtený z následujícího vztahu:

$$\sum_{i=1}^{m}\left(i-\frac{t_i}{\widehat{\tau}}\right)=0 \tag{2.38}$$

Odtud vyjádříme $\widehat{\tau}$:

$$\begin{aligned}\widehat{\tau}&=\frac{\sum_{i=1}^{m} t_i}{\sum_{i=1}^{m} i}\\ &=2\frac{\sum_{i=1}^{m} t_i}{m(m+1)}\end{aligned} \tag{2.39}$$

Potom odhad přímky $\widehat{p}$:

$$\begin{aligned}\widehat{p}(t)&=E(\widehat{N}(t))\\ &=\frac{t}{\widehat{\tau}}\\ &=t\cdot\frac{m(m+1)}{2\sum_{i=1}^{m} t_i}\end{aligned} \tag{2.40}$$

Dále navrhneme míru přijatelnosti takového modelu. Čebyševova nerovnost 2. typu tvrdí, že pro libovolnou náhodnou veličinu X se střední hodnotou μ a rozptylem σ^2 platí ([17], str. 137, vztah (1)):

$$P(|X-\mu|\geq\varepsilon\sigma)\leq\frac{1}{\varepsilon^2} \tag{2.41}$$

Přepíšeme pro uvedený model:

$$\forall t > 0: \quad P\left(|N(t) - \frac{t}{\tau}| \geq \varepsilon\sqrt{\frac{t}{\tau}}\right) \leq \frac{1}{\varepsilon^2} \tag{2.42}$$

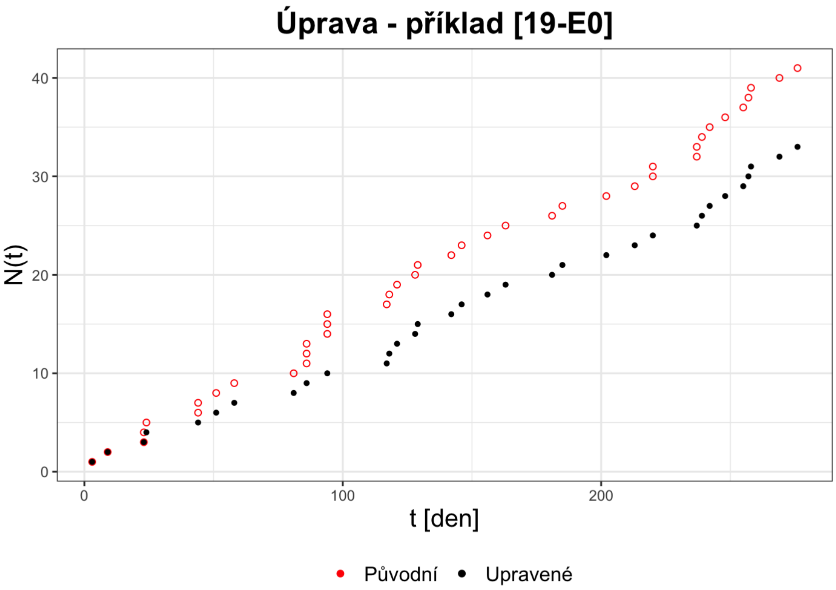
Potom jako *míru přijatelnosti* můžeme použit počet bodů ležící mimo „parabolické pásmo“:

$$\widehat{p}(t) \pm \varepsilon\sqrt{\frac{t}{\widehat{\tau}}}, \qquad \varepsilon > 0 \tag{2.43}$$

Nakonec upravíme data z pohledu toho, že v každém kole k danému datumu může sázející využít maximálně jednu arbitrážní příležitost. Tedy pokud se k danému datumu arbitrážních příležitostí vyskytne více než 1, sázející vybere tu nejlepší arbitrážní příležitost. Potom k sestavení pravděpodobnostně statistického modelu, resp. k odhadu parametru modelu τ, použijeme soubor dat v následujícím tvaru:

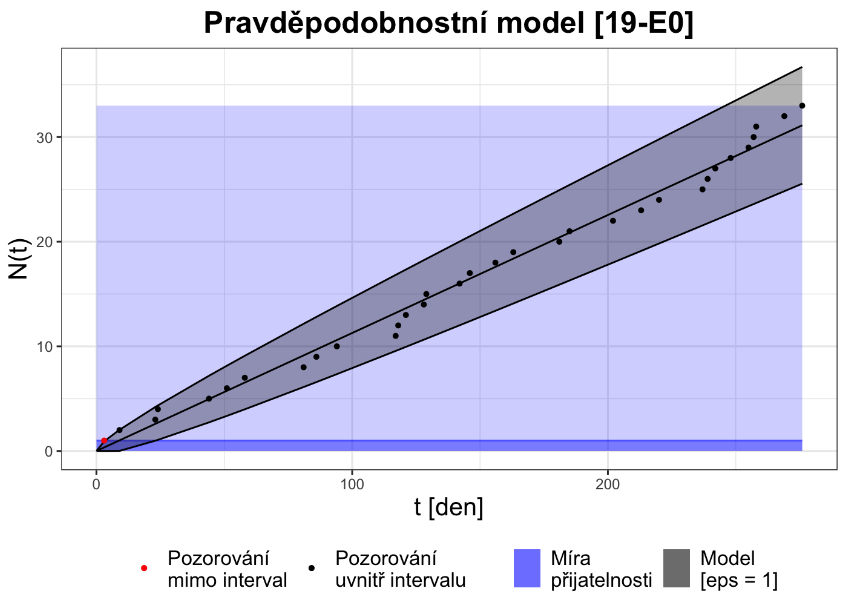
t	t_1	...	t_i	...	t_l
$N(t)$	1	...	i	...	l

kde l je počet všech *„využitelných“ arbitrážních příležitostí* během sezóny dané ligy. Na obrázku 2.5 vidíme příklad úpravy dat pro ligu Premier League sezóny 2018/2019.

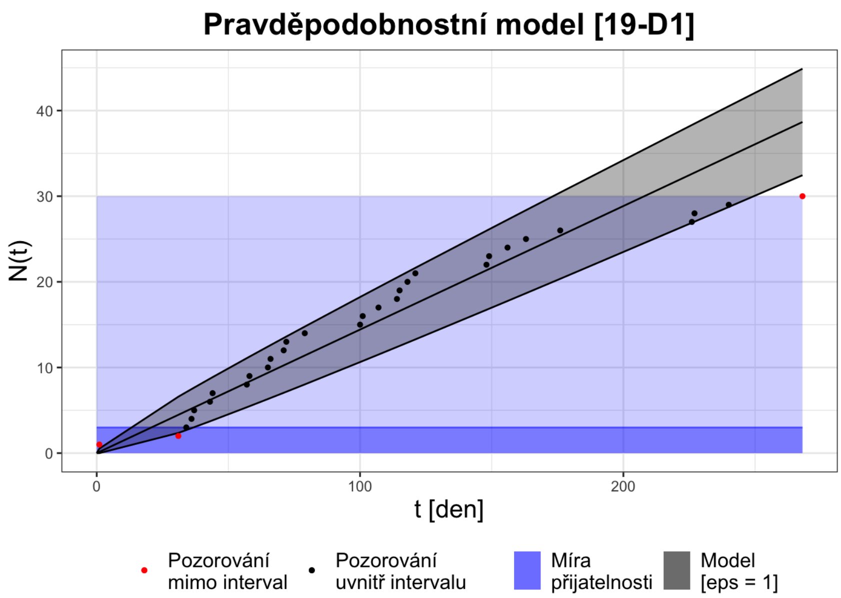


Obrázek 2.5: Příklad úpravy dat [2018/2019 – Premier League]

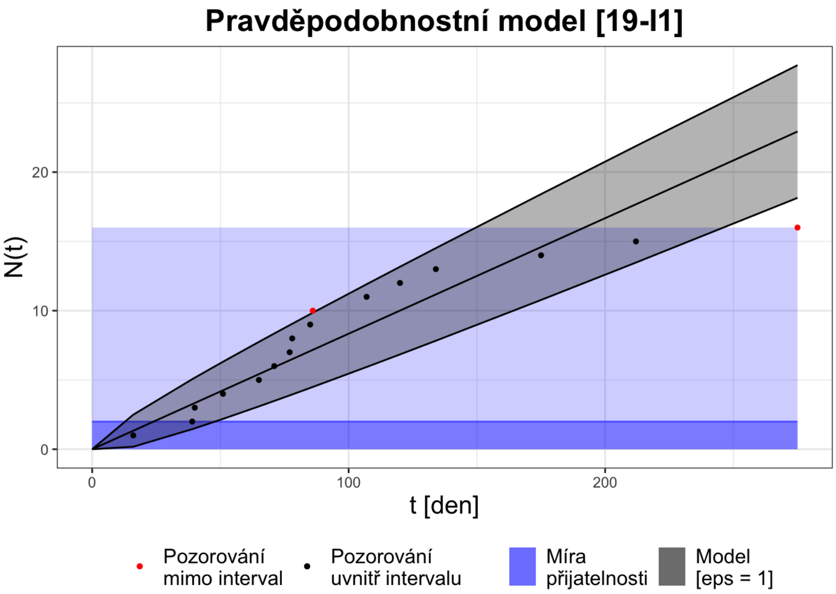
Výsledné grafy pravděpodobnostně statistických modelů pro nejvyšší fotbalové ligy sezóny 2018/2019 jsou k dispozici na obrázcích 2.6a až 2.6d. Grafy modelů pro nejvyšší fotbalové ligy předchozích sezón jsou k dispozici v příloze B1. Grafy jsou generovány pomocí přílohy A3. Vygenerované obrázky se automaticky uloží do složky „`img/model/`“ ve formátu `*.png`.



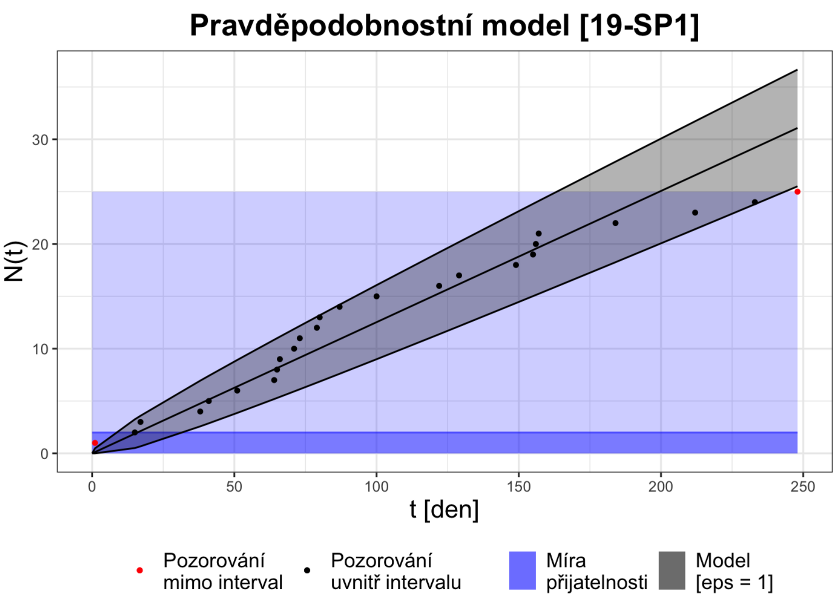
(a) 2018/2019 – Premier League



(b) 2018/2019 – Bundesliga 1



(c) 2018/2019 – Serie A

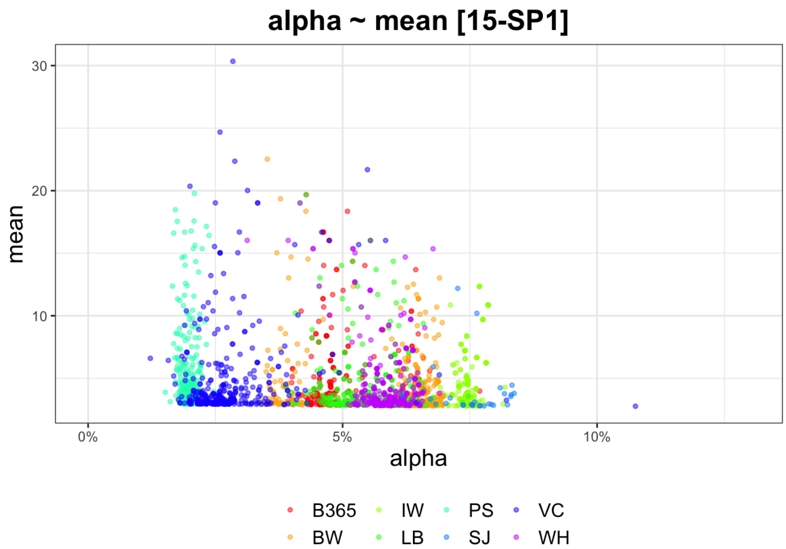


(d) 2018/2019 – La Liga Primera Division

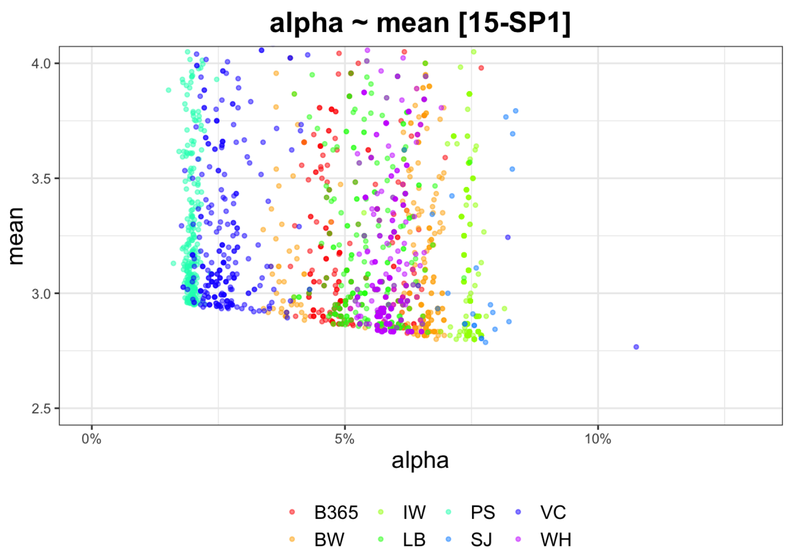
Obrázek 2.6: Pravděpodobnostně statistické modely nejvyšších fotbalových lig sezóny 2018/2019

2.4 Konkurenceschopnost

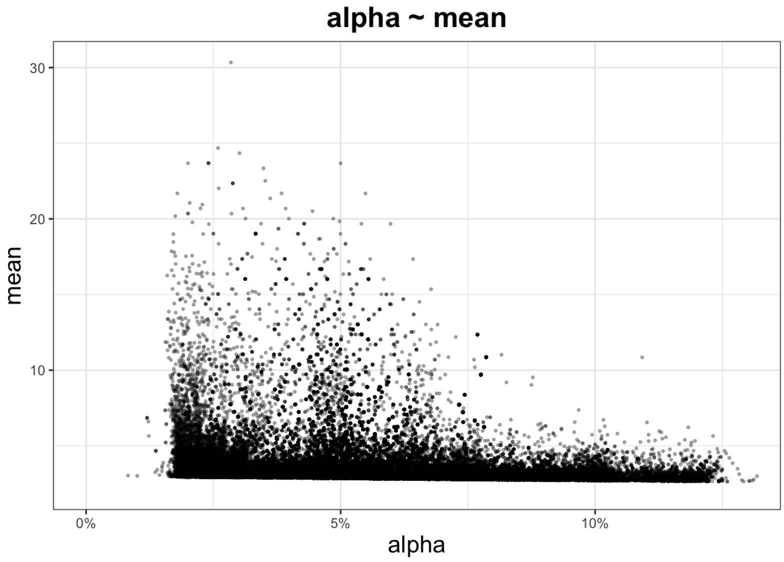
Sázkovou kancelář, která má nejvýhodnější, tj. nejvyšší kurzy na trhu, lze považovat jako nejvíce konkurenceschopnou. Cílem je tedy najít takový ukazatel, pomocí kterého lze odhadnout konkurenceschopnost sázkových kanceláří. Příkladem takové míry konkurenceschopnosti mohou být aritmetické průměry kurzů $mean$ jednotlivých zápasů, ale vzhledem k tomu, že kurzy se mohou pohybovat (teoreticky) v intervalu $(1, \infty)$, při porovnání bookmakerů v rámci několika zápasů taková míra může být zkreslená. Naším cílem je najít univerzálnější míru, pomocí které bychom mohli porovnat konkurenceschopnost kanceláří.



