

TDK-dolgozat

Gyarmati László

Egy Általános Thurstone Módszer és Előnyei - Sporteredmények Kiértékelésére Alkalmazva

Szerző: Gyarmati László 1. évfolyamos mesterszakos
programtervező informatikus

Témavezető: Dr. Mihálykóné dr. Orbán Éva egyetemi
docens, Dr. Mihálykó Csaba egyetemi docens

Készült: Pannon Egyetem Műszaki Informatikai Kar
Matematikai Tanszékén

Tartalomjegyzék

Köszönetnyilvánítás.....	Hiba! A könyvjelző nem létezik.
1. Bevezetés.....	5
2. Páros összehasonlítási módszerek.....	7
2.1. AHP	8
2.2. LLSM	10
2.3. Sorösszeg módszer	13
2.4. Thurstone módszer	15
2.4.1. Általánosított Thurstone modell.....	17
2.4.2. Jövőbeli valószínűségek számítása Thurstone módszerrel	23
2.5. A loglikelihood függvény maximumhelyének keresése.....	24
2.6. Ranglisták felállításának informatikai háttere	25
2.6.1. A Nelder-Mead optimalizáló.....	26
2.6.2. A loglikelihood függvény.....	28
2.7. Spearman-rangkorreláció	31
2.8. Konfidenciaintervallum számítása	32
3. Alkalmazás.....	33
3.1. Labdarúgás és kézilabda.....	33
3.2. Az NBI-es futball csapatok rangsorolása	35
3.2.1. Eredmények kialakulásának valószínűsége.....	37
3.2.2. Jövőbeli valószínűségek.....	38
3.3. A női kézilabda Bajnokok Ligájában játszó csapatok rangsorolása.....	40
3.3.1. Várható erősségek megadása, rangsorok a csoportkörben.....	40
3.3.2. A csoportok összefésülése.....	44
3.3.3. A továbbjutó csapatok és a négyes döntő előrejelzése.....	48

4.	Összefoglalás.....	54
5.	Jegyzékek	56
5.1.	Irodalomjegyzék	56
5.2.	Ábrajegyzék.....	57
5.3.	Táblázatjegyzék.....	57
6.	Mellékletek.....	59
6.1.	Táblázat mellékletek.....	59

1. Bevezetés

Minden egyes döntés előtt mérlegelünk, megpróbáljuk a lehető legjobb választást meghozni, gondolatainkkal a jövőbe utazni és előre jelezni, mi is fog történni, hogyan is érdemes dönteni, hogy az később megtérüljön. Érvényes ez például, amikor különböző termékeket szeretnénk különböző szempontok figyelembevételével rangsorolni, de nincs ez máshogy a sportok esetén sem, ahol még ha nem is fogadunk, mégis szeretnénk előre látni, hogy melyik csapat fog továbbjutni, melyik az erősebb és mennyivel, és nagy örömmel tölt el bennünket, ha a megérzéseink, előrejelzéseink igaznak bizonyulnak.

TDK dolgozatomban egy ilyen, az összehasonlítandó objektumok rangsorolásához, erősségük mértékének meghatározásához, illetve későbbi összehasonlításaik kimenetelének előrejelzéséhez (például sportban) alkalmas teljesen általános módszert vezetek be.

Célul tűzöm ki a módszer már egyszerűbb változataiban is jelentkező előnyös tulajdonságainak bemutatását valós adatokra alkalmazva a módszert. Többek között megmutatom, hogy számszerűsíteni lehet azt, ha előnyben van bizonyos szituációban egy összehasonlítandó objektum; hogy a módszerrel bizonyos feltételek mellett nemteljes összehasonlítások esetében is lehet érvényes rangsort felállítani; hogy össze lehet fésülni különböző részsorokat; hogy megbízható előrejelzést lehet adni jövőbeli összehasonlítások kimenetelére, például mérkőzések eredményére.

Ehhez a dolgozatom második felében felhasználok a női kézilabda Bajnokok Ligája (BL) 2020/2021-es szezonjának, illetve a magyar labdarúgó NBI 2019/2020-as bajnokságának eredményeit.

TDK dolgozatomban több, különböző specifikációjú Thurstone módszert alkalmazok, amikkel kapott rangsorokat, erősségeket összehasonlítom a sorösszeg [6] és a logaritmikus legkisebb négyzetek módszerével (Logarithmic Least Squares Method, LLSM) [1] kapott eredményekkel, rámutatva ezáltal arra, hogy valóban

Bevezetés

érdemes használni a közel sem egyszerű Thurstone módszert az objektumok rangsorának megállapításához, eredmények előrejelzéséhez.

A dolgozat eredményeinek egy részét publikálni szeretnénk [9, 10, 12].

Remélem, dolgozatomban sikerül az alkalmazásokon keresztül is bemutatni a Thurstone modellben rejlő plusz lehetőségeket a többi módszerrel összehasonlítva, és rávilágítani, hogy jobban alkalmazható egyrészt előny számszerűsítésére és használatára, másrészt jövőbeli összehasonlítások eredményei valószínűségeinek megadására, harmadrészt csoportok, részrangsorok összefésülésére.

2. Páros összehasonlítási módszerek

Mindig is megvolt az igény arra, hogy objektumokat rangsoroljunk, megállapítani az erősségeiket, ezek alapján sorba rendezni őket és így a jövőbeli összehasonlításokra előrejelzést adni. Nagyon könnyű egy-két objektumot összehasonlítani, de mikor tíz vagy annál több objektumot kell összevetni, akkor ez nagyon nehézkes, sőt lehetetlen feladat. Ezen probléma megoldásaként alakultak ki a páros összehasonlítási módszerek [1, 6, 7, 8, 9, 10, 11, 12, 14, 15], melyek esetén páronként hasonlítjuk össze az objektumokat és ezen összehasonlítások eredményéből megfelelő kiértékeléssel egy matematikailag alátámasztott rangsort kapunk.

A páros összehasonlításoknak több módját ismerjük, a következő alfejezetekben a teljesség igénye nélkül röviden négy módszert mutatok be előnyeikkel és hátrányaikkal egyetemben, hogy ez által vázlatos képet lehessen alkotni az egyes módszerek tulajdonságairól, erősségeiről, gyengeségeiről. Később ezek közül a sorösszeg és az LLSM módszert részletesebben is kifejtem, illetve a Thurstone módszer egy új, teljesen általános változatát megadom és elemzem, illetve bizonyos speciális eseteit felhasználva bemutatom egyes jó tulajdonságait valós adatokra alkalmazva.

2.1. AHP

Páros összehasonlítások során leggyakrabban az Analitikus Hierarchia Folyamat (Analytical Hierarchy Process, AHP [14]) elnevezésű módszert alkalmazzák. Ebben az összehasonlítás eredményeit 1, 3, 5, 7, 9-es számokkal, illetve ezek reciprokaival adjuk meg. Általában a 9 a sokkal-sokkal jobbat, 7 a sokkal jobbat, 5 a lényegesen jobbat, 3 a jobbat jelenti, míg 1 esetén azonos erősségűeknek tekintjük az összehasonlítás során a két objektumot. Ezeknek a számoknak a reciprokait használjuk a sokkal-sokkal rosszabb, sokkal rosszabb stb. esetben. Itt az egyes számoknak az értelme az, hogy az egyik objektum hányszor jobb a másiknál. Az összehasonlítások szimmetriája miatt, ha megvan az $a_{i,j}$ elem, akkor értelemszerűen ennek reciproka lesz az $a_{j,i}$ elem. Amennyiben több összehasonlítás (k db) van két objektum között, akkor az egyes összehasonlítások értékeit összeszorozzuk és vesszük a szorzat k -dik gyökét (mértani közép). Ez azért szerencsés, mert megőrzi a reciprocitás tulajdonságát.

Mivel a fenti kategóriák nem egzaktul megadottak, így a szubjektíven megadott mérőszámok vita tárgyát képezhetik. Példának okáért vegyük a következőt: 100 meccsből az egyik csapat a másik ellenében megnyer 65-öt, akkor most ezt a sokkal jobbnak megfelelő 7-es számmal jelöljük, vagy a lényegesen jobbnak megfelelő 5-tel jelöljük? Pedig ez még egy egyszerűbb, számszerűsíthető eset, képzeljük el, hogyan és hol húzható meg a sokkal jobb és a lényegesen jobb határa két joghurt összehasonlításakor, nem is beszélve arról, hogy ez hogy tehető meg, ha több véleményező értékkel.

Miután minden egyes páros összehasonlításra kiszámoltuk a saját határaink által megszabott értékeket, ezekkel feltöltjük az összehasonlítások mátrixát. Majd megkeressük a legnagyobb sajátértékhez tartozó egyre normált sajátvektorát az A mátrixnak, ez adja meg az objektumok súlyait és ezeken keresztül a rangsort (minél nagyobb a súly annál előrébb áll az objektum a rangsorban). A Perron-Frobenius tétel biztosítja, hogy mindig lesz egyértelműen meghatározható súlyvektor.

A módszer előnye, hogy kis számításigényű, gyorsan megállapítható a rangsor, viszont ez a módszer csak akkor használható, ha minden egyes

összehasonlítás megtörtént, ami kevés összehasonlítandó objektum esetén még lehetséges, vagy például egy körmérkőzéses bajnokság esetén, de sok objektum esetében, vagy egy kieséses bajnokságban már nem, így ilyen esetekben, nemteljes összehasonlításkor nem alkalmazható ez a módszer. További hátrány, hogy nem ad valószínűséget a jövőbeli összehasonlítások/mérkőzések kimenetelére, továbbá nem alkalmas az előny kezelésére, ezek mellett nem egzaktul meghatározott a kategóriákba való besorolás, ami nehezíti a módszer alkalmazhatóságát és annak pontosságát. Hátránya még a módszernek, hogy akkor ad igazán jó kiértékelési eredményt, ha a vélemények konzisztensek, ami gyakran nem teljesül, különösen nem sport eredmények esetén. Hányszor tapasztalunk a sportban igen meglepő eredményeket, például váratlan vereségeket, körbeveréseket!

2.2. LLSM

A logaritmus legkisebb négyzetek módszere (LLSM) [1] hasonlóan az AHP módszerhez, egy reciprokok szimmetrikus mátrixból indul ki, ami akár az AHP mátrix is lehet. A kiértékelés módszere azonban más, amelyben az AHP hátrányos tulajdonságaiból jónéhány már nem jelenik meg.

A kiértékelés során egy olyan $\underline{u}$ egyre normált, csupa pozitív koordinátájú sorvektort keressük, amely minimalizálja a logaritmus négyzetösszeg függvényt.

A logaritmus négyzetösszeg függvény teljes összehasonlítás esetén a következő:

$$\sum_{j=1}^n \sum_{i=1}^n \left(\ln a_{i,j} - \ln \left(\frac{u_i}{u_j} \right) \right)^2, \quad (1)$$

ahol $a_{i,j}$ jelenti az (i,j) pár összehasonlításának eredményét és u_i az i -dik csapat súlyát. Legyen i és j , $(i < j)$ csúcsok, ekkor a közöttük behúzott élt jelöljük $e(i,j)$ -vel, és ezek halmazát pedig jelölje H .

Az $A=(a_{i,j})$ mátrix elemeit Bozóki és társai a következőképpen adták meg [1]-ben: az (i,j) pár esetén, az i által nyert mérkőzések számát osztották a j által nyert mérkőzések számával, ez a hányados az $a_{i,j}$ eleme a mátrixnak. Például ha i 6-szor volt jobb j -nél és j 2-szer volt jobb i -nél, akkor az $a_{i,j} = 3$. Értelemszerűen ekkor $a_{j,i}$ az $a_{i,j}$ reciproka, vagyis $a_{j,i} = 1/3$. A módszerben, ha az (i,j) pár sosem lett összehasonlítva egymással, akkor az adott mátrixelemet elhagyták. Ehhez hasonlóan, ha például i 7-szer volt jobb j -nél, továbbá j 0-szor volt jobb i -nél, akkor sem adható meg az $a_{i,j}$ elem, mert a $7/0$ hányados nem létezik. Az ilyen eseteket többféleképpen kezelhetjük, de mindegyik módszer torzítást visz be az adatokba. Természetesen az is torzítás, ha elhagyjuk ennek az összehasonlításnak az eredményét.

Egy másik megközelítés lehet, hogy az AHP módszere szerint számoljuk az elemeket, i minden győzelméhez 3-t rendelünk (esetleg nagyobb arányú győzelmek esetén 5, 7 vagy 9-es szám is lehetséges, megállapodás szerint), és minden vereségéhez $1/3$ (vagy $1/5$, $1/7$, $1/9$) szorzótényezőt rendelünk, az esetleges

döntetleneknél pedig 1 a faktor értéke, majd ezeknek vesszük a mértani közepét. Így nem veszítünk el összehasonlításokat és nem torzítjuk az adatokat, és csak ott nem létezik az adott $a_{i,j}$ elem, ahol nem volt mérkőzés a két csapat között. Ennek az elem-megadási módszernek és az LLSM módszernek a kombinációját használjuk majd, és ezt nevezzük a későbbiekben AHP-LLSM módszernek.

Ahol az $a_{i,j}$ elem létezik valamilyen módszer segítségével meghatározva, ott $a_{i,j} > 0$ (és ekkor persze $a_{j,i} > 0$), és ez esetben van él i és j között; ahol nem létezik az $a_{i,j}$ ott nincs él i és j között.

Az $\underline{u} = (u_1, u_2, \dots, u_n)$ változóiban, ahol $\sum_{i=1}^n u_i = 1$ és $u_i > 0$, az alábbi függvényt minimalizáljuk:

$$\min_{\substack{\underline{u} \in R_+^n \\ \sum_{i=1}^n u_i = 1}} \sum_{e(i,j) \in H} \left(\ln a_{i,j} - \ln \left(\frac{u_i}{u_j} \right) \right)^2 \quad (2)$$

A feladat megoldása szempontjából elengedhetetlen, hogy a minimumhely létezzen és egyértelmű legyen.

A függvény minimumhelyének létezésére és annak egyértelműségére Bozóki és társai adtak szükséges és elégséges feltételt. A következő tételt mondták ki [1]:

1. Tétel

Az A kétdimenziós mátrix legyen az összehasonlítások mátrixa. Az A mátrixot reprezentáljuk egy gráffal, ebben a gráfban akkor van él i és j között, ha $a_{i,j} > 0$. Ha az előző szabály által képzett gráf összefüggő, akkor van a minimumfeladatnak megoldása és az egyértelmű.

Mivel a módszer nem tartalmaz valószínűségi számítási háttérrel, a módszer alkalmazásakor Bozóki és társai nem tudtak olyan eljárást mutatni, amelynek köszönhetően a jövőbeli mérkőzések kimenetele eredményének valószínűségét meg lehetne adni.

Az AHP-LLSM módszer óriási előnye, hogy nemteljes összehasonlítás esetén is használható. További nagyon jó tulajdonsága, hogy az optimalizálást lineáris egyenletrendszer megoldására lehet visszavezetni, ezért könnyen és

gyorsan számolható. Sajnos azonban, mint előbb említettem, nem képes valószínűséget adni a jövőbeli mérkőzések kimenetelére, továbbá az előnyt sem képes kezelni.

Megjegyezzük, hogy függetlenül attól, hogy hogyan számítjuk ki az A mátrix elemeit, az LLSM módszerrel meghatározhatóak a súlyok a fenti tétel értelmében, ha a tételben megadott egyszerű feltétel teljesül. Mivel a feltétel szükséges és elégséges, amennyiben a feltétel nem teljesül, nem lehet a kiértékelést elvégezni.

2.3. Sorösszeg módszer

Ennél a módszernél minden egyes objektum esetén azt adjuk meg, hogy az összehasonlítások hányadrésztében volt jobb a többiekénél. Jelölje n_i a „megnyert” összehasonlítások számát, o_i pedig az összes összehasonlítás számát i objektum esetén, ekkor n_i/o_i hányados adja meg az objektum erősségét. Ez az erősséget jellemző szám nullánál nem kisebb, egynél nem nagyobb. Ekkor a csapatok erejének mérőszámát az átlagos győzelmi arányuk adja meg. Ez a szám majdnem minden esetben képezhető.

Ezen módszer esetén a legegyszerűbb a feltétel, melynek teljesülése esetén érvényes rangsort ad a modell. Ez a feltétel a következő: mindegyik objektum legalább egy másikkal legyen összehasonlítva. Ez a feltétel szükséges is: ha valamelyik objektum esetén nincs összehasonlítás, akkor az az objektum nem kiértékelhető, így rangsor nem adható.

A módszer előnye, hogy gyorsan és könnyen számolható, matematikában nem jártas emberek is képesek a kiértékelés eredményét ellenőrizni, és ahogy fentebb említettem, egy nagyon egyszerű feltétel teljesülése esetén ad sorrendet.

Hátránya a módszernek, hogy ez sem képes jövőbeli összehasonlítások kimenetelére valószínűséget adni, továbbá, hogy az összes többi objektum ellen elért összegzett eredmény alapján adja meg az erősséget, amiből adódóan 3 megnyert összehasonlítás akár a legjobb, akár a legrosszabb objektum ellen ugyanannyit ér. Ezenkívül, mivel a módszer arányszámmal számol, így a 2/5 és 20/50 között nem tesz különbséget, pedig statisztikailag a több összehasonlítás biztosabb, mint a kevesebb. Nem képes az előny kimutatására sem. Hátránya még, hogy két nem azonos erősségű, egymástól független csoport összefésült rangsorát csak sok, a két csoport tagjait összevető összehasonlítással lehet megadni, ellenkező esetben nem képes helyes sorrendet szolgáltatni. Ezt később az egyik alkalmazás során részletesebben tárgyalom, megmutatva, hogy ebből a szempontból is a Thurstone módszer sokkal jobb.

