

BUDAPESTI CORVINUS EGYETEM
EÖTVÖS LORÁND TUDOMÁNYEGYETEM

Antal Edina

HALANDÓSÁGI MODELLEK ÖSSZEHASONLÍTÁSA ÉS ALKALMAZÁSA

MSc Szakdolgozat

Biztosítási és pénzügyi matematika MSc
Aktuárius specializáció

Témavezető:

Dr. Kovács Erzsébet

Operációkutatás és Aktuáriustudományok Tanszék



Budapest, 2018

Köszönetnyilvánítás

Köszönöm témavezetőmnek, Dr. Kovács Erzsébetnek, hogy mindig szakított rám időt. Türelemmel és hozzáértéssel segítette a munkámat, hogy szakdolgozatom szakmailag és formailag is helytálljon.

Köszönöm barátaimnak, hogy egyetemi éveim alatt támogattak és minden lehetséges módon segítettek.

Végül, de nem utolsó sorban köszönetet kell mondanom családomnak, hisz nélkülük nem jutottam volna el ideig.

Tartalomjegyzék

Bevezetés	5
1. Halandósági modellek elmélete	7
1.1. Halandósági tábla előrejelzése	7
1.2. Lee Carter modell	7
1.2.1. Interpretáció	7
1.2.2. Paraméterbecslés	10
1.2.3. $\hat{\kappa}_t$ kiigazítása	13
1.2.4. Előrejelzés	13
1.3. Cairns-Blake-Dowd modell	14
1.3.1. GAPC modellcsalád	14
1.3.2. Cairns-Blake-Dowd modell felírása	16
1.3.3. Paraméterbecslés	16
1.3.4. Előrejelzés	17
1.4. Összehasonlítás	18
2. Magyarország regionális szintű halandóságának elemzése és modellezése	20
2.1. Adatok	22
2.2. Referencia megyék halandósági adatainak vizsgálata	23
2.3. Referencia országok meghatározása	26
2.4. Referenciaország halandóságának előrejelzése	28
2.5. Projekció és megyei előrejelzés	32
2.5.1. Projekció	32
2.5.2. Eredmények	34
3. Következtetések és továbbfejlesztési lehetőségek	39

4. Függelék	43
4.1. Forráskód	43
4.1.1. Magyarország halandósági rátáinak LC és CBD modellel történő előrejelzésnek forráskódja	43
4.1.2. Országok és megyék halandóságának összehasonlítása	44
4.1.3. Horvátország halandósági rátáinak előrejelzésének forráskódja	54
4.1.4. Megyei és horvát női halandósági ráták vizsgálata	57

Bevezetés

Minden embert foglalkoztat a kérdés, hogy meddig fog élni. Természetesen nincs megfelelő válasz egyének szintjén, de a kérdés az aktuáriusok számára is lényeges egy-egy állományra tekintve. A nyugdíjrendszerek és biztosítások, főleg életbiztosítások, a halandósági rátákon alapulnak, melyek megmutatják, hogy mekkora valószínűséggel halunk meg a különböző években, adott korban. A populációkra meg tudjuk adni ezeket a rátákat a halottak számát és a kitettséget ismerve. Ehhez természetesen ismernünk kell ezt a két adatot, ami legfeljebb csak a jelen pillanatig tudunk konkrétan megmondani, tehát csak a múltra és a jelenre tudjuk meghatározni ezt a rátát. Ezzel szemben az aktuáriusok számára az is fontos, hogy a jövőben várhatóan hogyan alakul a halandóság, hisz állandóan változik. Az utóbbi két évszázadban folyamatosan javult. Ez az orvostudomány és közegészség fejlődésével, és az életmód javulásával magyarázható. Az aktuáriusok a kérdés megválaszolása érdekében, azaz a halandóság előrejelzéséhez halandósági modelleket hoztak létre, melyek dinamikus, jövőbeli halandósági táblákat generálnak. Ezek a modellek a múlt adataiból jeleznek előre, akár több év távlatába is. A legtöbb modell a korábbi adatokból megfigyelt trend segítségével jelez előre, az általam vizsgált két modell is így működik.

A modellek, melyeket a szakdolgozatom részletesebben bemutat, a Lee-Carter (1.2 alfejezet) és a Cairns-Blake-Dowd (1.3 alfejezet) modell. Több mint két évtizede [1992] D.Lee és Lawrence Carter publikált egy cikket, melyben bemutattak egy halandósági modellt, ami nem csak életkortól, de a naptári évektől függő változókat is tartalmazott. A modell kiválóan leírta az 1900-1989 közötti Egyesült Államokbeli halandósági ráták alakulását. Ennek köszönhetően hamar elismerést nyert a szakmai körökben. A Lee-Carter modellt napjainkban is gyakran alkalmazzák, de az idő előrehaladtával publikálásra kerültek újabb modellek, illetve a tudomány fejlődésével több fejlesztési lehetőségre is fény derült.

A másik modell, amit használtam, számos évvel később [2006] publikálta A.J.G Cairns, D. Blake és K.Dowd. A Lee-Carter modellhez képest újításokat, új paramétert tartalma-

zott, egy újabb naptári évtől függő tényezőt, ezért kéttényezős modellek közé soroljuk.

Az egyes országokban készített halandósági táblákat és a modellek alkalmazását vizsgálva érdekes tényre lettem figyelmes. Az Egyesült Királyságban a halandósági táblákat az egyes tagállamokra is elkészítik, és az így kapott halandósági ráták lényegesen eltérnek. Felmerült a kérdés, hogy Magyarországon a régiók között is fennáll-e ez az eltérés. Az adatok előzetes vizsgálata után arra a következtetésre jutottam, hogy ha nem is olyan jelentősen, de egyértelműen különböznek az egyes megyék halandóságai.

Szakedolgozatom céljaul tűztem ki Magyarország regionális halandóságának előrejelzését, a megyei adatok hiányossága ellenére. Több szempontból is fontos lehet a pontos regionális előrejelzés a nyugdíjszolgáltatók, a biztosítók, az egészségügyi szolgáltatók és a megyei önkormányzatok számára is. Az előrejelzés és annak eredményei a 2. fejezetben olvashatóak.

1. fejezet

Halandósági modellek elmélete

1.1. Halandósági tábla előrejelzése

Az életbiztosítások világában a halandóság előrejelzése az egyik legfontosabb téma, mivel a halandósági ráták elengedhetetlenek a biztosítási díjak kalkulációjához. A Központi Statisztikai Hivatal minden évben elkészíti az országos halandósági táblát, mely fontos a díjkalkulációhoz, de sok más szempontot is figyelembe kell venni. Egyrészt, a biztosított halandóság lényegesen eltér a néphalandóságtól, hisz akik biztosítást kötnek általában jobban szituáltak és tudatosabban élnek, ami alacsonyabb halandóságot eredményez. Másrészt, az életbiztosítási termékek fejlesztése során nem csak az adott év halandóságát kell figyelembe venni, hanem kalkulálni kell a jövőbeni halandóság javulásával. A probléma megoldására többféle előrejelzési modellt is létrehoztak.

1.2. Lee Carter modell

1992-ben Ronald D.Lee és Lawrence R.Carter publikálta a mára nagyon népszerűvé vált Lee-Carter modellt. A modell segítségével előző évi halandósági táblákból dinamikus halandósági táblákat tudunk előrejelezni.

1.2.1. Interpretáció

Legyen q_{xt}^c az x évesek t . évben vett központi halandósági rátája, ahol $x \in \{1, 2, \dots, \omega\}$, $\omega \geq 2$ és $t \in \{1, 2, \dots, T\}$, $T \geq 2$, $1, 2, \dots, T$ egymást követő naptári évek. Mivel a szerzők központi halandóságot használnak a modellben, ezért a fejezetben csak központi halandósági ráták

szerepelnek, ezért a c indexet elhagyom.

A modellt a következő egyenlet írja le:

$$\ln q_{xt} = \alpha_x + \beta_x \kappa_t + \epsilon_{xt}, \quad (1.1)$$

A formulában két korévtől függő paraméter van, az α_x és a β_x . Egy olyan paramétert is tartalmaz, a κ_t -t, mely az időbeli fejlődést prezentálja. κ_t -t mortalitási indexnek hívjuk.

Természetesen ez a modell se tudja tökéletesen leírni a halandósági ráták fejlődését, ezért a hatásokat, melyeket a modell nem magyaráz meg, az ϵ_{xt} hibatag tartalmazza. Feltételezzük a hibatagról, hogy eloszlása $\sim N(0, \sigma^2)$ minden x -re és t -re. Ekkor minden x és t esetén a logaritmusos halandósági ráták függetlenek egymástól és eloszlásuk $\sim N(a_x + b_x k_t, \sigma^2)$ minden x -re és t -re. Probléma, hogy a paraméterezés nem egyértelmű, mivel

$$(\alpha'_x, \beta'_x, \kappa'_t) = (\alpha_x, k\beta_x, \frac{\kappa_t}{k}), \quad \forall k \neq 0 \text{ és } k \in \mathbb{R} \quad (1.2)$$

vagy

$$(\alpha''_x, \beta''_x, \kappa''_t) = (\alpha_x - k\beta_x, \beta_x, \kappa_t + k), \quad \forall k \in \mathbb{R}. \quad (1.3)$$

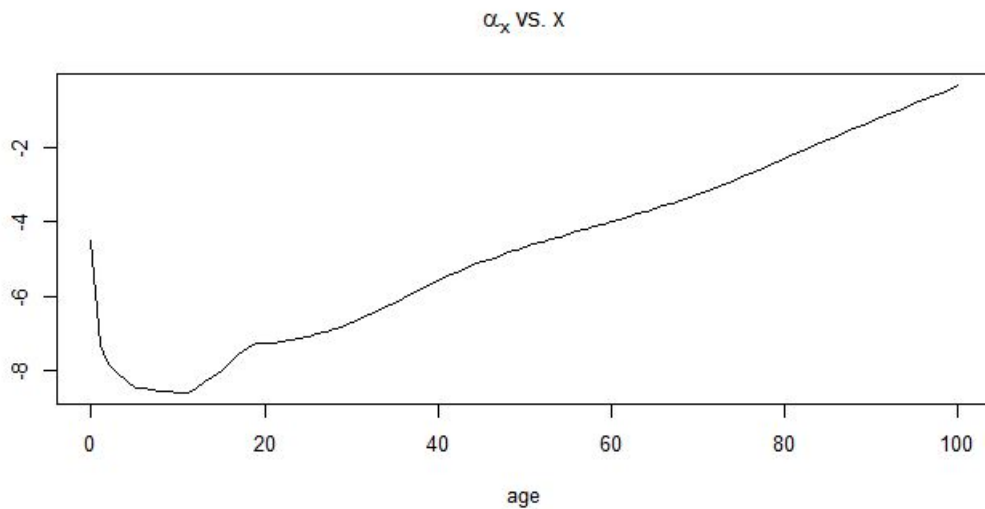
Az invariáns transzformációk elkerülése érdekében, a modell kidolgozói két megszorítást javasoltak:

$$\sum_{x=1}^{\omega} \beta_x^2 = 1 \text{ és } \sum_{t=1}^T \kappa_t = 0. \quad (1.4)$$

Felmerül a kérdés, hogy hogyan lehet értelmezni a paramétereket, melyik milyen hatást ír le.

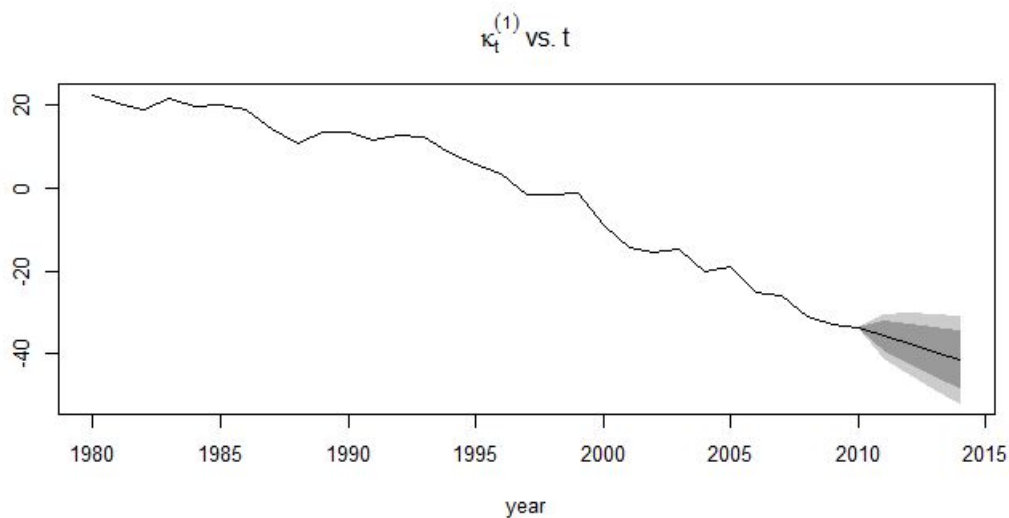
α_x : Átlagos log-mortalitás, azaz az egyes korévekhez tartozó log-mortalitási ráták átlaga a megfigyelt években. Megmutatja, hogy egyes korévben, a megfigyelt időszakban átlagosan mekkora valószínűséggel haltak meg (természetesen exponenciális transzformáció után). Tehát az $A = [\alpha_x]$ vektor az átlagos halandóság alakulását mutatja meg. Ahogy az 1.1 ábrán is látszik, az α_x -nek olyan alakja van, mint a hagyományos halandósági rátáknak.

κ_t : Mortalitási index, az egyetlen naptári évtől függő változó. Megmutatja, hogy az évek során a halandóság milyen mértékben fejlődik. Ez általában egy csökkenő sorozat, bár háborús időben egy rövid ideig növekedhet. Az 1.2 ábrán Magyarország mortalitási indexét láthatjuk 1980-2010 között. Az ábrát a *mortality.org* adatait felhasználva készítettem. Ez



1.1. ábra. α_x 1980-2010-es magyar adatok alapján
mortality.org adatai alapján saját számítások

is csökkenő tendenciájú, és az ábrán a következő négy év előrejelzése is látható, ami szintén további csökkenő tendenciát mutat, tehát az előrejelzés alapján a mortalitás egyre javul.



