

A DIFERENCIÁLSZÁMÍTÁS ALKALMAZÁSA A DEMOGRÁFIAI STATISZTIKÁBAN

APPLICATION OF DIFFERENTIAL CALCULATION IN DEMOGRAPHIC STATISTICS

Tóth Attila ^{1*}, Csáky Antal ²,

¹Pedagógusképző Intézet, Közép-európai Tanulmányok Kara, Nyitra, Szlovákia

²Pedagógusképző Intézet, Közép-európai Tanulmányok Kara, Nyitra, Szlovákia

Kulcsszavak:

Trendvonal.
Demográfiai statisztika.

Keywords:

Trendline.
Demographic statistics.

Cikktörténet:

Beérkezett 2018. július 10.
Átdolgozva 2018. augusztus 15.
Elfogadva 2018. október 01.

Összefoglalás

A differenciálszámítás alkalmazható a demográfiai számításokban, a trendvonalak meghatározására alkalmas, segítségével a csekélyebb változások sebessége is kimutatható. A változások sebességének változása a szociális intézkedésekkel, megélhetési küszöbvel determinált, összehasonlítható a szomszédos országokban.

Abstract

Differential calculations can be applied in demographic calculations. Can not be used only for defining trine line, but also for measuring speed of the changes. The speed of the changes is determined by social dispositions and living thresholds; the data can be compared with data of neighboring countries.

1. Bevezetés

A munkában rámutatunk arra, hogyan alkalmazható a differenciálszámítás a demográfiai számításokban. A deriválás nemcsak a trendvonal levezetésére alkalmas, hanem segítségével a csekélyebb változások sebessége is kimutatható. A változások sebességének változása a szociális intézkedésekkel, megélhetési küszöbvel determinált, összehasonlítható a szomszédos országokban. A kutatás tartalmaz egy trendvonal „anomáliát” is. Végsősorban rávilágítunk arra, hogy a családpolitikára érdemes odafigyelni, hiszen a népesség mértékének változása a sebességgel nyomban kimutathatók.

2. Deriválás a trendvonalak levezetéséhez

A trendvonalak meghatározásánál nem a véletlenszerű, hanem a meghatározó jellegű, determinisztikus komponenssel számolnak. Miszerint

$$y_i = \eta_i + \varepsilon_i \quad (1).$$

* Tóth Attila. Tel.: +421376316210
E-mail cím: atoth2@ukf.sk

Ha a véletlenszerű komponens nincs olyan nagy befolyással az eseményeinkre, akkor a legjobban befolyásoló tényező a feltehetően lineáris iránynál a következő egyenlettel jellemezhető:

$$\eta_i = \beta_0 + \beta_1 x_i, \quad (2)$$

illetve az idősoroknál az x helyett a t idő található. A trend bizonyos szempontból a jellegzetes „közép”vonal. A matematikában és más műszaki ágazatokban általában a szórásnégyzettel dolgoznak, így a szórásnégyzetet számítják ki a valós pontok, és a hozzájuk tartozó trendvonal pontjainak a különbségére, amit reziduumoknak (becsült maradéktag) nevezünk. Ha lineáris trendvonalról van szó, akkor gyakorlatilag a b_0 az analitikus geometriában az eltolás, a b_1 pedig az iránytényező, tehát, hogy milyen irányban halad az idősor vizsgált értékei, ha nem változnak a körülmények, a trend. A szórásnégyzetet a következő egyenlettel írhatjuk fel:

$$S = S(b_0, b_1) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2 \quad (3),$$

ahol \hat{y}_i a trendvonal pontjait jelenti. Mikor kapjuk a legkisebb távolságot az összes ponttól? Ha alkalmazzuk a differenciálszámítás azon tulajdonságát, hogy a szélsőértéket a derivált 0-helyén kell keresni: eszerint

$$\frac{\partial S}{\partial b_0} = 0 \text{ illetve } \frac{\partial S}{\partial b_1} = 0. \quad (4)$$

Az eltolás és iránytényező szerinti deriválás segítségével kapjuk meg a legkisebb távolságokat a valós pontoktól, hiszen a második deriváció mindkét esetben negatív előjelű. A deriváció szabályainak megfelelően a következő egyenleteket kapjuk:

$$\frac{\partial S}{\partial b_0} = 2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i)(-1) = 0 \quad (5);$$

illetve

$$\frac{\partial S}{\partial b_1} = 2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i)(-x_i) = 0 \quad (6),$$

amelyekből két egyenletet kapunk, két ismeretlennel, ha adottak az n, y_i, x_i adatok:

$$\sum_{i=1}^n y_i = n b_0 + b_1 \sum_{i=1}^n x_i \quad (7)$$

$$\sum_{i=1}^n y_i x_i = b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n (x_i)^2 \quad (8)$$

ezen egyenletekben az y függését vizsgáljuk az x -től, illetve az idősorok esetében az x helyett t van, tehát időbeli függőségről van szó.

