

## **VLIV COVIDU NA VÝSLEDKY SPORTOVNÍCH UTKÁNÍ: ZHORŠILY PRÁZDNÉ OCHOZY BILANCI DOMÁCÍCH TÝMŮ?**

### **IMPACT OF COVID ON RESULTS OF SPORTS MATCHES: DID SPECTATORS ABSENCE DECREASE THE ADVANTAGE OF HOME TEAMS?**

**Petr Volf**

*Adresa:* Institute of Information Theory and Automation AS CR, Pod vodárenskou věží 4, 18208 Praha 8

*E-mail:* [volf@utia.cas.cz](mailto:volf@utia.cas.cz)

**Abstrakt:** Součástí modelů, kterými se snažíme popsat a predikovat vývoj a výsledky sportovních utkání, bývá i komponenta či parametr zohledňující výhodu domácího prostředí. Budu se zabývat fotbalem či hokejem, ale tento jev se týká i jiných sportů, a nejen míčových. Tato výhoda může souviset se znalostí prostředí, s psychickým nastavením („doma by se mělo vyhrát“), ale také ji vytvářejí fandící diváci. A právě „doba covidová“ umožňuje porovnání, jak silný tento vliv je. Bez diváků či se silným omezením jejich počtu se hrála celá sezóna 2020/2021. Porovnávám proto úspěšnost domácích týmů v této sezóně s dvěma sezónami před i po. A to dost jednoduchým způsobem. Nemodeluji síly obran či útoků jednotlivých týmů, jak se to běžně dělá, srovnávám prostě jen počty domácích výher (proher, remíz), nebo počty vstřelených a obdržených branek, po jednotlivých kolech soutěže, bez rozlišování týmů. Ale přesto nejprve připomenu základní modely pro statistický popis sportovních utkání.

**Abstract:** The home advantage is a phenomenon having evidently a strong impact on results of sports matches. Therefore the “home advantage” parameter or component is standardly a part of statistical models describing (and forecasting) sports results. I will concentrate here to the soccer and ice-hockey, however this phenomenon concerns other sports as well. One initiator of this advantage is the presence of home spectators. That is why the time of Covid pandemic, when at least one season was played without (or with strong limitation of) spectators, gives a chance to compare the „Covid” season with the others and, possibly, to discover a significant decrease of home advantage impact. This paper uses just rather simple statistical tools, nevertheless, with the aid of basic statistical tests the changes of home teams results (regarding both the gained points and scored goals) have been detected.

## 1. Úvod, modely zohledňující domácí výhodu

Mám teď na mysli sporty jako fotbal, hokej, kde výsledek je určen na základě počtu branek. Cílem je zpravidla modelovat výsledky jednotlivých utkání, za účelem jejich predikce, tj. i s ohledem na spolehlivé sázení na ně. Takže je třeba zahrnout parametry jednotlivých týmů zohledňující jejich sílu, i další faktory mající vliv na zápas. Nejjednodušší by bylo vytvořit např. trinomičtý model pro výsledky výhra, remíza, prohra, třeba logistického typu, který umožňuje popsat i onen vliv různých okolností. Ale od začátku byly také vyvíjeny modely pro skóre zápasů. Samozřejmě, předpověď skóre moc spolehlivá není, hlavním cílem je ale stále předpovídat vítěze.

Nejjednodušší model je třeba následující založený na dvou nezávislých Poissonových náhodných veličinách (např. Maher, 1986, model 2): Označme  $X_{ij}$  a  $Y_{ij}$  počet gólů domácích a hostů v zápase mezi týmem  $i$  (domáci) a  $j$  (hosté). Pak

$$\begin{aligned} X_{ij} &\sim \text{Poisson}(\lambda_1 = \alpha_i \beta_j \gamma), \\ Y_{ij} &\sim \text{Poisson}(\lambda_2 = \alpha_j \beta_i), \end{aligned} \tag{1}$$

kde  $\alpha_i, \beta_i > 0$ ,  $\alpha_i$  charakterizuje útočnou sílu týmu  $i$ ,  $\beta_i$  pak sílu obrany. No a  $\gamma$  je parametr domácí výhody. Parametry  $\alpha_i, \beta_i$  musí být vhodně normovány (kvůli jednoznačnosti). Takže i zde se pochopitelně počítá s výhodou domácího prostředí, očekává se  $\gamma > 1$ . V článku Maher (1986), kde zpracovává data z ročníku 1971/1972 anglické Premier Division, přímo odhad  $\gamma$  není uveden, dá se spočítat z jeho tabulek, vychází kolem 1,38.

Maher si také všímá nesouladu četností výsledků s malým počtem gólů s nezávislým modelem, a navrhuje použít 2-rozměrný Poissonův model. Ten je vybudován takto: Nechť  $X_1, X_2, X_3$  jsou 3 nezávislé veličiny s Poissonovým rozdělením s parametry  $\lambda_1, \lambda_2, \lambda_3$ . Pak 2-rozměrné Poissonovo rozdělení je tvořeno veličinami  $Y = X_1 + X_3, Z = X_2 + X_3$ . Jejich kovariance je  $\lambda_3$ , a mají marginální Poissonova rozdělení s parametry  $\lambda_1 + \lambda_3, \lambda_2 + \lambda_3$ . Z konstrukce plyne, že korelace je nezáporná. Maherův odhad korelace je 0,2, a model prokazuje skutečně lepší fit.

