

Óbudai Egyetem

Doktori (PhD) értekezés

tézisfüzete



**Az extrémérték-statisztika használata új metrikák kifejlesztésére
a cukorbetegség kockázatértékeléséhez és kísérleti adatok
elemzéséhez**

Szigeti Mátyás

Témavezető:

Ferenci Tamás

**Alkalmazott Informatikai és Alkalmazott Matematikai
Doktori Iskola**

Budapest, 2023

1. A kutatás előzményei

A statisztikai elemzés a modern orvosi kutatások alapvető részét képezi. Más tudományágakhoz hasonlóan az ebben alkalmazott módszerek is folyamatosan fejlődnek és új kihívásokkal néznek szembe. A disszertációm két példával szolgál erre.

A fő része az 1-es típusú cukorbetegségben szenvedő betegek kockázatainak extrémérték-elmélettel (extreme value statistics, EVS) történő értékelésére szolgáló, újszerű megközelítésen alapuló metrikák kifejlesztését mutatja be. Ez a statisztika egy kevésbé közismert ága, amely a ritka, szokatlanul alacsony vagy magas értékű megfigyelések eloszlásaira fókuszál. Ez kézenfekvően alkalmazható az orvostudományban [1], ahol az extrémértékek természetesen fontos szerepet játszanak, mivel a betegségek gyakran felfoghatóak olyan szélsőséges állapotként, melyekben egyes biomarkerek abnormálisan alacsony vagy magas szintjében manifesztálódnak.

Ennek ellenére legjobb tudomásom szerint e tanulmányt megelőzően mindössze két példa volt az EVS-módszerek alkalmazására a klinikai orvostudományban: egy elemzés a koleszterin szintekről [2] és egy tanulmány a tüdőgyulladásos és influenzás halálesetekről [3]; mindkettő aránylag a közelmúltban került publikálásra. Ellenben más területeken, ahol – az orvostudományhoz hasonlóan – az ilyen ritka, szélsőséges eseményeknek elsöprő hatása van [4] elterjedt az EVS alkalmazása, például az építészetben [5, 6, 7], időjárás- és klímaelemzésben [8, 9, 10] vagy sport és pénzügyi statisztikában [11, 12, 13].

A megfelelő mennyiségű adat hiánya komoly akadályt jelent az EVS szélesebb körű alkalmazása szempontjából orvosbiológiai területen, mivel a legtöbb esetben a fontos biomarkerek olyanok, amelyekhez valamiféle mintavétel szükséges és a kapott mintát sokszor összetett és költséges folyamatban elemzik ki, ami jelentősen korlátozza a megfigyelések számát. Ráadásul a szélsőséges események ritkaságuk miatt a teljes mintának csak egy töredékét teszik ki, így sokkal kisebb a tényleges, használható elemszám, ezért az elemzés rendkívüli nehézséget jelenthet.

A mérés technika fejlődése a diabetológiát ebből a szempontból kivételessé teszi. A folyamatos glükózmonitorozás (CGM) széles körű elterjedésével a gyakori (jellemzően 5 perces mintavételi idővel történő) és hosszabb távú (akár hetekig vagy hónapokig tartó, klinikai gyakorlatban is elérhető) mérések lehetővé váltak és aránylag régen bekerültek a gyakorlatba [14, 15].

A folyamatos glükózmonitorozás eredményeit feldolgozó jelenlegi gyakorlat számos olyan metrikát használ, amelyek lehetővé teszik a több napon vagy héten keresztül gyűjtött CGM-mérések gyors értékelését [16, 17, 18, 19, 20, 21]. Ezek a metrikák azonban többnyire az általános vércukorszint-variabilitásra összpontosí-

tanak, nem pedig a szélsőséges értékekre, ami nem feltétlenül ugyanaz: egy beteg variabilitása lehet nagyon magas, még akkor is, ha a vércukorszintje soha nincs a szélsőséges tartományban, vagy ennek az ellenkezője is előfordulhat (bár nagyon valószínűtlen), hogy a beteg sok időt tölt a szélsőséges tartományban, nagyon alacsony variabilitás mellett. Így ezek a metrikák nem igazán alkalmasak ennek a szempontnak és a magas vércukorszint (hiperglikémia), még kevésbé a hozzá kapcsolódó (rejtett) kockázatok jellemzésére. Még azok a metrikák is, amelyek figyelembe veszik az extrémításokat (mint például a hiperglikémiás határ felett töltött idő aránya), általában matematikailag nagyon egyszerű – és ad hoc – mutatók, amelyek nem tartalmazzák az extrémítások viselkedésére vonatkozó statisztikai ismereteket.

Ezekkel a metrikákkal ellentétben az EVS lehetővé teszi annak valószínűségének becslését, hogy egy jövőbeli mérés meghalad egy bizonyos küszöbértéket (ami a hiperglikémia szempontjából releváns tényező), még akkor is, ha ilyen értéket eddig még soha nem figyeltek meg a mintában. Ennek segítségével kiszámítható annak valószínűsége, hogy a beteg vércukorszintje egy adott időintervallumban (pl. 1 évben) meghalad egy bizonyos küszöbértéket és közvetlenül megadható az intervallumban a küszöbérték fölött várhatóan eltöltött idő is. Ezek együttesen felvetik annak lehetőségét, hogy az EVS-en alapuló metrikák alkalmasabbak a cukorbetegséggel kapcsolatos kockázatok pontosabb jellemzésére. Ez a megközelítés, amely közvetlenül az extrémértékekkel foglalkozik, megalapozottabb, kifinomultabb statisztikai módszereken nyugszik [11, 3, 22, 23].

