



A földrengésből származó károk becslésének lehetőségei

Kovács E., Lipovits Á.

Pannon Egyetem, Matematikai és Számítástechnikai Tanszék, 8200 Veszprém, Egyetem u. 10.

ÖSSZEFOGLALÁS

A cikkünk az évenként magyarországi földrengésszám vizsgálatának néhány lehetőségét mutatja be. Célunk az volt, hogy egy érdekes adatsoron szemléltessük az MCMC módszer hasznosságát, melyet egyre gyakrabban alkalmaznak az utóbbi években. Megmutattuk, hogy a leggyakrabban használt eloszlások nem illeszkednek jól adatainkra. A gyakoriságot elfogadható módon közelítettük keverék-Poisson eloszlással. A rejtett Markov-modellből kapott becslések viszont azt mutatják, hogy az éves földrengésszámok nem függetlenek egymástól. Saját fejlesztésű program segítségével tudjuk az adatokat szűrni, és a következő 20 év forgatókönyveit szimulálni. A program által előállított jövőbeli események például a biztosításmatematikai számításokban is felhasználhatóak.

(Kulcsszavak: földrengés, keverék Poisson-eloszlás, rejtett Markov-modell)

ABSTRACT

Estimation possibilities of damages after earthquakes

E. Kovács, Á. Lipovits

Pannon University, Department of Mathematics and Computing, H-8200 Veszprém, Egyetem u. 10.

In our essay we tried to present some possible approaches of the annual earthquake numbers' statistical investigation, happening in Hungary. We aimed to demonstrate on an interesting series of data the usefulness of the MCMC method, more and more applied in the last years. We pointed out that the most often applied distributions 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 each other. We developed computer application for collecting data and simulation scenarios of earthquake numbers for the next 20 years. Future events might be easily simulated with the presented program, applying for insurance calculations for example.

(Keywords: earthquakes, mixed Poisson distribution, hidden Markov model)

BEVEZETÉS

A biztosítók részéről felmerül az igény, hogy egy bizonyos időegység (például egy év) alatt földrengésből származó kárukat valamilyen módon reálisan megbecsüljék. Ehhez megfelelő adatbázisra, jól dokumentált épületekre van szükség. (Fontos tudni az épületek helyét, korát, típusát, sérülékenységét, bár ezek az adatok még a biztosítóknál regisztrált épületek esetén sem állnak mindig rendelkezésre.)

Megkülönböztetjük a földrengés-veszélyeztettség és a földrengés-kockázat fogalmát. Az előbbi esetén egy adott helyen azt vizsgálják, hogy meghatározott időn – pl.

100 éven – belül mekkora valószínűséggel következik be egy adott magnitúdónál nagyobb rengés. Földrengés–kockázat viszont az anyagi veszteséget méri, így ezt befolyásolja az adott területen lévő épületek száma, értéke, műszaki állapota, illetve az épületek típusa is. Magyarország mindkét szempontból enyhe/közepes „kategóriába” tartozik.

Egy biztosítótársaságot nyilván az utóbbi érdekli. A veszteségek becsléséhez kiindulhatunk az elmúlt évtizedek adataiból (figyelembe véve többek között az inflációt, az új épületek minőségi paramétereit, a biztosítási állomány struktúráját), de ekkor eltekintենnk annak a kockázatától, hogy a korábbiaknál jóval nagyobb erősségű rengés is bekövetkezhet egy nagyvárosban, például Budapesten vagy annak közelében. Egy másik megközelítés, hogy több ezer forгатókönyvet generálunk, ahol egy forгатókönyv egy időegységre (például évre) szól. „Kisorsoljuk” – számítógéppel generáljuk – az év során bekövetkező rengések számát, minden egyes rengés helyét (szélességi és hosszúsági körrel megadva), a hipocentrum mélységét (kilométerben), a rengés erősségét (magnitúdóval vagy maximális intenzitással jellemezve). Majd minden egyes rengés esetén megbecsüljük a veszteségeket, figyelembe véve a biztosított épületek távolságát az epicentrumtól. (Az összes épület kárát nézzük minden rengés esetén, és ezeket összegezzük.) Így több ezer „év” „tapasztalatából” becslés adható a kockázat eloszlásfüggvényére, illetve más jellemzőire, például kvantilisre vagy várható értékére. Ehhez viszont szükséges tudni – sok más mellett vagy előtt –, hogy hány rengésre számíthatunk egy év alatt. Elég az olyan erősségű rengéseket vizsgálni, amelyek a tapasztalatok szerint már okozhatnak károkat. Az éves magyarországi földrengésszámot, mint valószínűségi változót vizsgáljuk, és számítógépes modellezését mutatjuk be. (Megjegyezzük, hogy ebben a modellben a következő rengés helye, mélysége, erőssége stb. is valószínűségi változó lesz.)