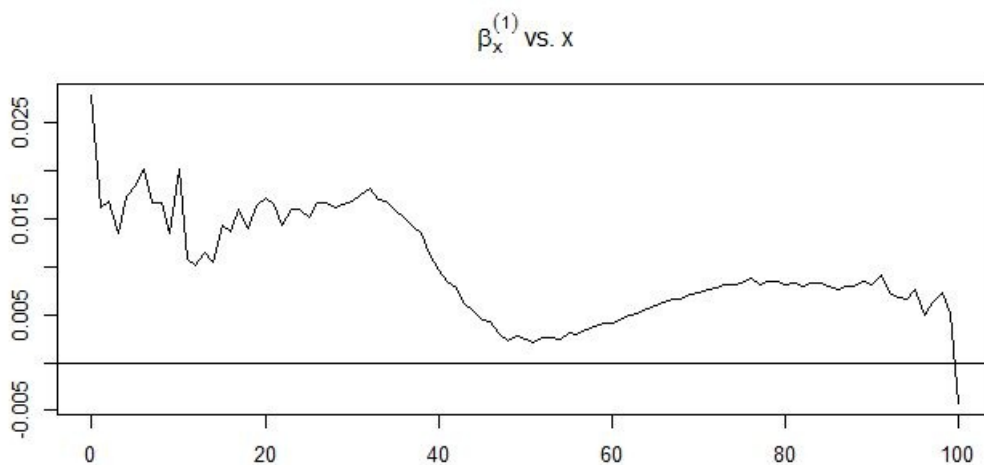
1.2. ábra. κ_t 1980-2010-es magyar adatok alapján.
mortality.org adatai alapján saját számítások

β_x : A mortalitási indexre való érzékenységet prezentálja a log-mortalitási rátáknak, vagyis azt fejezi ki, hogy ha a mortalitási index egy egységgel változik, akkor a log-

mortalitási ráta β_x egységgel változik, azaz

$$\mathbb{E}(\ln q_{x,t+1} - \ln q_{x,t}) = \beta_x(\kappa_{t+1} - \kappa_t). \quad (1.5)$$

Az 1.2 ábrán láthattuk, hogy κ_t -nek csökkenő trendje van, illetve az 1.3 ábrán észrevehetjük, hogy végig pozitív a β , így a kettő szorzata pozitív, tehát a mortalitási ráta csökken. Az is látható, hogy 50-es korévben befolyásolja legkevésbé a κ a log-mortalitási rátát, mivel itt majdnem 0 a β .



1.3. ábra. β_x 1980-2010-es magyar adatok alapján
mortality.org adatai alapján saját számítások

1.2.2. Paraméterbecslés

Legyen θ az α , β és κ paraméterek halmaza. A paraméterek becsléséhez maximum likelihood módszert szokás használni. Mivel tudjuk, hogy normális eloszlású az $\ln q_{xt}$, ezért

$$l(\theta) = -\frac{1}{2}\omega T \ln(2\pi) - \omega T \ln \sigma - \frac{1}{2}\sigma \sum_{x=1}^{\omega} \sum_{t=1}^T (\ln q_{xt} - \alpha_x - \beta_x \kappa_t)^2 \quad (1.6)$$

egyenletet kell maximalizálni paraméterenként, ahol ω a legmagasabb korév. Az egyenletet a θ paramétervektorra úgy tudjuk maximalizálni, tetszőleges $\sigma > 0$ esetén, ha $\sum_{x=1}^{\omega} \sum_{t=1}^T (\ln q_{xt} - \alpha_x - \beta_x \kappa_t)^2$ -t minimalizáljuk. Legyen $m(\theta) = \sum_{x=1}^{\omega} \sum_{t=1}^T (\ln q_{xt} - \alpha_x - \beta_x \kappa_t)^2$ a továbbiakban.

A $\frac{\partial m}{\partial \alpha_x} = \frac{\partial m}{\partial \beta_x} = \frac{\partial m}{\partial \kappa_t} = 0$ feltételek mellett a paraméterekre a következő egyenleteket kapjuk:

$$\frac{\partial m}{\partial \alpha_x} = -2 \sum_{t=1}^T (\ln q_{xt} - \alpha_x - \beta_x \kappa_t) \quad (1.7)$$

$$\frac{\partial m}{\partial \beta_x} = -2 \sum_{t=1}^T (\ln q_{xt} - \alpha_x - \beta_x \kappa_t) \kappa_t \quad (1.8)$$

$$\frac{\partial m}{\partial \kappa_t} = -2 \sum_{x=1}^T (\ln q_{xt} - \alpha_x - \beta_x \kappa_t) \beta_x \quad (1.9)$$

$\hat{\alpha}_x$ becslése

Az 1.7 egyenletből következik, hogy

$$\sum_t \alpha_x = \sum_t \ln q_{xt} - \sum_t \beta_x \kappa_t. \quad (1.10)$$

A $\sum_t \kappa_t = 0$ feltétel segítségével a következőt kapjuk:

$$\sum_t \alpha_x = \sum_t \ln q_{xt}, \quad (1.11)$$

így az α_x becslése:

$$\hat{\alpha}_x = \frac{1}{T} \sum_t \ln q_{xt}, \quad (1.12)$$

$\hat{\beta}_x$ és $\hat{\kappa}_t$ becslése

Legyen M mátrix a következő:

$$M_{xt} = \ln q_{xt} - \alpha_x = \beta_x \kappa_t, \quad (1.13)$$

ahol x a korosztály, t az adott év.

Az M_{xt} segítségével a következő egyenleteket tudjuk felírni, melyekben szerepel β_x és κ_t :

$$\sum_t M_{xt} \kappa_t = \sum_t (\beta_x \kappa_t) \kappa_t = \beta_x \sum_t \kappa_t^2 \quad (1.14)$$

$$\sum_x M_{xt} \beta_x = \sum_x (\beta_x \kappa_t) \beta_x = \kappa_t \sum_x \beta_x^2 = \kappa_t \quad (1.15)$$

Az így kapott két egyenletet a következő módon is felírhatjuk:

$$M\kappa = b\beta \quad (1.16)$$

és

$$M^T\beta = \kappa, \quad (1.17)$$

ahol $b = \sum_t \kappa_t^2$.

A 1.17 egyenletet szorozzuk balról M -el, így kapjuk a következőt:

$$MM^T\beta = M\kappa = b\beta, \quad (1.18)$$

azaz a korábbi feltételek mellett β az MM^T mátrix b sajátértékéhez tartozó sajátvektor.

$$m(\theta) = \sum_{t,x} (\ln q_{xt} - \hat{\alpha}_x - \beta_x \kappa_t)^2 \quad (1.19)$$

$$= \sum_{t,x} (M_{xt} - \beta_x \kappa_t)^2 \quad (1.20)$$

$$= \sum_{t,x} M_{xt}^2 + \sum_{t,x} (\beta_x \kappa_t)^2 - 2 \sum_t \kappa_t \left(\sum_x \beta_x M_{xt} \right) \quad (1.21)$$

$$= \sum_{t,x} M_{xt}^2 + 1 \times b - 2 \sum_t \kappa_t^2 \quad (1.22)$$

$$= \sum_{t,x} M_{xt}^2 + b - 2b \quad (1.23)$$

$$= \sum_{t,x} M_{xt}^2 - b. \quad (1.24)$$

Így $l(\theta)$ maximalizálásához $m(\theta)$ -t minimalizáljuk, amihez b -t maximalizálni kell.

Ezt követően megkapjuk β_x és κ_t becslését:

$\hat{\beta}$ = a MM^T mátrix maximális sajátértékéhez tartozó sajátvektor,

$\hat{\kappa} = M^T\hat{\beta}$.

A feladat megoldásról Vékás [2016]-ban részletesebben lehet olvasni.

1.2.3. $\hat{\kappa}_t$ kiigazítása

A szerzők egy kiigazítást javasolnak a κ_t -re, mivel a becslés minden megfigyelést azonos súllyal vett figyelembe, viszont az egyes naptári években és korcsoportokban a kitettség és halálesetek száma nagyságrendileg eltérhet.

A következő egyenletrendszert javasolták a mortalitási index kiigazítására az egyes naptári években:

$$\hat{\kappa}_t^{adj} = \{\hat{\kappa}_t \in \mathbb{R}^T : \sum_x D_{xt} = \sum_x E_{xt} e^{\hat{\alpha}_x - \hat{\beta}_x \hat{\kappa}_t}\}, \quad (1.25)$$

ahol D_{xt} az t . évben x évesen meghaltak száma, E_{xt} a t . évben az x évesek kitettsége.

1.2.4. Előrejelzés

Az előrejelzés készítéséhez az egyetlen időtől függő paramétert kell előrejelezni, a $\hat{\kappa}_t^{adj}$ -t. Az 1.2.4. fejezetben csak kiigazított mortalitási indexek fognak szerepelni, ezért elhagyom az adj indexet. A szerzők arra a következtetésre jutottak, hogy az ARIMA(0,1,0) folyamat a legalkalmasabb a $\hat{\kappa}_t$ alakulásának leírásához.

$$\hat{\kappa}_t = \hat{\kappa}_{t-1} + s + \varepsilon_t \quad t \in \{2, 3, \dots, T\}, \quad (1.26)$$

ahol $\hat{\kappa}_1$ ismert; s az eltolás, úgynevezett drift paraméter; $\varepsilon_t \sim N(0, \sigma^2)$ eloszlású, független hibtagok, ahol $\sigma^2 > 0$.

Az 1.26 egyenletből és az ε -ok eloszlásából következik, hogy a κ_t független növekményű, és

$$\hat{\kappa}_t - \hat{\kappa}_{t-1} \sim N(s, \sigma^2) \quad t \in \{2, 3, \dots, T\}, \quad (1.27)$$

Az s drift paraméterre, a hiba varianciájára és a trend sztenderd hibájára a következő egyenletek megoldását kapjuk maximum likelihood módszer segítségével:

$$\hat{s} = \frac{1}{T-1} \sum_{t=2}^T (\hat{\kappa}_t - \hat{\kappa}_{t-1}) = \frac{\hat{\kappa}_T - \hat{\kappa}_1}{T-1}. \quad (1.28)$$

$$\hat{\sigma}^2 = \frac{1}{T-1} \sum_{k=1}^{T-1} (\hat{\kappa}_{k+1} - \hat{\kappa}_k - \hat{s})^2, \quad (1.29)$$

$$\text{var}(\hat{s}) = \frac{\hat{\sigma}}{T-1}. \quad (1.30)$$

A $(T + h)$ -adik időpontra az előrejelzést iterációval tudjuk kiszámolni:

$$\mathbb{E}(\hat{\kappa}_{T+h}) = \mathbb{E}(\hat{\kappa}_T + \sum_{k=1}^h (\hat{\kappa}_{T+k} - \hat{\kappa}_{T+k-1})) = \hat{\kappa}_T + \sum_{k=1}^h \mathbb{E}(\hat{\kappa}_{T+k} - \hat{\kappa}_{T+k-1}) = \hat{\kappa}_T + h\hat{s}, h \in \{1, 2, \dots\}. \quad (1.31)$$

Visszahelyettesítve a modellbe, megkapjuk a log-mortalitást a $(T + h)$. időpontban.

$$\ln \hat{q}_{x,T+h} = \hat{\alpha}_x + \hat{\beta}_x(\hat{\kappa}_T + h\hat{s}) \quad x \in \{2, 3, \dots, \omega\}, h \in \{1, 2, \dots\}. \quad (1.32)$$

1.3. Cairns-Blake-Dowd modell

A Cairns-Blake-Dowd modellről Andrew J. G. Cairns, David Blake és Kevin Dowd írt a 2006-ban megjelent *A two-factor model for stochastic mortality with parameter uncertainty* című cikkükben. A modellt évekkel a Lee-Carter modell után publikálták, ezért eddig nem vált annyira ismertté, mint az LC modell. Másrészt, mivel a Lee-Carter modellt több mint húsz éve publikálták, több kritikus is javításokat tanácsolt. 2016-ban Villegas és szerzőtársai javasoltak egy általánosított korcsoport-időszak-kohorsz (Generalized Age-Period-Cohort, azaz GAPC) modelleszaládot, melynek a Cairns-Blake-Dowd modell is tagja, így a modellt a GAPC modelleszaládon keresztül mutatom be.

1.3.1. GAPC modelleszalád

Továbbra is q_{xt}^c , illetve q_{xt}^0 az x évesek t . évben vett központi, illetve kezdeti halandósági rátája, ahol $x \in \{1, 2, \dots, \omega\}$, $\omega \geq 2$ és $t \in \{1, 2, \dots, T\}$, $T \geq 2$, $1, 2, \dots, T$ egymást követő naptári évek.

D_{xt} a t . évben x évesen meghaltak száma, E_{xt} a t . évben az x évesek kitettsége, ám a Lee-Carter modellel szemben az egyik újítása a GACP modelleszaládnak, hogy a halottak számára, mint valószínűségi változó tekint. A következőekben legyen \tilde{D}_{xt} a halottak számának valószínűségi változója.

A kitettségi adatok típusától függően Poisson vagy binomiális eloszlású a \tilde{D}_{xt} , azaz

$$\tilde{D}_{xt} \sim \text{Poisson}(E_{xt}^c q_{xt}^c) \quad \text{vagy} \quad \tilde{D}_{xt} \sim \text{Binom}(E_{xt}^0 q_{xt}^0),$$

és függetlenek $\forall x$ -re és $\forall t$ -re.

Konkrétan a Cairns-Blake-Dowd modellt tekintve, a modell kezdeti kitettséget használ, így a \tilde{D}_{xt} binomiális eloszlású. Tehát a Lee-Carter és a Cairns-Blake-Dowd modell már a használt adatoknál eltér egymástól, hisz a Lee-Carter modell központi kitettséget alkalmaz.

Az általánosított modellt a következő egyenlet írja le:

$$g(q_{xt}) = \eta_{xt}, \quad (1.33)$$

ahol η_{xt} az adott modell szimmetrikus komponense, g a pozitív valós számokon értelmezett, folytonosan differenciálható, monoton növényő linkfüggvény.

Az általános szimmetrikus komponens :

$$\eta_{xt} = \alpha_x + \sum_{i=1}^N \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x}, \quad (1.34)$$

ahol csak úgy mint a Lee-Carter modell esetében α_x és β_x korosztálytól függő, κ_t naptári évtől függő változó, viszont ebben a modellben megjelent a γ , mely a kohorszhatástól függő változó. A kohorszhatás az azonos évben születettek, a kohorsz tagjainak, alul vagy felül becslése, mely hiba egész élettartamukon végigkíséri a generációt.

A paramétereket a Lee-Carter modellhez hasonlóan interpretáljuk:

α_x : Korcsoport-hatás, azaz az adott korban, a megfigyelt időszakban átlagosan mekkora valószínűséggel haltak meg.