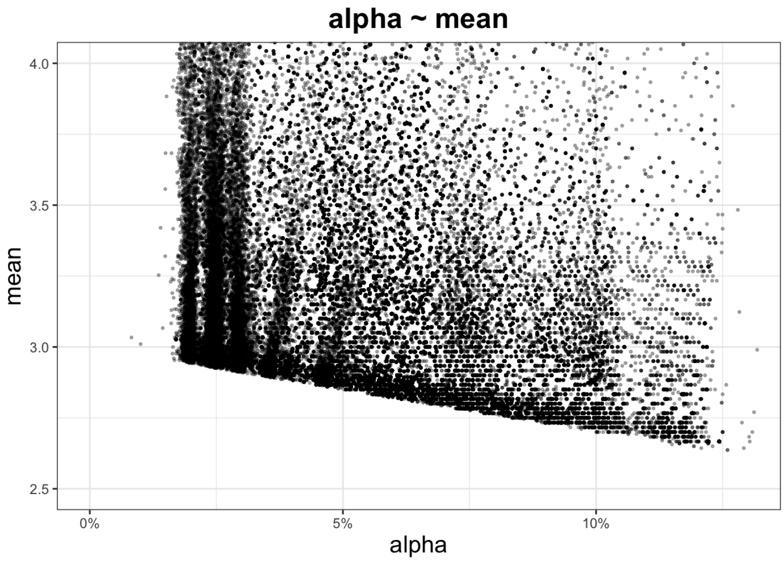
(a) 2014/2015 – La Liga Primera Division



(b) Přiblížená oblast grafu 2.7a



(c) Všechny zvolené ligy a sezóny



(d) Přiblížená oblast grafu 2.7c

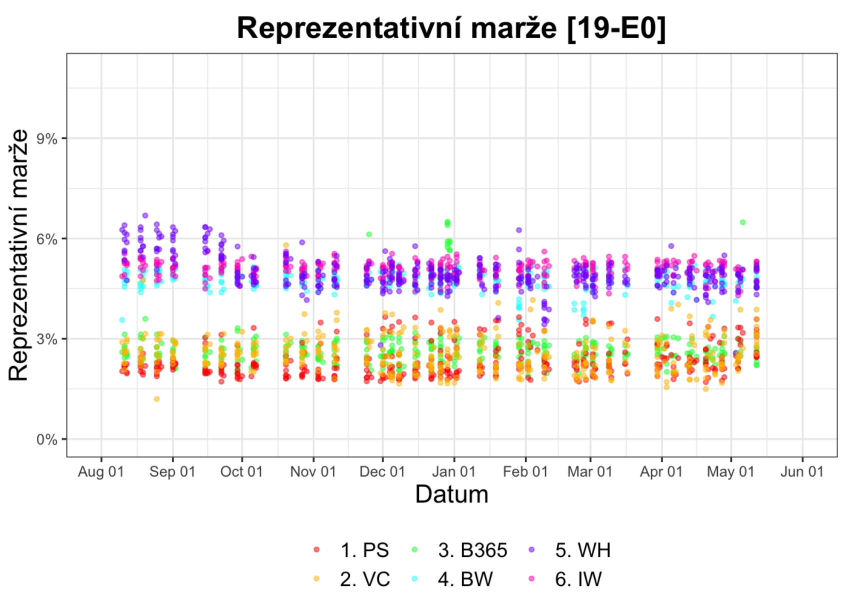
Obrázek 2.7: Závislost $\alpha \sim mean$

Víme, že odhad reprezentativní marže α (stejně jako $mean$) závisí na vypsaných kurzech, viz vztah (2.13). Dále prozkoumáme závislost těchto dvou ukazatelů α a $mean$. Na obrázku 2.7a vidíme bodový graf dané závislosti pro ligu La Liga Primera Division sezóny 2014/2015. Přiblížená oblast daného grafu je k dispozici na obrázku 2.7b. Z posledního obrázku je vidět, že tento graf má klesající tvar, tj. se zvětšující reprezentativní marží se snižují průměry kurzů. Dále ověříme přítomnost takové zákonitosti mezi ukazateli pro všechny zvolené kombinace lig, sezón a bookmakerů. Na obrázku 2.7c vidíme bodový graf závislosti těchto ukazatelů, jejichž hodnoty jsou vypočteny ze všech zvolených dat. Přiblížená oblast daného grafu je k dispozici na obrázku 2.7d. Z posledního obrázku vidíme, že zmíněná zákonitost platí i pro všechny zvolené sezóny a ligy najednou. Grafy jsou generovány pomocí přílohy A4. Vygenerované obrázky se automaticky uloží do složky „`img/alpha-mean/`“ ve formátu `*.png`.

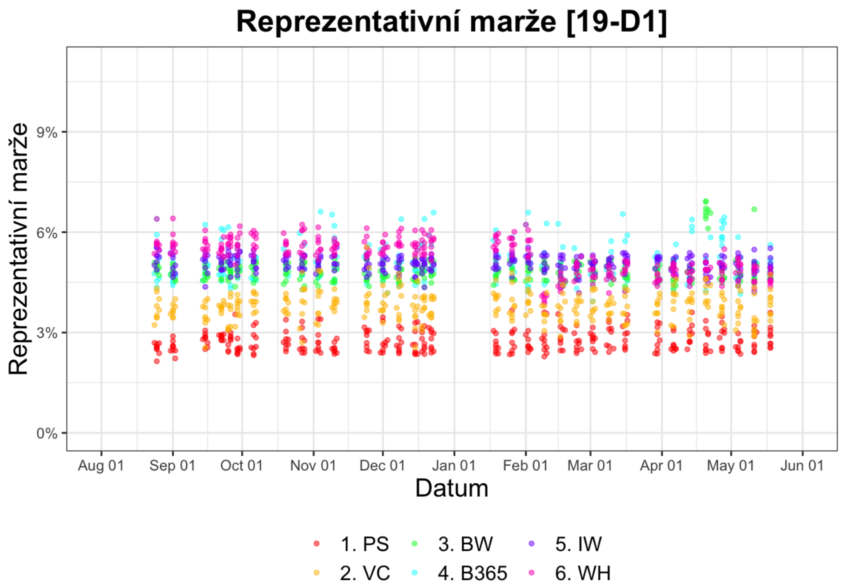
Potom jako *míru konkurenceschopnosti* lze používat reprezentativní marži, resp. její odhad z vypsaných kurzů. Tedy čím je nižší hodnota reprezentativní marže, tím jsou vyšší vypsané kurzy, což odpovídá vyšší konkurenceschopnosti dané sázkové kanceláře.

Na obrázcích 2.8a až 2.8d vidíme bodové grafy vývoje reprezentativní marže pro nejvyšší fotbalové ligy sezóny 2018/2019. Grafy vývoje reprezentativní marže pro dané fotbalové ligy předchozích sezón jsou k dispozici v příloze B2. Seřazení bookmakerů je provedeno vzestupně dle aritmetických průměrů reprezentativních marží.

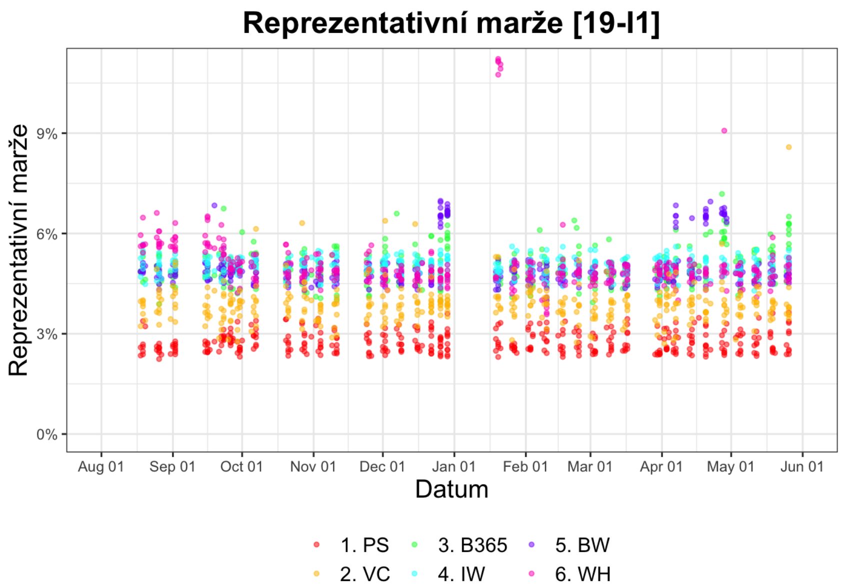
Z grafů vidíme, že na základě aritmetických průměrů reprezentativních marží všech zvolených bookmakerů v nejvyšších ligách je stabilně nejdominantnější sázková kancelář Pinnacle. Druhé místo na trhu pevně zaujímá kancelář VC Bet.



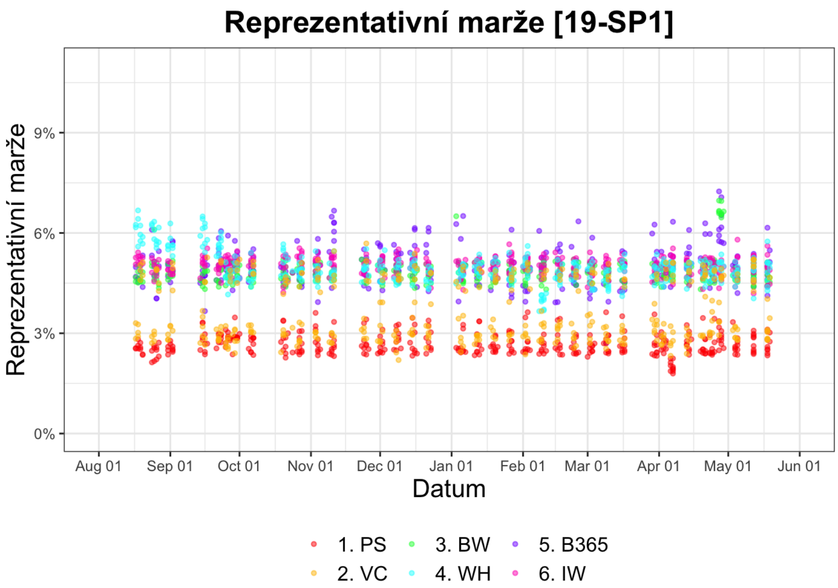
(a) 2018/2019 – Premier League



(b) 2018/2019 – Bundesliga 1



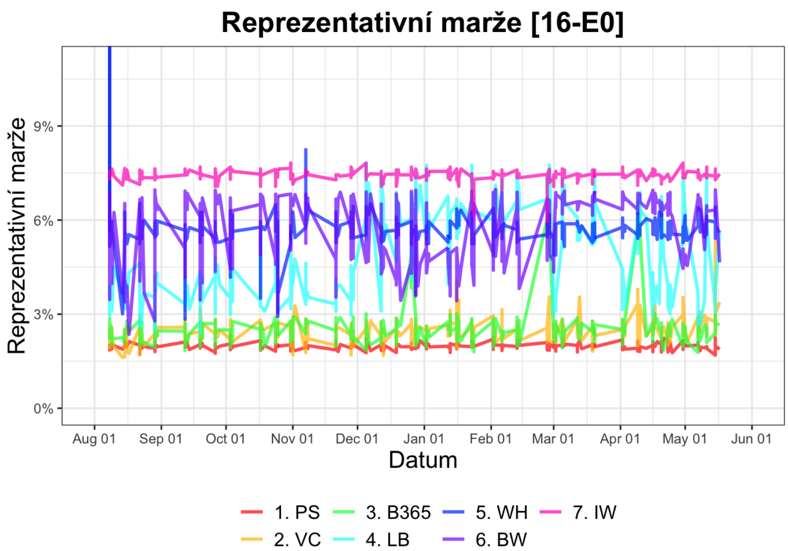
(c) 2018/2019 – Serie A



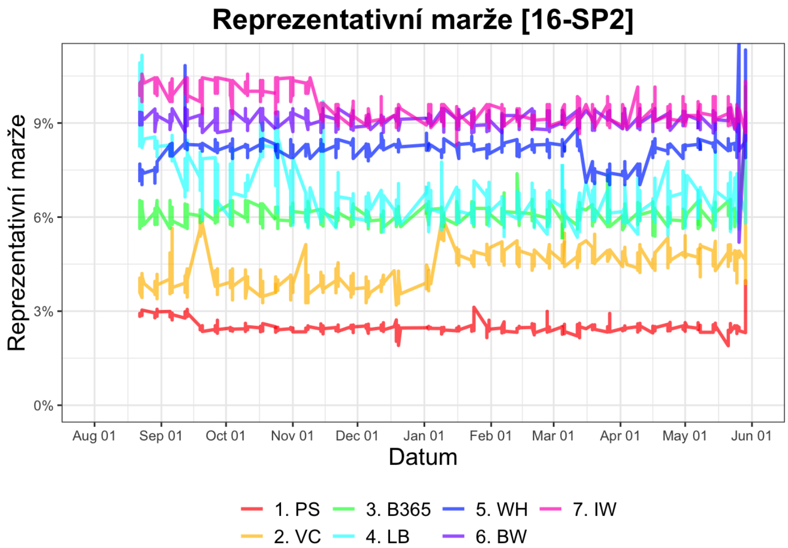
(d) 2018/2019 – La Liga Primera Division

Obrázek 2.8: Vývoj reprezentativní marže nejvyšších fotbalových lig sezóny 2018/2019

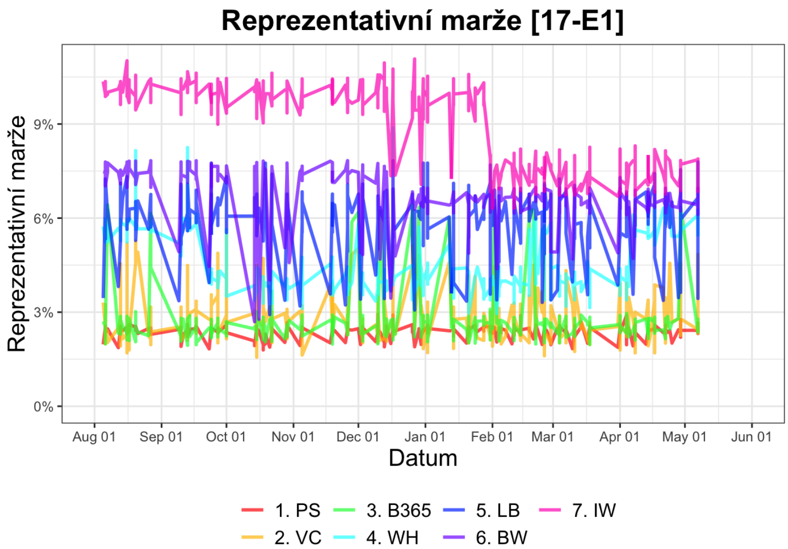
Vývoj reprezentativních marží zobrazíme pomocí čárových grafů, některé z nich jsou k dispozici na obrázcích 2.9a až 2.9d. Z těchto grafů je viditelná změna charakteru marží. Některé příklady změn maržových strategií jsou uvedeny v tabulce 2.2. Důvody změn charakteru marží navrhneme k možnému dalšímu zkoumání. Grafy vývoje reprezentativní marže jsou generovány pomocí přílohy A4. Vygenerované obrázky se automaticky uloží do složky „`img/marze/`“ ve formátu `*.png`.