ANYAG ÉS MÓDSZER

Program az adatok rendezésére és szűrésére

Az elmúlt 1600 év magyarországi rengéseinek rendezése, az utóregések kiszűrése erre a célra készített számítógépes programokkal történt, az adatok vizsgálata pedig az R–programcsomag valamint a modellekhez írt programok segítségével. (Megjegyezzük, hogy az északi szélesség 45,5- 49 fok és a keleti hosszúság 16-23 fok közötti területet értjük „Magyarország” alatt.) A magyarországi rengésekről (részben interaktív módon) részletes adatokat nyújt a www.georisk.hu, illetve a www.foldrenges.hu honlap. *Zsíros és mtsai.* (1988) valamint *Tóth és mtsai.* (1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004) munkáiban részletes adatsorokat találhatunk. A Pannon–régió, illetve a Kárpát-medence szeizmicitását pedig – több más szerző és tanulmány mellett – *Tóth és mtsai* (2002) cikke, illetve *Zsíros* (2000) tekinti át nagyon részletesen.

Adattisztítást végeztünk, mely során az utóregéseket kiszűrtük. Mindezt azért tettük, hogy a szűrés után megmaradt rengéseket függetlennek tekinthessük, ami a statisztikai vizsgálatokat némileg megkönnyíti. (A rengést követő utóregések erőssége – magnitúdója vagy maximális intenzitása – például szoros kapcsolatban van a főrengéssel.) Az „utórengésre” univerzálisan elfogadott egzakt definíció nincs, bár az utóregések számára vonatkozó Omori–formula (*Omori*, 1894) több mint 100 éves. Így az alábbi munka–definíciót adtuk: nevezzük utórengésnek azt a rengést (R), 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

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

ahol az első koordináta a szélességi, a második a hosszúsági kört jelenti (fokban). Beszélhetünk előregésekről is, de ezeket ebben a modellben nem definiáltuk.

A továbbiakban „regés” alatt az utóregések elhagyása után megmaradt regéseket értettük, amelyek legalább 2,9 magnitúdójúak. Az ennél kisebb magnitúdójú regések esetén minimális az esélye, hogy komoly kár következik be. A 3-as magnitúdónál kisebb regések maximális intenzitása – a tapasztalatok szerint – nagyon ritkán éri el a IV-es fokozatot, így komoly károkat nem okoznak. A *Munich Re Group* (2004) CD-je szerint a kárárány várható értéke még egy V-ös intenzitású regés esetén is 0,1%-nál kisebb. (Az intenzitás skála a pusztítás mértékét jelzi, I-től XII-ig változik. Az I-es erősségű regés alig érzékelhető, a XII-es a teljes pusztítást jelenti az épített környezetben. Erről is pontos adatokat tartalmaz a CD.)

1. ábra

Az adatokat feldolgozó program

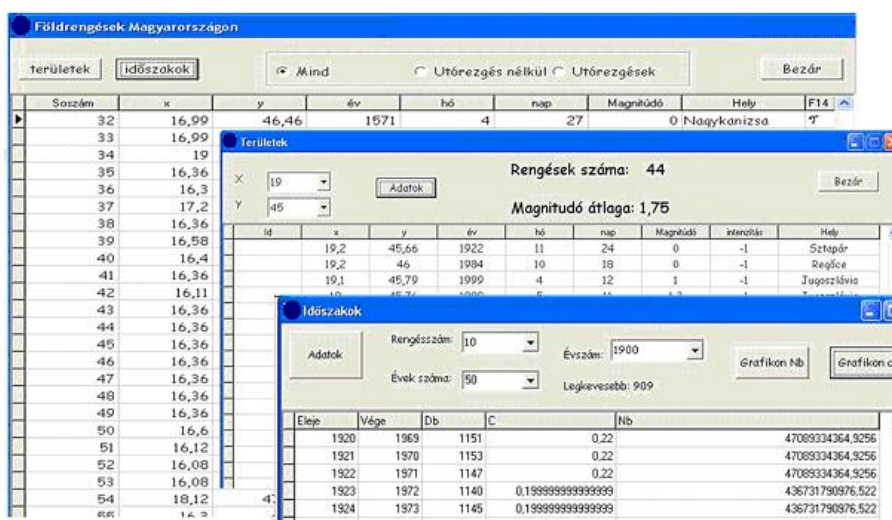


Figure 1: The application for the data processing

Számítógépes programunk (1. ábra) gyűjti, illetve kiszűri bármely terület és bármely időintervallum utóregéseit, így az utóregések számának eloszlását is vizsgálni tudjuk. Ez azért érdekes, mert tudnunk kell, hogy egy legenerált – szimulált – regést hány regés követ majd a főregés környékén. (Ezek is okozhatnak károkat.) Az utóregések száma is valószínűségi változó ebben a modellben. Az eloszlásukat azonosítani kell, és minden főregéshez ebből az eloszlásból rendelünk (generálunk) utóregés számot. Munkánkban elsősorban az évenkénti főregések számát kívánjuk vizsgálni, csupán néhány megjegyzést teszünk az utóregések számával kapcsolatban.

Az 1900 és 2003 közötti adatok esetén a következő gyakoriságokat kaptuk az utóregések számára vonatkozóan (2. ábra). Látható, hogy 144 regés esetén volt, 861 esetén nem volt utóregés.