2. Célkitűzések

E tanulmány fő célja a jelenlegi gyakorlatban CGM-adatok értékelésére használt hagyományos metrikák helyett új, a glükózmérések extrémértékeire összpontosító, EVS-megközelítésen alapuló, remélhetőleg pontosabb, könnyen érthető metrikák kidolgozása, melyek robusztusabb matematikai alapokon nyugszanak. Ennek érdekében az első lépés az elméleti háttér vizsgálata, a két fő EVS-módszer, úgymint a „küszöb feletti csúcsok” (peak over threshold, POT) és a „blokkmaximumok” (block maxima, BM) megközelítések értékelése és összehasonlítása volt a konkrét probléma megoldására való alkalmasság tekintetében.

Az előzetes vizsgálatok után az új, EVS-alapú metrikák alkalmazhatósága került pontos kidolgozásra és bemutatásra a REPLACE-BG kísérlet 14,8 millió mérést tartalmazó adatsorából, amely 226 fő 1-es típusú cukorbetegtől származott. Tudomásom szerint ez az első alkalmazása az EVS-alapú elemzésnek ezen a területen,

ezzel lehetővé téve a hagyományos és új metrikák átfogó összehasonlítását.

Ezen eredmények felhasználásával betegenkénti kockázatértékelre került sor, klinikailag fontos glükózsintek elérésének valószínűségét és a felettük eltöltött várható időt felhasználva. A metrikák közti korreláció meghatározása mellett további cél volt az EVS-elemzéssel azonosított legmagasabb kockázatú betegcsoportot összehasonlítása a hagyományos metrikák alapján meghatározott legmagasabb kockázatú csoporttal. Továbbá a tervek között szerepelt az elemzés bővítése nem-stacioner modellek alkalmazásával, ami lehetővé tette betegszintű klinikai jellemzők hozzáadását a modellhez, megbecsülve azok hatását a glükóz szintre. Ezen kívül cél volt megvizsgálni a statisztikai függetlenség kérdését a CGM idősoros adatai és az EVS elemzés tekintetében, vizsgálva azok validitását.

A részletesen bemutatott EVS-hez kapcsolódó elemzés mellett egy másik klinikai vizsgálat adatainak a regressziós modellezése, illetve annak statisztikai szempontjai is bemutatásra kerülnek. Ebben az esetben is fontos szerepet játszott a statisztikai függetlenség kérdése, és ez tette szükségessé komplex módszerek alkalmazását. Ebben a részben a RECITAL vizsgálat elsődleges elemzésének statisztikai aspektusai kerülnek bemutatásra, amely egy modern, multicentrikus, randomizált, kontrollált vizsgálat volt.

3. Vizsgálati módszerek

Az EVS fent említett két fő megközelítése számos hasonlóságot mutat, de a legfontosabb az, hogy mindkettő másodlagos mintát használ, amely a nyers adatokból kerül mintavételezésre és nevük tükrözi, hogy ez a másodlagos mintavétel miként történik. A POT megközelítés egy választott küszöbértéket használ és az efeletti megfigyelt értékeket választja be másodlagos mintába, míg a BM megközelítés pedig az adatokat egyenlő méretű, nem átfedő megfigyelési blokkokra osztja fel és minden egyes blokk maximumát (vagy minimumát) veszi másodlagos mintának az elemzéshez.

Történetileg a BM megközelítést fedezték fel és használták először, amikor a extrémértékek viselkedése először formálisan leírásra került Ronald Fisher és Leonard Henry Caleb Tippett által 1928-ban [24]. Megállapításaikat később Boris Vladimirovich Gnedenko 1943-ban bizonyítani tudta [25]. Ezek együttesen alkotják az úgynevezett Fisher-Tippett-Gnedenko-tételt, amely kimondja, hogyha léteznek olyan konstansok, amelyekkel független, azonos eloszlású valószínűségi változók maximuma lineárisan transzformálható úgy, hogy ez a renormált valószínűségi változó

egy nem elfajult eloszláshoz konvergál midőn a mintanagyság tart a végtelenbe, akkor ennek a határeloszlásnak az alábbiak egyikének kell lennie:

$$F(x) = \begin{cases} \exp \left[- \left\{ 1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right\}_+^{-1/\xi} \right] & \text{if } \xi \neq 0 \\ \exp \left[- \exp \left\{ - \left(\frac{x-\mu}{\sigma} \right) \right\} \right] & \text{if } \xi = 0 \end{cases}$$

Itt $\mu \in \mathbb{R}$ a hely-, $\sigma > 0$ a skála- és $\xi \in \mathbb{R}$ az alakparaméter. Szerencsére ezek az eloszlások zártak lineáris transzformációra nézve, így ahelyett, hogy azt mondanánk, hogy a transzformált maximum konvergál, a fenti megfogalmazással azt is mondhatjuk, hogy a maximum konvergál a fenti eloszlások valamelyikéhez az általánosság sérülése nélkül, mivel a hely- és skálaparamétereket amúgyis a mintából kell becsülni.