$\kappa_t^{(i)}$: Mortalitási index, a halandóság fejlődését mutatja meg az egyes naptári években. N db időszorral modellezzik.

$\beta_x^{(i)}$: Életkorfüggő érzékenység, amely $i = 0$ esetben a kohorszhatásra való érzékenységet, $i \in \{1 \dots N\}$ az adott mortalitási indexre való érzékenységet prezentálja. Amennyiben a mortalitási index vagy a kohorszhatás egy egységgel nő, logaritmikus linkfüggvény esetén m_{xt}^c , logit linkfüggvény esetén $\frac{m_{xt}^0}{1-m_{xt}^0}$ értéke $e^{b_x^{(i)}}$ -szeresére változik.

γ_{t-x} : Kohorszhatás, $t - x$ időszakban született kohorsz halandóságának átlagtól való eltérését prezentálja.

1.3.2. Cairns-Blake-Dowd modell felírása

A GAPC modell felírását követően rátérhetünk a Cairns-Blake-Dowd modell konkrét paramétereire.

Linkfüggvény:

A Cairns-Blake-Dowd modell kezdeti kitettséget használ, ez azonban a linkfüggvényt is meghatározza, mivel központi kitettség esetén Poisson eloszlású \tilde{D}_{xt} -nél a linkfüggvény a logaritmus függvény, kezdeti kitettség esetén binomiális eloszlású \tilde{D}_{xt} -nél logit függvény a linkfüggvény. (Blake és Hunt [2014])

Szimmetrikus komponens:

A Cairns-Blake-Dowd modell szimmetrikus komponense a következő:

$$\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)}, \quad (1.35)$$

tehát láthatjuk, hogy Cairns-Blake-Dowd modell esetében az $\alpha_x = 0$, $N = 2$, $\beta_x^{(1)} = 1$, $\beta_x^{(2)} = (x - \bar{x})$ és $\beta_x^{(0)} = 0$, illetve az $\bar{x} = \left(\frac{1+\omega}{2}\right)$.

Tehát a Cairns-Blake-Dowd modell:

$$\ln\left(\frac{q_{xt}}{1 - q_{xt}}\right) = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)}, \quad (1.36)$$

vagyis a modellben nem szerepel se korcsoport-hatás (α_x), se kohorszhatás (γ_{t-x}). A Cairns-Blake-Dowd modell esetén, a Lee-Carter modellel ellentétben, nincsenek megszorítások. (Vékás [2016])

1.3.3. Paraméterbecslés

A GAPC modellcsalád esetén a paraméterbecslés maximum likelihood elvvel végezhető el, ahol a Poisson, illetve a binomiális eloszlású \tilde{D}_{xt} -ek esetén különböző log-likelihood függvényt kell maximalizálni. Továbbra is θ az 1.2.2. fejezetben definiált paramétervektor.

Poisson eloszlás esetén a log-likelihood függvény:

$$l(\theta) = \sum_{x=1}^{\omega} \sum_{t=1}^T \mathbb{I}_{xt} (-E_{xt}^c q_{xt}^c + D_{xt} (\ln E_{xt}^c + \ln q_{xt}^c) - \ln(D_{xt}!)), \quad (1.37)$$

ahol \mathbb{I}_{xt} indikátor változó.

Binomiális eloszlás esetén, mely a Cairns-Blake-Dowd modell esetében releváns, a log-likelihood függvény:

$$l(\theta) = \sum_{x=1}^{\omega} \sum_{t=1}^T \mathbb{I}_{xt} \left(\ln \left(\frac{E_{xt}^0}{D_{xt}} \right) + D_{xt} \ln q_{xt}^0 + (E_{xt}^0 - D_{xt}) \ln(1 - q_{xt}^0) \right). \quad (1.38)$$

A feladatot több numerikus módszerrel meg lehet oldani, mint például Newton-módszerrel, vagy Általánosított lineáris modellel (GLM). (Bohuns és szerzőtársai [2002]); (Villegas és szerzőtársai [2016])

1.3.4. Előrejelzés

A mortalitási index előrejelzését, akárcsak a Lee-Carter modell esetében, véletlen bolyongással modellezzük, viszont a GAPC modellcsalád esetén a mortalitási index többdimenziós, pontosabban egy adott modell esetén N dimenziós, így többdimenziós véletlen bolyongást alkalmazunk. A Cairns-Blake-Dowd modellben két mortalitási index szerepel, így kétdimenziós véletlen bolyongással modellezzük. (Villegas és szerzőtársai [2016])

Legyen $\bar{\kappa}_t$ egy kétdimenziós vektor, mely a mortalitási indexeket foglalja magába, azaz

$$\bar{\kappa}_t = \begin{pmatrix} \kappa_t^{(1)} \\ \kappa_t^{(2)} \end{pmatrix}. \quad (1.39)$$

Ekkor a többdimenziós véletlen bolyongás:

$$\bar{\kappa}_t = \bar{\kappa}_{t-1} + s + \varepsilon_t, \quad (1.40)$$

ahol $\bar{\kappa}_1$ a kezdőérték vektor ismert, s továbbra is a trendparaméter vektor, ε_t független hibatagok, $\varepsilon_t \sim N(0, \Sigma^2)$ eloszlású, ahol Σ pozitív definit kovarianciamátrix.

A GAPC modellcsalád tagjainak előrejelzése során a kohorszhatást prezentáló γ_{t-x} -et is sztochasztikusan modellezni kell. Általában ARIMA(p,q,d) folyamattal szokták modellezni. Mivel a Cairns-Blake-Dowd modellben nem szerepel a kohorszhatás, így ennek a tagnak az előrejelzését nem ismertetem.

A mortalitási index független növekményű, többdimenziós normális eloszlású valószínűségi változó, s várható értékkel:

$$\bar{\kappa}_t - \bar{\kappa}_{t-1} \sim N(s, \Sigma^2) \quad t \in \{2, 3, \dots, T\}, \quad (1.41)$$

így a várhatóérték vektor és a kovarinaciámátrix maximum likelihood módszer általi becslései:

$$\hat{s} = \frac{1}{T-1} \sum_{t=2}^T (\bar{\kappa}_t - \bar{\kappa}_{t-1}) = \frac{1}{T-1} (\bar{\kappa}_T - \bar{\kappa}_1), \quad (1.42)$$

$$\hat{\Sigma} = \frac{1}{T-1} \sum_{t=2}^T (\bar{\kappa}_t - \bar{\kappa}_{t-1} - \hat{s})(\bar{\kappa}_t - \bar{\kappa}_{t-1} - \hat{s})^T. \quad (1.43)$$

Ezt követően fel tudjuk írni a $(T+h)$. időszakbeli mortalitási index becslését:

$$\mathbb{E}(\bar{\kappa}_{T+h}) = \mathbb{E}(\bar{\kappa}_T + \sum_{k=1}^h (\bar{\kappa}_{T+k} - \bar{\kappa}_{T+k-1})) = \bar{\kappa}_T + \sum_{k=1}^h \mathbb{E}(\bar{\kappa}_{T+k} - \bar{\kappa}_{T+k-1}) = \bar{\kappa}_T + h\hat{s}, h \in \{1, 2, \dots\}. \quad (1.44)$$

Így a jövőbeli, $(T+h)$. időszakban a halandósági ráta becslése a GAPC modelleszalád esetében (a kohorszhatás előrejelzését követően) :

$$\hat{q}_{x,T+h} = g^{-1}(\hat{\eta}_{x,T+h}) \quad (1.45)$$

Amennyiben konkrétan a Cairns-Blake-Dowd modellel akarunk előrejelezni, a halandósági ráták előrejelzése a következő:

$$\hat{q}_{x,T+h} = \frac{e^{(\bar{\kappa}_T^{(1)} + h\hat{s}^{(1)} + (x-\bar{x})(\bar{\kappa}_T^{(2)} + h\hat{s}^{(2)}))}}{e^{(\bar{\kappa}_T^{(1)} + h\hat{s}^{(1)} + (x-\bar{x})(\bar{\kappa}_T^{(2)} + h\hat{s}^{(2)}))} + 1}. \quad (1.46)$$

1.4. Összehasonlítás

A modellek megismerése után röviden összehasonlítom a két modellt.

- **Bemeneti adatok:** A két modell már input adatok szintjén eltér, mivel a Lee-Carter modell központi halandósági rátákat használ, ameddig a Cairns-Blake-Dowd modell kezdeti halálozás számot és kitettséget alkalmaz, ami praktikusabb, mivel a legtöbb adatbázisban a kezdeti adatok elérhetőek.
- **Halálozások száma:** A Lee-Carter modell a halottak számát egy determinisztikus változónak tekinti, miközben a Cairns-Blake-Dowd modellben D_{xt} egy $Binom(E_{xt}q_{xt})$ eloszlású valószínűség változó, és a halálozások száma a valószínűségi változó megvalósult értéke.

- **Linkfüggvény:** A linkfüggvény terén is eltérnek a modellek, mivel a Lee-Carter modellnek logaritmikus a linkfüggvénye, a Cairns-Blake-Dowd modell, pedig logit linkfüggvényt használ.
- **Korcsoport-hatás:** A Lee-Carter modell tartalmaz korcsoport-hatást reprezentáló változót, az α_x -et, mely az átlagos log-mortalitási ráta. Ezzel ellentétben a Cairns-Blake-Dowd modellben nincs változó, mely megmutatná a vizsgált időszakbeli átlagos halandóságot.
- **Mortalitási indexek száma:** A halandóság időbeli fejlődését a Lee-Carter modellben egyetlen változó reprezentálja, ezzel szemben Cairns-Blake-Dowd modell esetén két mortalitási index is szerepel a modellben, mely által több számolást igényel a modell használata.
- **Mortalitási indexek előrejelzése:** Mind a két modell ARIMA(0,1,0) folyamatot alkalmaz a mortalitási index előrejelzéséhez.
- **Kohorszhatás:** Egyik modell se tartalmazza a γ változót, ezáltal egyik modell se veszi figyelembe a kohorszhatást. Ez egy lehetséges továbbfejlesztési lehetősége a két modellnek.

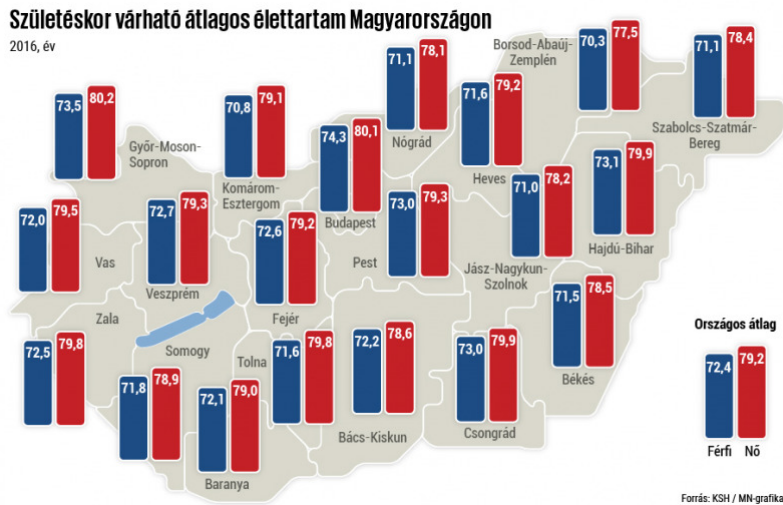
2. fejezet

Magyarország regionális szintű halandóságának elemzése és modellezése

Az országok halandóságának alakulásáról gyakran publikálnak az adott országokban. Az Egyesült Királyságban is többen foglalkoztak ezzel a területtel. Ezekben a tanulmányokban nem csak az egész szigetország előrejelzésével találkozunk, hanem régiókra bontva, külön Anglia és Wales, Skócia és Észak-Írország jövőbeli halandóságát is előre jelzik, mivel a három régió halandósága lényegesen eltér egymástól. A bontás segítségével az egyes területek halandóságát sokkal pontosabban meg lehet határozni. Anglia és Wales halandósága kifejezetten jobb, mint Észak-Írországé és Skóciáé. Skócia jelentősen rosszabb, mint a másik két régió. A szakirodalmak ezt a jelentős eltérést a társadalmi-gazdasági különbségeknek, az eltérő oktatási rendszernek és más társadalmi hatásoknak tulajdonítják. (Maccheroni és Nocito [2016])

Ezek után helytálló gondolat, hogy talán Magyarországon is érvényesülnek azok a társadalmi hatások, mely eltérő halandóságot eredményeznek Magyarország különböző területein.

Ezt az ötletet a születéskor várható élettartam vizsgálata is igazolta. (2.1 ábra) A nők és férfiak között is jelentős különbség van, több mint öt év, de nem csak nemenkénti bontásban, hanem megyénként is felfedezhetünk különbségeket. Magyarország nyugati részétől kelet felé haladva egyre alacsonyabb a várható élettartam. Természetesen ez nem törvénytörő, hisz Hajdu-Bihar megye várható élettartama nők és férfiak terén is jobb



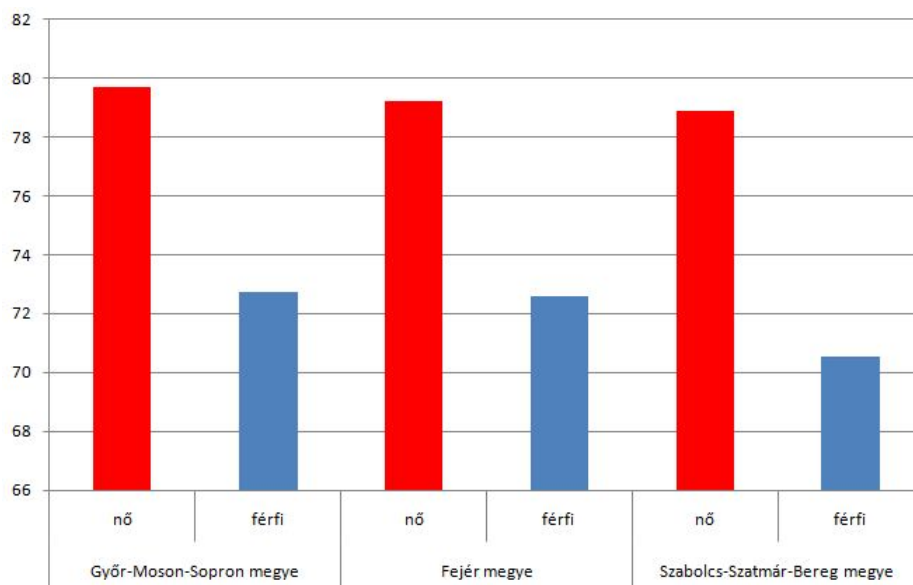
2.1. ábra. Országos születéskor várható élettartam

<https://mno.hu/eletmod>

Vas megyénél. Ez magyarázható azzal, hogy Hajdu-Bihar megyében található országunk második legnagyobb és legfejlettebb városa, Debrecen, ezáltal gazdaságilag előrehaladottabb, mint a környező megyék. Ezzel ellentétben, ha régióként nézzük, jól látszik, hogy Észak-Magyarország és Észak-Alföld (Nógrád, Borsod, Heves, Szabolcs-Szatmár-Bereg, Jász-Nagykun-Szolnok megye) születéskor várható élettartama jelentősen alacsonyabb, mint a Nyugat-Dunántúli régióban (Győr-Moson-Sopron, Vas, Zala megye).