2. ábra

Utóregések gyakorisága

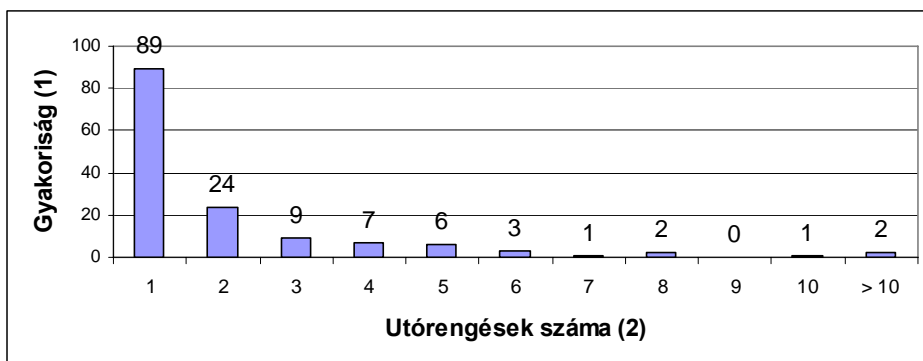


Figure 2: The frequency of the aftershocks

Frequency(1), Aftershocks(2)

Mivel itt az utóregések utóregései is szerepelnek, ezért egy másik adatsort is megnéztünk. Ha csak azokat az 1900 és 2003 közötti rengéseket nézzük, amelyek – fenti definíciónk alapján – nem utóregések, akkor 597 rengést nem követett utóregés. A 109 rengés hozott létre utóregést (3. ábra). Vagyis körülbelül minden hetedik rengés után kell a modellben utóregést generálni, és ha kell, akkor kb. 2/3 eséllyel egy utóregést. (Megjegyezzük, hogy ekkor egy utóregést is utóregés követ hozzávetőlegesen 1/7 valószínűséggel.)

3. ábra

A nem utóregések utóregéseinek gyakorisága

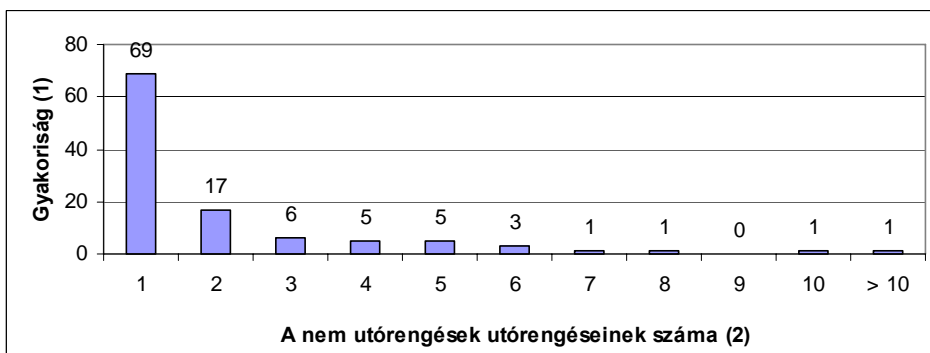


Figure 3: The frequency of the aftershocks of non-aftershocks

Frequency(1), Aftershocks of non-aftershocks(2)

Modellek a főrengések számának vizsgálatára

Megnéztük, hogy hány olyan rengésről tudunk az alábbi időintervallumokban (50 év hosszúak), amelyek Magyarországon történtek, magnitúdójuk pedig legalább 4, de kisebb mint 5 (vagyis jelentős rengések) (4. ábra).

4. ábra

A 4 és 5 magnitúdó közötti rengések gyakorisága

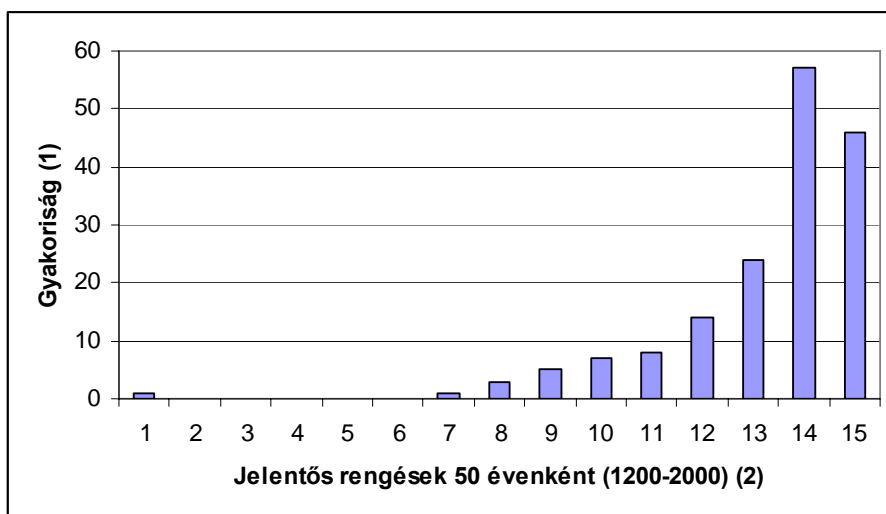


Figure 4: The frequency of the shocks with a magnitude between 4 and 5

Frequency(1), Shocks with a magnitude between 4 and 5(2)

Szembeötlő a megfigyelések számának gyors növekedése. Hasonló eredményeket találhatunk Zsíros (2000) tanulmányában. 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. Mivel az észlelések számának nagyfokú emelkedése a technikai fejlődés következménye, ezért úgy határoztunk, hogy csak 1900-tól tekintjük az adatokat. (Ezzel persze sok adatot figyelmen kívül hagyunk. Érdekes eljárást ír le Dargahi–Noubary (2002), az Észak–amerikai földrengésszám vizsgálatára, amellyel esetleg a korábbi adatok is hasznosíthatók lehetnek.)