Differenciálszámítással van tehát levezetve a legkisebb szórásnégyzet, és eszerint számítják a lineáris trendvonalakat. A kvadratis trendvonalak számítására pedig már a görbületet is beszámítják, ahol három egyenletet kapunk, három ismeretlennel. Az idősorokra így ezt kapjuk:

$$\sum_{i=1}^n y_i = T b_0 + b_1 \sum_{i=1}^n t_i + b_2 \sum_{i=1}^n (t_i)^2 \quad (9)$$

$$\sum_{i=1}^n y_i t_i = b_0 \sum_{i=1}^n t_i + b_1 \sum_{i=1}^n (t_i)^2 + b_2 \sum_{i=1}^n (t_i)^3 \quad (10)$$

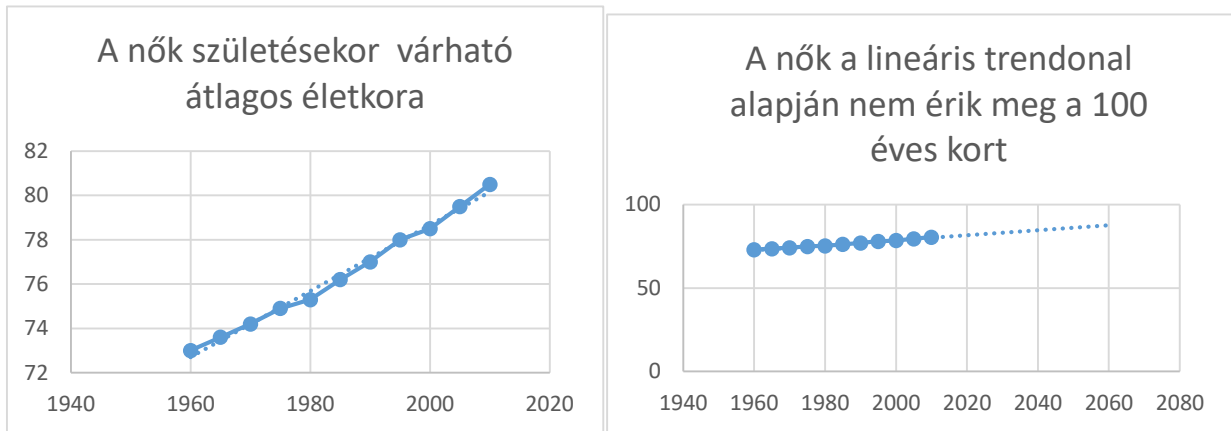
$$\sum_{i=1}^n y_i t_i^2 = b_0 \sum_{i=1}^n (t_i)^2 + b_1 \sum_{i=1}^n (t_i)^3 + b_2 \sum_{i=1}^n (t_i)^4 \quad (11)$$

A számítások akkor a legpontosabbak, ha a reziduumoktól való eltérés nulla, vagy a nullához közeli érték,

$$\sum_{i=1}^n \varepsilon = \sum_{i=1}^n y_i - \hat{y}_i = 0. \quad (12)$$

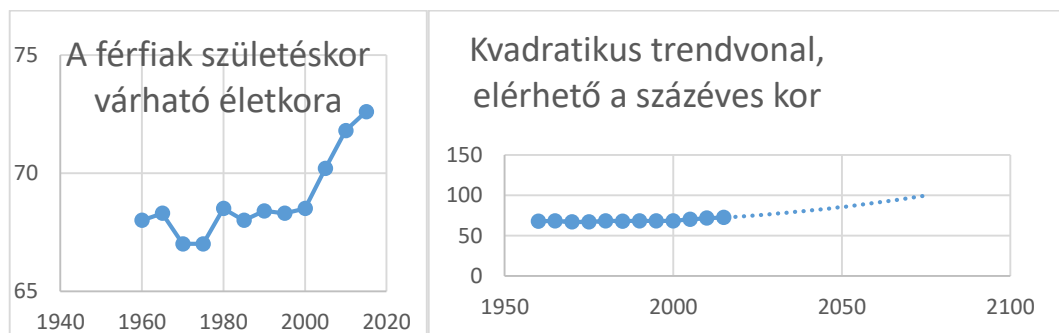
3. Az idősorok trendvonalainak az alkalmazása a demográfiai számításokban

Néhány gyakorlati példában bemutatjuk a trendvonalak klasszikus számítását a demográfiai statisztikában.



