

MODELLEK A MAGYARORSZÁGI ÉVES FÖLDRENGÉSSZÁMOK VIZSGÁLATÁRA

KOVÁCS ELŐD, ARATÓ MIKLÓS, LIPOVITS ÁGNES

A magyarországi éves földrengésszámokat vizsgáltuk. Célunk olyan sztochasztikus modell kiválasztása volt, amely megfelelően illeszkedik a megfigyelt adatokra, és alkalmas a jövőbeli folyamatok véletlen szimulálására. Vizsgálatunk során csak azokat a rengéseket néztük, amelyeknek magnitúdója legalább 2,9. Az adatok közül elhagytuk az – általunk adott definíció alapján kiszűrt – utórengéseket, majd megnéztük, hogy a XX. század során az egyes években hány rengés volt. Több modellt is illesztettünk az így kapott adatsorra: a biztosításmatematikában leggyakrabban használt kárszámeloszlásokat, keverék Poisson-eloszlásokat és rejtett Markov-modellt (HMM=Hidden Markov Model). Bemutatjuk a becslési módszereket és eredményeket, valamint néhány szimulációt.

1. Bevezetés

A katalógusokban a földrengéseket jellemző adatok általában: a rengés helye (3 dimenziós adat), ideje, erőssége, amelyek kiegészülhetnek más adatokkal. A magyarországi rengésekről (részben interaktív módon) részletes adatokat nyújt a www.georisk.hu, illetve a www.foldrenges.hu honlap. A [8]–[18] munkákban részletes adatsorokat találhatunk. A Pannon-régió, illetve a Kárpát-medence szeizmicitását pedig – több más szerző és tanulmány mellett – Tóth L. és társai cikke [18], illetve Zsíros T. [17] tekinti át nagyon részletesen.

1.1. A rengésszámról

A földrengésszámmal kapcsolatos egyik első vizsgálat Omori nevéhez fűződik. Az 1891-ben, Japánban, Nobiban történt rendkívül erős – 8-as magnitúdójú – földrengést követő utórengéseket vizsgálta. A főrengés óta eltelt napok t száma és az egy napra eső utórengések $n(t)$ száma között az alábbi kapcsolatot találta: $n(t) = K(t + c)^{-1}$, megfelelő K és c konstanssal (Omori, 1894). Később 44 utórengés-sorozat vizsgálatát követően Utsu ezt a következőképpen módosította: $n(t) = K(t + c)^{-p}$, ahol a p konstans gyakran 1-nél kissé nagyobb, általában 1 és 1,5 közé esik [3]. Utóbb más, több paraméteres illesztések is születtek. Megfelelő skálák birtokában – amelyek korábban nem léteztek – a XX. század első felében kezdték vizsgálni a rengés energiája és a rengésszám közötti kapcsolatot (Wadati, 1932, Gutenberg és Richter, 1936). Az energiát erg mértékegységben mérve

(1 erg = 10^{-7} joule) Gutenberg és Richter (1956) a $\log E = 1,5 \cdot M + 11,8$ definíciót adta a rengés M magnitúdójára. (Ha az energiát joule-ban mérjük, akkor – természetesen – a $\log E = 1,5M + 4,8$ összefüggést kapjuk.) Ezekkel a jelölésekkel korábbi eredményeiket az alábbi formában írhatták fel: $\log(N(M)) = a - bM$, ahol $N(M)$ jelöli az M -nél erősebb rengések számát (adott nagyságú területen és időintervallumban), a és b pedig a területre jellemző állandó. Az a konstans értéke (természetesen) függ a vizsgált terület és időintervallum nagyságától, illetve hosszától. A megfigyelt területhez tartozó b konstans értéke általában 0,6 és 1,1 közé esik (Zsíros munkájában 0,5 és 1,5 közöttinek mondja [17]), sok terület esetén közel 1-hez [7, 3]. Ezek a konstansok egy Magyarország nagyságú területen nagyon eltérőek lehetnek.

1.2. Statisztikai megközelítés

A földrengések eloszlásának statisztikai vizsgálata a XX. század közepén vett nagy lendületet, bár a legelső, legegyszerűbb modellek már jóval korábban megszülettek. A modellek egy része az időbeli, másik része a térbeli eloszlásra koncentrált, majd születtek tér-idő modellek is.

Az időbeli eloszlás vizsgálatára elsőként stacionárius Poisson-folyamatot alkalmaztak (Schuster, 1897) [3]. Eszerint egy $[t; t + \Delta t]$ időintervallumban egy adott területen bekövetkező földrengések száma csak Δt -től függő $\xi_{\Delta t}$ valószínűségi változó. Annak valószínűsége, hogy éppen n rengés történik az adott területen és időben:

$$P(\xi_{\Delta t} = n) = (\nu \cdot \Delta t)^n \cdot \exp(-\nu \cdot \Delta t) / n!.$$

Ebben a modellben csupán a ν intenzitás becslésére van szükség, amely megfelelően nagy T hosszúságú időintervallumot tekintve N/T , ahol N az időintervallumra eső rengések száma. Később inhomogén Poisson-folyamatokkal is közelítették a rengésszámokat.

Az időbeli eloszlást vizsgáló modellek másik csoportja a felújítási folyamatok apparátusát használja. Ezek két rengés között eltelt idő jellemzőit, például várható értékét vizsgálják.

1.3. A cikkben tárgyalt modellek

Tanulmányunkban az éves rengésszámot, mint valószínűségi változót vizsgáltuk. Először feltételeztük, hogy az évenkénti rengések számai független, azonos eloszlású valószínűségi változók, és eloszlásukat kívántuk azonosítani, paramétereiket becsülni.

Miután kiderült, hogy az $(a, b, 0)$ – vagyis a Poisson, binomiális, illetve negatív binomiális – eloszlások nem illeszkednek jól, Poisson eloszlások keverékeit vizsgáltuk és paramétereiket becsültük. Ezekre két Poisson keveréke esetén:

$$P(\xi_i = n) = p \cdot e^{-\lambda_1} \cdot \frac{\lambda_1^n}{n!} + (1 - p) \cdot e^{-\lambda_2} \cdot \frac{\lambda_2^n}{n!}, \quad n \geq 0 \quad (1.1)$$

illetve három keverékét nézve:

$$P(\xi_i = n) = p \cdot e^{-\lambda_1} \cdot \frac{\lambda_1^n}{n!} + q \cdot e^{-\lambda_2} \cdot \frac{\lambda_2^n}{n!} + (1 - p - q) \cdot e^{-\lambda_3} \cdot \frac{\lambda_3^n}{n!}, \quad n \geq 0,$$

ahol az i -edik év rengéseinek számát jelöli ξ_i .

Ezt követően lemondunk a függetlenség feltételezéséről. A keverék modellek egyfajta általánosítását tekintettük, ahol az egymást követő évek rengésszámai, mint valószínűségi változók, függenek egymástól. Egy olyan, kétállapotú modellben végeztünk paraméterbecslést, amelyben – az állapotokhoz kölcsönösen egyértelműen hozzárendelt – intenzitás paraméterek egy általunk nem megfigyelhető, stacionárius Markov-lánc szerint változnak. Feltételeztük, hogy az intenzitás paraméterek ismeretében az egyes évek rengései feltételesen független, Poisson-eloszlású valószínűségi változók, és ez a feltételes eloszlás csak az adott év állapotától függ.

A modell struktúra tehát a következő [20, 23, 24]: a diszkrét idejű sztochasztikus modellt egy rejtett, meg nem figyelhető $\{\mathbb{Z}_i\}$ Markov-lánc és egy megfigyelhető $\{\xi_i\}$ folyamat alkotja, melyekre

1. ha \mathbb{Z}_i adott, akkor ξ_i -k feltételesen függetlenek, valamint
2. ha \mathbb{Z}_i adott, ξ_k feltételes eloszlása csak \mathbb{Z}_k -tól függ minden k esetén.

Esetünkben ξ_k -k feltételes eloszlása Poisson, \mathbb{Z}_k -k pedig $\{0, 1\}$ értékkészletű valószínűségi változók. Ezt a modellt kétállapotú rejtett Markov-modellnek vagy keverék Markov-modellnek hívjuk (Hidden Markov Model, Markov Mixture Model, speciális esetben Poisson Hidden Markov Model).

Ebben a modellben az intenzitás paramétereket és az átmenet valószínűségeket becsültük.

2. Az adatok

Az elmúlt 1600 év magyarországi rengéseinek rendezése, az utórezgések kiszűrése erre a célra készített számítógépes programmal, az adatok vizsgálata pedig az R programcsomag segítségével történt. (Megjegyzés: az északi szélesség 45,5 – 49 fok és a keleti hosszúság 16 – 23 fok közötti területet értettük „Magyarország” alatt.)

Vizsgálatunk során nem az összes ismert rengést tekintettük. Csupán a főrengéseket (mainshock) néztük, kiszűrtük az utórengéseket (aftershock). (Megemlítjük, hogy az előrengések – foreshock – azonosításával egyáltalán nem foglalkoztunk). A fenti fogalmakra univerzálisan elfogadott egzakt definíció nincs [3], bár az utórengések számára vonatkozó Omori-formula (1894), amelyet már említettünk, több mint 100 éves. Így az alábbi munka-definíciót adtuk: nevezzük utórengésnek azt az R rengést, amelyhez találunk olyan rengést, hogy

- legfeljebb 30 nappal R előtt történt,
- epicentruma R epicentrumához „nagyon közel” van és
- magnitúdója R magnitúdójánál nagyobb.

