# Markov lánc Monte Carlo módszerek és alkalmazásuk

Vasas Krisztina

2007.05.17

Altenburger Gyula Szimpózium

# Markov láncok

- Irreducibilitás : egy osztály van
- Aperiodocitás : minden elem periódusa 1
- Visszatérőség : minden elem az osztályban visszatérő

# Markov-láncok stacionárius eloszlása

- Ha a lánc aperiodikus és irreducibilis, pozitív visszatérő akkor létezik $\pi(y)$ stacionárius eloszlás

$$\lim_{n\to\infty} P(X_n = y \mid X_0 = y) = \pi(y) = \frac{1}{m(y)}$$

- Ha egy ergodikus lánc kielégíti a „detailed balance" feltételt azaz:

$$K(x, y)\pi(x) = K(y, x)\pi(y), \quad \forall x, y \in \textit{állapottér}$$

- akkor a láncot reverzibilisnek nevezzük

# Egy gyakorlati probléma

- Bayesi módszerek a statisztikában:

$$\pi\left(x \mid y\right) \propto L(y \mid x)\, p(x)$$

- arányossági tényező:

$$\int L(y \mid x)\, p(x)\, dx$$

- Ezt nehéz (vagy nem is lehet) analitikus módszerekkel kiszámolni
- Ugyanígy, ha x többdimenziós, a marginális posteriorok, vagy az $E\left(\theta\left(X\right) \mid y\right)$ várható érték kiszámítása is bonyolult

# Megoldások

- Numerikus integrálás, approximáció
- A posterior eloszlásból veszünk mintát, majd a kiszámolni kívánt mennyiségeket ennek segítségével becsüljük: ez a Markov lánc Monte Carlo módszer lényege
- Az első alkalmazás: atomok energiaszintjének meghatározása (1953,Metropolis), majd 1970-ben statisztikai problémák megoldása

# Az MCMC-algoritmusok általános elve I.

- Legyen $\pi(x)$ egy „céleloszlás”, amelyből mintát kívánunk venni (közvetlenül nem tudunk)
- Generáljunk egy Markov-láncot, amely aperiodikus és irreducibilis, és amelynek stacionárius eloszlása éppen $\pi(x)$
- Ha a generált sor elég hosszú, akkor a tagok tekinthetők egy $\pi(x)$ eloszlásból való független mintának

# Az MCMC-algoritmusok általános elve II.

- Ebből a mintából becsülhetjük a céleloszláshoz kapcsolódó mennyiségeket, pl. a várható értéket:

$$\frac{1}{n}\sum_{t=1}^{n}\theta\left(X^{t}\right)\underset{n\to\infty}{\longrightarrow}E_{\pi}(\theta(X))$$

- ahol $X^t$ egy realizáció

# Hogyan generáljunk „jó” láncot?

- Jelölje K(x,y) az átmenetkernelt,azaz ha most x-ben van a lánc, akkor a feltételes eloszlását a lánc következő (y) állapotának, feltéve a jelenlegi állapotot
- Ahhoz, hogy $\pi(x)$ stac.eloszlású láncot generáljunk, olyan K(x,y) kell, amire: $K\pi = \pi$
- Ha a lánc reverzibilis $(\pi(x)K(x,y) = \pi(y)K(y,x))$
- akkor minden olyan K(x,y)-nak, ami teljesíti a feltételt, $\pi(x)$ a stacionárius eloszlása

# Metropolis-Hastings algoritmus

- Legyen $x=(x_1,\ldots,x_k)$ az állapotvektor , amelynek az i. komponensét újítjuk fel
- $q(x, y)$ tetszőleges eloszlás, amelyből generálunk egy $\mathrm{y}=(x_1,\ldots,x_{i-1}, y_i, x_{i+1},\ldots,x_k)$ elemet
- Ezt az elemet

$$\alpha(x, y)=\min\left(1, \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}\right)$$

- valószínűséggel elfogadjuk, mint a lánc következő elemét, ha nem fogadjuk el, akkor a lánc ugyanabban az állapotban marad

# Átmenetvalószínűségek a M-H által generált láncban

- Legyen $P(x, A)$ annak valószínűsége, hogy a lánc következő állapota az A halmazba esik, feltéve, hogy most x-ben van.

