

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra informatiky a výpočetní techniky

Diplomová práce

Statistická analýza výsledků sportovních soutěží

Prohlášení

Prohlašuji, že jsem diplomovou práci vypracoval samostatně a výhradně s použitím citovaných pramenů.

V Plzni dne _____

Tomáš Lindauer

Poděkování

Velmi rád bych poděkoval panu RNDr. Petru Stehlíkovi, Ph.D. za odborné vedení, řadu užitečných materiálů a informací a za čas, který této práci věnoval. Dále bych rád poděkoval společnosti Trefík.cz za poskytnutí dat potřebných k vypracování této diplomové práce.

Abstrakt

Cílem této práce je statistické zpracování výsledků soutěží v ledním hokeji. Pomocí základní statistické analýzy identifikujeme rozdíly mezi jednotlivými soutěžemi. Zjišťujeme, zda se jev averze ke ztrátě, pocházející z ekonomické teorie rozhodování, objevuje také v oblasti sportu (v ledním hokeji). Navrhli jsme hypotézu, kde testujeme, zda je počet vstřelených gólů ve všech třetinách stejný. Druhá námi navržená hypotéza porovnává dvě skupiny utkání. Testujeme, zda je průběh zápasu ve třetí třetině rozdílný pro utkání s remízovým a neremízovým stavem po dvou odehraných třetinách. Výsledky obou těchto hypotéz se pokoušíme vysvětlit pomocí averze ke ztrátě. Poslední navržená hypotéza se zabývá změnou pravidel v české extralize. Od sezony 2000/2001 je zaveden tříbodový systém ohodnocení vítězství týmu v základní hrací době. Zkoumáme, zda měla tato změna vliv na herní projev týmů v poslední třetině.

Klíčová slova: averze ke ztrátě, analýza hokejových soutěží, třífaktorová ANOVA, testování hypotéz, vliv změny pravidel

Abstract

The aim of this thesis is a statistical analysis of ice hockey leagues' results. We identify the differences among the major leagues. We examine if the phenomenon of loss aversion (coming from decision theory) appears also in sports area (ice hockey). We formulate a hypothesis, whether the number of goals in each period of the game is the same. Second hypothesis compares two groups of matches. We test if there is a difference in goal dynamics of third period between the tied and untied matches after second period of the game. The results of both hypotheses could be explained by loss aversion. The last hypothesis deals with the rule changes in Czech Extraleague. There is a three point system established since season 2000/2001 for a team that wins in regular time of the game. We try to analyze if this change of rules has an influence on team tactics in the last period of the game.

Key words: Loss aversion, analysis of Ice Hockey Leagues, Three-Way ANOVA, hypotheses testing, influence of rules changes

Obsah

1	Úvod	1
2	Popis dat	2
2.1	Datový soubor	2
2.2	Rozdíly v bodovém ohodnocení	5
2.3	Základní statistiky	8
2.3.1	Vstřelené góly	8
2.3.2	Góly po minutách	10
2.3.3	Vyrovnanost soutěží	13
3	Averze ke ztrátě	18
3.1	Nadační efekt	18
3.2	Náklonnost k současnému stavu	19
3.3	Averze ke ztrátě	21
4	Rozdíl gólů mezi třetinami	24
4.1	Shoda počtu branek v jednotlivých třetinách	25
4.2	Regrese	27
4.2.1	Lineární regrese	27
4.2.2	Nelineární regrese	30
5	Vliv systému bodování	33
5.1	Remíza vs. Neremíza (vše)	33
5.1.1	Srovnání systémů pro remízová utkání	34
5.1.2	Srovnání systémů pro neremízová utkání	35
5.2	Remíza vs. Neremíza (o jeden gól)	36
5.2.1	Srovnání systémů pro remízová utkání	37
5.2.2	Srovnání systémů pro neremízová utkání	37
5.3	Remíza vs. Neremíza (skupiny)	38
5.3.1	Srovnání systémů pro remízová utkání	40
5.3.2	Srovnání systémů pro neremízová utkání	41

5.4	Testování závislosti ppsti gólu na bodování	43
6	Dynamika zápasu ve třetí třetině	45
6.1	Předpoklady pro analýzu rozptylu	45
6.1.1	Nezávislost dat	45
6.1.2	Homoskedasticita	45
6.1.3	Test normality	47
6.2	t-test pro jednotlivé soutěže	48
6.3	Třífaktorová ANOVA	52
6.3.1	Model 3Faktorové - ANOVA	52
7	Závěr a shrnutí dosažených výsledků	55

1 Úvod

Velmi zajímavou oblastí pro statistickou analýzu jsou výsledky sportovních utkání. Tato oblast je obzvláště přitažlivá, jelikož obsahuje velké množství dat, která jsou nezbytná pro testování ekonomických teorií. V této práci jsme se zaměřili na jev z teorie rozhodování. Zkoumáme zajímavosti v průběhu utkání v ledním hokeji a jejich souvislost s averzí ke ztrátě.

Pro analýzu výsledků a identifikaci jevů je nutné mít dostatečně rozsáhlou databázi utkání. Náš datový soubor se skládá z podrobných záznamů o více než 26 000 zápasech. Více než polovina zápasů je z kanadsko-americké NHL. Ostatní utkání jsou z evropských lig včetně české nejvyšší soutěže a také mezinárodních utkání. Podrobné seznámení s datovým souborem a jeho základní statistický popis je uveden v Kapitole 2.

V následující kapitole se seznámíme s pojmem averze ke ztrátě používaným v teorii rozhodování. Tento pojem je spjat s dalšími jevy objevujícími se v ekonomii. Jedná se o nadační efekt a náklonnost k aktuálnímu stavu. Všechny tyto pojmy jsou vysvětleny na konkrétních příkladech. V závěru kapitoly jsou uvedeny námi navržené hypotézy, které jsou spojeny s averzí ke ztrátě a její existencí v utkáních ledního hokeje.

Další část navazuje na Kapitulu 2, která popisuje datový soubor. Objevili jsme zde totiž velmi zajímavý jev týkající se rozložení počtu vstřelených branek v průběhu utkání. Porovnáváme počty branek v jednotlivých třetinách a také testujeme závislost průměrného počtu gólů na minutě v utkání. Pomocí námi navrženého modelu nelineární regrese zobrazíme průběh vstřelených branek v závislosti na probíhající minutě utkání.

V Kapitole 5 se zabýváme vlivem změny bodovacího systému na průběh poslední třetiny utkání. Testujeme, zda se v české extralize změnil průměrný počet branek vstřelených ve třetí třetině zápasu (mezi 41. a 55. minutou) po zvýšení bodového ohodnocení za vítězství v základní hrací době z dvou na tři body. Zkoumáme také, zda má bodový systém vliv na pravděpodobnost, že ve třetí třetině padne branka.

V této části navazujeme na teorii z Kapitoly 3, jelikož pomocí averze ke ztrátě odůvodníme výsledky hypotéz. Porovnáváme průběh třetích třetin u zápasů, kde je stav po čtyřiceti minutách remízový a utkání, kde je stav neremízový. Použijeme také třífaktorovou analýzu rozptylu k rozpoznání faktorů, které mají významný vliv na průměrný počet vstřelených branek ve třetí třetině zápasu.

V závěru stručně popíšeme výsledky jednotlivých hypotéz. Zároveň také shrneme zajímavé jevy v utkáních, které jsme v průběhu práce objevili.

2 Popis dat

2.1 Datový soubor

Datový soubor byl složen z údajů o zápasech ze světových soutěží v ledním hokeji. Celkově se jednalo o 26 313 zápasů. Byla použita data z české hokejové první ligy, české extraligy, slovenské extraligy, švédské Elitserien, finské SM liiga, kanadsko-americké NHL a také údaje z mezistátních utkání. Poděkování patří společnosti Trefík.cz, která nám poskytla tuto rozsáhlou databázi. Data byla získána ve formátu *.csv* a zpracována pomocí výpočetního programu Matlab¹. Veškeré výpočty a výstupy (tabulky, grafy) byly tedy provedeny v tomto softwaru. O zápasech byly zaznamenávány tyto údaje:

- v rámci jaké ligy byl zápas odehrán
- zda se jednalo o zápas v základní hrací části nebo playoff
- číslo kola v jakém se zápas odehrál
- datum utkání
- rozdělení na domácí a hostující tým
- výsledek utkání tzn. počet vstřelených branek domácího a hostujícího týmu
- časy vstřelení každé branky v utkání
- ukončení zápasu v základní hrací době či v prodloužení
- rozhodnutí zápasu v trestných střeleních

¹verze *R2009b*

Country	League	ID League	ID Match	Playoff 0/1	Round	Neutral site	Date	Home Team	ID Home Team	Away Team	ID Away Team	Goals home	Goals away
CZE	1st League	73837	5257190	0	10	0	14.10.2000	dř"stdř" n. Labem	433931	Opava	430999	4	3
CZE	1st League	73837	5257192	0	11	0	18.10.2000	Kadadř"n	437449	dř"stdř" n. Labem	433931	2	4
CZE	1st League	73837	5257224	0	15	0	4.11.2000	Hvdř"zda Brno	437556	Tdř"ebdř" dř"	436200	4	4
CZE	1st League	73837	5257300	0	26	0	13.12.2000	Hvdř"zda Brno	437556	Beroun	430985	5	3
CZE	Extraleague - Q	80667	5637721	1	-1	0	23.3.2002	Kladno	430995	Liberec	430996	1	4
CZE	Extraleague - Q	80667	5637938	1	-1	0	24.3.2002	Kladno	430995	Liberec	430996	2	5
CZE	Extraleague - Q	80667	5637942	1	-1	0	27.3.2002	Liberec	430996	Kladno	430995	2	4
CZE	Extraleague - Q	80667	5637943	1	-1	0	28.3.2002	Liberec	430996	Kladno	430995	3	0
CZE	Extraleague - Q	80667	5643500	1	-1	0	31.3.2002	Kladno	430995	Liberec	430996	1	2
FIN	SM Liiga	73969	5307675	0	1	0	16.9.1999	Kiekko Espoo	431033	Ilves Tampere	431029	2	4
FIN	SM Liiga	73969	5307676	0	1	0	16.9.1999	Jokerit Helsinki	431030	Lahti Pelicans	442582	1	4
FIN	SM Liiga	73969	5307677	0	1	0	16.9.1999	Jyp Jyvďř"skyldř"	431031	Hďř"meenlinna	431027	3	5
FIN	SM Liiga	73969	5307678	0	1	0	16.9.1999	Lukko Rauma	431034	Saipa	433950	4	1
FIN	SM Liiga	73969	5307679	0	1	0	16.9.1999	Tappara Tampere	431035	dř"ssdř" t Pori	431026	2	4
GLO	Baltica Cup	75041	5370579	0	1	1	17.12.1997	Russia	431265	Sweden	431267	0	0
GLO	Baltica Cup	75042	5370805	0	2	1	16.12.1998	Finland	431258	Czech Republic	431257	0	3
GLO	Baltica Cup	75043	5370955	0	1	1	16.12.1999	Czech Republic	431257	Sweden	431267	2	0
GLO	Baltica Cup	75043	5370962	0	5	1	20.12.1999	Sweden	431267	Finland	431258	0	0
GLO	Baltica Cup	80342	5626439	0	3	1	22.12.2001	Russia	431265	Czech Republic	431257	0	4
United States	NHL	77274	5423490	0	0	0	4.10.1996	St. Louis	431118	Colorado	431103	4	2
United States	NHL	77274	5423491	0	0	0	4.10.1996	Edmonton	431106	Buffalo	431100	4	3
United States	NHL	77274	5423492	0	0	0	4.10.1996	Los Angeles	431109	NY Islanders	431112	1	0
United States	NHL	77274	5423493	0	0	0	5.10.1996	Boston	431099	NY Rangers	431113	4	4
United States	NHL	77274	5423494	0	0	0	5.10.1996	Montreal	431110	Ottawa	431114	3	3
SVK	Extraliga	74530	5478179	0	1	0	7.9.1999	Liptovskďř" Mikulďř"	433217	Slovan Bratislava	433214	2	4
SVK	Extraliga	74530	5478180	0	1	0	7.9.1999	Spďř"skďř" Novďř" Ves	433220	Zvolen	433221	0	4

Obrázek 2.1: Ukázka datového souboru - první část

Goals Home in Detail	Goals Home Times													
4	16	19	43	58	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	10	28	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
4	18	33	36	46	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
5	17	23	25	29	60	-1	-1	-1	-1	-1	-1	-1	-1	-1
1	37	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	33	42	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	1	56	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
3	16	27	43	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
1	40	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	48	54	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
1	57	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
3	12	28	60	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
4	19	22	51	56	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	16	55	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	27	40	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
4	10	19	26	44	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
4	9	24	33	44	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
1	29	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
4	17	20	29	34	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
3	1	30	40	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	12	51	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

Obrázek 2.2: Ukázka datového souboru - druhá část

Dále v textu budou používány zkratky pro názvy národních soutěží. Jejich přehled je v Tabulce 2.1.

Název soutěže	Zkratka
Kanadsko-americká Národní Hokejová Liga	NHL
Česká extraliga	CZE
První česká liga	CZE1
Švédská Elitserien	SWE
Finská SM-liiga	FIN
Slovenská extraliga	SVK
Mezinárodní utkání	INT

Tabulka 2.1: Přehled zkratk názvů soutěží

Data pro jednotlivé národní ligy byla sbírána v různých časových intervalech. Jsou zde tedy rozdíly v počtech utkání mezi soutěžemi. Přehled od kdy do kdy byla data pro jednotlivé soutěže sbírána, počty utkání a jejich podíl v celkovém datovém balíku je zobrazen v Tabulce 2.2.

Nejvíce dat bylo nasbíráno v kanadsko-americké NHL, kde byla data zaznamenávána od sezony 1996/1997 až do sezony 2010/2011². Celkově včetně playoff se jedná o 14 746 utkání.

Pro českou extraligu začala být data sbírána také v sezoně 1996/1997 nepřetržitě až do sezony 2010/2011 včetně zápasů playoff. Od roku 2001 do roku 2011 byly také zaznamenávány barážové zápasy o udržení v extralize³. Celkový počet zaznamenaných zápasů z české extraligy je 5 866.

Dále byla sledována také druhá česká nejvyšší soutěž. Data z první české hokejové ligy včetně utkání playoff jsou k dispozici od sezony 2000/2001 do sezony 2010/2011. Datový soubor se skládá z 3 495 utkání.

Čtvrtou nejpočetnější skupinu tvoří data z mezistátních utkání. Jedná se o zápasy národních týmů v časovém rozmezí od roku 1997 do roku 2011. Utkání byla hrána v rámci světových šampionátů, olympijských her a turnajů evropské hokejové tour. Celkově se jedná o data z 864 zápasů.

Pro slovenskou extraligu ledního hokeje a švédskou Elitserien byla data sbírána pouze ve dvou sezonách a to v sezoně 1999/2000 a 2000/2001. V obou případech byla data sbírána i pro zápasy ve vyřazovacích bojích. Soubor dat pro švédskou hokejovou soutěž obsahuje 573 utkání a pro slovenskou extraligu 432 zápasů.

²V NHL byla sezona 2004/2005 zrušena z důvodu chybějící kolektivní smlouvy.

³S výjimkou sezony 2006/2007, kdy bylo z extraligy vyloučeno mužstvo HC Vsetín a postup do nejvyšší české hokejové soutěže slavil vítěz finále 1.ligy HC Slovan Ústí nad Labem.

Nejméně dat bylo nasbíráno v nejvyšší finské hokejové lize, jelikož data byla sbírána pouze v sezoně 1999/2000. Opět včetně zápasů playoff. Celkový počet zaznamenaných utkání je 337.