Mivel a sorösszeg módszer eredményeit is össze szeretném hasonlítani az AHP-LLSM módszer által kapott súlyokkal, így a kapott arányokat (az objektumok erősségének vektorát) egyre kell normálni.

2.4. Thurstone módszer

A páros összehasonlítások esetén gyakran alkalmazott módszer a Thurstone módszer [15]. A módszer azon alapszik, hogy minden egyes kiértékelni kívánt objektum mögé látens valószínűségi változót rendelünk, vagyis a módszer valószínűségszámítási háttérrel rendelkezik.

Mikor két objektumot összehasonlítunk, akkor arról döntünk, hogy az összehasonlítás eredménye melyik kategóriába esik. Ezek a kategóriák a következők lehetnek: lényegesen jobb, jobb, egyforma, rosszabb, lényegesen rosszabb (esetleg egyéb további kategóriák). A kategóriáknak egy-egy intervallumot feleltetünk meg, melyek egymást kizáró intervallumok. A döntés meghozatalakor egyúttal voltaképpen arról (is) döntést hozunk, hogy az objektumok mögött megbújó valószínűségi változók különbsége melyik intervallumba esik.

Jelöljük $\underline{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ vektorral az objektumok mögött meghúzódó látens valószínűségű változókat.

Az i -edik objektum esetén ξ_i a látens valószínűségi változó, és m_i -vel jelöljük a várható értékét. Ez a várható érték adja meg az egyes objektum (átlagos) erősségét, ezért ezeket kell megbecsülnünk. Ezt a maximum likelihood módszer segítségével tesszük meg. Az $\underline{m} = (m_1, m_2, \dots, m_n)$ jelenti az objektumok várható erősségének vektorát, amely koordinátáit csökkenő sorba rendezve megadja az objektumok (csökkenő) sorrendjét.

A Thurstone által eredetileg kidolgozott módszer két kategóriát enged meg: jobbat és rosszabbat [15]. Vagyis i és j objektum összehasonlítása esetén végső soron arról történik döntés, hogy a $\xi_i - \xi_j$ különbség nagyobb vagy kisebb nullánál. Ha az i rosszabb j -nél, akkor az előbb említett különbség kisebb, mint nulla, ha viszont i jobb j -nél az adott összehasonlításban, akkor nagyobb, mint nulla a különbség.

Ezen klasszikus eset lehetséges változatai közül a Thurstone V modellt vettük alapul a modelljeinkhez. Ez azt feltételezi, hogy a ξ_i -k független, azonos

szórással rendelkező normális eloszlású valószínűségi változók. Feltehetjük továbbá az általánosság további megszorítása nélkül, hogy minden ξ_i szórása $1/\sqrt{2}$, tehát $\xi_i - \xi_j$ szórása $\sqrt{2} \cdot 1 / \sqrt{2} = 1$ lesz, míg a várható értéke $m_i - m_j$ lesz.

Normális eloszlású valószínűségi változókat használva a számításaimban, és Φ -vel jelölve a standard normális eloszlásfüggvényt, a következő egyenlőségek érvényesek:

$$\begin{aligned} P(i \text{ rosszabb, mint } j) &= P(0 < \xi_j - \xi_i) = \Phi(0 - (m_i - m_j)) \text{ és} \\ P(i \text{ jobb, mint } j) &= P(0 < \xi_i - \xi_j) = \Phi(0 - (m_j - m_i)) = \\ &1 - \Phi(0 - (m_i - m_j)). \end{aligned} \quad (3)$$

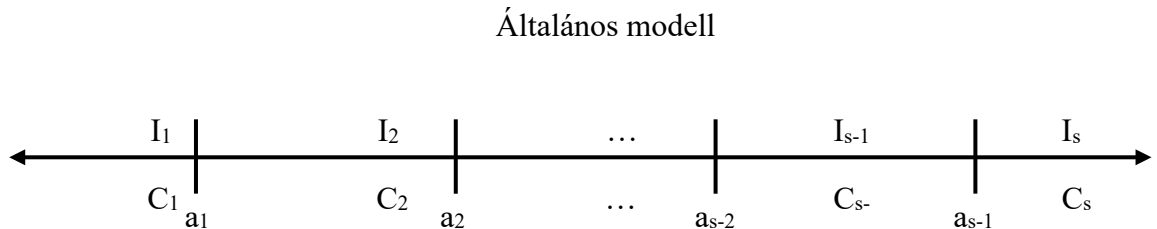
Thurstone a valószínűségeket a relatív gyakoriságokkal közelítette és a paramétereket a legkisebb négyzetek módszerével becsülte. Ez a módszer direkt módon több kategóriára nem általánosítható. Amint korábban írtam, Thurstone klasszikus módszere csak két kategóriát engedett meg. Azonban ki lehet terjeszteni a módszert több kategóriára oly módon, hogy minden egyes kategóriához egy intervallumot határozunk meg, és más paraméterbecslési módszert alkalmazunk. Ennek köszönhetően több kategóriát is tudunk alkalmazni. Ezt a következő alfejezetben részletesen kifejtem.

2.4.1. Általánosított Thurstone modell

Ebben az alfejezetben bemutatom a Thurstone módszer általánosításait, melyek nem csak az opciók számát érintik, de az előny és hátrány kezelését is. Ebben az alfejezetben tárgyalt általánosításokhoz [9] ad kiindulási alapot.

A ξ_i –kre vonatkozóan a korábban leírt tulajdonságokat itt is feltételezzük. Legyen $s \geq 2$ a kategóriák száma. A döntési kategóriákat, vagy más néven opciókat (például lényegesen rosszabb, rosszabb, egyforma, jobb, lényegesen jobb stb.) sorra $C_1, C_2, \dots, C_s$ jelölje, a nekik megfeleltetett intervallumokat pedig $I_1, I_2, \dots, I_s$. Az opciók kölcsönösen kizárják egymást, ennek megfelelően bármely két intervallum metszete üres, továbbá az összes intervallum uniója a valós számok halmazát adja. Ha s páros, akkor nincs középső, úgymond egyforma vagy döntetlen opció, míg s páratlan esetén van egyforma vagy döntetlen opció.

1. ábra Általános modell intervallumai és döntései



Feltesszük, hogy a döntések függetlenek, így ekkor a következőképpen írható fel modellhez tartozó likelihood függvény:

$$L(A|m_1, \dots, m_n, I_1, \dots, I_s) = \prod_{k=1}^s \prod_{i=1}^{n-1} \prod_{j=i+1}^n \left(P(\xi_i - \xi_j \in I_k) \right)^{A_{i,j,k}} \quad (4)$$

1.

ahol $A_{i,j,k}$ jelenti azt, hogy az i és j objektumok összehasonlítása során hányszor született C_k opciójú döntés. Ekkor $\xi_i - \xi_j = m_i - m_j + \eta_{i,j}$, ahol $\eta_{i,j}$ -k normális eloszlások használata esetén egymástól független standard normális eloszlású valószínűségi változók, összhangban az előző alfejezetben leírtakkal, s

mint ahogy a dolgozatban az alkalmazások során használjuk ezeket. Azonban lehetnek az $\eta_{i,j}$ -k általánosabb, háromszor folytonosan differenciálható, szigorúan logkonkáv sűrűségfüggvénnyel rendelkező valószínűségi változók is, amelyeknek F_{ij} -vel jelölhetjük az eloszlásfüggvényét.

Ha feltesszük, hogy a döntések egymástól függetlenek, ahogy azt előzőleg írtam, akkor az $A_{i,j,k}$ -k által meghatározott mintának a valószínűségét adja meg a likelihood függvény.

Az intervallumokat azok kezdő- és végpontjaik határozzák meg: $-\infty = a_0 < a_1 < a_2 < \dots < a_{s-1} < a_s = \infty$. Mivel a döntések szimmetrikusak így $a_i = -a_{s-i}$.

Normális eloszlásokat használva a következőképpen írjuk fel a valószínűségeket, a normális eloszlás eloszlásfüggvényét, Φ -t felhasználva:

$$P(\xi_i - \xi_j \in I_k) = \Phi(a_k - (m_i - m_j)) - \Phi(a_{k-1} - (m_i - m_j)) \quad (5)$$

Az általános s opciós likelihood függvény:

$$L(A|m_1, \dots, m_n, I_1, \dots, I_s) = L(A|m_1, \dots, m_n, a_1, \dots, a_{s-1}) = \prod_{k=1}^s \prod_{i=1}^{n-1} \prod_{j=i+1}^n \left(\Phi(a_k - (m_i - m_j)) - \Phi(a_{k-1} - (m_i - m_j)) \right)^{A_{i,j,k}} \quad (6)$$

$s = 2$ esetén a klasszikus „eredeti” Thurstone módszert adja a modell.

A paraméterbecslést a maximum likelihood módszer ([13]) segítségével végezzük. Ez azt jelenti, hogy keressük azon $\underline{m}$ és $\underline{a}$ vektorokat, amelyek esetén a (6) függvény maximális, azaz a likelihood függvény maximalizálása m_i -kben és a_i -kben történik.

Feltételezve, hogy az első és utolsó intervallum kivételével ugyanakkora d távolság van az egymást követő intervallum végpontok között, tehát $a_k - a_{k-1} = d$, $k = 2, 3, \dots, s-1$, $0 < d$, akkor a fent leírt likelihood függvénynek létezik egyértelmű maximuma, az alábbi 2. Tétel feltételeinek teljesülése esetén [11]:

2. Tétel

Tegyük fel, hogy van legalább egy (i_1, j_1) pár, amelyre $0 < A_{i_1, j_1, k}$ valamely $k = 2, \dots, s-1$ esetén. Továbbá van legalább egy olyan (i_2, j_2) pár is, melyre igaz, hogy $0 < A_{i_2, j_2, l}$ és $0 < A_{i_2, j_2, k}$, ahol $1 \leq l < k - 1 \leq s - 1$. Legyen a G gráfunk a következő: csúcsai az $1, 2, \dots, n$ objektumok. Az i és a j csúcs akkor van összekötve, ha vagy $0 < A_{i, j, k}$ valamely $1 < k < s$ esetén, vagy $0 < A_{i, j, l}$ és $0 < A_{i, j, s}$ teljesül. Rögzítsük az m_l értékét nullának. Ha az így kapott G gráf összefüggő, akkor a likelihood függvénynek van maximuma és a maximumhely egyértelmű.

Azért fontos, hogy létezzen maximumhelye a függvénynek, mert különben nem lehet kiértékelni a likelihood függvény segítségével a bemeneti adatokat, míg az egyértelműség azért fontos, mert ha a függvény több helyen is felvenné a maximumát, akkor nem tudnánk egyértelmű rangsort felállítani és egyértelmű erősségeket megadni.

Az itt bemutatott modell az általános s opciós, de az objektumok esetleges (helyzeti) előnyét nem figyelembe vevő modell volt.

Azonban előfordulhat, hogy az egyik félnek előnye van az összehasonlítás során, például sakkban (fehérrel játszás esetén jobbak az esélyeink nyerni, mint feketével), vagy labdarúgás esetén a hazai pálya bírhat ilyen jelentőséggel. Ha az előnyt is ki akarjuk mutatni, akkor újabb megközelítést kell kidolgozni. Ilyen volt a háromopciós konstans előnyt alkalmazó modell [8]-ban. Ekkor minden csapat esetén ugyanakkora előnyt feltételezünk. Ennél általánosabb, ha az előny mértéke más és más lehet a különböző csapatok esetén. Ilyen a háromopciós csapatonkénti előnyt alkalmazó modell [9]-ben.

A [9]-ben szereplő elgondolás általánosítását fogom bemutatni.

Képzeld el, hogy mobiltelefonokat hasonlítunk össze. Ekkor, ha az összehasonlítás előtt egy meggyőző reklámot láttunk az egyik telefonról, akkor az előnyben van az összehasonlítás során, míg ellenkezőleg, ha botrány tört ki a termék körül, akkor hátrányban van. Ha egyik sem történt, akkor (az ő szempontjából) semleges helyzetben hasonlítjuk össze egy másik telefonnal,

amivel kapcsolatban ugyanezek a dolgok az összehasonlítás előtt szintén megtörténhettek. Így a döntéskor kilencféle szituáció képzelhető el, annak megfelelően, hogy a két összehasonlítandó telefon milyen helyzetben van.

Így a likelihood függvény s opció esetén ebben a modellben így írható fel:

$$L(A|m_1, \dots, m_n, e_1, \dots, e_n, h_1, \dots, h_n, I_1, \dots, I_s) = \prod_{(p,q) \in \{-1,0,1\}^2} \prod_{k=1}^s \prod_{i=1}^{n-1} \prod_{j=i+1}^n \left(P \left(\xi_i^{(p)} - \xi_j^{(q)} \in I_k \right) \right)^{A_{i,j,k}^{(p,q)}} \quad (7)$$

Most $\xi_i^{(p)} - \xi_j^{(q)} = m_i + f_i^{(p)} - (m_j + f_j^{(q)}) + \eta_{i,j}$, ahol $f_i^{(p)} = e_i$, ha $p = 1$, $f_i^{(p)} = 0$, ha $p = 0$, $f_i^{(p)} = -h_i$, ha $p = -1$. Itt e_i jelentse az előny mértékét, h_i a hátrány mértékét, vagyis $p = 1$ esetén az i objektum előnyben van, $p = -1$ esetén hátrányban, míg $p = 0$ -nál semleges helyzetben. Hasonlóan $f_j^{(q)}$ esetében. Míg $A_{i,j,k}^{(p,q)}$ annak a darabszámát adja meg, ahány k -adik típusú döntés született a (p, q) -nak megfelelő szituációban, amikor i -t j -vel hasonlítottuk össze.

Ezek alapján a következőképpen írható fel a likelihood függvény standard normális eloszlásfüggvényt alkalmazva:

$$L(A|m_1, \dots, m_n, e_1, \dots, e_n, h_1, \dots, h_n, a_0, \dots, a_s) = \prod_{(p,q) \in \{-1,0,1\}^2} \prod_{k=1}^s \prod_{i=1}^{n-1} \prod_{j=i+1}^n \left(\Phi \left(a_k - (m_i + f_i^{(p)} - (m_j + f_j^{(q)})) \right) - \Phi \left(a_{k-1} - (m_i + f_i^{(p)} - (m_j + f_j^{(q)})) \right) \right)^{A_{i,j,k}^{(p,q)}} \quad (8)$$

A modell sok új paraméter bevezetését igényli, s ezeket szintén a maximum likelihood becsléssel becsülhetjük. Látható, hogy háromszor három, azaz kilenc eset van az összehasonlítások körülményét illetőleg. Mi ebből a dolgozat második felében, az alkalmazások egyikében az előny-hátrány és hátrány-előny párost alkalmazzuk, míg a másik példában a semleges-semleges párost.

Azt a modellt, amikor $e_i = h_i$, de ezek az értékek csapatonként mások lehetnek, csapatonkénti előnyös modellnek nevezzük. Ez abban az esetben, ha

semleges eset az összehasonlításokkor nem lép fel, nem jelenti az általánosság megszorítását. Ha van semleges eset is, akkor ez megszorítást, egyszerűsítést jelent.

Konstans előnyös modellnek azt fogjuk nevezni, amikor egy további megszorítást alkalmazunk: ilyenkor $e_i = h_i = e$. A korábbiakat figyelembe véve az előnyt nem alkalmazó modell esetén nyilván $e_i = h_i = 0$.

Látható, hogy érdemes lehet használni a komplexebb modellt, amennyiben lehetőség van rá. Például, ha nem tudjuk, hogy vajon van-e előny, akkor kimutatható az is, ha ténylegesen nincs előny. Amennyiben mégis van előny, úgy a modell képes a meglévő előny meglétét jelezni és számszerűsíteni is. Természetesen a komplexebb modellek esetében a több paraméter és az összetettebb kiértékelési feltételek miatt általában nagyobb számú és alkalmas kategória-megoszlású összehasonlításra (adatra) van szükség.

A komplexebb modelleknek hátránya, hogy nehezebb biztosítani, hogy létezzen és egyértelmű legyen a maximumhely. Általánosan sajnos még nem ismert az ezzel kapcsolatos tétel. A csapatonkénti előnyös módszer esetén azonban sikerült elégséges feltételt adni a maximum létezésére és egyértelműségére, s ez szerepel is [9]-ben. A következő állítás igaz:

3. Tétel

Legyenek F_{ij} -k olyan eloszlásfüggvények, amelyekre $0 < F_{ij}(x) < 1$ teljesül, háromszor folytonosan differenciálhatók R -n, valamint sűrűségfüggvényük nullára szimmetrikus és logaritmus szigorúan logkonkáv. Legyen $2 < s$. Tegyük fel, hogy van olyan $i_1 < j_1$ pár, amelyre $0 < A_{i_1, j_1, k}^{(1, -1)} + A_{i_1, j_1, k}^{(-1, 1)}$, valamely $1 < k < s$ esetén, valamint, van olyan $i_2 < j_2$ és $k < l-1$ pár, amelyre $0 < A_{i_1, j_1, k}^{(1, -1)} \cdot A_{i_1, j_1, l}^{(1, -1)}$ vagy $0 < A_{i_1, j_1, k}^{(-1, 1)} \cdot A_{i_1, j_1, l}^{(-1, 1)}$ teljesül. Legyen $\underline{e} = \underline{h}$. Rögzítsük az m_1 értékét nullának. Amennyiben a G gráf (később definiálom) összefüggő és nem faszerkezetű, akkor a függvény $0 < d = a_k - a_{k-1}$, $a_k = -a_{s-k}$ $k = 2, \dots, s-1$, $a_0 = -\infty$ feltételek mellett felveszi a maximumhelyét $\underline{m}$, $\underline{d}$ és $\underline{e}$ változóiban és az egyértelmű.