$$P(x, A) = \int_A K(x, y)dy + \chi_A(x)r(x)$$

- Itt $K(x, y) = q(x, y)\alpha(x, y)$ , $r(x) = 1 - \int_\Omega q(x, y)\alpha(x, y)dy$
- P(x,A) kielégíti a detailed balance feltételt $\pi(x)$-re vonatkozóan
- $\pi(x)$-et elég konstans szorzó erejéig ismerni
- Ha q(x,y) szimmetrikus, egyszerű az elfogadási valószínűség

# A Gibbs-sampler

- M-H speciális esete, akkor alkalmazható, ha a láncelemek feltételes sűrűségfüggvényei ismertek

$$\pi\left(y_i \mid x_{(i)}\right) = \pi\left(y_i \mid x_j \forall j \neq i\right)$$

- $$q(x, y) := \begin{cases} \pi\left(y_i \mid x_{(i)}\right) & ha\ x_{(i)} = y_{(i)}, i = 1 \ldots k \\ 0 & egyébként \end{cases}$$
- Ekkor

$$\alpha(x, y) = \frac{\pi(y)\pi\left(y_i \mid x_{(i)}\right)}{\pi(x)\pi\left(x_i \mid y_{(i)}\right)} = \frac{\pi(y)\pi\left(y_i \mid y_{(i)}\right)}{\pi(x)\pi\left(x_i \mid x_{(i)}\right)} = \frac{\pi\left(y_{(i)}\right)}{\pi\left(x_{(i)}\right)} = 1$$

# Egyéb algoritmusok

- Random-walk
- $$q(x, y) = f(y - x)$$
- Független sampler

$$q(x, y) = f(y)$$

- azaz a lánc pillanatnyi állapotától függetlenül választjuk a következőt
- Kombinált samplerek

# Felmerülő kérdések és problémák

- Hány iterációt hajtsunk végre a felújítások során?
- Mikortól lesz a lánc stacionárius?
- Milyenek az egyes algoritmusok keverési tulajdonságai?
- Mi legyen a kezdő állapot?
- Melyik algoritmus a „(leg)jobb”?

# Modellválasztási probléma

- Olyan modellek közül próbálunk választani, amelyekben a paraméterek száma maga is paraméter, azaz dimenziók között „ugrálunk”
- Pl. a lehetséges modellek: $\{\mathcal{M}_k, k \in \mathcal{K}\}$, ahol $\mathcal{M}_k$ paramétervektora $\theta_k \in \Re^{n_k}$
- Megoldás: „reversible jump” MCMC
- Green, 1995

# RJMCMC I.

- Legyen most $\pi(dx)$ mérték, $q(x,dy)$ egy magfüggvény, amelyből a jelölt elemet választjuk, illetve definiáljunk mozgástípusokat (m=1,2,…) amellyel egy dimenzióból a másikba kerülünk
- Ha x-ben vagyunk éppen, akkor legyen az együttes eloszlása $q_m(x,dy)$ annak, hogy y a következő lehetséges jelölt és m a kiválasztott mozgás típusa

$$\alpha_m(x,y) = \min\left(1, \frac{\pi(dy)q_m(y,dx)}{\pi(dx)q_m(x,dy)}\right)$$

# A RJMCMC megvalósítása I.

- Tegyük fel, hogy az *m* mozgási típus segítségével jutunk el **x**-ből **y**-ba (y magasabb dimenzióban van)
- Mi lehet az *m?*Válasszunk egy x-től független **u** véletlen vektort, és legyen **y** ennek és **x**-nek determinisztikus függvénye (y(**x**,**u**))
- **y**-ból **x**-be az inverz transzformáció segítségével juthatunk

# Példa a mozgástípusok definíciójára

- Két lehetséges altér:

$$C_1 = \{1\} \times \Re \text{ és } C_2 = \{2\} \times \Re^2$$

- Hogyan jutunk $\{2, \theta_1, \theta_2\}$-ből vissza $C_1$-be? Például az $f(\theta_1,\theta_2)=½(\theta_1+\theta_2)$ transzformációval.
- Azaz: $\{2, \theta_1, \theta_2\} \rightarrow \left\{1, \frac{1}{2}(\theta_1 + \theta_2)\right\}$
- És vissza $\{1, \theta\}$ -ból ? Választunk egy u véletlen vektort függetlenül θ-tól, és $\theta_1 = \theta + u$ illetve $\theta_2 = \theta - u$.