V dalším článku Dixon, Coles (1997) analyzují sezónu 1995/1996 a 4 nejvyšší anglické fotbalové soutěže (92 týmů). Parametr domácí výhody  $\gamma$  je opět přibližně 1,38, a zároveň ukazují, že je prakticky konstantní během celé sezóny. To znamená, že také sledovali, jak se parametry mění během roku. Nesoulad dat s nezávislým Poissonovým modelem řeší pak použitím „inflated” modelu, čili směsí základního modelu s distribucí dávající jinou váhu

několika výsledkům, (konkrétně, pravděpodobnosti 0:0, 1:1 jsou v „nezávislém“ modelu podceněny, naopak 0:1, 1:0 se vyskytují méně než z něj vyplývá).

Rodina modelů je opravdu bohatá. Třeba v Volf (2009) jsou modelem pro vývoj skóre 2 zároveň běžící čítací procesy, které se vzájemně ovlivňují prostřednictvím kovariát (v rámci Coxova regresního modelu, jehož parametry také mohou být proměnné v čase). Kovariáty odrážejí stavy týmů i průběžný stav zápasu.

Pokud jde třeba o lední hokej, tam vychází parametr domácí výhody nižší. Např. při analýze dat z české hokejové Extraligy v ročníku 2007/2008 jsem kdysi (opět v rámci modelu 2 nezávislých Poissonových rozdělení pro počet gólů) odhadl  $\gamma = 1,23$  (1,17, 1,29) (Volf, ICPM 2008). Porovnání domácí výhody u více sportů uvidíme hned v další části, v tabulce na Obr. 1.

## 2. Jak jednoduše měřit domácí výhodu?

Inspirací může být článek Pollard (1986). Ten se soustřeďuje právě na kvantifikaci výhody domácího prostředí, v rámci určité soutěže či ročníku, bez toho, aby modeloval vývoj utkání, sílu týmů a pod. Na začátku formuluje několik možných příčin takové výhody, některé se dají měřit, některé hůř:

1. Divácká kulisa, povzbuzování.
2. Znalost domácího prostředí (domácí hráči mají i zázemí, na které jsou zvyklí).
3. Taktika, domácí chtějí vyhrát.
4. Strannost rozhodčích (i nechtěná, způsobená atmosférou, reakcí diváků).
5. Další psychické faktory (+ i –, třeba i zvýšená nervozita).
6. Únava hostů z cesty, atd.

Dále se Pollard zabývá poměrně jednoduchou metodou: prostě spočte relativní četnost výher domácích týmů, nebo v případě, kdy je víc možností než prostá výhra/prohra, spočte poměrnou část bodů získaných domácími týmy. A pokud je statisticky významně nad 50 %, je to důkaz o existenci domácí výhody a zároveň i její číselné ohodnocení. Pro některé soutěže jsou jeho výsledky v následující tabulce.

V podstatě použijí také takové jednoduché porovnávání, nyní s cílem porovnat bilanci domácích za Covidu (tj. v podstatě bez diváků) s normálními sezónami. Ještě se zmíním o jedné studii vyhodnocující domácí výhodu, od autorů z KMA ZČU v Plzni (Marek, Vávra, 2020). Ti navrhli použít rozdíly

Table 1. Home advantage in different professional team sports.

Sport	League	Seasons	Record of home teams				Home advantage		
			Played	Won	Drawn	Lost	Per cent	s.e.(%)	P*
North America									
Baseball	AL and NL	1982-84	6316	3388	—	2928	53.6	0.6	<0.001
US football	NFL	1982-84	574	314	3	257	55.0 <sup>b</sup>	2.1	<0.01
Ice hockey	NHL	1981/2-83/4	2520	1328	361	831	59.9 <sup>c</sup>	0.9	<0.001
Basketball	NBA	1981/2-83/4	2829	1790	—	1039	63.3	0.9	<0.001
Soccer	NASL	1982-84	512	334	—	178	65.2	2.1	<0.001
England									
Cricket	County Championship	1981-83	478	162	189	127	56.1 <sup>b</sup>	2.9	<0.05
Soccer	FL Division 1	1981/2-83/4	1386	695	350	341	63.9 <sup>d</sup>	1.5	<0.001

\*One-sided test of proportion; null hypothesis of no home advantage.

<sup>b</sup>Drawn games excluded.

<sup>c</sup>Two points for a win, one for a draw.

<sup>d</sup>Three points for a win, one for a draw.

Obrázek 1: Tabulka 1 z Pollard (1986).

skóre ze zápasů každé dvojice týmů,  $i, j$ , doma i venku, a vytvořit trinomicovou veličinu  $C$  s hodnotami  $-1, 0, 1$  takto:

$$C_{ij} = \text{sign}[(h_i - a_j) - (a_i - h_j)],$$

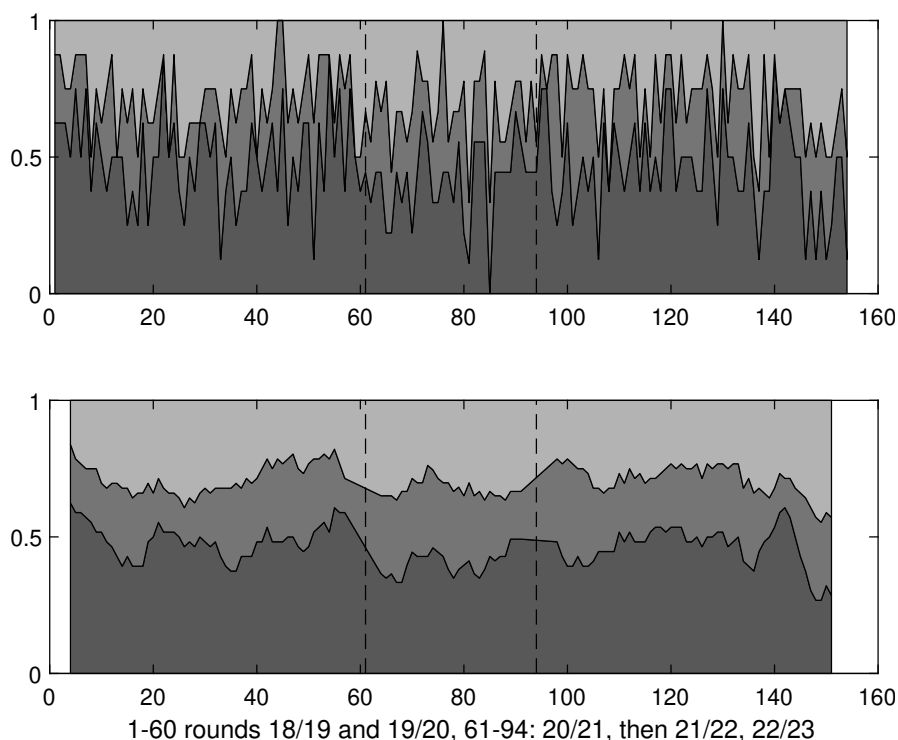
kde  $h_i, a_i$  jsou počty gólů vstřelených týmem  $i$  doma (home) a venku (away) v tomto „dvojzápase“ týmů  $i, j$ . Tyto hodnoty pak sečetli a zprůměrovali za jednotlivý ročník soutěže, pak případně za více ročníků téže soutěže. Zpracovali takto dost velké množství dat, z 19 evropských fotbalových soutěží (i nižších), za ročníky 2007/2008 až 2016/17, počet dvojic zápasů byl přes 36 000. Je otázka, zda a proč je tento způsob bilance výhodný. Metoda používá poměrně detailní informaci (skóre zápasů), kterou ale nevyužije naplno. Pro domácí výhodu zde svědčí, když je střední hodnota veličiny  $C$  významně větší než nula.

### 3. Analýza vlivu Covidu na fotbalovou Fortuna ligu

Nejdřív se podívám na českou 1. fotbalovou ligu. Budu porovnávat ročníky 2018/2019 až 2022/2023 bez 2020/2021 právě s tímto ročníkem. A to sice základní části těchto ročníků, bez nadstaveb. Výsledky zápasů z oněch ročníků lze nalézt na [www.sport.cz/sekce/fotbal-ceska-1-liga-vysledky-4](http://www.sport.cz/sekce/fotbal-ceska-1-liga-vysledky-4).

Jaro 2020 se vlastně dohrálo s diváky, neboť liga byla od poloviny března do konce května přerušena, zatímco ročník 2020/2021 začal s omezením na 2000 diváků, v září se rychle přešlo na maximum 500 a v říjnu byly všechny soutěže přerušeny. Od listopadu se pokračovalo, ale kompletně bez diváků, k pozvolnému uvolnění došlo až v květnu 2021. Nadstavba se už nehrála.

Každá základní část měla 16 týmů, hrál každý s každým doma i venku, takže bylo 30 kol po osmi zápasech. Až právě na sezónu 2020/2021, ta měla 18



Obrázek 2: Proporce domácích výher (spodní, nejtmaší pole), remíz, proher (horní, nejsvětější pole) po jednotlivých kolech v sezónách 2018/2019 až 2022/2023. Dole je totéž vyhlazeno klouzavým oknem šířky 7.

týmů, takže se hrálo 34 kol po 9 zápasech. Nezapomínám se jednotlivými týmy, nejmenší jednotkou je pro mne 1 odehrané kolo, takže bilanci kol 2020/2021 jsem přepočítal jakoby na 8 zápasů. Ještě zde podotknu, že grafy mají formu časových řad (časem je index kola, což většinou odpovídá periodě 1 týden), ale jde v podstatě o realizaci nezávislých náhodných veličin (stejně rozdělených v každé z obou skupin). To bylo v obou skupinách otestováno testy nezávislosti (série nad a pod mediánem, série nahoru a dolů).

Nejdřív jsem se zabýval počty výher domácích, remíz a výher hostů v jednotlivých kolech. Graf těchto počtů je na Obrázku 2. V dolní části je vždy zprůměrováno nejbližších 7 kol. Z obrázku se nedá usoudit na nějaký výrazný

pokles bilance domácích, to se dá udělat až pomocí porovnání proporcí, nejlépe analýzou kontingenční Tabulky ??.

	Výhry	Remízy	Prohry	Součty	Proporce		
Ostatní	458	227	275	960	0,4771	0,2364	0,2865
2020/2021	126	80	100	306	0,4118	0,2614	0,3268
Součty	584	307	375	1266			

Tabulka 1: Četnosti a relativní četnosti domácích výher, remíz, proher.

Z této tabulky nedostaneme statisticky významný výsledek,  $p$ -hodnota příslušného chí-kvadrát testu nezávislosti v kontingenční tabulce je 0,1341. Pokud však vezmeme domácí výhry proti ostatním výsledkům, rozdíl proporcí 0,4771 proti 0,4118 už je statisticky významný, test nezávislosti v  $2 \times 2$  kontingenční tabulce má  $p$ -hodnotu 0,0459. Takže lze aspoň takto považovat tezi o zhoršené bilanci v domácích zápasech za prokázanou.