A tételben szereplő G gráfot a következőképp definiáljuk. A gráf csúcsai legyenek az értékelni kívánt objektumok, ekkor az i -edik és j -edik csúcs ($i < j$) akkor van összekötve, ha van köztük összehasonlítás mind i előnyös, mind i hátrányos helyzete esetén:

$$0 < A_{i,j,1}^{(1,-1)} \text{ és } 0 < A_{i,j,s}^{(1,-1)}, \text{ vagy } 0 < A_{i,j,k}^{(1,-1)} \text{ valamely } 1 < k < s \quad (9)$$

és

$$0 < A_{i,j,1}^{(-1,1)} \text{ és } 0 < A_{i,j,s}^{(-1,1)}, \text{ vagy } 0 < A_{i,j,k}^{(-1,1)} \text{ valamely } 1 < k < s \quad (10)$$

tehát előnyben vagy egy nem szélső döntés, vagy mindkét szélső döntés megszületett az összehasonlítások során, továbbá hátrányban is hasonlóan vagy egy nem szélső döntés, vagy mindkét szélső döntés megszületett az összehasonlítások során.

2.4.2. Jövőbeli valószínűségek számítása Thurstone módszerrel

A Thurstone-féle módszerek kiemelkednek a páros összehasonlítási módszerek közül, ugyanis ezekkel a módszerekkel a jövőbeli összehasonlítások kimenetelére is lehet valószínűséget adni. Erre az előzőleg tárgyalt módszerek (AHP, LLSM, sorösszeg módszer) közül egyik sem képes.

Háromopciós modellekre mutatom be azt, hogyan lehet a fent említett valószínűségeket megadni. A többi esetben értelemszerűen hasonlóképp történik a kiszámítás. Most és a későbbiekben is $F_{ij} = \Phi$.

i rosszabb, mint j esetén a valószínűség

$$\Phi\left(a_1 - (m_i + f_i^{(p)} - (m_j + f_j^{(q)}))\right), \quad (11)$$

i egyforma j -vel esetén a valószínűség

$$\begin{aligned} & \Phi\left(a_2 - (m_i + f_i^{(p)} - (m_j + f_j^{(q)}))\right) \\ & - \Phi\left(a_1 - (m_i + f_i^{(p)} - (m_j + f_j^{(q)}))\right), \end{aligned} \quad (12)$$

i jobb, mint j esetén a valószínűség

$$1 - \Phi\left(a_2 - (m_i + f_i^{(p)} - (m_j + f_j^{(q)}))\right). \quad (13)$$

A háromopciós modell esetén $a_1 = -a_2$. Amennyiben m_i és f_i helyett a becsült értékeket használjuk, akkor becslést kapunk a valószínűségekre.

2.5. A loglikelihood függvény maximumhelyének keresése

A likelihood függvény helyett a logaritmusát használtam. Ezáltal a soktényezős szorzat összeggé szelődül, aminek köszönhetően könnyebben lehet optimalizálni. A két függvény maximumának értéke bár nem azonos, de a logaritmus függvény monotonitása következtében a maximumhelyük ugyanott található, és számunkra ez a lényeges.

Numerikusan optimalizáláshoz, maximumhely kereséséhez a sok alkalmazásban bizonyított Nelder-Mead módszert használtam, melynek segítségével gyorsan és hatékonyan kaptam meg a megoldást.

A függvény optimalizálása közepesen nagy számításigénnyel rendelkezik, a Nelder-Mead módszernek köszönhetően 10-20 objektum esetén 2500-5000 iteráció alatt 3-4 tizedesjegy pontossággal megadja a maximumhelyét a függvénynek. Erről még az elkövetkezendő fejezetekben részletesebben írok.

Egy várható értéket rögzítenünk (például nullának), mivel az optimalizálás a várható értékek különbségéről, az $m_i - m_j$ -ről dönt. Ezt az általánosság megszorítása nélkül megtehetjük.

Ahhoz, hogy a módszert össze tudjuk hasonlítani az AHP-LLSM-mel és a sorösszeggel, súlyvektort kell képeznünk az $\underline{m}$ vektorból. Ezt így tehetjük meg:

$$w_i = \frac{e^{m_i}}{\sum_{j=1}^n e^{m_j}}, \quad (14)$$

ahol w_i jelenti a súlyvektor i -edik koordinátáját, m_j pedig az $\underline{m}$ vektor j -edik koordinátáját.

Könnyen látható, hogy a w_i súlyok pozitívak és összegük 1. A súlyokat előállító transzformáció szigorúan monoton növekedő, ezért ugyanazt a sorrendet adja, mint a várható értékek vektora. A (14) transzformáció azt is biztosítja, hogy a súlyok értékét nem befolyásolja, hogy melyik m_i -t rögzítettük és mennyinek.

2.6. Ranglisták felállításának informatikai háttere

A módszereket alkalmazó programot C# .NET Core-ban írtam meg. A nyelv és a keretrendszer mellett azért esett a választásom, mert gyors, hatékony kód megírását teszi lehetővé, továbbá magasszintű programozási nyelv, mindazonáltal rengeteg függvénykönyvtár segíti a munkát.

Első lépésben letöltöttem az all sports api segítségével az NBI mérkőzéseit, továbbá a női kézilabda eredményeket [4]. JSON-ben voltak eltárolva az adatok, így Newtonsoft.Json függvénykönyvtárat használtam az adatok parsolásához.

A lementett adatokat SQLite adatbázisban tároltam el a későbbi használatra, továbbá csv reprezentációt is létrehoztam a bemeneti adatokból. Adatbázist azért használtam, hogy átláthatóan struktúráltan egy fájlban tárolhassam a bemeneti adatokat. Az SQLite ingyenes adatbáziskezelő program elégséges funkciót kínált számomra. Programomban a C# SQLite.Linq segítségével írtam és olvastam az adatbázist. Jó strukturáltsága miatt az adatbázis könnyen bővíthető volt új adatokkal. Az SQL nyelv segítségével a fentebb említett függvénykönyvtáron át kommunikáltam az adatbázissal a programból.

Az optimalizálás volt az egyik legnehezebb megoldásra váró feladata a programozásnak. Többtíz változós függvény szélsőértékének keresése nem egyszerű, ehhez, mint korábban említettem, a Nelder-Mead módszert hívtam segítségül. A Nelder-Mead módszer mellett a szimplex módszert is kipróbáltam [3], de végül a Nelder-Mead módszer mellett döntöttem, ugyanis jelentősen kevesebb iterációval, illetve idő alatt volt képes a szélsőérték megtalálására. A módszert a későbbiekben részletezem. A program futási idejére nagy hangsúlyt fektettem, ezért fix iterációs szám helyett előre meghatározott pontosság elérése esetén állítottam le a futtatást. Ez jelentős tempókülönbséggel szolgált a Nelder-Mead optimalizálások esetén, így mindössze egy-két másodpercbe telik az összes módszer lefuttatása a labdarúgás vagy kézilabda felhasznált adatai esetén.

A loglikelihood függvényre is kitérek későbbiekben, amely minden egyes Thurstone módszer lelkét adja.

Az eredményeket strukturáltan egy mappaszerkezetbe JSON szerializáló segítségével elmentettem. Innen töltöttem vissza őket, ha a megjelenítéshez szükséges word-ös táblázatokat akartam generálni.

MigraDoc-ot hívtam segítségül, hogy megjelenítsem a táblázatokat. Ezen függvénykönyvtárral C# kódból lehet rtf dokumentumot létrehozni, amelyekből szükség esetén pdf-et is készíthettem.

Az által, hogy kódból generáltam a táblázatokat, nem csak a megbízhatóságot tudtam növelni, de a táblázatok elkészítésének idejét is csökkenteni tudtam. A Thurstone módszer m_i paramétereinek megjelenítésén kívül a jövőbeli mérkőzések eredményén át, a súlyok és a konfidenciaintervallumok megadásán túl, a rangkorrelációs táblázatok és az egyes mérkőzések eredményeinek előrejelzését is C# programból valósítottam meg.

Az egész programot általánosítottam, jól paraméterezhető függvényekbe osztottam az egyes funkcionálisokat, továbbá megoldottam, hogy mindössze két gomb lenyomására a táblázatok nagy része elkészüljön.

2.6.1. A Nelder-Mead optimalizáló

A módszer a [4]-ben található módon implementáltam.

A programomban alkalmazott globális szélsőérték-kereső minden egyes lépéssel közelebb kerül a megoldáshoz. Teszteltem az optimalizálót és 10000 iteráció esetén maximum egy százados volt az eltérés 50 változós függvény esetén a valóságos szélsőértékhez.

Az algoritmus következőképpen keresi a szélsőértéket: n dimenzióban felvesszünk $n+1$ pontot úgy, hogy ne egy n dimenziós elfajuló szimplexet határozzanak meg a pontok. Ezek a pontok az iterációk alatt egyre közelebb kerülnek a szélsőértékhez. Legvégül a szélsőérték ε sugarú környezetében helyezkednek el a pontok. A loglikelihood függvény maximumát keressük, de ha annak -1-szeresét vesszük, akkor minimumát keressük a függvénynek. Az általam megírt kód minimumot keres a -1-szeres függvényértékek között.

A következő oldalon található az optimalizáló kódja, melyet nem csak a loglikelihood függvény optimalizálására, hanem az AHP-LLSM értékeinek kiszámítására is alkalmaztam. Kellő számú iteráció vagy pontosság mellett megadja a loglikelihood függvény maximumát, már amennyiben van maximuma a függvénynek.

A változók nevei és használatuk célja sorra a következő: *IterCount* megadja a maximális iterációszámot, illetve ha ez 0, akkor 5 tizedesjegy pontossáig keresi a szélsőértékét a függvénynek. A *full* tömb az *AllMiPoints* osztályokból áll, annak *point* vektorában található az m_i , b_i (esetenként e_i és h_i), továbbá a d_i változók, a *results* adattagja a számított függvényértéket adja vissza.

Az optimalizáló futási sebessége függ a függvény számítási komplexitásától (hány változós a függvény, milyen nehezen számítható műveleteket tartalmaz), továbbá attól, hogy hány tizedesjegy pontosságra törekszünk.

Páros összehasonlítási módszerek

A futtatásaim alkalmával a legkomplexebb modellek esetén is elégségesnek bizonyult a konstans 10000 iteráció, avagy az 5 tizedesjegy pontosságig számolt szélsőérték hely, ekkor a futási idő sosem volt több egy-két másodpercnél az összes módszerre.

2. ábra Nelder-Mead optimalizáló

```
while ((IterCount!=0 || (Math.Round(full[0].result,5)!=Math.Round(full[N].result,5) || double.IsInfinity(Math.Abs(full[0].result)))) && (IterCount==0 || g<IterCount))
{
    g++;
    if (g % 10000 == 0)
    {
        Console.WriteLine(g);
        double[] xc = new double[N];
        for (int i = 0; i < N; i++)
        {
            double sum = 0;
            for (int j = 0; j < N; j++)
            {
                sum += full[j].point[i];
            }
            xc[i] = sum / N;
        }
        double[] xr = new double[N];
        for (int i = 0; i < N; i++)
        {
            xr[i] = 2 * xc[i] - full[N].point[i];
        }
        double new_res = modelType == 0 ? Functions.LogLikeliHoodFunction(xr, matrix, teamCount, advantagesType, optionsCount, isNormal) : Functions.LLSM(xr, matrix, teamCount, sportType);
        if (new_res <= full[0].result)
        {
            double[] xe = new double[N];
            for (int i = 0; i < N; i++)
            {
                xe[i] = xc[i] + 1.0 * (xr[i] - xc[i]);
            }
            double new_res2 = modelType == 0 ? Functions.LogLikeliHoodFunction(xe, matrix, teamCount, advantagesType, optionsCount, isNormal) : Functions.LLSM(xe, matrix, teamCount, sportType);
            if (new_res2 < full[0].result)
            {
                temp.result = new_res2;
                temp.point = xe;
                full[N] = temp;
            }
            else
            {
                temp.result = new_res;
                temp.point = xr;
                full[N] = temp;
            }
        }
        else if (full[0].result < new_res && new_res < full[N - 1].result)
        {
            temp.result = new_res;
            temp.point = xr;
            full[N] = temp;
        }
        else if (full[N - 1].result <= new_res)
        {
            double[] xz = new double[N];
            if (full[N].result < new_res)
            {
                for (int i = 0; i < N; i++)
                {
                    xz[i] = 1.6 * xc[i] - 0.6 * xr[i];
                }
            }
            else
            {
                for (int i = 0; i < N; i++)
                {
                    xz[i] = 0.4 * xc[i] + 0.6 * xr[i];
                }
            }
            double new_res3 = modelType == 0 ? Functions.LogLikeliHoodFunction(xz, matrix, teamCount, advantagesType, optionsCount, isNormal) : Functions.LLSM(xz, matrix, teamCount, sportType);
            if (new_res3 < new_res && new_res3 < full[N].result)
            {
                temp.result = new_res3;
                temp.point = xz;
                full[N] = temp;
            }
            else
            {
                for (int i = 1; i < N + 1; i++)
                {
                    double[] xh = new double[N];
                    for (int j = 0; j < N; j++)
                    {
                        xh[j] = 0.5 * full[0].point[j] + 0.5 * full[i].point[j];
                    }
                    double new_res4 = modelType == 0 ? Functions.LogLikeliHoodFunction(xh, matrix, teamCount, advantagesType, optionsCount, isNormal) : Functions.LLSM(xh, matrix, teamCount, sportType);
                    temp.result = new_res4;
                    temp.point = xh;
                    full[i] = temp;
                }
            }
        }
    }
    Array.Sort<AllMiPoints>(full, (x, y) => x.result.CompareTo(y.result));
}
```

2.6.2. A loglikelihood függvény

Ebben az alfejezetben bemutatom az ötopcios csapatonkénti előnyt és hátrányt alkalmazó modell esetén a loglikelihood függvény kódját és ezt meg is magyarázom. Az egyszerűség kedvéért csak azt mutatom be, amikor a mintafüggvény az előny-hátrány és hátrány-előny esetében alkalmazható, további

kódrészletek segítségével megoldható, hogy a semleges és egyéb eshetőségeket is kezelje.

A módszer esetén használt A mátrixot, az összehasonlítások eredményeinek mátrixát a *matrix* nevű változó tárolja. Ez egy háromdimenziós mátrix, amelynek utolsó koordinátája dönt arról, hogy milyen helyzetben történt az összehasonlítás, azaz a (p,q) -ról, illetve arról, hogy milyen eredmény született az összehasonlítás során.

A bemeneti paramétereket a *points* nevű vektor tárolja, mely az m_i -ket, e_i -ket, h_i -ket, illetve a határoló pontokat tartalmazza.

A normális, illetve más típusú eloszlásfüggvény értékeinek számításához külső függvénytárakat használtam, így a program futtatásakor lehetett választani, hogy melyik eloszlásfüggvényt szeretnénk használni a számításokhoz.

A kódrészleten látható, hogy végigmegyünk a lehetséges (i,j) párosokon, melyekhez kiszámoljuk a határoló pontokat, amelyek sorra a következők: $a10, a11, \dots, a15, a20, a21, \dots, a25$. Az a betű utáni első szám, ha 1-es, akkor az i -edik csapat előnyben van, míg ha 2-es, akkor hátrányból indul.

Miután megkaptuk a határoló pontokat, kiszámoljuk az egyes intervallumokba esés valószínűségeit, majd ezeknek a logaritmusát vesszük, és ezeket változóknak tároljuk el, amelyeknek nevei *homebigLose*, *homesmallLose*, ..., *bigawayWin*.

A ciklusunk végén, amikor elfogytak a csapatpárok, a függvény visszaadja a loglikelihood függvény adott ponthoz tartozó értékét.

Teszteket végeztem a loglikelihood függvény és az optimalizálás pontosságát illetően. Előre beállított m_i -k, e_i -k alapján generáltam az A mátrixot, melyhez átlagosan 20000 mérkőzést generáltam egy-egy (i,j) páros esetén. Összevettem a beállított és a kiszámított m_i -ket és azt tapasztaltam, hogy 4 tizedesjegy pontosságra megegyeztek. Ez a teszt jól mutatja, hogy a módszer és az optimalizáló nagy pontosság elérésére alkalmas megfelelő mennyiségű adat mellett.