A fent bemutatott eloszlást általánosított extrémérték-eloszlásnak (generalized extreme value, GEV) nevezzük. A ξ értéke alapján három speciális esetet fed le, amelyek a következők:

1. Frechet eloszlás ($\xi > 0$):

$$F_\alpha(x) = \exp \left\{ - \left(\frac{x-u}{\sigma} \right)^{-\alpha} \right\}$$

2. Weibull eloszlás ($\xi < 0$):

$$F(x) = \exp \left\{ - \left(- \left(\frac{x-u}{\sigma} \right)^\alpha \right) \right\}$$

3. Gumbel eloszlás ($\xi = 0$):

$$F(x) = \exp \left\{ - \exp \left(\frac{x-u}{\sigma} \right) \right\}$$

Míg ez a leírás valószínűségelméleti szempontból már alkalmas elemzésre, a statisztikai vizsgálatban a maximumokat kell megbecsülni. Szükség van egy mintára, vagyis nem vehetjük egyszerűen a teljes sorozat maximumát és erre válaszul jelent meg először a BM megközelítés, vagyis a blokkmaximumok használata a maximumok eloszlásának leírására.

A GEV eloszlást három paraméter írja le: alak, skála és hely. Amikor egy modellt illesztünk a rendelkezésre álló empirikus adatokra (mintára), akkor ezeknek a

paramétereknek az értékeit – és annak bizonytalanságát – kell megbecsülni. Számos statisztikai módszer használható, de a gyakorlatban leggyakrabban a maximum likelihood becslés, az L-momentum és a Bayes-módszer használatos [26, 27]. A maximum likelihood becslés relatíve jó tulajdonságai és egyszerűsége miatt gyakran használt módszer és nagy adathalmazok esetén is alkalmazható, ahol más, számításigényesebb módszerek nem használhatóak. Az L-momentum módszer a valószínűséggel súlyozott momentumok lineáris kombinációin alapul, míg a bayesi megközelítés a jól ismert Bayes-tételt használja a paraméterek poszterior eloszlásának meghatározására a minta információinak és egy prior eloszlásnak a felhasználásával.

A másik fő megközelítést alkalmazva a választott küszöbérték feletti másodlagos minta (POT) aszimptotikusan az úgynevezett általánosított Pareto-eloszlást (generalised Pareto distribution, GPD) eloszlást követi, amelyet James Pickands III fél évszázaddal Fisher és Tippett munkája után, 1975-ben mutatott be először [28]. Kimutatta, hogy ezeknek a extrémértékeknek a viselkedése a POT másodlagos mintavétel után a következő valószínűségi törvényt követi:

$$P(X - u < y | X > u) \approx 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-1/\xi},$$

ahol ξ az alak- és σ a skálaparaméter.

4. Új tudományos eredmények

4.1. 1. téziscsoport

Tézis 1.1 Olyan új megközelítést dolgoztam ki, amely a diabetológia hagyományos mérési módszerei által használt variabilitás helyett a folyamatos glükózmonitorozás (CGM) méréseinek maximumait helyezi a középpontba a extrémérték-elmélet (EVS) alkalmazásával. Szimulált adatok használatával bemutattam, hogy az EVS a küszöb feletti csúcsérték (POT) megközelítéssel használható a CGM-görbék jellemzésére és klinikailag releváns metrikák előállítására a betegszintű kockázatok jellemzésére azáltal, hogy megbecsüljük annak valószínűségét, hogy egy új csúcsérték meghalad egy bizonyos küszöbértéket, azonban a küszöbérték megválasztása problematikus lehet.

A tézishez kapcsolódó publikáció: [29].

Tézis 1.2 Első alkalommal alkalmaztam az EVS blokkmaximum (BM) megközelítést a CGM-görbék jellemzésére. Bebizonyítottam, hogy klinikailag is releváns becsléseket adhat, amelyek metrikákként használhatók a betegszintű kockázatok értékelésére és számos előnnyel rendelkezik a POT-módszerrel szemben. A módszer segítségével választott küszöbértékek átlépésének valószínűsége és a felettük töltött várható idő kiszámítható, úgy, hogy ezek a küszöbértékek akár túl is lehetnek a CGM-mérések tartományán. Az elemzés valós életből származó adathalmazon történt.

A tézishoz kapcsolódó publikációk: [29, 30].

Ezzel az elemzéssel bizonyítást nyert az EVS és a BM megközelítés alkalmazásának életképessége a glükózsabályozás minőségének a CGM görbe alapján történő értékelésére és így az ennek tulajdonított kockázat jellemzésére is bizonyos mértékig. Világossá vált azonban, hogy a BM megközelítésnek fontos előnye van a POT megközelítéssel szemben és klinikailag relevánsabb metrikákhoz vezethet, hiszen közvetlenül mutatja bizonyos extrém értékek elérésének valószínűségét egyes betegek esetében, valamint az ezek felett eltöltött időt, még akkor is, ha a küszöb a megfigyelt tartományon túl van. A POT lehetővé teszi bizonyos extrémértékek elérésének valószínűségére vonatkozó becslések kiszámítását, de nem veszi figyelembe a extrémértékek időbeni előfordulását, ezért kevésbé alkalmas betegszintű értékelésre.