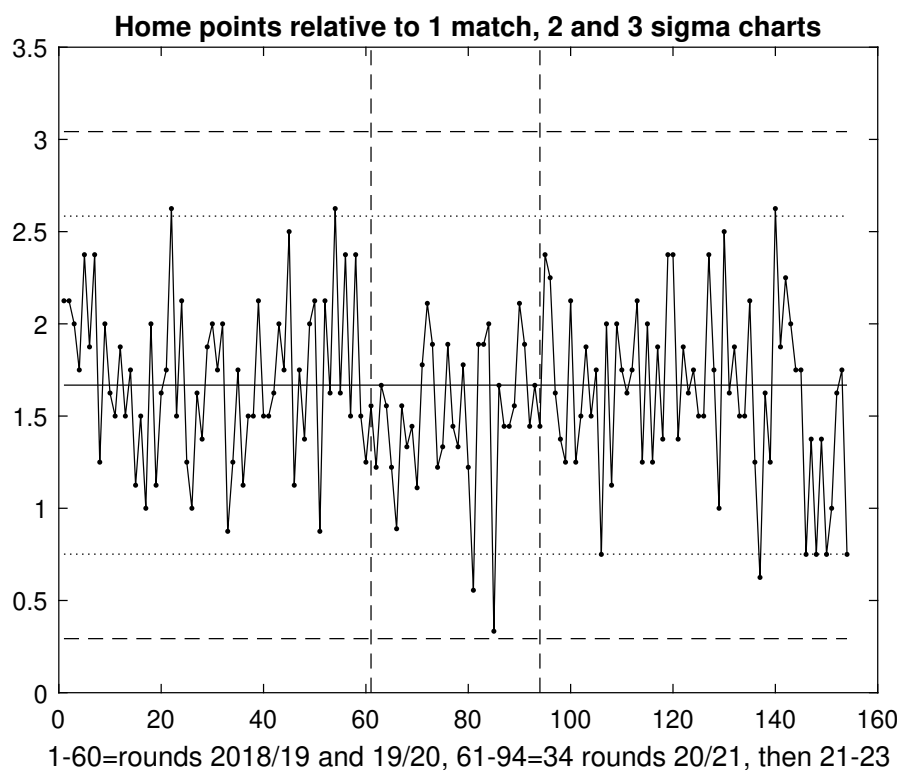
Pokud vycházím jen z bilance výher, remíz a proher, mohu je ohodnotit body (3 body za výhru, 1 bod za remízu, 0 za prohru) a tedy porovnávat body získané domácími týmy v obou obdobích. Obrázek 3 ukazuje, jak to vypadalo: V každém kole jsem spočítal počet bodů získaných domácími, relativně na 1 zápas. Takto získaná veličina má základní charakteristiky v Tabulce 2, a test normalitu jejího rozdělení nezamítá (použil jsem Lillieforsův test, tj. „zprísněný“ Kolmogorovův Smirnovův test, zvláště pro sezónu 2020/2021 a pro ostatní ročníky).

Sezóna	Průměr	Směr. odchylka	Medián	$n$ je počet kol
Ostatní sezóny	1,6677	0,4582	1,6250	120
2020/2021	1,4967	0,3966	1,5000	34

Tabulka 2: Charakteristiky počtu bodů získaných domácími za 1 zápas.

Na Obrázku 3 vidíme průměr i kontrolní linie (ve vzdálenosti 2 a 3 směrodatné odchylky od průměru) odpovídající této veličině v sezónách mimo 2020/2021. Pokud udělám  $t$ -test hypotézy rovnosti středních hodnot, příslušná  $p$ -hodnota je 0,0501. „Hrubší“ Wilcoxonův test na rovnost mediánů pak dá  $p$ -hodnotu 0,0674, čili v obou případech je výsledek na hranici zamítnutí hypotézy rovnosti.

Kromě počtu výher či bodů jsem se také podíval na skóre zápasů, zase vždy dohromady přes celé kolo. Počet gólů vstřelených domácími a hosty

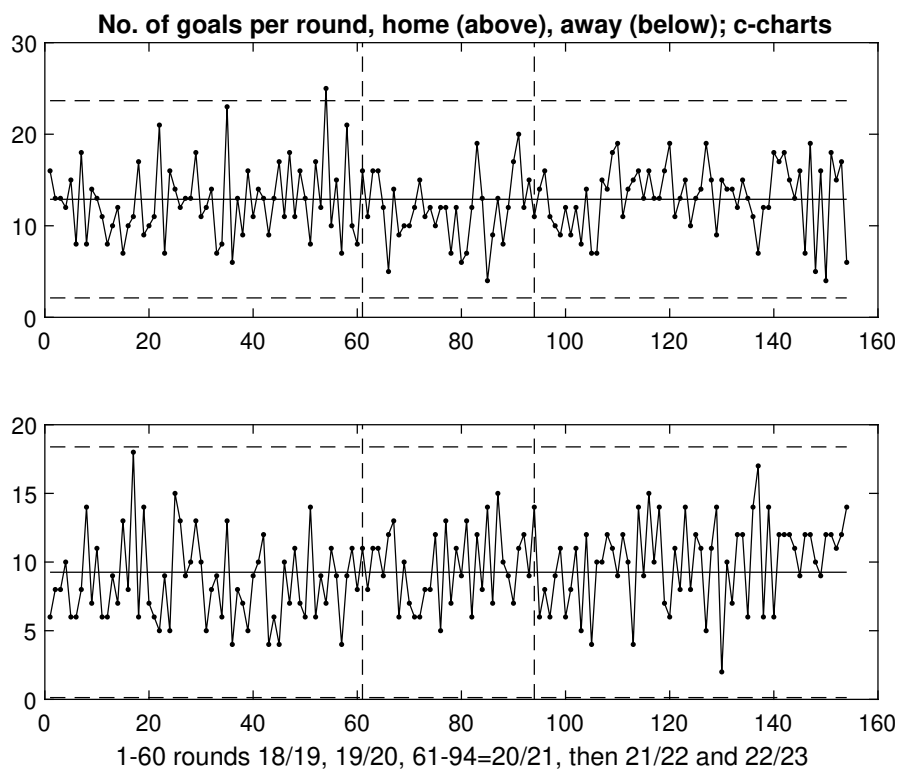


Obrázek 3: Průměrný počet bodů získaný domácím týmem za 1 zápas, po jednotlivých kolech.