# Az RJMCMC megvalósítása II.

- Ha $\pi(dx)q_m(x,dy)$ létezik véges sűrűségfüggvénye valamilyen (szimmetrikus) mértékre nézve, akkor az (m,y) párt előállíthatjuk az együttes sűrűségfüggvényükből
- Ekkor

$$\alpha(x,y) = \min\left(1, \frac{p(y)r_m(y)}{p(x)r_m(x)q(u)} \left| \frac{\partial y}{\partial(x,u)} \right| \right)$$

# Egy térstatisztikai probléma megoldása MCMC segítségével

- Cél : épületbiztosítások kárszámadatainak modellezése Poisson-modellel, illetve magyarországi települések veszélyességi kategóriákba sorolása
- 3222 település adatai állnak rendelkezésre
- Ezekből 168 kistérséget képeztünk
- Térbeli összefüggést mutató, „számláló” típusú adatok

# Poisson-modell

- Az alkalmazott modell: $y_i \approx \frac{\left(\lambda_{z_i} E_i\right)^{y_i}}{y_i!} e^{\lambda_{z_i} E_i}$
- $y_i$: az i. kistérséghez tartozó kárszám (Poisson-elo. $\lambda_z E_i$ várható kárszámmal)
- $z_i$ : az i. kistérség milyen veszélykategóriába tartozik, értékei természetes számok 1 és 6 között
- Minden kategóriához tartozik egy (napi) kárintenzitás – ez a $\lambda_j$
- $E_i$ : „módosított tartam”, azaz a kockázatban töltött idő napokban, módosítva az épülettípus, falazat és tető típusának hatásával

# Paraméterbecslés

- Paraméterek: a $\underline{z}=(z_1,\ldots,z_{168})$ és a $\underline{\lambda}=(\lambda_1,\ldots,\lambda_6)$ intenzitásparaméter-vektor és a $\psi$ interakcióparaméter
- Feladat: a paraméterek együttes a posteriori eloszlásából generáljuk a paraméterek becsléseit

# A priori sűrűségfüggvények I.

- Prior a $\underline{z}$-re:

$$p(z \mid \psi) = \frac{e^{\psi U(z)}}{e^{\theta_k(\psi)}}$$

- ahol $U(z) = \sum_{i \sim i'} \chi_{\{z_i = z_{i'}\}}$ i~i’ a szomszédsági relációt jelöli. Szomszédnak nevezünk két települést, ha 30 km-nél közelebb vannak egymáshoz. U tehát azt számolja meg, hogy egy adott szomszédsági struktúrában hány azonos kategóriába tartozó szomszéd van.
- $$\theta_k(\psi) = \log\left( \sum_{z \in \{1,\ldots,6\}^{168}} e^{\psi U(z)} \right)$$

# A priori sűrűségfüggvények II.

- $\Psi$ nemnegatív paramétert interakció-paraméternek nevezzük, a $\psi$=0 érték a kockázati osztályokba való független besorolást jelenti ({1,..,k}-n egyenletes eloszlás szerint), a térbeli függőség nő, ha $\psi$ nő, a priori eloszlását egyenletesnek tekintettük a {0,0.1,…,1} halmazon
- Prior $\lambda_j$-re: $\lambda_j \sim \Gamma(\alpha, \beta)$ $\quad \alpha = 1$ és $\beta = \dfrac{\sum y_i}{\sum E_i}$

- Az együttes a priori sűrűségfüggvény

$$p(\underline{\lambda} \mid \alpha, \beta) = k!\,\chi_{\{\lambda_1 < \lambda_2 < \ldots < \lambda_6\}} \prod_{j=1}^{6} \frac{\beta^{\alpha} \lambda_j^{\alpha-1}}{\Gamma(\alpha)} e^{\beta \lambda_j}$$