4.2. 2. Téziscsoport

Tézis 2.1 Az EVS blokkmaximum (BM) megközelítést a REPLACE-BG kísérlet 226 betegéből álló, nagy mintáján alkalmaztam, amelynek CGM-görbéi több mint 14,8 millió mérést tartalmaztak. Először hasonlítottam össze az EVS-metrikákat a széles körben használt hagyományos metrikákkal a hiperglikémia betegszintű kockázatértékeléséhez. Általánosságban relatíve gyenge vagy mérsékelt korrelációt volt megfigyelhető az EVS és a hagyományos metrikák között.

A tézishoz kapcsolódó publikációk: [29, 30, 31].

Tézis 2.2 Az új, EVS metrikák szerinti legmagasabb kockázatú betegek a hagyományos metrikák szerint nem értek el kimagasló pontszámokat. A mérések további vizsgálata azt mutatta, hogy ezeket erősen befolyásolta a CGM szenzor mérés határa okozta szaturáció. A további elemzés kimutatta, hogy az EVS-metrikák érzékenyebbek voltak a szaturációs szintek szimulált csökkenésére.

A tézishoz kapcsolódó publikációk: [31].

Tézis 2.3 A regressziós elemzésekhez hasonlóan magyarázóváltozókat is hozzá lehet adni az EVS modellekhez, hogy megvizsgáljuk azok hatását a modellezn kívánt eredményre. Megvizsgáltam a testtömegindex (BMI) hatását a vércukormaximumokra és statisztikailag szignifikáns hatást találtam: nagyobb testtömegindex a vércukorszintek óránkénti maximumainak magasabb értékeivel járt együtt.

A tézishoz kapcsolódó publikációk: [31, 32].

A REPLACE-BG adatainak felhasználása, amely több mint 14,8 millió valós CGM mérést tartalmaz, hatalmas lökést adott az EVS módszerek ilyen elemzésekre való fejlesztésére és alkalmazhatóságának bizonyítására irányuló erőfeszítéseknek. Legjobb tudásom szerint ez az elemzés és annak eredményei voltak az első olyan esetek, amikor az EVS-t a diabetológiában alkalmazták és különösen nagy adathalmazon. Az új EVS metrikák eredményei összehasonlításra kerültek a CGM mérések értékelésére és betegsúly jellemzésére használt legfontosabb és legszélesebb körben használt hagyományos metrikákkal. Bebizonyosodott, hogy az EVS lehetővé teszi a CGM mérések jellemzését a hiperglikémiás kockázat szempontjából a relevánsabb, extrém értékekre összpontosítva és lényegi, klinikailag könnyen értelmezhető, betegsúly metrikák létrehozására.

A CGM-mérések validitása nagyszámú, rendelkezésre álló megerősítő vércukorméréssel került ellenőrzésre, különös tekintettel a felső detektációs limitre 400 mg/dl-nél, ami a CGM szenzor fontos korlátjának bizonyult és a mért értékek telítődését okozta. Alacsonyabb detektációs limit – vagyis alacsonyabb telítődési pontok – hatásának szimulálásával bemutatásra került, hogy általánosságban az új EVS-metrikák érzékenyebbek bizonyultak erre a hatásra a hagyományos metrikákhoz képest.

Ezenfelül egy másik típusú, extrémérték-megközelítést alkalmazó elemzéssel a különböző egyéni klinikai jellemzők vagy kezelések hatását is pontosabban és prakti-

kusabban lehet értékelni nem-stacioner modellek segítségével. Ennek az alkalmazása ebben az esetben statisztikailag szignifikáns összefüggést mutatott ki a magasabb testtömegindex és a magasabb glükózmaximumok között.

4.3. 3. Téziscsoport

Tézis 3.1 Egy kevert hatású regressziós modellezési stratégiát mutattam be, mint számos alternatívájánál megalapozottabb és alkalmasabb megoldást, egy három szintű, klaszterezett, hierarchikus kísérlet adatszerkezetének modellezésére. Ennek révén megbízható statisztikai elemzésre került sor egy modern, multicentrikus, randomizált, kontrollált vizsgálatról, amely az első olyan nagyszabású vizsgálat volt, ami a rituximab hatékonyságát vizsgálta a ciklofoszfamidral összehasonlítva a kötőszöveti betegséggel társuló intersticiális tüdőbetegség kezelésében.

A tézishoz kapcsolódó publikációk: [33, 34].

