Kehl Dániel

Mintaelemszám tervezés Likert-skálás lekérdezések esetén klasszikus és bayesi keretek között

PhD értekezés tézisei

Témavezető: dr. Rappai Gábor

Tartalomjegyzék

1) A kutatás célja

A „Mintaelemszám tervezés Likert-skálás lekérdezések esetén klasszikus és bayesi keretek között” című dolgozatomban a mintatervezés egyik szűk, mégis nagy jelentőséggel bíró részterületét járom körül és próbálom újszerű eredményekkel kiegészíteni.

Hagyományosan a mintaelemszám tervezés a becslés standard hibájából, pontosabban a hibahatárból indul ki. Az aránybecslés esetén a mintaelemszámra rendezett képletben a z-érték és a kívánt hibahatár mellett az ismeretlen sokasági arány szerepel a független változóként. A leggyakoribb feltevés, hogy erről nem tudunk semmit a vizsgálat előtt, így a lehető legrosszabb esettel számolunk, ami gyakran túlbecsüli a szükséges mintaelemszámot, mégis a gyakorlatban hatékony segítséget jelent ez a megközelítés. A várható érték becslése esetén a sokasági arány helyett a sokasági variancia szerepel a megfelelő képletben. Az irodalomban javaslatként korábbi kutatási eredmények alapján meghatározott variancia szerepeltetése, illetve előzetes minta alapján történő becslés merül fel általánosan.

Dolgozatomban érintem a kétkimenetelű, aránybecslésre vezető esetet is, a fő hangsúly azonban azon van, hogy a sokasági variancia, illetve a sokasági szórás bizonyos, kevés kimenetellel rendelkező esetekben (pl. Likert-skála) hogyan becsülhető hatékonyan, amennyiben a kutatás végzőjének van előzetes elképzelése az eloszlás várható alakjáról. Olyan eloszlásokat definiálok, melyek a gyakorlati esetek nagyobb részét lefedik, majd ezekre meghatározom a varianciát, illetve az ebből származó becsült szükséges mintaelemszámot. Foglalkozom azzal a kérdéssel, hogy a lehetséges összes eloszlás figyelembe vétele esetén mekkora a szükséges mintaelemszám várható értéke.

Alapvetően tehát abból indulok ki, hogy a felmérést végző bizonyos előzetes, ha úgy tetszik, prior ismeretekkel rendelkezik a vizsgált sokaságról, ami megnyilvánulhat közvetlenül a variancia ismeretében is, azonban jellemzőbb, hogy magáról az eloszlásról van ismeretünk. A bayesi statisztika gondolati kerete (többek között) abban tér el a klasszikus statisztikáétól, hogy feltételezi előzetes információk létezését. A mintabeli információk és az a priori ismeretek ötvözésével előálló ún. poszterior eloszlás mindkét fajta tudást figyelembe veszi. Ennek alapján a bayesi statisztika megfelelő keretet nyújt egy olyan szituáció kezeléséhez, ahol az előzetes ismeretek felhasználása kerül szóba.

A bayesi statisztika és ökonometria napjainkban reneszánszát éli a nemzetközi szakirodalomban, hazánkban azonban csupán néhány tanulmány, illetve műhely foglalkozik vele. A módszertan térnyerésének több oka van, többek között az, hogy a Bayes-tétel gyakorlatilag önmagában elegendő az elméleti keret megértéséhez. Sok esetben lehetőség van (sőt, szükséges) külső információk ellenőrzött, dokumentált módon modellbe való építésére. Szintén okként említhetjük, hogy az elemzés végeredményeként előálló poszterior eloszlás önmagában a paraméterek eloszlását írja le, így konfidencia intervallumok előállításához nincs szükségünk kiegészítő feltevések (pl. normalitás) használatára. A bayesi módszerek nehézségét az okozza, hogy a (jellemzően sokdimenziós) poszterior eloszlásban rejlő információk prezentálása, megértése nem triviális, leggyakrabban szimulációs technikákat kell segítségül hívnunk. Ezen szimulációs technikák matematikai alapjai hosszabb történetre tekintenek vissza, azonban sokáig hiányzott az eljárásokhoz szükséges számítási, számítástechnikai háttér. Amióta mindkét alapfeltétel adott, a terület fejlődése töretlen. A bayesi megközelítésben lehetőségünk van tehát az előzetes ismeretek és a mintából származó információk egyesítésére. Éppen ez az a két módszer, amelyet a szakirodalom a várható érték becslés hibahatárából adódó mintanagyság képletben a szórás közelítésére javasol!