# Az a posteriori együttes sűrűségfüggvény

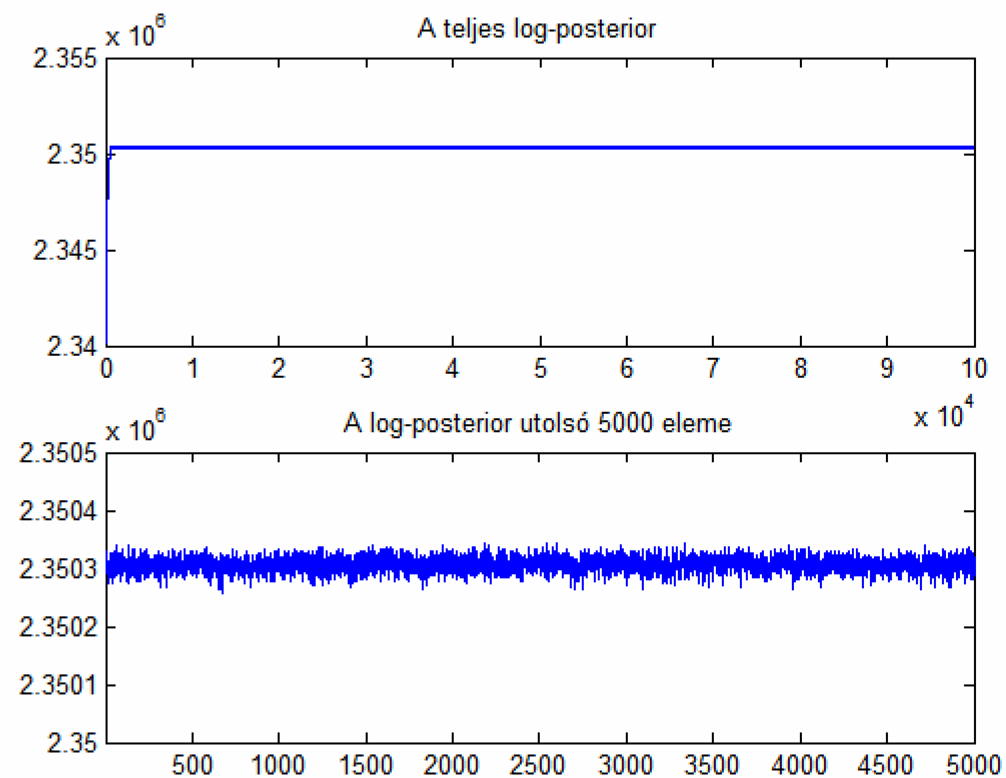
- $$p\left(\underline{z}, \underline{\lambda, \psi \mid y}\right) \propto p(\psi) p\left(\underline{\lambda} \mid \alpha, \beta\right) p\left(\underline{z} \mid \psi\right) p\left(y \mid \underline{z}, \underline{\lambda}\right)$$
- A $\underline{z}$ vektort Gibbs-sampling segítségével becsüljük, ehhez a „full conditional”:
$$P\left(z_i = j \mid \psi, y\right) \propto \lambda^{y_i} e^{-\lambda_j E_i} e^{\psi n_{ij}}$$
- $n_{ij}$ : i azon szomszédainak száma, akik szintén j osztályba vannak sorolva
- Ψ felújítása random-walk MCMC-vel, ±0.01-os perturbációkkal

# Az intenzitásparaméter felújítása

- Metropolis-Hastings algoritmus
- A „jelölt” eloszlás normális, ebből λ logaritmusát generáljuk, amit exp. függvénybe helyettesítünk
- Az elfogadási valószínűség:

$$\min\left(1, \prod_{j=1}^{6}\left(\frac{\lambda'_j}{\lambda_j}\right)^{\alpha+\sum_{i:z_i=j} y_i} e^{\left\{\left(\lambda'_j-\lambda_j\right)\left(\beta+\sum_{i:z_i=j} E_i\right)\right\}}\right)$$