Magyarország fejlettségének három szintjét határoztam meg: fejlett, átlagos és elmaradott. A fejlett megye az ország nyugati területeiről került kiválasztásra. Győr-Moson-Sopron megyével reprezentáltam ezt a csoportot. Átlagosnak olyan megyét választottam, mely egy közepes fejlettségi szintet képvisel, így a második megye Fejér megye lett az ország középső régiójából. Az elmaradottabb megyét pedig a keleti határhoz közeli megyék közül válogattam, így esett a döntés Szabolcs-Szatmár-Bereg megyére.

A választott megyékben 2014-es adatok alapján megnéztem a várható élettartamot. Továbbra is jelentős különbséggel talákoztam férfi-nő bontásban, de megyei szinten is szemmel látható különbség volt a várható élettartamok között. (2.2. ábra) A Szabolcs-Szatmár-Bereg megyei férfiak két évvel korábban halnak meg, mint győri társaik. Nők esetében kicsit kisebb a különbség, de még így is egy évvel kevesebb a nyugati nők várható élettartama, mint a szabolcsi nőké. Fejér megye választása valóban megfelel az átlagos régiók reprezentálásához, hisz mind nők, mind a férfiak esetében a másik két megye közé



2.2. ábra. A három megyében a születéskor várható élettartam 2014-es adatok alapján KSH adatai alapján saját számítás

esik a várható élettartam szempontjából.

A három megye nevének további rövidítése:

1. Szabolcs-Szatmár-Bereg megye: Szabolcs megye
2. Fejér megye: Fejér megye
3. Győr-Moson-Sopron megye: Győr megye

2.1. Adatok

A megyei halandósági adatokért a Központi Statisztikai Hivatalhoz fordultam. A 2014-es, 2015-ös és 2016-os évre a megyei halandósági táblák férfi-nő koréves bontása helyett csak korcsoportos bontásban álltak rendelkezésemre. A Központi Statisztikai Hivatal 1971-es közleményében szereplő 1959/60-as év halandósági tábláját is felhasználtam (Pallós [1971]). Az országok halandósági adatai a <http://www.mortality.org> honlapról gyűjtöttem ki. Harminckilenc ország több évnyi adatából dolgoztam, melyek szintén szerepeltek öt éves korcsoportos bontásban.

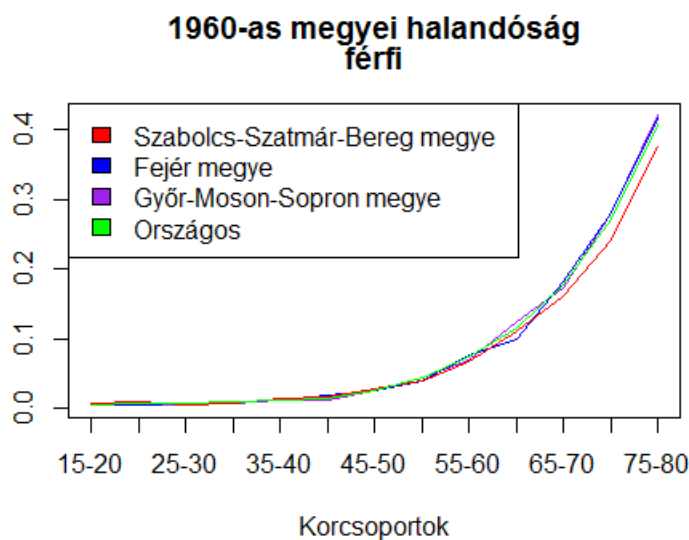
A kapott adatok esetén nem az egész halandósági táblát vettem figyelembe, csak a 15-80 év közötti korosztályokat vizsgáltam. A választásom azért esett erre a tartományra, mert 15-80 éves korig voltak megfelelő adatok, melyek rendelkezésemre álltak. Az országok halandósági adatait 2000-tól 2015-ig terjedő időszakon vizsgáltam.

2.2. Referencia megyék halandósági adatainak vizsgálata

Magyarország és a megyék halandóságának összehasonlítása során láthatóvá vált, hogy régiók halandósága nagy mértékben eltér egymástól és a teljes Magyarországra vonatkozó halandóságtól is.

Először az 1960-as országos, illetve megyei halandósági rátákat hasonlítottam össze.

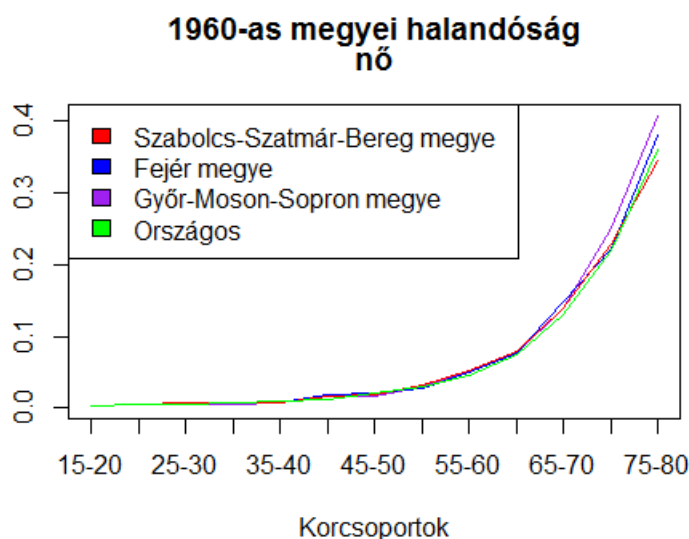
Az előzetes sejtésem szerint Győr megye az átlag alatti halandósággal rendelkezett volna, Fejér megye az átlaghoz közelivel, vagyis az országossal, és Szabolcs megye átlag felettiel. A sejtés nem igazolódott be 1960-ban se a férfi (2.3 ábra), se a női (2.4 ábra) halandósági rátákat nézve.



2.3. ábra. 1960-as megyei és magyarországi férfi halandósági ráták
KSH adatai alapján saját számítás

A férfi halandósági táblát tekintve a győri és az országos halandóság hasonló volt.

A 40-45, 55-60 és 65-70 éves korcsoportokban a győrieknek kisebb volt a halandósági rátája az országosnál, de a többi korévben minimális volt a különbség. A nők esetében a győri halandóság 65-70 éves korcsoportig inkább megegyezett, de 70 év felett jelentősen magasabb volt az országosnál.



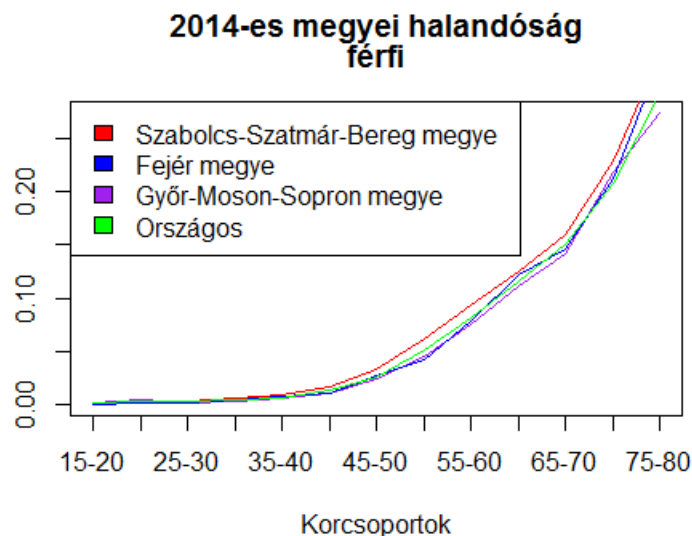
2.4. ábra. 1960-as megyei és magyarországi női halandósági ráták
KSH adatai alapján saját számítás

Fejér megye halandósága a férfiaknál 60-65 éves korcsoportnál az országos halandóság alá esett, de a továbbiakban végig felette volt. Nőknél sokkal jobban együtt mozgott a halandóság az országoséval, ám az idősebb korosztályban, 70-85 éves korban ismét magasabb volt, mint az országos.

Szabolcs megye okozta a legnagyobb meglepetést, mivel halandósága megegyezett, néhol meghaladta, de több esetben, főleg idősebb korban elmaradt az országos halandóságtól. Ez főleg férfiaknál 65 éves kortól figyelhető meg. 65 éves korig a szabolcsi ráták együtt mozognak az országossal, ám 65 éves kortól lényegesen jobbak. A nőknél ez nem ilyen szembeűnő. Itt csak a 75-80 éves korcsoportban jobb a halandóság az országoshoz képest. Napjaink sztereotípiái alapján nem voltam meggyőzve arról, hogy ezek a megfigyelések még most is helytállóak, ezért a 2014-es halandóságot is megvizsgáltam.

Ebben az esetben már beigazolódott az előzetes sejtésem.

A Győr megyei férfiak esetében már 45 éves kortól felfigyelhetünk a javulásra, igaz 75-80 éves korcsoportnál kicsit meghaladja az országos halandóságot, de nem nagy mértékben. (2.5



2.5. ábra. 2014-es megyei és magyarországi férfi halandósági ráták
KSH adatai alapján saját számítás

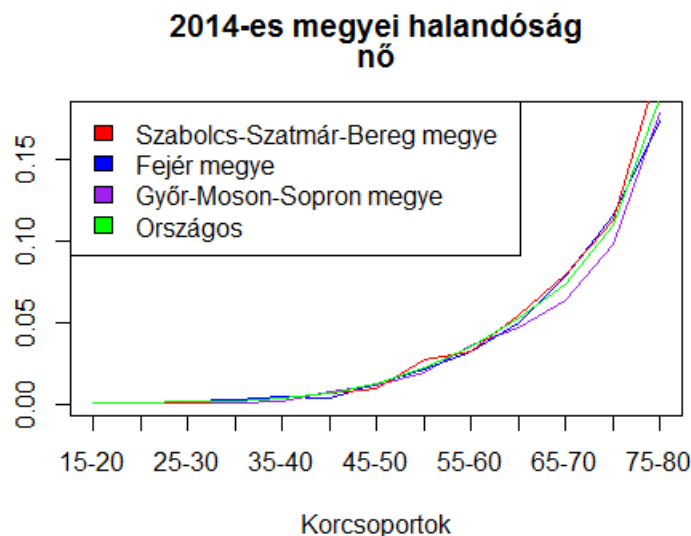
ábra) A nőknél Győr megye 60 éves korig együtt mozog a magyarországi halandósággal, majd jelentősen javul az idősebb korosztályban. (2.6 ábra)

Fejér megyei férfiak esetében az országoshoz közel van a halandóság, ám jelentősebb az eltérés az 50-55 és 60-65 éves korosztálynál. Az előbbinél jobb, utóbbinál rosszabb volt a halandóság, de ezt követően visszatér az országoshoz. A 75-80 éveseknél szintén rosszabb a Fejér megyei halandóság. (2.5 ábra) A nőknél is hasonlóak a megfigyeléseim. Végig közel azonos a halandóság az országoséval, kivéve 75-80 évesek esetén, ahol kicsit jobb az átlagnál. (2.6 ábra)

Szabolcs megye esetében a férfiak már 30 éves kortól kezdve jelentősen rosszabb halandósággal rendelkeznek, mint a többi megye. (2.5 ábra) A nőknél ez a különbség nem ilyen nagy, de itt is rosszabb a halandóság az országosnál 50-55 éves kortól kezdődően. 45-50 és 55-60 éves korcsoportban minimálisan jobb, de nem lényegesen. (2.6 ábra)

Az eltérő megfigyelések érdekes problémát vetnek fel. Milyen társadalmi-gazdasági változások következtek be 1960-2014 között Magyarországon? A szakdolgozatomban nem célja ezt a kérdést megválaszolni, de ez további érdekes kutatási téma lehet.

A két év halandóságának vizsgálatát követően egyértelművé vált, hogy valóban eltérés



2.6. ábra. 2014-es megyei és magyarországi női halandósági ráták
KSH adatai alapján saját számítás

van a megyék között, még ha nem is olyan jelentősen, mint az Egyesült Királyságban.

Ezek a megfigyelések vezettek az ötlethez, hogy teljes magyarországi előrejelzés helyett, régiók szintjén dolgozzak, a pontosabb előrejelzés érdekében.

Az ötletet sajnos nem lehetett egyszerűen a gyakorlatba ültetni, mivel a megyei adatok igen korlátozottan álltak rendelkezésemre. A továbbhaladás érdekében olyan halandósági idősort kellett keresni, mely hasonló a megyei adatokhoz, és megfelelő mennyiségű adat elérhető.

2.3. Referencia országok meghatározása

A *mortality.org* oldalán fellelhető harminckilenc országból kettőt ki kellett zárni, mivel Magyarországhoz nem hasonlítottam saját megyéit, illetve Norvégia adatai hiányosak voltak.

Először a lehetséges referenciaországok körének leszűkítésére a halandóságok közötti négyzetes hibát vizsgáltam meg.

A következő halandóságokat hasonlítottam a 37 országéhoz:

1. 2014-es Szabolcs-Szatmár-Bereg megyei női halandóság

2. 2014-es Fejér megyei női halandóság
3. 2014-es Győr-Moson-Sopron megyei női halandóság
4. 2014-es Szabolcs-Szatmár-Bereg megyei férfi halandóság
5. 2014-es Fejér megyei férfi halandóság
6. 2014-es Győr-Moson-Sopron megyei férfi halandóság

Megnéztem, hogy 2014-es megyei halandóság és az országok halandósága közötti négyzetes hiba melyik országban, melyik évben a legkisebb:

$$\sum_{15-20}^{75-80} (q_{megye} - q_{ország})^2 \quad (2.1)$$

Mind a női és férfi halandóság, mind a három megye ugyanahhoz az országhoz, Horvátország adataihoz közelített. Ez az eredmény várható volt. A társadalmi eltérések mellett sincs akkora különbség megyéink között, mint a környező országok és a megyéink halandósága között, hisz hiába különböző az egyes régiókban élő emberek életvitele, társadalmi szintje, de nem térnek el annyira egymástól, mint az országoké.