Munka-definíció: az $A(x_1, y_1)$ és $B(x_2, y_2)$ pontot térben „nagyon közelinek” tekintjük, ha

$$\left((x_1 - x_2)^2 + 1,89(y_1 - y_2)^2 \right)^{\frac{1}{2}} \leq \frac{1}{2},$$

ahol az első koordináta hosszúsági, a második a szélességi kört jelenti (fokban).

A definíciónk szerinti utóregéseket töröltük a listából. (Először csak megjelöltük a fenti kritériumnak eleget tevőket, utána töröltük a megjelölteket. Módszerünkkel az utóregések utóregéseit is töröltük. Másik listát eredményezne az azonnali törlés.) A továbbiakban „regés” alatt az utóregések elhagyása után megmaradt regéseket értjük, amelyek legalább 2,9 magnitúdójúak.

Az ennél kisebb magnitúdójú regések esetén minimális az esélye, hogy komoly kár következik be, ugyanis a 3-as magnitúdónál kisebb regések maximális intenzitása – a tapasztalatok szerint – nagyon ritkán éri el a IV-es fokozatot, így komoly károkat nem okoznak. Az intenzitás skála a pusztítás mértékét jelzi 12 fokozat segítségével (I-től XII-ig változik). Ez utóbbi már teljes pusztulást jelent az épített környezetben, az előbbi viszont alig érzékelhető. A Munich Re Group CD-je [1] szerint a kárány várható értéke még egy V-ös intenzitású regés esetén is 0,1%-nál kisebb. Megjegyezzük, hogy több intenzitás skála ismert, a közöttük lévő kapcsolatról áttekintést ad T. Utsu munkája [2].

3. A megfigyelt regések számának és erősségének alakulása az évszázadok során

A regisztrált magyarországi földrengések száma rendkívül gyors növekedést mutat, ennek oka – minden bizonnyal – a technikai eszközök fejlődése. Nézzünk két példát ennek szemléltetésére!

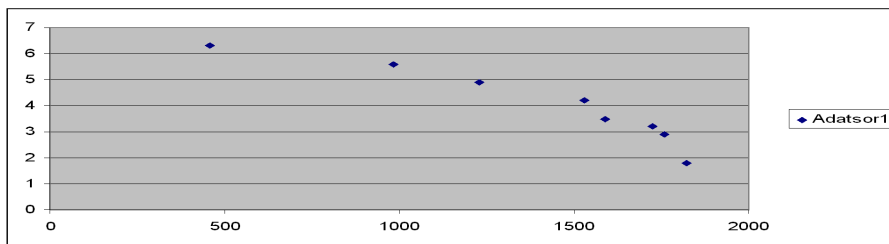
3.1. Csökkenő sorrendben tekintsük, hogy mely évből származik az első regisztráció az egyes magnitúdó értékekből! Vagyis az előzőnél kisebb magnitúdó-érték hol jelenik meg leghamarabb Magyarországon:

1. táblázat.

<i>Magnitúdó</i>	6,3	5,6	4,9	4,2	3,5	3,2	2,9	1,8
<i>Év</i>	456.	984.	1230.	1528.	1590.	1723.	1757.	1822.

Adatainkat az 1. ábrán látható grafikonon is megjelenítettük.

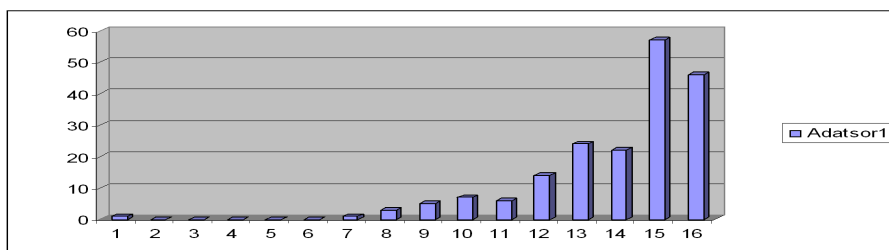
3.2. Nézzük meg, hány olyan rengésről tudunk az alábbi időintervallumokban (50 év), amelyek Magyarországon történtek, magnitúdójuk pedig legalább 4, de kisebb mint 5 (vagyis jelentős regések). (2. táblázat, 2. ábra)



1. ábra. A korábbiaknál kisebb magnitúdójú rengések megjelenése

2. táblázat.

Év	1200-1250	1250-1300	1300-1350	1350-1400
Esemény	1	0	0	0
Év	1400-1450	1450-1500	1500-1550	1550-1600
Esemény	0	0	1	3
Év	1600-1650	1650-1700	1700-1750	1750-1800
Esemény	5	7	6	14
Év	1800-1850	1850-1900	1900-1950	1950-2000
Esemény	24	22	57	46



2. ábra. 4 és 5 magnitúdó közötti erősségű rengések észlelési gyakorisága 50 évenként

Szembeötlő a megfigyelések számának gyors növekedése. Hasonló eredményeket találhatunk Zsíros tanulmányának [17] 5. fejezetében. A legnagyobb változás éppen a XX. század elején történt, amikor közel egy tucat megfigyelő-állomásból álló hálózatot építettek ki Magyarország akkori területén, mellyel az ország az észlelés terén a világ élvonalába került [18].

4. Két lehetséges megközelítés

4.1. Megtehetjük, hogy az évenkénti (magyarországi) földrengésszám eloszlásának vizsgálatokor csupán az elmúlt kb. száz év megfigyeléseit vesszük figyelembe, „kidobva” a korábbi rengéseket. E mellett szól, hogy ezek az adatok jóval részletesebbek és megbízhatóbbak a korábbiaknál. Munkánkban ezt az utat választottuk, vállalva rengeteg adat mellőzését.

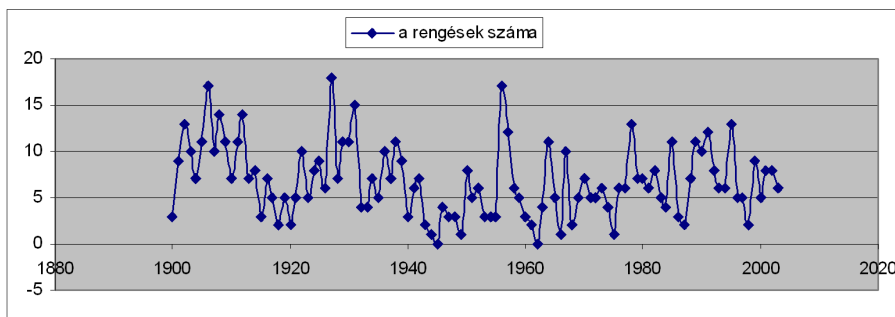
4.2. A másik út, hogy speciális statisztikai módszerekkel „összehasonlíthatóvá” tesszük a különböző korok megfigyeléseit. A kisebb magnitúdójú rengésekről csak nagyon hiányosak az adataink. Így az elmúlt 1600 év összes – utórengések nélküli – eseményét tekintve minden évben egy bizonyos „küszöb” (threshold) feletti adatokkal kell dolgoznunk [4], amikor az éves rengésszámot vizsgáljuk, illetve amikor egy-egy hipotézist ellenőrizzük. Ráadásul a feltételezett „küszöb” – részben társadalmi okokból – területenként és koronként is eltérő. Feltételezzük, hogy a „küszöb” alatti rengéseket „nem látjuk”, róluk semmilyen adatunk nincs. És abban sem lehetünk biztosak, hogy a választott „küszöb” felett minden (megtörtént) rengést ismerünk. A threshold módszert földrengések vizsgálatára többek között Dargahi-Noubary alkalmazta [4], egyszerre becslve egy adott időszakban a földrengések számát és az erősséget jellemző paramétert. Ezt az utat egy másik munkánkban akarjuk bemutatni.

5. A XX. századi adatok vizsgálata, modellek

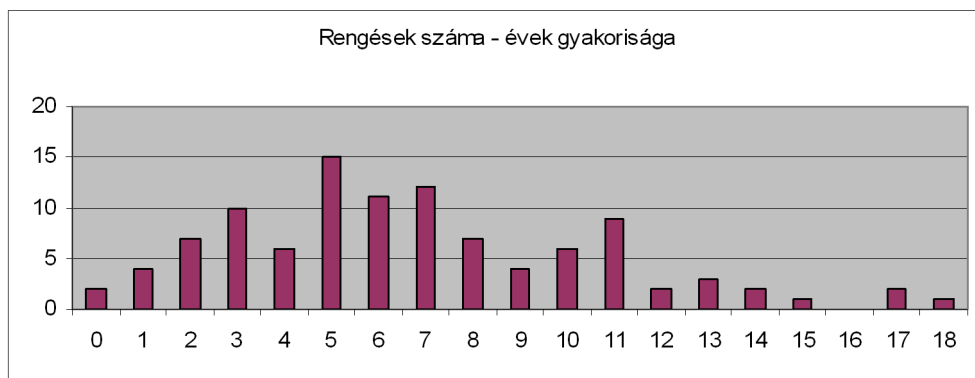
1900-tól 2003-ig az egyes években regisztrált legalább 2,9 magnitúdójú nem-utórengések számát a 3. ábrán mutatjuk meg.