A XX. századi adatok vizsgálata

1900-tól 2003-ig az egyes években regisztrált legalább 2,9 magnitúdójú nem-utórengések számának alakulását vizsgáltuk (5. ábra).

Az 1900. és 2003. között (104 év) megnéztük, hogy hány olyan év volt, amikor a földrengések száma 0, 1, 2 stb. volt (6. ábra).

5. ábra

Évenkénti rengésszámok (1900-2003)

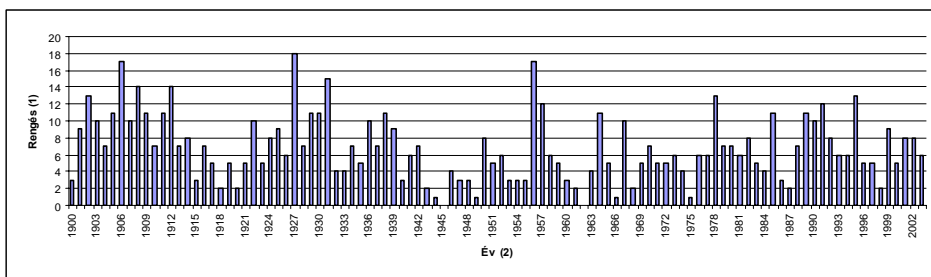


Figure 5: The annual shocks

Shock(1), Year(2)

6. ábra

Évi rengések számának gyakorisága

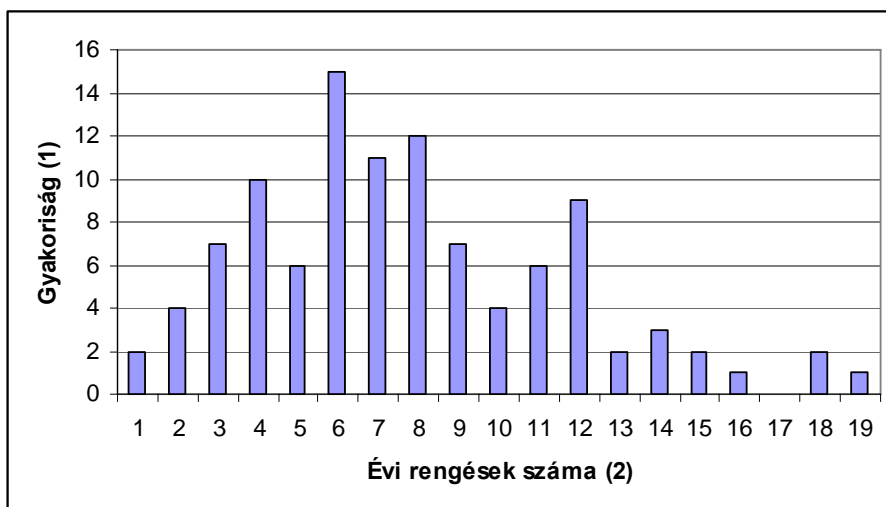


Figure 6: The frequency of the number of the annual shocks

Frequency (1), The number of the annual shocks(2)

Modellek

1. Első modellünkben az évenkénti rengések számait független, azonos eloszlású valószínűségi változónak tekintettük.

1.1. A biztosításmatematikában a kárszámok modellezésére leggyakrabban alkalmazott (a,b,0) eloszlások illeszkedését vizsgáltuk, vagyis a Poisson-, negatív

binomiális és binomiális eloszlást (Kagan and Jackson, 2000). Ezekre $P(\xi=n)=(a+b/n) \cdot P(\xi=n-1)$, $n=1,2,3,\dots$ (Példánkban ξ az éves rengésszám.)

A Biztosításmatematika című jegyzet (Arató, 1997) gondolatmenetét követve kiderült, hogy ezek az eloszlások nem illeszkedtek adatainkra. Ezért keverék Poisson–eloszlásokat tekintettünk.

1.2.1. Feltételeztük, hogy az évenkénti rengések számai: $\xi_{1900}, \xi_{1901}, \dots, \xi_{2003}$ függetlenek, azonos eloszlásúak, és $P(\xi_i=n)=p \cdot e^{-\lambda_1} \cdot \lambda_1 n/n! + (1-p) \cdot e^{-\lambda_2} \cdot \lambda_2 n/n!$ (minden i -re). A likelihood függvényt az R-programcsomagok 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 9,129, tehát az illeszkedés nagyon jó, a hipotézist elfogadtuk.