1. ábra. A szlovák statisztikai hivatal adatai alapján számított lineáris trendvonal és prognózis

Az 1. ábrán jól látható, hogy a nők a születéskor számított átlagos életkora, amely növekedő tendenciát mutat. A prognózis szerint azonban ez a növekedés lassú, 40-50 év múlva lépi túl a kilencvenes küszöböt.



2. ábra A szlovák statisztikai hivatal adatai alapján számított kvadratikus trendvonal és prognózis

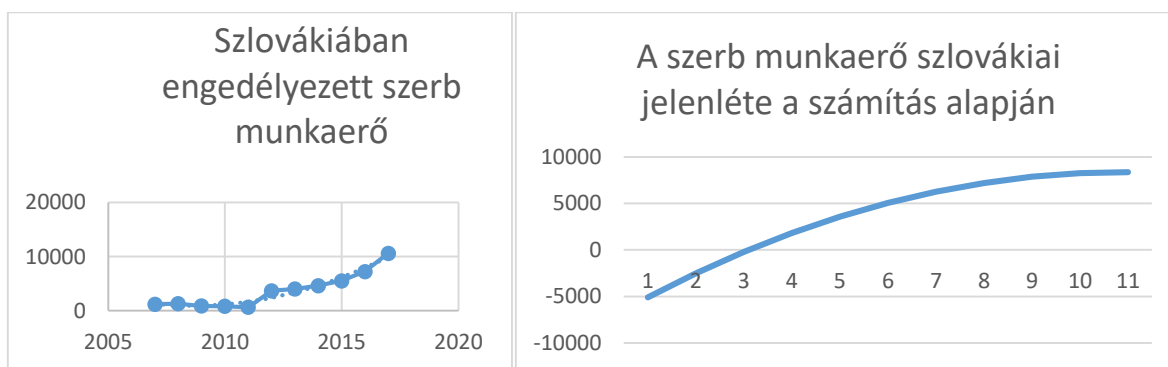
A második ábra szerint, míg mindeztáig a férfiak a statisztika alapján „belehaltak“ feleségük elvesztésébe, úgy néz ki a prognózis szerint, hogy megedződnek és így megelőzik a nők átlagos életkorát, és túlélnek őket. A második ábra bal oldalán jól látható, hogy parabolikus a trend, tehát nyilván kvadratikus trendvonal számításával próbálkozunk. A számítások pontossága az ötévenként megadott adathalmazra $\sum_{i=1}^{12} \varepsilon = 0$ a férfiak esetében, $\sum_{i=1}^{12} \varepsilon = 3,610^{-10}$ a nők esetében. Megkérdőjelezendő azonban az, hogy a trend valóban megmarad-e. Húsz évvel ezelőtt még fel sem tételeztünk volna a mára már elért magas életkort, az orvostudomány fejlődése azonban erre rácsáfolt.

4. A deriváció alkalmazása a változások mérésére

Az előzőekben rámutattunk a lineáris trend és kvadratikus trend különbözőségére, amit a demográfiai számítások átlagos életkorra vonatkozó adatok alapján dolgoztunk fel. A következőkben pedig bemutatjuk azokat a számításokat, amelyek nem várt eredményeket hoztak a kvadratikus egyenlet sor megoldásaként. Érdekes módon éppen a Szlovákiában hivatalosan tartózkodó vendégmunkások esetében a nyilvántartottak száma a 3. ábra szerint másodfokú polinom függvényre enged következtetni. A számítások viszont erre rácsáfoltak, hiszen egy fordított parabola jelenik meg, ahol az eltolás negatív előjelű, van kicsúcsosodás, majd hanyatlás is. Ha jól