5.1. Első modellünkben feltételeztük, hogy az évenkénti rengések számai: $\xi_{1900}, \xi_{1901}, \dots, \xi_{2003}$ függetlenek és azonos eloszlásúak. Az alábbi grafikon mutatja, hogy 1900 és 2003 között (104 év) hány olyan év volt, amikor a földrengések száma 0, 1, 2, \dots , 17, 18 volt. (4. ábra)

Az évenkénti rengések száma eloszlásának azonosítása során próbálkozhatunk először a Poisson-, negatív binomiális, illetve binomiális eloszlással, lásd például Kagan és Jackson cikkét [5]. Megjegyezzük, hogy ezek az eloszlások a biztosítás-



3. ábra. Rengésszámok 1900 és 2003 között (az utóregések nélkül)



4. ábra. Az évenkénti rengések számainak gyakoriságai

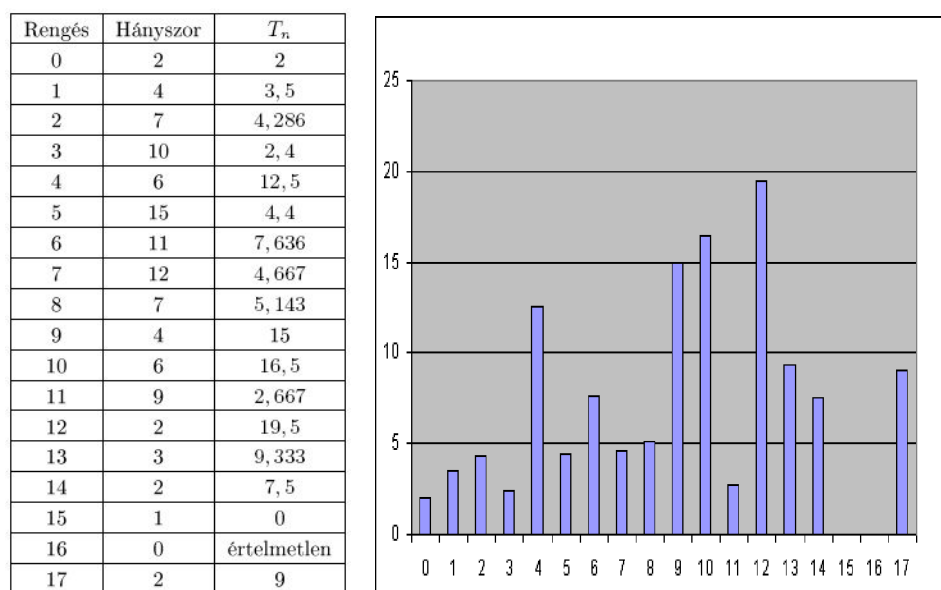
matematikában gyakran alkalmazott $(a, b, 0)$ eloszlások [6], melyekre

$$P(\xi = n) = (a + b/n) \cdot P(\xi = n - 1), \quad n = 1, 2, 3, \dots$$

Megnéztük, hogy adataink ilyen típusúak-e. K_n -nel jelöltük azoknak az éveknek a számát, amikor n rengés volt ($n = 0, 1, 2, \dots$), és a

$$T_n = (n + 1) \cdot K_{n+1}/K_n \quad (n = 0, 1, 2, \dots)$$

mennyiségeket is tekintettük (ahol értelmezhető volt). Így a következő adatsorokat kaptuk, a T_n értékeket pedig grafikonon is ábráztuk:



5. ábra. A T_n értékek

Mivel a T_n sorozat nem (kvázi) monoton növekedő, nem monoton csökkenő és közel állandónak sem mondható, ezért a fent említett három eloszlást rögtön elvetettük [6]. Ellenőriztük mégis azt a hipotézist, hogy ξ_i ($i = 1900, 1901, \dots, 2003$) független, Poisson-eloszlású (azonos paraméterrel) χ^2 -próbával is. Ekkor a

$$P(\xi_i = n) = \lambda^n \cdot e^{-\lambda}/n!$$

hipotézist elfogadva a paraméterbecslés: $\hat{\lambda} = 706/104 = 6,788$. A 3. táblázat mutatja a kategóriákat, a (tényleges) megfigyelések számát (melyeket $N(n)$ jelöl) és a becsült valószínűségekkel számolt értékeket.

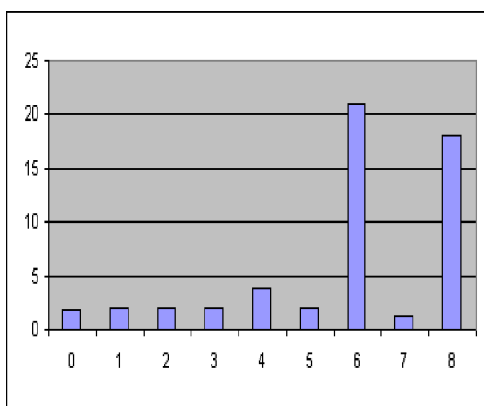
3. táblázat

n	≤ 1	2	3	4	5	6	7	8	9	≥ 10
$N(n)$	6	7	10	6	15	11	12	7	4	26
Becsült értékek	0,91	2,70	6,11	10,36	14,07	15,92	15,44	13,10	9,88	15,50

A becült paraméterek száma 1, így a szabadsági fok $10 - 1 - 1 = 8$. A χ^2 statisztika értéke 55,35. A hipotézist elvetjük, mert a χ^2 eloszlás kvantilise 99%-os szinten 20,09 (a szignifikancia szint közel 1).

Megvizsgáltuk a gyakoriságokat abban az esetben is, amikor a magnitúdó legalább 3,5, illetve legalább 4. Az alábbi, összesen 263 rengést feldolgozó adatsorokat kaptuk a 3,5-ös küszöb esetén. A T_n értékeket grafikonon is ábrázoltuk:

Rengés(n)	Gyakoriság(K_n)	T_n
0	14	1,7857
1	25	1,92
2	24	2
3	16	2
4	8	3,75
5	6	2
6	2	21
7	6	1,3333
8	1	18
10	0	

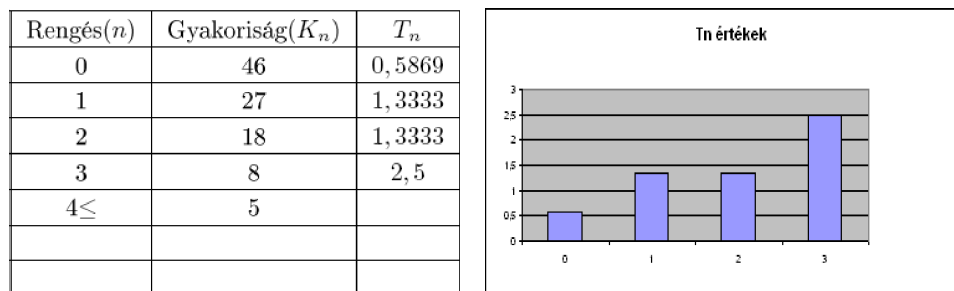


6. ábra. A T_n értékek a legalább 3,5 magnitúdójú rengések esetén

A 7. ábrán látható adatokat pedig a 4 magnitúdós küszöbnél kaptuk (109 rengés).

Az első esetben a három említett eloszlás nem kerülhet szóba. A második esetben a növekedés negatív binomiális eloszlást sejtet, amit megerősít, hogy a várható érték becslése (1,048) kisebb, mint a szórásnégyzet becslése (1,22²) [6]. Korábbi földrengésszám vizsgálatok miatt is erre az eloszlásra gondolhatunk [5]. 5 kategóriát alkalmazva a szabadsági fok 2, a χ^2 -statisztika értéke 8,1 lett. A hipotézist elvetjük, mert a χ^2 eloszlás kvantilise 95%-os szinten 5,99 (a szignifikancia szint 98,26%).

5.2. Legteljesebb (legalább 2,9 magnitúdójú rengéseket tartalmazó) adatsorunkat vizsgáljuk ismét. Az előző eredmények miatt (mivel az illeszkedések nem voltak megfelelők) az eloszlást 2 Poisson keverékével közelítettük. Feltételeztük, hogy az évenkénti rengések számai: $\xi_{1900}, \xi_{1901}, \dots, \xi_{2003}$ függetlenek és az (1.1) képletnek megfelelő azonos eloszlásúak.



7. ábra. A T_n értékek a legalább 4 magnitúdójú rengések esetén

A likelihood függvényt az R programcsomag segítségével, közelítéssel maximalizáltuk. A közelítés a következő eredményt adta: $\hat{p} = 0,56$, $\hat{\lambda}_1 = 4,3657$, $\hat{\lambda}_2 = 9,88$. A χ^2 statisztika értéke ebben az esetben 9,129, tehát az illeszkedés nagyon jó, a hipotézist elfogadhatjuk. (A kategóriák száma 10, a becsült paramétereké 2, így a szabadsági fok 7. A 7 szabadsági fokú χ^2 eloszlás eloszlásfüggvénye a 9,129 helyen 75,65 %).