3. ábra LogLikeliHood függvény érdemi része

```
double h0 = points[teamCount * 2];
double h1 = points[teamCount * 2 + 1];
for (int i = 0; i < teamCount - 1; i++)
{
    for (int j = i + 1; j < teamCount; j++)
    {
        double adv_b = (points[teamCount*2+2 + j] + points[teamCount + i]);
        double disadv_b = (points[teamCount + j] + points[teamCount*2+2 + i]);
        double a10;
        double a11;
        double a12;
        double a13;
        double a14;
        double a15;
        double a20;
        double a21;
        double a22;
        double a23;
        double a24;
        double a25;
        if (isNormal)
        {
            a10 = normal.CumulativeDistribution(Double.NegativeInfinity);
            a11 = normal.CumulativeDistribution(-Math.Exp(h0) - Math.Exp(h1) + points[j] - points[i] - adv_b);
            a12 = normal.CumulativeDistribution(-Math.Exp(h0) + points[j] - points[i] - adv_b);
            a13 = normal.CumulativeDistribution(Math.Exp(h0) + points[j] - points[i] - adv_b);
            a14 = normal.CumulativeDistribution(Math.Exp(h0) + Math.Exp(h1) + points[j] - points[i] - adv_b);
            a15 = normal.CumulativeDistribution(Double.PositiveInfinity);

            a20 = normal.CumulativeDistribution(Double.NegativeInfinity);
            a21 = normal.CumulativeDistribution(-Math.Exp(h1) - Math.Exp(h0) - Math.Exp(h1) + points[j] - points[i] + disadv_b);
            a22 = normal.CumulativeDistribution(-Math.Exp(h0) + points[j] - points[i] + disadv_b);
            a23 = normal.CumulativeDistribution(Math.Exp(h0) + points[j] - points[i] + disadv_b);
            a24 = normal.CumulativeDistribution(Math.Exp(h1) + Math.Exp(h0) + points[j] - points[i] + disadv_b);
            a25 = normal.CumulativeDistribution(Double.PositiveInfinity);
        }
        else
        {
            a10 = logistic.DistributionFunction(Double.NegativeInfinity);
            a11 = logistic.DistributionFunction(-Math.Exp(h0) - Math.Exp(h1) + points[j] - points[i] - adv_b);
            a12 = logistic.DistributionFunction(-Math.Exp(h0) + points[j] - points[i] - adv_b);
            a13 = logistic.DistributionFunction(Math.Exp(h0) + points[j] - points[i] - adv_b);
            a14 = logistic.DistributionFunction(Math.Exp(h0) + Math.Exp(h1) + points[j] - points[i] - adv_b);
            a15 = logistic.DistributionFunction(Double.PositiveInfinity);

            a20 = logistic.DistributionFunction(Double.NegativeInfinity);
            a21 = logistic.DistributionFunction(-Math.Exp(h0) - Math.Exp(h1) + points[j] - points[i] + disadv_b);
            a22 = logistic.DistributionFunction(-Math.Exp(h0) + points[j] - points[i] + disadv_b);
            a23 = logistic.DistributionFunction(Math.Exp(h0) + points[j] - points[i] + disadv_b);
            a24 = logistic.DistributionFunction(Math.Exp(h0) + Math.Exp(h1) + points[j] - points[i] + disadv_b);
            a25 = logistic.DistributionFunction(Double.PositiveInfinity);
        }
        double homebigLose = Math.Log(a11 - a10);
        double homesmallLose = Math.Log(a12 - a11);
        double homedraw = Math.Log(a13 - a12);
        double homesmallWin = Math.Log(a14 - a13);
        double homebigWin = Math.Log(a15 - a14);

        double awaybigLose = Math.Log(a21 - a20);
        double awaysmallLose = Math.Log(a22 - a21);
        double awaydraw = Math.Log(a23 - a22);
        double awaysmallWin = Math.Log(a24 - a23);
        double awaybigWin = Math.Log(a25 - a24);

        if ((matrix[i][j][3] > 0))
        {
            sum += (matrix[i][j][3]) * homesmallWin;
        }
        if ((matrix[i][j][4] > 0))
        {
            sum += (matrix[i][j][4]) * homebigWin;
        }
        if ((matrix[i][j][2] > 0))
        {
            sum += (matrix[i][j][2]) * homedraw;
        }
        if ((matrix[i][j][0] > 0))
        {
            sum += (matrix[i][j][0]) * homebigLose;
        }
        if ((matrix[i][j][1] > 0))
        {
            sum += (matrix[i][j][1]) * homesmallLose;
        }

        if ((matrix[i][j][8] > 0))
        {
            sum += (matrix[i][j][8]) * awaysmallWin;
        }
        if ((matrix[i][j][9] > 0))
        {
            sum += (matrix[i][j][9]) * awaybigWin;
        }
        if ((matrix[i][j][7] > 0))
        {
            sum += (matrix[i][j][7]) * awaydraw;
        }
        if ((matrix[i][j][5] > 0))
        {
            sum += (matrix[i][j][5]) * awaybigLose;
        }
        if ((matrix[i][j][6] > 0))
        {
            sum += (matrix[i][j][6]) * awaysmallLose;
        }
    }
}
```

2.7. Spearman-rangkorreláció

A Spearman rangkorreláció általánosan bevett mérőszám két rangsor közti összefüggőség mérésére [2]. Ezt a módszert akkor lehet alkalmazni, amikor mindkét ranglistát azonos elemek alkotják. A kapott érték egy korrelációs együttható, ahol az adatokat az objektumok helyezései szolgáltatják. A kapott együttható -1 és 1 közötti szám. Ha a rangkorreláció értéke 1, akkor a két ranglista azonos, ha -1, akkor épp ellentétes. Amennyiben a rangkorreláció közel van egyhez, akkor a két sorrend egymástól nem tér el lényegesen, míg ha -1 körüli értéket kapunk, akkor a két sorrend nagyjából egymás ellentettje.

A rangkorrelációs együttható a következő képlet alapján számítandó:

$$r_s = \frac{\sum_i ((rg(x_i) - rga(x))(rg(y_i) - rga(y)))}{\sqrt{\sum_i (rg(x_i) - rga(x))^2} \sqrt{\sum_i (rg(y_i) - rga(y))^2}}. \quad (15)$$

Ha adott a sorba rendezendő elemek halmaza, a képletben x és y is az elemek helyezéseit tartalmazza a kétfajta sorba rendezés esetén. $rga(x)$ x rangjainak középértékét jelenti, míg $rg(x_i)$ az i -edik objektum rangját jelenti az x ranglistán; $rg(y_i)$ és $rga(y)$ esetén hasonlóan.

2.8. Konfidenciaintervallum számítása

Ismeretlen paraméterek becslésére statisztikában gyakran intervallumot adunk meg [13]. $1-\alpha$ szintű konfidenciaintervallumnak nevezik azt az intervallumot, ami a becslendő, ismeretlen paramétert $1-\alpha$ valószínűséggel tartalmazza. Egy esemény bekövetkezési valószínűségére az alábbi módon szerkeszthető $1-\alpha$ szintű konfidenciaintervallum:

$$\left[\frac{k}{n} - \frac{u_\alpha \sqrt{\frac{k}{n} \left(1 - \frac{k}{n}\right)}}{\sqrt{n}} ; \frac{k}{n} + \frac{u_\alpha \sqrt{\frac{k}{n} \left(1 - \frac{k}{n}\right)}}{\sqrt{n}} \right], \quad (16)$$

ahol n jelöli a kísérletek számát, k pedig jelenti azt, hogy hányszor következik be az esemény a kísérletek során, u_α pedig a standard normális eloszlástáblázatából kiolvasott kritikus érték, amelyre $u_\alpha = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$. Amennyiben $1 - \alpha = 0,99$, akkor $u_\alpha = 2,5783$. Ezen konfidenciaintervallumot fogom használni arra, hogy meghatározzam azt, hogy a győzelem, a döntetlen és a vereség valószínűsége milyen mértékben térhet el a relatív gyakoriságoktól 0,99 valószínűséggel.

3. Alkalmazás

3.1. Labdarúgás és kézilabda

Európában a két legnépszerűbb sportág a labdarúgás és a kézilabda. A teljesen általános modell speciális változatainak segítségével ennek a két sportágnak az eredményeit felhasználva szeretném bemutatni az elkövetkezendő alfejezetekben, hogy a Thurstone módszer jelentősen kiemelkedik a többi páros összehasonlítási módszer közül. Labdarúgás esetén az OTP BANK Liga (NBI) 2019/2020-as idényét választottam, a kézilabda esetén pedig a női Bajnokok Ligája 2020/2021-es eredményein keresztül mutatom be a Thurstone módszer előnyeit. Ennek érdekében összevetjük a módszereket az AHP-LLSM és a sorösszeg módszerekkel.

A páros összehasonlításokat a lejátszott mérkőzések jelentik. A győzelem „jobb”, a vereség „rosszabb”, a döntetlen „egyforma” eredményt jelent a páros összehasonlításokor. Az AHP-LLSM alkalmazásánál olyan problémába ütköztünk, hogy bár az NBI-es mérkőzések esetén teljes az összehasonlítás, de 66 páros esetén 17-ben az egyik csapat minden meccsét megnyerte a másikkal szemben. A kézilabda BL 8 csapatos csoportkörei esetében se volt jobb a helyzet, az összehasonlítandó 56 páros esetén 34-ben a páros egyik tagja ugyancsak minden pontot begyűjtött, ráadásul a szintén vizsgált nyolcaddöntős meccsek esetében az immár 16 összehasonlítandó csapatból alkotható párokra vonatkozó teljes összehasonlítás egyáltalán nincs már meg. Emiatt az AHP módszer szóba se jöhetett. Az LLSM módszer esetén, ha az egymás ellen szerzett pontok hányadosával akartuk volna meghatározni az A mátrix elemeit, akkor vagy rengeteg összehasonlítást el kellett volna hagynunk, ami információ veszteséssel járna, vagy sokszor kellett volna mesterségesen torzítani az adatokat. Ezért azt választottuk, hogy a korábban bemutatott AHP-LLSM kombinált módszert alkalmaztuk, amely képes ezeket a problémákat kezelni. Így nem volt gond, hogy nincsen meg minden összehasonlítás (azaz nem játszott minden csapat mindegyikkel), illetve nem szükséges mesterséges adatokkal belenyúlni az eredményekbe. A mátrix elemeit az

AHP által alkalmazott módszerrel alakítottuk ki, a győzelem 3-at, a döntetlen 1-et, a vereség $1/3$ -t ért, majd ezeket a számokat összeszoroztuk, és ennek a szorzatnak vettük az n -edik gyökét, ahol n jelenti azt, hogy hány összehasonlítás (azaz meccs) volt a két csapat között, majd az LLSM módszer segítségével számoltuk ki a csapatok súlyait. A módszer alkalmazhatóságának 2. tételbeli feltétele mindig fennállt. A szintén használt sorösszeg módszer esetében ilyen jellegű problémák nem voltak, minden csapat több összehasonlításban is részt vett, így az alkalmazhatóság feltétele minden esetben teljesült.

A meccsek eredményeként kialakuló győzelem, döntetlen, vereség lehetőségek miatt háromopciós modelleket alkalmaztunk. A használt háromopciós Thurstone modelleket a következőképp jelöljük:

Az előny nélküli modellt, amikor a $(p, q) = (0, 0)$, és az $e_i = h_i = 0$ feltételezéssel éltünk, 3_N-nel jelöltük. A konstans előnyös modellt, ahol a (p, q) $(1, -1)$ és $(-1, 1)$ lehetett, és $e_i = h_i = e$ volt, 3_B-vel, míg a csapatonkénti előnyös modellt, amikor is (p, q) $(1, -1)$ és $(-1, 1)$ lehetett, és $e_i = h_i = e_i$ volt, 3_C-vel jelöltük. A két utóbbit csak a futball eredmények elemzésekor használtuk, mert a pandémia miatt a kézilabda BL meccseit gyakran nézők nélkül, vagy csak korlátozott nézőszám mellett, illetve esetenként harmadik helyszínen játszották, így a hazai pálya előnye gyakorlatilag sokszor eltűnt, vagy esetlegessé vált ezekben az esetekben. Nagyon nehéz lett volna különbséget tenni így a mérkőzések között. Továbbá, (és ez legalább akkora súllyal számított a választásunkkor) a mérkőzésszám arányosan lényegesen kevesebb volt, így statisztikailag megbízhatatlanabb lett volna a maximum likelihood becslés.

3.2. Az NBI-es futball csapatok rangsorolása

NBI-es bajnokság 12 csapat között zajlott körmérkőzésesen. Mindenki játszott mindenkivel 3 mérkőzést, vagyis 198 mérkőzést játszottak összesen. A mérkőzéseket az ellenfelek egyikének stadionjában játszották. Ennek okán érdemes (és mint látni fogjuk hasznos) volt alkalmazni az előnyt beépítő modelleket is.

Ebben az alfejezetben a [9]-es cikk eredményeit bővítem ki és mutatom be.

1. táblázat A Thurstone modellek várható értékei és előny-mértékei

	Hivatalos eredmények		3_N		3_B		3_C		
Határ	-		0,33442		0,33739		0,35014		
Előny	-		-		0,07204		-		
Helyezés	Csapatnév	Pontok	Csapatnév	m_i	Csapatnév	m_i	Csapatnév	m_i	e_i
1	Ferencváros	76	Ferencváros	1,97798	Ferencváros	2,00753	Ferencváros	2,09698	0,3987
2	MOL Fehérvár	63	MOL Fehérvár	1,58292	MOL Fehérvár	1,60281	MOL Fehérvár	1,60428	-0,10105
3	Puskás Akadémia	54	Puskás Akadémia	1,37085	Puskás Akadémia	1,39858	Puskás Akadémia	1,40809	-0,1685
4	Mezőkövesd	50	Mezőkövesd	1,24511	Mezőkövesd	1,25475	Mezőkövesd	1,31699	-0,21669
5	Honvéd	44	Honvéd	1,09658	Honvéd	1,1118	Zalaegerszeg	1,12629	0,16819
6	Újpest	43	Zalaegerszeg	1,08491	Zalaegerszeg	1,09679	Honvéd	1,11468	-0,15589
7	Zalaegerszeg	43	Újpest	1,02981	Újpest	1,03938	Újpest	1,07338	-0,09228
8	Kisvárd	42	Paks	0,99847	Paks	1,01842	Paks	1,03872	0,2742
9	Diósgyőr	41	Kisvárd	0,98	Kisvárd	0,99193	Kisvárd	1,0351	0,2891
10	Paks	41	Diósgyőr	0,9357	Diósgyőr	0,95705	Diósgyőr	0,99845	0,12559
11	Debrecen	39	Debrecen	0,88289	Debrecen	0,88326	Debrecen	0,84506	0,28992
12	Kaposvár	14	Kaposvár	0	Kaposvár	0	Kaposvár	0	0,22304

Az 1. Táblázatban láthatóak a hivatalos eredmények, illetve az előnyt alkalmazó, illetve előnyt nem alkalmazó Thurstone modellek által adott várható erősségek, előny-mértékek.

A táblázatban látható, hogy a bajnokcsapat, a Ferencváros, jelentősen erősebb a többi csapatnál, illetve az utolsó Kaposvár csapata meg jelentősen gyengébb a többi csapatnál. Az 5-11-dik csapat között nincs nagy eltérés erősségek tekintetében. Amint látható, a Thurstone modellek között csak a 3_C modell esetén más a sorrend, a másik két módszer azonos sorrendet ad. Az eltérés a Honvéd és a Zalaegerszeg helycseréjéből származik, amely abból adódik, hogy a Zalaegerszeg otthoni pályán erősebb, míg a Honvéd nem, így helyet cserél az amúgy hasonló erősségű két csapat a rangsorban. A hivatalos ranglistától való eltérések: az Újpest és Zalaegerszeg helycseréje, amely csapatoknak ugyan ugyanannyi pontja van, de Zalaegerszeg erősebb csapatok ellen ért el több győzelmet és döntetlent. Másik jelentős különbség az, hogy a Paks nyolcadik helyen van a ranglistánkon a tizedik hely helyett. Ez részben abból adódik, hogy a Paks képes volt nyerni a második helyezett Mol Fehérvár ellen. A táblázatban a „határ” jelenti a döntetlen (egyforma) felső határát (a $0 < d$ paraméter becsült értékét), illetve annak ellentettje az alsó határát. Látható, hogy a csapatok összesítve a hazai pályát előnyként élik meg. Csapatonként vizsgálva az előny meglétét, az FTC odahaza nem is kapott ki, így pozitív nagy szám az előnykezelésének mértéke, míg Mezőkövesd idegenben volt jobb, így negatív szám az előnykezelésének mértéke (odahaza 7-szer nyert, 3-szor döntetlent ért el, 7-szer kikapott, idegenben 4-szer kikapott, 5 döntetlent ért el, 7-szer nyert). A modell szépen visszaadja az egyes csapatoknál, hogy idegenben, vagy otthon erősebbek és számszerűsíti is ezt. Erre, mint korábban írtam, nincs mód az AHP, LLSM, sorösszeg módszerek esetén.

A mellékletben található M1. táblázatban láthatóak a súlyai az egyes módszereknek, a sorösszeg esetén a döntetlenre 1 pontot adtam, míg a győzelemért 2-t (labdarúgás esetén hivatalosan 3 pont jár a győzelemért, 1 a döntetlenért, 0 a vereségért, ekkor viszont egy meccs végeredményétől függ az összesen kiosztott pontok száma, ami nincs igazán összefüggésben a csapatok erősségeivel). Az AHP-

LLSM módszernél a mátrix elemeit az előbbi alfejezetben tárgyalt módon számítottam.

Látható, hogy a két módszer azonos sorrendet ad a 3_N, 3_B módszerekkel, a 3_C-től az eltérés az előbbieken említett Honvéd, Zalaegerszeg helycsere. A 2. táblázat tartalmazza a rangkorrelációs értékeket.

2. táblázat Rangkorrelációs értékek egyes módszerek esetén

	3_N	3_B	3_C	sorösszeg	AHP-LLSM
3_N	1	1	0,99301	1	1
3_B	1	1	0,99301	1	1
3_C	0,99301	0,99301	1	0,99301	0,99301
sorösszeg	1	1	0,99301	1	1
AHP-LLSM	1	1	0,99301	1	1

3.2.1. Eredmények kialakulásának valószínűsége

A Thurstone módszerek segítségével becsülhetjük az egyes eredmények kialakulásának valószínűségét és ezt összevethetjük a relatív gyakorisággal. Ez a lehetőség nem adatik meg a többi módszer esetén. Az alábbi táblázatban láthatóak az eredmény kialakulásának valószínűségei, a relatív gyakoriságok, illetve a hozzájuk tartozó konfidenciaintervallumok.

3. táblázat A 3_B és 3_C módszerek által adott valószínűségek a győzelem, döntetlen és vereség kialakulására és a relatív gyakoriságok az összes csapatot figyelembevéve

Típus	Relatív gyakoriság	Konfidenciaintervallum	3_B	3_C
Győzelem	0,43434	[0,34361;0,52508]	0,43586	0,43484
Döntetlen	0,22222	[0,14612;0,29833]	0,22027	0,22182
Vereség	0,34343	[0,25651;0,43036]	0,34387	0,34334

Amint a táblázatban látható, mindkét módszer nagyon pontosan visszaadja az egyes eredmények relatív gyakoriságát. A mellékletben található M2. táblázatban megadtam az egyes csapatok különböző típusú eredményeinek relatív gyakoriságát, illetve a módszerek által adott valószínűségeket, továbbá zárójelben az eltérést abszolútértékben. A 3_B modell esetén az összes valószínűségre vonatkozó átlagos eltérés 0,06725 volt, tehát a módszer 6,725%-ban nem tudta visszaadni a relatív gyakoriság mértékét, a 3_C esetén azonban ez csak 0,04844 volt, ami jelentősen jobb, tehát újfent bizonyítást nyert, hogy érdemes használni a komplexebb modelleket/módszereket.