Liga	Od roku	Do roku	Počet zápasů	Podíl z celkem
NHL	1996	2011	14 746	56,04 %
CZE	1996	2011	5 866	22,29 %
CZE1	2000	2011	3 495	13,28 %
INT	1997	2011	864	3,28 %
SWE	1999	2001	573	2,19 %
SVK	1999	2001	432	1,64 %
FIN	1999	2000	337	1,28 %
Celkem			26 313	100 %

Tabulka 2.2: Přehled soutěží a počtu utkání

2.2 Rozdíly v bodovém ohodnocení

Bodová ohodnocení se týká utkání, která byla odehrána v tzv. **základní části sezony**. Po této části sezony se tabulka rozdělí na dvě části. Týmy z horní části tabulky postupují do playoff a pro týmy v druhé části tabulky zpravidla sezona končí. Určení týmu, který by měl po sezoně sestoupit do nižší soutěže je velmi specifické, jelikož liga jako například NHL je uzavřená soutěž, takže týmům nehrozí sestup do jiné soutěže. Naproti tomu v české extralize se hraje tzv. **baráž** o účast v příštím ročníku extraligy. Od sezony 1996/1997 do sezony 2006/2007 hrál baráž automaticky tým, který skončil po základní části sezony na posledním místě tabulky. Od sezony 2007/2008 hrají poslední čtyři týmy tzv. **playout**, kde se utkají mezi sebou čtyřikrát každý s každým. Poslední z této minitabulky (započítávají se do ní body z průběhu celé sezony) hraje baráž s vítězem první ligy o účast v příštím ročníku extraligy.

Bodování se mezi jednotlivými národními ligami liší, jelikož každá liga má svá vlastní pravidla, jimiž se řídí. Ale bodové ohodnocení se v minulosti měnilo i v rámci jednotlivých lig v závislosti na vývoji pravidel moderního hokeje. Z historického hlediska byl až do konce devadesátých let ve všech zkoumaných ligách zápas ukončen po odehrání základní hrací doby. To se ale změnilo na přelomu tisíciletí.

Jako první ze zmíněných soutěží zavedli změnu v bodování švédská nejvyšší soutěž Elitserien a kanadsko-americká NHL. V NHL bylo zavedeno tzv. prodloužení, které nastává v případě, že je stav utkání po základní hrací době (60 minut) nerozhodný, již v sezoně 1982/1983. Ve švédské soutěži zavedli odlišný bodovací systém na rozdíl od NHL. Za výhru v základní hrací době získává tým 3 body. Pokud je po základní hrací době stav utkání nerozhodný, tak nastává pětiminutové prodloužení. Jestliže se jednomu z týmů podaří v prodloužení skórovat, tak utkání končí. Vítěz získává 2 body, poražený 1 bod. V případě, že nepadne v prodloužení branka, tak přichází na řadu série samostatných nájezdů, které rozhodnou o vítězi utkání. Vítěz získává 2 body a poražený 1 bod. Utkání tak má vždy svého vítěze. Přehled bodovacích systémů je v Tabulce 2.3.

Liga	Od roku	Do roku	Počet sezon	V/Vp/Pp/R/P
NHL	1996	1999	3	2/2/0/1/0
NHL	1999	2004	5	2/2/1/1/0
NHL	2004	2011	7	2/2/1/-/0
SWE	1999	2001	2	3/2/1/-/0

Tabulka 2.3: Přehled bodovacích systémů soutěží NHL a Elitserien

V NHL bylo zavedeno odlišné bodování. V případě výhry v základní hrací době získává vítězný tým 2 body a poražený žádný. Pokud je po šedesáti minutách stav nerozhodný, přichází prodloužení. Až do sezony 1998/1999 byl tým, který vstřelil v prodloužení branku jako první odměněn dvěma body a poražený nezískal žádný. V následující sezoně, ale přišla změna v bodovém odměňování a tým, kterému se povedlo vstřelit branku získal opět 2 body, ale poražený obdržel 1 bod. Na rozdíl od švédské Elitserien při nerozhodném stavu po prodloužení nepřicházejí samostatné nájezdy, nýbrž je utkání ukončeno s tím, že každý tým získává jeden bod. Tato pravidla platila v NHL 5 let. Poté vedení kanadsko-americké NHL také přistoupilo k zavedení samostatných nájezdů v případě, že stav utkání je nerozhodný i po prodloužení. Body jsou rozděleny stejně jako v případě rozhodnutí zápasu během prodloužení. Vítěz tedy získává 2 body a poražený 1 bod. Toto bodové ohodnocení platí až do současnosti.

Nejvyšší česká a slovenská liga následovaly tuto změnu v bodovém ohodnocení s ročním zpožděním. Přehled těchto bodovacích systémů je v Tabulce 2.4. Od sezony 2000/2001 se tedy zavedlo prodloužení, ve kterém vítěz získal 2 body a poražený 1 bod. V případě remízy i po pětiminutovém prodloužení bylo utkání ukončeno a oba týmy získaly po jednom bodu. Pokud byl zápas již rozhodnut v základní hrací době, tak vítěz získal 3 body a poražený žádný. Datový soubor obsahuje data z následujících sezon už jen pro

Liga	Od roku	Do roku	Počet sezon	V/Vp/Pp/R/P
CZE	1996	2000	4	2/-/-/1/0
CZE	2000	2006	6	3/2/1/1/0
CZE	2006	2011	5	3/2/1/-/0
SVK	1999	2000	1	2/-/-/1/0
SVK	2000	2001	1	3/2/1/1/0

Tabulka 2.4: Přehled bodovacích systémů české a slovenské extraligy

českou extraligu. Zde došlo ještě k jedné změně v sezoně 2006/2007. A to ke stejné jako v případě NHL. Byla zrušena remíza. Pokud nepadla branka v prodloužení, tak přichází na řadu samostatné nájezdy. Vítěz opět získává 2 body a poražený jeden bod. Bodové ohodnocení za vítězství v základní hrací době zůstalo stejné. Toto bodové ohodnocení je platné až do současnosti.

Ve druhé české nejvyšší soutěži došlo ke stejným změnám jako v nejvyšší soutěži. Přehled těchto bodovacích systémů je v Tabulce 2.5. Jen tyto změny přicházely s ročním zpožděním. Od sezony 2001/2002 se hrálo prodloužení a od sezony 2007/2008 byly zavedeny po prodloužení samostatné nájezdy a tedy byla zrušena remíza.

Liga	Od roku	Do roku	Počet sezon	V/Vp/Pp/R/P
CZE1	2000	2001	1	2/-/-/1/0
CZE1	2001	2007	6	3/2/1/1/0
CZE1	2007	2011	4	3/2/1/-/0

Tabulka 2.5: Přehled bodovacích systémů české první ligy

Z finské nejvyšší soutěže máme záznamy o utkáních pouze ze sezony 1999/2000. V této sezoně bylo bodování následující. Zápas byl ukončen po základní hrací době, nebylo tedy prodloužení, ani samostatné nájezdy. Vítěz získal 2 body a poražený žádný. V případě remízy si týmy spravedlivě rozdělily po jednom bodu. Přehled bodovacího systému je v Tabulce 2.6.

Liga	Od roku	Do roku	Počet sezon	V/Vp/Pp/R/P
FIN	1999	2000	1	2/-/-/1/0

Tabulka 2.6: Přehled bodovacího systému finské SM-liigy

2.3 Základní statistiky

2.3.1 Vstřelené góly

V této části jsou provedeny základní popisné statistiky pro jednotlivé ligové soutěže. Výsledky jsou shrnuty v tabulkách a grafech. Jako ukázkový vzorek byly zvoleny základní statistiky celého utkání. Základní statistiky pro jednotlivé třetiny jsou uvedeny v příloze diplomové práce na CD. Data byla rozdělena podle národních soutěží. V Tabulce 2.7 je přehled základních statistik vstřelených gólů pro jednotlivé soutěže.

Liga	Počet zápasů	Průměr gólů	Sm. odchylka gólů	Medián gólů
CZE	5 866	5,48	1,77	5
CZE1	3 495	5,46	1,82	5
FIN	337	6,15	1,94	6
SWE	573	5,88	1,88	6
NHL	14 746	5,37	1,65	6
SVK	432	6,20	2,17	6
INT	864	5,62	2,10	5,5
Celkem	26 313	5,43	1,74	5

Tabulka 2.7: Základní statistiky vstřelených gólů pro jednotlivé soutěže

Zápasy jsme také zkoumali také z pohledu domácího a hostujícího týmu. V Tabulce 2.8 jsou zobrazeny základní statistiky vstřelených gólů pro domácí tým v rámci jednotlivých soutěží.

Liga	Počet zápasů	Počet gólů	Průměr gólů	Sm. odchylka gólů	Medián gólů
CZE	5866	18 325	3,12	1,84	3
CZE1	3495	10 917	3,12	1,91	3
FIN	337	1 154	3,42	2,04	3
SWE	573	1 950	3,40	1,98	3
NHL	14 746	41 469	2,81	1,67	3
SVK	432	1 622	3,75	2,37	3
INT	864	2 922	3,38	2,23	3
Celkem	26 313	78 801	2,97	1,80	3

Tabulka 2.8: Základní statistiky vstřelených gólů pro domácí tým

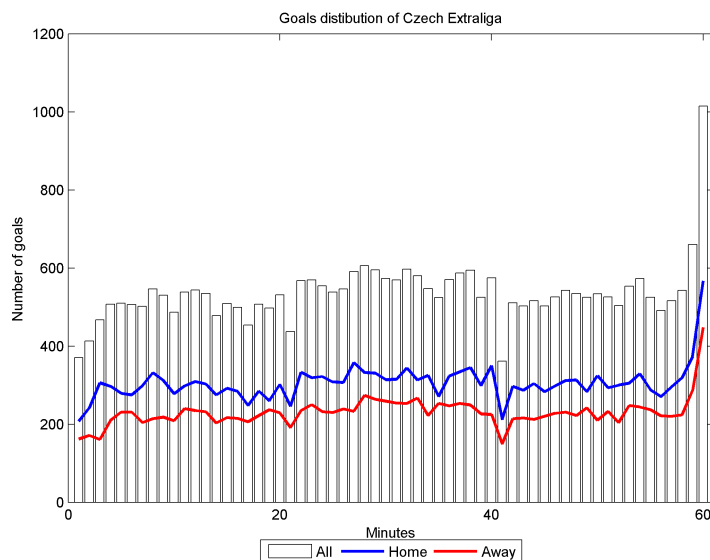
V následující tabulce jsou zobrazeny základní statistiky vstřelených gólů pro hostující tým rozdělené podle soutěží.

Liga	Počet zápasů	Počet gólů	Průměr gólů	Sm. odchylka gólů	Medián gólů
CZE	5 866	13 813	2,35	1,61	2
CZE1	3 495	8 174	2,34	1,63	2
FIN	337	920	2,73	1,77	2
SWE	573	1 418	2,47	1,65	2
NHL	14 746	37 749	2,56	1,62	2
SVK	432	1 055	2,44	1,70	2
INT	864	1 937	2,24	1,80	2
Celkem	26 313	64 802	2,46	1,63	2

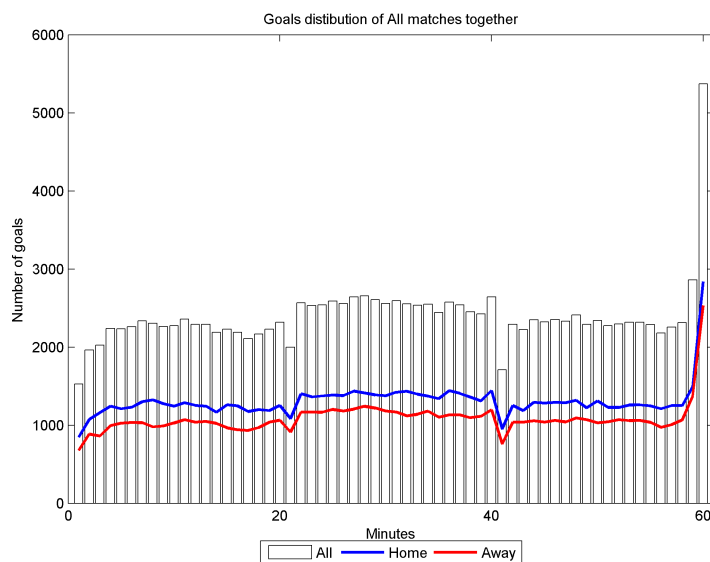
Tabulka 2.9: Základní statistiky vstřelených gólů pro hostující tým

2.3.2 Góly po minutách

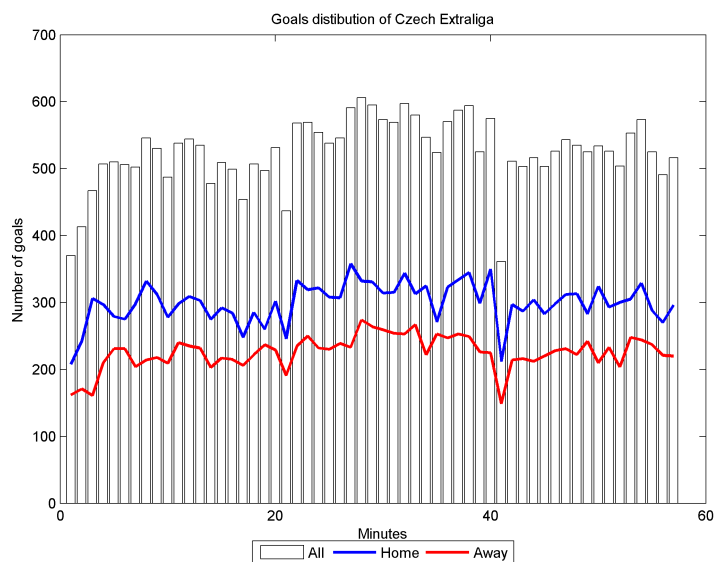
Zkoumali jsme rozložení počtu branek v jednotlivých minutách. Zde jsme narazili na zajímavý jev, kdy nejvíce branek padá v průběhu druhé třetiny. Neuvažujeme závěrečné 3 minuty utkání, jelikož v nich může docházet k ne-standardní herní situaci tzv. **hry bez brankáře**. Jde o odvolání brankáře a jeho náhradu hráčem do pole. Tým se tedy snaží těžit z početní výhody jednoho hráče navíc a vstřelit branku. Ve všech dostupných ligách se tento jev objevil. Pro ukázkou jsme zvolili data z české extraligy a souhrnná data ze všech soutěží dohromady. Rozdělení branek po minutách pro tyto soutěže je zobrazeno v následujících obrázcích. Grafy pro ostatní soutěže jsou přiloženy v příloze diplomové práce na CD.



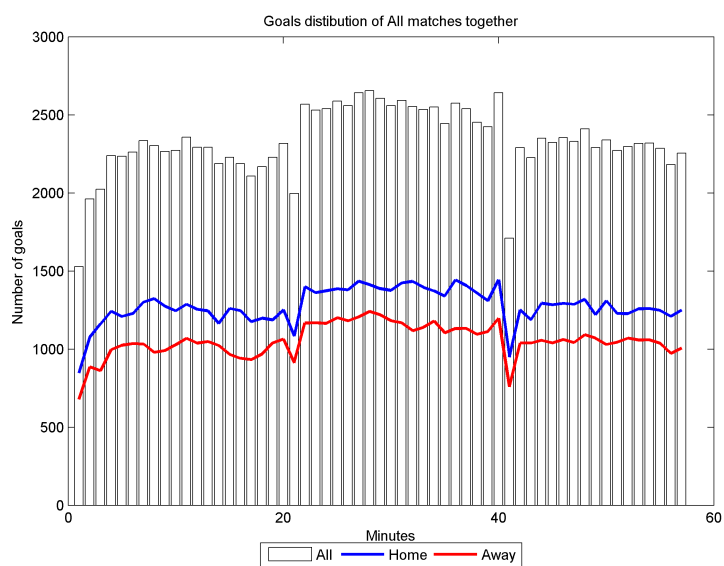
Obrázek 2.3: Rozdělení branek do 60. minuty v české extralize



Obrázek 2.4: Rozdělení branek do 60.minuty ve všech soutěžích



Obrázek 2.5: Rozdělení branek do 57.minuty v české extralize



Obrázek 2.6: Rozdělení branek do 57.minuty ve všech soutěžích

2.3.3 Vyrovnanost soutěží

Zjišťovali jsme také vyrovnanost jednotlivých soutěží. Vypočítali jsme pravděpodobnosti výhry domácího celku, remízy a výhry hostujícího týmu (prohry domácího celku). Používali jsme výsledek v normální hrací době. Pokud v zápase měly týmy po 60 minutách stejný počet branek a nastalo prodloužení, tak se tento zápas započítal jako remízový bez ohledu na výsledek prodloužení.