# A log-posterior grafikonja (100000 iteráció)

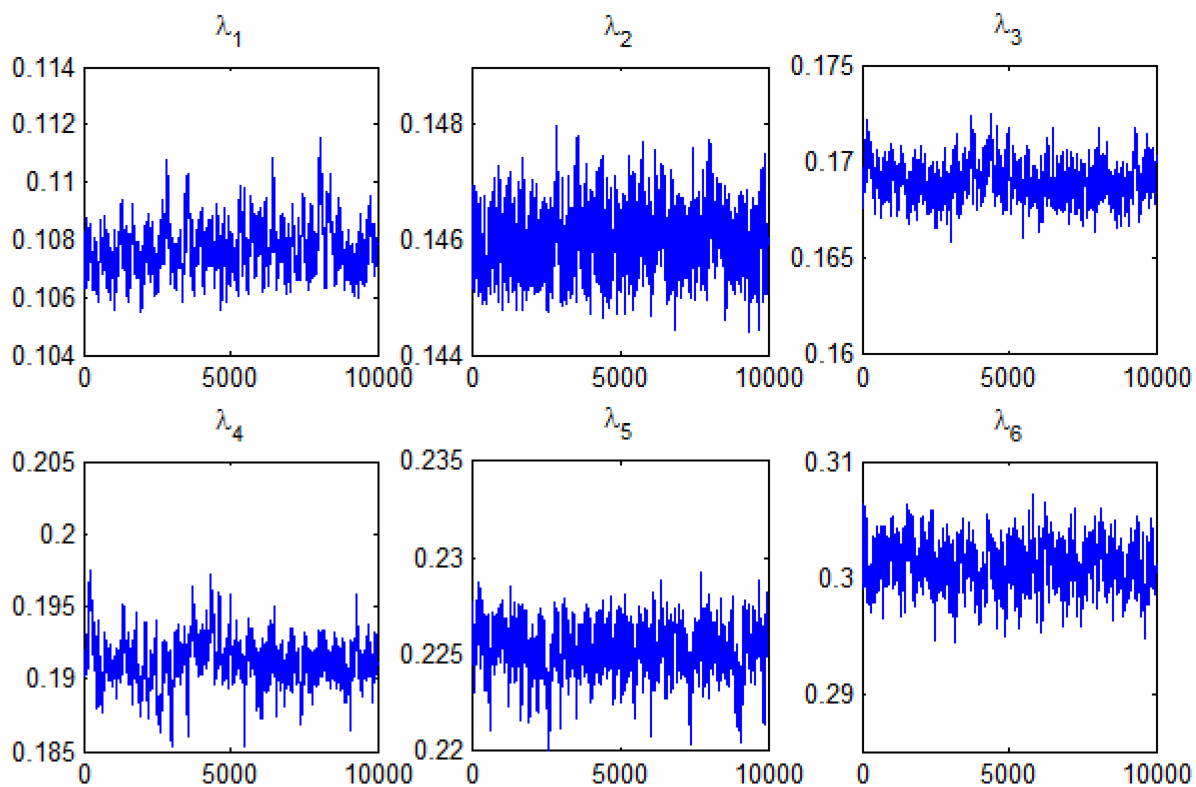


# Az intenzitásparaméterek

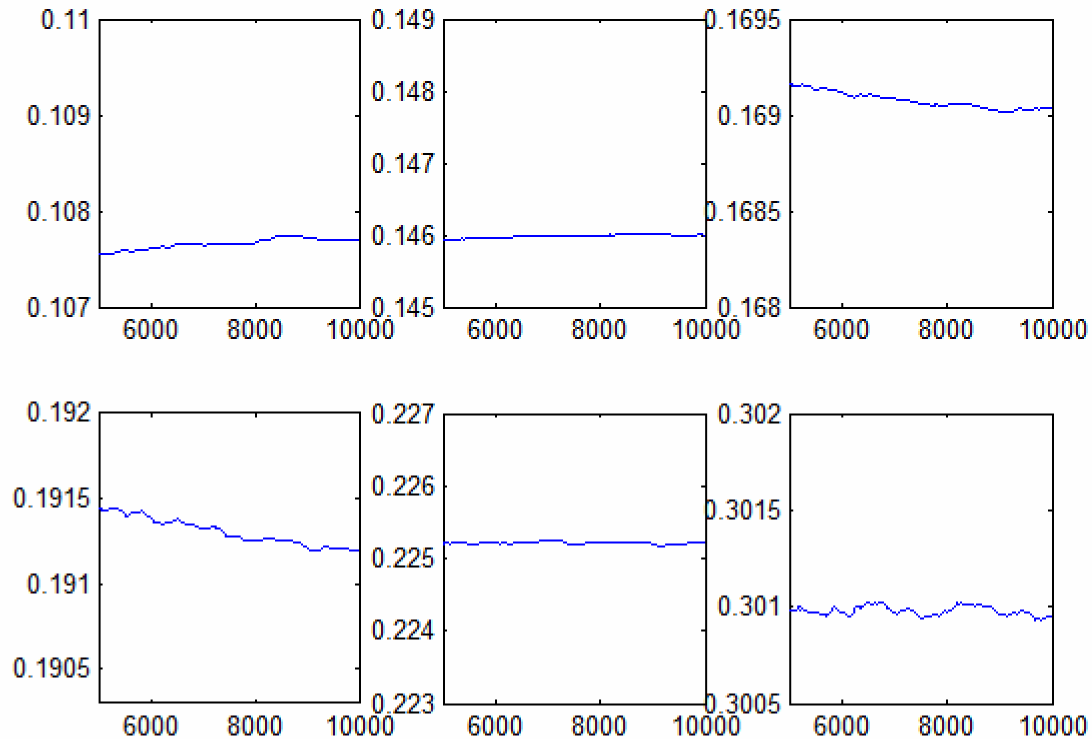
- A paraméterek autokorrelációfüggvénye gyorsan lecseng nullához, ez jelzi, hogy a lánc elemei konvergálnak a kívánt eloszláshoz.
- Az utolsó 5000 iterációban már nagyon keveset változnak a paraméterek értékei, ami szintén a jó konvergenciára utal