Ezután nem meglepő, hogy a több paramétert tartalmazó 3 Poisson-eloszlású valószínűségi változó keverékével dolgozó modell még jobb illeszkedést mutatott. Csak szemléltetésül közöljük a maximum likelihood becslés eredményeit: $\hat{p} = 0,1033$, $\hat{q} = 0,5915$, $\hat{\lambda}_1 = 1,8528$, $\hat{\lambda}_2 = 5,5832$, $\hat{\lambda}_3 = 10,7964$.

5.3. Az egyes évek (utóregések nélküli) rengéseinek számát függetlennek tekintve azt tapasztaltuk, hogy a keverék Poisson-eloszlás igen jól illeszkedik az adatainkra. Mégis úgy gondoljuk (alapvetően a földrengésekkel foglalkozó külföldi irodalom alapján), hogy vizsgálnunk kell olyan modellt is, amelyben az egymást követő évek rengéseinek számai kapcsolatban állnak egymással, függenek egymástól.

Olyan modellt tekintünk, amelyben az egyes évek megfigyeléseinek (rengéseinek) számai az állapotok ismeretében feltételesen független, Poisson-eloszlású valószínűségi változók, az állapotok viszont Markov-láncot alkotnak. Így egy rejtett Markov-láncot kell vizsgálni.

6. Rejtett Markov-modell

6.1. Modell struktúra. A következő modellt építjük fel. Legyenek $\mathbb{Z}_1, \mathbb{Z}_2, \dots, \mathbb{Z}_N$ indikátor valószínűségi változók. \mathbb{Z}_m ($m = 1, 2, \dots, N$)-re úgy tekintünk, mint az m -edik „év” állapotára. Legyen λ_0 a 0 állapot, λ_1 pedig az 1-es állapot paramétere ($\lambda_0, \lambda_1 > 0$). Tegyük fel, hogy az állapotok homogén Markov-láncot alkotnak:

$$P(\mathbb{Z}_{m+1} = j | \mathbb{Z}_m = i, \mathbb{Z}_{m-1} = z_{m-1}, \dots, \mathbb{Z}_1 = z_1, \mathbb{Z}_0 = z_0) = \\ = P(\mathbb{Z}_{m+1} = j | \mathbb{Z}_m = i) = p_{ij}, \quad i, j \in \{0, 1\}, m \in \{0, 1, \dots, N-1\}.$$

Így az átmenet-valószínűség mátrix:

$$\Pi = (p_{ij})_{i,j \in \{0,1\}} = \begin{bmatrix} p_{00} & 1 - p_{00} \\ p_{10} & 1 - p_{10} \end{bmatrix}.$$

Jelölje ξ_m az m -edik év rengéseinek (megfigyeléseinek) számát. Feltesszük, hogy $\xi_1, \xi_2, \dots, \xi_N$ feltételesen függetlenek a $\mathbb{Z}_1 = z_1, \mathbb{Z}_2 = z_2, \dots, \mathbb{Z}_N = z_N$ feltételre vonatkozóan, továbbá ξ_m feltételes eloszlása Poisson, és csak \mathbb{Z}_m -től függ minden m esetén ($m \in \{1, 2, \dots, N\}$), vagyis

$$P(\xi_m = k_m | \mathbb{Z}_1 = z_1, \mathbb{Z}_2 = z_2, \dots, \mathbb{Z}_m = i, \dots, \mathbb{Z}_N = z_N) = \frac{\lambda_i^{k_m}}{k_m!} e^{-\lambda_i}$$

és

$$P(\xi_1 = k_1, \xi_2 = k_2, \dots, \xi_N = k_N | \mathbb{Z}_1 = z_1, \dots, \mathbb{Z}_N = z_N) = \prod_{j=1}^N \frac{\lambda_{z_j}^{k_j}}{k_j!} e^{-\lambda_{z_j}}$$

Látható, hogy az (1.1) képlettel meghatározott keverék Poisson-modell jelen modellünk speciális esete a $\mathbf{p}_{00} = \mathbf{p}_{10} = p$ paraméterválasztással. Célunk megbecsülni a $\mathbb{Z}_1, \mathbb{Z}_2, \dots, \mathbb{Z}_N$ értékeket, λ_0 és λ_1 -et, valamint a \mathbf{p}_{00} és \mathbf{p}_{10} átmenet-valószínűségeket.

6.2. Bayes-i megközelítés. Bayes-i megközelítéssel dolgozunk. Feltesszük, hogy paramétereink a priori eloszlása a következő:

$$\mathbf{p}_{00} \sim E(0, 1), \quad \mathbf{p}_{10} \sim E(0, 1), \quad \lambda_0 \sim \Gamma(\alpha_0, \beta_0), \quad \lambda_1 \sim \Gamma(\alpha_1, \beta_1).$$

Adottnak feltételezzük az α_0, β_0 és α_1, β_1 paramétereket.

A Bayes-formulából

$$f(y|\underline{x}) = \frac{f(\underline{x}|y) \cdot f(y)}{\int_A f(\underline{x}|y) f(y) dy} \alpha f(\underline{x}|y) \cdot f(y),$$

ahol $A = \{\text{az összes lehetséges } y \text{ érték}\}$. Itt az α a Bayes-i statisztikában szokásos jelölés, azaz azt jelenti, hogy a két mennyiség konstans szorzótól eltekintve egyenlő. Ezért most

$$f_{\mathbf{p}_{00}, \mathbf{p}_{10}, \lambda_0, \lambda_1 | \underline{\xi}}(p_{00}, p_{10}, \lambda_0, \lambda_1 | \underline{k}) \alpha \\ \alpha P(\underline{\xi} = \underline{k} | \mathbf{p}_{00} = p_{00}, \mathbf{p}_{10} = p_{10}, \lambda_0 = \lambda_0, \lambda_1 = \lambda_1) \cdot f_{\mathbf{p}_{00}, \mathbf{p}_{10}, \lambda_0, \lambda_1}(p_{00}, p_{10}, \lambda_0, \lambda_1).$$

Ennek a függvénynek a maximumhelyét, illetve maximumhelyeit keressük.

Megjegyzés. $\underline{\xi} = \underline{k}$ feltétel a $\xi_1 = k_1, \xi_2 = k_2, \dots, \xi_N = k_N$ feltételt rövidíti.

Amennyiben a $\mathbf{p}_{00}, \mathbf{p}_{10}, \lambda_0, \lambda_1$ valószínűségi változók függetlenségét feltételezzük, akkor a fentiek így írhatók:

$$P(\underline{\xi} = \underline{k} | \mathbf{p}_{00} = p_{00}, \mathbf{p}_{10} = p_{10}, \lambda_0 = \lambda_0, \lambda_1 = \lambda_1) \cdot \chi_{\{p_{00} \in [0,1]\}} \cdot \chi_{\{p_{10} \in [0,1]\}} \cdot \lambda_0^{\alpha_0 - 1} \cdot e^{-\beta_0 \cdot \lambda_0} \cdot \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} \cdot \lambda_1^{\alpha_1 - 1} \cdot e^{-\beta_1 \cdot \lambda_1} \cdot \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} \alpha$$

$$\alpha P(\underline{\xi} = \underline{k} | \mathbf{p}_{00} = p_{00}, \mathbf{p}_{10} = p_{10}, \lambda_0 = \lambda_0, \lambda_1 = \lambda_1) \cdot \lambda_0^{\alpha_0 - 1} \cdot \lambda_1^{\alpha_1 - 1} \cdot e^{-\beta_0 \lambda_0 - \beta_1 \lambda_1} = g(p_{00}, p_{10}, \lambda_0, \lambda_1).$$

Ha pedig $\mathbb{Z}_1, \mathbb{Z}_2, \dots, \mathbb{Z}_N$ értékeket is becsülni kívánjuk:

$$\begin{aligned} f_{\mathbb{Z}_1, \mathbb{Z}_2, \dots, \mathbb{Z}_N, \mathbf{p}_{00}, \mathbf{p}_{10}, \lambda_0, \lambda_1 | \underline{\xi}}(z_1, z_2, \dots, z_N, p_{00}, p_{10}, \lambda_0, \lambda_1 | \underline{k}) &= \\ &= P(\underline{\xi} = \underline{k} | \mathbb{Z}_1 = z_1, \mathbb{Z}_2 = z_2, \dots, \mathbb{Z}_N = z_N, \\ &\quad \mathbf{p}_{00} = p_{00}, \mathbf{p}_{10} = p_{10}, \lambda_0 = \lambda_0, \lambda_1 = \lambda_1) \cdot \\ &\quad \cdot P(\mathbb{Z}_1 = z_1, \mathbb{Z}_2 = z_2, \dots, \mathbb{Z}_N = z_N | \mathbf{p}_{00} = p_{00}, \\ &\quad \mathbf{p}_{10} = p_{10}, \lambda_0 = \lambda_0, \lambda_1 = \lambda_1) \cdot \\ &\quad \cdot f_{\mathbf{p}_{00}}(p_{00}) \cdot f_{\mathbf{p}_{10}}(p_{10}) \cdot f_{\lambda_0}(\lambda_0) \cdot f_{\lambda_1}(\lambda_1) / P(\underline{\xi} = \underline{k}) = \\ &= \frac{g(z_1, z_2, \dots, z_N, p_{00}, p_{10}, \lambda_0, \lambda_1)}{P(\underline{\xi} = \underline{k})} \alpha g(z_1, z_2, \dots, z_N, p_{00}, p_{10}, \lambda_0, \lambda_1), \end{aligned}$$

ahol

$$f_{\mathbf{p}_{00}}(p_{00}) = \chi_{\{p_{00} \in [0,1]\}}, \quad f_{\mathbf{p}_{10}}(p_{10}) = \chi_{\{p_{10} \in [0,1]\}}$$

az $f_{\lambda_0}(\lambda_0)$ és az $f_{\lambda_1}(\lambda_1)$ függvény a gamma eloszlás sűrűségfüggvénye, a $P(\underline{\xi} = \underline{k})$ csupán egy normalizáló konstans.