Egyszerű megfontolás alapján is várható volt, hogy egy szomszédos ország lesz a referenciaország, hisz a szomszédokhoz a leghasonlóbb az életvitelünk, a természeti adottságok, illetve a magyarok száma is jelentősebb, mint a távolabbi országokban. Abban az esetben, ha nem ilyen reális eredményt kaptam volna, további módszereket alkalmaztam volna a referenciaország kiválasztásához.

Egy újabb módszerrel pontosítani lehetett az illeszkedést. A referencia országhoz egy évet is választottam, mely évben a legkisebb volt a négyzetes hiba. Ezáltal a két idősort eltolva illesztettem egymáshoz. A továbbiakban csak eltolásként hivatkozok a műveletre. Ez az évszám már eltért a különböző megyék, különböző nemek esetében.

Mivel a halandóság folyamatosan javul, ezért várható volt, hogy az elmaradottabb ország egy korábbi évhez, a fejlettebb ország, pedig napjainkhoz fog illeszkedni. (2.1 táblázat) A várakozásokat mind Szabolcs megye, mind Győr megye teljesítette. Szabolcs megyei nők halandósága a horvátországi nők 2009-ben jegyzett halandóságához közelített, ameddig a győriek a 2013-as halandósághoz. Jól látszik, hogy Győr megye halandósága sokkal

megye	nő	féfi
Szabolcs-Szatmár-Bereg	2009	2006
Fejér	2014	2009
Győr-Moson-Sopron	2013	2014

2.1. táblázat. 2014-es megyei halandóságának horvátországi referencia évei
KSH és *mortality.org* adatai alapján saját számítás

jobban együtt mozgott a horváttal. Ez magyarázható azzal, hogy Győr-Moson-Sopron és Horvátország földrajzilag is sokkal közelebb van egymáshoz, mint Szabolcs-Szatmár-Bereg megye és a referencia ország, és gazdaságilag is sokkal hasonlóbbak, fejlettebbek. Fejér megye képviselte a középmezőnyt. A férfiakat nézve teljes mértékben a besorolásnak megfelelően viselkedett, hisz Szabolcs megyéhez egy korábbi, Fejér megyéhez egy későbbi, de mégis öt évvel korábbi év tartozott, mint Győr megyéhez. A nők esetében ez már nem mondható el. A nőknél Fejér megye és Győr megye szerepet cserélt, és az eredmények azt sugallták, hogy Fejér megye halandósága alacsonyabb Győr megye halandóságánál, pedig a korábbiakban láthattuk hogy ez nem teljesen igaz. (2.6 ábra) A magyarázat, hogy Győr megye női halandósága nem egyértelműen jobb Fejér megyéjénél. Vannak korcsoportok, ahol az átlagos megye a jobb. Másrészt, Horvátország 2014-es és 2015-ös halandósága kis mértékben tér el.

2.4. Referenciaország halandóságának előrejelzése

A megfelelő referenciaország és év kiválasztását követően elvégeztem a megyék előrejelzését. A pontosság érdekében a korábban említett két modellt, a Lee-Carter és Cairns-Blake-Dowd modellt használtam. Megnéztem, hogy a horvát adatok esetén az egyes nemeknél az egyes korévekben melyik modell illeszkedik jobban.

A *mortality.org* honlapon 2002-től 2015-ig szerepelnek a horvát halandósági táblák. Ahhoz, hogy megállapítsam melyik korcsoportra melyik modell illik jobban, a 2002-2010-es adatokból előrejeleztem 2011-2015-re Lee-Carter és Cairns-Blake-Dowd modellel is. Az előrejelzést a korosztályos adatokkal végeztem, hisz ezáltal pontosabban ki tudtam választani a modelleket.

Két információs kritériumot néztem meg. Az egyik az Akaike-féle információs krité-

rium, a másik Bayes-féle információs kritérium. Mind a két kritérium a jól illeszkedő, nem túlillesztett modellek esetében kisebb számot ad. Ennek következtében modellválasztás esetén az AIC-t és BIC-t minimalizálni szokták.

férfi	AIC	BIC
LC	5407,257	6036,17
CBD	6809,701	6899,546

2.2. táblázat. Horvát férfi halandóság előrejelzésének információs kritériumai
mortality.org adatai alapján saját számítás

A horvát férfiak adatainál mind a két információs kritérium a Lee-Carter modell esetében kisebb volt (2.2 táblázat). Amennyiben két modellt vizsgálunk és a modellek AIC-inek eltérése elhanyagolható, akkor a két modell egyformán jó. Ebben az esetben az eltérés több, mint 1000 az AIC, a BIC esetében, pedig több mint 800 volt, így a két modell nem lehetett egyformán jónak mondani. (2.2 táblázat)

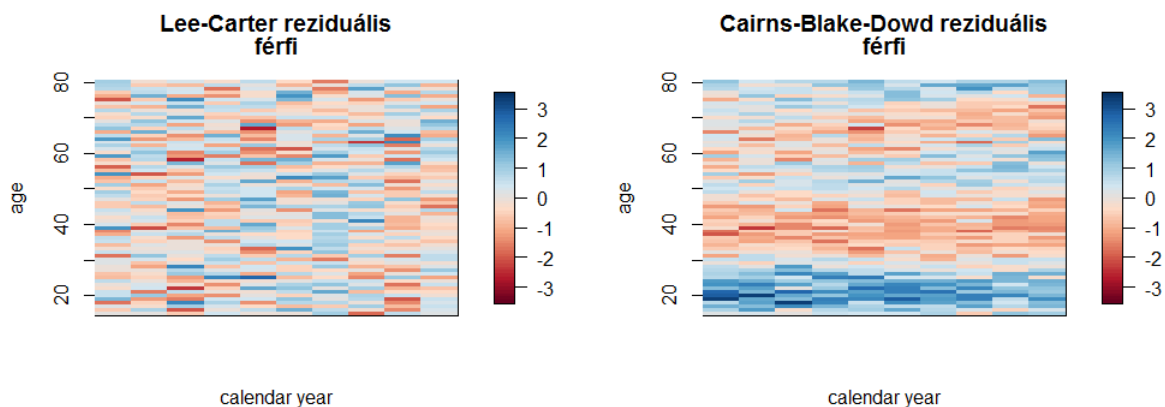
A nők esetében is hasonló helyzettel találkoztam (2.3 táblázat). Itt is a Lee-Carter modell sokkal jobban teljesített, mind a Cairns-Blake-Dowd modell, és itt a különbség lényegesebb volt. Az információs kritériumokból arra következtettem, hogy mind a két esetben a Lee-Carter modell illeszkedik jobban.

nő	AIC	BIC
LC	4812,746	5441,66
CBD	7086,829	7176,674

2.3. táblázat. Horvát női halandóság előrejelzésének információs kritériumai
mortality.org adatai alapján saját számítás

Nem csak az információs kritériumokat, de a reziduálisokat is megvizsgáltam a négy előrejelzés esetében. Hőtérkép alapján figyeltem meg a modellek illesztését, mely jól mutatja, hogy az illesztett érték hol magasabb, vagy alacsonyabb. A 2.7 és 2.8 ábrák alapján mind a két esetben a Lee-Carter modell illeszkedett jobban, hisz a hőtérképen a 0 értéket a fehér szín jelzi, és a Lee-Carter modell reziduálisainak hőtérképe világosabb a

Cairns-Blake-Dowd modell reziduálisainál, vagyis az, hogy egyenletesebben világosabb ábrát kaptunk a férfiak és nők esetén az Lee-Carter modellnél, azt jelenti, hogy a reziduálisok nulla közeli, tehát az illeszkedés pontosabb. Mintázatot se fedeztem fel. Ez arra utal, hogy a kohorszhatás se jelenik meg a Lee-Carter modell esetén. Ezzel szemben a Cairns-Blake-Dowd modell reziduálisai mind a két esetben egyértelmű kohorszhatást mutatnak. A férfiaknál 2002-ben 20 éves kortól indulva a modell alul illeszti a halandóságot, 40 éves kortól pedig felül illeszti.

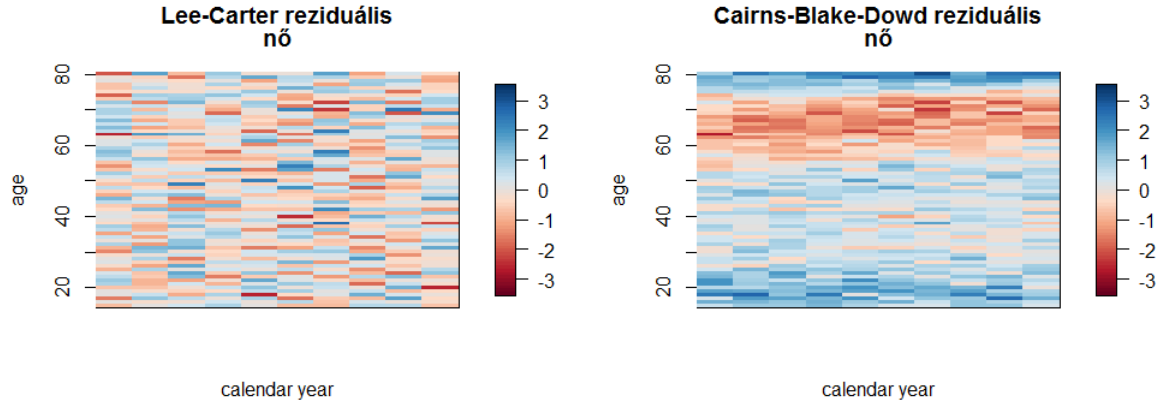


2.7. ábra. Horvát férfi halandósági előrejelzés reziduálisai
mortality.org adatai alapján saját számítás

Az 2.7 ábrán a kohorszhatást jelző átlós egyenes meredeksége nem észrevehető, hisz csak nyolc évnnyi idősorra illesztettünk, de több év esetén láthatóvá válna, ahogy az illesztési hiba végigvonulna az adott generáción. Ötven éves kor felett a modell reziduálisai sokkal jobban közelítenek a nullához, ezáltal a szakirodalmakban leírt megfigyelések is beigazolódottak, miszerint a Cairns-Blake-Dowd modell főleg az idősebb korosztályra illeszkedik jobban. (Vékás [2016]); (Carins, Blake és Dowd [2006]) A nők esetében már nem volt releváns a megfigyelés. A nőknél 60–70-es korosztálynál túlillesztett a modell, 20 éves kor környékén pedig, a férfiaknál kisebb mértékben, de alulbecsli a halandósági rátákat. A nullához közeli értékeket 30 – 50 kor között veszik fel a hibatagok.

Az eddigi két vizsgálat az előrejelzett adatokat újrailleszti a rendelkezésre álló adatokra, ám mint előrejelzést is szerettem volna megvizsgálni. Ez volt az oka, hogy nem a teljes idősből készítettem az előrejelzést, csak nyolc évnnyi adatból, mivel az előrejelzett öt évet és a valós öt évnnyi adatot hasonlítottam egymáshoz.

Az adatok összehasonlításához illeszkedés vizsgálatot használtam, a 2.2 képletet hasz-



2.8. ábra. Horvát női halandósági előrejelzés reziduálisai
mortality.org adatai alapján saját számítás

nálva:

$$\chi_t^{m,g^2} = \sum_{x=15}^{80} \frac{(D_{xt}^g - E_{xt}^g \tilde{q}_{xt}^{m,g})^2}{\tilde{q}_{xt}^{m,g} (1 - \tilde{q}_{xt}^{m,g}) E_{xt}^g} \quad (2.2)$$

ahol g jelöli a nemet, m a modellt és t az évet. Vagyis $g = \text{nő vagy férfi}$, $m = LC \text{ vagy } CBD$, $t = 2011, 2012, 2013, 2014, 2015$. D_{xt} , E_{xt} a valós horvát halottak száma és kitettség a t . évben x éves korban, \tilde{q}_x pedig az előrejelzett horvát ráták, a megfelelő években x éves korban.

Ezáltal húsz illeszkedés vizsgálatot végeztem el, és a 20 vizsgálatból egyetlen egy illeszkedést se tudtam elfogadni 5%-os szignifikancia szint mellett az összes életkort nézve.

A modelleket tovább alkalmaztam korévenként, és így 1 szabadságfokú χ^2 tesztet használtam. Ezáltal már konkrét korcsoportoknál is meg tudtam mondani, hogy melyik modell illeszkedik jobban. Az illeszkedés vizsgálat eredménye igazolta, amit a reziduálisok vizsgálat során láttam. A Lee-Carter modell összességében mind a nők mind a férfiak esetén az összes évben jobb illeszkedést eredményezett, ám korévenként nézve voltak időszakok, ahol a Cairns-Blake-Dowd modell jobb volt. A férfiak esetén, mint korábbi vizsgálatok esetén is, az idősebb korosztályban, konkrétan 50 – 85 év között jobban illeszkedett a CBD. A nőkre vonatkozó adatok szintén igazolták a reziduálisoknál megfigyelteket, mivel a CBD ebben az esetben a 25 – 45 közötti korosztályokra illeszkedett jobban.

A megfigyelések alapján az alkalmazott előrejelzési struktúrát 2.4 táblázat tartalmazza.

	nő	férfi
15-20	LC	LC
20-25	LC	LC
25-30	CBD	LC
30-35	CBD	LC
35-40	CBD	LC
40-45	CBD	LC
45-50	LC	LC
50-55	LC	CBD
55-60	LC	CBD
60-65	LC	CBD
65-70	LC	CBD
70-75	LC	CBD
75-80	LC	CBD

2.4. táblázat. Korcsoportok előrejelzéséhez használt struktúrája
mortality.org adatai alapján saját számítás

Az előrejelzést követően az öt éves korcsoportos q_x -eket a 2.3 képlet alapján határoztam meg:

$${}_5q_x = \frac{{}_5m_x}{1 - \frac{5}{2}m_x}, \quad (2.3)$$

ahol ${}_5m_x$ 5 éves korcsoportok esetén az x korcsoportoz tartozó halálozási ráta.

2.5. Projekció és megyei előrejelzés

2.5.1. Projekció

Az előrejelzett korcsoportos halandósági ráták és a valós megyei ráták között még mindig jelentős különbség volt, ezért további technikákat próbáltam alkalmazni a pontosabb előrejelzés érdekében.

Arra gondoltam, hogy a statisztikában szokásos módon összevetem az eltérések hányadosát és különbségét, melyet a továbbiakban hányados és különbség projekciónak fogok nevezni.