A dolgozatban a célom a mintaelemszám tervezésének hatékony módszerét kidolgozni. A klasszikus mellett a bayesi statisztikai gondolatkört is felhasználom, valamint összevetem az általuk szolgált eredményeket. Azt várom, hogy a munkám olyan kézzel fogható eljárásokkal gazdagítja a szakirodalmat, amelyeket a gyakorlati statisztika képes felhasználni. A bayesi statisztika alapgondolatainak és fő módszereinek áttekintését jelen dolgozattól függetlenül is fontosnak tartom.

2) Az értekezés felépítése

Az értekezés összesen nyolc fő fejezetet tartalmaz, amelyből négy alkotja a munka anyagának döntő részét. A disszertáció több éves munka eredménye, így az egyes fejezetek a korábbi években megjelent – jellemzően egyszerzős – cikkekben már publikálásra kerültek.

A bevezető fejezet után a második azzal foglalkozik, hogy a címben szereplő Likert-skálás kérdések esetén milyen statisztikai mutatók alkalmazhatók (Kehl, 2011). Az évtizedeken át zajló vita még a mai napig is tart, azt azonban megállapíthatjuk, hogy a legtöbb kutató egyetért abban, hogy átlag és szórás mutatókat csak olyan esetben számítunk, amikor a változó értékek ekvidisztansok, tisztán ordinális esetben ezek a mutatók nem alkalmazhatók. Az alkalmazók többsége Likert-skála használata esetén feltételezi ennek a jellemzőnek a teljesülését.

A harmadik fejezetben a mintaelemszám tervezés klasszikus statisztikai szemléletben történő vizsgálatával foglalkozom (Kehl-Rappai, 2006). A fejezet elején a Likert-skála eredetét, fő felhasználási területeit, majd a mintaelemszám tervezés általános módszerét mutatom be. Előre definiált eloszlások esetén határozok meg szórásjellemzőket, amelyek segítségével az adott eloszlású véletlen változóhoz tartozó szükséges mintaelemszámot becsülöm. Megtörténik egyfajta érzékenységvizsgálat és a különböző feltételezések mellett nyert elemszámok összehasonlítása, valamint a minden lehetséges mintára vonatkozó várható érték meghatározása (Kehl, 2007) is.

A negyedik fejezetben a bayesi gondolkodás alapjait foglalom össze, valamint kitérek azokra az alapvető és összetettebb módszerekre is, melyek a poszterior eloszlás összefoglalását segítik (Kehl, 2012). A szakasz gyakorlatilag egy általános bevezető a bayesi statisztikába, majd az azt kiegészítő szimulációs (MC és MCMC) technikákat 1-1 rövid példával kiegészítve mutatom be. A szimulációk futtatásához az R környezetet használom, a Függelék a szükséges kódokat tartalmazza, így az anyag a későbbiekben egy magyar nyelvű, bayesi statisztikával, szimulációval foglalkozó kurzus alapja is lehet.

Az ötödik fejezet a bayesi megközelítés gyakorlati megvalósítását mutatja be a dolgozat fő problémáján keresztül két alfejezetben. Elsőként definiálom a kétváltozós esetben könnyen alkalmazható konjugált prior eloszlást, amely rugalmasan képes a rendelkezésre álló előzetes információk leírására. A poszterior pedig tartalmazza az esetleges előzetes mintavételi eredményeket is, így a két tudás kombinációjaként értelmezhető. A második – jóval rövidebb – alpont a kettőnél több kimenetellel rendelkező változók esetét mutatja be. A fejezet rövidségét az teszi lehetővé, hogy a kétváltozós eset gyakorlatilag gond nélkül általánosítható többváltozósra. A binomiális eloszlás helyét a multinomiális, a bétáét a Dirichlet veszi át, így a folyamat gyakorlatilag teljes mértékben megegyezik a kétváltozós esetben leírttal.