Ennek az $f_{\mathbb{Z}_1, \mathbb{Z}_2, \dots, \mathbb{Z}_N, \mathbf{p}_{00}, \mathbf{p}_{10}, \lambda_0, \lambda_1 | \underline{\xi}}(z_1, z_2, \dots, z_N, p_{00}, p_{10}, \lambda_0, \lambda_1 | \underline{k})$ függvénynek a maximumhelyét keressük. A függvény tényezőit külön-külön vizsgáljuk meg!

$$\begin{aligned} P(\underline{\xi} = \underline{k} | \mathbb{Z}_1 = z_1, \mathbb{Z}_2 = z_2, \dots, \mathbb{Z}_N = z_N, \\ \mathbf{p}_{00} = p_{00}, \mathbf{p}_{10} = p_{10}, \lambda_0 = \lambda_0, \lambda_1 = \lambda_1) = \\ = \prod_{i=1}^N \left[\frac{(\lambda_0 \cdot (1 - z_i) + \lambda_1 \cdot z_i)^{k_i}}{k_i!} e^{-(\lambda_0(1-z_i) + \lambda_1 \cdot z_i)} \right], \end{aligned}$$

valamint

$$\begin{aligned}
 & P(\mathbb{Z}_1 = z_1, \mathbb{Z}_2 = z_2, \dots, \mathbb{Z}_N = z_N | \mathbf{p}_{00} = p_{00}, \mathbf{p}_{10} = p_{10}, \\
 & \quad \quad \quad \lambda_0 = \lambda_0, \lambda_1 = \lambda_1) = \\
 & = P(\mathbb{Z}_N = z_N | \mathbb{Z}_{N-1} = z_{N-1}, \dots, \mathbb{Z}_1 = z_1, \mathbf{p}_{00} = p_{00}, \mathbf{p}_{10} = p_{10}, \\
 & \quad \quad \quad \lambda_0 = \lambda_0, \lambda_1 = \lambda_1) \cdot \\
 & \cdot P(\mathbb{Z}_{N-1} = z_{N-1} | \mathbb{Z}_{N-2} = z_{N-2}, \dots, \mathbb{Z}_1 = z_1, \mathbf{p}_{00} = p_{00}, \mathbf{p}_{10} = p_{10}, \\
 & \quad \quad \quad \lambda_0 = \lambda_0, \lambda_1 = \lambda_1) \cdot \\
 & \dots \cdot P(\mathbb{Z}_2 = z_2 | \mathbb{Z}_1 = z_1, \mathbf{p}_{00} = p_{00}, \mathbf{p}_{10} = p_{10}, \lambda_0 = \lambda_0, \lambda_1 = \lambda_1) \cdot \\
 & \cdot P(\mathbb{Z}_1 = z_1 | \mathbf{p}_{00} = p_{00}, \mathbf{p}_{10} = p_{10}, \lambda_0 = \lambda_0, \lambda_1 = \lambda_1) = \\
 & \quad = p_{00}^{\ell_{00}} \cdot (1 - p_{00})^{\ell_{01}} \cdot p_{10}^{\ell_{10}} \cdot (1 - p_{10})^{\ell_{11}} \cdot \\
 & \cdot P(\mathbb{Z}_1 = z_1 | \mathbf{p}_{00} = p_{00}, \mathbf{p}_{10} = p_{10}, \lambda_0 = \lambda_0, \lambda_1 = \lambda_1),
 \end{aligned}$$

ahol ℓ_{00} jelöli a 0 állapotból 0 állapotba, ℓ_{01} jelöli a 0 állapotból 1-es állapotba, ℓ_{10} az 1-es állapotból 0 állapotba, ℓ_{11} pedig az 1-es állapotból 1-es állapotba történő átmenetek számát.

Megjegyzés. $\ell_{00} + \ell_{01} + \ell_{10} + \ell_{11} = N - 1$.

Megjegyezzük, hogy az utolsó tényező esetén \mathbb{Z}_1 eloszlása már független a feltételben szereplő eseményektől.

6.3. Az a posteriori eloszlások.

6.3.1. Először a λ_0 valószínűségi változó a posteriori eloszlását vizsgáljuk.

$$\begin{aligned}
 f_{\lambda_0 | \underline{\xi}, \underline{\mathbb{Z}}}(\lambda_0 | \underline{k}, \underline{z}) &= \frac{g(z_1, z_2, \dots, z_N, \lambda_0, \lambda_1, p_{00}, p_{10})}{p(\underline{\xi} = \underline{k})} = \\
 &= \frac{1}{P(\underline{\xi} = \underline{k})} \cdot K \cdot f_{\lambda_0}(\lambda_0) \cdot \prod_{i: z_i=0}^{\frac{((1-z_i) \cdot \lambda_0)^{k_i}}{k_i!}} \cdot e^{-\lambda_0},
 \end{aligned}$$

vagyis a nullás állapotokat tekintjük csak. K nem tartalmaz λ_0 -t, ahogy a $\frac{1}{P(\underline{\xi} = \underline{k})}$ és a $k_i!$ tényezők sem, szorzatukat C jelöli. Ekkor

$$f_{\lambda_0 | \underline{\xi}, \underline{\mathbb{Z}}}(\lambda_0 | \underline{k}, \underline{z}) = C \cdot f_{\lambda_0}(\lambda_0) \cdot \lambda_0^{\sum_{i: z_i=0} k_i} \cdot e^{-\lambda_0 \cdot \sum_{i=1}^N (1-z_i)}.$$

Mivel $\lambda_0 \sim \Gamma(\alpha, \delta)$, és így sűrűségfüggvénye

$$f_{\lambda_0}(\lambda_0) = \frac{\delta^\alpha}{\Gamma(\alpha)} \cdot \lambda_0^{\alpha-1} \cdot e^{-\delta \cdot \lambda_0} \quad (\lambda_0 > 0),$$

akkor

$$\begin{aligned} f_{\lambda_0|\underline{\xi}, \underline{Z}}(\lambda_0|\underline{k}, \underline{z}) &= h(\lambda_0) = C \cdot \frac{\delta^\alpha}{\Gamma(\alpha)} \cdot \lambda_0^{\alpha-1+\sum_{i:z_i=0} k_i} \cdot e^{-\lambda_0\left(\delta+\sum_{i=1}^N 1-z_i k_i\right)} = \\ &= B \cdot \lambda_0^{\alpha-1+\sum_{i:z_i=0} k_i} \cdot e^{-\lambda_0\left(\delta+\sum_{i=1}^N (1-z_i)\right)}. \end{aligned}$$

Tehát azt kaptuk, hogy λ_0 a posteriori eloszlása:

$$\lambda_0|\underline{\xi} = \underline{k}, \underline{Z} = \underline{z} \sim \Gamma\left(\alpha + \sum_{i:z_i=0} k_i, \delta + \sum_{i=1}^N (1-z_i)\right).$$

Ugyanezt mondhatjuk el λ_1 a posteriori eloszlásáról (α' és δ' , illetve $\alpha' + \sum_{i:z_i=1} k_i$ és $\delta' + \sum_{i=1}^N z_i$ paraméterrel.) [22]

6.3.2. Áttérünk a \mathbf{p}_{00} valószínűségi változó a posteriori eloszlásának vizsgálatára:

$$\begin{aligned} f_{\mathbf{p}_{00}|\underline{\xi}, \underline{Z}}(p_{00}|\underline{k}, \underline{z}) &= \frac{g(z_1, z_2, \dots, z_N, \lambda_0, \lambda_1, p_{00}, p_{10})}{P(\underline{\xi} = \underline{k})} = \\ &= C \cdot f_{\mathbf{p}_{00}}(p_{00}) \cdot p_{00}^{\ell_{00}} \cdot (1-p_{00})^{\ell_{01}}, \end{aligned}$$

Megkaptuk, hogy az a posteriori eloszlás $\beta(\ell_{00} + 1, \ell_{01} + 1)$. Hasonlóan adható meg \mathbf{p}_{01} a posteriori eloszlása is, amely $\beta(\ell_{10} + 1, \ell_{11} + 1)$ lesz.

6.4. MCMC (Markov Chain Monte Carlo) módszer alkalmazása.

A maximumhely keresést az MCMC (Markov Chain Monte Carlo) módszerrel végzzük (lásd például a [21] összefoglaló cikket).