1.2.2. Megvizsgáltuk az illeszkedést abban az esetben is, amikor 3 Poisson–eloszlású valószínűségi változó keverékének tekintettük az éves rengésszámot. Ekkor feltételeztük, hogy az évenkénti rengések számai: $\xi_{1900}, \xi_{1901}, \dots, \xi_{2003}$ függetlenek, azonos eloszlásúak, és $P(\xi_i=n)=p \cdot e^{-\lambda_1} \cdot \lambda_1 n/n! + q \cdot e^{-\lambda_2} \cdot \lambda_2 n/n! + (1-p-q) \cdot e^{-\lambda_3} \cdot \lambda_3 n/n!$ ($i=1900, 1901, \dots, 2003$). Az R-programmal a paraméterekre a maximum likelihood becslés $\hat{p}=0,1033$, $\hat{q}=0,5915$, $\hat{\lambda}_1=1,8528$, $\hat{\lambda}_2=5,5832$, $\hat{\lambda}_3=10,7964$ lett. Itt is elvégeztük χ^2 -próbát. Az eredmény 3,456 lett, ami nagyon jó illeszkedést jelent.

2. A két Poisson keverékének modelljét általánosítva rejtett Markov–modellt vizsgáltunk.

Legyen Z_1, Z_2, \dots, Z_N valószínűségi változó, melyek értékészlete $\{0,1\}$. Z_m ($m=1,2,\dots,N$) -re úgy tekintünk, mint az m -edik „év” állapotára. Legyen λ_0 a 0-s állapot, λ_1 pedig az 1-es állapot paramétere ($\lambda_0, \lambda_1 > 0$). Tegyük fel, hogy

$$P\left(\mathbf{Z}_{m+1} = j \mid \mathbf{Z}_m = i, \mathbf{Z}_{m-1} = z_{m-1}, \dots, \mathbf{Z}_1 = z_1, \mathbf{Z}_0 = z_0\right) =$$

$$P\left(\mathbf{Z}_{m+1} = j \mid \mathbf{Z}_m = i\right) = p_{ij}, \quad \left(i, j \in \{0,1\}, m \in \{0,1,\dots,N-1\}\right),$$

nem függ m -től.

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 (valószínűségi változó) az m -edik év rengéseinek (megfigyeléseinek) számát. Feltesszük, hogy $\xi_1, \xi_2, \dots, \xi_N$, feltételesen független a $Z_1=z_1, Z_2=z_2, \dots, Z_N=z_N$, feltételre vonatkozóan, továbbá ξ_m feltételes eloszlása Poisson, és csak Z_m -től függ minden m esetén ($m \in \{1,2,\dots,N\}$), vagyis

$$P(\xi_m = k_m \mid \mathbf{Z}_1 = z_1, \mathbf{Z}_2 = z_2, \dots, \mathbf{Z}_m = i, \dots, \mathbf{Z}_N = z_N) = \frac{\lambda_i^{k_m}}{k_m!} e^{-\lambda_i} \quad \text{és}$$

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

Feladatunk az, hogy becsljük Z_1, Z_2, \dots, Z_N értékét, λ_0 és λ_1 és értékét, valamint a p_{00} és p_{10} átmenet–valószínűségeket.

Bayes–i megközelítés

Bayes–i megközelítéssel dolgozunk. Feltesszük, hogy paramétereink a priori eloszlása a következő:

$$\begin{aligned} p_{00} &\sim E(0,1), \\ p_{10} &\sim E(0,1), \\ \lambda_0 &\sim \Gamma(\alpha_0, \beta_0), \\ \lambda_1 &\sim \Gamma(\alpha_1, \beta_1), \end{aligned}$$

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

A Bayes–formulából

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

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 k -val jelölve (vektorba rendezve) az 1900 és 2003 közötti 104 év megfigyeléseit a feltételes sűrűségfüggvény:

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

ahol:

$f_{p_{00}}(p_{00}) = \mathcal{X}_{\{p_{00} \in [0,1]\}}$, $f_{p_{10}}(p_{10}) = \mathcal{X}_{\{p_{10} \in [0,1]\}}$, az $f_{\lambda_0}(\lambda_0)$ és az $f_{\lambda_1}(\lambda_1)$ függvény a gamma eloszlás sűrűségfüggvénye, a $P(\xi = k)$ csupán egy normalizáló konstans.

Ennek az $f_{\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_N, \mathbf{p}_{00}, \mathbf{p}_{10}, \lambda_0, \lambda_1 | \xi}(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N, p_{00}, p_{10}, \lambda_0, \lambda_1 | k)$ függvénynek a maximumhelyét keressük. A függvény tényezőit külön–külön megvizsgáltuk:

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

Valamint

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

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

Az a posteriori eloszlások

Először a λ_0 valószínűségi változó a posteriori eloszlását vizsgáltuk. Azt kaptuk, hogy

$$f_{\lambda_0 | \underline{\xi}, \underline{z}}(\lambda_0 | \underline{k}, \underline{z}) = \frac{g(z_1, z_2, \dots, z_N, \lambda_0, \lambda_1, p_{00}, p_{10})}{P(\underline{\xi} = \underline{k})} =$$

$$C \cdot f_{\lambda_0}(\lambda_0) \cdot \lambda_0^{\sum_{i=1}^N k_i} \cdot e^{-\lambda_0 \sum_{i=1}^N (1-z_i)}$$