A hatodik fejezet az értekezés eredményeinek rövid összefoglalását adja és további potenciális kutatási területeket azonosít, míg az utolsó két pont a főszövegbe terjedelmüknél fogva nem illő levezetéseket, bizonyos eloszlások jellemzőit, illetve a már említett programkódokat tartalmazza. A dolgozatot részletes felhasznált irodalomjegyzék zárja.

3) Tézisek

A fenti célokkal párhuzamosan megfogalmaztam azokat a hipotéziseket, melyek a kutatás fő motivációját adták. A hipotézisek szorosan kötődnek a disszertáció egyes fejezeteihez, a velük kapcsolatosan levont konklúzióimat az összefoglaló fejezetben tárgyalom. A hipotézisek tehát a következők:

1. A Likert-skálás lekérdezések segítségével nyert változók esetén a módszerválasztás különös jelentőséggel bír.
2. Amennyiben rendelkezünk külső információval a sokasági eloszlásról, az hatékonyan alkalmazható az előzetesen kalkulált szükséges mintaelemszám csökkentésére Likert-skálás kérdések esetén is.
3. Adott kívánt hibahatár és megbízhatósági szint esetén meghatározható a szükséges mintaelemszám várható értéke.
4. Az előzetes információk és egy esetleges előzetes mintavétel adatainak összesítése a bayesi keretrendszerben probléma nélkül megoldható. A bayesi módszertannal kiszámított szükséges mintaelemszám értékek a klasszikus statisztikai módszerekkel számítható értékekkel összhangot mutatnak.
5. Az előzetes információk bizonytalansága és az előzetes minta mintavételi hibája figyelembe vehető a bayesi gondolatvilágban, ami egyértelműen előnyt jelent a klasszikus megközelítéshez képest.

4) Eredmények

Doktori disszertációmban a mintavétel egyik fontos részterületével, a mintaelemszám tervezésével foglalkoztam. Az első fejezetben a bevezetés mellett bemutattam a kutatási irányokat és a dolgozat szerkezetét. A második fejezetben a mérési skálák rendszerét, illetve az ehhez kapcsolódó tudományos vitát tekintettem át azzal a céllal, hogy jobban megértsem a statisztikai műveletek alkalmazhatóságának feltételeit Likert-skálás lekérdezések segítségével nyert változók esetén. A harmadik fejezetben a mintaelemszám tervezés hagyományos, aránybecslésre épülő módszerének bemutatása mellett kidolgoztam a Likert-skálák esetére alkalmazható képleteket, eljárásokat. A fejezet konklúziójaként azt találtam, hogy az általam javasolt mintaelemszám tervezés mellett szükség szerint előzetes mintavételt kell végrehajtani. A két információforrás összefogása legegyszerűbben a bayesi keretek között oldható meg, melynek legfontosabb fogalmairól és eljárásairól rövid összefoglalót adtam. A komplex, mindkét információforrást figyelembe vevő bayesi eredmények leírása adja a dolgozat ötödik fejezetét.

Az alábbiakban a bevezetőben már taglalt hipotézisekre adott válaszaimat foglalom össze röviden.

1. **A Likert-skálás lekérdezések segítségével nyert változók esetén a módszerválasztás különös jelentőséggel bír.**

A mérési skálák elmélete (Stevens, 1946, 1955) és az azzal kapcsolatos tudományos vita (Michell, 1986) rávilágít arra, hogy a komoly irodalmi csatározás ellenére továbbra is több vélemény, ha úgy tetszik iskola létezik. A klasszikus nominális-ordinális-intervallum-arány skálarendszer hiányosságait többen jelezték, az egyik – témakörünk szempontjából jelentős – kritika az, hogy adott változók bizonyos esetekben nehezen sorolhatók be a fenti kategóriákba (Velleman-Wilkinson, 1993). A Likert-skála kapcsán felvetődik a kérdés, hogy ordinális, vagy intervallum skála erősségű eredményeket kapunk-e felhasználásával. Mindet azért fontos, mert a leggyakrabban alkalmazott statisztikai módszertanok csak intervallum, vagy arány skálán értelmezhetőek. A Likert-skálás alkalmazások esetén a leggyakrabban praktikus szempontok érvényesülnek, törekedni kell arra, hogy a változóértékek valóban ekvidisztansok legyenek. Amennyiben ez a lekérdezés formájával (megfelelő kérdésfeltevés, kategórianevek) nem biztosítható, úgy ordinális skálán mért változóként kell kezelnünk az eredményeinket.