během jednotlivých kol ukazuje Obrázek 4 (opět, počty gólů v každém kole v sezóně 2020/2021 byly brány jako zaokrouhlení 8/9 skutečného počtu). Obě skupiny dat (2020/2021 a ostatní) v obou grafech lze modelovat jako nezávislé Poissonovy náhodné veličiny. Tabulka 3 ukazuje jejich charakteristiky.

Sezóna	Průměr doma	Průměr venku	Std doma	Std venku
Ostatní	12,8917	9,2583	3,9211	3,2032
2020/2021	11,7647	9,7059	3,7178	2,7362

Tabulka 3: Charakteristiky počtu vstřelených gólů za jedno kolo.



Obrázek 4: Počty gólů vstřelených domácími (nahore) a hosty (dole) v jednotlivých kolech.

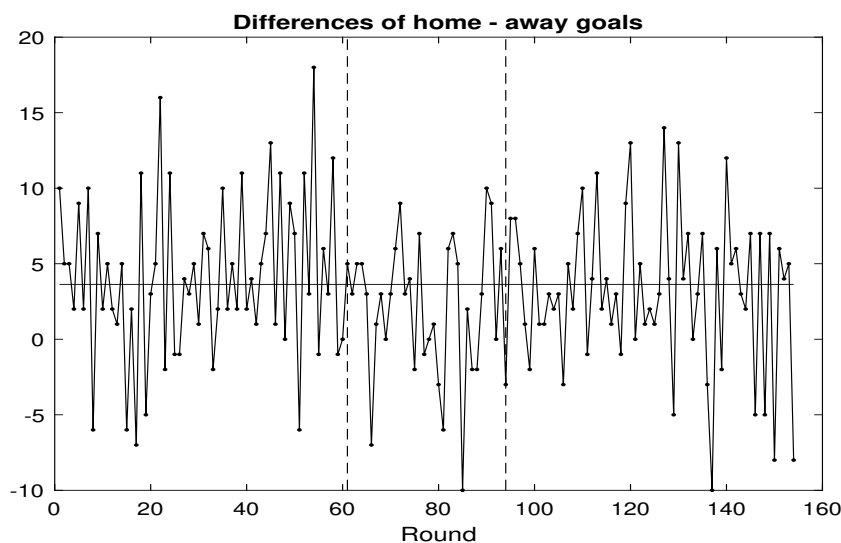
Obrázek 4 sice naznačuje jakési rozdíly, snížené počty gólů doma a zvýšené venku, výraznější je to pro domácí góly, zatímco pro góly venku se nepodařilo prokázat statisticky významnou odlišnost. Jednostranné testy rovnosti středních hodnot na základě CLV použité pro Poissonovo rozdělení vedly k následujícím  $p$ -hodnotám:

$$p = 0,0471 \text{ pro góly domácích, } p = 0,2287 \text{ pro góly hostů.}$$

Obrázek 4 jsem doplnil kontrolními liniemi odpovídajícími řadě Poissonových veličin, z dat bez ročníku 2020/2021.

Jak jsem zmínil, někteří autoři používají 2-rozměrné Poissonovo rozdělení, pro góly vstřelené a obdržené během jednoho zápasu. Ta používaná verze má jedno omezení, a to že korelace je kladná. Zde šlo o počty gólů (domácích





Obrázek 5: Rozdíly počtu gólů vstřelených domácími a hosty v jednotlivých kolech.

a hostů) vždy za celé kolo, korelace byla jednak hodně malá (statisticky nevýznamná), a navíc záporná.