1. $\mathbb{Z}_1, \mathbb{Z}_2, \dots, \mathbb{Z}_N$ értékeit $\{0, 1\}$ -en diszkrét egyenletes eloszlásból, \mathbf{p}_{00} és \mathbf{p}_{10} értékét $E(0; 1)$ eloszlásból generáljuk, λ_0 -at $\Gamma\left(3; \frac{1}{1.5}\right)$, λ_1 -et $\Gamma\left(5; \frac{1}{2}\right)$ eloszlásból választjuk ki. Az így kapott értékek lesznek iterációs eljárásunk kezdőértékei.

(Bár választásainkat a paraméterek esetén „indokolni” nem tudjuk, a következő megfontolásból születtek: a keverék Poisson-modell esetén a várható értékek becslései $\hat{\lambda}_1 = 4,38$ és $\hat{\lambda}_2 = 9,88$ voltak; most pedig $3 \cdot \frac{1}{1.5} = 4,5$ és $5 \cdot \frac{1}{2} = 10$ a fenti értékekhez közeleiek.)

2. A $\{1, 2, \dots, N\}$ -ből visszatévéses mintavétellel kiválasztjuk i_1, i_2, \dots, i_N -t. Jelölje $g^*(i_1)$ azt a függvényértéket, melyet úgy kapunk, hogy z_{i_1} -et $1 - z_{i_1}$ -re változtatjuk, a többi argumentumot változatlanul hagyjuk.

Ha $\frac{g^*(i_1)}{g(z_1, \dots, z_N, \lambda_0, \lambda_1, p_{00}, p_{00})} > 1$, akkor z_{i_1} -et $1 - z_{i_1}$ -re változtatjuk.

Ha $\frac{g^*(i_1)}{g(z_1, z_2, \dots, z_N, \lambda_0, \lambda_1, p_{00}, p_{10})} = a < 1$, akkor „ a ” valószínűséggel kicseréljük, $1 - a$ valószínűséggel meghagyjuk z_{i_1} értékét.

Hasonlóan járunk el i_2, i_3, \dots, i_N esetén is. (Ez a lépés a Metropolis–Hastings algoritmusnak felel meg.)

3. Új értéket generálunk λ_0 a posteriori eloszlásából, jelölje ezt λ_0^* . λ_0 -t λ_0^* -ra változtatjuk, a többi argumentumot változatlanul hagyjuk.

4. Hasonlóan járunk el az aktuális λ_1, p_{00} és p_{10} értékkel szemben is. (A 3. és 4. lépést Gibbs-lépésnek nevezzük.)

5. A 2–4. pontban leírtakat megfelelően sokszor – néhány tízezerszer – megismételjük.

6. A λ_0 és λ_1 , valamint a p_{00} és p_{10} becslése az algoritmus során használt értékek átlaga lesz. (Az első néhány ezer adatot – az iterációk mintegy tizedrészét – nem vesszük figyelembe.)

7. A Z_i ($i = 1, 2, 3, \dots, N$) becslése az az állapot, melyben a rendszer az algoritmus során többször tartózkodott.

Az iterációs becslési eljárás elvégzéséhez programot dolgoztunk ki, amelynek alkalmazásával a következő eredményeket kaptuk:

5 000 figyelembe nem vett iteráció után 100 000 iterációt végeztünk. Az alábbi táblázatban láthatóak becsléseink. Az „összesített” becslés az első („Átlag”) sorban látható, a további sorok 100–100 iteráció során kapott átlagot mutatják (vagyis csupán 1 200 iteráció eredményét tartalmazza az 1–12. sor, az első viszont 100 000-ét). Az utolsó két oszlop azt jelzi, hogy az utolsó 2 évben az iterációk hány százalékában volt az egyes állapotban a rendszer.

	10	11	p00	p10	≥103	≥104
Átlag	4,182047578	9,595539041	0,732472625	0,272438237	73,505	53,007
1.	4,051910094	9,457883732	0,701145060	0,281892006	68	57
2.	4,404168520	9,884150196	0,759503698	0,313682174	68	42
3.	4,123858765	9,403907699	0,719902759	0,247088922	61	57
4.	4,164482130	9,424047812	0,746502007	0,235290183	92	70
5.	4,305467640	9,759515695	0,743700170	0,282466375	72	51
6.	4,122284966	9,526731139	0,716296275	0,283679996	86	48
7.	4,035598721	9,356752326	0,730958985	0,251609114	88	69
8.	4,064423719	9,425677438	0,702309667	0,264378470	83	56
9.	4,156852810	9,540712504	0,745310557	0,257037332	87	65
10.	4,137226495	9,399213003	0,749120459	0,234497373	82	71
11.	3,920489485	9,383593921	0,671171453	0,269699750	79	50
12.	3,925001458	9,083686216	0,699076408	0,235576461	99	77

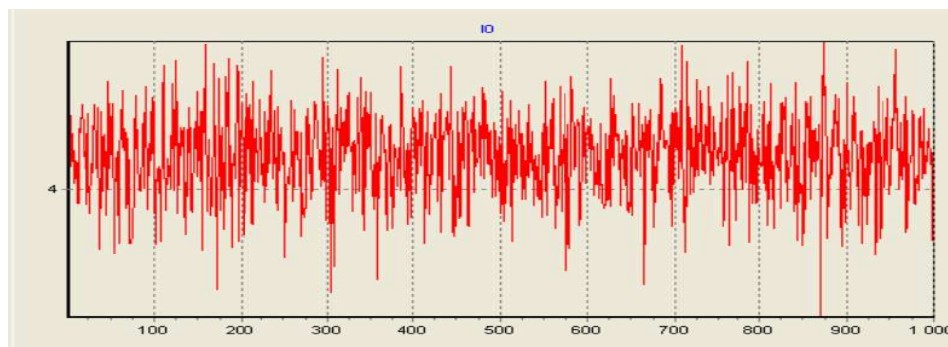
Megjegyezzük, hogy az előbbivel megegyező feltételekkel történő másik két futtatás eredménye (a fenti sorrendben):

λ_0 becslése	λ_1 becslése	p_{00} becslése	p_{10} becslése	Z_{103} becslése	Z_{104} becslése
4,2667	9,5401	0,2717	0,7385	72,29	51,81
4,1848	9,5976	0,2727	0,7333	73,64	53,06

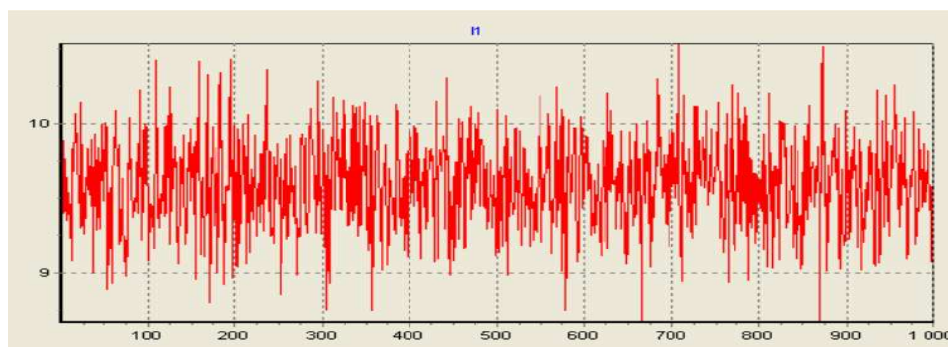
Látható, hogy p_{00} és p_{10} becslései igen távol állnak egymástól és a futási eredmények alapján elutasítható az a hipotézis, hogy egyenlőek (már említettük, hogy ez felelne meg a keverék Poisson-eloszlás (1.1) modelljének).

λ_0 és λ_1 a priori eloszlásának paramétereit megváltoztatva hasonló becsléseket kaptunk.

Az iterációkat 100-asával csoportosítva λ_0 és λ_1 becslült értékeit ábrázolja az alábbi két grafikon:

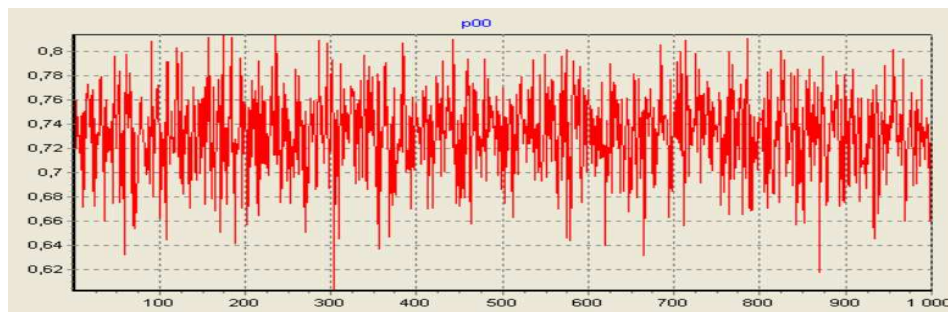


8. ábra. A λ_0 becslései az iterációkat százával csoportosítva

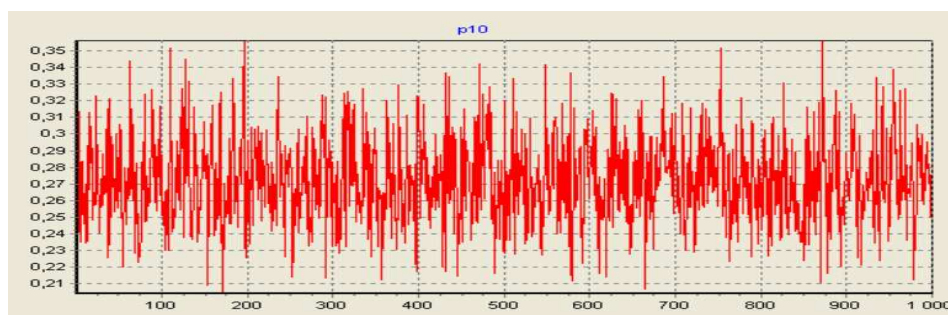


9. ábra. A λ_1 becslései az iterációkat százával csoportosítva

Az átmenet valószínűségek becslése pedig a következőképpen alakult \mathbf{p}_{00} , illetve \mathbf{p}_{10} esetén:



10. ábra. A \mathbf{p}_{00} becslései az iterációkat százasaival csoportosítva



11. ábra. A \mathbf{p}_{10} becslései az iterációkat százasaival csoportosítva

7. A becslések eredményeinek összehasonlítása, előrejelzések

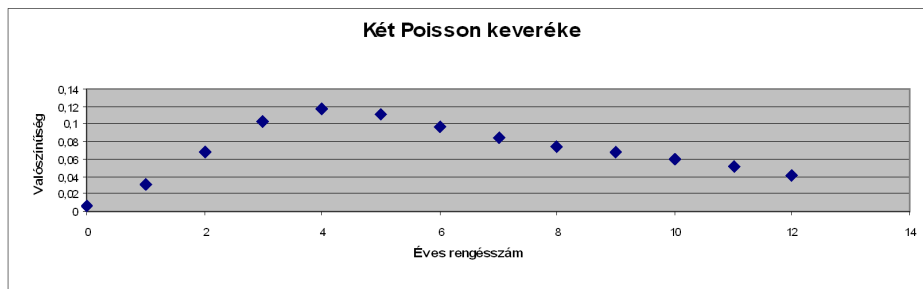
7.1. Két Poisson-eloszlású valószínűségi változó keverékének eloszlása esetén a becslésekre a következők adódtak: $\hat{p} = 0,56$, $\hat{\lambda}_1 = 4,3657$, $\hat{\lambda}_2 = 9,88$. A 12. ábra mutatja a becsült paraméterekkel a rengésszám eloszlását.

Adatainkat grafikonon is ábrázoltuk.

Látható, hogy négy rezgésre számíthatunk legnagyobb eséllyel, illetve például annak esélye, hogy legalább 10 rengés lesz egy év során 0,2396 e modell szerint. Az illesztett eloszlás mediánja 6.

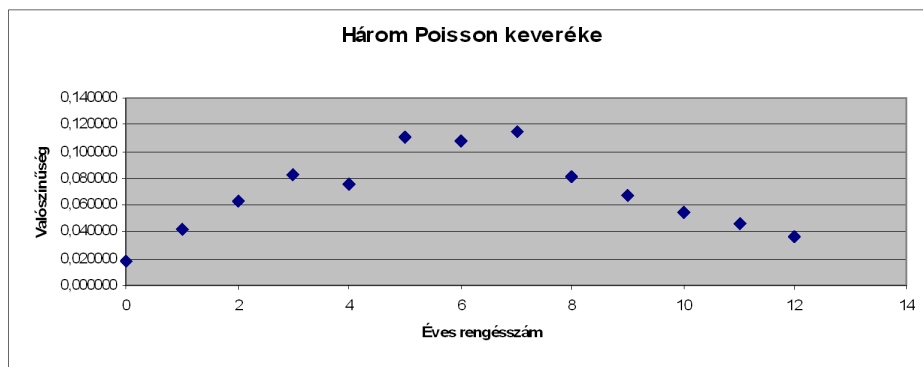
7.2. Három Poisson-eloszlású valószínűségi változó keveréke esetén $\hat{p} = 0,1033$, $\hat{q} = 0,5915$, $\hat{\lambda}_1 = 1,8528$, $\hat{\lambda}_2 = 5,5832$, $\hat{\lambda}_3 = 10,7964$ adódott. Az eloszlást és a hozzá tartozó grafikon a 13. ábra mutatja.

Gyakoriság	0	1	2	3	4	5	6
Valószínűség	0,0071	0,0313	0,0689	0,1023	0,1166	0,1117	0,0975
Gyakoriság	7	8	9	10	11	12	> 12
Valószínűség	0,0837	0,074	0,067	0,0599	0,0514	0,0414	0,0872



12. ábra. Két Poisson keverékének eloszlása a becsült paraméterekkel

Gyakoriság	0	1	2	3	4	5	6
Valószínűség	0,018422	0,042497	0,062834	0,083002	0,076011	0,111148	0,108231
Gyakoriság	7	8	9	10	11	12	> 12
Valószínűség	0,115015	0,080752	0,066643	0,055092	0,045506	0,036976	0,097869



13. ábra. Három Poisson keverékének eloszlása a becsült paraméterekkel

A legnagyobb valószínűséggel hét rengés történik évente a modell szerint. (Öt és hat rengésnek is szinte ugyanennyi az esélye.) Az éves rengésszám mediánjának becslése e modellben is 6. Legalább 10 rengés lesz 0,2354 valószínűséggel, így várhatóan kb. 4 évente lesz legalább 10 rengés.

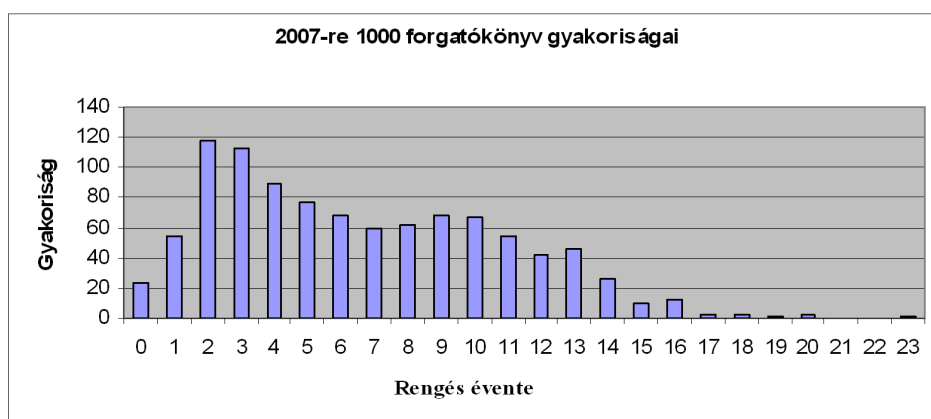
7.3. Az MCMC módszer egyik előnye, hogy segítségével a következő évek történéseit is szimulálni tudjuk. A becslési iteráció minden (vagy például minden 100.) lépésénél a pont aktuális paraméterekkel a feladatnak megfelelő számú év eredményét sztochasztikusan generáljuk. Ezáltal a módszer alapja lehet más vizsgálatoknak is (például földrengések modellezésének).

Az alábbi táblázat soraiban egy-egy előrejelzés látható 2004-től 2023-ig. A paraméterek becslése után állapotokat generáltunk 20 évre a megfelelő átmenetvalószínűségekkel, majd az állapothoz tartozó (szintén becsült) paraméterrel Poisson-eloszlásból adatokat generáltunk:

5	11	6	9	12	5	3	0	4	3	11	6	3	11	2	11	2	4	12	2
9	6	4	4	3	0	6	4	9	5	8	4	7	11	1	3	4	16	10	3

A 20 év rengésszámainak generálását 1000-szer végeztük el. A 2007-es évre (amely a negyedik generált adatot jelenti) az alábbiakat mondhatjuk:

7.3.1. A rengésszám gyakorisága a következőnek adódott 1000 forgatókönyv vizsgálata során:



14. ábra. A 2007-es évre generált adatok gyakorisága 1000 forgatókönyv esetén

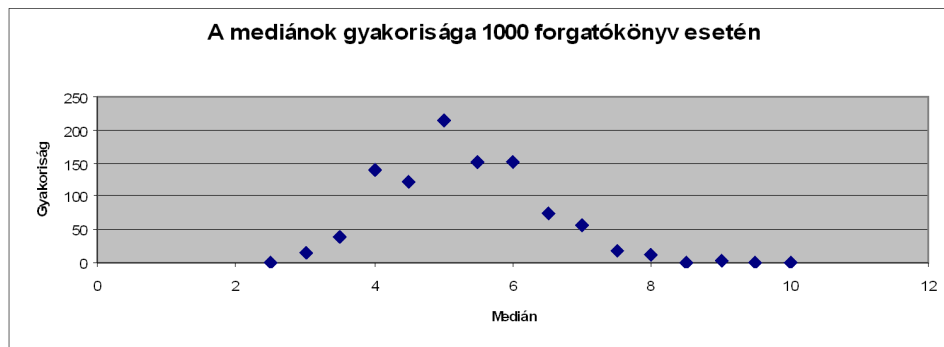
7.3.2. Az említett évre a rengésszám kvantilisére 99%-os szinten 16, 95%-os szinten pedig 13 adódott. Meg kell jegyezni, hogy a keverék Poisson-eloszlású

modellben a becült paraméterekkel nagyon hasonló értékek jönnek ki.