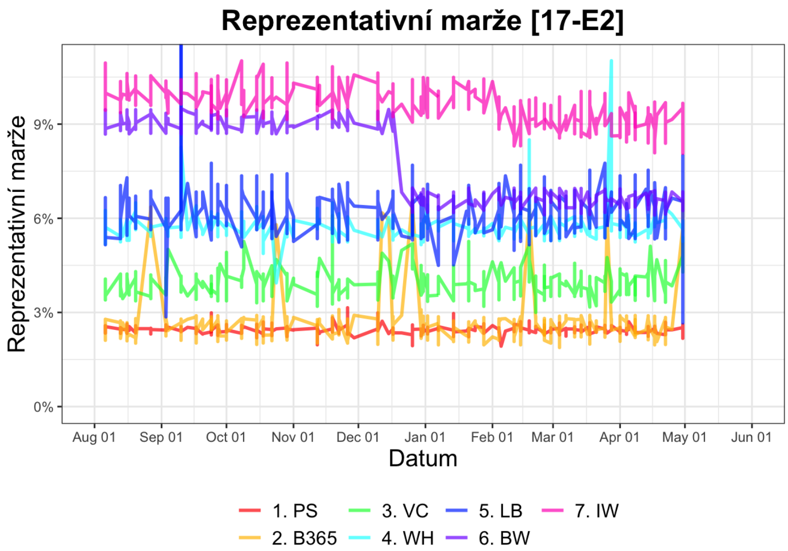
(a) 2015/2016 – Premier League



(b) 2015/2016 – La Liga Segunda Division



(c) 2016/2017 – Championship



(d) 2016/2017 – League 1

Obrázek 2.9: Příklady změn maržových strategií

Obrázek	Sezóna	Liga	Bookmaker	Období změny
2.9a	2015/2016	Premier League	Bet&Win	od začátku března
2.9b	2015/2016	La Liga Segunda Division	VC bet	od začátku ledna
2.9c	2016/2017	Championship	Interwetten	od začátku února
2.9d	2016/2017	League 1	Bet&Win	od 2. poloviny prosince

Tabulka 2.2: Příklady změn maržových strategií

2.5 Strategie sázení

V dané podkapitole budou rozebrány některé strategie sázení při výskytu arbitrážní příležitosti, konkrétně:

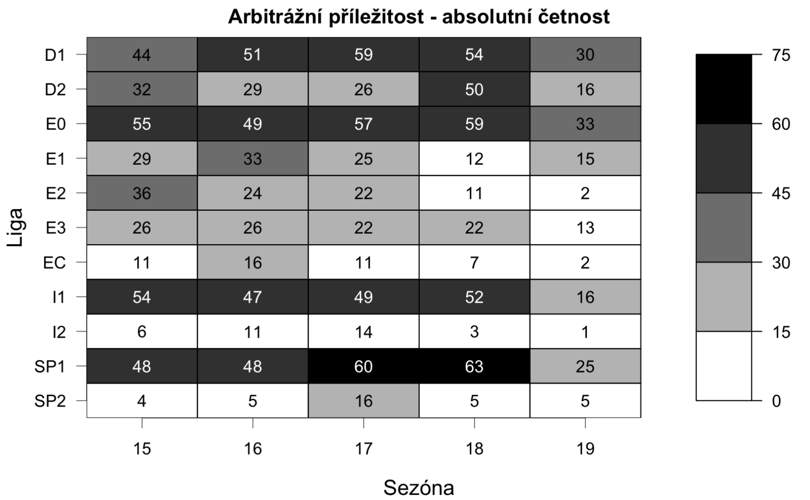
1. *Maximální garance*
2. *Garance + podíl na výsledku*
3. *Garance + outsider*
4. *Garance + favorit*

Teorie k popisu strategií sázení byla čerpána především z [4]. Všechny obrázky z dané podkapitoly jsou vygenerovány pomocí přílohy A5. Vygenerované obrázky se automaticky uloží do složky „`img/strategie/`“ ve formátu `*.png`.

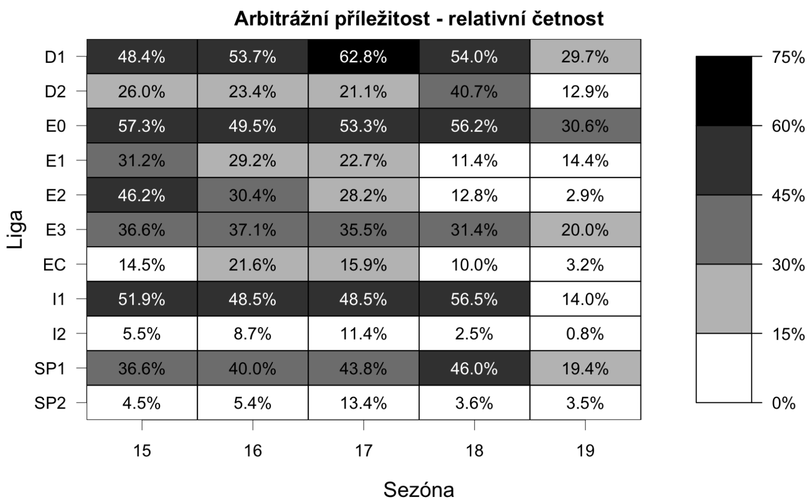
Tabulka absolutních četností „využitelných“ arbitrážních příležitostí je k dispozici na obrázku 2.10. Relativní četnosti získáme tím, že absolutní četnosti vydělíme počtem hracích dnů. Tabulka relativních četností „využitelných“ arbitrážních příležitostí je k dispozici na obrázku 2.11. Z daných tabulek vidíme, že největší počet arbitrážních příležitostí obvykle patří nejvyšším ligám:

E_0	Premier League
D_1	Bundesliga 1
I_1	Serie A
SP_1	La Liga Primera Division

Takový jev lze objasnit tím, že vyšší ligy jsou populárnější mezi sázejícími než ty nižší, což odpovídá vyšší konkurenci mezi sázkovými kancelářemi.



Obrázek 2.10: Tabulka absolutních četností „využitelných“ arbitrážních příležitostí



Obrázek 2.11: Tabulka relativních četností „využitelných“ arbitrážních příležitostí

2.5.1 Maximální garance

Při výskytu arbitrážní příležitosti strategie zajišťuje sázejícímu výnos bez ohledu na výsledek zápasu. Platí, že výnos ze sázení na danou hru při správném rozdělení prostředků je vždy stejný, nezávisle na tom, jaká z událostí nastala.

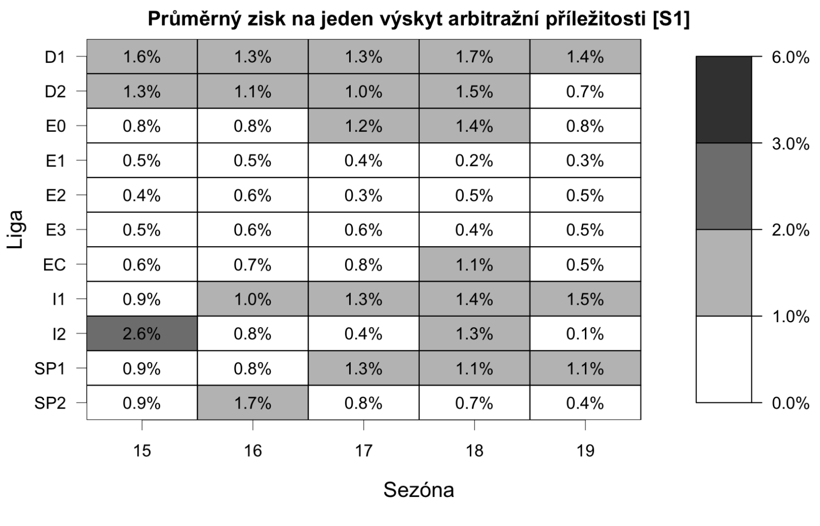
Volba bezrizikového poměru vsázených prostředků m_i na i-tý výsledek:

$$m_i = \left(K_i \sum_{j=1}^{M} \frac{1}{K_j} \right)^{-1}, \quad i = 1, ..., M \tag{2.44}$$

Výnos při takové volbě poměrů vsázených prostředků:

$$K_i m_i = \left(\sum_{j=1}^{M} \frac{1}{K_j} \right)^{-1}, \quad i = 1, ..., M \tag{2.45}$$

Tabulka průměrného zisku na jeden hrací den s výskytem (nejlepší k danému dnu) arbitrážní příležitosti pro danou strategii je k dispozici na obrázku 2.12. Pomocí této tabulky a tabulky absolutních četností „využitelných" arbitrážních příležitosti můžeme dopočítat celkový možný garantovaný zisk při sázení na jednotlivou ligu, a to s podmínkou, že vsazená částka na všechny zápasy bude stejná. Jinými slovy, zisk ze sázení nebude využit k následujícímu sázení. Například při sázení na zápasy ligy Premier League během sezóny 2018/2019 by z 1 Kč vsazené na počátku sezóny mohl mít sázející na konci sezóny garantovaný zisk ve výši $1 \cdot 0.008 \cdot 59 = 0.472$ Kč.



Obrázek 2.12: Tabulka průměrného zisku na hrací den s výskytem arbitrážní příležitosti pro strategii „Maximalní garance"

2.5.2 Garance + podíl na výsledku

Při výskytu arbitrážní příležitosti strategie zajišťuje sázejícímu výnos bez ohledu na výsledek zápasu. Platí, že výnos ze sázení na danou hru při správném rozdělení prostředků záleží na tom, jaká z událostí nastala. Tedy čím je vyšší kurz na jednotlivý výsledek zápasu, tím je větší možný výnos ze sázení.

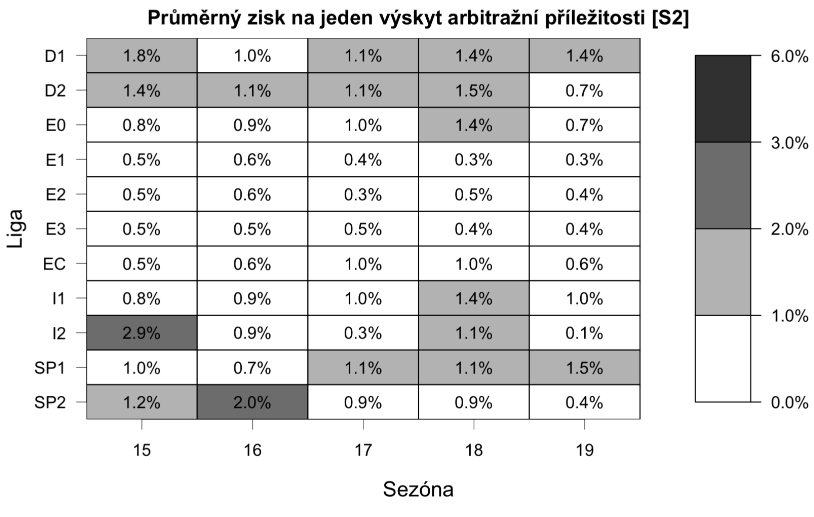
Volba bezrizikového poměru vsázených prostředků m_i na i-tý výsledek:

$$m_i = \frac{1}{K_i} + \frac{1}{M}\left(1 - \sum_{j=1}^{M} \frac{1}{K_j}\right), \quad i = 1, ..., M \tag{2.46}$$

Výnos při takové volbě poměrů vsázených prostředků:

$$K_i m_i = 1 + \frac{K_i}{M}\left(1 - \sum_{j=1}^{M} \frac{1}{K_j}\right), \quad i = 1, ..., M \tag{2.47}$$

Tabulka průměrného zisku na jeden hrací den s výskytem (nejlepší k danému dnu) arbitrážní příležitosti pro danou strategii je k dispozici na obrázku 2.13. Například za zmíněných výše podmínek, při sázení na zápasy ligy Premier League během sezóny 2018/2019, by z 1 Kč vsazené na počátku sezóny mohl mít sázející na konci sezóny garantovaný zisk ve výši $1 \cdot 0.007 \cdot 59 = 0.413$ Kč.



Obrázek 2.13: Tabulka průměrného zisku na hrací den s výskytem arbitrážní příležitosti pro strategii „Garance + podíl na výsledku“

2.5.3 Garance + outsider

Strategie zajišťuje sázejícímu největší možný výnos v případě výhry outsidera (s pohledu sázkové kanceláře) a (místo ztráty) nulový zisk při nastání jiné události.