2. **Amennyiben rendelkezünk külső információval a sokasági eloszlásról, az hatékonyan alkalmazható az előzetesen kalkulált szükséges mintaelemszám csökkentésére Likert-skálás kérdések esetén is.**

A hagyományosan alkalmazott (Rappai- Pintér, 2007) – aránybecslés hibahatárából kiinduló – mintaelemszám tervezés esetén is lehetőség van külső információk felhasználására, ekkor a sokasági arányról élhetünk feltételezéssel. Ezzel analóg módon alkalmazható külső információ abban az esetben, ha várható értékre vonatkozó becslést kívánunk végrehajtani. Ebben az esetben a sokasági varianciáról kell tudással rendelkeznünk. Mivel ez a gyakorlatban ritkán fordul elő, azt a megközelítést alkalmaztam, hogy nem a varianciáról, hanem az azt implikáló eloszlásról van sejtésünk. Amennyiben kevés kimenetelű a változónk, úgy definiálhatók olyan tipikus eloszlások, melyek esetén a variancia könnyen meghatározható. A varianciák segítségével számított mintaelemszámok azt mutatják, hogy a szükséges mintaelemszám akár tízszeres is lehet a különböző eloszlások között. Amennyiben tehát rendelkezünk információval, annak tervezésbe való beépítése jelentős megtakarításokat jelenthet a mintaelemszám tekintetében. Az alábbi táblázat a szükséges mintaelemszámokat mutatja különböző hibahatárok és az előre definiált tipikus eloszlások mentén.

1. táblázat: Szükséges mintaelemszámok várható értéke néhány Likert-skála esetén, előre adott hibahatárok mellett, $1-\alpha=0,955$

$\Delta^{(5)}$	Extrém két-móduszú	Fordított normális	Fordított piramis	Egyenletes	Piramis	Kvázi normális	Extrém egy-móduszú
0,02	40 000	28 298	24 444	20 000	13 333	7 553	4 161
0,04	10 000	7 074	6 111	5 000	3 333	1 888	1 040
0,1	1 600	1 132	978	800	533	302	166
0,2	400	283	244	200	133	76	42

A táblázatban foglalt eredmények alapján a rendelkezésre álló információk beépítése szükségesnek és hasznosnak ítélhető.

3. Adott kívánt hibahatár és megbízhatósági szint esetén meghatározható a szükséges mintaelemszám várható értéke.

Abban az esetben, ha nincs előzetes elképzelésünk a válaszlehetőségek közötti megoszlásról, azt feltételezzük, hogy minden lehetőségnek azonos a valószínűsége, úgy a mintaelemszámok várható értékére (és eloszlására) vagyunk kíváncsiak. Egy konkrét eset bemutatása után általános esetben is sikerül meghatároznom a keresett várható értékre vonatkozó – meglepően egyszerű – képletet.

A varianciára vonatkozó várható érték

$$E\left(\sigma_{n,k}^2\right)=\frac{k(k-1)(n-1)}{12n}$$

módon számítható. Amennyiben a várható értéket a korrigált varianciára írjuk fel, abban az esetben a képlet tovább egyszerűsödik, így a szükséges mintaelemszám várható értéke:

$$E(n)=\frac{k(k-1)}{3\Delta^2}$$

Ekkor a várható érték néhány Likert-skála és hibahatár esetén:

2. táblázat: Szükséges mintaelemszámok ötfokozatú Likert-skála, relatív hibahatár és különféle eloszlás-típusok esetén, $1-\alpha=0{,}955$

Δ	$k=5$	$k=7$	$k=9$	$k=10$
0,005	266 667	560 000	960 000	1 200 000
0,010	66 667	140 000	240 000	300 000
0,050	2 667	5 600	9 600	12 000
0,100	667	1 400	2 400	3 000