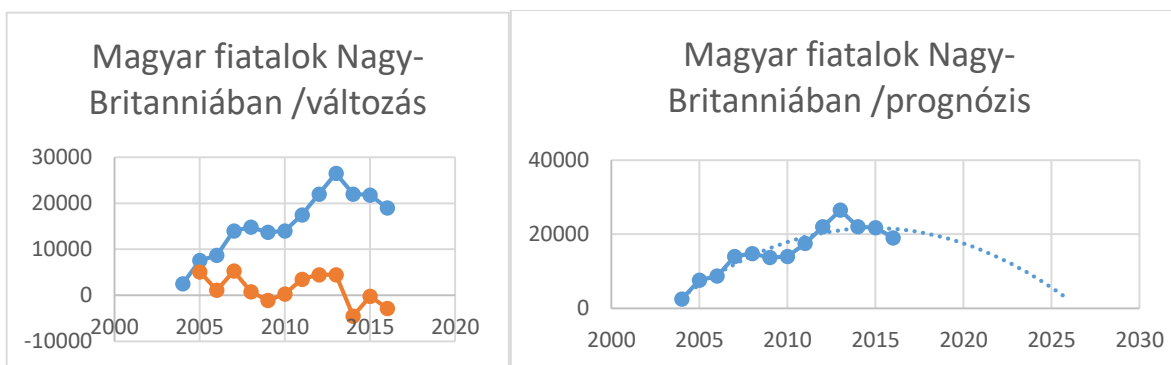
belegondolunk ennek az interpretációjába, akkor talán azt várnánk el, hogy a gazdasági célból útnak eredt fiatalok egyszer csak a saját hazájukban is találjanak majd megfelelő munkahelyet, ami a megélhetésüket is biztosítja. Az adatsor viszont, ha nem vesszük figyelembe a kezdeti értékeket növekedő típusú parabolát mutat, az elvártat. Illetve politikailag nem elvártat. A szerbiai fiatal férfiak a nyelvrokonság miatt is vonzóan találják Szlovákiát.

Hiszen az lenne a jó, ha mindenki a saját hazájában találna meg megélhetését, odahaza alapítana családot. Ugyanez megvizsgálható a magyar fiatalok esetében is, akik az angoloknál találtak sokkal jobban kifizetődő munkát. Pedig nem mindenki tud megfelelő szinten angolul. Megfigyelendő, hogy a magyar fiatalok jelenléte Nagy-Britanniában is fordított trendvonalat mutat, tehát meg kellene keresni a módját annak, hogy otthon maradjanak. Az okok keresését pedig az útnak indulásuk okát kellene megkeresni. Ezt pedig úgy találhatnánk meg, ha megnéznénk az akkori politikai intézkedéseket, munkaerő piaci helyzetet, és sok kis apró más befolyásoló tényezőt. Talán ezekben az esetekben nem lehet figyelmen kívül hagyni a véletlenszerű hibafaktort (ϵ_i). Ha alkalmazzuk a differenciálszámítást geometriailag is, akkor a baloldali görbe derivációjával kapjuk a változás mértékét, ami gyakorlatilag évenkénti változás. A kivándorlás „sebességének” a görbéjéből pontosan kiolvasható mikor volt pontosan növekedő, illetve mikor volt csökkenő tendenciájú. Tehát a trendvonalak is a differenciálszámítás eredményeként alkalmazhatóak a demográfiai statisztikában, de a változás mértéke (numerikusan számítva a statisztikában lehet bázis index is, amelyet a kívánt évhez hasonlítunk pontos időtartamokban) alkalmazva a görbére rögtönzött sebességet mutat.



3. ábra A szlovák statisztikai hivatal adatai alapján számított kvadratikus trendvonal és prognózis

A harmadik ábrán megfigyelhető a különbség, míg a program automatikus válasza a kért másodfokú polinómfüggvényre felfelé ívelő parabolát mutat, ugyanakkor a numerikus számítás eredményeként a jobboldali ábra jön elő. Érdekes az is, hogy az első néhány év elhagyásával a klasszikus számítás is felfelé ívelő parabolát mutat. Vizsgáltuk a parabolikus átbillenést, azaz, hogy mikor kezd a három egyenletrendszer megoldása nem lefelé, hanem felfelé ívelővé válni. A klasszikus számítás ugyanis a baloldali ábrát úgy érzékeli, mint egy nagy fordított parabola.