Ennek gyakorlati haszna is jelentkezik. Azzal ugyanis, hogy ki tudjuk mutatni az előnyt, illetve ezt számszerűsíteni is tudjuk, minden egyes csapatra pontosabb előrejelzést tudunk adni például egy otthon vagy idegenben lejátszott mérkőzés esetén, mint ezt részletesebben a következő alfejezetben bemutatom.

3.2.2. Jövőbeli valószínűségek

Ahogy említettem korábban, a Thurstone módszerek képesek a jövőbeli mérkőzések eredményének valószínűséget megadni, míg a többi módszer nem. Ezen jövőbeli mérkőzések eredményeinek valószínűségét a 3_C modell esetén az M3. táblázat tartalmazza. Itt nem vesszük figyelembe azt, hogy odahaza vagy idegenben játszották a mérkőzést, azaz nem vesszük figyelembe az előny/hátrány mértékét. Az M4. táblázat az idegenben játszás esetén tartalmazza a valószínűségeket, tehát e_i -vel csökkentjük a várható erősségét a csapatnak (az ellenfelét növeljük e_j -vel). Az M5. táblázat a hazai pályán való játszás esetén mutatja az egyes eredmények valószínűségeit, itt e_i -vel növeljük a csapat erősségét (az ellenfelét csökkentjük e_j -vel). A táblázatokban az értékek százalékban vannak megadva, az első sor a győzelemre vonatkozik, míg a második a döntetlen valószínűségére, a vereség pedig ezek alapján számítható, mert összesen 100%-nak kell lennie a három érték összegének. Jól látható például az FTC esetén, hogy ha figyelembe vesszük az előnyt, akkor kimutatható, hogy odahaza jóval nagyobb valószínűséggel nyer, mint amikor idegenben játszik.

Alkalmazás

Annak köszönhetően, hogy valószínűségek adhatók meg a jövőbeli mérkőzések eredményeire, figyelembe véve még az előny/hátrány mértékét is, pontosabban lehet előre jelezni két csapat egymás elleni eredményét, tehát jobb lesz a jövőbeli mérkőzések várható eredményének előrejelzése, amit fogadások nyereségesebbé tételében is fel lehet használni.

3.3. A női kézilabda Bajnokok Ligájában játszó csapatok rangsorolása

A női kézilabda Bajnokok Ligájában először két csoportban folynak a mérkőzések. Itt körmérkőzéses rendszer van, majd innen továbbjutva egyenes kieséssel menetelnek a döntőig a csapatok. A koronavírus következtében azonban nem sikerült megrendezni egyes mérkőzéseket, így azok eredményeit mesterségesen eldöntötték, a pontokat odaítélték a vétlen csapatnak. A lejátszott és a le nem játszott mérkőzések eredményeit fogom elemezni ebben az alfejezetben. A csapatok neve az EHF weboldala által használt nevekkel egyezik meg (ezért írom például rövid ő-vel a Györi Audi ETO KC nevét).

A [10] cikk eredményeinek bővítése, elmélyítése található ebben és a következő alfejezetekben, amelyekben végig a 3_N módszer szerint számoltam.

3.3.1. Várható erősségek megadása, rangsorok a csoportkörben

Az alábbi, 4. táblázatban láthatóak az A csoport eredményei, ahol a „Tényleges” jelenti a ténylegesen lejátszott meccsek alapján a pontokat és az erősségeket, az „Összes” jelenti azt, amikor a ténylegesen lejátszott meccsek mellett a pandémia miatt nem lejátszott, de a vétlen csapatnak odaítélt mérkőzéseket is beszámítjuk a pontok, illetve az erősségek kialakításába. E szerint alakult ki egyébként a hivatalos EHF eredmény is. A „Várható” pedig arra vonatkozik, hogy ha a Thurstone módszerből származó valószínűségeket figyelembe vesszük, és a nem lejátszott meccseket így „pótoljuk”, akkor várható értékben milyen eredmények, pontszámok születtek volna, és ezek alapján milyen erősségek számíthatók az egyes csapatoknak. Ezt a „pótlást” a következőképpen tettük meg: véletlenszerűen, de a tényleges esetnek megfelelő várható értékek szerinti valószínűségek alapján dönt el, hogy egy-egy mérkőzés generált eredménye döntetlen, vereség vagy győzelem lett az egyik, illetve a másik csapat szempontjából. Ezt egymilliószor, egymástól függetlenül megismételtük, és az így kapott pontok és várható értékek átlagát vettük. Ezeket a várható pontszámokat és

várható értékeket a többi módszer (AHP-LLSM, sorösszeg) segítségével nem lehetett volna megadni.

Ugyanezen eredményeket a B csoport esetében az 5. táblázatban találjuk.

4. táblázat Az A csoport csapatainak erősségei és pontjai

Típusok		Tényleges		Várható		Összes	
Határ		0,17021		0,17206		0,15436	
Helyezés	Csapatnév	m_i	Pontok	m_i	Pontok	m_i	Pontok
1	Rostov-Don	1,84598	19	1,87066	21,74481	1,84421	21
2	Metz Handball	1,67112	18	1,68940	19,17710	1,83578	20
3	Vipers Kristiansand	1,44305	14	1,45936	17,35871	1,40503	16
4	FTC-Rail Cargo Hungaria	1,31122	16	1,32693	16	1,3892	16
5	CSM Bucuresti	1,2587	13	1,27377	15,34918	1,48992	17
6	Team Esbjerg	0,91373	10	0,92827	11,24325	1,05393	12
7	RK Krim Mercator	0,59941	7	0,61139	7,75675	0,60122	7
8	SG BBM Bietigheim	0	3	0	3,37020	0	3

Az A csoportban a le nem játszott meccsek következtében odaítélt pontok legnagyobb nyertese láthatóan a CSM Bucuresti csapata volt. Ennek köszönhetően az előtte lévő két csapatot is megelőzte az „Összes” szerint a „Ténylegeshez” képest pontszámban és erősségben. A kapott pontokon túl még az is számított, hogy a nála jelentősen erősebb Rostov-Don ellen lett az egyik odaítélt győzelme. A Team Esbjerg csapata is gazdagabb lett két odaítélt ponttal, de ezeket a pontokat a nála gyengébb RK Krim Mercator ellen kapta, amit nagy valószínűséggel amúgy is megszerzett volna, ha lejátszották volna a mérkőzést. Ezért az erőssége nem sokkal nőtt. Mivel a Rostov-Don ugyan szerzett odaítélt két pontot, de egy odaítélt veresége is volt a CSM Bucuresti csapatától, így erőssége összességében csökkent, és ennek köszönhetően a Metz Handball csapata majdnem be is érte erősség tekintetében, amely szintén kapott két odaítélt pontot. A Vipers Kristiansand pedig a fő vesztese lett az odaítéléseknek, így a „Ténylegeshez” képest kettővel hátrébb

csúszott az „Összes” szerinti sorrendben. Később ki is derült, hogy tényleges erő tekintve jobb az őt a hivatalos eredmény szerint megelőző CSM Bucurestinél, illetve FTC-Rail Cargo Hungariánál is.

A B csoport eredményei az alábbi táblázatban láthatóak:

5. táblázat A B csoport csapatainak erősségei és pontjai

Típus		Tényleges		Várható		Összes	
Határ		0,29938		0,33917		0,24326	
Helyezés	Csapatnév	m_i	Pontok	m_i	Pontok	m_i	Pontok
1	Györi Audi ETO KC	2,75571	24	2,78552	24	2,28354	24
2	CSKA	2,71809	23	2,74467	23	2,26729	23
3	Brest Bretagne Handball	2,08197	17	2,11088	18,62931	1,50622	17
4	Odense Håndbold	1,54553	13	1,56867	13	1,15546	13
5	Buducnost	1,36652	12	1,38806	12	0,99022	12
6	SCM Ramnicu Valcea	1,14625	6	1,16609	10,28657	0,75066	10
7	BV Borussia 09 Dortmund	1,0214	7	1,03827	8,81128	0,67483	9
8	HC Podravka Vegeta	0	2	0	2,27284	0	4

A B csoportban, részint azért, mert kevesebb volt az odaítélt mérkőzés, részint azért, mert azok sokkal kiegyenlítettebb módon alakultak, sorrendváltás nem történt. A nem lejátszott mérkőzések okán odaítélt pontok nyertese a HC Podravka Vegeta csapata volt. Ez a csapat jelentősen le volt maradva a többi csapattól erősségek tekintetében, de a „Tényleges” szerint hatodik SCM Ramnicu Valcea csapata ellen odaítélt győzelem hatására erősödött, így az előtte álló csapat várható értékétől vett eltérés jelentősen csökkent.

A mellékletben található M6., M7., M8. és M9. táblázatokból látható, hogy az egyes páros összehasonlítási módszerek esetén szinte mindig azonos sorrend született az A és a B csoport esetén is, mind a ténylegesen lejátszott, mind az összes meccs figyelembevételével.

A fent említett táblázatok szerint a 3_N módszer egyértelműen meg tudta határozni a ténylegesen lejátszott meccsek alapján mindkét csoportban az első kettő és az utolsó kettő csapatot. Vagyis megmaradhatott volna az eredeti kiírás, miszerint a csoportok első két helyezettje automatikusan, play-offs során játszott meccs nélkül bejut a legjobb nyolc közé, míg az utolsó kettő kiesik a további küzdelmekből, és csak a 3-6. helyezettek játszanak a nyolc közé jutásért. Így elkerülhetők lettek volna a felesleges A1-B8, A2-B7 és B1-A8, B2-A7 párharcok, amikből kivétel nélkül az elől álló csapatok jutottak tovább, az eredeti kiírásnak megfelelő módon kialakítva a legjobb nyolc csapat mezőnyét.

3.3.2. A csoportok összefésülése

A szurkolók hagyományos „játéka”, hogy két párhuzamosan futó csoportot és csapataikat összehasonlítsák, melyik az erősebb, illetve mely csapatok hogyan viszonyulnak egymáshoz. Ebben az alfejezetben ezt matematikailag alátámasztott módon fogjuk megcsinálni, megmutatva, hogy lehetséges megbízható módon összefésülni a két csoport nyolc-nyolc csapatát. Közvetlenül a csoportkör után az EHF szakértői csapata is összefésülte a csapatokat, elemezve és sorrendbe állítva a szerintük legjobb nyolc csapatot, amivel egyúttal indirekt módon megadták, hogy a nyolcaddöntő párosításai esetében mely csapatokat tartanak továbbjutóknak [5]. Ezt fogjuk EHF ranglistának nevezni és előrejelzésnek tekinteni.

Ahhoz, hogy a két csoport összefésülését a Thurstone és az AHP-LLSM módszerek meg tudják valósítani, szükség volt a nyolcaddöntő esetén két mérkőzésre, melyekkel összekötjük a két csoportot. E nélkül ugyanis nincsen semmilyen információnk a két csoport csapatainak egymáshoz viszonyított erősségéről, mivel korábban nem volt mérkőzés a két csoport között. A nyolcaddöntő első köréből a Győr - Bietigheim és a Rostov-Don - Podravka mérkőzés eredményét is figyelembe véve kiszámítottuk a csapatok sorrendjét és az erősségeiket, és az alábbi 6. táblázatban adjuk meg azokat:

6. táblázat Az EHF ranglista és a 3_N modellel számított várható erősségek 16 csapatra

	EHF ranglista	3_N	
Határ	-	0,2049	
Helyezés	Csapatnév	Csapatnév	m_i
1	Győri Audi ETO KC	Győri Audi ETO KC	2,68924
2	CSKA	CSKA	2,6539
3	Rostov-Don	Rostov-Don	2,30083
4	Metz Handball	Metz Handball	2,10873
5	FTC-Rail Cargo Hungaria	Brest Bretagne Handball	2,03352
6	Brest Bretagne Handball	Vipers Kristiansand	1,87772

Alkalmazás

7	CSM Bucuresti	FTC-Rail Cargo Hungaria	1,7426
8	Odense Håndbold	CSM Bucuresti	1,6879
9	-	Odense Håndbold	1,52054
10	-	Buducnost	1,34445
11	-	Team Esbjerg	1,33073
12	-	SCM Ramnicu Valcea	1,12667
13	-	RK Krim Mercator	1,01071
14	-	BV Borussia 09 Dortmund	1,00993
15	-	SG BBM Bietigheim	0,39024
16	-	HC Podravka Vegeta	0

Az alábbi, 7. táblázatban szerepelnek a súlyok, illetve a sorrendek az egyes módszerek szerint, a résztvevő 16 csapat esetén:

7. táblázat Az EHF ranglista és az egyes módszerek által adott sorrendek 16 csapatra

	EHF ranglista	3_N		sorösszeg		AHP-LLSM	
Helyezés		Csapatnév	Súly	Csapatnév	Súly	Csapatnév	Súly
1	Györi Audi ETO KC	Györi Audi ETO KC	0,15382	Györi Audi ETO KC	0,10947	Györi Audi ETO KC	0,10956
2	CSKA	CSKA	0,14848	CSKA	0,10376	CSKA	0,10581
3	Rostov- Don	Rostov-Don	0,10431	Rostov-Don	0,10203	Rostov-Don	0,10222
4	Metz Handball	Metz Handball	0,08608	Metz Handball	0,08745	Brest Bretagne Handball	0,08043
5	FTC-Rail Cargo Hungaria	Brest Bretagne Handball	0,07984	Brest Bretagne Handball	0,08259	Metz Handball	0,07504

Alkalmazás

6	Brest Bretagne Handball	Vipers Kristiansand	0,06832	Vipers Kristiansand	0,08038	Vipers Kristiansand	0,07503
7	CSM Bucuresti	FTC-Rail Cargo Hungaria	0,05969	FTC-Rail Cargo Hungaria	0,07218	FTC-Rail Cargo Hungaria	0,06539
8	Odense Håndbold	CSM Bucuresti	0,05651	CSM Bucuresti	0,06842	CSM Bucuresti	0,06108
9	-	Odense Håndbold	0,0478	Odense Håndbold	0,05865	Odense Håndbold	0,05326
10	-	Buducnost	0,04008	Buducnost	0,05414	Buducnost	0,04973
11	-	Team Esbjerg	0,03954	Team Esbjerg	0,04858	Team Esbjerg	0,04969
12	-	SCM Ramnicu Valcea	0,03224	SCM Ramnicu Valcea	0,0379	SCM Ramnicu Valcea	0,04385
13	-	RK Krim Mercator	0,02871	BV Borussia 09 Dortmund	0,03684	BV Borussia 09 Dortmund	0,04002
14	-	BV Borussia 09 Dortmund	0,02869	RK Krim Mercator	0,03401	RK Krim Mercator	0,03525
15	-	SG BBM Bietigheim	0,01544	SG BBM Bietigheim	0,01458	SG BBM Bietigheim	0,02773
16	-	HC Podravka Vegeta	0,01045	HC Podravka Vegeta	0,00902	HC Podravka Vegeta	0,0259

Látható, hogy mindhárom módszer szinte azonos eredményt adott, egy-két helyen van csak helycsere a rangsorokban. Az első három helyezést illetőleg megegyezik mind a három módszer és az EHF ranglista is. A 6-12. és a 15-16. helyezést tekintve is közmegegyezés van a három módszer esetén. A legnagyobb eltérés a módszerek eredményei és az EHF ranglista helyezései között van. Az is észrevehető, hogy míg a Thurstone módszer a Györi Audi ETO KC fölényét jelentősnek tartja, addig a többi modellben ezt a súlyok nem támasztják alá. Feltűnő, hogy az EHF ranglistán a Vipers Kristiansand nem fért be a legjobb nyolc közé,

míg a másik három szerint a 6. a sorban, míg épp ellenkezőleg, az FTC-Rail Cargo Hungaria az EHF szerint 5., míg a három módszer szerint csak 7..

A fenti eredmények alapján felmerül a kérdés, hogy vajon érdemes-e használni az egyszerű sorösszeg vagy az AHP-LLSM módszer helyett a bonyolultabb Thurstone módszert az összefésülésre. A válaszunk igen, és ezt az alábbi példával indokoljuk.

Legyen két fiktív csoportunk, A és B (mint a Bajnokok Ligájában), de most legyen jelentősen erősebb az A csoport a B-nél. Az A csoport és B csoport csapatainak számai és erősségei az M10. táblázatban találhatóak. Ezen erősségek alapján körmérkőzéses rendszerben az A csoportban és a B csoportban is 100-100 mérkőzést véletlenszerűen lejátszattam, majd párba rendeztem a csapatokat úgy, hogy az A csapat első helyezettje a B csapat utolsójával játszik, a második az utolsó előttivel és így tovább, éppen úgy, ahogy ez a Bajnokok Ligájában is történt. Ez esetben is 100 mérkőzést játszottam az erősségeknek megfelelően generálva a mérkőzések eredményeit a párokba állított két-két csapat között. Az egyes módszerek által felállított rangsorok az M11. táblázatban találhatóak.

Az erősségek alapján a sorrend a következő lenne: az élen az A csoport első négy csapata helyezkedne el, majd felváltva a B és az A csoport csapatai jönnének és végül a B csoport utolsó négy csapata zárja a rangsort. Jól látható, hogy míg a Thurstone módszer szépen visszaadta ezt a várható sorrendet, addig a másik két módszer erre nem volt képes. Az AHP-LLSM módszer majdnem jó eredményt adott összefűzésre, csak egy-két helyen hibázott a teszt során a rangsor megállapításában. Vagyis több csoport összefésülése esetén a Thurstone módszert érdemes és kell alkalmazni.