4. **Az előzetes információk és egy esetleges előzetes mintavétel adatainak összesítése a bayesi keretrendszerben probléma nélkül megoldható. A bayesi módszertannal kiszámított szükséges mintaelemszám értékek a klasszikus statisztikai módszerekkel számítható értékekkel összhangot mutatnak.**

A bayesi megközelítés (Gelman et al., 2004) egyrészt prior, másrészt mintabeli információk meglétét feltételezi. A klasszikus eredményekkel való összehasonlítás úgy lehetséges, hogy a priorba beépítjük az arányról, vagy az eloszlás formájáról rendelkezésünkre álló adatokat, majd ezt az eloszlást tekintjük egyben poszteriornak is. A két megközelítés különbsége, hogy a bayesi esetben egy konkrét érték helyett egy eloszlás az eredmény, melynek várható értéke azonban a két vizsgált esetben közelítően megegyezik. A bayesi megközelítés egyszerűségét az adja, hogy a Bayes-tétel alkalmazásával gyakorlatilag tetszőleges probléma felírható. A dolgozatban bemutatom, hogy milyen módon található a mintavétel problémájához konjugált prior, majd különböző paraméterezések mellett milyen módon nyerhető ki a poszterior információ. Nem informatív prior alkalmazásával és adatok hozzáadása nélkül az eredményeink a klasszikus várható érték meghatározásához hasonló értéket mutatnak (v. ö. 2. táblázat, $k=5$ oszlop).

3. táblázat: Szükséges mintaelemszámok várható értéke ötfokozatú Likert-skála esetén, előre adott hibahatárok mellett, bayesi közelítésben, $1-\alpha=0{,}955$

Δ	$E(n)$
0,005	266 616,5
0,010	66 669,5
0,050	2 667,0
0,100	666,5

5. **Az előzetes információk bizonytalansága és a mintavételi hiba figyelembe vehető a bayesi gondolatvilágban, ami egyértelműen előnyt jelent a klasszikus megközelítéshez képest.**

Az előzetes információk a bayesi keretek között egy megfelelő eloszláscsaládba tartozó eloszlás paraméterein keresztül építhetők a modellbe. Az ehhez társuló, mintabeli adatokat leíró likelihood segítségével meghatározott poszterior szintén egy eloszlás, azaz nem csupán egy pontbecslés, definíciójánál fogva tartalmazza a variancia és ezzel együtt a szükséges mintaelemszám bizonytalanságát. A kétkimenetelű esetben a poszterior a konjugált prior esetén

$$p(\theta|y)=\frac{\frac{1}{\mathrm{B}(\underline{\alpha},\underline{\beta})}\theta^{\underline{\alpha}+k-1}(1-\theta)^{\underline{\beta}+n-k-1}}{\frac{\mathrm{B}(\underline{\alpha}+k,\underline{\beta}+n-k)}{\mathrm{B}(\underline{\alpha},\underline{\beta})}}=\frac{1}{\mathrm{B}(\underline{\alpha}+k,\underline{\beta}+n-k)}\theta^{\underline{\alpha}+k-1}(1-\theta)^{\underline{\beta}+n-k-1},$$

Likert-skála esetén pedig

$$p(\boldsymbol{\theta}|y)\propto p(y|\boldsymbol{\theta})\,p(\boldsymbol{\theta})\propto\prod_{j=1}^{k}\theta_j^{n_j}\times\prod_{j=1}^{k}\theta_j^{\underline{\alpha}_j-1}=\prod_{j=1}^{k}\theta_j^{\underline{\alpha}_j+n_j-1}\sim Dir(\overline{\boldsymbol{\alpha}}).$$

Az ismeretlen (ténylegesen szükséges mintaelemszám) sokasági paramétert nagy valószínűséggel tartalmazó intervallum meghatározása bayesi értelemben magától értetődő.