Pravděpodobnost výhry domácího celku se pohybovala nad hranicí 50 %. Jen ve finské lize a NHL se domácí týmy dostaly pod tuto hranici vyhraných utkání. Pravděpodobnost remízy se pohybovala přibližně od 0,1 do 0,2. Pouze u NHL byla zjištěna vyšší pravděpodobnost remízy a to 0,224. V případě výhry hostujícího týmu byla opět nejvyšší pravděpodobnost zjištěna u NHL 0,343. V ostatních soutěžích je pravděpodobnost výhry hostujícího týmu v rozmezí od 0,259 do 0,320.

Pomocí Chí-kvadrát testu jsme otestovali hypotézu: pravděpodobnost výhry v základní hrací době pro domácí tým je v NHL stejná jako ve všech ostatních soutěžích. Tuto hypotézu jsme zamítli na hladině významnosti $\alpha = 5\%$, jelikož p-hodnota je rovna nule. Výsledek testu je zobrazen v Tabulce 2.10. Popis tohoto testu lze najít v [8] v Kapitole 4.10 a 4.11.

H_0	$\pi_{NHL} = \pi_{Ostatni}$ ⁴	
Vyrovnanost soutěží	NHL	Ostatní soutěže
Počet vítězství domácích	6 387	6 130
Počet neproher hostů	8 359	5 437
Chí-kvadrát statistika	243,6	
P-hodnota	0	
Hladina významnosti	5 %	
Nulová hypotéza	zamítnuta	

Tabulka 2.10: Přehled výsledku Chí-kvadrát testu vyrovnanosti NHL

Z tohoto můžeme soudit, že v kanadsko-americké NHL nehraje domácí prostředí tak významnou roli jako v ostatních soutěžích. Jelikož domácí tým vyhrává s mnohem nižší pravděpodobností než je tomu v jiných ligách. Přehled pravděpodobností pro jednotlivé ligy je zobrazen v Tabulce 2.11

⁴ π označuje rozdělení náhodné veličiny

Liga	Pravděpodobnost výhry domácích	Pravděpodobnost remízy	Pravděpodobnost výhry hostů
CZE	0,522	0,204	0,274
CZE1	0,520	0,188	0,292
FIN	0,478	0,202	0,320
SWE	0,590	0,099	0,311
NHL	0,433	0,224	0,343
SVK	0,579	0,162	0,259
INT	0,583	0,144	0,273
Celkem	0,476	0,208	0,316

Tabulka 2.11: Pravděpodobnosti výhry, remízy a prohry domácího celku

Zaznamenávali jsme i počet utkání, ve kterých došlo k rozhodnutí zápasu až v prodloužení a určili jsme poměr vítězných prodloužení mezi domácím a hostujícím týmem. Tyto údaje jsou v Tabulce 2.12.

Liga	Počet prodloužení	Výhra domácí	Výhra hosté	Ppst výhry domácí	Ppst výhry hosté
CZE	746	374	372	0,501	0,499
CZE1	529	299	230	0,565	0,435
FIN	16	9	7	0,563	0,437
SWE	57	29	28	0,509	0,491
NHL	2 227	1 140	1 087	0,511	0,489
SVK	25	19	6	0,760	0,240
INT	64	32	32	0,500	0,500
Celkem	3 664	1 902	1 762	0,52	0,48

Tabulka 2.12: Přehled údajů o prodlouženích

Výhodu domácího prostředí můžeme také prozkoumat porovnáním pravděpodobností výher domácích oproti výhrám hostů. V Tabulce 2.13 jsou vypočteny podmíněné pravděpodobnosti výher domácího týmu o jeden, dva, tři,... a o šest a více branek za podmínky, že domácí tým zvítězil. Stejně podmíněné pravděpodobnosti jsme vypočetli také pro výhry hostujícího týmu, které jsou uvedeny v Tabulce 2.14.

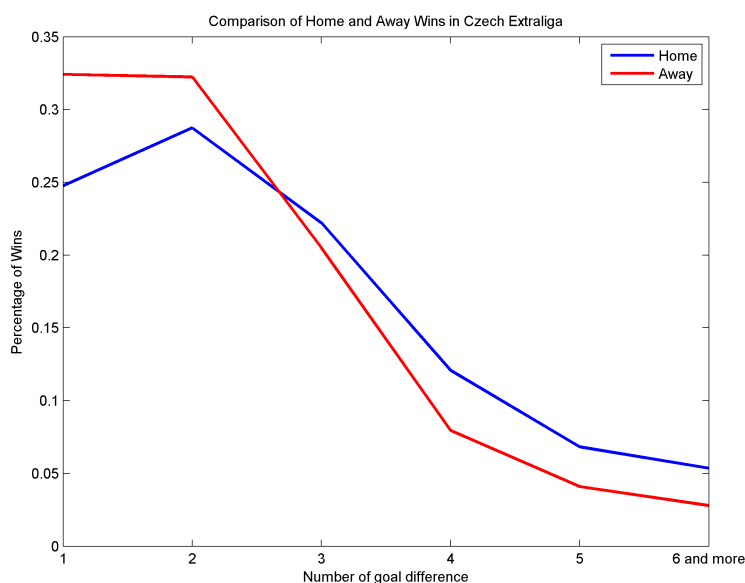
Liga/Rozdíl golů	1	2	3	4	5	6 a více
CZE	0,248	0,287	0,222	0,121	0,068	0,054
CZE1	0,229	0,307	0,209	0,117	0,066	0,073
FIN	0,217	0,193	0,273	0,137	0,068	0,112
SWE	0,231	0,240	0,272	0,112	0,071	0,074
NHL	0,302	0,288	0,245	0,099	0,039	0,027
SVK	0,180	0,224	0,196	0,168	0,084	0,148
INT	0,232	0,222	0,188	0,135	0,067	0,155
Celkem	0,270	0,284	0,232	0,111	0,053	0,050

Tabulka 2.13: Pravděpodobnosti výhry domácího týmu o 1,2,3,4,5 a o šest a více branek

Liga/Rozdíl golů	1	2	3	4	5	6 a více
CZE	0,324	0,322	0,205	0,080	0,041	0,028
CZE1	0,330	0,316	0,219	0,075	0,042	0,019
FIN	0,287	0,333	0,204	0,093	0,037	0,046
SWE	0,281	0,303	0,258	0,101	0,028	0,028
NHL	0,328	0,296	0,239	0,082	0,031	0,023
SVK	0,268	0,330	0,161	0,134	0,063	0,045
INT	0,284	0,250	0,212	0,097	0,089	0,068
Celkem	0,324	0,303	0,228	0,082	0,037	0,025

Tabulka 2.14: Pravděpodobnosti výhry hostujícího týmu o 1,2,3,4,5 a o šest a více branek

Podmíněná pravděpodobnost výhry hostujícího týmu o jeden gól je ve všech soutěžích větší než výhra domácího týmu o jeden gól. Tento poměr se s rostoucím počtem branek obrací ve prospěch domácího týmu, kde u výher o šest a více branek je pravděpodobnost dvakrát vyšší u domácího týmu oproti hostujícímu týmu. Jako vzor jsme vybrali českou extraligu. Na Obrázku 2.7 jsou zobrazeny podíly výher pro domácí a hostující tým v závislosti na počtu gólů, o které tým vyhrál.



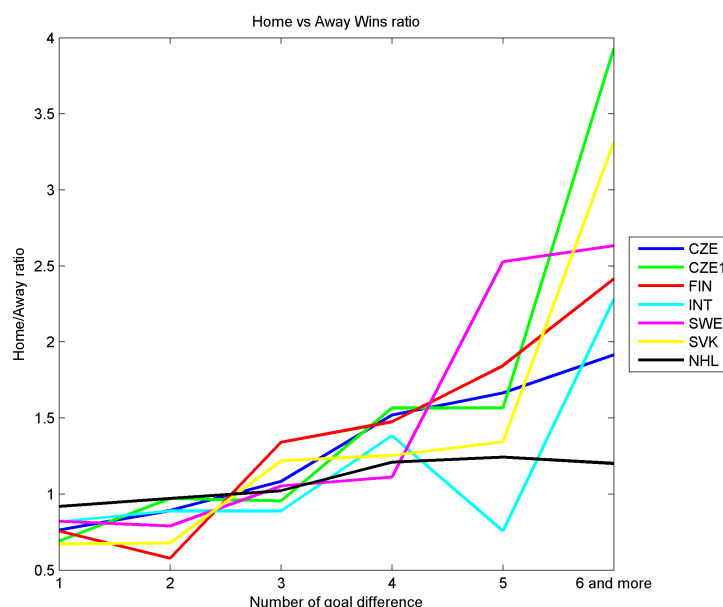
Obrázek 2.7: Porovnání podmíněných pravděpodobností výher domácího a hostujícího týmu v závislosti na počtu gólů, o které vyhrály v české extralize

Přehled poměru podmíněných pravděpodobností výher o jeden, dva, tři, čtyři, pět a šest a více branek domácího týmu oproti výhrám hostujícího týmu je zobrazen v Tabulce 2.15.

Liga/Rozdíl golů	1	2	3	4	5	6 a více
CZE	0,76	0,89	1,08	1,52	1,67	1,91
CZE1	0,69	0,97	0,95	1,57	1,57	3,93
FIN	0,76	0,58	1,34	1,48	1,84	2,41
SWE	0,82	0,79	1,05	1,11	2,53	2,63
NHL	0,92	0,97	1,02	1,21	1,24	1,20
SVK	0,67	0,68	1,22	1,25	1,34	3,32
INT	0,82	0,89	0,89	1,38	0,76	2,28
Celkem	0,83	0,94	1,02	1,35	1,46	2,00

Tabulka 2.15: Poměr mezi podmíněnými ppstmi výher domácího a hostujícího týmu o 1,2,3,4,5 a o šest a více branek

Zajímavé je, že například v české extralize se poměr mezi podmíněnými pravděpodobnostmi výher domácích a hostů zvyšuje z přibližně 0,76 (v případě výher o jednu branku) až do 1,92 (v případě výher o šest a více branek). Ale u NHL je toto rozmezí pouze od 0,92 do 1,20. Což opět vypovídá o tom, že domácí prostředí nemá v NHL na výsledek utkání takový vliv. Srovnání soutěží pomocí poměru mezi podmíněnými pravděpodobnostmi výher domácích a hostujících je zobrazeno na Obrázku 2.8.



Obrázek 2.8: Srovnání soutěží pomocí poměru mezi podmíněnými pravděpodobnostmi výher domácích a hostů

3 Averze ke ztrátě

Averze ke ztrátě (Loss aversion) je ekonomický pojem z oblasti teorie rozhodování. Popisuje náklonnost jednotlivců k silnějšímu vnímání ztráty než stejného zisku. Lidé mají tendenci se více vyhýbat ztrátám (za cenu nižšího zisku), než riskovat ve prospěch vyššího zisku (s možností větší ztráty) [4].

Při rozhodování lidí za bezrizikové volby dochází k zvláštním řetězcům chování. Skutečnost, že lidé často požadují mnohem více za odevzdání předmětu než by byli schopni zaplatit pro jeho získání. Tento jev se nazývá **nadační efekt** (endowment effect). Druhou zvláštností je silná tendence jedinců k setrvání ve stávajícím stavu, jelikož nevýhody opuštění tohoto stavu se jeví větší než výhody. Tento efekt nazýváme **náklonnost k současnému stavu** (status quo bias). V následujících podkapitolách jsou jednotlivé pojmy vysvětleny a popsány na příkladech.

3.1 Nadační efekt

Ztráta užitku spojená se vzdáním se hodnotného předmětu je větší než získaný užitek spojený s přijetím tohoto předmětu. Tato nesrovnalost je označována jako nadační efekt. Bylo provedeno několik experimentů na prokázání tohoto jevu. Následující experiment provedli a popsali D. Kahneman, J. L. Knetsch a R.H. Thaler ve svém odborném článku [3].

77 studentů bylo náhodně rozděleno do tří skupin. První skupina, Prodejci, byla obdarována hrníčkem na kávu a dotázána, zda by byli ochotni prodat tento hrníček za jakoukoliv částku na škále od \$0,25 do \$9,25. Druhá skupina Kupců byla dotázána, zda by byli ochotni koupit daný hrníček ve stejné množině cen. Třetí skupina, nazývaná Voliči, nedostala hrníček, ale mohla si zvolit na každé úrovni z cenové škály, zda si vyberou hrníček nebo dané množství peněz.

Všimněme si, že Prodejci a Voliči jsou v objektivně stejné situaci. Rozhodují se na každé cenové úrovni mezi hrníčkem a daným množstvím peněz. Nicméně Voliči se chovali mnohem více jako Kupci než jako Prodejci. Průměrné zjištěné ceny byly: Prodejci: \$7,12; Voliči: \$3,12; Kupci: \$2,87. Ceny prodávajících jsou přibližně 2,5krát větší než ceny, za které jsou kupující ochotni hrneček koupit. Tato nesrovnalost je tedy nazývána nadační efekt.

3.2 Náklonnost k současnému stavu

Zachování stávajícího stavu je jednou z možností týkajících se problémů rozhodování. Na Obrázku 3.1 je zobrazeno několik možností. Pokud se osoba, která se rozhoduje, nachází v referenčním bodě t , tak nedělá rozdíl mezi x a y . Preferuje ale x před y z referenčního bodu x , a y před x z bodu y . Obrázek 3.1 a následující příklad náklonnosti k současnému stavu je možné najít v odborném článku [4].

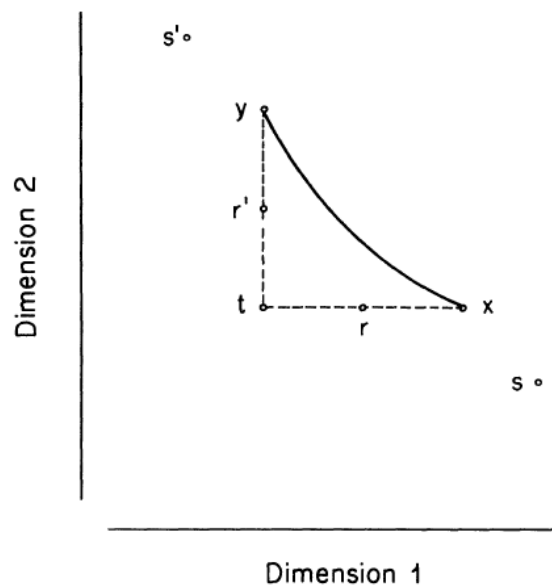


FIGURE II
Multiple Reference Points for the Choice Between x and y

Obrázek 3.1: Množina referenčních bodů pro volbu mezi x a y

Vysokoškoláci ze dvou různých tříd byli požádáni o vyplnění krátkého dotazníku. Studenti z první třídy dostali jako okamžitou odměnu zdobený hrníček. Studenti z druhé třídy obdrželi velkou tabulku švýcarské čokolády. Na konci lekce byl oběma skupinám studentů ukázán alternativní dárek a byla jim povolena možnost vyměnit dárek, který dostali, za ten druhý. Museli jen zvednout kartičku s nápisem "Obchod". Ačkoliv transakční náklady spojené s výměnou byly malé, tak přibližně 90 % účastníků zůstalo u dárku, který obdrželi.

V článku [3] autoři zmiňují jiný experiment zkoumající náklonnost k současnému stavu. Státy New Jersey a Pennsylvania nabízejí volbu mezi dvěma

typy automobilového pojištění: levnější pojistka, která omezuje právo žaloby, a dražší pojistka, která podporuje neomezené právo žaloby. Motoristům v New Jersey je jako základní nabízena levnější pojistka s možností získat neomezené právo žaloby za příplatek. Nicméně v Pennsylvánii je základní možností dražší typ pojistky s možností změny na levnější typ. Dvě skupiny byly dotázány na výběr mezi těmito dvěma alternativami pojistek. První skupině byl prezentován plán z New Jersey a druhé plán z Pennsylvánie. Z těch, kterým byl nabídnut plán z New Jersey, si jen 23 % zvolilo zakoupení práva žaloby. Zatímco 53 % ze skupiny, které byl nabídnut plán z Pennsylvánie, zůstalo u zachování tohoto práva.