Označíme K_i kurz na výhru outsidera, tj. nejvyšší možný kurz.[5] Volba bezrizikového poměru vsázených prostředků m_i na výhru outsidera:

$$m_i = \frac{1}{K_i} + \left(1 - \sum_{j=1}^{M} \frac{1}{K_j}\right), \quad i = 1, ..., M \tag{2.48}$$

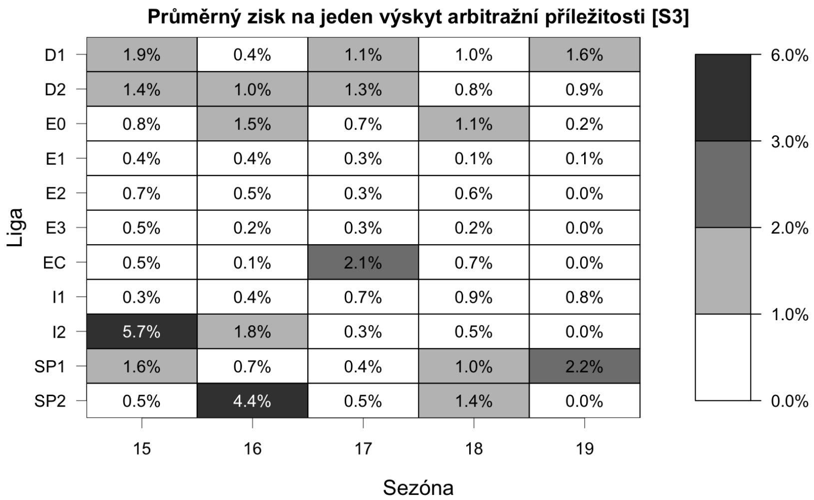
Potom výhra outsidera znamená pro sázejícího výnos:

$$K_i m_i = 1 + K_i \left(1 - \sum_{j=1}^{M} \frac{1}{K_j}\right), \quad i = 1, ..., M \tag{2.49}$$

Volba bezrizikového poměru vsázených prostředků m_j na výsledek $j \neq i$:

$$m_j = \frac{1}{K_j} \tag{2.50}$$

Tabulka průměrného zisku na jeden hrací den s výskytem (nejlepší k danému dnu) arbitrážní příležitosti pro danou strategii je k dispozici na obrázku 2.14. Například za zmíněných výše podmínek, při sázení na zápasy ligy Premier League během sezóny 2018/2019, by z 1 Kč vsazené na počátku sezóny mohl mít sázející na konci sezóny garantovaný zisk ve výši $1 \cdot 0.002 \cdot 59 = 0.118$ Kč.



Obrázek 2.14: Tabulka průměrného zisku na hrací den s výskytem arbitrážní příležitosti pro strategii „Garance + outsider“

2.5.4 Garance + favorit

Strategie zajišťuje sázejícímu největší možný výnos v případě výhry favorita (s pohledu sázkové kanceláře) a (místo ztráty) nulový zisk při nastání jiné události.

[5] Kurzy na remízu v tomto případě neuvažujeme.

Označíme K_i kurz na výhru favorita, tj. nejnižší možný kurz.[5] Volba bezrizikového poměru vsázených prostředků m_i na výhru favorita:

$$m_i = \frac{1}{K_i} + \left(1 - \sum_{j=1}^{M} \frac{1}{K_j}\right), \quad i = 1, ..., M \tag{2.51}$$

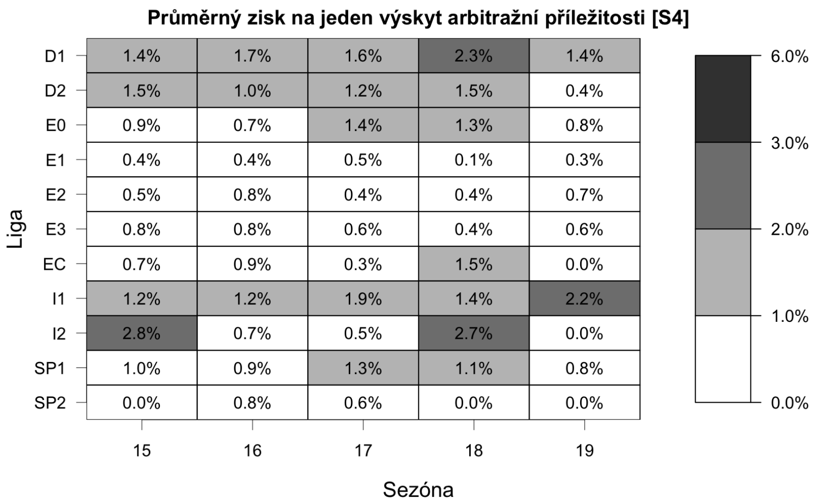
Potom výhra favorita znamená pro sázejícího výnos:

$$K_i m_i = 1 + K_i \left(1 - \sum_{j=1}^{M} \frac{1}{K_j}\right), \quad i = 1, ..., M \tag{2.52}$$

Volba bezrizikového poměru vsázených prostředků m_j na výsledek $j \neq i$:

$$m_j = \frac{1}{K_j} \tag{2.53}$$

Tabulka průměrného zisku na jeden hrací den s výskytem (nejlepší k danému dnu) arbitrážní příležitosti pro danou strategii je k dispozici na obrázku 2.15. Například za zmíněných výše podmínek, při sázení na zápasy ligy Premier League během sezóny 2018/2019, by z 1 Kč vsazené na počátku sezóny mohl mít sázející na konci sezóny garantovaný zisk ve výši $1 \cdot 0.008 \cdot 59 = 0.472$ Kč.



Obrázek 2.15: Tabulka průměrného zisku na hrací den s výskytem arbitrážní příležitosti pro strategii „Garance + favorit“

2.6 Chování sázkových kanceláří

V této podkapitole prozkoumáme chování sázkových kanceláří SK při vypsání kurzů. Pomocí relativních četností bodově odhadneme následujících pravděpodobnosti:

$$
\begin{aligned}
P(H_{SK}|H) &= P(SK \text{ tipovala vítězství domácích}|\text{domácí zvítězili})\\
P(H_{SK}|D) &= P(SK \text{ tipovala vítězství domácích}|\text{remíza})\\
P(H_{SK}|A) &= P(SK \text{ tipovala vítězství domácích}|\text{hosté zvítězili})
\end{aligned}
$$

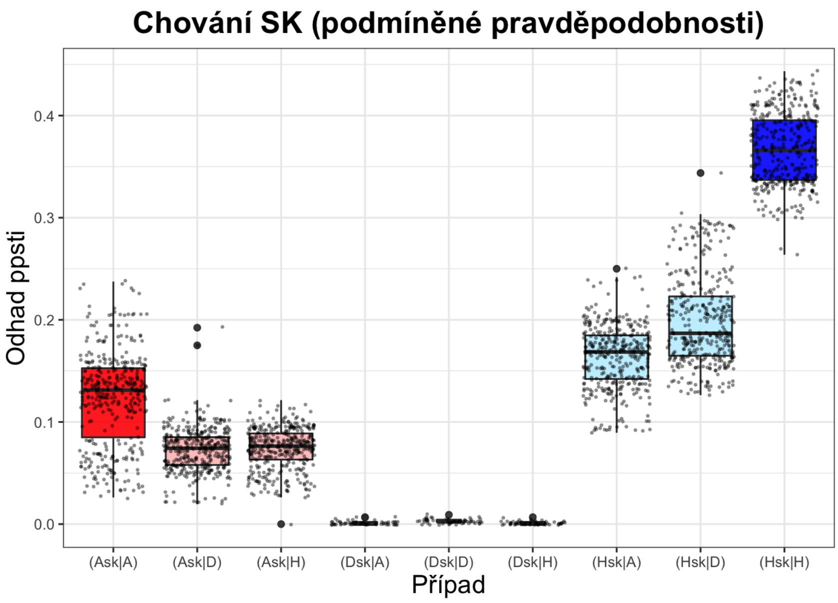
$$
\begin{aligned}
P(D_{SK}|H) &= P(SK \text{ tipovala remízu}|\text{domácí zvítězili})\\
P(D_{SK}|D) &= P(SK \text{ tipovala remízu}|\text{remíza})\\
P(D_{SK}|A) &= P(SK \text{ tipovala remízu}|\text{hosté zvítězili})
\end{aligned}
$$

$$
\begin{aligned}
P(A_{SK}|H) &= P(SK \text{ tipovala vítězství hostů}|\text{domácí zvítězili})\\
P(A_{SK}|D) &= P(SK \text{ tipovala vítězství hostů}|\text{remíza})\\
P(A_{SK}|A) &= P(SK \text{ tipovala vítězství hostů}|\text{hosté zvítězili})
\end{aligned}
$$

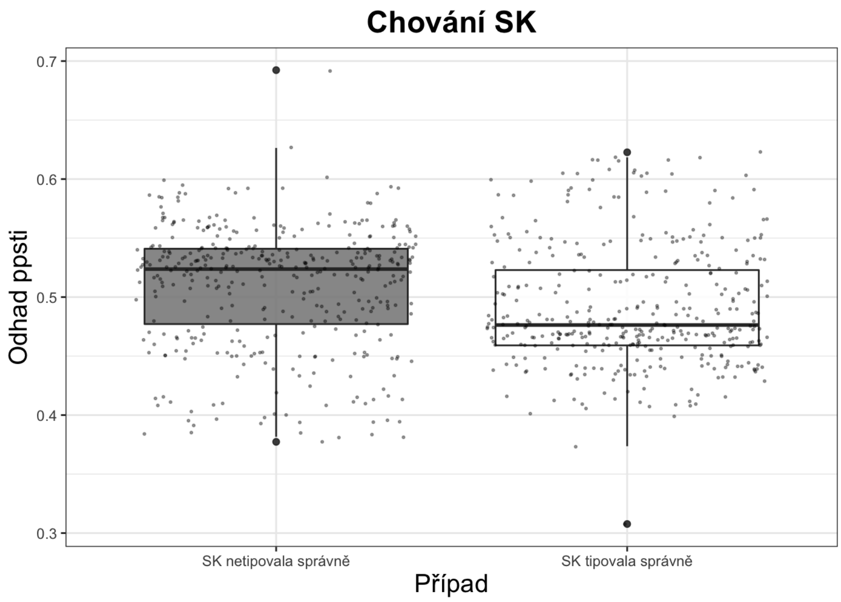
Z toho pak dopočteme:

$$
\begin{aligned}
P(SK \text{ tipovala správně}) &= P(H_{SK}|H) + P(A_{SK}|A) + P(D_{SK}|D)\\
P(SK \text{ netipovala správně}) &= 1 - P(SK \text{ tipovala správně})
\end{aligned}
$$

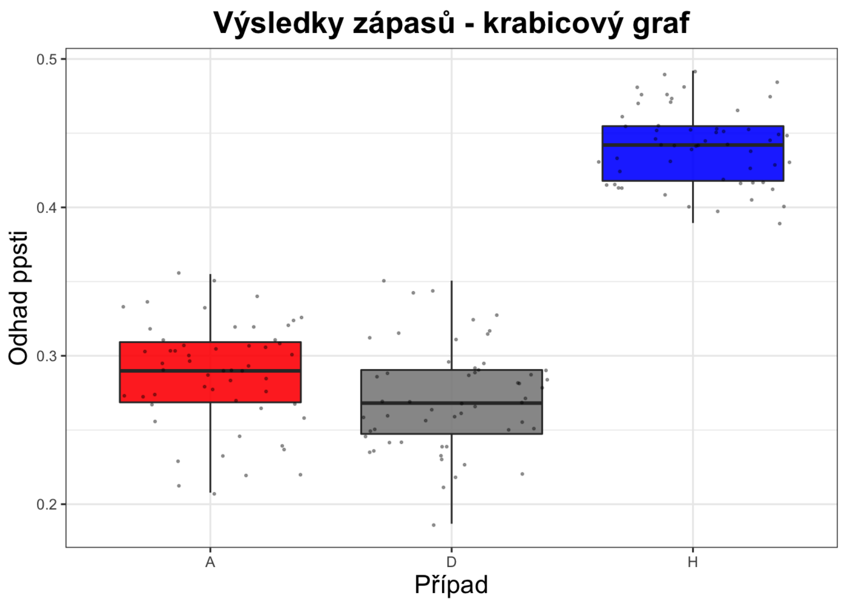
Na obrázcích 2.16 a 2.17 jsou k dispozici krabicové grafy odhadů těchto pravděpodobností, vypočtených ze všech dostupných dat. Pro porovnání vykreslíme krabicový graf 2.18 výsledků všech zápasů.



Obrázek 2.16: Chování sázkových kanceláří (podmíněné pravděpodobnosti)



Obrázek 2.17: Chování sázkových kanceláří



Obrázek 2.18: Výsledky zápasů

Šíře takové „krabice" odpovídá počtu pozorování, spodní a horní hranice reprezentují třetí $Q3$ a první $Q1$ kvartily, linie uvnitř „krabice" označuje medián $Q2$, nejnižší hodnota spodní úsečky odpovídá $Q1 - 1.5 \cdot IQR$, nejvyšší hodnota horní úsečky odpovídá $Q3 + 1.5 \cdot IQR$, kde $IQR = Q3 - Q1$ je kvartilové rozpětí, hodnoty ležící mimo reprezentují odlehlé hodnoty ([18], str. 234-238). Grafy jsou generovány pomocí přílohy A6. Vygenerované obrázky se automaticky uloží do složky „`img/chovani_sk/`" ve formátu `*.png`.

Z výsledků lze vypozorovat následující zákonitost:

- Sázková kancelář častěji preferuje výhru domácích, než výhru hostů
- Ve většině případů sázková kancelář nepreferuje remízu

Preference výhry domácích je vzhledem k převážné většině vítězství domácích týmů zřejmá, což (na první pohled) nelze říci o remízách. Zdánlivě častější preference remízy by mohla snížit chybovost při odhadu výsledků zápasů. Ale kvůli tomu, že cílem sázkové kanceláře je maximalizace zisku, nikoliv predikce výsledků, při stanovování kurzů na jednotlivé zápasy sázková kancelář nejvíce míří na odhad poměru vsazených prostředků, nikoliv na odhad pravděpodobnosti nastání jednotlivých událostí. Tedy takovou preferenci výsledků lze vysvětlit preferencí sázejících.

Uvedeme příklad toho, že sázková kancelář nemusí prodělat i když neuhádne výsledek zápasu. Zápas, který se ukončil remízou:

Liga	**Datum**	**Domácí**	**Hosté**	**H**	**D**	**A**
Premier League	22/04/2019	Chelsea	Burnley	1.25	6.25	15.00

Podle vzorce (2.11) by odhad poměru sázek, předpokládaný bookmakerem, měl být 77.92 % na domácí, 15.58 % na remízu a 6.49 % na hosty. Při takovém rozdělení prostředků se zisk sázkové kanceláře rovná 0.0260 Kč ze vsazené 1 Kč bez ohledu na to, který z výsledků nastal. V realitě se předpokládaný a reálný poměr vsazených prostředků může pravděpodobně lišit. Jedním z řešení takového problému je možnost sázkové kanceláře měnit vypsané kurzy na základě již reálného poměru vsazených prostředků.