Ez a rész egy más környezetben összpontosít a statisztikai függetlenség problémáira, ami az EVS-elemzéshez használt idősoros adatokkal kapcsolatban is vizsgálatra került. Mivel ebben a klinikai kísérletben a végpontok folytonos változók voltak, melyeknél minden betegnél több mérést végeztek a követés során, ezeket nem lehetett függetlenként kezelni és emiatt a közönséges regressziós modellek és statisztikai módszerek nem lennének alkalmasak az elemzésre. További csavar, hogy ezek a megfigyelések egy harmadik szinten, a betegeket felvett kórház szerint is klaszterbe sorolhatók (amelynek annál is inkább fontosabb szerepe lehet, mivel a légzésfunkciós méréseket spirométerekkel végezték), ami nem tette lehetővé az elemzéshez az ismételt méréseknél általában használt egyéb statisztikai módszerek alkalmazását. További előnye a bemutatott megközelítésnek, hogy lehetővé teszi az összes rendelkezésre álló adat felhasználását, beleértve a hiányzó megfigyeléseket tartalmazó betegeket is, valamint az egyenlőtlen csoportméretek kezelését, ami valós klinikai körülmények között elkerülhetetlen. Lehetővé teszi továbbá a módszer további magyarázóváltozók hozzáadását a modellhez. Ez nagyobb rugalmasságot és robusztusságot kölcsönöz az elemzésnek, ami rendkívül hasznos a klinikai vizsgálatokban, ahol ideális esetben az elemzés részleteit előre szükséges megtervezni, úgy, hogy a statisztikusnak semmilyen vagy korlátozott ismeretei vannak csak a tényleges adatokról. Az előnyök mellett a kevert hatású modellek matematikai háttere és a lineáris regresszióhoz képesti különbségei is bemutatásra kerültek egy példán keresztül, amely részletes magyarázatot adott arra, hogy hogyan képes modellezni ezeket

az összefüggéseket, amikor a megfigyelések függetlenségére vonatkozó feltételek nem teljesülnek.

5. Az eredmények hasznosítási lehetősége

Az EVS és annak BM megközelítése révén olyan új metrikákat kerültek kifejlesztésre, amelyek könnyen értelmezhető eszközök a magas glükózsinték betegsúlyú kockázatértékelésében, mivel közvetlenül megbecsülhetik bizonyos, klinikailag fontos küszöbértékek felett eltöltött időt, még akkor is, ha a mintában ezeket a küszöbértékeket soha nem érték el a betegek. Ezek matematikailag megalapozottabbak és kifinomultabbak, mint a – többnyire egyszerű és ad hoc – klasszikus metrikák és ami fontosabb, hogy legnagyobb kockázatúként más betegeket azonosítanak az új és a hagyományos metrikák. További fontos eredmény, hogy a szenzorok detektációs limitje lényeges szerepet játszhat, mivel a legmagasabb kockázatú betegek EVS metrikákon keresztül történő értékelését erősen befolyásolta a szenzor felső detektációs limitje.

Ezen felül az elemzés tartalmazza a testtömegindex és a glükózmérések óránkénti maximumai közötti kapcsolat értékelését, nem-stacioner modellek alkalmazásával. Ez az elemzés kimutatta, hogy statisztikailag szignifikáns kapcsolat van a magasabb testtömegindex és magasabb glükózmáxima között.

Az elemzési folyamat során egy váratlan hiba került felfedezésre az elemzéshez használt statisztikai szoftvercsomagban, amely fals eredményekhez vezetett az egyik hagyományos metrika kiszámításakor. Ennek messzemenő következményei voltak, mivel ez a csomag aránylag széles körben elterjedt a glikémiás variabilitás elemzésére, és a hiba legalább két publikált klinikai kísérlet elsődleges eredményeit befolyásolta. A csomag készítője megerősítette ezeket a megállapításokat és egy későbbi verzióban kijavította a hibát. A két publikált vizsgálat szerzőit értesítettem minderről.

A munka utolsó része egy másik, eltérő orvosbiológiai alkalmazását mutatja be klinikai kísérletek eredményeinek statisztikai modellezésének. A bemutatott kísérlet elemzése a multi-centrikusság révén, illetve az ismételt mérések miatt háromszintű, hierarchikus, kevert hatású regressziós modell használatát igényelte. Ennek a módszernek a háttere, az alternatívái és a választás okai kerültek bemutatásra. A kérdéssel kapcsolatos probléma a statisztikai függetlenség volt, amely az EVS elemzésben is megjelent, de egészen más összefüggésben. A bemutatott munka egy klinikai kísérlet valódi elsődleges elemzése volt, így a vonatkozó irányelvek követelményeinek teljesítése mellett az elemzéssel kapcsolatos legtöbb döntést már korán,

az adatok korlátozott ismeretében vagy teljes hiányában kellett meghozni, ami a statisztika más alkalmazásaihoz képest jelentős nehézségeket jelentett. Így az alkalmazott regressziós modellek rugalmassága és robusztussága fontos szerepet játszott ebben a kérdésben. Végül ez az elemzés egy jól megalapozott, tudományos bizonyítékot szolgáltatott egy potenciálisan életveszélyes tüdőbetegség új kezelésének hatékonyságáról és biztonságosságáról a standard, jelenleg alkalmazott kezeléssel összehasonlítva.