3.3 Averze ke ztrátě

Averze ke ztrátě poukazuje na tendenci lidí silně upřednostňovat vyhnutí se ztrátě oproti získání výnosu. Nadační efekt a efekt náklonnosti k současnému stavu se projevují v asymetrickém tvaru S-křivky užitekovej funkce, která je nazývána averze ke ztrátě. Existující empirické studie naznačují, že poměr strmostí užitekovej funkce mezi dvěma oblastmi pro malé nebo průměrné výnosy a ztráty peněz je kolem 2:1 [4]. Užitekovej funkce je naznačena v Obrázku 3.2.

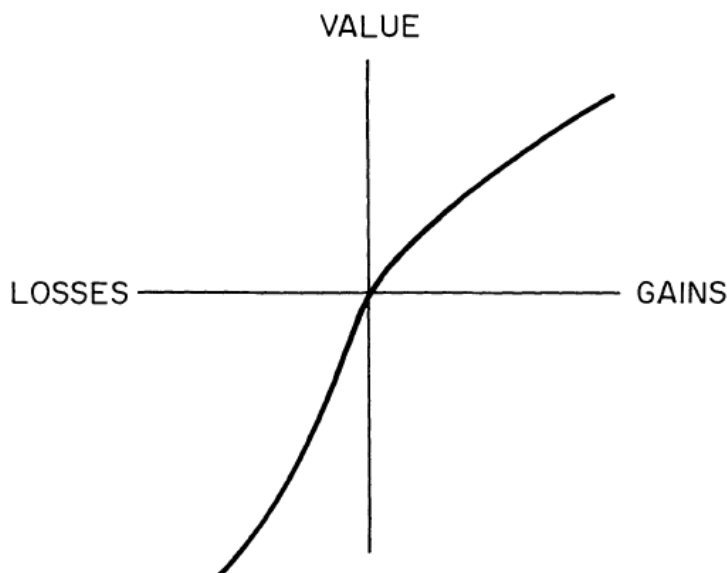


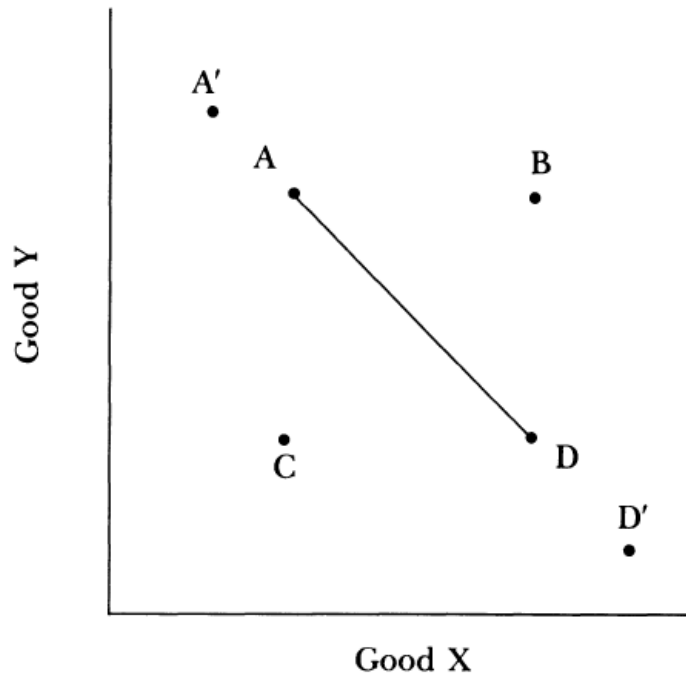
FIGURE I
An Illustration of a Value Function

Obrázek 3.2: Typická užitekovej funkce

Možnosti obchodu a jiné transakce jsou hodnoceny jako zisky a ztráty vztahující se k neutrálnímu referenčnímu bodu. Na Obrázku 3.3 je vysvětlen přístup osoby rozhodující se mezi dvěma možnostmi.

Osoba má na výběr mezi stavem A , kde má více zboží Y a méně zboží X , a stavem D , kde má naopak více zboží X a méně zboží Y . Na obrázku jsou zobrazeny čtyři různé referenční body. Osoba volí mezi dvěma zisky, jestliže referenční bod je C . Nebo naopak mezi dvěma ztrátami, jestliže referenční bod je B , a dvěma různými výměnami, jestliže referenční body

Multiple reference points for the choice between A and D



Obrázek 3.3: Množina referenčních bodů pro volbu mezi body A a D

jsou A , resp. D . Řekněme, že zboží Y je hrneček a zboží X jsou peníze. Referenční body pro Prodejce a Kupce z předchozího příkladu v Podkapitole 3.1 jsou A a C . Averze k riziku vyjadřuje, že rozdíl mezi stavem, kdy vlastníme hrneček a kdy ho nevlastníme je větší z bodu A než z C , což vysvětluje rozdílné množství peněz, které subjekty přikládají za těchto podmínek hrnečku.

Uvedeme si příklad z [3]. Představme si, že jako část profesního tréninku nám byla přidělena brigáda. Tento trénink končí a my musíme hledat zaměstnání. Zvažujeme dvě možnosti. Ve většině aspektů jsou jako naše brigáda kromě množství osobních kontaktů a času stráveným dojížděním do a z práce. K porovnání těchto dvou prací a té současné slouží Tabulka 3.1.

Práce	Kontakt s ostatními	Čas dojíždění
Současná práce	izolován po dlouhé úseky	10 min.
Práce A	kontakt s ostatními je limitován	20 min.
Práce D	mírně společenský	60 min.

Tabulka 3.1: Přehled možností prací

Možnosti A a D jsou hodnoceny z referenční pozice současné práce, která je lepší z hlediska dojíždění, ale s nízkým počtem osobních kontaktů (bod A' v Obrázku 3.3). V další verzi problému máme stejné pracovní možnosti, ale referenční práce zahrnuje "mnohem příjemnější společenské prostředí a 80 minut dojíždění denně do práce", což je bod D' na Obrázku 3.3. V první verzi problému si práci A vybralo 70 % dotázaných. Ve druhé verzi jen 33 %. Z toho plyne, že lidé jsou citlivější na oblast, ve které ztrácejí vzhledem k jejich referenčnímu bodu.

Druhý příklad averze ke ztrátě je opět z odborného článku [3]. Dotázaným bylo řečeno, že byli vystaveni vzácné smrtelné nemoci a nyní čelí šanci 0,1 %, že do čtrnácti dnů zemřou bezbolestnou smrtí. Musejí se rozhodnout, kolik by byli schopni zaplatit za vakcínu, aby byla okamžitě zakoupena. Stejně osoby byly také dotázány, jakou kompenzaci by požadovaly za účast v lékařském experimentu, ve kterém by čelily 0,1% šanci rychlé a bezbolestné smrti. Většina dotázaných byla schopna zaplatit za vakcínu částku o řád vyšší než by požadovala za účast v lékařském experimentu.

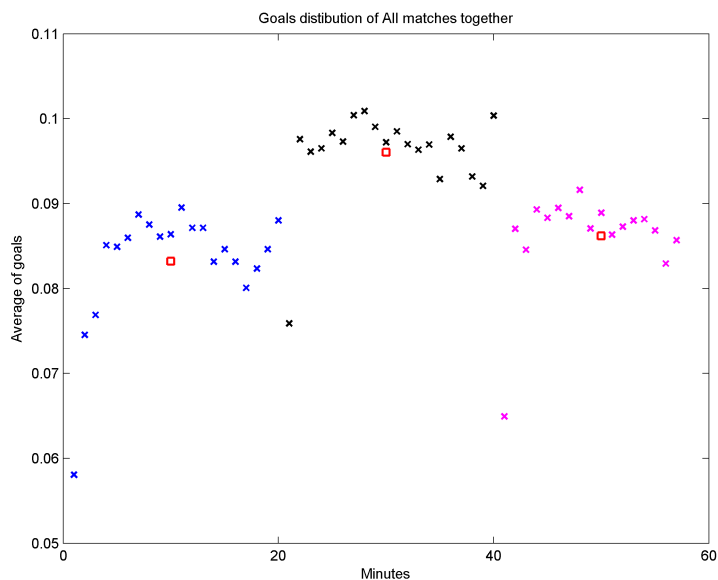
Averzí ke ztrátě v oblasti sportu se taktéž zabývá několik odborných článků. D. G. Pope a M. E. Schweitzer publikovali v roce 2011 článek o averzi ke ztrátě v golfu [7]. Další dva články o averzi ke ztrátě jsou spjaty s ledním hokejem, konkrétně s NHL. A. N. Banerjee, J. F. Swinnen a A. Weersink se zabývali změnami pravidel a strategií týmů [1] a D. Coates s B. R. Humphreyssem zkoumali návštěvnost utkání v souvislosti s vyrovnaností týmů [2]. Dále je v přípravě článek o averzi k riziku a ztrátě v oblasti sportu, jehož autory jsou L. J. A. Lenten, J. Libich a P. Stehlík [6].

V této práci jsme formulovali dvě hypotézy, které by mohly poukázat na projevení se averze ke ztrátě v ledním hokeji. První hypotéza se týká vyššího počtu vstřelených branek v průběhu druhé třetiny. Tento jev můžeme vysvětlit pomocí averze ke ztrátě, jelikož týmy nastupují do poslední třetiny za určitého stavu a v určité pozici. Pokud tým nastupuje v pozici vedoucího týmu, tak se bude pravděpodobně více soustředit na obrannou hru na úkor útočení a možnosti své vedení ještě navýšit. Vyhýbá se tedy ztrátě bodů za vítězství na úkor vylepšení své pozice v utkání. Tudíž ve třetí třetině padne méně branek. Případem, kdy je stav po dvou třetinách remízový, se zabývá další hypotéza. Zkoumali jsme průměrný počet gólů vstřelených ve třetí třetině u zápasů, kde je stav neremízový, resp. remízový. Předpokládáme, že v remízových utkáních se týmy snaží vyhnout ztrátě jednoho bodu za remízu raději než možnosti získat bonusový bod v případě vstřelení branky a vítězství v základní hrací době. Opět se tedy jedná o jev vysvětlitelný pomocí averze ke ztrátě.

4 Rozdíl gólů mezi třetinami

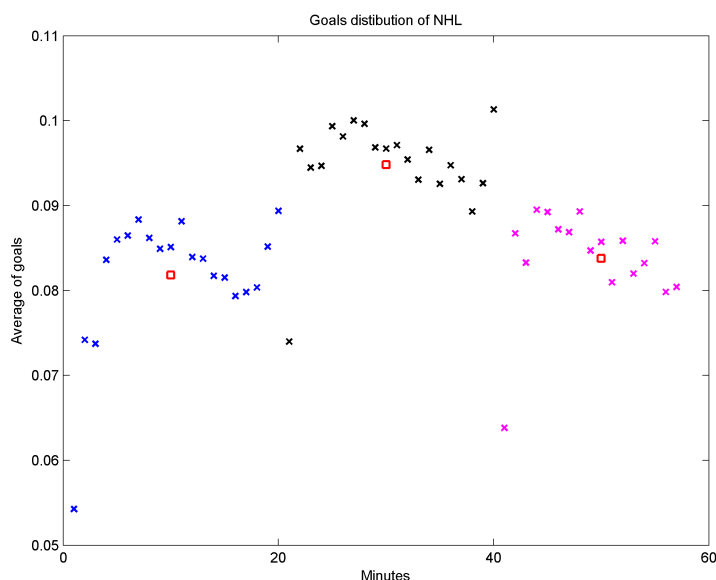
V Podkapitole 2.3.2 jsme identifikovali zajímavý jev. Jednalo se o vyšší počet vstřelených branek v průběhu druhé třetiny v porovnání s dalšími dvěma třetinami. Tento jev prozkoumáme blíže v této kapitole.

Vykreslili jsme si průměrný počet branek, které padnou v každé minutě utkání. Neuvažujeme poslední tři minuty zápasu, jelikož v závěru třetí třetiny dochází k nestandardní herní situaci tzv. hry bez brankáře. Na Obrázku 4.1 je zobrazeno rozložení průměrného počtu gólů v každé minutě a průměrný počet gólů v každé třetině přepočtený na jednu minutu. Jedná se o data všech soutěží dohromady.



Obrázek 4.1: Průměrný počet branek v každé minutě ve všech ligách dohromady

Dále jsme na ukázkou vybrali data z kanadsko-americké NHL, která jsou zobrazena na Obrázku 4.2. Grafy pro ostatní soutěže jsou přiloženy na CD u diplomové práce.



Obrázek 4.2: Průměrný počet branek v každé minutě v NHL

4.1 Shoda počtu branek v jednotlivých třetinách

Hypotézu o tom, že v každé třetině padá stejný počet branek jsme testovali pomocí chí-kvadrát testu. Popis tohoto testu lze najít v [8] v Kapitole 4.10 a 4.11. Jelikož ve třetí třetině neuvažujeme poslední tři minuty utkání, tak z vlastní podstaty hypotézy musíme zanedbat také poslední tři minuty v první a druhé třetině. Testovali jsme tři skupiny dat: českou extraligu, NHL a všechna dostupná data dohromady. V Tabulce 4.1 je uveden přehled počtů gólů v daných intervalech pro jednotlivé skupiny.

Skupina	Počet zápasů	Počet gólů od 1-17 min	Počet gólů od 21-37 min	Počet gólů od 41-57 min
CZE	5 866	8 395	9 551	8 745
NHL	14 746	20 365	23 786	21 005
Všechna data	26 313	37 075	43 029	38 551

Tabulka 4.1: Přehled počtu branek v daných intervalech

U všech námi testovaných skupin jsme zamítli nulovou hypotézu na hladině významnosti 5 % o tom, že pravděpodobnost stejného počtu vstřelených branek ve třetině nezávisí na pořadí třetiny. Výsledek Chí-kvadrát testu pro českou extraligu je zobrazen v Tabulce 4.2.

H_0	$\pi_{1.tr} = \pi_{2.tr} = \pi_{3.tr}$ ¹		
Česká extraliga	1.třetina	2.třetina	3.třetina
Počet gólů	8 395	9 551	8 745
Očekávaný počet gólů	8 897	8 897	8 897
Chí-kvadrát statistika	39,1		
P-hodnota	$3,28 \cdot 10^{-9}$		
Hladina významnosti	5 %		
Nulová hypotéza	zamítnuta		

Tabulka 4.2: Přehled výsledku Chí-kvadrát testu shodnosti počtu gólů ve třetinách pro českou extraligu

Výsledek Chí-kvadrát testu pro NHL je zobrazen v následující tabulce.

H_0	$\pi_{1.tr} = \pi_{2.tr} = \pi_{3.tr}$ ¹		
NHL	1.třetina	2.třetina	3.třetina
Počet gólů	20 365	23 786	21 005
Očekávaný počet gólů	21 718,67	21 718,67	21 718,67
Chí-kvadrát statistika	149,4		
P-hodnota	0		
Hladina významnosti	5 %		
Nulová hypotéza	zamítnuta		

Tabulka 4.3: Přehled výsledku Chí-kvadrát testu shodnosti počtu gólů ve třetinách pro NHL

¹ π označuje rozdělení náhodné veličiny

Výsledek Chí-kvadrát testu pro všechny soutěže dohromady je zobrazen v následující tabulce.

H_0	$\pi_{1.tr} = \pi_{2.tr} = \pi_{3.tr}$ ¹		
Všechna data	1.třetina	2.třetina	3.třetina
Počet gólů	37 075	43 029	38 551
Očekávaný počet gólů	39 551,67	39 551,67	39 551,67
Chí-kvadrát statistika	239,3		
P-hodnota	0		
Hladina významnosti	5 %		
Nulová hypotéza	zamítnuta		

Tabulka 4.4: Přehled výsledku Chí-kvadrát testu shodnosti počtu gólů ve třetinách pro všechna data

4.2 Regrese

V této části jsme se zaměřili na závislost počtu branek na minutě zápasu. Vytvořili jsme dvě hypotézy, které testujeme v následujících dvou podkapitolách.