$$P(\xi_i \leq 13) = 94,4\%, \quad P(\xi_i \leq 14) = 96,6\%,$$

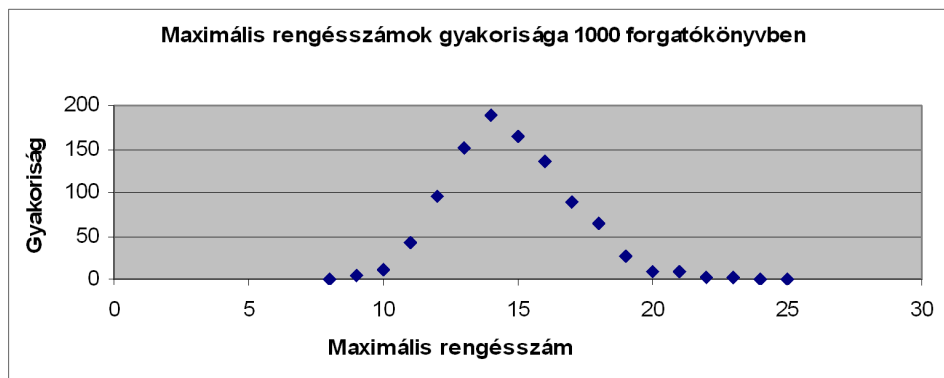
$$P(\xi_i \leq 16) = 98,9\%, \quad P(\xi_i \leq 17) = 99,4\%$$

7.3.3. A 20 év mediánjának gyakorisága a 15. ábrán látható. A legtöbbször az 5, utána 4, 5, 5,5 és 6 jött ki.



15. ábra. A mediánok gyakorisága 1000 forgatókönyv esetén

7.3.4. A 20 év generált rengéseit tartalmazó 1000 forgatókönyvben a maximális rengésszámokat vizsgálva az esetek több mint felében legalább 13 volt a maximális éves rengésszám. Leggyakrabban 14-et kaptunk maximumnak, és nem elhanyagolható (0,05 és 0,1 közötti) azoknak az éveknek a relatív gyakorisága, amikor 17 vagy 18 eseményt generáltunk maximumisan (20 év alatt).



16. ábra. A 20 évenkénti maximális rengésszámok gyakoriságai 1000 forgatókönyv esetén

8. Összefoglalás

Dolgozatunkban megpróbáltuk bemutatni a Magyarországon bekövetkező éves földrengésszámok statisztikai vizsgálatának néhány lehetséges megközelítését. Céljaink közé tartozott, hogy az utóbbi években egyre népszerűbb MCMC-módszer hasznosságát egy érdekes adatsoron szemléltessük.

Megállapítottuk, hogy a leggyakrabban alkalmazott eloszlások nem illeszkednek jól adatainkra. A keverék Poisson-eloszlással a gyakoriságokat megfelelően közelítettük, de a rejtett Markov-modell becslésénél kapott eredmények azt sejtetik, hogy az éves rengésszámok nem függetlenek egymástól.

A kidolgozott programmal a jövőbeli események könnyen szimulálhatók, felhasználhatók például biztosítási számításokhoz. A továbbiakban tervezzük kisebb területek földrengésszámainak leírását tér-idő modellek segítségével is, de úgy gondoljuk, hogy az eddig kapott eredmények segítséget nyújthatnak ezekhez a vizsgálatokhoz is.

Köszönjük Szeidovitz Győző segítségét, akitől több hasznos tanácsot kaptunk a földrengések szakirodalmával kapcsolatban. Köszönjük továbbá a lektor hasznos észrevételeit.

Hivatkozások

- [1] MUNICH RE GROUP: *World of Natural Hazards*. CD (2004)
- [2] UTSU, T.: *Relationships between Magnitude Scales*. In: International Handbook of Earthquake and Engineering Seismology, Chapter 44, Academic Press, (2002)
- [3] UTSU, T.: *Statistical Features of Seismicity*. In: International Handbook of Earthquake and Engineering Seismology, Chapter 43, Academic Press, (2002)
- [4] DARGAHI-NOUBARY, G.R.: *The use of modern statistical theories in assessment of earthquake hazard, with application to quiet regions of eastern North America*. Soil Dynamics and Earthquake Engineering **22** (2002), 361–369.
- [5] KAGAN, Y.Y., JACKSON, D.D.: *Probabilistic earthquake forecasting*. Geophys. J. In. **143** (2000), 438–453.
- [6] ARATÓ, MIKLÓS: *Általános Biztosításmatematika*. ELTE Eötvös Kiadó, Budapest (1997)
- [7] VERE-JONES, D.: *Forecasting earthquakes and earthquake risk*. International Journal of Forecasting **11** (1995), 503–538.
- [8] TÓTH, L., MÓNUS, P., ZSÍROS, T.: *Hungarian Earthquake Bulletin, 1996*. GeoRisk, Budapest (1997)
- [9] TÓTH, L., MÓNUS, P., ZSÍROS, T.: *Hungarian Earthquake Bulletin, 1997*. GeoRisk, Budapest (1998)
- [10] TÓTH, L., MÓNUS, P., ZSÍROS, T.: *Hungarian Earthquake Bulletin, 1998*. GeoRisk, Budapest (1999)

- [11] TÓTH, L., MÓNUS, P., ZSÍROS, T.: *Hungarian Earthquake Bulletin, 1999*. GeoRisk, Budapest (2000)
- [12] TÓTH, L., MÓNUS, P., ZSÍROS, T., KISZELY, M., KOSZTYU, Z.: *Hungarian Earthquake Bulletin, 2000*. GeoRisk, Budapest (2001)
- [13] *Hungarian Earthquake Bulletin 2001*. GeoRisk, Budapest (2002)
- [14] *Hungarian Earthquake Bulletin, 2002*. Georisk, Budapest (2003)
- [15] *Hungarian Earthquake Bulletin, 2003*. Georisk, Budapest (2004)
- [16] ZSÍROS, T., MÓNUS, P., TÓTH, L.: *Hungarian earthquake catalog (456-1986)*. MTA GGKI, Budapest (1988)
- [17] ZSÍROS, T.: *A Kárpát medence szeizmicitása és földrengés veszélyessége: Magyar földrengés katalógus (456-1995)*. MTA GGKI, Budapest (2000)
- [18] TÓTH, L., MÓNUS, P., ZSÍROS, T., KISZELY, M.: *Seismicity in the Pannonian Region – earthquake data, EGU Stephan Mueller Special Publication Series. 3* (2002), 9–28.
- [19] JACKSON, D.D., KAGAN, Y.Y.: *Testable earthquake forecast for 1999*. Seim. Res. Let. **70** (1999), 393–403.
- [20] RYDÉN, T.: *Hidden Markov models*. In Encyclopedia of Actuarial Science, Vol 2 (J. Teugels and B. Sundt, eds.), Wiley (2004), 821–827.
- [21] BROOKS, S.P.: *Markov Chain Monte Carlo Method and Its Application, The Statistician* Vol. **47**, No. **1** (1998), 69–100.
- [22] CHIB, S.: *Calculating posterior distributions and modal estimates in Markov mixture models*. Journal of Econometrics **75** (1996), 79–97.
- [23] PAROLI, R., REDAELLI, G., SPEZIA, L.: *Poisson Hidden Markov Models for Time Series of Overdispersed Insurance Counts*
- [24] EPHRAIM, Y., MERHAV, N.: *Hidden Markov Processes*. IEEE Transactions on Information Theory, Vol. **48**, No. **6** (2002)

(Beérkezett: 2006. június 8.)

KOVÁCS ELŐD

Veszprémi Egyetem

Matematikai és Számítástechnikai Tanszék

8200 Veszprém, Pf. 158.

kovacse@almos.vein.hu

ARATÓ MIKLÓS

Eötvös Loránd Tudományegyetem

Természettudományi Kar, Valószínűségelméleti és Statisztika Tanszék

Alkalmazott Matematikai Lapok (2008)

1117 Budapest, Pázmány Péter s. 1/c
aratonm@ludens.elte.hu

LIPOVITS ÁGNES
Veszprémi Egyetem
Matematikai és Számítástechnikai Tanszék
8200 Veszprém, Pf. 158.
lipovitsa@szt.vein.hu

MODELLING OF ANNUAL FREQUENCY OF EARTHQUAKES IN HUNGARY

ELŐD KOVÁCS, MIKLÓS ARATÓ, ÁGNES LIPOVITS

Some possible approaches of the annual earthquake numbers' statistical investigation happening in Hungary is tried to be presented in the essay. We aimed to demonstrate an interesting series of data the usefulness of the MCMC method, which is getting more and more applied in the last years.

We pointed out that the distributions applied most often do not apt well to our data. We approached the frequencies with the mixed Poisson distribution acceptably, but the results, received by the evaluation of the hidden Markov model, suggest that the annual numbers of frequencies are not independent of one another.

Future events might be easily simulated with the presented program, applying for insurance calculations for example.