Irodalmi hivatkozások listája

- [1] Stephen J Roberts. “Extreme value statistics for novelty detection in biomedical data processing”. *IEE Proceedings-Science, Measurement and Technology* 147.6 (2000), 363–367. old.
- [2] P. Bermudez és Zilda Mendes. “Extreme Value Theory in Medical Sciences: Modeling Total High Cholesterol Levels”. *Journal of statistical theory and practice* 6 (2012. szept.), 468–491. old. DOI: 10.1080/15598608.2012.695673.
- [3] Maud Thomas, Magali Lemaitre, Mark L Wilson, Cécile Viboud, Youri Yordanov, Hans Wackernagel és Fabrice Carrat. “Applications of extreme value theory in public health”. *PloS one* 11.7 (2016), e0159312.
- [4] RR Kinnison. *Applied extreme-value statistics*. Techn. jel. Pacific Northwest Lab., Richland, WA (USA), 1983.
- [5] Michel K Ochi. “Principles of extreme value statistics and their application”. *Paper of the Society of Naval Architects and Marine Engineers, SNAME, 1981* (1981).
- [6] Erwan Le Roux, Guillaume Evin, Nicolas Eckert, Juliette Blanchet és Samuel Morin. “Non-stationary extreme value analysis of ground snow loads in the French Alps: a comparison with building standards”. *Natural Hazards and Earth System Sciences* 20.11 (2020), 2961–2977. old.
- [7] Jonathan Auerbach és Phyllis Wan. “Forecasting the urban skyline with extreme value theory”. *International Journal of Forecasting* 36.3 (2020), 814–828. old.
- [8] Richard W Katz. “Statistics of extremes in climate change”. *Climatic change* 100.1 (2010), 71–76. old.

- [9] Georgia Lazoglou és Christina Anagnostopoulou. “An overview of statistical methods for studying the extreme rainfalls in Mediterranean”. *Multidisciplinary Digital Publishing Institute Proceedings* 1.5 (2017), 681. old.
- [10] Erin Towler, Balaji Rajagopalan, Eric Gilleland, R Scott Summers, David Yates és Richard W Katz. “Modeling hydrologic and water quality extremes in a changing climate: A statistical approach based on extreme value theory”. *Water Resources Research* 46.11 (2010).
- [11] Manfred Gilli és tsai. “An application of extreme value theory for measuring financial risk”. *Computational Economics* 27.2 (2006), 207–228. old.
- [12] MB Adam és Jonathan Angus Tawn. “Modelling record times in sport with extreme value methods”. *Malaysian Journal of Mathematical Sciences* 10.1 (2016), 1–21. old.
- [13] Serguei Y Novak. *Extreme value methods with applications to finance*. CRC Press, 2011.
- [14] SF Clarke és JR Foster. “A history of blood glucose meters and their role in self-monitoring of diabetes mellitus”. *British journal of biomedical science* 69.2 (2012), 83–93. old.
- [15] Bruce W Bode. “Clinical utility of the continuous glucose monitoring system”. *Diabetes Technology & Therapeutics* 2.1, Supplement 1 (2000), 35–41. old.
- [16] F John Service, George D Molnar, John W Rosevear, Eugene Ackerman, Lael C Gatewood és William F Taylor. “Mean Amplitude of Glycemic Excursions, a Measure of Diabetic Instability”. *Diabetes* 19.9 (1970), 644–655. old. ISSN: 0012-1797. DOI: 10.2337/diab.19.9.644. eprint: <https://diabetes.diabetesjournals.org/content/19/9/644.full.pdf>. URL: <https://diabetes.diabetesjournals.org/content/19/9/644>.
- [17] C.M. McDonnell, S.M. Donath, S.I. Vidmar, G.A. Werther és F.J. Cameron. “A Novel Approach to Continuous Glucose Analysis Utilizing Glycemic Variation”. *Diabetes Technology & Therapeutics* 7.2 (2005). PMID: 15857227, 253–263. old. DOI: 10.1089/dia.2005.7.253. eprint: <https://doi.org/10.1089/dia.2005.7.253>. URL: <https://doi.org/10.1089/dia.2005.7.253>.
- [18] T Battelino, T Danne, RM Bergenstal, SA Amiel, R Beck, T Biester, E Bosi, BA Buckingham, WT Cefalu, KL Close és tsai. “Doyle FJ 3rd, Garg S, Grunberger G, Heller S, Heinemann L, Hirsch IB, Hovorka R, Jia W, Kordonouri O, Kovatchev B, Kowalski A, Laffel L, Levine B, Mayorov A, Mathieu