4.2.1 Lineární regrese

Zápas jsme rozdělili na třetiny. Pomocí lineární regrese jsme testovali v každé třetině zvlášť závislost průměrného počtu branek na minutě třetiny. Vysvětlovaná proměnná je tedy průměrný počet branek a vysvětlující proměnná je minuta ve třetině. Regresní model je tedy $y = \beta_0 + \beta_1 \cdot x + \epsilon_i$ a testujeme hypotézu H_0 :

$$\beta_1 = 0$$

Hypotézu H_0 zamítáme na hladině významnosti $\alpha = 5 \%$ v případě třetí třetiny ve švédské lize a také v první a třetí třetině ve finské lize. V těchto případech je průměrný počet branek závislý na minutě třetiny a to tak, že s přibývajícím odehraným časem se zvyšuje průměrný počet vstřelených branek. Ve všech ostatních soutěžích v každé třetině je průměrný počet gólů

¹ π označuje rozdělení náhodné veličiny

nezávislý na minutě třetiny. Za zmínku stojí výsledek lineární regrese pro druhou třetinu švédské ligy, kde jako v jediném případě se s přibývajícím časem průměrný počet branek snižuje. Tato závislost se ale ukázala jako statisticky nevýznamná. V následujících tabulkách je zobrazen odhadů koeficientu β_1 lineární regrese pro každou třetinu ve všech testovaných soutěžích.

	CZE	CZE1	NHL	FIN
1.třetina	$5,35 \cdot 10^{-4} *$	$3,00 \cdot 10^{-4}$	$5,48 \cdot 10^{-4} *$	$1,66 \cdot 10^{-3} **$
2.třetina	$3,09 \cdot 10^{-4}$	$8,80 \cdot 10^{-5}$	$1,43 \cdot 10^{-4}$	$2,48 \cdot 10^{-4}$
3.třetina	$6,75 \cdot 10^{-4} *$	$6,43 \cdot 10^{-4} *$	$4,79 \cdot 10^{-5}$	$1,72 \cdot 10^{-3} **$

Tabulka 4.5: Přehled odhadů koef. β_1 lin. regrese v každé třetině.

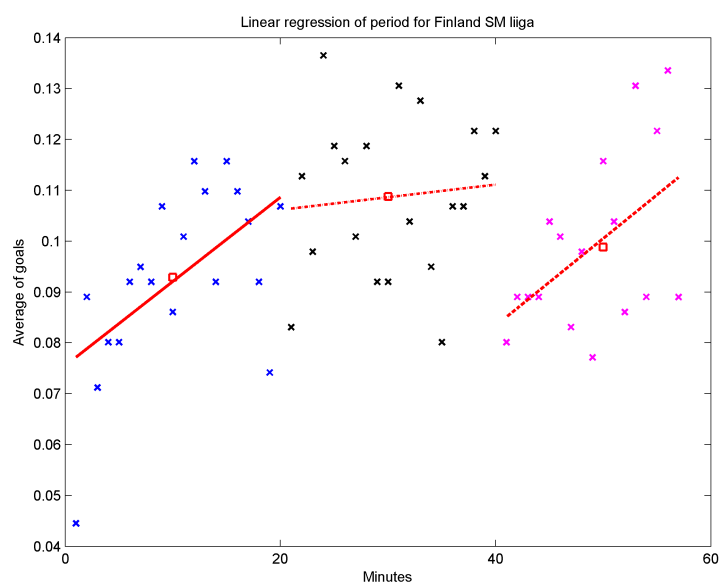
Vysvětlivky: * p = 0,1; ** p = 0,05; *** p = 0,01

	SWE	SVK	INT	Vše
1.třetina	$2,49 \cdot 10^{-4}$	$1,47 \cdot 10^{-3} *$	$1,07 \cdot 10^{-4}$	$5,20 \cdot 10^{-4} *$
2.třetina	$-1,20 \cdot 10^{-4}$	$3,29 \cdot 10^{-4}$	$2,04 \cdot 10^{-4}$	$1,74 \cdot 10^{-4}$
3.třetina	$1,16 \cdot 10^{-3} **$	$6,81 \cdot 10^{-4}$	$-5,10 \cdot 10^{-5}$	$3,20 \cdot 10^{-4}$

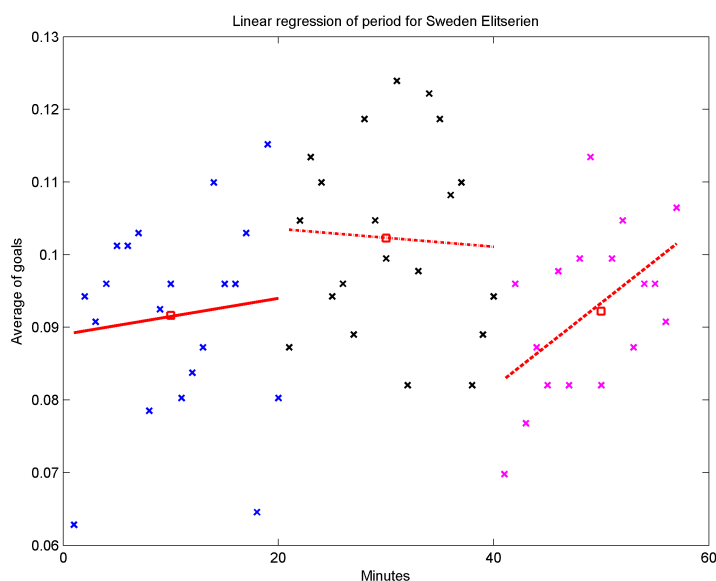
Tabulka 4.6: Přehled odhadů koef. β_1 lin. regrese v každé třetině.

Vysvětlivky: * p = 0,1; ** p = 0,05; *** p = 0,01

Zobrazení lineární regrese v jednotlivých třetinách pro finskou a švédskou ligu jsou na Obrázku 4.3 a 4.4.



Obrázek 4.3: Lineární regrese třetin pro finskou ligu



Obrázek 4.4: Lineární regrese třetin pro švédskou ligu

4.2.2 Nelineární regrese

V této části jsme již zápas na třetiny nerozdělovali, ale zkoumali jsme závislost průměrného počtu branek na minutě utkání až do 57. minuty zápasu. Navrhli jsme model nelineární regrese v tomto tvaru:

$$y = \beta_0 + \beta_1 \cdot e^{-\beta_2(x-\beta_3)^2}$$

Tento model se ukázal jako nevhodný pro data z finské a slovenské soutěže, jelikož výstupní hodnoty nenabývaly konečných čísel. Z tohoto důvodu jsme tyto soutěže zanedbali. Odhad koeficientů pro ostatní soutěže byl proveden v programu Matlab a odhady koeficientů jsou uvedeny v následující tabulce.

	CZE	CZE1	NHL	SWE	INT	Vše
β_0	0,0862	0,0858	0,0821	0,0921	0,0864	0,0845
β_1	0,0143	0,0136	0,0165	0,0157	0,0229	0,0157
β_2	0,0189	0,0165	0,0112	0,0206	0,0151	0,0138
β_3	30,4583	30,3042	30,0911	30,6561	31,1986	30,1937

Tabulka 4.7: Odhady koeficientů pro jednotlivé soutěže

Hypotézu H_0 :

$$\beta_1 = \beta_2 = \beta_3 = 0$$

jsme zamítli na základě intervalů spolehlivosti na hladině významnosti $\alpha = 5\%$ u NHL, u mezinárodních utkání a u skupiny slučující data ze všech soutěží. To znamená, že všechny parametry jsou statisticky významné. V případě české extraligy, české první ligy a švédské ligy byl statisticky nevýznamný na stejné hladině významnosti parametr β_2 . Intervaly spolehlivosti koeficientů pro českou extraligu, první ligu a švédskou ligu jsou uvedeny v Tabulce 4.8.

	CZE		CZE1		SWE	
	dolní odhad	horní odhad	dolní odhad	horní odhad	dolní odhad	horní odhad
β_0	0,0835	0,0889	0,0828	0,0887	0,0875	0,0968
β_1	0,0081	0,0206	0,0072	0,0200	0,0045	0,0269
β_2	-0,0019	0,0396	-0,0031	0,0360	-0,0161	0,0573
β_3	27,9877	32,9289	27,4882	33,1203	26,7836	34,5286

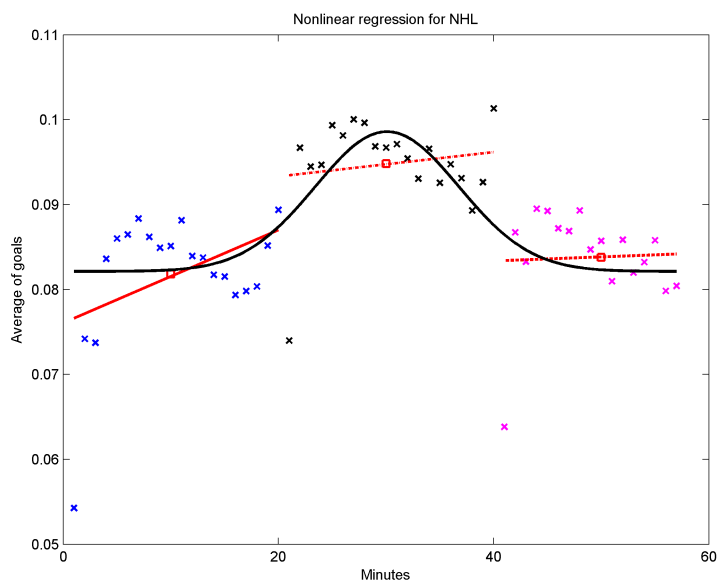
Tabulka 4.8: Intervaly spolehlivosti koeficientů pro CZE, CZE1 a SWE na hladině významnosti 5 %

Intervaly spolehlivosti koeficientů pro ostatní soutěže jsou uvedeny v Tabulce 4.9.

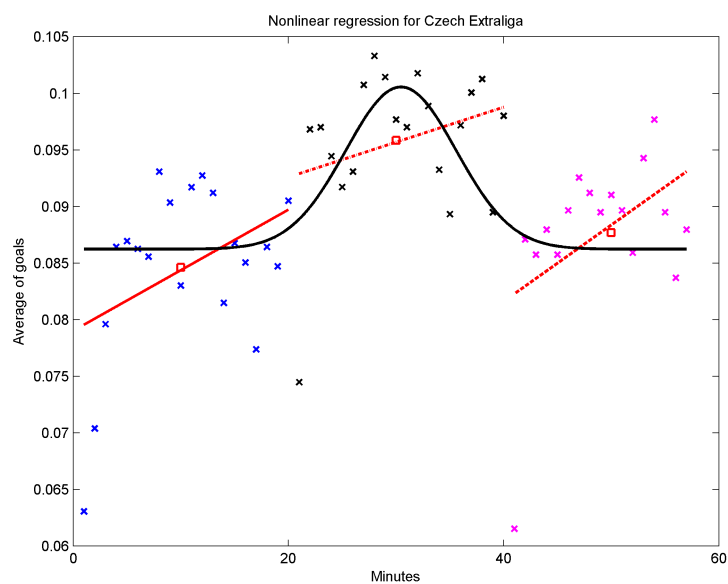
	NHL		INT		Vše	
	dolní odhad	horní odhad	dolní odhad	horní odhad	dolní odhad	horní odhad
β_0	0,0792	0,0850	0,0828	0,0900	0,0819	0,0870
β_1	0,0112	0,0217	0,0153	0,0304	0,0106	0,0208
β_2	0,0018	0,0207	0,0023	0,0279	0,0022	0,0255
β_3	27,8412	32,3411	29,1267	33,2705	28,0679	32,3194

Tabulka 4.9: Intervaly spolehlivosti koeficientů pro NHL, INT a všechna data na hladině významnosti 5 %

Pro ukázkou nelineární regresní funkce jsme zvolili data z kanadsko-americké NHL a české extraligy. Funkce jsou zobrazeny v následujících dvou grafech. Grafy pro ostatní soutěže jsou uvedeny v přílohách na přiloženém CD.



Obrázek 4.5: Nelineární regrese pro kanadsko-americkou NHL



Obrázek 4.6: Nelineární regrese pro českou extraligu

5 Vliv systému bodování

V této kapitole se zaměřujeme na to, zda bodovací systém má vliv na motivaci týmů rozhodnout utkání v základní hrací době. Použili jsme data z české extraligy. V sezoně 2000/2001 došlo v české extralize ke změně v oceňovacím systému. Změna se týkala zavedení tzv. prodloužení viz část 2.2. Změnil se ale také počet bodů, které obdržel tým za vítězství v základní hrací době. Do této změny pravidel získal tým za vítězství v základní hrací době 2 body, kdežto po této změně byl vítězný tým odměněn třemi body.

Vybrali jsme zápasy z české extraligy, které byly sehrány v rámci základní části sezony. To znamená, že jsme vyloučili utkání hraná v play-off a v baráži o extraligu. Námi vybraná data jsme rozdělili do dvou skupin. V první skupině byla utkání z let 1996 - 2000, kdy platil dvoubodový systém oceňování. Do druhé skupiny patřila utkání z let 2000 - 2011, kdy už byl zaveden třibodový systém oceňování. U obou skupin jsme rozdělili data na utkání, kde byla po dvou třetinách remíza a na ostatní utkání, kde nebyl stav utkání po 40 minutách remízový. Zkoumali jsme, zda je statisticky významný rozdíl mezi průměrným počtem vstřelených branek ve třetí třetině u dvoubodového a třibodového systému. Ve třetí třetině jsme se zaměřili na časový úsek mezi 41. minutou a 55. minutou. Opět z důvodu možnosti hry bez brankáře, která v posledních minutách utkání nastává. Zkoumali jsme tři různé hypotézy týkající se průměrného počtu vstřelených branek ve třetí třetině, které jsou popsány v následujících podkapitolách. V závěrečné podkapitole jsme testovali pomocí chí-kvadrát testu, zda pravděpodobnost, že padne ve třetí třetině gól závisí na bodovacím systému.

5.1 Remíza vs. Neremíza (vše)

V obou bodovacích systémech jsme porovnávali dvě skupiny zápasů. První skupinu tvořila utkání, kde byl remízový stav po čtyřiceti minutách a druhou utkání, kde byl stav po čtyřiceti minutách neremízový tzn. jeden z týmů je ve vedení. V případě obou skupin jsme nerozlišovali počet dosažených branek. Zápasy s výsledky 0:0, 1:1, 2:2 atd. jsme všechny zařadili do skupiny remízových utkání. Stejný přístup byl uplatněn ve druhé skupině, kam patřila utkání s jakýmkoli rozhodným výsledkem po dvou třetinách. V následujících dvou tabulkách jsou shrnuty údaje o utkáních v obou bodovacích systémech.

Dvoubodový systém	Remíza	Neremíza
Počet utkání	243	953
Průměr gólů v rozmezí 41-55 min.	1,189	1,475
Směr. odchylka gólů	1,131	1,186
Průměr domácích gólů v rozmezí 41-55 min.	0,712	0,878
Směr. odchylka domácích gólů	0,904	0,937
Průměr hostujících gólů v rozmezí 41-55 min.	0,477	0,597
Směr. odchylka hostujících gólů	0,688	0,769

Tabulka 5.1: Dvoubodový systém

Tříbodový systém	Remíza	Neremíza
Počet utkání	905	3159
Průměr gólů v rozmezí 41-55 min.	1,238	1,328
Směr. odchylka gólů	1,073	1,101
Průměr domácích gólů v rozmezí 41-55 min.	0,697	0,754
Směr. odchylka domácích gólů	0,842	0,868
Průměr hostujících gólů v rozmezí 41-55 min.	0,540	0,574
Směr. odchylka hostujících gólů	0,695	0,730

Tabulka 5.2: Tříbodový systém

5.1.1 Srovnání systémů pro remízová utkání

Stanovili jsme dvě hypotézy. V případě remízových utkání jsme testovali, zda se změnil průměrný počet vstřelených branek po zavedení tříbodového systému. Předpokládáme, že zvýšením bodového ohodnocení za vítězství v základní hrací době vzroste motivace týmů rozhodnout utkání v průběhu třetí třetiny a tím i počet vstřelených branek. Tuto hypotézu otestujeme oboustranným t-testem středních hodnot pro dva libovolné nezávislé výběry (viz [8] Kapitola 4.6.3 t-test pro dva libovolné nezávislé výběry).

$$H_0 : \mu_{dvoj} = \mu_{troj}$$

Hypotézu na hladině významnosti $\alpha = 5\%$ nezamítáme. Průměrný počet branek vstřelených v rozmezí 41. až 55. minuty v zápasech, kde je stav po dvou třetinách remízový, je u dvoubodového a tříbodového systému shodný. Výsledky hypotézy jsou shrnuty v Tabulce 5.3.