5) További kutatási területek

További kutatási területként két irányt vázolok fel röviden. Egyrészt a becslés témaköréből kilépve a hipotézisellenőrzés esetén felhasználható eljárások kidolgozása a következő célom. Az ilyen jellegű vizsgálatok figyelembe veszik az elérni kívánt α szignifikancia szint mellett a másodfajú hiba elkövetésének valószínűségét, illetve a próba erejét is. További kutatást indukál az egyszerű véletlen mintavételtől eltérő, például rétegzett mintavétel esetén történő mintaelemszám meghatározás.

A másik fontos kutatási irány a bayesi statisztika és ökonometria alkalmazásainak megismerése és lehetőség szerinti fejlesztése. A klasszikus statisztikai eszköztár valamennyi módszere becsülhető bayesi értelemben is, sok esetben – megfelelő nem informatív priort használva – pedig elkerülhetőek identifikációs problémák is a segítségével. A bayesi ökonometria gazdag nemzetközi irodalma jó alapot nyújt ehhez a jövőbeli kutatási irányhoz.

6) Felhasznált hivatkozások

- Kehl Dániel, Rappai Gábor (2006): Mintaelemszám tervezése Likert-skálát alkalmazó lekérdezésekben, Statisztikai Szemle, 84, 9, 848-875
- Kehl Dániel (2007): A szükséges mintaelemszám várható értéke Likert-skálás lekérdezések esetén, Egy életpálya három dimenziója: tanulmánykötet Pintér József emlékére, 63-77, PTE-KTK
- Kehl Dániel (2011): Skálák és statisztikák: a méréselméletről és történetéről, Statisztikai Szemle, 89, 10-11, 1057-1080
- Kehl Dániel (2012): Monte Carlo módszerek a statisztikában, Statisztikai Szemle, 90, 5, megjelenés alatt
- Gelman, A., Carlin, J. B., Stern H. S., Rubin, D. B. (2004): Bayesian Data Analysis, Chapman & Hall/CRC
- Michell, J. (1986): Measurement Scales and Statistics: A Clash of Paradigms, Psychological Bulletin, 100, 3, 398-407
- Rappai Gábor, Pintér József (szerk.) (2007): Statisztika, PTE KTK Kiadó, Pécs
- Stevens, S. S. (1946): On the Theory of Scales of Measurement, Science, 103, 677-680
- Stevens, S. S. (1955): On the Averaging of Data, Science, 121, 113-116
- Velleman, P. F., Wilkinson, L. (1993): Nominal, Ordinal, Interval, and Ratio Typologies Are Misleading, The American Statistician, 47, 1, 65–72

7) A szerző eddigi tudományos teljesítménye

Az értekezéshez közvetlenül kapcsolódó publikációk listája:

- Kehl Dániel, Rappai Gábor (2006): Mintaelemszám tervezése Likert-skálát alkalmazó lekérdezésekben, **Statisztikai Szemle**, 84, 9, 848-875
- Kehl Dániel (2007): A szükséges mintaelemszám várható értéke Likert-skálás lekérdezések esetén, Egy életpálya három dimenziója: tanulmánykötet Pintér József emlékére, 63-77, PTE-KTK
- Balogh L. Dániel, Kehl Dániel (2009): Effects of Rounding on Descriptive Statistical Measures, Challenges for Analysis of the Economy, the Businesses, and Social Progress, Szeged
- Kehl Dániel (2009): On the History of Measurement, Challenges for Analysis of the Economy, the Businesses, and Social Progress, Szeged
- Kehl Dániel (2011): Skálák és statisztikák: a méréselméletről és történetéről, **Statisztikai Szemle**, 89, 10-11, 1057-1080
- Kehl Dániel (2012a): Szemelvények a Markov lánc Monte-Carlo módszerek történetéből (cikk ismertetés), **Statisztikai Szemle**, 90, 4, 352-354
- Kehl Dániel (2012b): Monte Carlo módszerek a statisztikában, **Statisztikai Szemle**, 90, 6, 521-543

Egyéb, lektorált folyóiratban megjelent publikációk:

- Kehl Dániel, Sipos Béla (2007): A gazdasági növekedés ciklikus változása az USA-ban, **Fejlesztés és Finanszírozás**, 5, 4, 3-12