- C, Murphy HR, Nimri R, Nørgaard K, Parkin CG, Renard E, Rodbard D, Saboo B, Schatz D, Stoner K, Urakami T, Weinzimer SA, Phillip M. Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range”. *Diabetes Care* 42.8 (2019), 1593. old.
- [19] Lalo Magni, Davide M Raimondo, Chiara Dalla Man, Marc Breton, Stephen Patek, Giuseppe De Nicolao, Claudio Cobelli és Boris P Kovatchev. “Evaluating the efficacy of closed-loop glucose regulation via control-variability grid analysis”. *Journal of diabetes science and technology* 2.4 (2008), 630–635. old.
- [20] David Rodbard. “Evaluating quality of glycemic control: graphical displays of hypo-and hyperglycemia, time in target range, and mean glucose”. *Journal of diabetes science and technology* 9.1 (2014), 56–62. old.
- [21] Michelle Nguyen, Julia Han, Elias K Spanakis, Boris P Kovatchev és David C Klonoff. “A review of continuous glucose monitoring-based composite metrics for glycemic control”. *Diabetes technology & therapeutics* 22.8 (2020), 613–622. old.
- [22] Daniel Cooley, Douglas Nychka és Philippe Naveau. “Bayesian spatial modeling of extreme precipitation return levels”. *Journal of the American Statistical Association* 102.479 (2007), 824–840. old.
- [23] Lee Fawcett és David Walshaw. “Estimating return levels from serially dependent extremes”. *Environmetrics* 23.3 (2012), 272–283. old.
- [24] R. A. Fisher és L. H. C. Tippett. “Limiting forms of the frequency distribution of the largest or smallest member of a sample”. en. *Mathematical Proceedings of the Cambridge Philosophical Society* 24.2 (1928. ápr.), 180–190. old. ISSN: 1469-8064, 0305-0041. DOI: 10.1017/S0305004100015681. (Elérés dátuma 2020. 02. 09.).
- [25] B. Gnedenko. “Sur La Distribution Limite Du Terme Maximum D’Une Serie Aleatoire”. *Annals of Mathematics* 44.3 (1943), 423–453. old. ISSN: 0003486X. URL: <http://www.jstor.org/stable/1968974>.
- [26] Paul Embrechts, Claudia Klüppelberg és Thomas Mikosch. *Modelling extremal events: for insurance and finance*. 33. köt. Springer Science & Business Media, 2013.
- [27] Laurens De Haan, Ana Ferreira és Ana Ferreira. *Extreme value theory: an introduction*. 21. köt. Springer, 2006.

- [28] James Pickands III és tsai. “Statistical inference using extreme order statistics”. *the Annals of Statistics* 3.1 (1975), 119–131. old.

A tézispontokhoz kapcsolódó tudományos közlemények

- [29] Mátyás Szigeti, Tamás Ferenci és Levente Kovács. “The use of peak over threshold methods to characterise blood glucose curves”. *2020 IEEE 14th International Symposium on Applied Computational Intelligence and Informatics (SACI)*. 2020, 199–000204. old. DOI: 10.1109/SACI49304.2020.9118838.
- [30] Mátyás Szigeti, Tamás Ferenci és Levente Kovács. “The use of block maxima method of extreme value statistics to characterise blood glucose curves”. *2020 IEEE 15th International Conference of System of Systems Engineering (SoSE)*. 2020, 433–438. old. DOI: 10.1109/SoSE50414.2020.9130427.
- [31] Mátyás Szigeti, Tamás Ferenci és Levente Kovács. “The Use of Extreme Value Statistics to Characterize Blood Glucose Curves and Patient Level Risk Assessment of Patients With Type I Diabetes”. *Journal of Diabetes Science and Technology* 17.2 (2023). PMID: 34814774, 400–408. old. DOI: 10.1177/19322968211059547.
- [32] Tamás Ferenci, Mátyás Szigeti és Levente Kovács. “Using non-stationary extreme value analysis to characterize blood glucose curves”. *2022 IEEE 20th Jubilee World Symposium on Applied Machine Intelligence and Informatics (SAMI)*. IEEE. 2022, 171–000176. old. DOI: 10.1109/SAMI54271.2022.9780743.
- [33] Toby M Maher, Veronica A Tudor, Peter Saunders, Michael A Gibbons, Sophie V Fletcher, Christopher P Denton, Rachel K Hoyles, Helen Parfrey, Elisabetta A Renzoni, Maria Kokosi, Matyas Szigeti és tsai. “Rituximab versus intravenous cyclophosphamide in patients with connective tissue disease-associated interstitial lung disease in the UK (RECITAL): a double-blind, double-dummy, randomised, controlled, phase 2b trial”. *The Lancet Respiratory Medicine* 11.1 (2023), 45–54. old. ISSN: 2213-2600. DOI: [https://doi.org/10.1016/S2213-2600\(22\)00359-9](https://doi.org/10.1016/S2213-2600(22)00359-9).
- [34] Vicky Tsipouri, Peter Saunders, Greg J. Keir, Deborah Ashby, Sophie V. Fletcher, Michael Gibbons, Matyas Szigeti, Helen Parfrey, Elizabeth A. Renzoni és Chris P. Denton. “Poster: Rituximab versus cyclophosphamide for the treatment of connective tissue disease associated interstitial lung disease

(RECITAL): a randomised controlled trial”. *4th International Clinical Trials Methodology Conference (ICTMC) and the 38th Annual Meeting of the Society for Clinical Trials*. *Trials*, 2017. máj., 130. old. DOI: 10.1186/s13063-017-1902-y.