Az, hogy a Bajnokok Ligája esetében mégis jól működött az AHP-LLSM és a sorösszeg módszer is, az annak köszönhető, hogy az EHF-nek sikerült nagyjából ugyanolyan erősségű A és B csoportba beosztani a 16 csapatot. Ez utóbbi abból is látszik, hogy a Thurstone módszer szerinti 16-os rangsorban az A csoportbeli csapatok helyezési sorszámainak összege majdnem megegyezik a B csoportéval, 69-67.

A Thurstone kiértékelési módszer előnye tehát, hogy az egymástól független csoportban játszó csapatokat már nagyon kevés egymás ellen játszott mérkőzés után megbízhatóan össze tudjuk hasonlítani és rangsorolni, ezáltal jogos szurkolói igényeket tudunk kielégíteni. A módszer alkalmas arra is, hogy utólag ellenőrizni tudjuk, hogy a csoportbeosztás kiegyensúlyozott volt-e.

3.3.3. A továbbjutó csapatok és a négyes döntő előrejelzése

Ebben az alfejezetben megmutatjuk, hogy a Thurstone módszer még legegyszerűbb változatában is nagyon hatékonyan képes a jövőbeli mérkőzések eredményét előrejelezni, akár számszerűsítve is a mérkőzések kimeneteleinek valószínűségét. Mondanunk sem kell, hogy ez értékes többletinformációt jelent a szakértőknek, a sportfogadásokban résztvevőknek vagy általában a szurkolóknak.

Első lépésként azt mutatjuk be, hogy milyen sikerrel volt előre jelezhető, hogy mely csapatok kerülnek a legjobb nyolc közé. A 8. táblázatban látható az első nyolc közé ténylegesen bejutott csapatok, illetve az egyes módszerek által a 7. táblázat szerint előrejelzetten az első nyolcba jutó csapatok listája. A nyolcaddöntő sorsolása az M12. táblázatban látható, minden egyes páros esetén a 7. táblázatban található ranglista szerint előbb álló csapat továbbjutását jeleztük előre:

8. táblázat A legjobb nyolc közé jutott és az előrejelzett csapatok listája

8 közé jutott	EHF ranglista	3_N	sorösszeg	AHP-LLSM
Györi Audi ETO KC	Györi Audi ETO KC	Györi Audi ETO KC	Györi Audi ETO KC	Györi Audi ETO KC
CSKA	CSKA	CSKA	CSKA	CSKA
Rostov-Don	Rostov-Don	Rostov-Don	Rostov-Don	Rostov-Don
Metz Handball	Metz Handball	Metz Handball	Metz Handball	Brest Bretagne Handball
Vipers Kristiansand	FTC-Rail Cargo Hungaria	Brest Bretagne Handball	Brest Bretagne Handball	Metz Handball
Brest Bretagne Handball	Brest Bretagne Handball	Vipers Kristiansand	Vipers Kristiansand	Vipers Kristiansand

Alkalmazás

CSM Bucuresti	CSM Bucuresti	FTC-Rail Cargo Hungaria	FTC-Rail Cargo Hungaria	FTC-Rail Cargo Hungaria
Buducnost	Odense Håndbold	CSM Bucuresti	CSM Bucuresti	CSM Bucuresti

Amint látható, a három páros összehasonlítási módszer mindegyike egyet tévedett, míg a hivatalos EHF előrejelzés két helyen is hibázott. A Buducnost győzelmét az FTC-Rail Cargo Hungaria ellen semelyik modell sem volt képes előre jelezni, továbbá a hivatalos előrejelzés a Vipers Kristiansand győzelmét sem látta előre.

A legjobb négy közé kerülő csapatok előrejelzéséhez már a nyolcaddöntő mindkét fordulója mérkőzéseinek eredményét felhasználtuk. Az így kapott súlyok az M13. táblázatban találhatók meg. Az M14. táblázatban látható a negyeddöntő sorsolása, amelyek alapján megadtuk a legjobb négy közé kerülő csapatot (mindig az M13. táblázatban előrébb álló csapatot jeleztük továbbjutónak).

A legjobb négy közé jutó csapatok listája, az egyes módszerek esetén az első négy közé várt csapatok listája, illetve az EHF ranglistája alapján a négy közé jutó csapatok listája a 9. táblázatban látható:

9. táblázat A legjobb négy közé (Final4) jutott és előrejelzett csapatok listája

4 közé jutott	EHF ranglista	3_N	sorösszeg	AHP-LLSM
Györi Audi ETO KC	Györi Audi ETO KC	Györi Audi ETO KC	Györi Audi ETO KC	Györi Audi ETO KC
CSKA	CSKA	CSKA	Rostov-Don	Rostov-Don
Brest Bretagne Handball	Rostov-Don	Rostov-Don	CSKA	CSKA
Vipers Kristiansand	Metz Handball	Brest Bretagne Handball	Brest Bretagne Handball	Brest Bretagne Handball

Amint kiolvasható a táblázatból, a Thurstone módszer, a sorösszeg és az AHP-LLSM is 3 bejutót jól jelzett előre, míg az EHF ranglistája csak kettőt tudott jól megjósolni. A Vipers Kristiansand csapata azonban meglepetést okozott:

bejutott a legjobb négy közé és ezt egyik módszer sem tudta megjósolni. Azonban az erősségét a három módszer jobban megbecsülte, mint a szakértők. Mindhárom alkalmazott módszer szerint ez a csapat a hatodik helyre került, míg az EHF ranglistán az első nyolc közé sem került be.

A Final4 május végén játszódó párharcainak az eredményeit is előre jelezhetjük. Ehhez azonban már a negyedöntő eredményeit is figyelembe vettem. Az így kapott rangsorokat és súlyokat az M15. táblázat tartalmazza az egyes módszerek szerint mind a 16 csapat eredményeit figyelembe véve. A rangsorban előrébb állót tekintettem helyezésben előbbre állónak. Ezek alapján az egyes módszerek segítségével, illetve az EHF ranglistája alapján jósolt helyezések, figyelembe véve az elődöntő sorsolását is (M16. táblázat) az alábbiak:

10. táblázat Az első négy csapat jósolt sorrendje az összes meccs eredményének figyelembe vételével és az EHF ranglista szerint

	EHF ranglista	3_N	sorösszeg	AHP-LLSM
Helyezés	Csapatnév	Csapatnév	Csapatnév	Csapatnév
1	Györi Audi ETO KC	Györi Audi ETO KC	Györi Audi ETO KC	Györi Audi ETO KC
2	CSKA	CSKA	CSKA	CSKA
3	Brest Bretagne Handball	Brest Bretagne Handball	Brest Bretagne Handball	Brest Bretagne Handball
4	Vipers Kristiansand	Vipers Kristiansand	Vipers Kristiansand	Vipers Kristiansand

Kiszámoltam akkor is a csapatok rangsorát, amikor csak a legjobb nyolc csapat egymás között lejátszott mérkőzéseinek eredményeit vesszük figyelembe, ennek a súlytáblázata a mellékletben található M17. táblázat. Ekkor ugyan kevesebb mérkőzést vehetünk figyelembe (igaz fele annyi csapat súlyait kell meghatározni), de ezek az igazán erős csapatok és a mérkőzések egy jó része időben közelebb zajlott a Final4-hoz, így várhatóan nagyobb a jósló erejük. Emiatt érdekes lehet az így kapott sorrend is.

Sajnos csak az első négy csapat egymás elleni eredményei alapján sem a Thurstone, sem a sorösszeg, sem az AHP-LLSM nem tudott volna sorrendet adni,

mert a Vipers Kristiansand a másik három csapat egyikével sem lenne ebben az esetben összehasonlítva, emiatt tovább szűkíteni a kört nem lehetett.

A legjobb nyolc csapat eredményei alapján az előre jelzett helyezések az elődöntő sorsolását figyelembe véve a következők:

11. táblázat Az első négy csapat jóslott sorrendje a legerősebb nyolc csapat eredményei figyelembevételével és az EHF ranglista szerint

	EHF ranglista	3_N	sorösszeg	AHP-LLSM
Helyezés	Csapatnév	Csapatnév	Csapatnév	Csapatnév
1	Györi Audi ETO KC	Vipers Kristiansand	Györi Audi ETO KC	Vipers Kristiansand
2	CSKA	Györi Audi ETO KC	Vipers Kristiansand	Györi Audi ETO KC
3	Brest Bretagne Handball	CSKA	CSKA	CSKA
4	Vipers Kristiansand	Brest Bretagne Handball	Brest Bretagne Handball	Brest Bretagne Handball

Látható, hogy míg az összes lejátszott meccset alapul véve minden módszer azonos sorrendet adott meg, addig, ha csak az első nyolc közé jutó csapatok közti eredményeket vesszük figyelembe, akkor csak az AHP-LLSM és Thurstone módszer adott meg egyforma sorrendet és a sorrendek erősen különböznek a 10. táblázatban szereplő sorrendektől. Például a Vipers Kristiansand itt két módszer szerint is megnyeri a Bajnokok Ligáját, míg a Győr csak a sorösszeg módszer szerint (és az EHF ranglista szerint).

Megjegyezzük, hogy az EHF ranglistát a négy közé jutáshoz, illetve a végső sorrend előrejelzéséhez csak a teljesség kedvéért vettük számításba. Nyilvánvaló, hogy mivel ez a ranglista csak a csoportkör eredményei alapján született, szükségképpen erős hátrányban van, így jóval korlátozottabb érvényességgel bír a többi módszerhez képest.

A végső sorrend előrejelzésénél természetesen többet is tudunk mondani, az egyes párharcok kimenetelének a valószínűségeit is meg tudjuk adni. Ezt nyilván ugyanígy korábban is meg tudtuk tenni, csak nem láttuk olyan érdekesnek, hogy be is mutassuk. Most azonban, mivel jövőbeli meccsekről van szó, különösen is érdekes lehet. Tehát nézzük meg, hogy milyen esélye van annak, hogy az elődöntőből az egyik vagy a másik csapat jut tovább. Számolhatunk az M15-ös táblázatot megadó várható értékek szerint (1. verzió), számolhatunk az M17-es táblázatot eredményező várható értékek szerint (2. verzió), de érdekes lehet számolni e kettő várható értékeinek az 1/2-1/2 súllyal tekintett kombinációja szerint is (3. verzió). A 12. táblázat mutatja az elől álló csapat továbbjutási esélyét az elődöntőből a három verzió szerint. A 10. és 11. táblázatnak megfelelően a győri csapat minden verzió szerint esélyesebb a továbbjutásra, csak az esélyei mások, míg a másik ágon már nem egyértelmű a helyzet. Az 1. verzió szerint a moszkvai csapat, a 2. és a 3. verzió szerint a norvég csapat esélye nagyobb a döntőbe jutásra. Ezért aztán két fináléra adunk előrejelzést, amit a 13. táblázatban lehet látni. Itt is az elől álló csapat győzelmi esélyét adtuk meg. Mindkét lehetséges forgatókönyv verzióinak többségében (összesen hatból ötben) a Győr esélyei a nagyobbak a végső győzelemre, de egy esetben nem, ráadásul ott a Vipers Kristiansand esélye jócskán nagyobbak mondható, mint a győri csapaté.

12. táblázat Az elődöntőből való továbbjutás esélyei verzióként

Az elődöntő meccsei	1. verzió	2. verzió	3. verzió
Győri Audi ETO KC - Brest Bretagne Handball	0.7144	0.8120	0.7660
Vipers Kristiansand – CSKA	0.3311	0.7435	0.5433

13. táblázat Várható döntő kimenetelének esélyei verzióként

Várható döntő	1. verzió	2. verzió	3. verzió
Győri Audi ETO KC – CSKA	0.6628	0.5744	0.6194
Győri Audi ETO KC - Vipers Kristiansand	0.8043	0.3203	0.5774

Alkalmazás

Bízunk benne, hogy a járványügyi helyzet májusra való javulása miatt már a helyszínen is szurkolni tudó hazai nézők segítségével a Győr megszerezheti hatodik alkalommal is a BL-győztes címet!

4. Összefoglalás

Dolgozatomban páros összehasonlítási módszereket mutattam be, ezeken belül a Thurstone módszerre fókuszálva.

Megkonstruáltam egy általános módszert, amely külön számszerűsíti az előnyt és a hátrányt minden egyes objektum esetén. Dolgozatomban három különféle, három opciót megengedő Thurstone modellt alkalmaztam, hogy ezek által rangsoroljam az egyes csapatokat és előre jelezzem a mérkőzések eredményeit, és a továbbjutásokat.

Dolgozatom második felének első részében bemutattam az egyes modellek alkalmazhatóságát az NBI-es futballbajnokságon keresztül. Prezentáltam a Thurstone modellben az előny kezelésének lehetőségét és pontosságát, a jövőbeli mérkőzésekre adott valószínűségek számolásának módját, továbbá bemutattam, hogy a modell nagyon pontosan képes visszaadni a csapatok közti finom különbségeket, esetleges kirívó eredményeket. Továbbá a modell nagy pontossággal adja meg a győzelem, a vereség és a döntetlen valószínűségeit akár egyes csapatokra bontva, kevés adat esetén is.

Kézilabda esetén rámutattam a Thurstone modell segítségével, hogy a pandémia miatt nem lejátszott meccsek nélkül is meg lehetett határozni a csapatok erősségét, amik sokszor valósabbnak minősültek, mint az odaítélt pontok alapján kialakult hivatalos sorrend szerintinek megfeleltethetők.

Összehasonlítottam a Thurstone módszert a többi módszerrel azon eshetőség esetén, amikor össze akarunk fésülni két csoportot és meg akarjuk adni a teljes rangsort és erősségeket, és úgy találtam, hogy a Thurstone módszer ezen a területen is kiemelkedik a többi általam bemutatott páros összehasonlítási módszer közül, mert képes kezelni a csoportok közti erősségből származó különbségeket, és ezt két eltérő erősségű csoport összefésülésénél is jól meg tudja adni.

Sikerült azt is bemutatnom, hogy a Thurstone módszer valóban alkalmas arra, hogy a kieséses szakaszban előrejelezzük a párharcok kimeneteleit nagy pontossággal.

Összefoglalás

Előrejelzést adtam a női Bajnokok Ligája első négy csapatának sorrendjére a Final4 bizonyos mérkőzéseinek esélyeit számszerűsítve a dolgozatomban.

Továbbá sikerült elérnem, hogy a program pár másodperc alatt képes legyen az összes módszert lefuttatni és megadni, az egyes módszerek szerint számolt erősségeket, súlyokat, illetve bizonyos Thurstone módszerek esetében releváns előnyhasználati mértéket.

Sikerült automatizálnom a programot, így a régebbi 20-30 gombnyomásos, fél órát is igénybe vevő generálások két gombnyomással, kevesebb mint tíz másodperc alatt befejeződnek.

Úgy gondolom, hogy bizonyítani tudtam a dolgozatomban, hogy a Thurstone modellből származtatott módszerek még a legegyszerűbb verzióikban is széleskörűen jól alkalmazható páros összehasonlítási módszerek, amelyek képesek az előnyt kezelni, jövőbeli valószínűségekkel számolni, így igen pontos előrejelzést tudnak adni a mérkőzések kimenetelére. Arra is rávilágítottam, hogy a modell alkalmas csoportok csapatainak összefésülésére és az egymáshoz viszonyított erősségek meghatározására. Ezen tulajdonságokban a Thurstone módszerek egyértelműen jobbnak bizonyultak a dolgozatomban használt, általánosan elfogadott többi módszernél.

A módszeremet szeretném továbbfejleszteni, hogy a jövőben alkalmas legyen arra, hogy a segítségével elvégzett statisztikai elemzésekből kapott előrejelzésekre alapozva fogadásokat lehessen nyereségessé tenni.

5. Jegyzékek

5.1. Irodalomjegyzék

- [1] Bozóki, S., Fülöp, J., Rónyai, L. (2010). On optimal completion of incomplete pairwise comparison matrices. *Mathematical and Computer Modelling*, 52(1), 318-333.
- [2] Fahrmeir, L., Heumann, C., Künstler, R., Pigeot, I., Tutz, G. (2004). *Statistik*, Springer
- [3] Hartung, F. (2004). *Bevezetés a numerikus analízisbe*, Veszprémi Egyetemi Kiadó, Veszprém
- [4] <https://allsportsapi.com/> (letöltés dátuma 2021. április 15.) All Sports API
- [5] <https://ehfcl.eurohandball.com/women/2020-21/news/en/gyor-top-power-ranking-going-into-play-offs/> (letöltés dátuma 2021. április 15.) Győr top power ranking going into play-offs
- [6] J. González-Díaz, Ruud Hendrickx, Edwin Lohmann, (2014). Paired comparisons analysis: an axiomatic approach to ranking methods, *Soc. Choice Welfare*. 42(1),139-169
- [7] Mihálykóné Orbán, É., Mihálykó, Cs., Kajtár, P., (2019). Általánosított Thurstone-módszer alkalmazásokkal, *Alk. Mat. Lapok*, 36(2), 255-262.
- [8] Mihálykóné Orbán, É., Mihálykó, Cs., (2021). Thurstone módszer általánosítása előnyök figyelembe vételére, *Sigma* 52(1), 47-60
- [9] Mihálykó, Cs., Mihálykóné Orbán, É., Gyarmati L., Thurstone módszer általánosításai, *Alk. Mat. Lapok*, (bírálat alatt)
- [10] Éva Orbán-Mihálykó, Csaba Mihálykó, László Gyarmati, The Evaluation of Sports Tournaments' Results Applying the Generalized Thurstone Method, *Operations Research Letters*, (bírálat alatt)
- [11] Orbán-Mihálykó, É., Mihálykó, C., Koltay, L. (2019). A generalization of the Thurstone method for multiple choice and incomplete paired comparisons. *Central European Journal of Operations Research*, 27(1), 133-159.
- [12] Orbán-Mihálykó É. , Mihálykó Cs., Gyarmati L., (2021). The evaluation of sports tournaments' results by Thurstone methods, in *Book of Abstracts of 91st online meeting of the EURO Working Group in Multi Criteria Decision Aiding (EWG-MCDA)*, pp. 32-33
- [13] Prékopa A. (1972). *Valószínűségelmélet műszaki alkalmazásokkal*. Műszaki Kiadó, Budapest
- [14] Saaty, T. L., (1990). How to make a decision: the analytic hierarchy process, *European Journal of Operational Research*, 48(1), 9-26.
- [15] Thurstone, L. L., (1927). A law of comparative judgement. *Psychol. Rev.* 34(4), 273-286.