	Dvoubodový systém	Třibodový systém
Počet utkání	243	905
Průměr gólů v rozmezí 41-55 min.	1,189	1,238
Směr. odchylka gólů	1,131	1,073
p-hodnota t-testu	0,275	
Hladina významnosti	5 %	
Hypotéza H_0	nezamítnuta	

Tabulka 5.3: Výsledky t-testu středních hodnot pro remízové zápasy

5.1.2 Srovnání systémů pro neremízová utkání

Druhá hypotéza testuje utkání, ve kterých není po dvou třetinách stav remízový. Testujeme, zda je průměrný počet vstřelených branek v rozmezí 41. minuty až 55. minuty shodný v obou bodovacích systémech. Předpokládáme, že by průměrný počet branek neměl být shodný. Tuto hypotézu otestujeme pomocí dvoustranného t-testu středních hodnot pro libovolné nezávislé výběry (viz [8] Kapitola 4.6.3 t-test pro dva libovolné nezávislé výběry).

$$H_0 : \mu_{dvoj} = \mu_{troj}$$

Tuto hypotézu na hladině významnosti $\alpha = 5 \%$ zamítáme. Rozdíl průměrného počtu branek vstřelených v rozmezí 41. až 55. minuty v zápasech, kde není po dvou třetinách remíza, mezi dvoubodovým a třibodovým systémem je statisticky významný. Výsledky hypotézy jsou shrnuty v Tabulce 5.4.

	Dvoubodový systém	Třibodový systém
Počet utkání	953	3159
Průměr gólů v rozmezí 41-55 min.	1,475	1,328
Směr. odchylka gólů	1,186	1,101
p-hodnota t-testu	$6,67 \cdot 10^{-4}$	
Hladina významnosti	5 %	
Hypotéza H_0	zamítnuta	

Tabulka 5.4: Výsledky t-testu středních hodnot pro neremízové zápasy

5.2 Remíza vs. Neremíza (o jeden gól)

Stejně jako v minulé kapitole jsme porovnávali dvě skupiny zápasů v obou bodovacích systémech. Skupina tzv. remízových utkání zůstala stejná, ale narozdíl od předchozí kapitoly skupina tzv. neremízových utkání se změnila. Vybrali jsme pouze ta utkání, kde domácí nebo hostující tým vede o jednu branku. V následujících dvou tabulkách jsou shrnuty údaje o utkáních v obou bodovacích systémech.

Dvoubodový systém	Remíza	Neremíza
Počet utkání	243	402
Průměr gólů v rozmezí 41-55 min.	1,189	1,413
Směr. odchylka gólů	1,131	1,177
Průměr domácích gólů v rozmezí 41-55 min.	0,712	0,843
Směr. odchylka domácích gólů	0,904	0,911
Průměr hostujících gólů v rozmezí 41-55 min.	0,477	0,570
Směr. odchylka hostujících gólů	0,688	0,751

Tabulka 5.5: Dvoubodový systém

Tříbodový systém	Remíza	Neremíza
Počet utkání	905	1475
Průměr gólů v rozmezí 41-55 min.	1,238	1,304
Směr. odchylka gólů	1,073	1,112
Průměr domácích gólů v rozmezí 41-55 min.	0,697	0,749
Směr. odchylka domácích gólů	0,842	0,856
Průměr hostujících gólů v rozmezí 41-55 min.	0,540	0,555
Směr. odchylka hostujících gólů	0,695	0,743

Tabulka 5.6: Tříbodový systém

5.2.1 Srovnání systémů pro remízová utkání

Opět jsme testovali, zda se zvýšil průměrný počet vstřelených branek po zavedení třibodového systému. Jelikož data v případě remízových utkání jsou shodná jako v Podkapitole 5.1, použili jsme postup z této kapitoly. Hypotéza a její výsledek je také identický. Hypotézu o shodnosti průměru vstřelených branek v tzv. remízových zápasech mezi oběma bodovacími systémy přijímáme na hladině významnosti 5 %. Pro podrobné výsledky viz Podkapitola 5.1.1.

5.2.2 Srovnání systémů pro neremízová utkání

Opět testujeme utkání, ve kterých není po dvou třetinách stav remízový. Ale tentokrát je rozdíl branek mezi domácím týmem a hostujícím týmem přesně jeden gól. Testujeme, zda je průměrný počet vstřelených branek v rozmezí 41. minuty až 55. minuty shodný v obou bodovacích systémech. Předpokládáme, že by průměrný počet branek neměl být shodný. Tuto hypotézu otestujeme pomocí dvoustranného t-testu středních hodnot pro libovolné nezávislé výběry (viz [8] Kapitola 4.6.3 t-test pro dva libovolné nezávislé výběry).

$$H_0 : \mu_{dvoj} = \mu_{troj}$$

Tuto hypotézu na hladině významnosti $\alpha = 5\%$ nezamítáme. Rozdíl průměrného počtu branek vstřelených v rozmezí 41. až 55. minuty v zápasech, kde po dvou třetinách vede jedno z mužstev o gól, mezi dvoubodovým a třibodovým systémem je statisticky nevýznamný. Výsledky hypotézy jsou shrnuty v Tabulce 5.7.

	Dvoubodový systém	Třibodový systém
Počet utkání	402	1475
Průměr gólů v rozmezí 41-55 min.	1,413	1,304
Směr. odchylka gólů	1,177	1,112
p-hodnota t-testu	0,0957	
Hladina významnosti	5 %	
Hypotéza H_0	nezamítnuta	

Tabulka 5.7: Výsledky t-testu středních hodnot pro neremízové zápasy

5.3 Remíza vs. Neremíza (skupiny)

V této části jsme se podrobněji zaměřili na výsledek po 40 minutách utkání. Ve skupině remízových utkání jsme rozlišovali mezi stavem 0:0, 1:1, 2:2, 3:3, 4:4 a 5:5 a více. Stejným způsobem jsme rozlišovali i stav utkání, kde po dvou třetinách jeden z týmů vedl. Rozlišovali jsme zda padly 2, 4, 6, 8, nebo 10 a více gólů. V případě, že padly 2 branky, tak přicházel v úvahu výsledek po dvou třetinách utkání buď 2:0 nebo 0:2. Pokud by byl výsledek 1:1, tak by byl zápas zařazen do skupiny remízových utkání.

Stejně jako v předchozích dvou částech jsme zjišťovali údaje o utkáních v obou systémech. Tyto údaje jsou shrnuty v následujících dvou tabulkách.

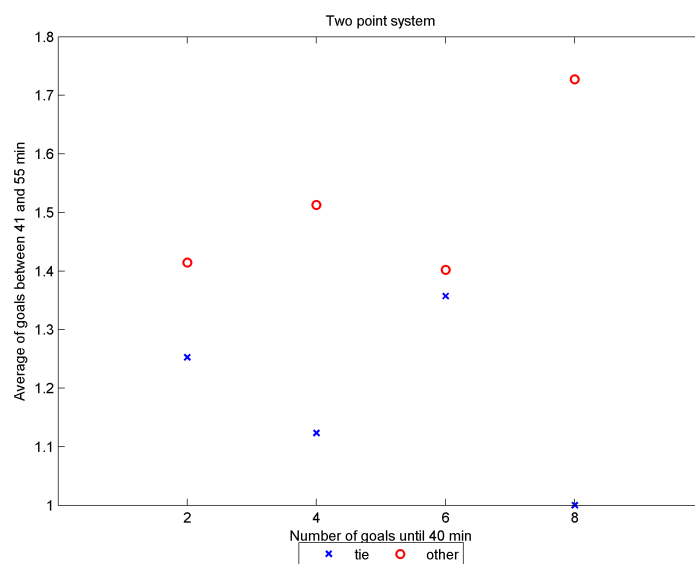
Dvoubodový systém	Počet utkání		Průměr gólů mezi 41.-55.min		Směrodat. odch. gólů	
	Počet gólů	Remíza	Jiný	Remíza	Jiný	Remíza
0	32	0	1	-	0,880	-
2	91	70	1,253	1,414	1,207	1,280
4	81	152	1,123	1,513	1,029	1,239
6	28	92	1,357	1,402	1,420	1,090
8	9	22	1	1,727	1	1,202
10+	2	8	2,5	1,25	0,707	1,281

Tabulka 5.8: Dvoubodový systém

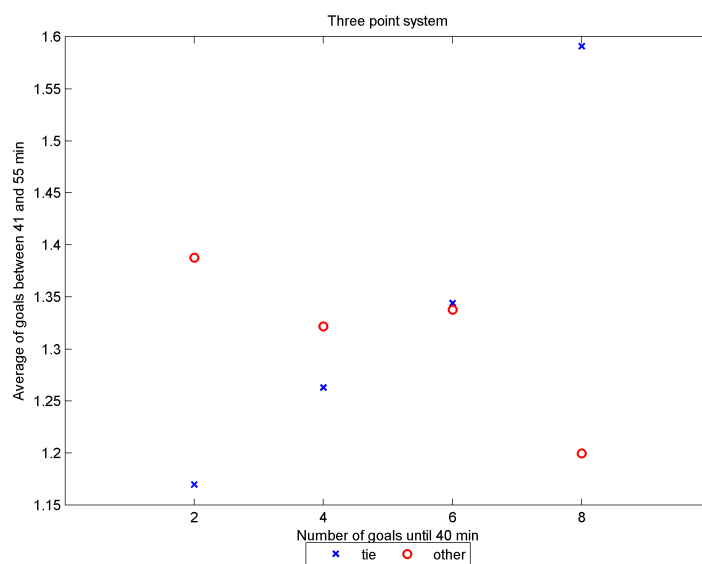
Tříbodový systém	Počet utkání		Průměr gólů mezi 41.-55.min		Směrodatná odch. gólů	
	Počet gólů	Remíza	Jiný	Remíza	Jiný	Remíza
0	122	0	1,262	-	1,141	-
2	400	343	1,170	1,388	1,007	1,053
4	266	494	1,263	1,322	1,098	1,013
6	93	225	1,334	1,338	1,147	1,154
8	22	45	1,591	1,200	1,221	0,991
10+	2	9	1	1,667	0	0,707

Tabulka 5.9: Tříbodový systém

Srovnání průměrného počtu gólů mezi 41. a 55. minutou mezi skupinou remízových a neremízových utkání pro jednotlivé bodovací systémy jsou zobrazeny v následujících dvou grafech.



Obrázek 5.1: Srovnání skupin u dvoubodového systému



Obrázek 5.2: Srovnání skupin u třibodového systému

5.3.1 Srovnání systémů pro remízová utkání

V této části jsme opět porovnávali oba systémy z hlediska průměrného počtu vstřelených branek ve třetí třetině v rozmezí mezi 41. a 55. minutou. V obou systémech jsme rozdělili remízová utkání podle výše remízy a ty poté testovali. Začínali jsme s remízami 0:0 a pokračovali až do remíz 5:5 a vyšších. Hypotéza

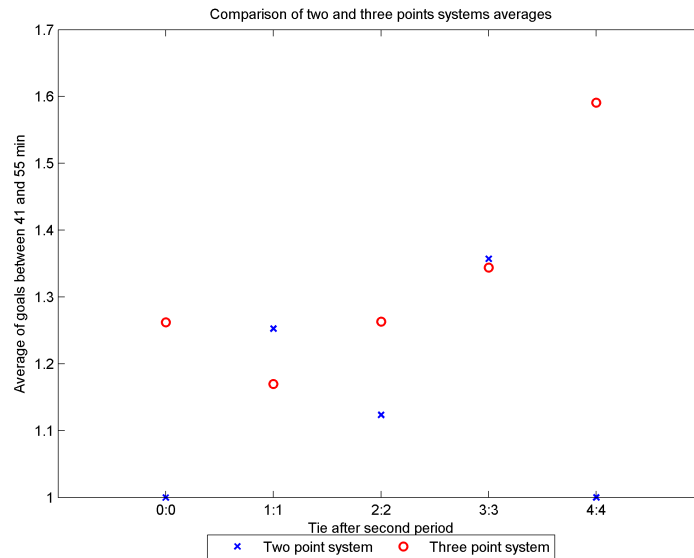
$$H_0 : \mu_{dvoj} = \mu_{troj}$$

byla opět testována pomocí dvoustranného t-testu středních hodnot pro libovolné nezávislé výběry (viz [8] Kapitola 4.6.3 t-test pro dva libovolné nezávislé výběry). Výsledky hypotéz pro jednotlivé skupiny remíz jsou shrnuty v Tabulce 5.10. Hodnoty ve sloupci hypotéza mohou nabývat jen hodnot nula a jedna. Nula znamená, že nulovou hypotézu přijímáme. V případě jednotky danou hypotézu zamítáme. Hypotézu jsme testovali na hladině významnosti $\alpha = 5 \%$.

Dvoubodový systém Stav utkání	Počet utkání		Průměr gólů mezi 41.-55.min		t-test středních hodnot	
	2-bod.	3-bod.	2-bod.	3-bod.	p-hodnota	H ₀
0:0	32	122	1	1,262	0,165	0
1:1	91	400	1,253	1,170	0,545	0
2:2	81	266	1,123	1,263	0,294	0
3:3	28	93	1,357	1,344	0,965	0
4:4	9	22	1	1,591	0,179	0
5:5+	2	2	2,5	1	0,205	0

Tabulka 5.10: Dvoubodový systém

Ve všech případech nezamítáme nulovou hypotézu. Rozdíl mezi průměrným počtem branek u dvoubodového a třibodového systému není statisticky významný. Porovnání bodovacích systémů pro jednotlivé skupiny remízových utkání je graficky zobrazeno na následujícím obrázku.



Obrázek 5.3: Srovnání bodovacích systémů pro jednotlivé skupiny remíz

5.3.2 Srovnání systémů pro neremízová utkání

V obou bodovacích systémech jsme rozdělili neremízová utkání do skupin podle počtu vstřelených branek do konce druhé třetiny. Vytvořili jsme tedy skupiny s 2, 4, 6, 8 a 10 a více brankami. Například do skupiny se čtyřmi brankami patřila utkání, kde byl stav po čtyřiceti minutách 4:0, 3:1, 1:3 nebo 0:4. Opět jsme testovali oba systémy z hlediska průměrného počtu vstřelených branek ve třetí třetině v rozmezí mezi 41. a 55. minutou. Hypotéza

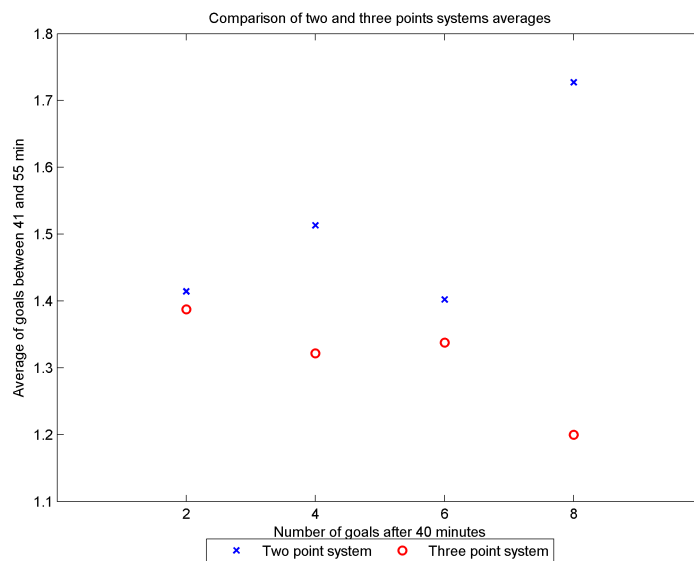
$$H_0 : \mu_{dvoj} = \mu_{troj}$$

byla opět testována pomocí dvoustranného t-testu středních hodnot pro libovolné nezávislé výběry (viz [8] Kapitola 4.6.3 t-test pro dva libovolné nezávislé výběry). V následující tabulce je přehled výsledků a p-hodnot t-testu pro jednotlivé skupiny.