További, a tézispontokhoz nem kapcsolódó tudományos közlemények

- [35] Manjit S Gohel, Jocelyn Mora, Matyas Szigeti, David M Epstein, Francine Heatley, Andrew Bradbury, Richard Bulbulia, Nicky Cullum, Isaac Nyamekye, Keith R Poskitt és tsai. “Long-term clinical and cost-effectiveness of early endovenous ablation in venous ulceration: a randomized clinical trial”. *JAMA surgery* 155.12 (2020), 1113–1121. old. DOI: 10.1001/jamasurg.2020.3845.
- [36] Ping-Tee Tan, Suzie Cro, Eleanor Van Vogt, Matyas Szigeti és Victoria R Cornelius. “A review of the use of controlled multiple imputation in randomised controlled trials with missing outcome data”. *BMC medical research methodology* 21.1 (2021), 1–17. old. DOI: 10.1186/s12874-021-01261-6.
- [37] Neil R Poulter, Christos Savopoulos, Aisha Anjum, Martha Apostolopoulou, Neil Chapman, Mary Cross, Emanuela Falaschetti, Spiros Fotiadis, Rebecca M James, Ilias Kannelos, Matyas Szigeti és tsai. “Randomized crossover trial of the impact of morning or evening dosing of antihypertensive agents on 24-hour ambulatory blood pressure: the HARMONY trial”. *Hypertension* 72.4 (2018), 870–873. old. DOI: 10.1161/HYPERTENSIONAHA.118.11101. URL: <https://www.ahajournals.org/doi/abs/10.1161/HYPERTENSIONAHA.118.11101>.
- [38] Mátyás Szigeti, Levente Kovács és Tamás Ferenci. “Stability of relative and absolute metrics: empirical evidence from pulmonology”. *2019 IEEE 17TH World Symposium on Applied Machine Intelligence and Informatics (SAMI 2019)*. 2019, 235–238. old. DOI: 10.1109/SAMI.2019.8782769.
- [39] Livia Priyanka Elek, Matyas Szigeti, Berta Erdélyi-Hamza, Mátyás Szigeti, Konstantinos N Fountoulakis és Xénia Gonda. “What you see is what you get? Association of belief in conspiracy theories and mental health during COVID-19.” *Neuropsychopharmacologia Hungarica* 24 (2022), 42–55. old. ISSN: 1419-8711. URL: <https://europepmc.org/article/med/35451591>.

- [40] Livia Priyanka Elek, Mátyás Szigeti, Berta Erdélyi-Hamza, Nikolett Beáta Vadon, Konstantinos N. Fountoulakis, Daria Smirnova és Xénia Gonda. “Association of lifestyle changes during the pandemic are associated depression and its distinct symptom clusters – consideration for prevention and intervention”. *Neuroscience Applied* (2023). ISSN: 2772-4085. DOI: 10.1016/j.nsa.2022.100818. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9789424/>.
- [41] Nikolett Beata Vadon, Livia Priyanka Elek, Matyas Szigeti, Berta Erdelyi-Hamza, Daria Smirnova, Konstantinos N Fountoulakis és Xenia Gonda. “Association between Lifestyle-and Circadian Rhythm-Related Changes, and Different Depression Symptom Clusters during COVID-19.” *Psychiatria Danubina* 34.Suppl 8 (2022. szept.), 81–89. old. ISSN: 0353-5053. URL: <https://europepmc.org/article/med/36170708>.
- [42] Péter Torzsa, László Kalabay, Dalma Dorottya Csatlós, Csenge Hargittay, Bernadett Márkus, András Mohos, Mátyás Szigeti, Tamás Ferenci, Verschoor Marjolein, Rozsnyai Zsafia, Gussekloo Jacobijn, K. E. Poortvliet Rosalinde és Streit Sven. “A nagyon idős és esendő állapotú betegek antihipertenzív kezelési gyakorlata az alapellátásban”. *Lege Artis Medicinae* 30 (2020), 111–121. old. ISSN: 0866-4811. DOI: 10.33616/lam.30.011. URL: http://real-j.mtak.hu/14007/7/LAM_2020_03.pdf#page=30.
- [43] Péter Torzsa, László Kalabay, Dalma Dorottya Csatlós, Csenge Hargittay, Bernadett Márkus, András Mohos, Mátyás Szigeti, Tamás Ferenci, Marjolein Verschoor, Zsófia Rozsnyai, Gussekloo Jacobijn, Rosalinde KE Poortvliet és Sven Streit. “Antihipertenzív kezelés: A családorvosok inkább kezelik az esendő pácienseket”. *Medical Tribune* 18 (2020), 39–41. old. ISSN: 1589-1283.
- [44] Tímea Vissi, Regina Szabó, Blanka Bágyi, Adél Göntér, Fanni Akkir, Mátyás Szigeti, Gabriella Erzsébet Papp, Éva Feketéné Szabó és Anna Kelemen. *Cerebrális parézissel élő gyermekek számára készült diagnózis specifikus életminőség felmérő kérdőív (CPQOL) magyar nyelven történő alkalmazása*. 2018.
- [45] Anita Zadori, Zsuzsanna Kis, Tibor Toth, Matyas Szigeti, Andras Temesvari, Geza Fontos, Noémi Nyolczas és Peter Andreka. “Long-Term Efficacy and Safety of Left Atrial Appendage Closure Procedures”. *International Heart Journal* 64.2 (2023), 188–195. old. DOI: 10.1536/ihj.22-639.