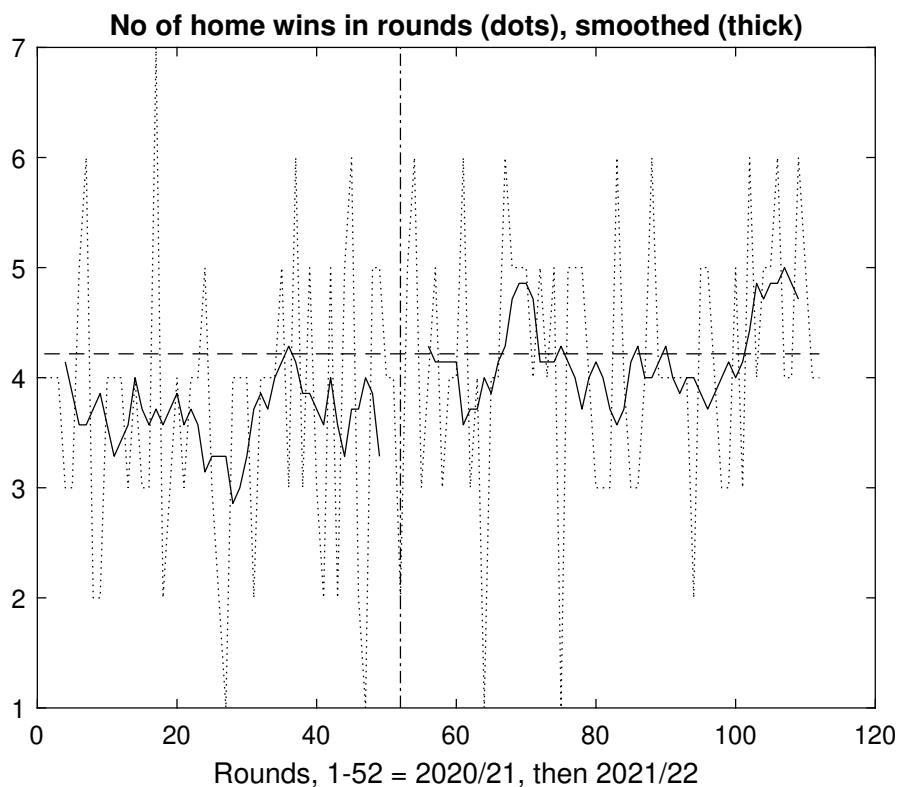
Ještě jsem zkusil i další model, a to rozdíl dvou nezávislých Poissonových rozdělení, výsledkem je tak zvané **Skellamovo rozdělení** (viz třeba opět Volf, ICPM 2008). Na Obrázku 5 jsou vidět rozdíly počtu gólů vstřelených domácími a hosty po jednotlivých kolech, průměry těchto rozdílů jsou 2,0588 pro 2020/2021 a 3,6434 pro ostatní (této hodnotě odpovídá i centrální linie v grafu).

Opět je možné použít alespoň přibližný test na rovnost příslušných středních hodnot, založený na CLV: Jsou-li  $X, Y$  nezávislé Poissonovy n.v. s parametry  $\lambda_1, \lambda_2$ , tak  $S = X - Y$  má  $E(S) = \lambda_1 - \lambda_2$  a rozptyl  $\text{var}(S) = \lambda_1 + \lambda_2$ , čili vhodná testovací veličina je

$$Z = (\bar{S}_1 - \bar{S}_2) / \sqrt{\text{var}(S_1)/n_1 + \text{var}(S_2)/n_2},$$

kde  $\bar{S}_1, \bar{S}_2$  jsou ony průměry rozdílů uvedené výše, odhady rozptylů jsou (jak naznačeno výše) součty oněch průměrů z Tabulky 3, zvlášť v řádku pro 2020/2021 a v řádku pro ostatní,  $n_1, n_2$  odpovídající počty hodnot, zde tedy počty kol v 2020/2021 a v ostatních sezónách.

Nyní takto provedený jednostranný test dal  $p = 0,0164$ , neboli prokázal, že v sezóně 2020/2021 se významně snížil rozdíl mezi počty gólů vstřelených domácími a hosty. Nicméně tento rozdíl je stále kladný, a to statisticky významně, test může opět být založen na podobné veličině  $Z_1 = (\bar{S}_1 - 0) / \sqrt{\text{var}(S_1)/n_1}$ , která má hodnotu 2,5908 a tedy odpovídající  $p$ -hodnota v rámci standardního normálního rozdělení je 0,0048.



Obrázek 6: Počty domácích výher v jednotlivých kolech hokejové extraligy (tečkovaně), vyrovnané pro každou sezónu klouzavým oknem šířky 7 (plné křivky).

#### 4. Výsledky pro hokejovou Tipsport extraligu

Pro českou hokejovou extraligu jsem provedl jen zkrácenou analýzu, na základě dat – výsledků zápasů, které jsou k dispozici na webových stránkách <https://www.sport.cz/sekce/hokej-extraliga-vysledky-23>.

Porovnával jsem tytéž sezóny jako pro fotbal, resp. jejich základní části. Česká hokejová extraliga má (zpravidla) 14 týmů, hraje se čtyřikrát každý s každým, tj. 52 kol po sedmi zápasech. Výjimkou byl ročník 2021/2022, kde bylo 15 týmů, takže se hrálo 60 kol (také po sedmi zápasech, 1 tým měl vždy volno). Základní část sezóny 2019/2020 se stihla dohrát před vypuknutím pandemie (play-off se pak už nehrálo). V sezóně 2020/2021 to bylo opět tak, že do října 2020 byly počty diváků omezeny, říjen se vynechal a od listopadu se hrálo zcela bez diváků.