Tříbodový systém	Počet utkání		Průměr gólů mezi 41.-55.min		t-test středních hodnot		
	Počet gólů	2-bod.	3-bod.	2-bod.	3-bod.	p-hodnota	H ₀
2		70	343	1,414	1,388	0,871	0
4		152	494	1,513	1,322	0,085	0
6		92	225	1,402	1,338	0,639	0
8		22	45	1,727	1,200	0,083	0
10+		8	9	1,250	1,667	0,216	0

Tabulka 5.11: Tříbodový systém

Ve všech případech nezamítáme nulovou hypotézu. Rozdíl mezi průměrným počtem branek u dvoubodového a tříbodového systému není statisticky významný. Porovnání bodovacích systémů pro jednotlivé skupiny neremízových utkání je graficky zobrazeno na následujícím obrázku.



Obrázek 5.4: Srovnání bodovacích systémů pro jednotlivé skupiny neremízových zápasů

5.4 Testování závislosti ppsti gólu na bodování

Testovali jsme hypotézu, zda existuje závislost mezi pravděpodobností padnutí branky ve třetí třetině (v rozmezí mezi 41. a 55. minutou) a bodovacím systémem. Provedli jsme dva chí-kvadrát testy. Popis chí-kvadrát testu lze najít v [8] v Kapitole 4.10 a 4.11. V prvním jsme testovali utkání, ve kterých byl stav po čtyřiceti minutách hry 1:1. Druhý test testoval zápasy, kde byl stav po dvou třetinách hry 2:2. Pro obě testované skupiny zápasů jsme v obou bodovacích systémech zjistili kolik bylo utkání, ve kterých padla branka ve třetí třetině a počet utkání, kde branka ve třetí třetině nepadla a provedli chí-kvadrát test.

Výsledek chí-kvadrát testu pro utkání se stavem 1:1 jsou uvedeny v Tabulce 5.12 a pro zápasy se stavem 2:2 v Tabulce 5.13.

H₀: padnutí branky ve 3. tř. nazávisí na bod. systému		
Remíza 1:1	Dvoubodový systém	Tříbodový systém
Počet zápasů, kde padl gól	29	117
Počet zápasů, kde nepadl gól	62	283
Chí-kvadrát statistika		0,243
P-hodnota		0,622
Hladina významnosti		5 %
Nulová hypotéza		nezamítnuta

Tabulka 5.12: Přehled výsledku chí-kvadrát testu nezávislosti ppsti padnutí gólů na bod. systému pro zápasy se stavem 1:1

H_0 : padnutí branky ve 3. tř. nazávisí na bod. systému		
Remíza 2:2	Dvoubodový systém	Tříbodový systém
Počet zápasů, kde padl gól	28	78
Počet zápasů, kde nepadl gól	53	188
Chí-kvadrát statistika		0,805
P-hodnota		0,370
Hladina významnosti		5 %
Nulová hypotéza		nezamítnuta

Tabulka 5.13: Přehled výsledku chí-kvadrát testu nezávislosti ppsti padnutí gólů na bod. systému pro zápasy se stavem 2:2

V obou případech jsme nezamítli nulovou hypotézu na hladině významnosti 5 % o tom, že pravděpodobnost padnutí gólu ve třetí třetině (v rozmezí mezi 41. a 55. minutou) nezávisí na bodovacím systému.

6 Dynamika zápasu ve třetí třetině

Tato kapitola se zabývá průběhem zápasu ve třetí třetině. Snažíme se zjistit, jestli je průběh třetí třetiny rozdílný u zápasů s remízovým a neremízovým stavem po dvou třetinách. Očekáváme, že v případě remízových utkání bude průměrný počet branek ve třetí třetině nižší než u neremízových utkání. Jelikož v utkáních, kde je po dvou třetinách remíza, očekáváme, že oba týmy budou útočit opatrněji ve snaze neobdržet branku a tím udržet alespoň remízový stav. Remíza totiž zajistí týmům jistotu zisku nejméně jednoho bodu. V případě, že stav není remízový očekáváme, že jeden z týmů musí začít útočit mnohem více, aby dosáhl alespoň na remízu a tím získá bodů. Z této situace, ale může těžit i soupeř tým, že útočí do otevřenější obrany. Za těchto předpokladů očekáváme vyšší průměrný počet branek ve třetí třetině než v remízových utkáních.

Pomocí třífaktorové ANOVA testu jsme zjišťovali, které z námi vybraných parametrů mají statisticky významný vliv na průměrný počet vstřelených branek ve třetí třetině. Zvolili jsme tři různé modely, které jsou blíže popsány v Podkapitole 6.3.

6.1 Předpoklady pro analýzu rozptylu

6.1.1 Nezávislost dat

Soubory dat splňují podmínku nezávislosti, protože každý zápas je unikátní zúčastněnými týmy a časem odehrání. Není možné, aby proti sobě hrály dva týmy ve stejném čase s rozdílným výsledkem.

6.1.2 Homoskedasticita

Druhým předpokladem pro analýzu rozptylu je shodnost rozptylu. Pro otestování tohoto předpokladu jsme použili Levene test shody rozptylu, který je součástí programu Matlab. Hypotézu o shodnosti rozptylu mezi skupinami remízových a neremízových utkáních jsme testovali na hladině $\alpha = 5\%$. Přehled p-hodnot testu pro jednotlivé skupiny ve všech soutěžích je uveden v Tabulkách 6.1 a 6.2.

Nulová hypotéza byla definována takto:

$$H_0 : \sigma_{remíza}^2 = \sigma_{jiné}^2$$

Levene test shody rozptylu mezi remíza vs. jiné				
Počet gólů	CZE	CZE1	NHL	Vše
2	0,2654	0,3207	0,9216	0,2543
4	0,6681	0,2098	0,0668	0,0086
6	0,5667	0,2804	0,1172	0,5186
8	0,7288	0,6981	0,1739	0,7940
10+	0,7098	0,3360	0,6380	0,8553

Tabulka 6.1: P-hodnoty Levene testu mezi skupinami remíza a neremíza

Ve všech případech kromě skupiny všech dat s počtem čtyř gólů po dvou třetinách přijímáme nulovou hypotézu o shodnosti rozptylu mezi remízovými a neremízovými utkáními.

Levene test shody rozptylu mezi remíza vs. jiné				
Počet gólů	FIN	SWE	SVK	INT
2	0,7593	0,2710	0,2128	0,9598
4	0,2204	0,4215	0,6432	0,0740
6	0,6234	0,2611	0,4688	0,1582
8	0,2208	-	0,6916	0,5524

Tabulka 6.2: P-hodnoty Levene testu mezi skupinami remíza a neremíza

Z Tabulky 6.2 vidíme, že ve všech případech přijímáme nulovou hypotézu o shodnosti rozptylu mezi remízovými a neremízovými utkáními. Nejsou zde uvedeny p-hodnoty testu pro skupinu s počtem deset a více branek za dvě třetiny, jelikož v soutěžích v této tabulce nebyl dostatečný počet utkání s tímto počtem branek.

6.1.3 Test normality

Posledním předpokladem je normalita dat. Pro testování normality jsme použili Lilliefors test, který je obsažen v softwaru Matlab. Data byla rozdělena na remízová a neremízová utkání a dále jsme je rozdělili podle počtu vstřelených branek do konce druhé třetiny. U všech těchto skupin byla testována normalita dat. V Tabulkách 6.3 a 6.4 je přehled výsledků normality remízových a neremízových utkání pro všechna data dohromady. Výsledky normality dat pro jednotlivé soutěže jsou uvedeny v příloze na CD. Nulová hypotéza byla definována takto:

$$H_0 : X \sim N(\mu, \sigma^2)$$

Lilliefors test normality dat				
Počet gólů	Počet zápasů	Testová statistika	Krit. hodnota	Hypotéza
2	2 383	0,226	0,019	1
4	1 870	0,225	0,021	1
6	625	0,240	0,036	1
8	125	0,214	0,080	1
10+	19	0,210	0,200	1

Tabulka 6.3: Lilliefors test normality pro remízová utkání

Lilliefors test normality dat				
Počet gólů	Počet zápasů	Testová statistika	Krit. hodnota	Hypotéza
2	2 258	0,224	0,019	1
4	3 153	0,231	0,016	1
6	1 510	0,233	0,023	1
8	393	0,229	0,045	1
10+	64	0,235	0,111	1

Tabulka 6.4: Lilliefors test normality pro neremízová utkání

Z výše uvedených tabulek vyplývá, že žádná ze skupin nesplňuje podmínku normality. Pokud jsou ale soubory dostatečně rozsáhlé ($n_i > 50$), tak se data na základě centrální limitní věty, chovají přibližně jako z normálního rozdělení (viz prezentace [5]). Při závěru o zamítnutí či nezamítnutí hypotézy H_0 musíme být opatrnější, pokud se p-hodnota pohybuje na hranici hladiny významnosti.

6.2 t-test pro jednotlivé soutěže

Použili jsme oboustranný t-test pro dva libovolné nezávislé výběry (viz [8] Kapitola 4.6.3 t-test pro dva libovolné nezávislé výběry) a testovali střední hodnoty vstřelených gólů ve třetí třetině v závislosti na dosavadním průběhu zápasu. Zkoumáme, jestli existuje rozdíl mezi zápasy, ve kterých je po dvou třetinách výsledek remízový a zápasy, ve kterých padl za dvě třetiny stejný počet gólů, ale stav zápasu není remízový. Například rozdíl mezi zápasy, kde je po dvou třetinách stav 2:2 a zápasy s výsledkem 3:1 nebo 4:0 resp. 1:3 a 0:4. Průměrný počet gólů ve třetí třetině opět zjišťujeme v rozmezí od 41. do 55. minuty z důvodu, že ke konci utkání dochází k nestandardní herní situaci hry bez brankáře. Tento test byl proveden v každé lize zvlášť. Poté jsme test provedli také na všechna námi dostupná data najednou. Nulová hypotéza byla tedy definována takto:

$$H_0 : \mu_{remiza} = \mu_{neremiza}$$

Na ukázkou jsme vybrali výsledné p-hodnoty t-testu pro data z NHL a české extraligy, které jsou uvedeny v Tabulkách 6.5 a 6.6. Tabulky s výsledky t-testu pro ostatní soutěže jsou uvedeny v přílohách na CD. Testovali jsme na hladině významnosti $\alpha = 5\%$.

NHL	Počet utkání		Průměr gólů mezi 41.-55.min		t-test středních hodnot p-hodnota Hyp.	
	Počet gólů	Remíza	Jiné	Remíza	Jiné	
2	1 370	1 243	1,237	1,238	0,983	0
4	1 123	1 750	1,192	1,315	0,003	1
6	371	790	1,267	1,251	0,815	0
8	66	197	1,045	1,406	0,017	1
10+	12	29	1,083	1,448	0,233	0

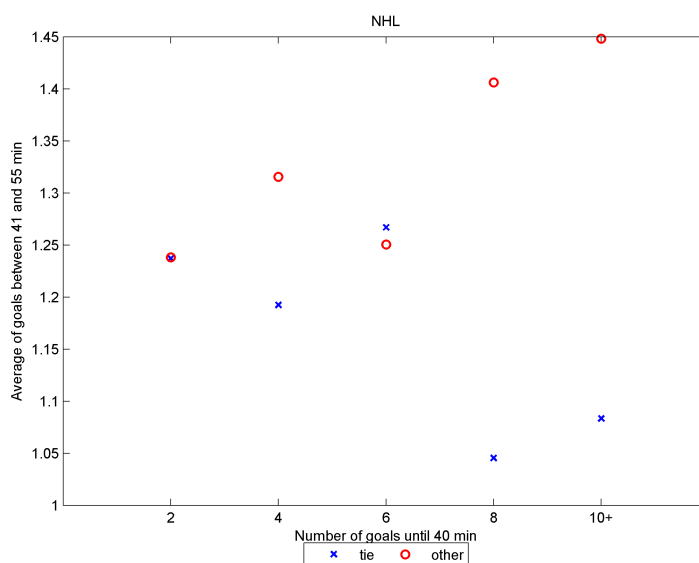
Tabulka 6.5: t-test středních hodnot gólů ve třetí třetině pro remízové a neremízové utkání v NHL

Česká extraliga Počet gólů	Počet utkání		Průměr gólů mezi 41.-55.min		t-test středních hodnot	
	Remíza	Jiné	Remíza	Jiné	p-hodnota	Hyp.
2	546	479	1,203	1,363	0,016	1
4	383	722	1,209	1,360	0,026	1
6	132	353	1,311	1,363	0,657	0
8	32	75	1,406	1,387	0,933	0
10+	4	17	1,750	1,471	0,621	0

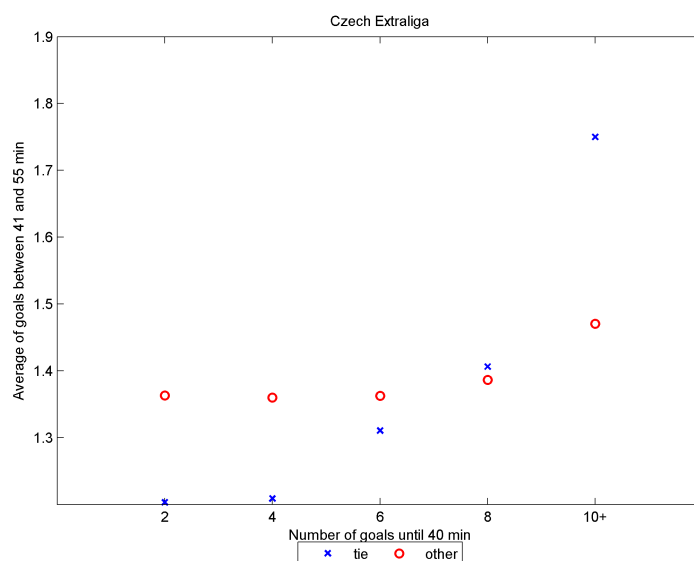
Tabulka 6.6: t-test středních hodnot gólů ve třetí třetině pro remízové a neremízové utkání v české extralize

Ve všech ostatních soutěžích jsme na hladině významnosti $\alpha = 5\%$ nezamítli nulovou hypotézu o shodě středních hodnot vstřelených branek ve třetí třetině v rozmezí 41. a 55. minuty v remízových a neremízových utkáních.

Srovnání průměrného počtu vstřelených gólů v rozmezí mezi 41. a 55. minutou mezi remízovými a neremízovými zápasy pro kanadsko-americkou NHL a českou extraligu je zobrazeno na Obrázcích 6.1 a 6.2.