Különbség projekció:

$$k(\tilde{q}_x) = \tilde{q}_x - \frac{(2014\tilde{q}_x - 2014q_x) + (2015\tilde{q}_x - 2015q_x)}{2}, \quad (2.4)$$

ahol \tilde{q}_x az előrejelzett horvát ráta, $2014\tilde{q}_x, 2015\tilde{q}_x$ a megfelelő megyében, az eltolást figyelembe véve, $2014, 2015$ -nek megfelelő horvát halandósági ráta, $2014q_x, 2015q_x$ a megfelelő megyei valós ráták. A projekció lényege, hogy a 2014-es és 2015-ös eltérések átlagával korigálja a későbbi rátákat is.

Hányados projekció:

$$h(\tilde{q}_x) = \frac{\tilde{q}_x}{\frac{(2014\tilde{q}_x/2014q_x) + (2015\tilde{q}_x/2015q_x)}{2}}, \quad (2.5)$$

Ez a projekció 2014-es és 2015-ös ráták arányának számtani átlagával korigál. Ebben az esetben nem csak a számtani, de a mértani átlagot is megvizsgáltam, de ez kevésbé vagy ugyanolyan jó eredményt adott.

A két projekciót a 2016-os adatokon ellenőriztem illeszkedés vizsgálattal. 5%-os szignifikancia szint mellett, a tizenkét illeszkedésből egyiket se fogadtam el, de ez várható volt, hisz a vizsgálat annyira érzékeny, hogy az előrejelzett és a valós adatok illeszkedését se fogadta el.

Ettől függetlenül a χ^2 -eket továbbra is használtam a projekciók kiválasztásához, mégpedig úgy, hogy a kisebb χ^2 -ű projekciót, vagyis a jobban illeszkedőt fogadtam el.

A hat esetből négyszer a hányados projekció volt a hatékonyabb, a maradék két esetben pedig elhanyagolható volt a két érték közötti különbség, ezért a második függvényt választottam, kivéve egy esetet: Győr-Moson-Sopron férfi rátáinál a 2014-es halandóság nagyon torzította az eredményt mind két függvénynél, ezért csak a 2015-ös arányt vettem figyelembe.

Természetesen ezek a vetítések nagyon kezdetlegesek és csak a jelenlegi esetben érvényesek. Hosszútávon nem alkalmazhatóak, mivel a különbség projekció negatív halandósági rátákat, a hányados pedig egynél nagyobb rátákat eredményezhet. Jelen helyzetben ilyen problémák nem merültek fel. Lehetséges továbbgondolása a vizsgálatnak egy általánosított projekció.

2.5.2. Eredmények

A valós adatokat és az előrejelzéseket megyénként és nemenként külön-külön ábrán szemléltettem. (2.9-2.14 ábrák)

Útmutatás az ábrákhoz:

- *Valós adatok:* A KHS-tól kapott valós megyei halandósági ráták.
- *Horvát ráták alapján:* A *mortality.org* honlapon szereplő halandósági rátákból számolt megyei halandóságok az eltolás és projekció segítségével.
- *Előrejelzett ráták alapján:* Az előrejelzett halandósági rátákból számolt megyei halandóságok az eltolás és projekció segítségével.

Győr-Moson-Sopron megye férfi

A győr-moson-soproni férfiak esete volt az egyetlen, ahol nem tudtam a projekciónál figyelembe venni a 2014-es adatokat, mivel, ahogy az 2.9 ábrán is látszik 70 – 75 éves korcsoportnál a 2014-es halandóság túl magas volt, ami torzította volna az eredményt. Ilyen problémák elkerülése miatt használtam két év átlagát, hogy a kisebb anomáliák ne hangsúlyosan szerepeljenek a projekcióban. Ebben az esetben nagyon jelentős volt az eltérés, ezért nem tudtam az átlaggal korrigálni. A másik eltérés, mely az ábrán is látványos, 2016-ban a 70 – 75-ös korosztály jelentősen rossz halandósága, ezért ebben a korcsoportban a 2016-os adat nem mérvadó. Mivel Győr megye a horvát adatokkal együtt mozgott, ezért nem volt időbeli eltolás, így nem volt lehetőség a horvát valós rátákból számolni, hisz az előrejelzett rátákat már alkalmazhattuk.

Győr-Moson-Sopron megye nő

A nők hasonló helyzetben voltak. Az eltolás itt már jelen volt, hisz a 2014-es győri adatok a 2013-as horvát adatoknak felelnek meg. Így is csak egy év volt, amire még voltak valós megyei adatok. 2016-hoz képest 75 – 80 és 50 – 55 éves korcsoportokban tér el jelentősebben az előrejelzés. Az előbbinél a modell alulbecsült, az utóbbinál pedig kisebb mértékben felülbecsült. A többi korévben nincs szembetűnő eltérés. (2.10 ábra)

Fejér megye férfi

A Fejér megyei férfiaknál már megjelennek a horvát adatokból számolt megyei előrejelzett ráták, mivel itt már az eltolás öt év. Ez azt jelentette, hogy a 2014-es Fejér megyei adatok a 2009-es horvát adatoknak felelnek meg, vagyis a 2020-as előrejelzett horvát adatok alapján a 2025-ös Fejér megyei halandóság előrejelzését kapjuk meg. Az egyes korcsoportokban egyenletesen javul a halandóság az idő előrehaladtával. Az előrejelzett adatok jól illeszkedtek a valós megyei adatokhoz, kivéve 60 – 65 éves korcsoportban. Ebben a korcsoportban a projekció és az eltolás ellenére is jelentősebben eltértek a horvát és megyei ráták.

Fejér megye nő

A Fejér megyei nőknél ismét nem találkozunk valós horvát adatokból előrejelzett megyei rátákkal, hisz újfent nincs eltolás ebben az esetben. (2.12 ábra) Mint a győri nőknél itt is a 75 – 80-as korcsoportban láthatunk egy jelentősebb eltérést. A 2016-ban az időskori halandóság jelentősen alacsonyabb volt, mint addig, de ezt a megyei adatok simítatlanságának tudtam be, hisz ilyen mértékű javulás az országos adatokon se látható.

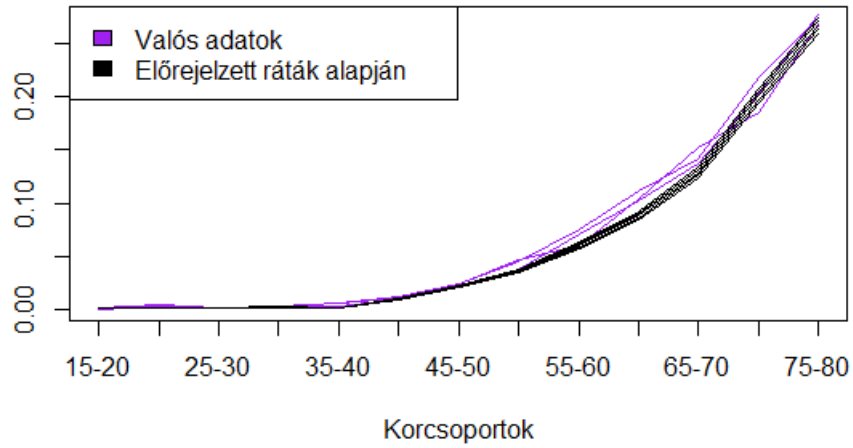
Szabolcs-Szatmár-Bereg megye férfi

A legjobb a Szabolcs-Szatmár-Bereg megyei férfi előrejelzés lett. A nyolc évnyi eltolás miatt a horvát valós adatokat itt tudtam a legjobban felhasználni. (2.13 ábra) Az eltolás miatt itt lehetett a legtovább előrejelezni, mégpedig 2028-ig. A halandóság egyenletes javulását a valós megyei adatoknál is észre lehetett venni.

Szabolcs-Szatmár-Bereg megye nő

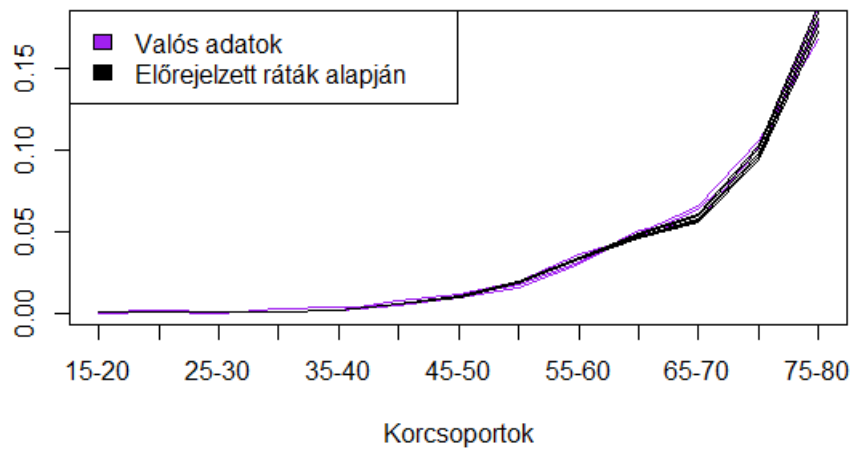
A nők esetében szintén felhasználtam a horvát valós adatokkal, mivel itt öt évnyi eltolás volt. A 2020-as adatokat 2025-ös előrejelzéshez tudtam felhasználni. A megyei adatok szintén egyenletesen javultak, csakúgy mint a férfiaknál, és simábbak voltak, mint más megyék adatai, kivéve a 50-55 éves korcsoportnál. Lényeges eltérés nem fordult elő.

Győr-Moson-Sopron megye férfi

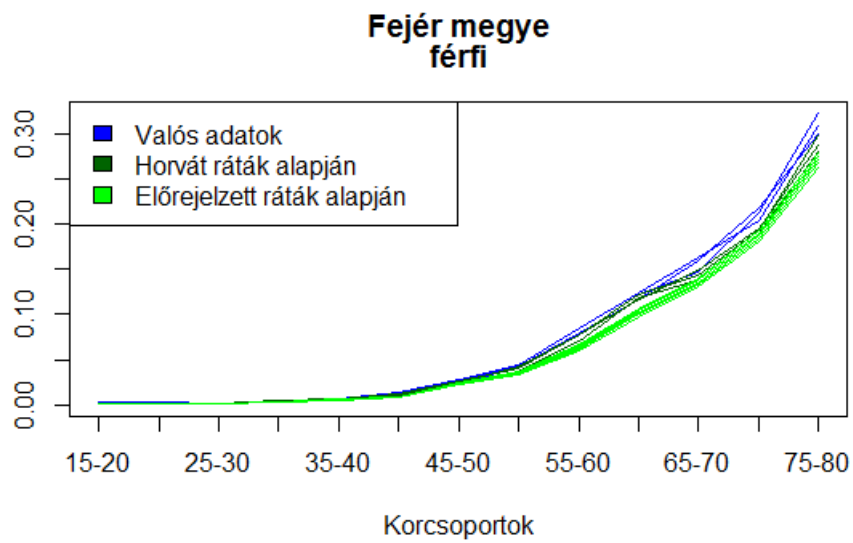


2.9. ábra. Győr-Moson-Sopron megye férfi előrejelzése
KSH és *mortality.org* adatai alapján saját számítás

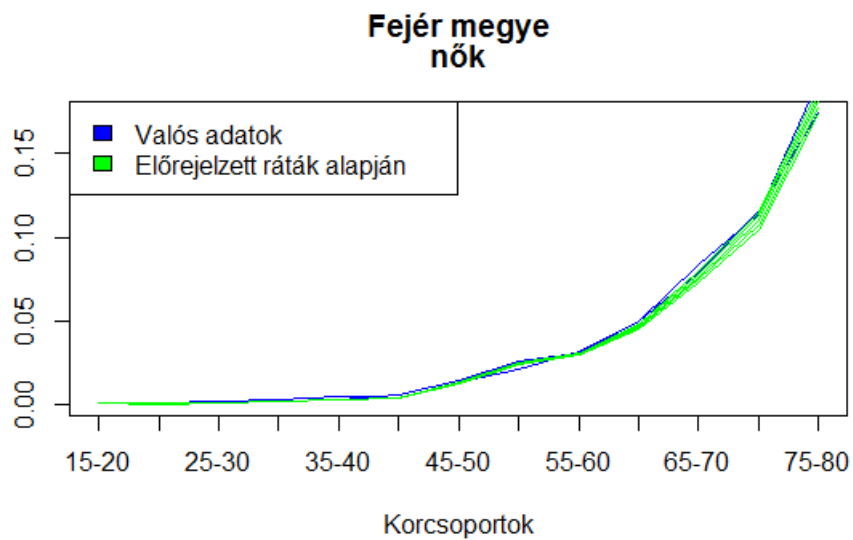
Győr-Moson-Sopron megye nők



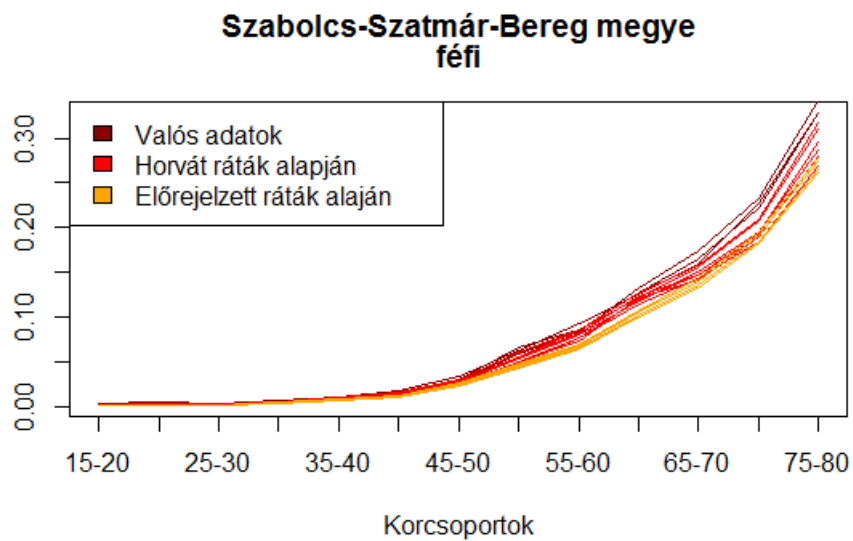
2.10. ábra. Győr-Moson-Sopron megye női előrejelzése
KSH és *mortality.org* adatai alapján saját számítás