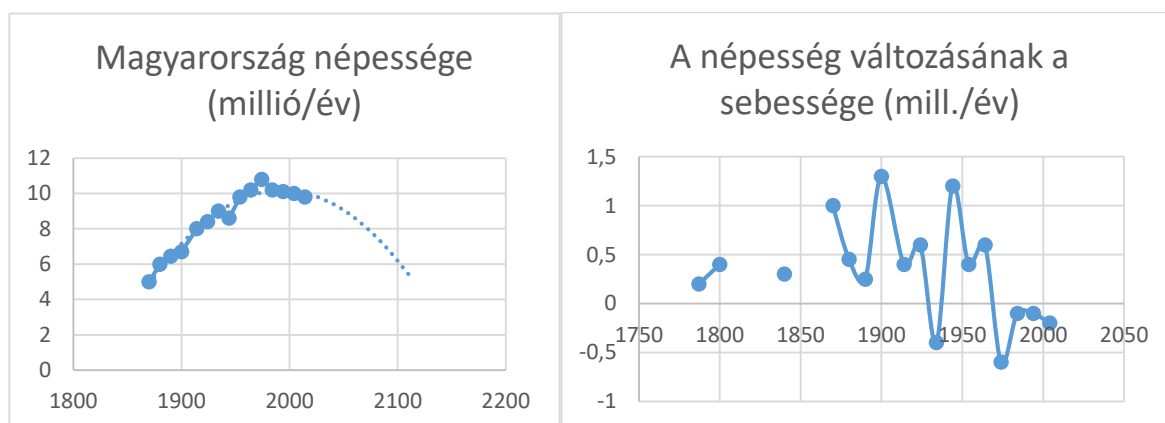


4. ábra A ONS 13.09.2017 Portfolio statisztikai adatai alapján feldolgozott változás és prognózis

Érdekes módon a magyar fiatalok angliai munkavállalása is ilyen jelenség, hiszen ha megvizsgáljuk a 4. ábra baloldalát, akkor az utolsó 4 év elhagyásával felfelé ívelő parabolikus összefüggést sejtethetnénk. Ha ebbe belegondolunk, a haza számára nem elfogadható, hogy gyakorlatilag a végtelenbe szaporodik a jobb megélhetésért kivándorlók száma. Nyilván az ország kilépése az EU-ból valószínűleg bizonyítani fogja a jobboldali ábrán látható prognózist. Érdekes megfigyelni a változás mértékét is a baloldali ábrán, ahol az alsó vonal éppen a változás mértéke néhány ezres nagyságrendben eléggé nagyoknak tűnik.

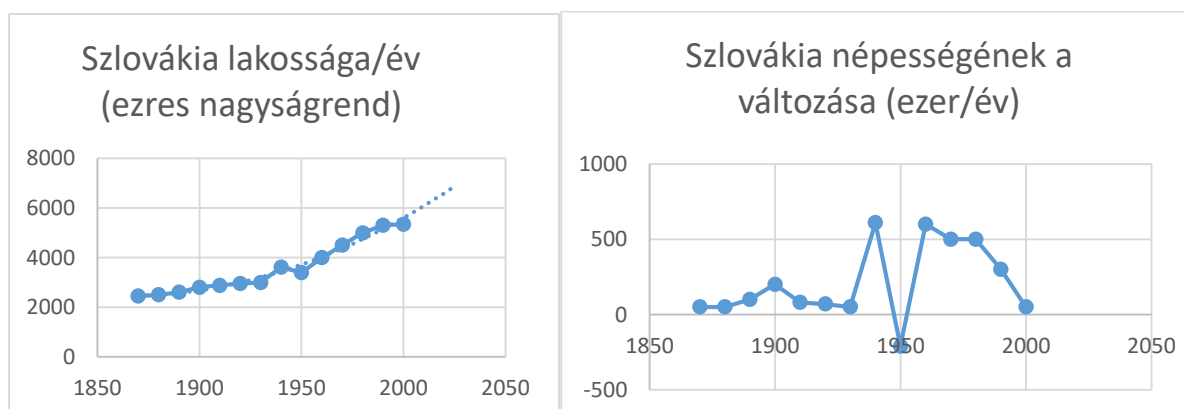
5. Össznépességi szám adatok

Ha már érintőlegesen foglalkoztunk a kivándorlás, bevándorlás illetve várható életkor problémakörrel vessünk pillantást az össznépességre. A V4 családkongresszuson publikált irodalom szerint a Magyarország jelenlegi területén a lélekszám a következő ábra szerint változik.