| | Átlagos paraméterérték |
|---|---|
| $\lambda_1$ | 0.1079 |
| $\lambda_2$ | 0.1461 |
| $\lambda_3$ | 0.1690 |
| $\lambda_4$ | 0.1910 |
| $\lambda_5$ | 0.2252 |
| $\lambda_6$ | 0.3008 |

# Az intenzitásparaméterek trajektóriái



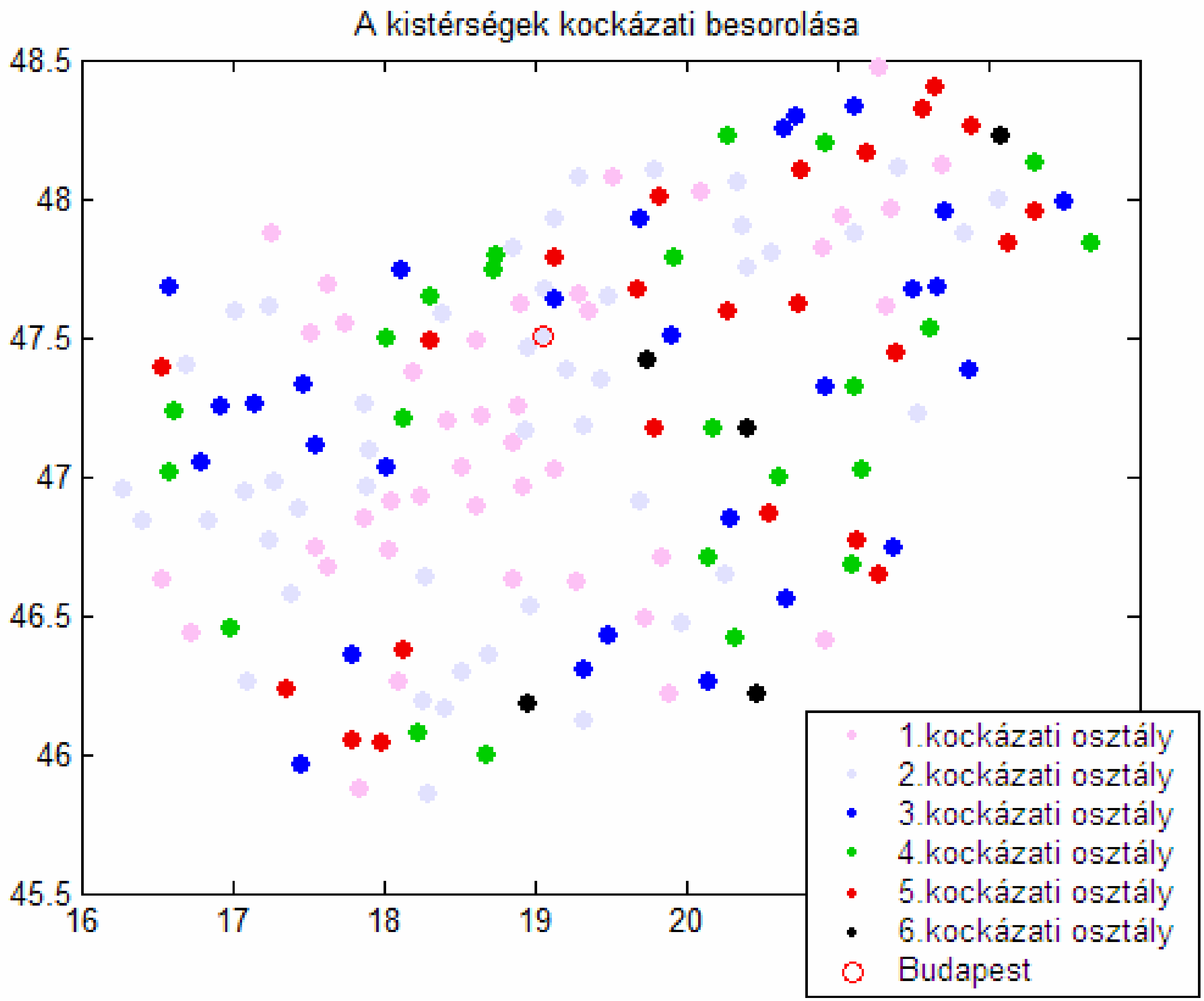
# Az intenzitásparaméterek trajektóriái az utolsó 5000 iterációban