Závěr

Během této práce byla rozebrána realita sázení na sportovní výsledky, konkrétně typy sázek, cíle sázkové kanceláře, výhody sázkové kanceláře a výskyty arbitrážních příležitostí. Dále byly uvedeny některé podmínky volby jednotlivých kurzů sázkovou kanceláří, odhad poměru vsázených prostředků, reprezentativní a individuální marže na základě vypsáných kurzů. V dalším kroku byly ukázány některé možné způsoby měření jistoty bookmakera s výsledkem zápasu, a to například pomocí ukazatele max_o. Čím je bližší tento ukazatel k 1, tím je jistější bookmaker s vypsanými kurzy. Dále byl navrhnut pravděpodobnostně statistický model výskytů arbitrážních příležitostí, a to s ohledem na datové „nedostatky“. Jako míra konkurenceschopnosti sázkové kanceláře byla odůvodněně zvolena reprezentativní marže. Čím je nižší hodnota reprezentativní marže, tím jsou vyšší vypsané kurzy, což odpovídá vyšší konkurenceschopnosti dané sázkové kanceláře. Následně byly probrány některé možné strategie sázení a možný zisk při využití těchto strategií. Nakonec bylo ukázáno, že cílem sázkové kanceláře je maximalizace zisku, proto při stanovování kurzů bookmaker nejvíce míří na odhad poměru vsazených prostředků, nikoliv na odhad pravděpodobnosti nastání jednotlivých událostí.

Literatura a další zdroje

[1] RStudio Download. *RStudio, PBC* [online]. 2020. Dostupné z: https://rstudio.com/products/rstudio/download/

[2] A Computational Analysis of Arbitrage Opportunities in Sports Gambling. *Brian Matejek*. Independent Work Report, Spring 2013. Dostupné z: https://pdfs.semanticscholar.org/a706/9f1df68a9f4bca0df3c14e242a2db6b80f2b.pdf

[3] Investments. *William F. Sharpe, Gordon J. Alexander, Jeffery V. Bailey*. Prentice Hall, Upper Saddle River. NewJersey 07458. ISBN 0-13-010130-3.

[4] Sports betting arbitrage in English football. *Patrice Marek, František Vávra*. 2019. Dostupné z: https://www.researchgate.net/publication/331199091_Sports_Betting_Arbitrage_in_English_Football

[5] Sázková (individuální) arbitráž. Pracovní text ZČU Plzeň, FAV/KMA. *František Vávra*. Září 2018.

[6] Historical Football Results and Betting Odds Data. *Football-Data.co.uk* [online]. 2020. Dostupné z: https://www.football-data.co.uk/data.php

[7] Notes. *Football-Data.co.uk* [online]. 2020. Dostupné z: https://www.football-data.co.uk/notes.txt

[8] Maržový model. Pracovní text ZČU Plzeň, FAV/KMA. *František Vávra*. Srpen 2019.

[9] Poznámky k měření toho, jak si je jistá sázková kancelář. Pracovní text ZČU Plzeň, FAV/KMA. *František Vávra*. Listopad 2019.

[10] Co to je „korelace" pro tuto úlohu. Pracovní text ZČU Plzeň, FAV/KMA. *František Vávra*. Listopad 2019.

[11] Základy počtu pravděpodobnosti a matematické statistiky. *Jaroslav Hátle, Jiří Likeš*. SNTL Praha 1974.

[12] The coefficients of correlation and determination as measures of performance in forecast verification. *Murphy, A. H.* Wea. Forecasting, vol. 10, 1995.

[13] Correlation and regression analysis: a historian's guide. *Archdeacon, Thomas J.* University of Wisconsin Press, 1994. ISBN 0-299-13650-7.

[14] Correlation (in statistics). *A.V. Prokhorov (originator)*. Encyclopedia of Mathematics. ISBN 0-299-13650-7 [online]. Dostupné z:
`http://www.encyclopediaofmath.org/index.php?title=Correlation_(in_statistics)&oldid=11629`

[15] Stochastic processes. *Sheldon M. Ross*. Wiley, 1996. ISBN 0-471-12062-6.

[16] Model. Pracovní text ZČU Plzeň, FAV/KMA. *František Vávra*. Březen 2020.

[17] „On Markov-Type Inequalities". *Steliga, Katarzyna; Szynal, Dominik*. International Journal of Pure and Applied Mathematics, 58 (2). ISSN 1311-8080 [online]. Dostupné z:
`http://ijpam.eu/contents/2010-58-2/2/2.pdf`

[18] A Modern Introduction to Probability and Statistics. *Dekking, F.M.* Springer, 2005. ISBN 1-85233-896-2.

[19] Inter-market Arbitrage in Sports Betting. *Egon Franck, Erwin Verbeek, Stephan Nuesch*. Working Paper #48, October 2009. Dostupné z:
`http://www.ncer.edu.au/papers/documents/WPNo48.pdf`

[20] Fair Bets in Sports Betting. *Jiří Jansa*. IES FSV UK, diplomová práce FSV UK 2012. Dostupné z:
`ies.fsv.cuni.cz/default/file/download/id/19308`

[21] A Sure Win Strategy for Football Betting. *Stanley Ong, Chris Kwek*. 101footballblueprint.com. Dostupné z:
`http://www.thelivebettingsystem.com/downloads/surewinstrategy_2009mar22.pdf`

Přílohy

Příloha A

A1 `1_notes.txt`
Soubor obsahuje popis fotbalových dat použitých ke zpracování

A2 `2_jistota.R`
Soubor obsahuje „korelační“ analýzu vybraných ukazatelů s cílem najít ukazatel popisující míru jistoty sázkové kanceláře s výsledkem zápasu, viz podkapitola 2.2.

A3 `3_model.R`
Soubor obsahuje výpočty využité k sestavení pravděpodobnostně statistického modelu výskytu arbitrážní příležitosti, viz podkapitola 2.3.

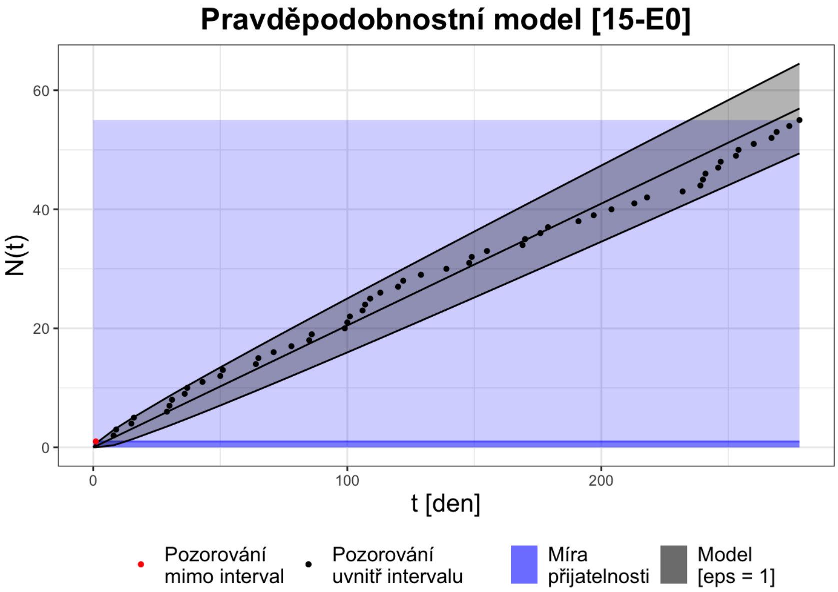
A4 `4_konkurence.R`
Soubor obsahuje analýzu vybraných ukazatelů s cílem najít ukazatel popisující míru konkurenceschopnosti sázkové kanceláře, viz podkapitola 2.4.

A5 `5_strategie.R`
Soubor obsahuje výpočet četností výskytů arbitrážních příležitostí a průměrného zisku na jeden výskyt arbitrážní příležitosti pro některé strategie sázení, viz podkapitola 2.5.

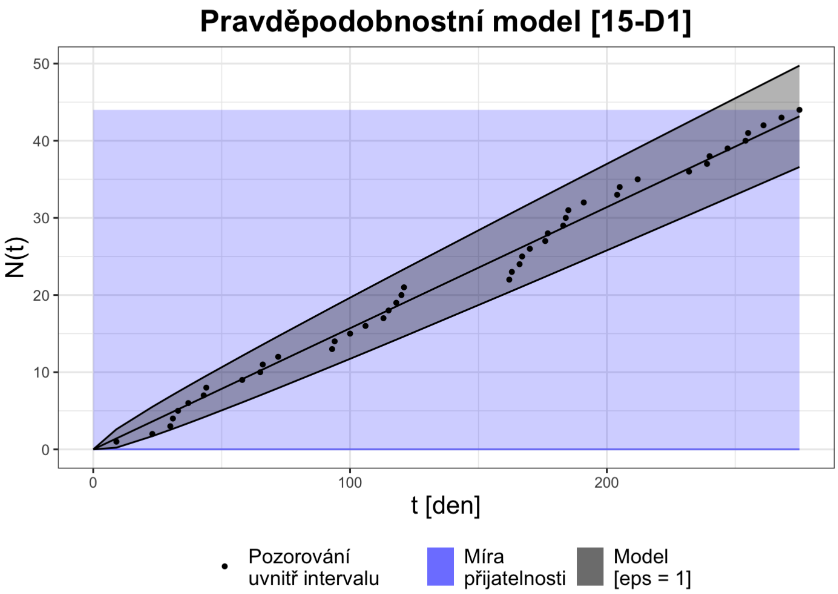
A6 `6_chovani_sk.R`
Soubor obsahuje bodové odhady pravděpodobností využitých k analýze chování sázkových kanceláří, viz podkapitola 2.6.

Příloha B

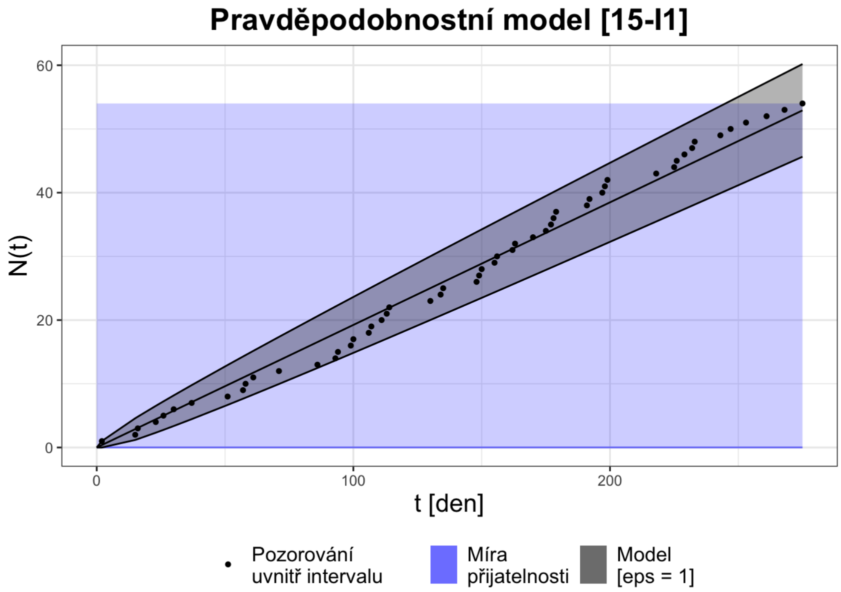
B1 Pravděpodobnostně statistické modely nejvyšších fotbalových lig