Mivel $\lambda_0 \sim \Gamma(\alpha, \delta)$, és így sűrűségfüggvénye $f_{\lambda_0}(\lambda_0) = \frac{\delta^\alpha}{\Gamma(\alpha)} \cdot \lambda_0^{\alpha-1} \cdot e^{-\delta \lambda_0}$ ($\lambda_0 > 0$), ezért

$$f_{\lambda_0 | \underline{\xi}, \underline{z}}(\lambda_0 | \underline{k}, \underline{z}) = h(\lambda_0) = C \cdot \frac{\delta^\alpha}{\Gamma(\alpha)} \cdot \lambda_0^{\alpha-1 + \sum_{i=1}^N k_i} \cdot e^{-\lambda_0 \left(\delta + \sum_{i=1}^N (1-z_i) \right)} =$$

$$= B \cdot \lambda_0^{\alpha-1 + \sum_{i=1}^N k_i} \cdot e^{-\lambda_0 \left(\delta + \sum_{i=1}^N (1-z_i) \right)}$$

Tehát azt kaptuk, hogy λ_0 a posteriori eloszlása:

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

Ugyanezt mondhatjuk el λ_1 a posteriori eloszlásáról (α' és δ' , illetve $\alpha' + \sum_{i=1}^N k_i$ és

$\delta' + \sum_{i=1}^N z_i$ paraméterrel).

A p_{00} valószínűségi változó a posteriori eloszlásának vizsgálatára során kiderült, hogy az

a posteriori sűrűségfüggvény $f_{p_{00} | \underline{\xi}, \underline{z}}(p_{00} | \underline{k}, \underline{z}) = \frac{g(z_1, z_2, \dots, z_N, \lambda_0, \lambda_1, p_{00}, p_{10})}{P(\underline{\xi} = \underline{k})} =$

$= C \cdot f_{p_{00}}(p_{00}) \cdot p_{00}^{\ell_{00}} \cdot (1-p_{00})^{\ell_{01}}$, ahol a C konstans.

Megkaptuk, hogy az a posteriori eloszlás $\beta(1_{00}+1, 1_{01}+1)$.

Hasonlóan adható meg p_{01} a posteriori eloszlása is, amely $\beta(1_{10}+1, 1_{11}+1)$ lesz.

MCMC (Markov Chain Monte Carlo) módszer alkalmazása

A maximumhely keresést az MCMC (Markov Chain Monte Carlo) módszerrel végezzük (Rydén, 2004).

1. Z_1, Z_2, \dots, Z_N , értékeit $\{0,1\}$ -en diszkrét egyenletes eloszlásból, p_{00} p_{00} és 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.

2. A $\{1, 2, \dots, N\}$ -ből visszatevéses 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 $\frac{g^*(i_1)}{g(z_1, z_2, \dots, z_N, \lambda_0, \lambda_1, p_{00}, p_{00})} = a < 1$, akkor „a” valószínűséggel kicseréljük, 1-a valószínűséggel meghagyjuk z_{i_1} értékét.
A program az „a” értékét kiszámolja, majd ezt egy (0,1)-ből generált számmal összehasonlítja. Ha a generált szám kisebb a-nál, csere történik, különben nem.
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.
Megjegyezzük, hogy exponenciális eloszlásból megfelelő számú „kisorsolt” érték összegeként generáltunk gamma eloszlásból értékeket.
4. Hasonlóan járunk el az aktuális λ_1 , p_{00} és p_{10} értékkel szemben is.
Megjegyezzük, hogy p_{00} esetén $l_{00} + l_{01} + 3$ darab értéket generáltunk egyenletes eloszlásból, majd az értékek rendezése után a program kiválasztotta az $l_{00} + 2$ - ediket. (A 3. és 4. lépést Gibbs-lépésnek nevezzük.)
5. A 2–4. pontban leírtakat megfelelően sokszor – néhány tízezerszer – megismételjük. Természetesen a lépésszám a programban beállítható.
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.

EREDMÉNY ÉS ÉRTÉKELÉS

Az iterációs becslési eljárás elvégzéséhez programot dolgoztunk ki (7. ábra), amelynek alkalmazásával az alábbi eredményeket kaptuk.

5000 figyelembe nem vett iteráció után 100000 iterációt végeztünk (8. ábra). Az „összesített” becslés az első („Átlag”) sorban látható, a további sorok 100–100 iteráció során kapott átlagot mutatják (vagyis csupán 1200 iteráció eredményét tartalmazza az 1–12. sor, az első viszont 100000-ét). Az utolsó két oszlop azt jelzi, hogy az utolsó 2 évben az iterációk hány százalékában volt az egyes állapotban a rendszer.

Az iterációkat 100-asával csoportosítva λ_0 becslést értékeit a program minden futtatásnál ábrázolja (9. ábra). (Hasonló grafikon készíthető λ_1 értékeiről.)