# A $z$ és $\psi$ paraméterek becslése

- „Veszélyességi” térkép: egy térképen beszínezzük a kistérségek régióközpontjait, minden veszélyosztálynak egy színt megfeleltetve. Az első kockázati osztály a legkevésbé veszélyes(lila), a 6. (fekete) a legveszélyesebb károkozás szempontjából
- A $\psi$ becslése 1 körül ingadozik
- A legkockázatosabb települések ÉK-en, illetve Pécs környékén találhatóak, Budapest kis mértékben kockázatosnak tekinthető

# Veszélyességi térkép



# A modell lehetséges javítása

- A kockázati osztályok számát nem mondjuk meg előre, hanem azt is paraméterként kezeljük. (RJMCMC)

# Egy rezsimváltó autoregresszív modell

$$Y_t = \begin{cases} Y_{t-1} + \varepsilon_{1,t}, & if \quad I_t = 0 \\ a(Y_{t-1} - c) + c + \varepsilon_{2,t}, & if \quad I_t = 1 \end{cases}$$

- $I_t$ az aktuális rezsim típusa:
  - 0 ha t a növekvő rezsimbe tartozik
  - 1 ha t a csökkenő rezsimbe tartozik.
- $I_t$ Markov-lánc a következő átmenetvalószínűség-mátrixxal:

$$P = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$$

- A generáló zajok feltételesen függetlenek:
  - $\varepsilon_{1,t} \sim \Gamma(\alpha,\lambda)$ a növekvő rezsimben (pozitív sokkok)
  - $\varepsilon_{2,t} \sim N(0, \sigma^2)$ a csökkenő rezsimben

# A modell stacionaritása

- A modell a következő sztochasztikus differenciaegyenlet alakjában írható: $$Y_t = a_t Y_{t-1} + b_t$$

- ahol $$a_t = 1\chi_{\{I_t=0\}} + a\chi_{\{I_t=1\}}$$

$$b_t = \varepsilon_{1,t}\chi_{\{I_t=0\}} + \varepsilon_{2,t}\chi_{\{I_t=1\}}$$

- Brandt(1996) tételéből következik, hogy az egyértelmű stacionárius megoldás :

$$Y_t = b_t + \sum_{i=1}^{\infty} a_t a_{t-1} \cdots a_{t-i-1} b_{t-i}$$

- mivel

$$E\left(\log\left(|a_0|\right)\right) < 0, \text{ and } E\left(\log\left(|b_0|^+\right)\right) < \infty$$

# A modell becslése

- A modell paraméterei: $\theta = (p_{00}, p_{11}, \alpha, \lambda, \eta, a, c)$, ahol $\eta=1/\sigma^2$ .
- A látens változók ($I_t$) a modell becslését megnehezítik.
- Kézenfekvő lehetőségek:
  - Maximum likelihood
    (A likelihood felírható, de csak rekurzív formában, a látens változók miatt)
  - **Markov Chain Monte Carlo (MCMC)**
  - Efficient Method of Moments (EMM).