- Kehl Dániel, Sipos Béla (2007): Évszázados trendek és hosszú ciklusok az Amerikai Egyesült Államokban, Kínában és a világgazdaságban, **Hitelintézeti Szemle**, 6, 4, 242-282
- Poór József, Karoliny Zsuzsa, Kehl Dániel, Farkas Ferenc (2009): Changes in Corporate Employment and HR Practice: An Analysis of 2007 Survey Data Based on the Life Cycle Model, **Zarzadzanie Zasobami Ludzkimi**, 1, 103-121
- Kehl Dániel, Sipos Béla (2009): A telítődési, a logisztikus és életgörbe alakú trendfüggvények becslése Excel parancsfájl segítségével, **Statisztikai Szemle**, 87, 4, 381-411
- Kehl Dániel, Sipos Béla (2009): Az egy főre jutó GDP változásának lehetséges hatásai, **Fejlesztés és Finanszírozás**, 7, 4, 43-52
- Kehl Dániel, Sipos Béla (2010): Regressziós modellek becslése és tesztelése Excel parancsfájl segítségével. (szoftverismertetés), **Statisztikai Szemle**, 88, 7-8, 833-855
- Kispál Z, Balogh D, Erdei O, Kehl Dániel, Juhász Z, Vastyán AM, Farkas A, Pintér AB, Vajda P (2011): Complications after bladder augmentation or substitution in children: a prospective study of 86 patients, **BJU International**, 108, 2, 282-289 (IF: 3,190)
- Komócsi A, Vorobcsuk A, Kehl Dániel, Aradi D (2012): Novel oral anticoagulants in patients with acute coronary syndromes: meta-analysis of randomized controlled trials, **Archives of Internal Medicine**, *megjelenés alatt* (IF: 10,639)

MTMT összefoglaló táblázat

Kehl Dániel munkásságának összefoglalása (2012.09.03.)

Közlemény típusok	Száma		Hivatkozások	
	Összesen	Részletezve	Független	Összes
I. Tudományos folyóiratcikk	12	---	---	---
teljes cikk, nemzetközi folyóiratban	---	2	1	2
teljes cikk, hazai idegen nyelvű folyóiratban	---	2	0	0
teljes cikk, hazai magyar nyelvű folyóiratban	---	8	15	17
II. Könyvek	0	---	---	---
a) Szakkönyv	0	---	---	---
Szakkönyv, idegen nyelvű	---	0	0	0
Szakkönyv, magyar nyelvű	---	0	0	0
b) Szerkesztett könyv	0	---	---	---
Szerkesztett könyv, idegen nyelvű	---	0	0	0
Szerkesztett könyv, magyar nyelvű	---	0	0	0
III. Könyvfejezet	1	---	---	---
Könyvfejezet, idegen nyelvű	---	0	0	0
Könyvfejezet, magyar nyelvű	---	1	0	0
IV. Proceedings+	3	---	0	0
Idegen nyelvű	---	1	0	0
Magyar nyelvű	---	2	1	1
Tudományos közlemények összesen (I-IV.)	**16**	**---**	**17**	**20**
Egyéb tudományos művek++	---	25	1	2

Összesített impakt faktor	13,829	---	---	---
Idézettség száma	---	---	18	22
Hirsch index	2	---	---	---

Oktatási művek				
Tankönyv	0	---	---	---
Idegen nyelvű	---	0	0	0
Magyar nyelvű	---	0	0	0
Tankönyvfejezet, idegen nyelvű	---	0	0	0
Tankönyvfejezet, magyar nyelvű	---	0	0	0

Ismeretterjesztő művek				
Könyvek	0	---	0	0
Folyóiratcikkek, könyvrészletek	0	---	0	0

Közéleti és nem besorolt művek				
Könyvek	0	---	0	0
Folyóiratcikkek, tanulmányok, könyvrészletek	0	---	0	0

Megjegyzések:

--- : Nem kitölthető mező

Teljes cikk: ismert lektoráltságú szakfolyóiratban megjelent cikk, az V. táblázatban használt specifikációval

+ Konferencia előadások folyóiratban vagy könyvben, absztraktok nélkül

++ Ide értve a nem-teljes folyóiratcikket és a nem ismert lektoráltságú folyóiratokban megjelent műveket