Az átmenet-valószínűségek becslését is ábrázolja a program (10. ábra). (p_{10} esetén is hasonló grafikont kapunk.)

7. ábra

Az előrejelzést végző program

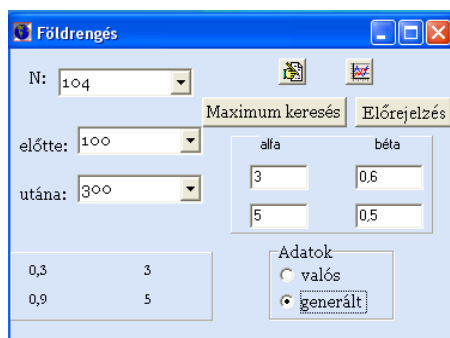


Figure 7: The prediction module

8. ábra

A szimulált paraméterek átlagai

Értékek						
	l0	l1	p00	p10	z103	z104
Átlag	4,357446886	9,882844204	0,747851043	0,307371721	66	45,66666666
1.	4,499083537	10,15187268	0,759383446	0,338632942	55	38
2.	4,415065812	9,912127269	0,766337362	0,293565960	62	45
3.	4,158191316	9,584532662	0,717832319	0,289916262	81	54

Figure 8: The mean of the simulated parameters

9. ábra

A λ_0 becslései az iterációkat százasaival csoportosítva

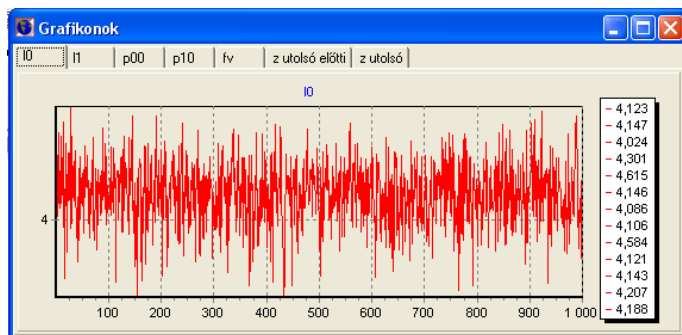


Figure 9: The estimated values for λ_0 , grouped by 100 iterations

10. ábra

A p_{00} becslései az iterációkat százasaival csoportosítva

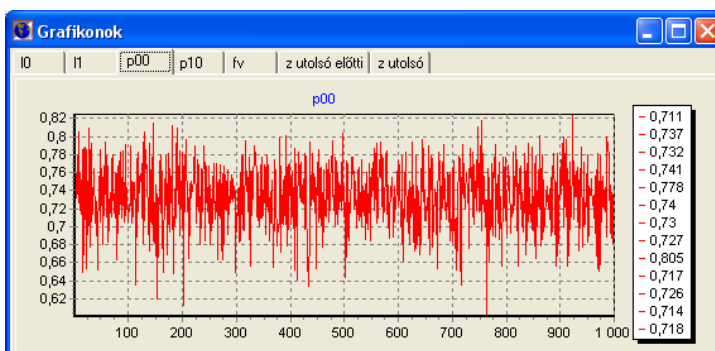


Figure 10: The estimated values for p_{00} , grouped by 100 iterations

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. A paraméterek becslése után állapotokat generáltunk 20 évre a megfelelő átmenet-valószínűségekkel, majd az állapothoz tartozó (szintén becslült) paraméterrel Poisson–eloszlásból adatokat generáltunk (11. ábra). (A program természetesen lehetővé teszi 20–nál több (vagy kevesebb) év vizsgálatát is.)

11. ábra

Előrejelzés

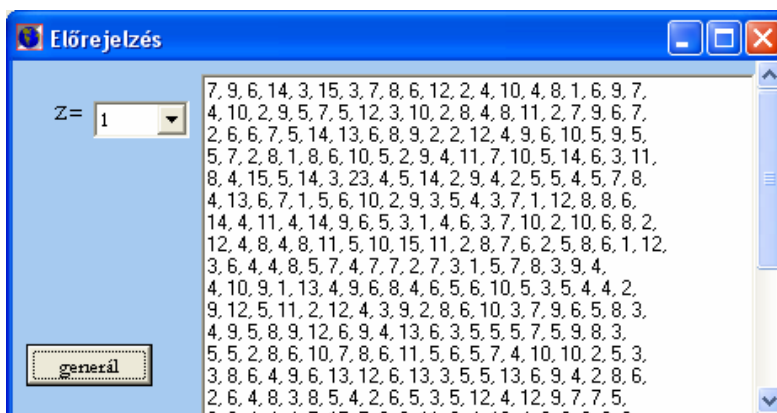


Figure 11: The forecast

1000 forgatókönyvet hoztunk létre, vagyis 1000–szer 20 év előrejelzését generáltuk le. 1000 előrejelzéshez jutottunk például a 2007–es évre vonatkozóan is (12. ábra).