5. ábra KSH a századok statisztikája adatai alapján feldolgozott prognózis és változás

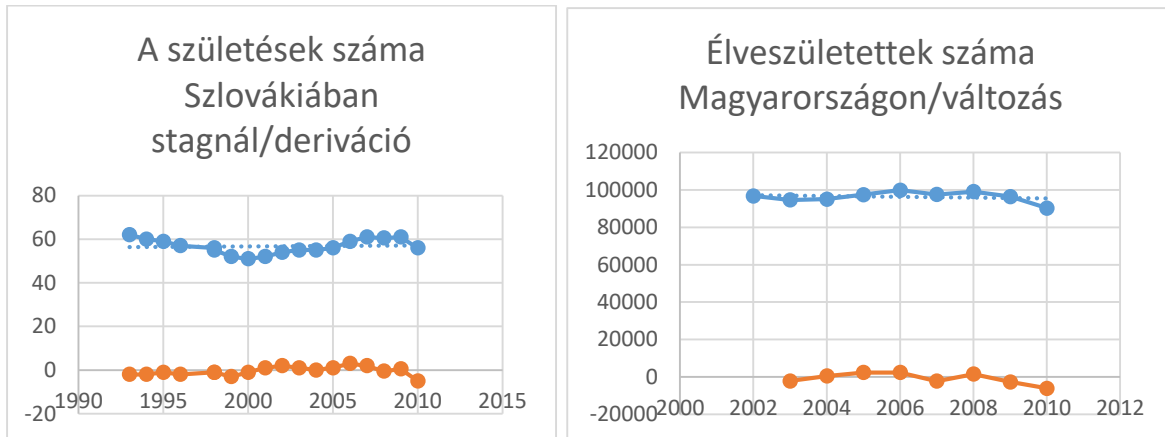
Az adatokból sajnos csökkenő tendenciájú parabola mutatkozik trendként, nyilvánvaló, hogy mindent meg kell tenni azért, hogy az 1985 óta évi 20 000 fővel ne fogyatkozzék a magyar hon. A változás mértéke a sebesség nagyon jól kiolvasható az egyes időszakokban az 5. ábra jobb felén. A történészeink elgondolkodhatnak az egyes időszakok különbségein



6. ábra Karásek Vladimír: Vývoj osídlenia Slovenska, Geografia statisztikai adatai alapján feldolgozott prognózis és változás

Az egyes időszakokban a két országban jól összehasonlítható a változás mértéke, az egyszerű derivált görbe segítségével. Nyilvánvalóan megerősíthető az ún. baby-boom, ami a háborúk után lenni szokott (még a legutolsó szerb-horvát háború után is kimutatható). A szlovákiai trendvonal biztatóan felfelé irányul népességi szempontból, de a változás mértéke itt is lefelé irányuló

tendenciát mutat. A hetvenes évek Husák gyermekiből a demokrácia üres bölcsői lettek. Az utolsó egy-két évtized gyermeáldását szemlélve annyi megállapítható, hogy stagnál mindkét országban.



7. ábra A PPT sűčasny populačny vjvoj v SR, es Csaladbarat fordulat kiadvany statisztikai adatai alapjan feldolgozva

A 7. ábrán összehasonlítható Magyarország és Szlovákia élveszületetteinek a száma, nyilván a szlovákiai összlakosság feleakkora nagyságú (10/5 millió).

6. Zárógondolat

A demográfiai statisztikában alkalmazott deriváció segítségével rámutattunk arra, hogy a deriváció segítségével meghatározott trendvonalak pontjai extrémén kicsi közelítéssel meghatározhatóak. A számításokban kapott reziduumok összege 10^{-10} nagyságrendű közelítést mutat a nullához, vagy pontosan 0 . A parabolák felfelé, vagy lefelé ívelő mivoltában valószínűleg más befolyásoló tényezőket is figyelembe kell venni. Az egyes kormányok családtámogatásához meg kell tehát vizsgálni az összes befolyásoló tényezőt, és azok befolyásolási mértékét. A deriváció másik alkalmazása a változás mértéke (ami lehet évenkénti, vagy 5 – 10 évenkénti rendszerességű), ami kimutathatja az egyes intézkedések hatékonyságát.

Irodalomjegyzék

- [1] ONS 13.09.2017 Portfolio
- [2] PPT sűčasny populačny vjvoj v SR, Infostat 2011
- [3] Csaladbarat fordulat, Emberi Erőforrások Minisztériuma, 2017
- [4] Karásek Vladimír: Vjvoj osidlenia Slovenska, Geografia
- [5] Origo, Népeségnövekedés, 2010.01.04
- [6] Bick, Alexander. 2016. «The Quantitative Role of Child Care for Female Labor Force Participation and Fertility.» Journal of the European Economic Association, im Erscheinen
- [7] Infostat, Prognóza strednej dĺžky žiovta, 2002
- [8] Majerová Silvia, Aká bude populácia Slovenska o 30 rokov,