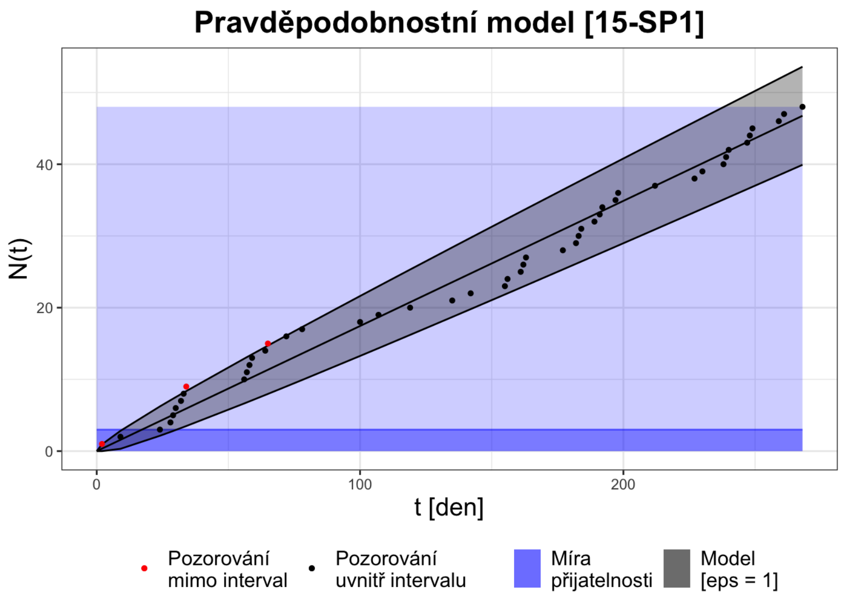
(a) 2014/2015 – Premier League



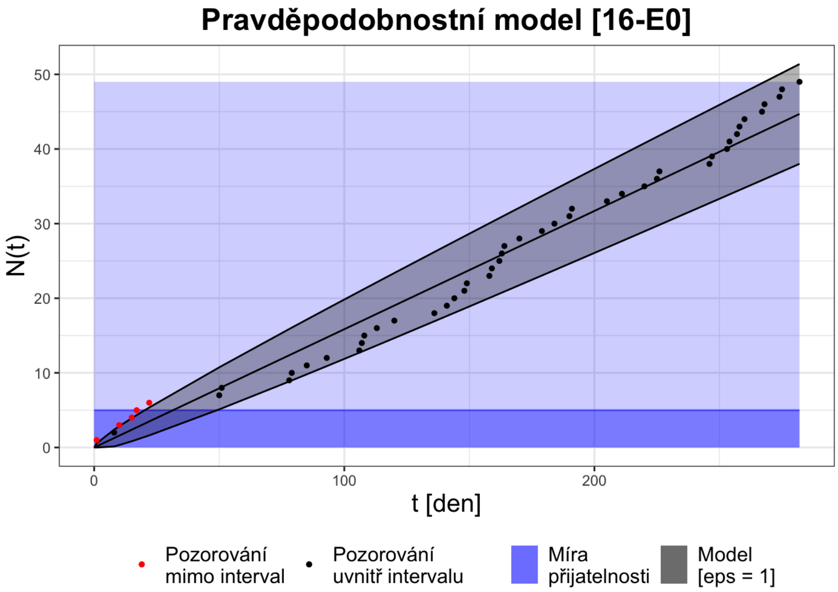
(b) 2014/2015 – Bundesliga 1



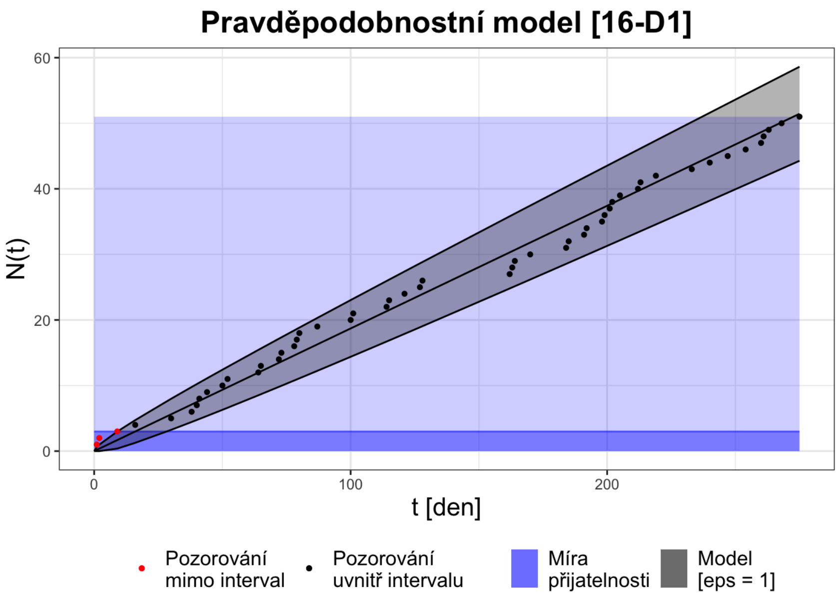
(c) 2014/2015 – Serie A



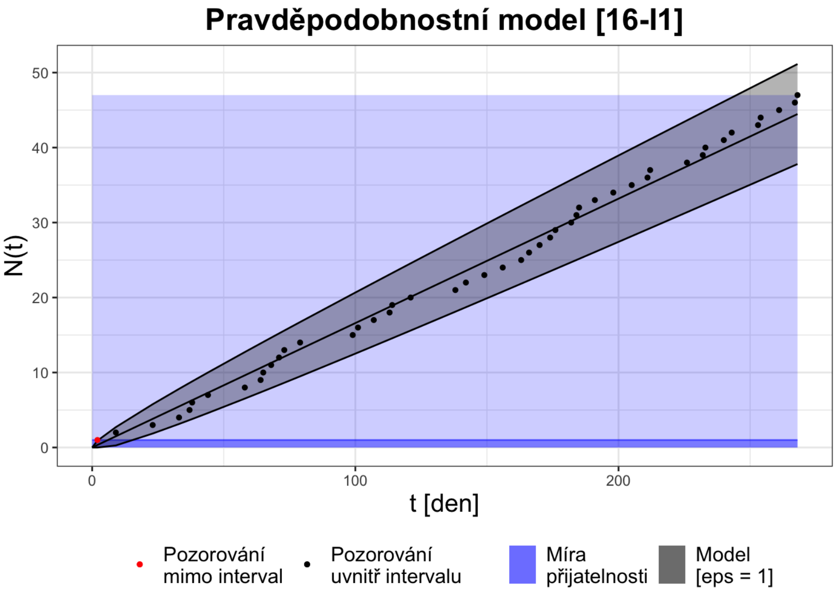
(d) 2014/2015 – La Liga Primera Division



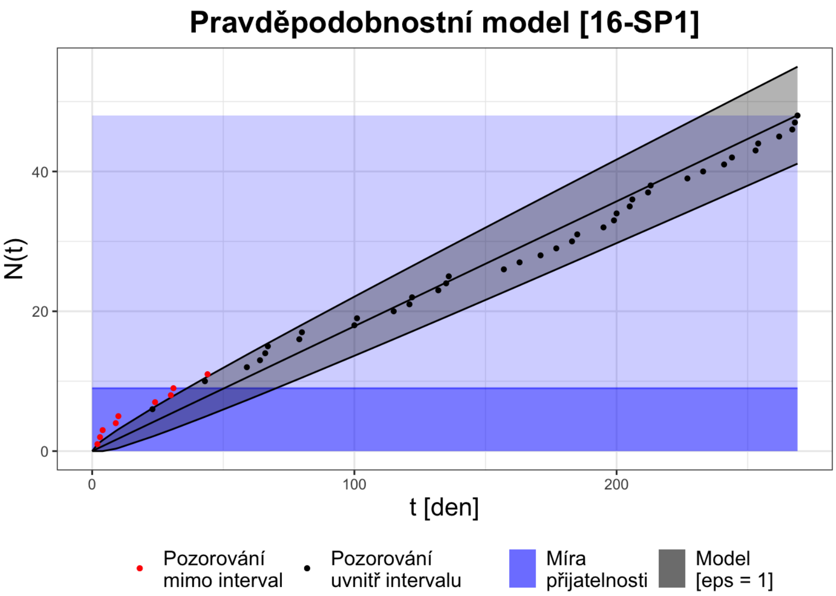
(e) 2015/2016 – Premier League



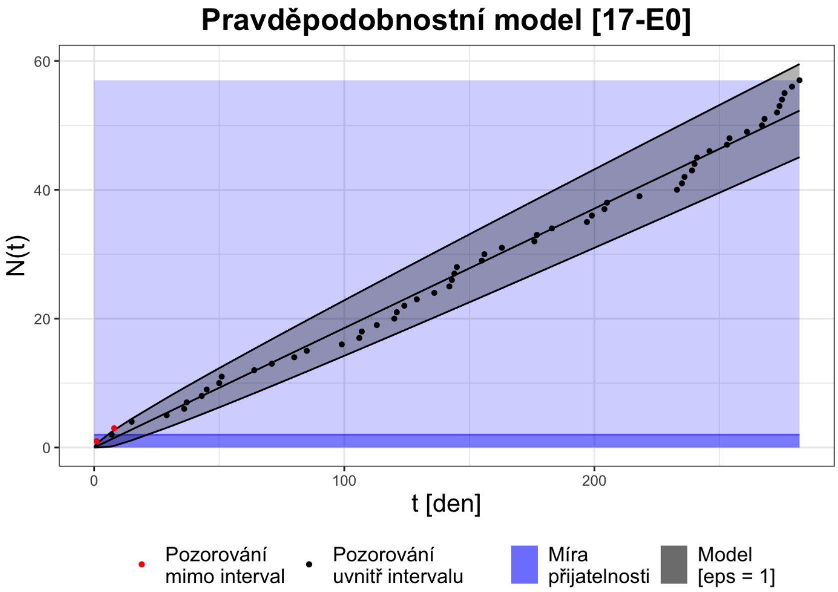
(f) 2015/2016 – Bundesliga 1



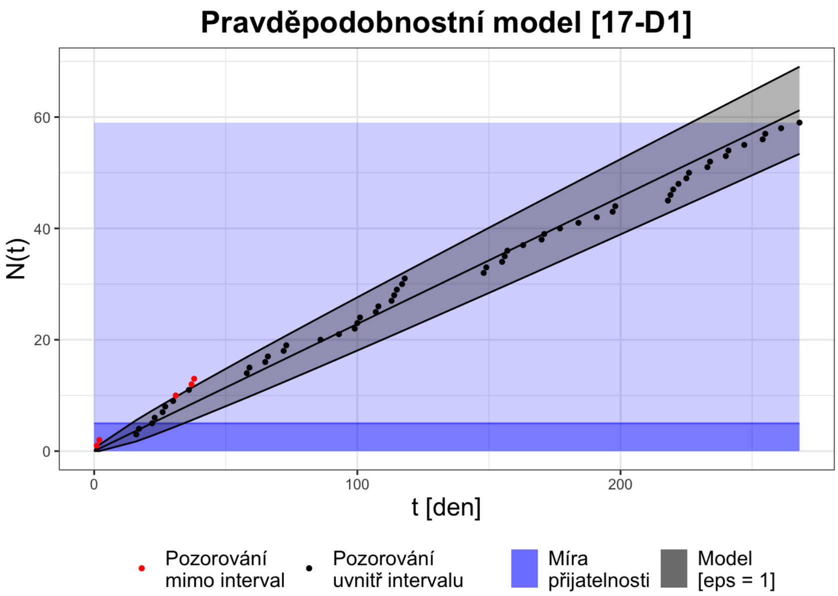
(g) 2015/2016 – Serie A



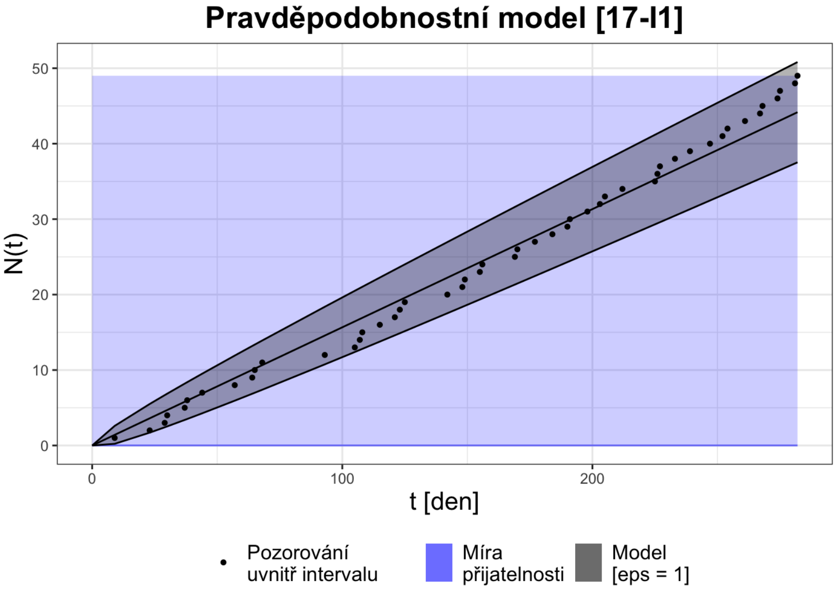
(h) 2015/2016 – La Liga Primera Division



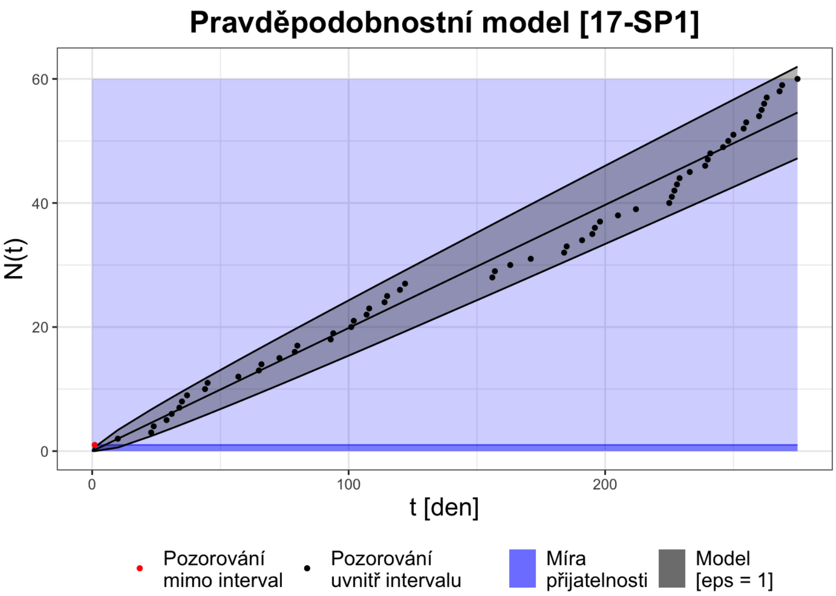
(i) 2016/2017 – Premier League



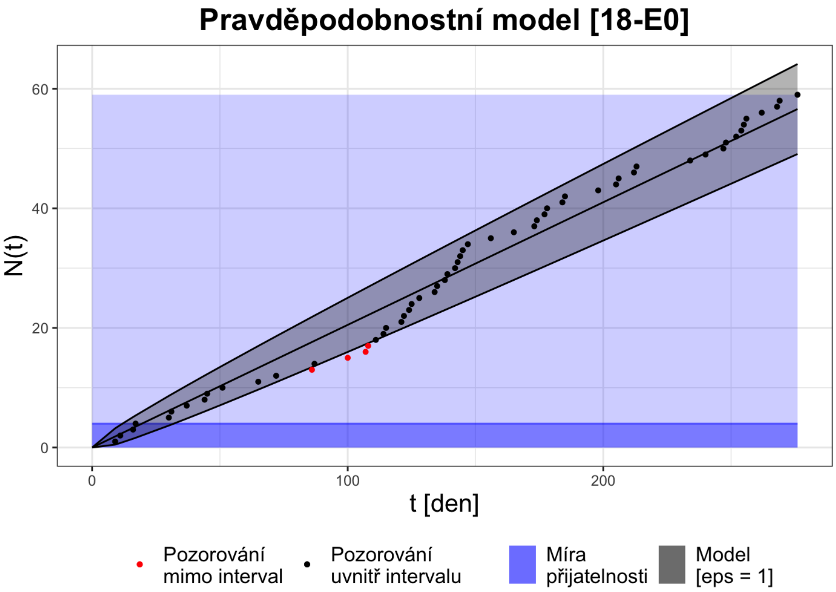
(j) 2016/2017 – Bundesliga 1



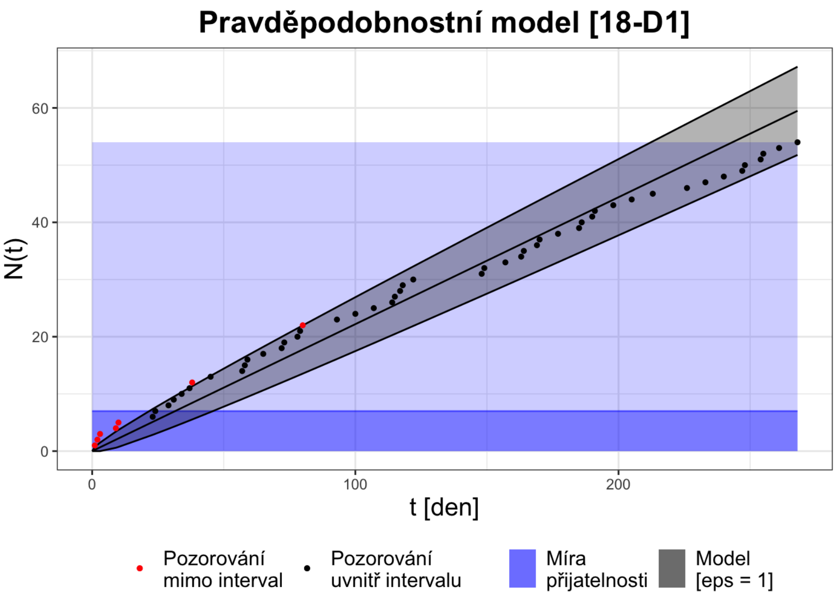
(k) 2016/2017 – Serie A



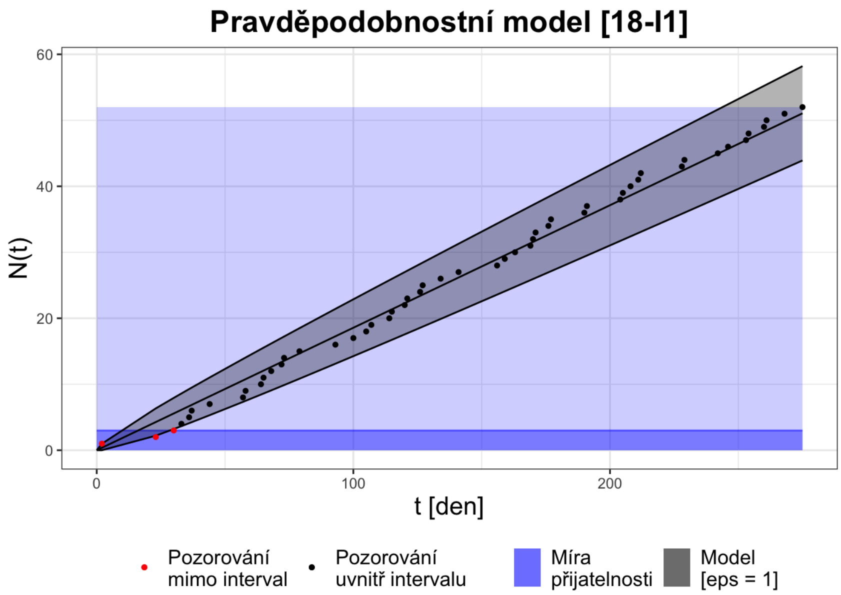
(l) 2016/2017 – La Liga Primera Division