12. ábra

A 2007-es évre generált adatok gyakorisága 1000 forgatókönyv esetén

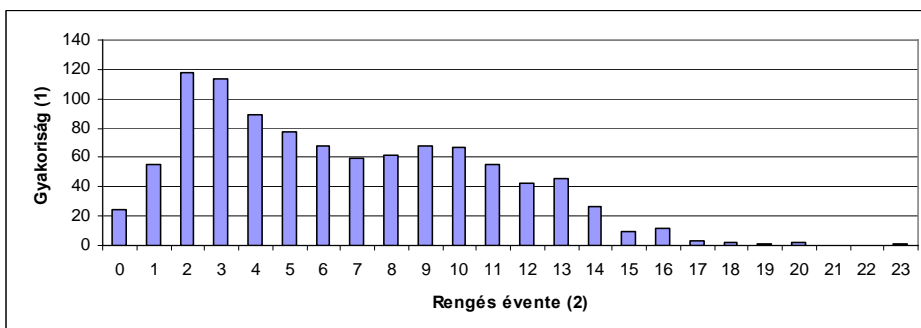


Figure 12: The frequency of the generated data in case of 1000 scenarios

Frequency(1), The number of the annual shocks(2)

ÖSSZEFOGLALÁS

Az egy év alatt Magyarországon bekövetkező éves földrengésszámot vizsgáltuk.

Adataink közül programunkkal kiszűrtük az utóregéseket. Kiderült, hogy Poisson-, negatív binomiális és binomiális eloszlás nem illeszkedik adatainkra. A keverék Poisson-eloszlás paraméterbecslését, majd az illeszkedésvizsgálatot végeztük el, az R-programot használva. Ezt követően az adatok egymástól való függését feltételező modellt, a rejtett Markov modellt (Hidden Markov Model) vizsgáltuk. Bayes-i megközelítéssel dolgoztunk, MCMC módszert alkalmaztunk kétállapotú rendszerben. A paraméterek becslése több ezer (esetleg százezer) csere után történik meg, amelyet programunk végez el. Megadtuk a paraméterbecsléseket és több lehetséges forgatókönyvet a következő évekre.

KÖSZÖNETNYILVÁNÍTÁS

Köszönetet mondunk Arató Miklósnak a dolgozat írásának minden stádiumában nyújtott odaadó segítségéért és javaslataiért.

IRODALOM

- Arató M. (1997). Általános Biztosításmatematika. ELTE Eötvös Kiadó : Budapest. 19-28.
- Dargahi–Noubary, G.R. (2002). 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. 361-369.
- Omori, T. (1894). On the aftershocks of earthquakes Journal of the College of Science of the Imperial University of Tokyo, 111-200.
- Kagan, Y.Y., Jackson, D.D. (2000). Probabilistic earthquake forecasting, Geophys.J.In., 143. 438-453.

- Munich Re Group (2004). World of Natural Hazards, CD
- Rydén, T. (2004). „Hidden Markov models” in Encyclopedia of Actuarial Science, 2. J. Teugels, B. Sundt, eds.: Wiley, 821–827.
- Tóth, L., Mónus, P., Zsíros, T. (1997). Hungarian Earthquake Bulletin, 1996. GeoRisk : Budapest
- Tóth, L., Mónus, P., Zsíros, T. (1998). Hungarian Earthquake Bulletin, 1997. GeoRisk : Budapest
- Tóth, L., Mónus, P., Zsíros, T. (1999). Hungarian Earthquake Bulletin, 1998. GeoRisk : Budapest
- Tóth, L., Mónus, P., Zsíros, T. (2000). Hungarian Earthquake Bulletin, 1999. GeoRisk : Budapest
- Tóth, L., Mónus, P., Zsíros, T., Kiszely, M., Kosztyu, Z. (2001). Hungarian Earthquake Bulletin, 2000. GeoRisk : Budapest
- Tóth, L., Mónus, P., Zsíros, T., Kiszely, M., (2002). Hungarian Earthquake Bulletin 2001. GeoRisk : Budapest
- Tóth, L., Mónus, P., Zsíros, T., Kiszely, M., Czifra, T. (2003). Hungarian Earthquake Bulletin 2002., Georisk : Budapest
- Tóth, L., Mónus, P., Zsíros, T., Kiszely, M., Czifra, T. (2004). Hungarian Earthquake Bulletin 2003., Georisk : Budapest
- Tóth, L., Mónus, P., Zsíros, T., Kiszely, M. (2002). Seismicity in the Pannonian Region – earthquake data, EGU Stephan Mueller Special Publication Series 3. 9–28.
- Zsíros, T., Mónus, P., Tóth, L. (1988). Hungarian earthquake catalog MTA GGKI : Budapest 456-1986.
- Zsíros T. (2000). A Kárpát medence szeizmicitása és földrengés veszélyessége: Magyar földrengés katalógus MTA GGKI : Budapest, 456-1995.

Levelezési cím (*Corresponding author*):

Kovács Előd

Pannon Egyetem, Matematikai és Számítástechnikai Tanszék,
8200, Veszprém, Egyetem u. 10.

*Pannon University, Department of Mathematics and Computing
H-8200, Veszprém, Egyetem u. 10.*

Tel.: 36-88-624-234

e-mail: kovacse@almos.uni-pannon.hu