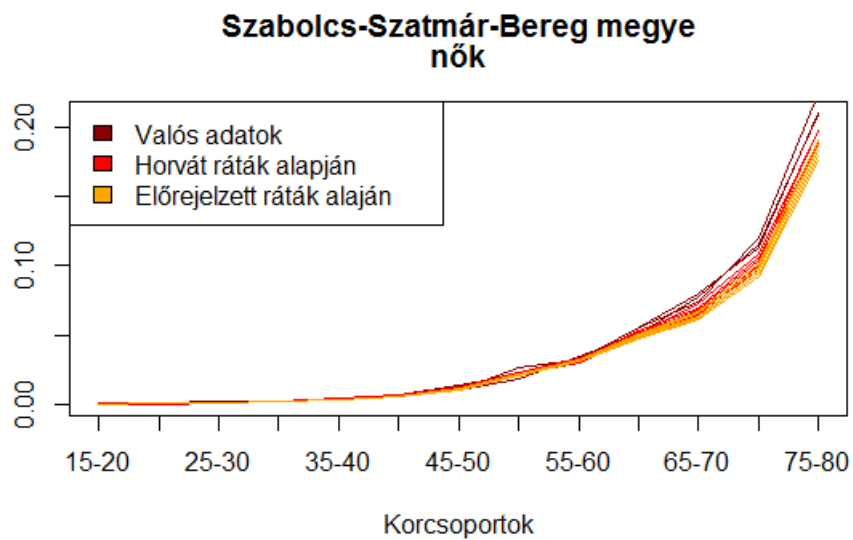
2.11. ábra. Fejér megye férfi előrejelzése
KSH és *mortality.org* adatai alapján saját számítás



2.12. ábra. Fejér megye női előrejelzése
KSH és *mortality.org* adatai alapján saját számítás



2.13. ábra. Szabolcs-Szatmár-Bereg megye férfi előrejelzése
KSH és *mortality.org* adatai alapján saját számítás



2.14. ábra. Szabolcs-Szatmár-Bereg megye női előrejelzése
KSH és *mortality.org* adatai alapján saját számítás

3. fejezet

Következtetések és továbbfejlesztési lehetőségek

A célom Magyarország regionális halandóságának előrejelzése volt, melyet hiányos adatok mellett próbáltam elvégezni. Mivel a megyei halandósági táblákból csak három évé állt rendelkezésemre, ezért előzetes vizsgálatok után a horvát halandósági rátákkal helyettesítettem, mely már hosszabb időszakra elérhető volt. Az adatokat két, a korábbi fejezetekben ismertetett modell segítségével előrejeleztem. Az előrejelzést tényadatokon ellenőriztem és arra a következtetésre jutottam, hogy a horvát adatokat egyik modell se írja le teljesen megfelelően, de az egész 15-85 éves korcsoportot nézve a Lee-Carter modell mégis jobb a Cairns-Blake-Dowd modellnél. Részletesebb vizsgálat során azt észleltem, hogy a Cairns-Blake-Dowd modell az egyes korcsoportoknál jobban illeszkedik, mint a Lee-Carter. A férfiaknál ez a korcsoport 50-től 85 éves korig tartott, a nőknél pedig 25-45 éves kor között illeszkedett jól. A férfiak illeszkedése a szerzők által is leírt észrevételt igazolja, miszerint a Cairns-Blake-Dowd modell főleg az idősebb korosztályban illeszkedik jobban. A nőknél nem igazolódott be ez az állítás, de így is használni tudtam a modellt.

A horvát halandósági rátákra különböző módszereket alkalmaztam, hogy jobban illeszkedjenek a megyei adatokhoz.

Az így kapott előrejelzett ráták és a megyei ráták illeszkedését nem tudtam ellenőrizni a megyei adatok hiánya miatt, de grafikus módon a trendet és a halandósági görbét vizsgálva hasonlóan viselkedtek, mint amit a konkrét megyei előrejelzéstől elvártam.

Mindezek ellenére a vizsgálatnak és előrejelzésnek több továbbfejlesztési lehetősége is van.

Egyrészt, a referenciaország, Horvátország helyett előfordulhat alkalmasabb ország a

megyék reprezentálására, de nekem csak a *mortality.org* által vizsgált országok halandósága állt rendelkezésemre.

Másrészt, hosszabb vizsgálat után lehet, hogy található alkalmasabb halandósági modell a horvát adatok előrejelzéséhez, mely jobban leírja a halandósági ráták alakulását, illetve a két modellre több kritikát is megfogalmaztak a szakirodalomban. (Vékás [2016]) Egyik ilyen továbbfejlesztési lehetősége a modelleknek, a kohorszhatást kezelő paraméterrel való bővítésük.

Harmadrészt, alkalmaztam két módszert, az eltolást és a projekciót, mely jobban hozzáilleszti a horvát adatokat a megyei adatokhoz. A két módszer, mint már a dolgozatban is megemlítettem, nagyon kezdetleges és hosszabb időtávban akár értelmetlen rátákat is eredményezhetnek, így javasolt a továbbfejlesztésük, mely lényegesen több évnnyi illesztésre is alkalmas.

Minden javítási igény mellett a megyei halandósági előrejelzések hasznosak mind a biztosítóknak, mind a megyei önkormányzatoknak. Amennyiben biztosítók figyelembe veszik a halandóságok közötti regionális különbségeket a long-term care szolgáltatásoknál vagy a járadék fizetéseknél, pontosabb becslést tudnak adni a jövőbeni hasonló jellegű szolgáltatásokra. Az önkormányzatok számára is fontos ez, hisz az előrejelzett halandósági ráták birtokában, a megfelelő szakemberek segítségével egy-egy város vagy megye populációjára vetítve meg lehet mondani a várható idősök, nyugdíjasok számát. A jövőbeli adatok tudatában sokkal könnyebb megszervezni például az idősök ellátását, becslést adni az idősök otthona iránti igényükre. Másrészt, a nyugdíjasok vásárlóerejének felmérésében is hasznos lehet, hisz az előrejelzéssel könnyebben lehet kalkulálni, pontosabb becslést lehet adni egyes termékek és egészségügyi szolgáltatások iránti jövőbeli keresletükre.

Irodalomjegyzék

- [1] Arató Miklós, Bozsó Dávid, Elek Péter, Zempléni András (2009) *Forecasting and Simulating Mortality Tables. Mathematical and Computer Modelling.*
- [2] Claudia Breakwell, Madhavi Bajekal (2001) *Health expectancies in the UK and its constituent countries*, Office for National Statistics. (www.researchgate.net/publication/7253968)
- [3] N. Brouhns, M.Deniut, I. Van Keilegom (2002) *A poisson log-bilinear regression approach to the construction of projected life table*, Insurance: Mathematics and Economics, 31:373-393.
- [4] Andrew J. G. Cairns, David Blake, Kevin Dowd (2006) *A two-factor model for stochastic mortality with parameter uncertainty: Theory and Calibration*, The Journal of Risk and Insurance, Vol. 73, No. 4, 687-718.
- [5] Andrew J.G. Cairns, David Blake, Kevin Dowd, Guy D. Coughlan, David Epstein, Marwa Khalaf-Allah (2010) *Insurance: Mathematics and Economics- Mortality density forecast: An analysis of six stochastic mortality models.* (www.elsevier.com/locate/ime)
- [6] Edmund Cannon (2009) *Estimation and pricing with the Cairns-Blake-Dowd model of mortality*, WP Number 65.
- [7] Ján Gogola *A comparison of Lee-Carter and Cairns-Blake-Dowd stochastic mortality model*, Mathematical Methods in Science and Mechanics.
- [8] J. Hunt, D. Blake (2014) *A general procedure for constructing mortality models*, North American Actuarial Journal, 18(1):116-138.

- [9] Kovács Erzsébet, Vékás Péter (2016) *Mortality and Longevity Risk, Chapter 9in: The Palgrave Handbook of Unconventional Risk Transfer, Springer, pp. 269-297.*
- [10] Központi Statisztikai Hivatal (2008) *A demográfiai folyamatok regionális különbségei, 1980-2007 .*
- [11] Li J. (2014) *A Quantitative Comparison of Simulation Strategies for Mortality Projection. Annals of Actuarial Science, Vol. 8. Part 2. pp.281-297.*
- [12] Carlo Maccheroni, Samuel Nocito (2016) *Backtesting the Lee-Carter and Cairns-Blake-Dowd Stochastic Mortality Models on Italian Death Rates. (www.mdpi.com)*
- [13] Pallós Emil (1971) *KSH Népeségtudományi Kutató Intézetének és a Magyar Tudományos Akadémia Demográfiai Bizottságának közleményei 31..*
- [14] P.J. Sweeting (2011) *A Trend-Change Extension of the Cairns-Blake-Dowd Model. Annals of Actuarial Science, Vol. 5. Part 2. pp. 143-162.*
- [15] Vékás Péter (2016) *Az élettartam-kozkázat modellezése.*
- [16] A. M. Villegas, P. Millossovich, V. Kaishev (2016) *StMoMo: An R Package for Stochastic Mortality Modelling.*

4. fejezet

Függelék

4.1. Forráskód

4.1.1. Magyarország halandósági rátáinak LC és CBD modellel történő előrejelzésnek forráskódja

```
library("StMoMo")
library("readxl")

Dxt<-read_excel("C:/Users/User/Desktop/Előadás/demogr_ext.xlsx",
  col_names=FALSE, sheet="dxt")

Ext2<-read_excel("C:/Users/User/Desktop/Előadás/demogr_ext.xlsx",
  col_names=FALSE, sheet="ext2")

Ext<-read_excel("C:/Users/User/Desktop/Előadás/demogr_ext.xlsx",
  col_names=FALSE, sheet="extc")

ages<-0:100
years<-1975:2015

LC<-lc()
CBD<-cbd(link ="log")
LCfit<-fit(LC, Dxt=Dxt, Ext=Ext, ages=ages, years=years,ages.fit=ages)
```

```

CBDfit<-fit(CBD, Dxt=Dxt, Ext=Ext2, ages=ages, years=years,ages.fit=ages)

AIC(LCfit)
AIC(CBDfit)
BIC(LCfit)
BIC(CBDfit)

LCres<-residuals(LCfit)
CBDres<-residuals(CBDfit)
plot(LCres, type="colourmap", reslim=c(-3.5,3.5))
plot(CBDres, type="colourmap", reslim=c(-3.5,3.5))

LCfor<-forecast(LCfit, h=4)
CBDfor<-forecast(CBDfit, h=4)
plot(LCfor)
plot(CBDfor)

LCfor[1]
CBDfor[1]

write.csv(LCfor[1], "C:/Users/User/Desktop/Előadás/forecastlc.csv")
write.csv(CBDfor[1], "C:/Users/User/Desktop/Előadás/forecastcbd.csv")

```

4.1.2. Országok és megyék halandóságának összehasonlítása

```

##### Országok #####
#####Lengyel
LTPOL <- ReadHMD("LT_f",countries="POL",interval="5x1","a-edda@hotmail.com","*****")
LTPOL$data
QxPOL<-matrix(data=NA, nrow =13 , ncol= 40)
k=0
for ( i in 1:40)
{

```

```

k=1974+i
f<-1:24
f<- LTPOL$data[which(LTPOL$data$Year==k),5]
QxPOL[,i]<-f[5:17]
}
#####Szlovákia
LTSK <- ReadHMD("LT_f",countries="SVK",interval="5x1","a-edda@hotmail.com","*****")
LTSK$data
QxSK<-matrix(data=NA, nrow =13 , ncol= 40)
k=0
for ( i in 1:40)
{
k=1974+i
f<-1:24
f<- LTSK$data[which(LTSK$data$Year==k),5]
QxSK[,i]<-f[5:17]
}
#####Szlovénia
LTSN <- ReadHMD("LT_f",countries="SVN",interval="5x1","a-edda@hotmail.com","*****")
LTSN$data
QxSN<-matrix(data=NA, nrow =13 , ncol= 40)
k=0
for ( i in 1:40)
{
k=1974+i
f<-1:24
f<- LTSN$data[which(LTSN$data$Year==k),5]
QxSN[,i]<-f[5:17]
}
#####Cseh
LTCZE <- ReadHMD("LT_f",countries="CZE",interval="5x1","a-edda@hotmail.com","*****")
LTCZE$data
QxCZE<-matrix(data=NA, nrow =13 , ncol= 40)
k=0

```

```

for ( i in 1:40)
{
  k=1974+i
  f<-1:24
  f<- LTCZE$data[which(LTCZE$data$Year==k),5]
  QxCZE[,i]<-f[5:17]
}

#####Ausztria
LTAUT <- ReadHMD("LT_f",countries="AUT",interval="5x1","a-edda@hotmail.com","*****")
LTAUT$data
QxAUT<-matrix(data=NA, nrow =13 , ncol= 40)
k=0
for ( i in 1:40)
{
  k=1974+i
  f<-1:24
  f<- LTAUT$data[which(LTAUT$data$Year==k),5]
  QxAUT[,i]<-f[5:17]
}

#####Horváth
LTHRV<- ReadHMD("LT_f",countries="HRV",interval="5x1","a-edda@hotmail.com","*****")
LTHRV$data
QxHRV<-matrix(data=NA, nrow =13 , ncol= 14)
k=0
for ( i in 1:14)
{
  k=2001+i
  f<-1:24
  f<- LTHRV$data[which(LTHRV$data$Year==k),5]
  QxHRV[,i]<-f[5:17]
}

#####UKRajna
AUTKR <- ReadHMD("LT_f",countries="UKR",interval="5x1","a-edda@hotmail.com","*****")

```

```

AUTKR$data
QxUKR<-matrix(data=NA, nrow =13 , ncol= 40)
k=0
for ( i in 1:40)
{
  k=1974+i
  f<-1:24
  f<- AUTKR$data[which(AUTKR$data$Year==k),5]
  QxUKR[,i]<-f[5:17]
}
##### Megyék #####
#####Fejér
####Lengyel

nhfnPOL<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhfnPOL[,i]<-(QxPOL[,i]-fn14)^2}

szumfnPOL<-1:40

for (j in 1:40)
{szumfnPOL[j]=0
for (i in 1:13)
{szumfnPOL[j]=szumfnPOL+nhfnPOL[i,j]}}
#####Szlovák
nhfnSK<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhfnSK[,i]<-(QxSK[,i]-fn14)^2}

szumfnSK<-1:40

for (j in 1:40)
{szumfnSK[j]=0
for (i in 1:13)