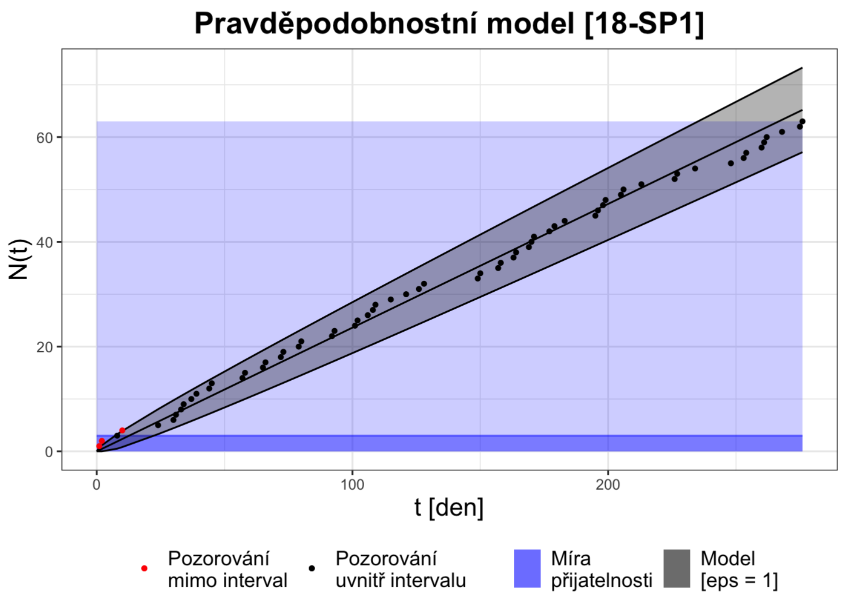
(m) 2017/2018 – Premier League



(n) 2017/2018 – Bundesliga 1



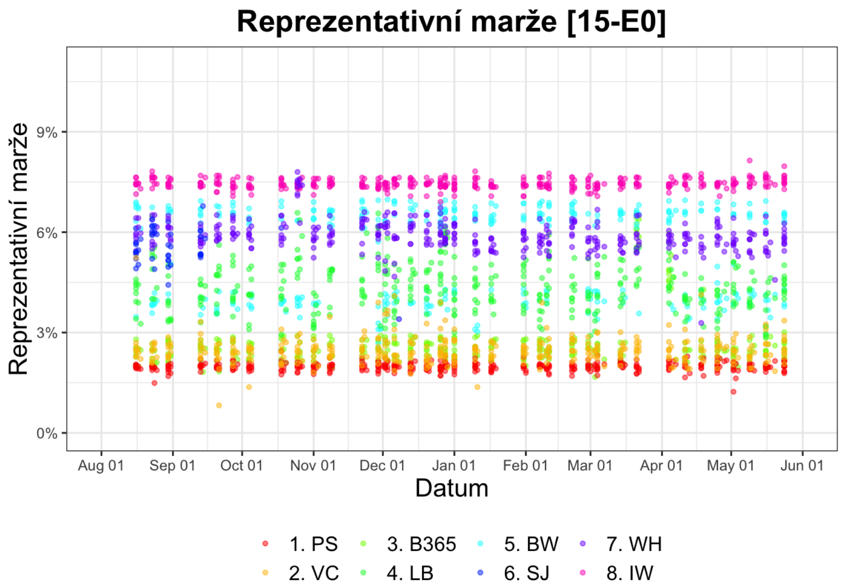
(o) 2017/2018 – Serie A



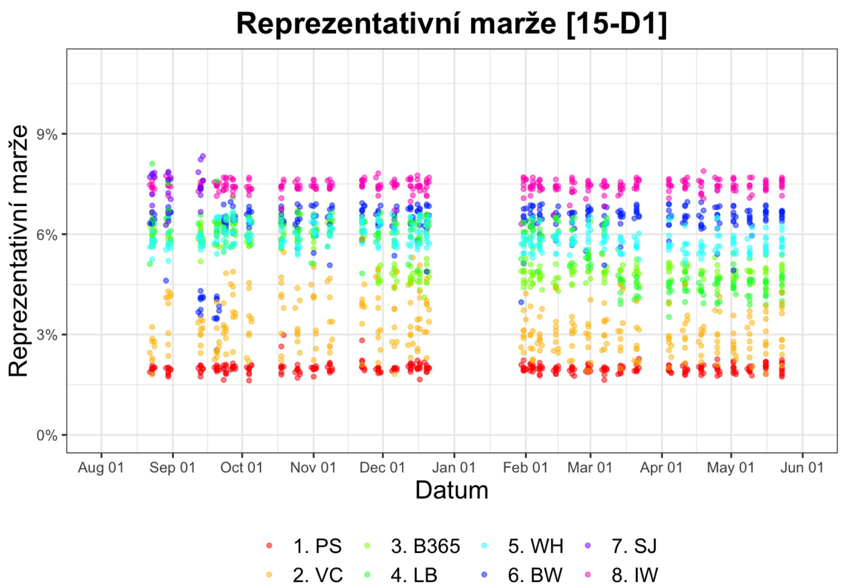
(p) 2017/2018 – La Liga Primera Division

Obrázek B1: Pravděpodobnostně statistické modely nejvyšších fotbalových lig

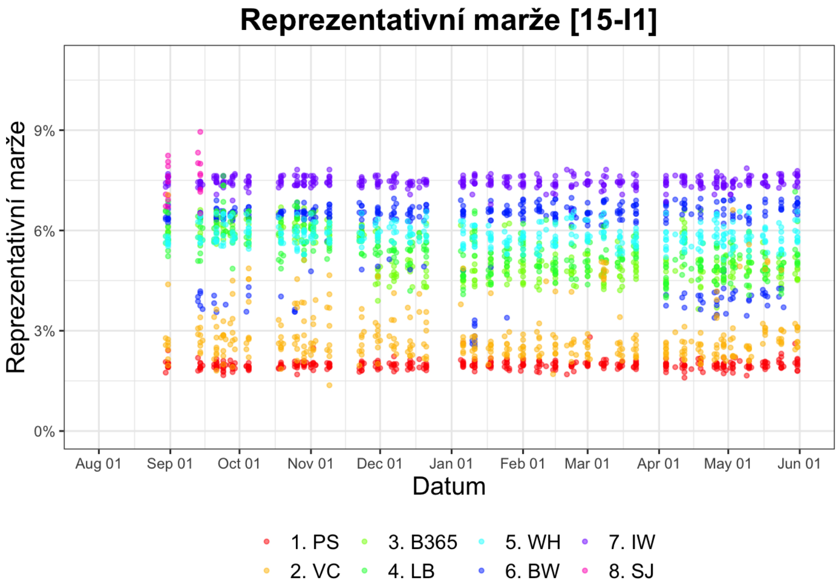
B2 Vývoj reprezentativní marže nejvyšších fotbalových lig



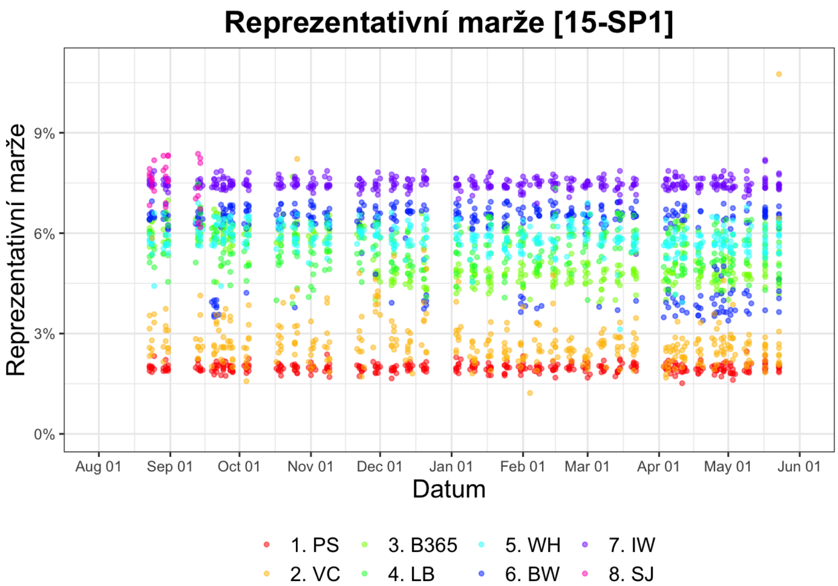
(a) 2014/2015 – Premier League



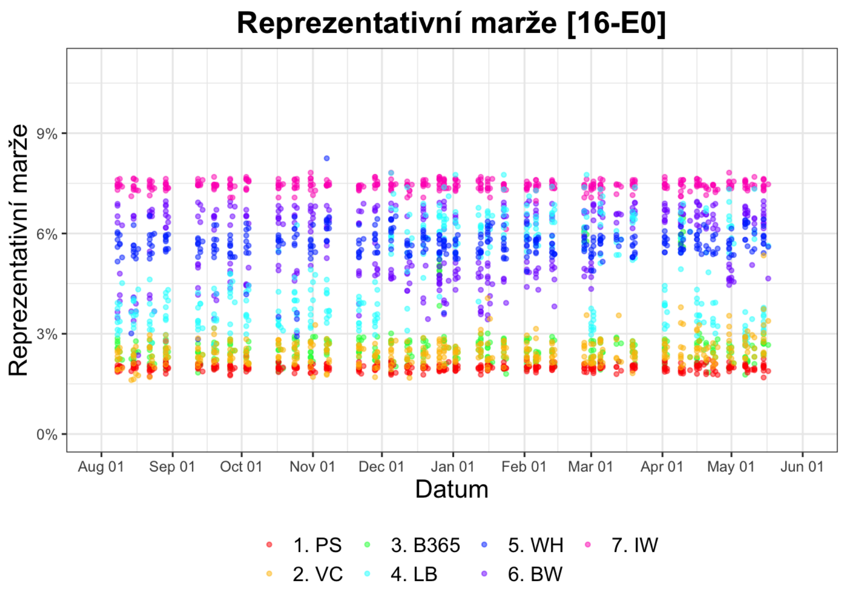
(b) 2014/2015 – Bundesliga 1



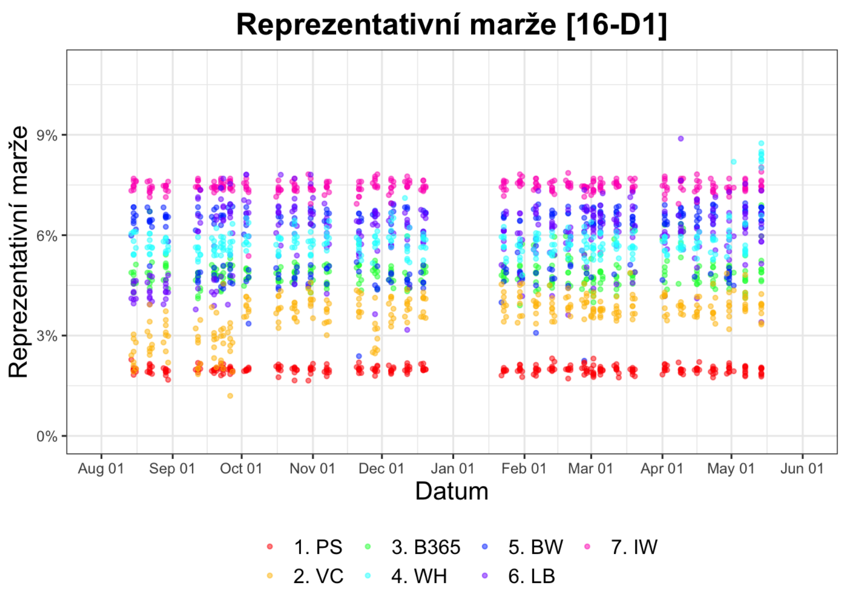
(c) 2014/2015 – Serie A



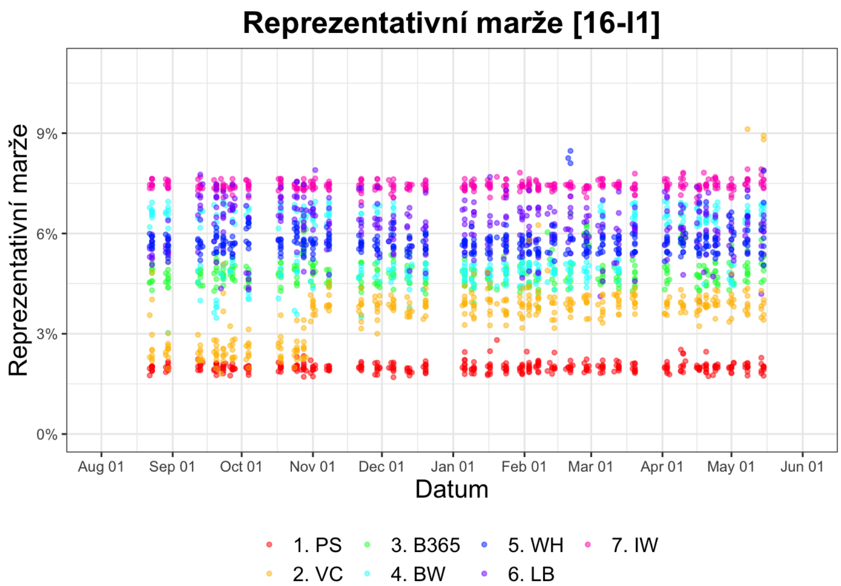
(d) 2014/2015 – La Liga Primera Division