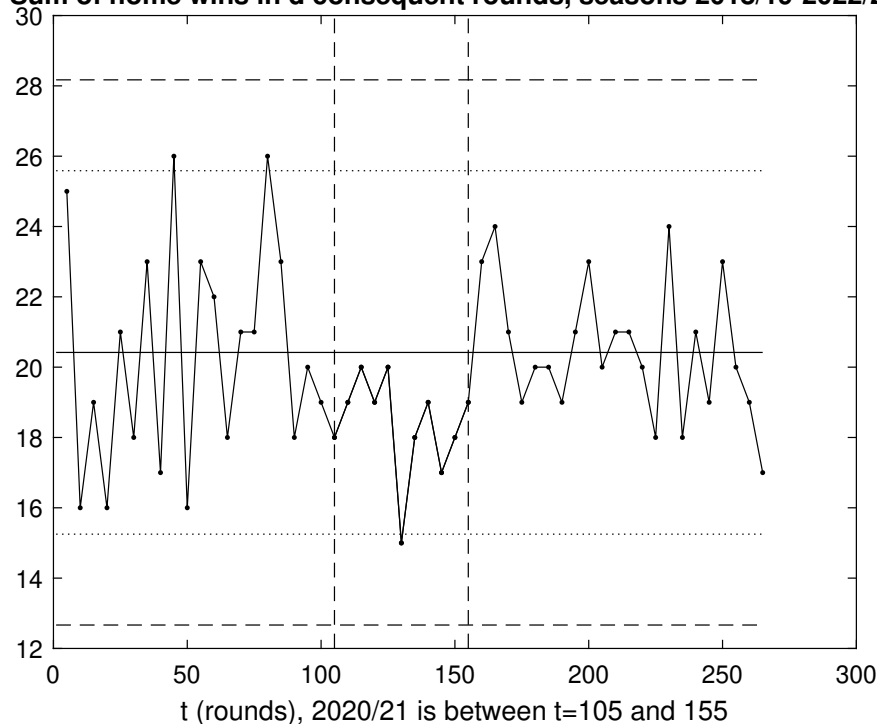
Porovnával jsem nyní jen počty vítězství domácích, po jednotlivých kolech (s tím, že jsem zahrnul i vítězství v prodloužení či po nájezdech). Obrázek 6 porovnává tyto počty v kolech sezón 2020/2021 a 2021/2022. Je vidět, že v průměru je počet vítězství domácích menší v „covidové“ sezóně. Průměry jsou 3,6538, 4,2167 (tomu odpovídá centrální linie v grafu). Po vydělení sedmi dostanu relativní četnosti 0,5220 a 0,6024. Test jejich rovnosti (v  $2 \times 2$  kontingenční tabulce, viz 1. část Tabulky 4) vede k zamítnutí této hypotézy, s  $p$ -hodnotou 0,0235.

Pokud porovnávám sezónu 2020/2021 se všemi čtyřmi ostatními, dostanu z oněch 4 průměr 4,0833 vítězství domácích na 1 kolo, tj. relativní četnost 0,5833, dohromady z 216 kol  $\times$  7 zápasů. I v příslušné kontingenční tabulce (2. část Tabulky 4) hypotézu nezávislosti zamítám, s  $p$ -hodnotou 0,0337.

Nabídnou ještě jedno porovnání: Sečetl jsem počty domácích vítězství vždy za 5 kol (tj. za kola 1–5, 6–10, atd.), čili z celkem 268 kol jsem dostal 53

Sezóna	Dom. výhry	Dom. prohry	Součty
2020/2021	190	174	364
2021/2022	253	167	420
Součty	443	341	784
2020/2021	190	174	364
Ostatní	882	630	1512
Součty	1072	804	1876

Tabulka 4: Četnosti domácích výher a proher.

Sum of home wins in  $d$  consequent rounds, seasons 2018/19-2022/23

Obrázek 7: Počty domácích vítězství za  $d = 5$  kol po sedmi zápasech. Linie jsou průměr a  $\pm 2$  sigma, 3 sigma pásy, počítané bez ročníku 2020/2021.

hodnot, z toho 10 hodnot odpovídalo sezóně 2020/2021. Výsledek je na Obrázku 7. Jeden bod tak odpovídá součtu 35 výsledků Bernoulliho pokusů. Při  $n = 35$  už s touto veličinou mohu zacházet jako s normální, a snažím se ukázat, že ve zkoumané sezóně má jiné charakteristiky. Graf to i potvrzuje. Průměr pro sezónu 2020/2021 byl 18,2690, pro ostatní 20,4186. Směrodatné odchylky byly 1,5055 a 2,5841, mediány 19 a 20.

$F$ -test zamítl rovnost rozptylů,  $s p = 0,0458$ . Hypotéza rovnosti středních hodnot byla pak testována pomocí přibližného  $t$ -testu pro různý rozptyl,  $s p = 0,0015$ , ale i testování shody mediánů testem Wilcoxonova prokázalo jejich nerovnost ( $s p = 0,0122$ ). Takže tímto způsobem se podařilo prokázat, že v sezóně 2020/2021 byla bilance vítězství domácích celků statisticky významně horší než v sezónách ostatních. Ale i za nepřítomnosti diváků působí

domácí prostředí motivačně, relativní četnost domácích vítězství byla stále nadpoloviční, i když pozorovaná proporce 0,5220, z 364 utkání, není statisticky významně odlišná od 0,5;  $p$ -hodnota jednostranného testu (v rámci CLV pro binomické rozdělení) byla 0,2.

## 5. Závěrečné poznámky

Průzkumová a grafická analýza ukazuje zpravidla na pokles úspěšnosti domácích týmů (i když úspěšnost je stále nadpoloviční). Výsledky statistických testů (hypotéz „sezóna 2020/2021 se neliší“ v určité charakteristice, se snahou je zamítnout) jsou většinou statisticky významné, a někdy jen „na hraně“ (typicky s  $p$ -hodnotou kolem 0,1). I takový výsledek je cenný, dá se interpretovat i tak, že domácí výhoda nespočívá jen v podpoře obecenstva na stadionu (zvláště když na většinu utkání české fotbalové ligy chodí běžně jen pár tisíc diváků). I Pollard (1986) ve svém rozboru uvádí diváky jen jako jeden z faktorů výhody domácího prostředí. A podobné závěry jsem našel i v práci Wunderlich et al. (2021), která už se taky zabývá důsledky nepřítomnosti diváků během pandemie:

„More than 40,000 matches before and during the pandemic, including more than 1,000 professional matches without spectators across the main European football leagues, have been analyzed. Results support the notion of a crowd-induced referee bias: the increased sanctioning of away teams disappears in the absence of spectators with regard to fouls ( $p < 0.001$ ), yellow cards ( $p < 0.001$ ), and red cards ( $p < 0.05$ ). Moreover, the match dominance of home teams decreases significantly as indicated by shots ( $p < 0.001$ ) and shots on target ( $p < 0.01$ ).“

Domnívám se, že počet faulů jednoho týmu souvisí s aktivitou týmu druhého (je víc obranných zákroků, i nečistých), není to tedy jen případný důsledek strannosti rozhodčích. Dále, ten významný pokles aktivity domácích týmů je trochu v rozporu s tím, že se to (podle autorů) neprojevovalo dostatečně ve výsledcích, počtu gólů a bodů:

„However, in terms of the home advantage itself, surprisingly, only a non-significant decrease has been found. Spectators thus do not seem to be the main driving factor of the home advantage.“

Soudím, že mně se, na „mých“ datech, podařilo prokázat statisticky významný pokles vlivu domácího prostředí na výsledky.