```

```

{szumfnSK[j]=szumfnSK+nhfnSK[i,j]}}
#####Szlóvén
nhfnSN<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhfnSN[,i]<-(QxSN[,i]-fn14)^2}

szumfnSN<-9:40

for (j in 9:40)
{szumfnSN[(j-8)]=0
for (i in 1:13)
{szumfnSN[(j-8)] =szumfnSN+nhfnSN[i,j]}}
#####Cseh
nhfnCZE<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhfnCZE[,i]<-(QxCZE[,i]-fn14)^2}

szumfnCZE<-1:40

for (j in 1:40)
{szumfnCZE[j]=0
for (i in 1:13)
{szumfnCZE[j]=szumfnCZE+nhfnCZE[i,j]}}
#####Ausztria
nhfnAUT<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhfnAUT[,i]<-(QxAUT[,i]-fn14)^2}

szumfnAUT<-1:40

for (j in 1:40)
{szumfnAUT[j]=0
for (i in 1:13)
{szumfnAUT[j]=szumfnAUT+nhfnAUT[i,j]}}

```



```

#####Horváth
nhfnHRV<-matrix(data=NA, nrow =13 , ncol= 14)
for (i in 1:14)
{nhfnHRV[,i]<-(QxHRV[,i]-fn14)^2}

szumfnHRV<-1:14

for (j in 1:14)
{szumfnHRV[j]=0
for (i in 1:13)
{szumfnHRV[j]=szumfnHRV+nhfnHRV[i,j]}}
#####Ukrajna
nhfnUKR<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhfnUKR[,i]<-(QxUKR[,i]-fn14)^2}

szumfnUKR<-1:39

for (j in 1:39)
{szumfnUKR[j]=0
for (i in 1:13)
{szumfnUKR[j]=szumfnUKR+nhfnUKR[i,j]}}
#####Szabolcs
####Lengyel
nhsznPOL<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhsznPOL[,i]<-(QxPOL[,i]-szn14)^2}

szumsznPOL<-1:40

for (j in 1:40)
{szumsznPOL[j]=0
for (i in 1:13)
{szumsznPOL[j]=szumsznPOL+nhsznPOL[i,j]}}

```

```

#####Szlovák
nhsznSK<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhsznSK[,i]<-(QxSK[,i]-szn14)^2}

szumsznSK<-1:40

for (j in 1:40)
{szumsznSK[j]=0
for (i in 1:13)
{szumsznSK[j]=szumsznSK+nhsznSK[i,j]}}
#####Szlovén
nhsznSN<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhsznSN[,i]<-(QxSN[,i]-szn14)^2}

szumsznSN<-9:40

for (j in 9:40)
{szumsznSN[(j-8)]=0
for (i in 1:13)
{szumsznSN[(j-8)] =szumsznSN+nhsznSN[i,j]}}
#####Cseh
nhsznCZE<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhsznCZE[,i]<-(QxCZE[,i]-szn14)^2}

szumsznCZE<-1:40

for (j in 1:40)
{szumsznCZE[j]=0
for (i in 1:13)
{szumsznCZE[j]=szumsznCZE+nhsznCZE[i,j]}}
#####Ausztria

```

```

nhsznAUT<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhsznAUT[,i]<-(QxAUT[,i]-szn14)^2}

szumsznAUT<-1:40

for (j in 1:40)
{szumsznAUT[j]=0
for (i in 1:13)
{szumsznAUT[j]=szumsznAUT+nhsznAUT[i,j]}}
#####Horváth
nhsznHRV<-matrix(data=NA, nrow =13 , ncol= 14)
for (i in 1:14)
{nhsznHRV[,i]<-(QxHRV[,i]-szn14)^2}

szumsznHRV<-1:14

for (j in 1:14)
{szumsznHRV[j]=0
for (i in 1:13)
{szumsznHRV[j]=szumsznHRV+nhsznHRV[i,j]}}
#####Ukrajna
nhsznUKR<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhsznUKR[,i]<-(QxUKR[,i]-szn14)^2}

szumsznUKR<-1:39

for (j in 1:39)
{szumsznUKR[j]=0
for (i in 1:13)
{szumsznUKR[j]=szumsznUKR+nhsznUKR[i,j]}}
#####Győr
####Lengyel

```

```

nhgynPOL<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhgynPOL[,i]<-(QxPOL[,i]-gyn14)^2}

szumgynPOL<-1:40

for (j in 1:40)
{szumgynPOL[j]=0
for (i in 1:13)
{szumgynPOL[j]=szumgynPOL+nhgynPOL[i,j]}}
#####Szlóvák
nhgynSK<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhgynSK[,i]<-(QxSK[,i]-gyn14)^2}

szumgynSK<-1:40

for (j in 1:40)
{szumgynSK[j]=0
for (i in 1:13)
{szumgynSK[j]=szumgynSK+nhgynSK[i,j]}}
#####Szlóvén
nhgynSN<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhgynSN[,i]<-(QxSN[,i]-gyn14)^2}

szumgynSN<-9:40

for (j in 9:40)
{szumgynSN[(j-8)]=0
for (i in 1:13)
{szumgynSN[(j-8)] =szumgynSN+nhgynSN[i,j]}}
#####Cseh
nhgynCZE<-matrix(data=NA, nrow =13 , ncol= 40)

```

```

for (i in 1:40)
{nhgynCZE[,i]<-(QxCZE[,i]-gyn14)^2}

szumgynCZE<-1:40

for (j in 1:40)
{szumgynCZE[j]=0
for (i in 1:13)
{szumgynCZE[j]=szumgynCZE+nhgynCZE[i,j]}}
#####Ausztria
nhgynAUT<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)
{nhgynAUT[,i]<-(QxAUT[,i]-gyn14)^2}

szumgynAUT<-1:40

for (j in 1:40)
{szumgynAUT[j]=0
for (i in 1:13)
{szumgynAUT[j]=szumgynAUT+nhgynAUT[i,j]}}
#####Horváth
nhgynHRV<-matrix(data=NA, nrow =13 , ncol= 14)
for (i in 1:14)
{nhgynHRV[,i]<-(QxHRV[,i]-gyn14)^2}

szumgynHRV<-1:14

for (j in 1:14)
{szumgynHRV[j]=0
for (i in 1:13)
{szumgynHRV[j]=szumgynHRV+nhgynHRV[i,j]}}
#####Ukrajna
nhgynUKR<-matrix(data=NA, nrow =13 , ncol= 40)
for (i in 1:40)

```

```

{nhgynUKR[,i]<-(QxUKR[,i]-gyn14)^2}

szumgynUKR<-1:39

for (j in 1:39)
{szumgynUKR[j]=0
for (i in 1:13)
{szumgynUKR[j]=szumgynUKR+nhgynUKR[i,j]}}

```

4.1.3. Horvátország halandósági rátáinak előrejelzésének forráskódja

```

DxHRV<-ReadHMD("Dx",countries="HRV",interval="5x1","a-edda@hotmail.com","*****")
ExHRV <- ReadHMD("Ex",countries="HRV",interval="5x1","a-edda@hotmail.com","****")

DxHRVn<-matrix(data=NA, nrow = 13 , ncol= 10)
k=0
for ( i in 1:10)
{
k=2005+i
DxHRVnseg<-1:24
DxHRVnseg <- DxHRV$data[which(DxHRV$data$Year==k),4]
DxHRVn[,i]<-DxHRVnseg[5:17]}
ExHRVn<-matrix(data=NA, nrow =13 , ncol= 10)
k=0
for ( i in 1:10)
{
k=2005+i
ExHRVnseg<-1:24
ExHRVnseg<- ExHRV$data[which(ExHRV$data$Year==k),4]
ExHRVn[,i]<-ExHRVnseg[5:17]}

```

```
ExHRVnk<-ExHRVn-DxHRVn/2
```

```
DxHRVf<-matrix(data=NA, nrow =13 , ncol= 10)
```

```
k=0
```

```
for ( i in 1:1)
```

```
{
```

```
  k=2005+i
```

```
  DxHRVfseg<-1:24
```

```
  DxHRVfseg<- DxHRV$data[which(DxHRV$data$Year==k),5]
```

```
  DxHRVf[,i]<-DxHRVfseg[5:17]}
```

```
ExHRVf<-matrix(data=NA, nrow =13 , ncol= 10)
```

```
k=0
```

```
for ( i in 1:10)
```

```
{
```

```
  k=2005+i
```

```
  ExHRVfseg<-1:24
```

```
  ExHRVfseg<- ExHRV$data[which(ExHRV$data$Year==k),5]
```

```
  ExHRVf[,i]<-ExHRVfseg[5:17]}
```

```
ExHRVfk<-ExHRVf-DxHRVf/2
```

```
ages<-c(1,2,3,4,5,6,7,8,9,10,11,12,13)
```

```
years<-2006:2015
```

```
LCHRVn<-lc()
```

```
CBDHRVn<-cbd(link ="log")
```

```
LCHRVf<-lc()
```

```
CBDHRVf<-cbd(link ="log")
```

```
LCfitHRVn<-fit(LCHRVn, Dxt=DxHRVn, Ext=ExHRVnk, ages=ages,  
  years=years,ages.fit=ages)
```

```
CBDfitHRVn<-fit(CBDHRVn, Dxt=DxHRVn, Ext=ExHRVn, ages=ages,  
  years=years,ages.fit=ages)
```

```
LCfitHRVf<-fit(LCHRVf, Dxt=DxHRVf, Ext=ExHRVfk, ages=ages,  
  years=years,ages.fit=ages)
```

```
CBDfitHRVf<-fit(CBDHRVf, Dxt=DxHRVf, Ext=ExHRVf, ages=ages,  
  years=years,ages.fit=ages)
```

```
AIC(LCfitHRVn)
```

```
AIC(CBDfitHRVn)
```

```
BIC(LCfitHRVn)
```

```
BIC(CBDfitHRVn)
```

```
AIC(LCfitHRVf)
```

```
AIC(CBDfitHRVf)
```

```
BIC(LCfitHRVf)
```

```
BIC(CBDfitHRVf)
```

```
LCresHRVn<-residuals(LCfitHRVn)
```

```
CBDresHRVn<-residuals(CBDfitHRVn)
```

```
LCresHRVf<-residuals(LCfitHRVf)
```

```
CBDresHRVf<-residuals(CBDfitHRVf)
```

```
plot(LCresHRVn, type="colourmap", reslim=c(-3.5,3.5))
```

```
plot(CBDresHRVn, type="colourmap", reslim=c(-3.5,3.5))
```

```
plot(LCresHRVf, type="colourmap", reslim=c(-3.5,3.5))
```

```
plot(CBDresHRVf, type="colourmap", reslim=c(-3.5,3.5))
```

```
LCforHRVn<-forecast(LCfitHRVn, h=4)
```

```
CBDforHRVn<-forecast(CBDfitHRVn, h=4)
```

```
LCforHRVf<-forecast(LCfitHRVf, h=4)
```

```
CBDforHRVf<-forecast(CBDfitHRVf, h=4)
```

```
plot(LCforHRVn)
```

```
plot(CBDforHRVn)
```

```
plot(LCforHRVf)
```



```
plot(CBDforHRVf)
```

```
write.csv(LCforHRVn[1], "C:/Users/User/Desktop/Szakdog/Relorejelzesek/LCforecastHRVn[1].csv")
write.csv(CBDforHRVn[1], "C:/Users/User/Desktop/Szakdog/Relorejelzesek/CBDforecastHRVn[1].csv")
write.csv(LCforHRVf[1], "C:/Users/User/Desktop/Szakdog/Relorejelzesek/LCforecastHRVf[1].csv")
write.csv(CBDforHRVf[1], "C:/Users/User/Desktop/Szakdog/Relorejelzesek/CBDforecastHRVf[1].csv")
```

4.1.4. Megyei és horvát női halandósági ráták vizsgálata

```
LTHRvN<- ReadHMD("LT_f",countries="HRV",interval="5x1","a-edda@hotmail.com","Edina000")
LTHRvN$data
QxHRVn<-matrix(data=NA, nrow =13 , ncol= 14)
k=0
for ( i in 1:14)
{
  k=2001+i
  f<-1:24
  f<- LTHRvN$data[which(LTHRvN$data$Year==k),5]
  QxHRVn[,i]<-f[5:17]
}
####
nhfnHRVn<-matrix(data=NA, nrow =13 , ncol= 14)
for (i in 1:14)
{nhfnHRVn[,i]<-(QxHRVn[,i]-fn14)^2}

szumfnHRVn<-1:14

for (j in 1:14)
{szumfnHRVn[j]=0
for (i in 1:13)
{szumfnHRVn[j]=szumfnHRVn+nhfnHRVn[i,j]}}
```

```

###
nhsznHRVn<-matrix(data=NA, nrow =13 , ncol= 14)
for (i in 1:14)
{nhsznHRVn[,i]<-(QxHRVn[,i]-szn14)^2}

szumsznHRVn<-1:14

for (j in 1:14)
{szumsznHRVn[j]=0
for (i in 1:13)
{szumsznHRVn[j]=szumsznHRVn+nhsznHRVn[i,j]}}
####
nhgynHRVn<-matrix(data=NA, nrow =13 , ncol= 14)
for (i in 1:14)
{nhgynHRVn[,i]<-(QxHRVn[,i]-gyn14)^2}

szumgynHRVn<-1:14

for (j in 1:14)
{szumgynHRVn[j]=0
for (i in 1:13)
{szumgynHRVn[j]=szumgynHRVn+nhgynHRVn[i,j]}}

fejér<-QxHRVn[,13]-fn14
győr<-QxHRVn[,12]-gyn14
szabolcs<-QxHRVn[,8]-szn14
abs(fejér)

fejérszazalek<-(abs(fejér))/fn14
szabolcsszazalek<-(abs(szabolcs))/szn14
győrszazalek<-(abs(győr))/gyn14

fnkul<-QxHRVn[,13]-fn14
fnhany<-QxHRVn[,13]/fn14

```

```
sznkul<-QxHRVn[,8]-szn14  
qk16<-QxHRVn[,10]-sznkul  
sznhany<-QxHRVn[,8]/szn14  
qh16<-QxHRVn[,10]/sznhany
```

```
sznkul15<-QxHRVn[,9]-szn15  
sznhany15<-QxHRVn[,9]/szn15
```

```
gynkul<-QxHRVn[,12]-gyn14  
qk16_gyn<-QxHRVn[,14]-gynkul  
gynhany<-QxHRVn[,12]/gyn14  
qh16_gyn<-QxHRVn[,14]/gynhany
```

```
ynkul15<-QxHRVn[,13]-gyn15  
gynhany15<-QxHRVn[,13]/gyn15  
gynkul16<-QxHRVn[,14]-gyn16  
gynhany16<-QxHRVn[,14]/gyn16
```