5.2. Ábrajegyzék

1. ábra Általános modell intervallumai és döntései.....	17
4. ábra Nelder-Mead optimalizáló.....	28
5. ábra LogLikeliHood függvény érdemi része.....	30

5.3. Táblázatjegyzék

1. táblázat A Thurstone modellek várható értékei és előny-mértékei	35
2. táblázat Rangkorrelációs értékek egyes módszerek esetén	37
3. táblázat A 3_B és 3_C módszerek által adott valószínűségek a győzelem, döntetlen és vereség kialakulására és a relatív gyakoriságok az összes csapatot figyelembevéve	37
4. táblázat Az A csoport csapatainak erősségei és pontjai	41
5. táblázat A B csoport csapatainak erősségei és pontjai	42
6. táblázat Az EHF ranglista és a 3_N modellel számított várható erősségek 16 csapatra.....	44
7. táblázat Az EHF ranglista és az egyes módszerek által adott sorrendek 16 csapatra.....	45
8. táblázat A legjobb nyolc közé jutott és az előrejelzett csapatok listája	48
9. táblázat A legjobb négy közé (Final4) jutott és előrejelzett csapatok listája	49
10. táblázat Az első négy csapat jóslott sorrendje az összes meccs eredményének figyelembe vételével és az EHF ranglista szerint	50
11. táblázat Az első négy csapat jóslott sorrendje a legerősebb nyolc csapat eredményei figyelembevételével és az EHF ranglista szerint.....	51
12. táblázat Az elődöntőből való továbbjutás esélyei verzióként	52
13. táblázat Várható döntő kimenetelének esélyei verzióként	52
M1. táblázat NBI csapatok súlyai egyes modellek esetén	60
M2. táblázat NBI csapatok egyes eredményeinek relatív gyakorisága, modellek által adott valószínűségek	61

M3. táblázat Jövőbeli mérkőzések eredményének valószínűsége előny figyelembevétele nélkül	64
M4. táblázat Jövőbeli mérkőzések eredményének valószínűsége idegenben	65
M5. táblázat Jövőbeli mérkőzések eredményének valószínűsége odahaza	66
M6. táblázat Az A csoport csapataira kapott súlyok az egyes módszerek esetén, csak a lejátszott meccseket figyelembe véve	67
M7. táblázat Az A csoport csapataira kapott súlyok az egyes módszerek esetén, a nem lejátszott meccseket is figyelembe véve.....	67
M8. táblázat A B csoport csapataira kapott súlyok az egyes módszerek esetén, csak a lejátszott meccseket figyelembe véve	68
M9. táblázat A B csoport csapataira kapott súlyok az egyes módszerek esetén, a nem lejátszott meccseket is figyelembe véve.....	68
M10. táblázat Generált erősségek	69
M11. táblázat Két csoport összefésülésekor kapott sorrendek és súlyok az egyes módszerek esetén	70
M12. táblázat A nyolcaddöntő párosításai	71
M13. táblázat Az egyes modellek esetén a csapatok súlyai a nyolcaddöntőig bezárólag	72
M14. táblázat A negyeddöntő párosításai	73
M15. táblázat A csapatok súlyai az egyes módszerek esetén a negyeddöntő mindkét fordulója meccseinek figyelembevételével.....	74
M16. táblázat Az elődöntő párosításai	75
M17. táblázat Az egyes modellek esetén a csapatok súlyai a legjobb nyolc csapat meccseinek figyelembevételével.....	75

6. Mellékletek

6.1. Táblázat mellékletek

M1. táblázat NBI csapatok súlyai egyes modellek esetén

Helyezés	3 N		3 B		3 C		sorösszeg		AHP-LLSM	
	Csapatnév	Súly	Csapatnév	Súly	Csapatnév	Súly	Csapatnév	Súly	Csapatnév	Súly
1	Ferencváros	0,18249	Ferencváros	0,18465	Ferencváros	0,19535	Ferencváros	0,13384	Ferencváros	0,14655
2	MOL Fehérvár	0,12293	MOL Fehérvár	0,12319	MOL Fehérvár	0,11935	MOL Fehérvár	0,11364	MOL Fehérvár	0,11481
3	Puskás Akadémia	0,09944	Puskás Akadémia	0,10044	Puskás Akadémia	0,09809	Puskás Akadémia	0,10101	Puskás Akadémia	0,09855
4	Mezőkövesd	0,08769	Mezőkövesd	0,08698	Mezőkövesd	0,08955	Mezőkövesd	0,09091	Mezőkövesd	0,08725
5	Honvéd	0,07559	Honvéd	0,0754	Zalaegerszeg	0,074	Honvéd	0,08081	Honvéd	0,07723
6	Zalaegerszeg	0,07471	Zalaegerszeg	0,07427	Honvéd	0,07315	Zalaegerszeg	0,08081	Zalaegerszeg	0,0772
7	Újpest	0,07071	Újpest	0,07013	Újpest	0,07019	Újpest	0,07828	Újpest	0,07489
8	Paks	0,06852	Paks	0,06867	Paks	0,0678	Paks	0,07576	Paks	0,07265
9	Kisvárd	0,06727	Kisvárd	0,06688	Kisvárd	0,06755	Kisvárd	0,07576	Kisvárd	0,07263
10	Diósgyőr	0,06436	Diósgyőr	0,06459	Diósgyőr	0,06512	Diósgyőr	0,07323	Diósgyőr	0,07045
11	Debrecen	0,06104	Debrecen	0,05999	Debrecen	0,05586	Debrecen	0,07071	Debrecen	0,06834
12	Kaposvár	0,02525	Kaposvár	0,0248	Kaposvár	0,02399	Kaposvár	0,02525	Kaposvár	0,03944

M2. táblázat NBI csapatok egyes eredményeinek relatív gyakorisága, modellek által adott valószínűségek

Típus	Csapatnév	Relatív gyakoriság	3_B	3_C
Otthoni Győzelem	Ferencváros	0,82353	0,76681 (0,05671)	0,84658 (0,02305)
	MOL Fehérvár	0,58824	0,62133 (0,0331)	0,55148 (0,03675)
	Puskás Akadémia	0,4375	0,54109 (0,10359)	0,45482 (0,01732)
	Mezőkövesd	0,41176	0,48379 (0,07202)	0,4057 (0,00607)
	Honvéd	0,35294	0,42731 (0,07436)	0,35166 (0,00128)
	Zalaegerszeg	0,3125	0,42144 (0,10894)	0,45301 (0,14051)
	Újpest	0,35294	0,39918 (0,04624)	0,35563 (0,00269)
	Paks	0,4375	0,39113 (0,04637)	0,45238 (0,01488)
	Kisvárd	0,4375	0,38102 (0,05648)	0,45562 (0,01812)
	Diósgyőr	0,4375	0,36781 (0,06969)	0,39223 (0,04527)
	Debrecen	0,41176	0,34036 (0,0714)	0,38442 (0,02734)
	Kaposvár	0,1875	0,08904 (0,09846)	0,11472 (0,07278)
Otthoni Döntetlen	Ferencváros	0,17647	0,14818 (0,02829)	0,1052 (0,07127)
	MOL Fehérvár	0,17647	0,20246 (0,02599)	0,22427 (0,0478)
	Puskás Akadémia	0,25	0,22302 (0,02698)	0,23905 (0,01095)
	Mezőkövesd	0,17647	0,23322 (0,05675)	0,24179 (0,06532)
	Honvéd	0,23529	0,23927 (0,00398)	0,23876 (0,00346)
	Zalaegerszeg	0,4375	0,23966 (0,19784)	0,23723 (0,20027)
	Újpest	0,23529	0,24068 (0,00539)	0,23861 (0,00332)
	Paks	0,25	0,24088 (0,00912)	0,23852 (0,01148)
	Kisvárd	0,25	0,24099 (0,00901)	0,23838 (0,01162)
	Diósgyőr	0,125	0,2409 (0,1159)	0,24034 (0,11534)
	Debrecen	0,29412	0,23984 (0,05428)	0,24325 (0,05087)
	Kaposvár	0,0625	0,15434 (0,09184)	0,17618 (0,11368)

Otthoni Vereség	Ferencváros	0	0,08501 (0,08501)	0,04822 (0,04822)
	MOL Fehérvár	0,23529	0,17621 (0,05909)	0,22425 (0,01105)
	Puskás Akadémia	0,3125	0,23589 (0,07661)	0,30613 (0,00637)
	Mezőkövesd	0,41176	0,28299 (0,12877)	0,35251 (0,05926)
	Honvéd	0,41176	0,33342 (0,07834)	0,40958 (0,00218)
	Zalaegerszeg	0,25	0,3389 (0,0889)	0,30976 (0,05976)
	Újpest	0,41176	0,36013 (0,05163)	0,40576 (0,006)
	Paks	0,3125	0,36799 (0,05549)	0,3091 (0,0034)
	Kisvárd	0,3125	0,37799 (0,06549)	0,306 (0,0065)
	Diósgyőr	0,4375	0,39129 (0,04621)	0,36743 (0,07007)
	Debrecen	0,29412	0,4198 (0,12568)	0,37233 (0,07821)
	Kaposvár	0,75	0,75661 (0,00661)	0,7091 (0,0409)
Idegen Győzelem	Ferencváros	0,5625	0,67544 (0,11294)	0,58745 (0,02495)
	MOL Fehérvár	0,5	0,5175 (0,0175)	0,56152 (0,06152)
	Puskás Akadémia	0,41176	0,43662 (0,02485)	0,50588 (0,09412)
	Mezőkövesd	0,4375	0,38126 (0,05624)	0,48539 (0,04789)
	Honvéd	0,375	0,32858 (0,04642)	0,38665 (0,01165)
	Zalaegerszeg	0,35294	0,32321 (0,02973)	0,29423 (0,05872)
	Újpest	0,375	0,30302 (0,07198)	0,35151 (0,02349)
	Paks	0,23529	0,29579 (0,0605)	0,23706 (0,00176)
	Kisvárd	0,29412	0,28676 (0,00736)	0,23215 (0,06197)
	Diósgyőr	0,29412	0,27505 (0,01907)	0,26319 (0,03092)
	Debrecen	0,25	0,25103 (0,00103)	0,1785 (0,0715)
	Kaposvár	0,05882	0,05197 (0,00685)	0,03663 (0,02219)
Idegen Döntetlen	Ferencváros	0,25	0,18819 (0,06181)	0,22732 (0,02268)
	MOL Fehérvár	0,375	0,22945 (0,14555)	0,22054 (0,15446)
	Puskás Akadémia	0,47059	0,23931 (0,23128)	0,23317 (0,23742)

	Mezőkövesd	0,3125	0,24095 (0,07155)	0,23728 (0,07522)
	Honvéd	0,25	0,23812 (0,01188)	0,24829 (0,00171)
	Zalaegerszeg	0,17647	0,23758 (0,0611)	0,243 (0,06653)
	Újpest	0,1875	0,23504 (0,04754)	0,24815 (0,06065)
	Paks	0,23529	0,23395 (0,00135)	0,23122 (0,00408)
	Kisvárd	0,11765	0,23244 (0,11479)	0,22987 (0,11222)
	Diósgyőr	0,17647	0,23025 (0,05378)	0,23726 (0,06079)
	Debrecen	0,0625	0,22486 (0,16236)	0,20922 (0,14672)
	Kaposvár	0,05882	0,11331 (0,05449)	0,09628 (0,03746)
Idegen Vereség	Ferencváros	0,1875	0,13637 (0,05113)	0,18523 (0,00227)
	MOL Fehérvár	0,125	0,25305 (0,12805)	0,21793 (0,09293)
	Puskás Akadémia	0,11765	0,32407 (0,20643)	0,26095 (0,14331)
	Mezőkövesd	0,25	0,37779 (0,12779)	0,27734 (0,02734)
	Honvéd	0,375	0,4333 (0,0583)	0,36506 (0,10553)
	Zalaegerszeg	0,47059	0,43921 (0,03138)	0,46277 (0,08777)
	Újpest	0,4375	0,46193 (0,02443)	0,40033 (0,03717)
	Paks	0,52941	0,47026 (0,05915)	0,53172 (0,00231)
	Kisvárd	0,58824	0,4808 (0,10743)	0,53799 (0,05025)
	Diósgyőr	0,52941	0,4947 (0,03471)	0,49955 (0,02986)
	Debrecen	0,6875	0,52412 (0,16338)	0,61228 (0,07522)
	Kaposvár	0,88235	0,83472 (0,04764)	0,86709 (0,01526)

M3. táblázat Jövőbeli mérkőzések eredményének valószínűsége előny figyelembevétele nélkül

	Ferencváros	MOL Fehérvár	Puskás Akadémia	Mezőkövesd	Zalaegerszeg	Honvéd	Újpest	Paks	Kisvárd	Diósgyőr	Debrecen	Kaposvár
Ferencváros		55,67 24,37	63,26 21,8	66,63 20,44	73,26 17,42	73,64 17,23	74,97 16,56	76,06 15,99	76,17 15,93	77,29 15,34	81,64 12,9	95,97 3,31
MOL Fehérvár	19,97 24,37		43,88 26,88	47,49 26,31	55,09 24,53	55,55 24,4	57,17 23,91	58,53 23,48	58,67 23,43	60,09 22,96	65,88 20,76	89,51 7,96
Puskás Akadémia	14,94 21,8	29,24 26,88		39,78 27,27	47,28 26,35	47,74 26,27	49,38 25,94	50,77 25,64	50,91 25,61	52,37 25,26	58,43 23,51	85,5 10,57
Mezőkövesd	12,92 20,44	26,19 26,31	32,95 27,27		43,67 26,9	44,12 26,84	45,76 26,61	47,14 26,38	47,28 26,35	48,74 26,08	54,85 24,6	83,32 11,91
Zalaegerszeg	9,33 17,42	20,38 24,53	26,37 26,35	29,43 26,9		36,75 27,37	38,31 27,34	39,64 27,28	39,78 27,27	41,2 27,16	47,25 26,36	78,12 14,89
Honvéd	9,14 17,23	20,05 24,4	25,99 26,27	29,03 26,84	35,88 27,37		37,87 27,35	39,2 27,3	39,34 27,29	40,75 27,2	46,79 26,44	77,77 15,08
Újpest	8,48 16,56	18,91 23,91	24,67 25,94	27,63 26,61	34,35 27,34	34,77 27,35		37,62 27,36	37,76 27,36	39,16 27,3	45,15 26,7	76,52 15,75
Paks	7,95 15,99	17,99 23,48	23,59 25,64	26,49 26,38	33,08 27,28	33,5 27,3	35,02 27,36		36,45 27,38	37,83 27,36	43,78 26,89	75,45 16,31
Kisvárd	7,9 15,93	17,9 23,43	23,48 25,61	26,37 26,35	32,95 27,27	33,37 27,29	34,89 27,36	36,18 27,38		37,7 27,36	43,64 26,91	75,33 16,37
Diósgyőr	7,37 15,34	16,95 22,96	22,37 25,26	25,18 26,08	31,63 27,16	32,05 27,2	33,54 27,3	34,81 27,36	34,95 27,36		42,2 27,07	74,16 16,97
Debrecen	5,46 12,9	13,36 20,76	18,06 23,51	20,55 24,6	26,39 26,36	26,77 26,44	28,15 26,7	29,33 26,89	29,45 26,91	30,73 27,07		68,97 19,43
Kaposvár	0,72 3,31	2,53 7,96	3,94 10,57	4,77 11,91	6,99 14,89	7,15 15,08	7,73 15,75	8,24 16,31	8,3 16,37	8,87 16,97	11,6 19,43	