Obrázek 6.1: Srovnání průměrného počtu gólů pro remízové a neremízové zápasy v NHL



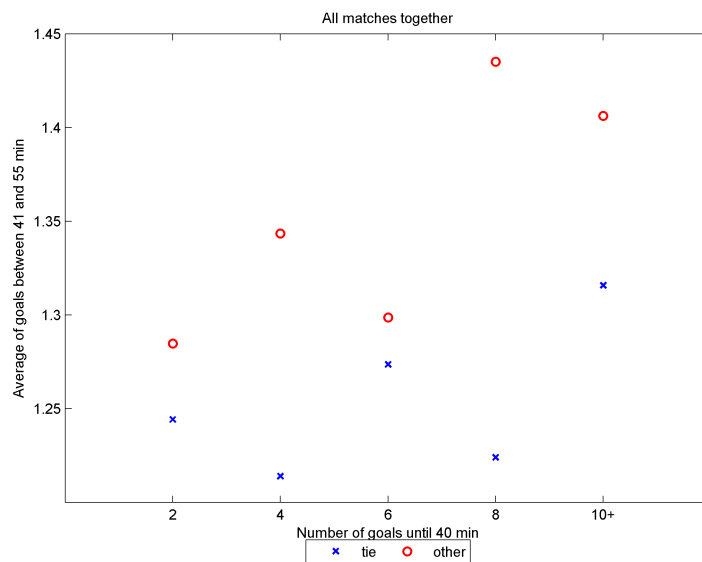
Obrázek 6.2: Srovnání průměrného počtu gólů pro remízové a neremízové zápasy v české extralize

V případě, že testujeme všechna námi získaná data najednou, tak zamítáme nulovou hypotézu o shodě středních hodnot vstřelených branek ve třetí třetině v rozmezí 41. a 55. minuty v remízových a neremízových utkáních u skupiny čtyř vstřelených branek za první dvě třetiny zápasu. Přehled výsledků t-testu pro všechna data dohromady je v zobrazen v následující tabulce.

Všechna data	Počet utkání		Průměr gólů mezi 41.-55.min		t-test středních hodnot	
	Počet gólů	Remíza	Jiné	Remíza	Jiné	p-hodnota
2	2 383	2 258	1,244	1,285	0,209	0
4	1 870	3 153	1,214	1,343	$6,26 \cdot 10^{-5}$	1
6	625	1 510	1,274	1,299	0,640	0
8	125	393	1,224	1,435	0,061	0
10+	19	64	1,316	1,406	0,728	0

Tabulka 6.7: t-test středních hodnot gólů ve třetí třetině pro remízové a neremízové utkání ve všech ligách

Srovnání průměrného počtu vstřelených gólů v rozmezí mezi 41. a 55. minutou mezi remízovými a neremízovými zápasy je zobrazeno na Obrázku 6.3.



Obrázek 6.3: Srovnání průměrného počtu gólů pro remízové a neremízové zápasy pro všechna data

6.3 Třífaktorová ANOVA

V třífaktorové analýze rozptylu zkoumáme vliv tří (námi zvolených) faktorů na měřenou veličinu. Jako měřenou veličinu jsme zvolili průměrný počet gólů ve třetí třetině zápasu. Opět jsme zvolili časový úsek ve třetí třetině od 41. do 55. minuty z důvodu možné hry bez brankáře v posledních pěti minutách. Navržený model je popsán v následující podkapitole.

6.3.1 Model 3Faktorové - ANOVA

Zkoumali jsme vliv následujících tří faktorů na průměrný počet branek vstřelených ve třetí třetině utkání.

Faktorem X_1 je typ soutěže. Rozlišujeme tedy mezi NHL, českou extraligou, českou první ligou, finskou SM-liigou, švédskou Elitserien, slovenskou extraligou a mezinárodními utkáními.

Faktorem X_2 je počet gólů po dvou třetinách utkání. Zaznamenávali jsme počet branek, který padl do konce druhé třetiny tzn. 0, 1, 2, 3, 4, 5, ... , 14.

V posledním faktoru X_3 rozlišujeme zápasy, které jsou po dvou třetinách utkání remízové a neremízové. Jsou tedy dvě úrovně, remíza a neremíza.

Výsledná tabulka třífaktorové ANOVY pro tento model je zobrazena níže.

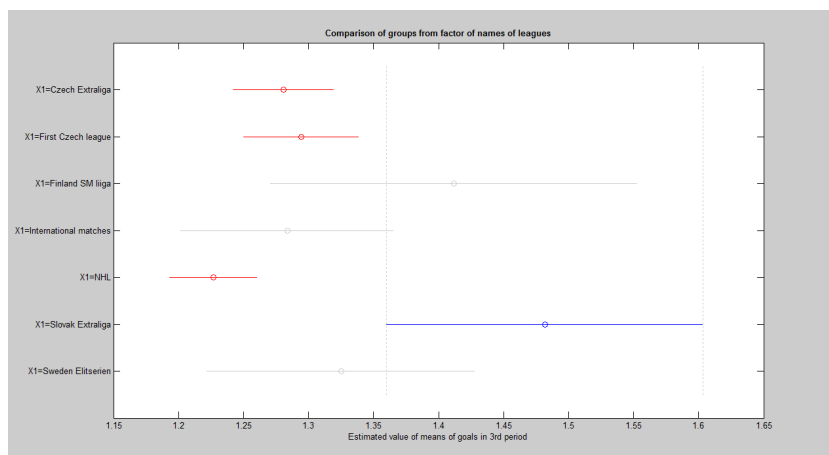
Zdroj	Součet čtverců	Stupně volnosti	Singul.	Průměr čtverců	F	Ppst > F
X1	54,58	6	0	9,097	7,273	$8,85 \cdot 10^{-8}$
X2	53,39	14	0	3,814	3,049	$9,7 \cdot 10^{-5}$
X3	1,72	1	0	1,716	1,372	0,241
Chyba	32 883,6	26 291	0	1,251		
Celkem	33 002,2	26 312	0			

Tabulka 6.8: Výsledná tabulka třífaktorové ANOVY pro navržený model

Na hladině významnosti $\alpha = 5\%$ zamítáme nulovou hypotézu, že všechny úrovně středních hodnot prvního a druhého faktoru jsou shodné. Pouze u třetího faktoru ji nezamítáme. Proto jsme provedli následnou analýzu, abychom zjistili kolik takových neshodných dvojic existuje.

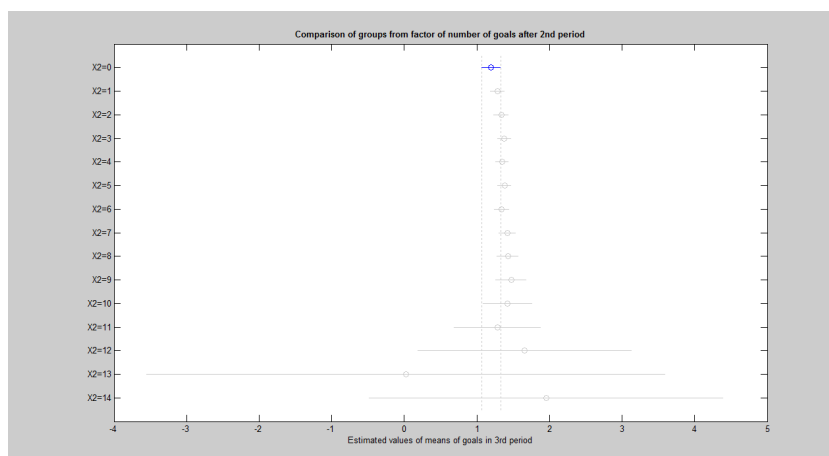
Post-hoc analýza byla provedena Bonferroniho testem na intervalu spolehlivosti $\alpha = 5\%$, který je v programu Matlab pro tyto účely zabudován. U prvního faktoru bylo odhaleno, že střední hodnoty počtu gólů ve třetí

třetině jsou statisticky různé mezi skupinami slovenská a česká extraliga, slovenská extraliga a NHL, slovenská extraliga a první česká liga a mezi finskou SM-liigou a NHL. Na Obrázku 6.4 jsou zobrazeny statisticky významné skupiny od slovenské extraligy.



Obrázek 6.4: Srovnání průměrného počtu gólů ve 3. třetině pro jednotlivé hokejové soutěže

V případě druhého faktoru jsme mezi skupinami neobjevili žádné statisticky významné rozdíly ve středních hodnotách počtu gólů ve třetí třetině. Porovnání jednotlivých skupin je zobrazeno v následujícím obrázku.



Obrázek 6.5: Srovnání průměrného počtu gólů ve 3. třetině pro skupiny podle počtu gólů po 2. třetině

Poznámka: Původně jsme navrhli tři různé modely třífaktorové ANOVY, ze kterých jsme posléze vybrali výše zmíněný model. Jelikož jsou ale skupiny faktoru X_2 seřazeny od 0 do 14 dochází ke ztrátě informace. To bychom mohli vylepšit použitím následujícího regresního modelu:

$$y = a \cdot X_2 + b \cdot X_3 + \underbrace{c \cdot SWE + d \cdot CZE + e \cdot NHL + \dots}_{X_1}$$

, kde

y ... průměrný počet branek ve třetí třetině

X_1 ... jsou logické proměnné $\{0,1\}$

X_2 ... počet gólů po dvou třetinách

X_3 ... logická proměnná $\{0,1\}$, remíza vs neremíza po dvou odehraných třetinách.

7 Závěr a shrnutí dosažených výsledků

Zpracovávali jsme data ze sportovních utkání v ledním hokeji, která nám byla poskytnuta z databáze společnosti Trefík.cz. Jedná se o data z evropských soutěží, kanadsko-americké NHL a mezistátních utkání v celkovém rozsahu 26 313 zápasů. Naším úkolem bylo identifikovat rozdíly mezi jednotlivými soutěžemi, zjistit zda změna pravidel v soutěži má vliv na průběh utkání a prozkoumat časovou dynamiku vývoje utkání ve třetí třetině.

V Podkapitole 2.3.3 jsme porovnávali vyrovnanost jednotlivých soutěží. Zjistili jsme, že v evropských soutěžích a mezistátních utkáních je výhoda domácího prostředí silnější než v případě NHL, jelikož pravděpodobnost výhry domácího týmu v základní hrací době v těchto skupinách není stejná. Větší vyrovnanost NHL je také patrná z Obrázku 2.8 na straně 17, kde jsou zobrazeny poměry mezi podmíněnou pravděpodobností výhry domácích (za podmínky jejich výhry) oproti podmíněné pravděpodobnosti výhry hostů (za podmínky výhry hostů). NHL se jeví jako nejvyrovnanější soutěž, poněvadž se tento poměr pohybuje pouze v rozmezí 0,92 až 1,2.

V Kapitole 4 jsme zkoumali zajímavý jev, a to počet branek v jednotlivých třetinách. Porovnali jsme počty branek v každé třetině a zjistili jsme, že existuje statisticky významný rozdíl v počtech vstřelených branek mezi jednotlivými třetinami. V Podkapitole 4.2.2 jsme vytvořili model nelineární regrese a na Obrázcích 4.5 a 4.6 ze strany 32 jsou zobrazeny zvýšené počty vstřelených branek ve druhé třetině pro NHL a českou extraligu. Pro jednotlivé třetiny v každé soutěži jsme provedli také lineární regresi, kde jsme zkoumali závislost průměrného počtu branek na minutě třetiny. Průměrný počet branek se ukázal nezávislý na minutě třetiny u všech testovaných třetin soutěží. Výjimkou je první a třetí třetina ve finské lize a třetí třetina ve švédské lize, kde s přibývajícím časem roste průměrný počet vstřelených branek.

Další kapitola se zabývala vlivem změny pravidel (bodového systému) na průběh poslední třetiny v české extralize. Zkoumali jsme rozdíl v průměrném počtu branek ve třetích třetinách u remízových a neremízových utkání (po dvou třetinách) v dvoubodovém a třibodovém systému. Průměrný počet branek ve třetí třetině v rozmezí 41. a 55. minuty se mezi dvoubodovým a třibodovým systémem neliší. Zápasů jsme ještě rozdělili podle výše remízy a testovali jednotlivé skupiny remíz obou bodovacích systémů. Průměrný počet branek ve třetí třetině se mezi systémy opět neliší. Můžeme tedy shrnout,

že v zápasech s remízovým stavem po dvou odehraných třetinách neměla změna bodovacího systému vliv na motivaci týmů rozhodnout utkání v poslední třetině ve svůj prospěch. V části 5.4 jsme ověřili, že pravděpodobnost padnutí gólu ve třetí třetině utkání nezávisí na bodovacím systému.

V poslední části jsme zkoumali průběh třetí třetiny mezi utkáními s remízovým a neremízovým stavem (po dvou třetinách). Testovali jsme střední hodnoty vstřelených gólů ve třetí třetině. Statisticky významné rozdíly vyšly pro skupinu čtyř vstřelených branek (tj. pro remízu 2:2 a neremízu 4:0, 3:1, 1:3 a 0:4) u NHL, české extraligy a všech dat dohromady. Tento výsledek je možné odůvodnit pomocí averze ke ztrátě, jelikož týmy raději důsledněji brání stav 2:2 a tím i získá alespoň jednoho bodu za remízu, než aby riskovaly hrát útočněji se snahou skórovat (a tím získat více bodů za vítězství), ale také možnost inkasovat branku a tím ztratit všechny body. V Podkapitole 6.3 jsme navrhli model třífaktorové analýzy rozptylu s faktory: typ soutěže, počet gólů po dvou třetinách a rozlišení stavu utkání po dvou třetinách (remíza a neremíza). Měřenou veličinou byl zvolen průměrný počet branek ve třetí třetině (rozmezí mezi 41. a 55. minutou). Statisticky významné rozdíly středních hodnot byly zjištěny u prvního a druhého faktoru. Následnou analýzou jsme u prvního faktoru identifikovali skupiny se statisticky významnými rozdíly v průměrném počtu gólů ve třetí třetině, a to mezi slovenskou extraligou a NHL, českou extraligou a českou první ligou a mezi finskou ligou a NHL. Stejnou analýzu jsme provedli i pro druhý faktor, ale neobjevili jsme žádné statisticky významné rozdíly středních hodnot počtu gólů mezi skupinami. Grafické znázornění těchto dvou analýz je na Obrázcích 6.4 a 6.5 na straně 53.

Shrnutí výsledků této práce. Pokud pomíneme poslední minuty zápasu, kdy je obvyklá hra bez brankáře zjistíme, že nejvíce branek padá během druhé třetiny. Tento jev si můžeme vysvětlit pomocí tzv. averze k riziku, kdy předpokládáme, že po první třetině týmy rozpoznají taktiku soupeře a snaží se tuto informaci ve druhé třetině zužitkovat ke vstřelení branek. V poslední třetině se již začne projevovat zmíněná averze ke ztrátě, jelikož tým, který vyhrává, se bude snažit více bránit, aby neztratil vedení. Tudíž v této části utkání padne méně gólů. Dále jsme zjistili, že změna pravidel bodového systému neovlivnila počet vstřelených branek ve třetí třetině v utkáních, v nichž byla po dvou třetinách remíza. Tím se nezvýšila motivace týmu hrát útočnějším stylem k rozhodnutí zápasu v základní hrací době. V určité skupině utkání jsme odhalili, že průměrný počet vstřelených branek ve třetí třetině je vyšší u zápasů s neremízovým stavem než u zápasů s remízovým stavem (po dvou třetinách). Tento jev jsme opět odůvodnili averzí ke ztrátě, kdy předpokládáme, že týmy se snaží vyhnout ztrátě jednoho bodu za remízu raději než šanci získat další bod v případě vstřelení branky a vítězství v zápase.

Literatura

- [1] *A. N. Banerjee, J. F. Swinnen: Skating on thin ice: rule changes and team strategies in the NHL*
Canadian Journal of Economics, 2007, Vol. 40, No. 2, str. 493-514
- [2] *D. Coates, B. R. Humphreys: Game Attendance and Competitive Balance in the NHL*
Working paper No.2011-08, University of Alberta, 2011
- [3] *D. Kahneman, J.L. Knetsch, R.H. Thaler: Anomalies: The Endowment Effect, Loss Aversion and Status Quo Bias*
Journal of Economic Perspectives, 1991, Vol. 5, No. 1, str. 193-206
- [4] *D. Kahneman, A. Tversky: Loss Aversion in Riskless Choice: A Reference Dependent Model*
Quarterly Journal of Economics 106, 1991, str. 1039-1061
- [5] *G. Klein, S. Lee, D. Monday: The Assumptions of Anova*
Prezentace je uvedena na příloženém CD, 2005
- [6] *L. J. A. Lenten, J. Libich, P. Stehlík: Risk and Loss Aversion in Sports*
V přípravě
- [7] *D. G. Pope, M. E. Schweitzer: Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes*
The American Economic Review 101, 2011, str. 129-157
- [8] *J. Reif: Metody matematické statistiky*
Západočeská univerzita, Plzeň 2004