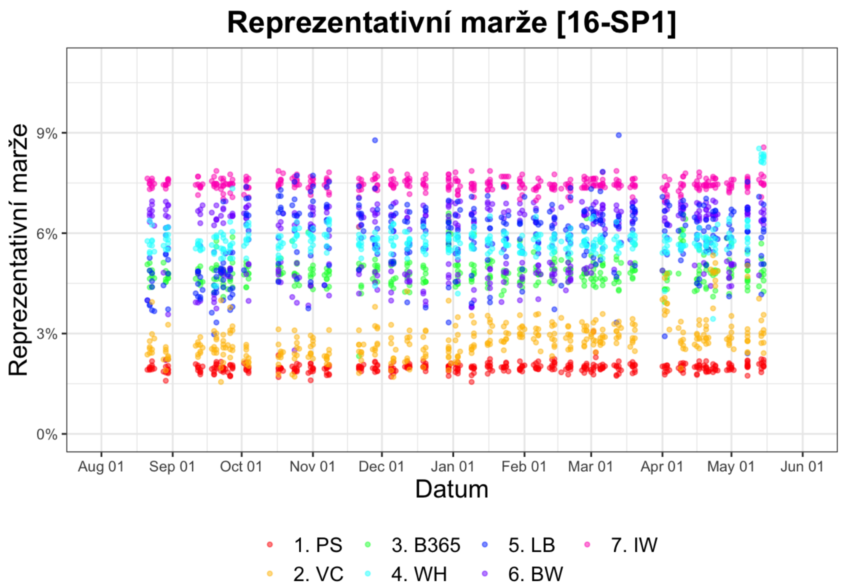
(e) 2015/2016 – Premier League



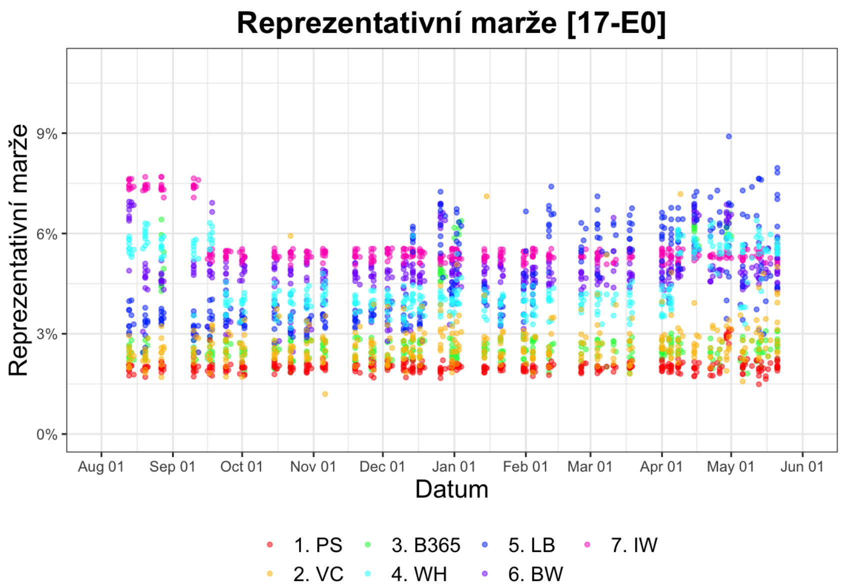
(f) 2015/2016 – Bundesliga 1



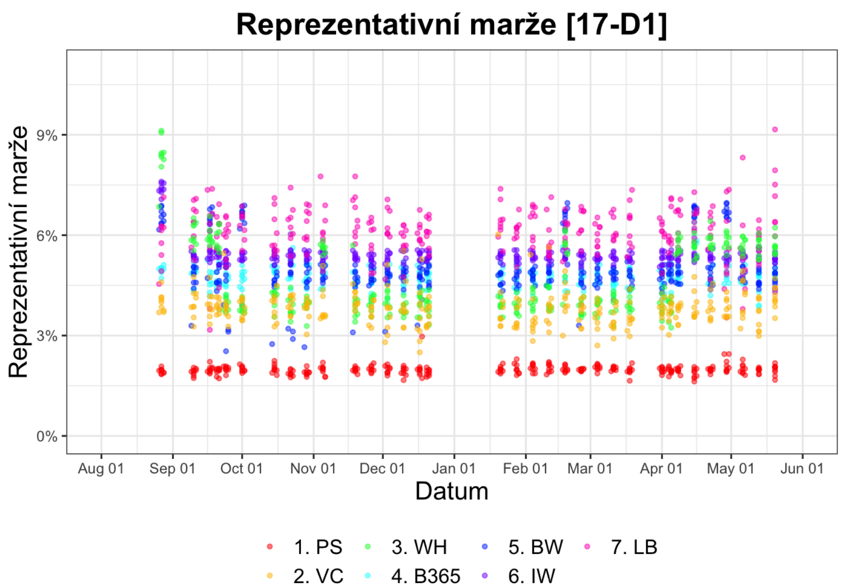
(g) 2015/2016 – Serie A



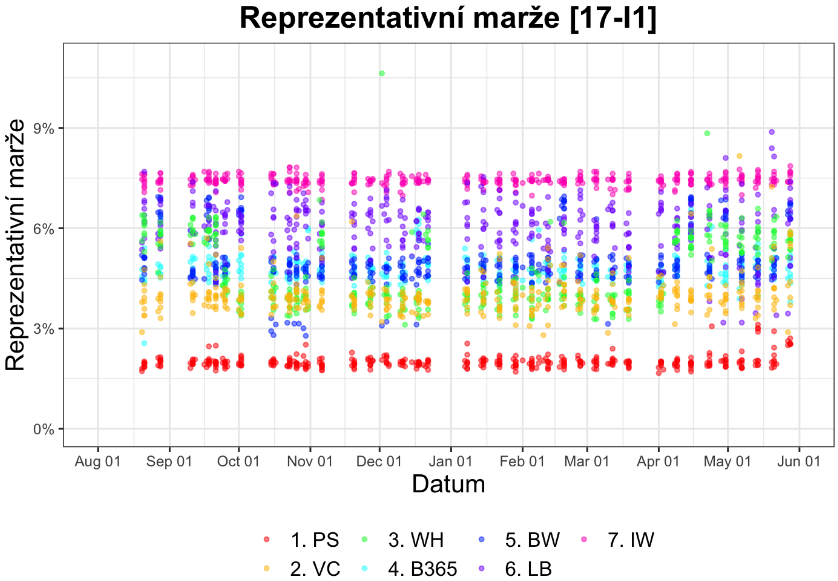
(h) 2015/2016 – La Liga Primera Division



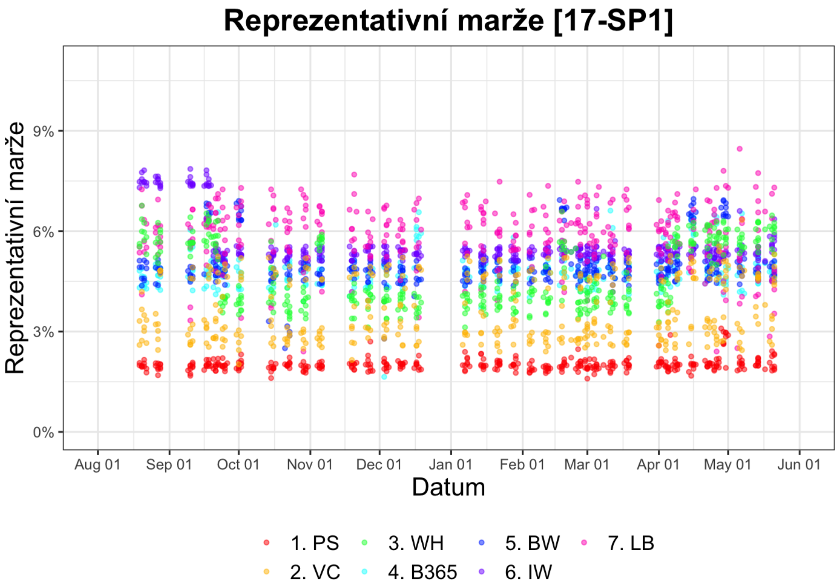
(i) 2016/2017 – Premier League



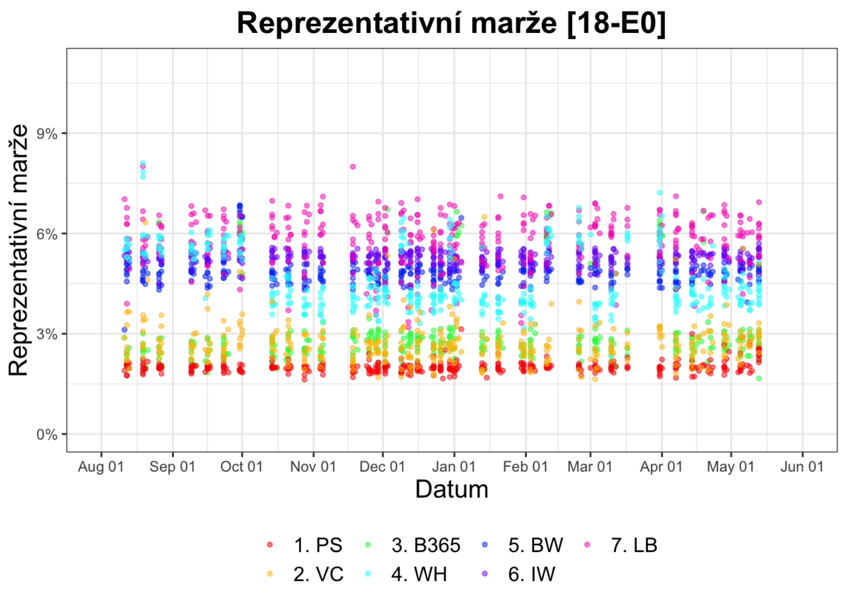
(j) 2016/2017 – Bundesliga 1



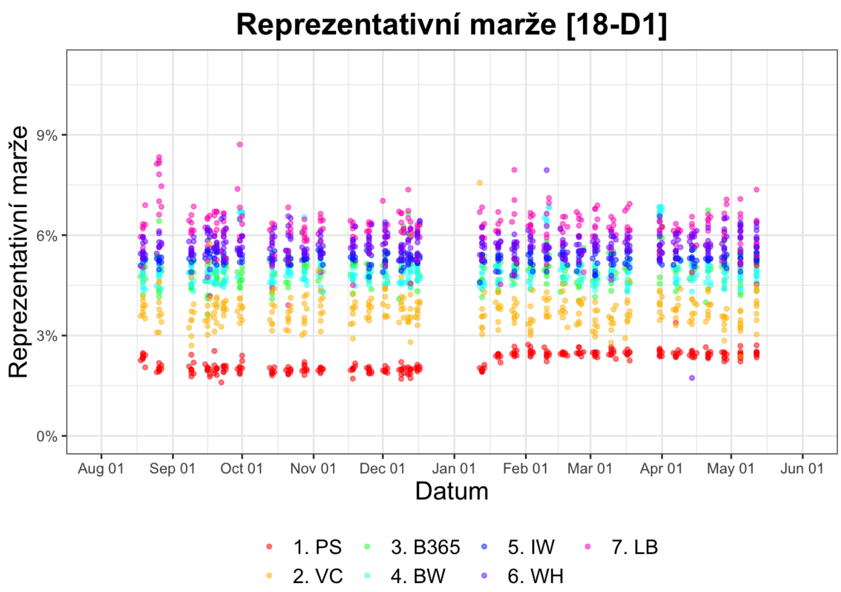
(k) 2016/2017 – Serie A



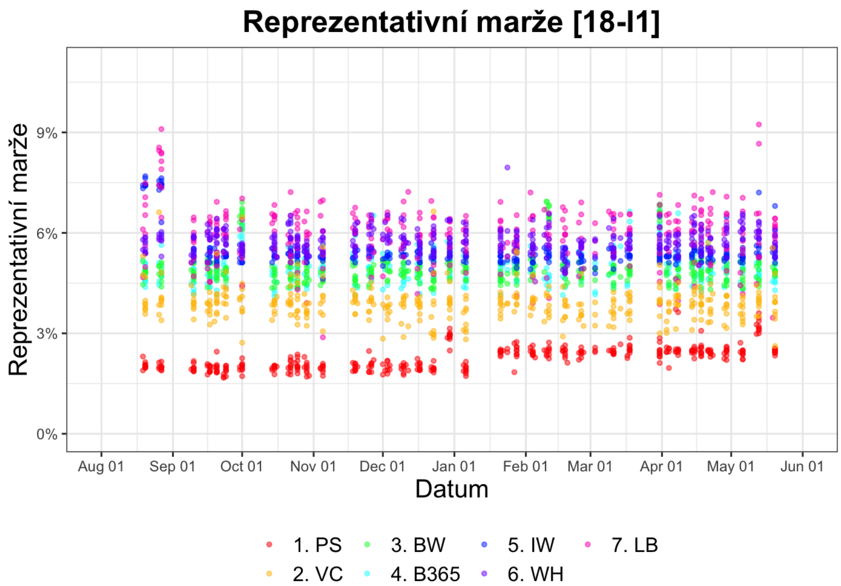
(l) 2016/2017 – La Liga Primera Division



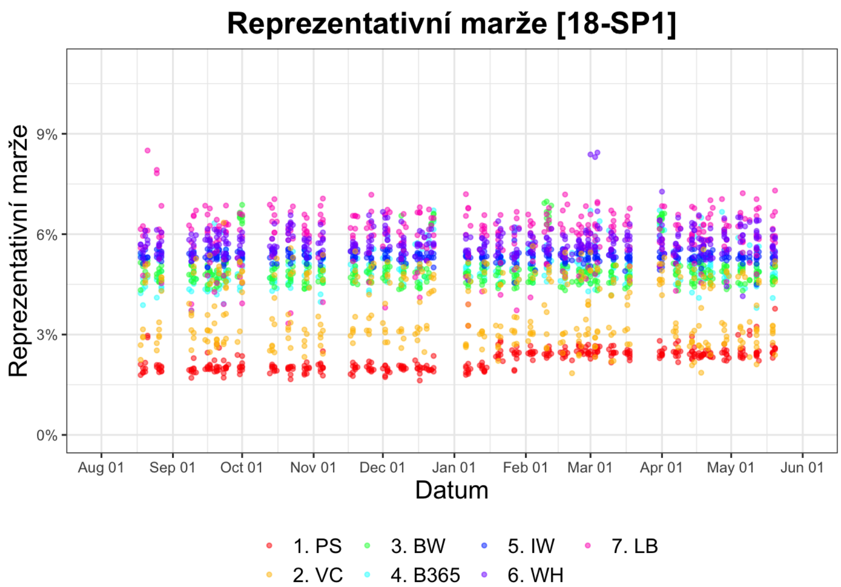
(m) 2017/2018 – Premier League



(n) 2017/2018 – Bundesliga 1



(o) 2017/2018 – Serie A



(p) 2017/2018 – La Liga Primera Division

Obrázek B2: Vývoj reprezentativní marže nejvyšších fotbalových lig