M4. táblázat Jövőbeli mérkőzések eredményének valószínűsége idegenben

	Ferencváros	MOL Fehérvár	Puskás Akadémia	Mezőkövesd	Zalaegerszeg	Honvéd	Újpest	Paks	Kisvárd	Diósgyőr	Debrecen	Kaposvár
Ferencváros		43,84 26,88	54,32 24,75	59,79 23,06	52,14 25,32	65,15 21,06	64,32 21,39	51,41 25,49	50,96 25,6	58,87 23,37	58,44 23,51	86,97 9,63
MOL Fehérvár	12,7 20,29		54,6 24,67	60,06 22,97	52,42 25,25	65,41 20,95	64,58 21,28	51,69 25,43	51,24 25,53	59,14 23,28	58,71 23,42	87,12 9,53
Puskás Akadémia	10,22 18,25	39,1 27,31		55,02 24,55	47,29 26,35	60,55 22,79	59,69 23,09	46,55 26,48	46,11 26,55	54,08 24,81	53,64 24,93	84,22 11,36
Mezőkövesd	9,47 17,56	37,46 27,36	47,77 26,26		45,58 26,64	58,89 23,36	58,02 23,64	44,85 26,74	44,41 26,81	52,37 25,26	51,94 25,37	83,16 12
Zalaegerszeg	2,95 8,8	18,53 23,74	26,38 26,36	31,12 27,11		36,29 27,38	35,45 27,37	24,04 25,77	23,69 25,67	30,29 27,02	29,91 26,97	64,99 21,12
Honvéd	5,76 13,32	28 26,67	37,48 27,36	42,86 27	35,42 27,37		47,58 26,3	34,73 27,35	34,32 27,34	41,93 27,1	41,51 27,14	75,72 16,17
Újpest	4,65 11,71	24,58 25,92	33,58 27,3	38,79 27,32	31,6 27,16	44,3 26,82		30,95 27,09	30,55 27,05	37,88 27,35	37,47 27,36	72,32 17,87
Paks	1,87 6,49	13,81 21,07	20,46 24,57	24,64 25,93	18,94 23,92	29,31 26,89	28,55 26,77		18,15 23,56	23,9 25,73	23,56 25,63	57,59 23,78
Kisvárd	1,79 6,3	13,41 20,79	19,94 24,36	24,06 25,78	18,44 23,7	28,67 26,79	27,92 26,66	17,96 23,46		23,32 25,56	22,99 25,46	56,86 24,01
Diósgyőr	2,42 7,73	16,34 22,62	23,67 25,67	28,18 26,71	22,01 25,14	33,14 27,28	32,33 27,22	21,47 24,95	21,14 24,83		27,02 26,49	61,78 22,35
Debrecen	1,1 4,49	9,71 17,78	15,04 21,86	18,53 23,74	13,8 21,06	22,55 25,32	21,88 25,09	13,39 20,78	13,15 20,61	17,9 23,44		49,28 25,97
Kaposvár	0,11 0,79	1,89 6,55	3,49 9,8	4,71 11,81	3,09 9,06	6,28 14	6,01 13,65	2,96 8,82	2,89 8,67	4,48 11,46	4,38 11,3	

M5. táblázat Jövőbeli mérkőzések eredményének valószínűsége odahaza

	Ferencváros	MOL Fehérvár	Puskás Akadémia	Mezőkövesd	Zalaegerszeg	Honvéd	Újpest	Paks	Kisvárd	Diósgyőr	Debrecen	Kaposvár
Ferencváros		67,01 20,29	71,53 18,25	72,97 17,56	88,25 8,8	80,92 13,32	83,64 11,71	91,64 6,49	91,92 6,3	89,84 7,73	94,41 4,49	99,11 0,79
MOL Fehérvár	29,28 26,88		33,6 27,31	35,18 27,36	57,73 23,74	45,32 26,67	49,5 25,92	65,12 21,07	65,8 20,79	61,04 22,62	72,51 17,78	91,56 6,55
Puskás Akadémia	20,93 24,75	20,73 24,67		25,97 26,26	47,26 26,36	35,16 27,36	39,12 27,3	54,97 24,57	55,7 24,36	50,66 25,67	63,09 21,86	86,7 9,8
Mezőkövesd	17,15 23,06	16,97 22,97	20,43 24,55		41,76 27,11	30,14 27	33,89 27,32	49,43 25,93	50,17 25,78	45,12 26,71	57,73 23,74	83,48 11,81
Zalaegerszeg	22,54 25,32	22,33 25,25	26,36 26,35	27,78 26,64		37,21 27,37	41,24 27,16	57,13 23,92	57,86 23,7	52,85 25,14	65,14 21,06	87,85 9,06
Honvéd	13,79 21,06	13,64 20,95	16,65 22,79	17,75 23,36	36,34 27,38		28,88 26,82	43,81 26,89	44,54 26,79	39,58 27,28	52,13 25,32	79,72 14
Újpest	14,29 21,39	14,13 21,28	17,22 23,09	18,33 23,64	37,18 27,37	26,12 26,3		44,69 26,77	45,42 26,66	40,44 27,22	53,02 25,09	80,34 13,65
Paks	23,1 25,49	22,89 25,43	26,97 26,48	28,4 26,74	50,19 25,77	37,91 27,35	41,96 27,09		58,58 23,46	53,58 24,95	65,82 20,78	88,22 8,82
Kisvárd	23,45 25,6	23,23 25,53	27,34 26,55	28,79 26,81	50,64 25,67	38,34 27,34	42,4 27,05	58,3 23,56		54,03 24,83	66,24 20,61	88,44 8,67
Diósgyőr	17,76 23,37	17,58 23,28	21,11 24,81	22,37 25,26	42,69 27,02	30,97 27,1	34,76 27,35	50,37 25,73	51,11 25,56		58,66 23,44	84,06 11,46
Debrecen	18,05 23,51	17,87 23,42	21,43 24,93	22,7 25,37	43,12 26,97	31,36 27,14	35,17 27,36	50,81 25,63	51,55 25,46	46,49 26,49		84,32 11,3
Kaposvár	3,4 9,63	3,34 9,53	4,42 11,36	4,84 12	13,89 21,12	8,11 16,17	9,8 17,87	18,63 23,78	19,13 24,01	15,87 22,35	24,75 25,97	

M6. táblázat Az A csoport csapataira kapott súlyok az egyes módszerek esetén, csak a lejátszott meccseket figyelembe véve

	3_N		sorösszeg		AHP-LLSM	
Helyezés	Csapatnév	Súly	Csapatnév	Súlys	Csapatnév	Súly
1	Rostov-Don	0,22317	Rostov-Don	0,19731	Rostov-Don	0,21407
2	Metz Handball	0,18684	Metz Handball	0,17255	Vipers Kristiansand	0,15188
3	Vipers Kristiansand	0,14857	Vipers Kristiansand	0,1586	Metz Handball	0,15185
4	FTC-Rail Cargo Hungaria	0,13058	FTC-Rail Cargo Hungaria	0,14242	FTC-Rail Cargo Hungaria	0,13245
5	CSM Bucuresti	0,12377	CSM Bucuresti	0,135	CSM Bucuresti	0,12356
6	Team Esbjerg	0,08764	Team Esbjerg	0,09586	Team Esbjerg	0,10063
7	RK Krim Mercator	0,06416	RK Krim Mercator	0,0671	RK Krim Mercator	0,07134
8	SG BBM Bietigheim	0,03527	SG BBM Bietigheim	0,03115	SG BBM Bietigheim	0,05422

M7. táblázat Az A csoport csapataira kapott súlyok az egyes módszerek esetén, a nem lejátszott meccseket is figyelembe véve

	3_N		Sorösszeg		AHP-LLSM	
Helyezés	Csapatnév	Súly	Csapatnév	Súly	Csapatnév	Súly
1	Rostov-Don	0,20528	Rostov-Don	0,1875	Rostov-Don	0,18789
2	Metz Handball	0,20341	Metz Handball	0,17857	Metz Handball	0,17526
3	CSM Bucuresti	0,14427	CSM Bucuresti	0,15179	CSM Bucuresti	0,14274
4	Vipers Kristiansand	0,1322	FTC-Rail Cargo Hungaria	0,14286	Vipers Kristiansand	0,13328
5	FTC-Rail Cargo Hungaria	0,13	Vipers Kristiansand	0,14286	FTC-Rail Cargo Hungaria	0,13324
6	Team Esbjerg	0,09314	Team Esbjerg	0,10714	Team Esbjerg	0,10125
7	RK Krim Mercator	0,05925	RK Krim Mercator	0,0625	RK Krim Mercator	0,07181
8	SG BBM Bietigheim	0,03246	SG BBM Bietigheim	0,02679	SG BBM Bietigheim	0,05455

M8. táblázat A B csoport csapataira kapott súlyok az egyes módszerek esetén, csak a lejátszott meccseket figyelembe véve

	3_N		sorösszeg		AHP-LLSM	
Helyezés	Csapatnév	Súly	Csapatnév	Súlys	Csapatnév	Súly
1	Györi Audi ETO KC	0,28933	Györi Audi ETO KC	0,22013	Györi Audi ETO KC	0,22161
2	CSKA	0,27778	CSKA	0,21095	CSKA	0,20689
3	Brest Bretagne Handball	0,14742	Brest Bretagne Handball	0,16792	Brest Bretagne Handball	0,15719
4	Odense Håndbold	0,08622	Odense Håndbold	0,11924	Odense Håndbold	0,10414
5	Buducnost	0,0721	Buducnost	0,11006	Buducnost	0,09727
6	SCM Ramnicu Valcea	0,05772	SCM Ramnicu Valcea	0,07704	SCM Ramnicu Valcea	0,08574
7	BV Borussia 09 Dortmund	0,05105	BV Borussia 09 Dortmund	0,0749	BV Borussia 09 Dortmund	0,07821
8	HC Podravka Vegeta	0,0184	HC Podravka Vegeta	0,01975	HC Podravka Vegeta	0,04895

M9. táblázat A B csoport csapataira kapott súlyok az egyes módszerek esetén, a nem lejátszott meccseket is figyelembe véve

	3_N		sorösszeg		AHP-LLSM	
Helyezés	Csapatnév	Súly	Csapatnév	Súly	Csapatnév	Súly
1	Györi Audi ETO KC	0,28095	Györi Audi ETO KC	0,21429	Györi Audi ETO KC	0,22452
2	CSKA	0,27613	CSKA	0,20536	CSKA	0,20968
3	Brest Bretagne Handball	0,12921	Brest Bretagne Handball	0,15179	Brest Bretagne Handball	0,13888
4	Odense Håndbold	0,09091	Odense Håndbold	0,11607	Odense Håndbold	0,10551
5	Buducnost	0,07713	Buducnost	0,10714	Buducnost	0,09854
6	SCM Ramnicu Valcea	0,06071	SCM Ramnicu Valcea	0,08929	SCM Ramnicu Valcea	0,08584
7	BV Borussia 09 Dortmund	0,05627	BV Borussia 09 Dortmund	0,08036	BV Borussia 09 Dortmund	0,08016
8	HC Podravka Vegeta	0,02869	HC Podravka Vegeta	0,03571	HC Podravka Vegeta	0,05687

M10. táblázat Generált erősségek

A csoport		B csoport	
Csapatnév	M _i	Csapatnév	M _i
1	2,75	9	1,75
2	2,5	10	1,5
3	2,25	11	1,25
4	2,0	12	1,0
5	1,75	13	0,75
6	1,5	14	0,5
7	1,25	15	0,25
8	1,0	16	0

M11. táblázat Két csoport összefésülésekor kapott sorrendek és súlyok az egyes módszerek esetén

	3_N		sorösszeg		AHP-LLSM	
Helyezés	Csapatnév	Súly	Csapatnév	Súly	Csapatnév	Súly
1	1	0,18803	1	0,10523	1	0,1259
2	2	0,13475	9	0,0993	2	0,10521
3	3	0,11111	2	0,0932	3	0,09369
4	4	0,08563	10	0,09102	4	0,0792
5	5	0,07112	3	0,08438	9	0,07848
6	9	0,06892	11	0,07641	10	0,07032
7	10	0,05819	4	0,07273	5	0,06944
8	6	0,05189	5	0,06313	11	0,05681
9	11	0,04352	12	0,06258	6	0,05508
10	7	0,04317	13	0,05094	7	0,04817
11	12	0,03353	6	0,04695	12	0,04696
12	8	0,03266	14	0,04172	8	0,04037
13	13	0,02647	7	0,03688	13	0,03972
14	14	0,02166	15	0,02859	14	0,03505
15	15	0,01601	8	0,02477	15	0,02892
16	16	0,01334	16	0,02219	16	0,02669

M12. táblázat A nyolcaddöntő párosításai

Nyolcaddöntő	
Csapat 1	Csapat 2
SG BBM Bietigheim	Györi Audi ETO KC
HC Podravka Vegeta	Rostov-Don
RK Krim Mercator	CSKA
BV Borussia 09 Dortmund	Metz Handball
Team Esbjerg	Brest Bretagne Handball
SCM Ramnicu Valcea	CSM Bucuresti
Vipers Kristiansand	Odense Håndbold
Buducnost	FTC-Rail Cargo Hungaria

M13. táblázat Az egyes modellek esetén a csapatok súlyai a nyolcaddöntőig bezárólag

	3_N		sorösszeg		AHP-LLSM	
Helyezés	Csapatnév	Súly	Csapatnév	Súly	Csapatnév	Súly
1	Györi Audi ETO KC	0,16065	Györi Audi ETO KC	0,11045	Györi Audi ETO KC	0,11313
2	CSKA	0,12412	Rostov-Don	0,10369	Rostov-Don	0,10012
3	Rostov-Don	0,10403	CSKA	0,09862	CSKA	0,0979
4	Brest Bretagne Handball	0,09221	Brest Bretagne Handball	0,08836	Brest Bretagne Handball	0,08835
5	Metz Handball	0,08305	Metz Handball	0,08739	Metz Handball	0,07299
6	Vipers Kristiansand	0,06317	Vipers Kristiansand	0,07768	Vipers Kristiansand	0,07101
7	FTC-Rail Cargo Hungaria	0,05606	FTC-Rail Cargo Hungaria	0,071	FTC-Rail Cargo Hungaria	0,06234
8	Odense Håndbold	0,05298	CSM Bucuresti	0,06762	CSM Bucuresti	0,05788
9	CSM Bucuresti	0,05242	Odense Håndbold	0,05917	Odense Håndbold	0,05691
10	Buducnost	0,04555	Buducnost	0,05523	Buducnost	0,05278
11	SCM Ramnicu Valcea	0,03804	Team Esbjerg	0,04208	SCM Ramnicu Valcea	0,04698
12	Team Esbjerg	0,03502	SCM Ramnicu Valcea	0,04208	Team Esbjerg	0,04576
13	RK Krim Mercator	0,03392	RK Krim Mercator	0,03787	BV Borussia 09 Dortmund	0,04141
14	BV Borussia 09 Dortmund	0,03158	BV Borussia 09 Dortmund	0,03682	RK Krim Mercator	0,03855
15	SG BBM Bietigheim	0,01557	SG BBM Bietigheim	0,01352	SG BBM Bietigheim	0,02715
16	HC Podravka Vegeta	0,01162	HC Podravka Vegeta	0,00842	HC Podravka Vegeta	0,02674

M14. táblázat A negyedöntő párosításai

Negyedöntő	
Csapat 1	Csapat 2
Buducnost	Győri Audi ETO KC
Vipers Kristiansand	Rostov-Don
CSM Bucuresti	CSKA
Brest Bretagne Handball	Metz Handball

M15. táblázat A csapatok súlyai az egyes módszerek esetén a negyeddöntő mindkét fordulója meccseinek figyelembevételével

	3_N		sorösszeg		AHP-LLSM	
Helyezés	Csapatnév	Súly	Csapatnév	Súly	Csapatnév	Súly
1	Györi Audi ETO KC	0,1726	Györi Audi ETO KC	0,11477	Györi Audi ETO KC	0,11292
2	CSKA	0,11339	Rostov-Don	0,09684	CSKA	0,09325
3	Brest Bretagne Handball	0,09797	CSKA	0,09684	Brest Bretagne Handball	0,09117
4	Rostov-Don	0,08782	Brest Bretagne Handball	0,09114	Rostov-Don	0,0892
5	Metz Handball	0,07725	Metz Handball	0,08177	Vipers Kristiansand	0,08103
6	Vipers Kristiansand	0,07326	Vipers Kristiansand	0,08177	Metz Handball	0,07098
7	CSM Bucuresti	0,05803	FTC-Rail Cargo Hungaria	0,07263	FTC-Rail Cargo Hungaria	0,06285
8	FTC-Rail Cargo Hungaria	0,05613	CSM Bucuresti	0,06859	CSM Bucuresti	0,06121
9	Odense Håndbold	0,05399	Odense Håndbold	0,06052	Odense Håndbold	0,05758
10	Buducnost	0,04385	Buducnost	0,05021	Buducnost	0,05268
11	SCM Ramnicu Valcea	0,03854	Team Esbjerg	0,04304	SCM Ramnicu Valcea	0,04721
12	Team Esbjerg	0,03576	SCM Ramnicu Valcea	0,04304	Team Esbjerg	0,04633
13	RK Krim Mercator	0,03318	RK Krim Mercator	0,03873	BV Borussia 09 Dortmund	0,04121
14	BV Borussia 09 Dortmund	0,03152	BV Borussia 09 Dortmund	0,03766	RK Krim Mercator	0,03867
15	SG BBM Bietigheim	0,01537	SG BBM Bietigheim	0,01383	SG BBM Bietigheim	0,02739
16	HC Podravka Vegeta	0,01133	HC Podravka Vegeta	0,00861	HC Podravka Vegeta	0,02633

M16. táblázat Az elődöntő párosításai

Elődöntő	
Csapat 1	Csapat 2
Györi Audi ETO KC	Brest Bretagne Handball
Vipers Kristiansand	CSKA

M17. táblázat Az egyes modellek esetén a csapatok súlyai a legjobb nyolc csapat meccseinek figyelembevételével

	3_N		sorösszeg		AHP-LLSM	
Helye zes	Csapatnév	Súly	Csapatnév	Súly	Csapatnév	Súly
1	Vipers Kristiansand	0,26994	Györi Audi ETO KC	0,20118	Vipers Kristiansand	0,22049
2	Rostov-Don	0,18311	Vipers Kristiansand	0,1857	Rostov-Don	0,16741
3	Györi Audi ETO KC	0,16927	CSKA	0,17023	Györi Audi ETO KC	0,15638
4	CSKA	0,14032	Rostov-Don	0,14444	CSKA	0,1379
5	Brest Bretagne Handball	0,06984	Brest Bretagne Handball	0,10833	Brest Bretagne Handball	0,08932
6	Metz Handball	0,06814	Metz Handball	0,08843	Metz Handball	0,08531
7	CSM Bucuresti	0,06681	CSM Bucuresti	0,07074	CSM Bucuresti	0,08336
8	Buducnost	0,03256	Buducnost	0,03095	Buducnost	0,05984