## Literatura

- [1] Dixon, M. J., Coles, S. G. (1997): Modelling association football scores and inefficiencies in the football betting market. *Appl. Stat.* **46**, 265–280. *cit. 4*
- [2] Hejdušek, M. (2007): *Matematické metody pro modelování a predikci výsledků sportovních utkání*. DP na KM FJFI ČVUT v Praze. *cit.*
- [3] Karlis, D., Ntzoufras, I. (2003): Analysis of sports data by using bivariate Poisson models. *J. R. Stat. Soc.* **D 52**, 381–394. *cit.*
- [4] Maher, M. J. (1982): Modelling association football scores. *Stat. Neerlandica* **36**, 109–118. *cit. 4*
- [5] Marek, P., Vávra, F. (2020): Comparison of Home Advantage in European Football Leagues. *Risks* **87**. <https://doi.org/10.3390/risks8030087>. *cit. 5*
- [6] Pollard, R. (1986): Home advantage in soccer: A retrospective analysis. *Journal of Sport Sciences* **4**, 237–248. *cit. 5, 6*
- [7] Rambousková, D. (2007): *Pravděpodobnostní modely a metody pro analýzu sportovních výsledků*. DP na KPMS MFF UK v Praze. *cit.*
- [8] Volf, P. (2008): On model for difference of two Poisson variables. *In: Proceedings of the ICPM 2008*, TU Liberec, 34–44. *cit. 5*
- [9] Volf, P. (2009): A random point process model for the score in sport matches. *IMA Journal of Management Mathematics* **20**, 121–131. *cit. 5*
- [10] Volf, P. (2017): Competing risk model with a non-traditional application. *Informační bulletin České statistické společnosti* **28**, 1, 1–11. *cit.*
- [11] Wunderlich F., Weigelt M., Rein R., Memmert D. (2021): How does spectator presence affect football? Home advantage remains in European top-class football matches played without spectators during the COVID-19 pandemic. *PLoS ONE* **16**, 3: e0248590. <https://doi.org/10.1371/journal.pone.0248590>. *cit. 15*

## Obsah

### Vědecké a odborné články

*Petr Volf*

Vliv Covidu na výsledky sportovních utkání:

Zhoršily prázdné ochozy bilanci domácích týmů? ..... 3

*Pavel Stríž*

Generování pseudonáhodných čísel po cifrách ..... 17

### Zprávy a informace

*Tomáš Chupáň*

Statistické dny 2022 a 2023 ..... 23

*Redakce*

Pozvánka na Statistické dny 2024 ..... 29

*Redakce*

Šťastný nový rok 2024! ..... 31

**Informační Bulletin České statistické společnosti** vychází čtyřikrát do roka v českém vydání. Příležitostně i mimořádné české a anglické číslo. Vydavatelem je Česká statistická společnost, IČ 00550795, adresa společnosti je Na padesátém 81, 100 82 Praha 10. Evidenční číslo registrace vedené Ministerstvem kultury ČR dle zákona č. 46/2000 Sb. je E 21214. Časopis je sázen v programu  $\text{T}_{\text{E}}\text{X}$ , ve formátu  $\text{LuaHBT}_{\text{E}}\text{X}$  s písmy balíku  $\text{C}_{\text{S}}\text{fonts}$ .

The Information Bulletin of the Czech Statistical Society is published quarterly.  
The contributions in the journal are published in English, Czech and Slovak languages.

**Předseda společnosti:** Mgr. Ondřej Vencálek, Ph.D., Katedra matematické analýzy a aplikací matematiky, Přírodovědecká fakulta Univerzity Palackého, 17. listopadu 12, 771 46 Olomouc, e-mail: [ondrej.vencalek@upol.cz](mailto:ondrej.vencalek@upol.cz).

**Redakce:** prof. RNDr. Gejza DOHNAL, CSc. (šéfredaktor), prof. RNDr. Jaromír ANTOCH, CSc., doc. RNDr. Zdeněk KARPÍŠEK, CSc., RNDr. Marek MALÝ, CSc., doc. RNDr. Jiří MIČHÁLEK, CSc., prof. Ing. Jiří MILITKÝ, CSc., doc. Ing. Iveta STANKOVIČOVÁ, PhD., Mgr. Ondřej VENCÁLEK, Ph.D.

**Redaktor časopisu:** Mgr. Ondřej VENCÁLEK, Ph.D., [ondrej.vencalek@upol.cz](mailto:ondrej.vencalek@upol.cz).  
Informace pro autory jsou na stránkách společnosti, <http://www.statspol.cz/>.

**DOI:** 10.5300/IB, <http://dx.doi.org/10.5300/IB>  
**ISSN 1210–8022 (Print), ISSN 1804–8617 (Online)**

Toto číslo bylo vtištěno s laskavou podporou Českého statistického úřadu.