# Az a posteriori sűrűségfüggvény

$$f\left(\underline{\theta} \mid \{Y_t\}\right) = \int f\left(\{I_t\}_{t=1}^{T}, \underline{\theta} \mid \{Y_t\}\right) d\{I_t\}_{t=1}^{T}$$

$$f\left(\{I_t\}_{t=1}^{T}, \underline{\theta} \mid \{Y_t\}\right) \propto f\left(\{Y_t\} \mid \{I_t\}_{t=1}^{T}, \underline{\theta}\right) f\left(\{I_t\}_{t=1}^{T} \mid \underline{\theta}\right) f\underline{(\theta)} \propto$$

$$\propto \prod_{t=1}^{T} \left[f_{\Gamma(\alpha,\lambda)}\left(Y_t - Y_{t-1}\right)\right]^{\chi\{I_t=0\}} \left[f_{N(0,\sigma^2)}(Y_t - c - a(Y_{t-1} - c))\right]^{\chi\{I_t=1\}} p_{00}^{n_{00}} p_{01}^{n_{01}} p_{10}^{n_{10}} p_{11}^{n_{11}} \cdot f(\underline{\theta})$$

Ahol például $n_{00}$=#($I_{t-1}$=0 and $I_t$=0).

# Markov Chain Monte Carlo módszer

- Az együttes posteriorból kell mintát vennünk, hogy meg kapjuk a becsléseket
- Erre a megfelelő módszer egy Markov-lánc készítése a következő stacionárius eloszlással: $f\left(\{I_t\}_{t=1}^{T}, \underline{\theta} \mid \{Y_t\}\right)$ .
- A rejtett állapotokat ($I_t$) a strukturális paramétekhez hasonlóan „felújítjuk”
- Gibbs-samplinget szertnénk használni, amikor csak lehet
- Konjugált priorok segítségével 6 paraméternek felírhatóak a megfelelő feltételes eloszlásai
- A Metropolis-Hastings lépés csak a növekvő rezsimben jelenlévő gamma eloszlás alakparaméterének (α) becslésekor szükséges

# Gibbs sampling

- Ha felírhatóak a teljes feltétele eloszlások:

$$\pi(\theta_j \mid \theta_1,\ldots,\theta_{j-1},\theta_{j+1},\ldots,\theta_n) = \frac{\pi(\theta_1,\ldots,\theta_{j-1},\theta_j,\theta_{j+1},\ldots,\theta_n)}{\int \pi(\theta_1,\ldots,\theta_{j-1},\theta_j,\theta_{j+1},\ldots,\theta_n)\,d\theta_j}$$

- a mintavétel a feltételes eloszlásokból realizálható:

$$\begin{aligned}
&\theta_1^{(t+1)} \sim \pi\left(\theta_1 \mid \theta_2^{(t)},\theta_3^{(t)},\ldots,\theta_n^{(t)}\right)\\
&\theta_2^{(t+1)} \sim \pi\left(\theta_2 \mid \theta_1^{(t+1)},\theta_3^{(t)},\ldots,\theta_n^{(t)}\right)\\
&\vdots\\
&\theta_n^{(t+1)} \sim \pi\left(\theta_n \mid \theta_1^{(t+1)},\theta_2^{(t+1)},\ldots,\theta_{n-1}^{(t+1)}\right)
\end{aligned}$$

# Metropolis-Hastings

- Válaszzunk egy Y jelöltet $q(y|\theta^{(t)})$ eloszlásból az iteráció minden lépésében
- A jelölt pontot fogadjuk el az alábbi valószínűséggel

$$\alpha\left(\theta^{(t)}, y\right) = \min\left(\frac{\pi(y)q\left(\theta^{(t)} \mid y\right)}{\pi\left(\theta^{(t)}\right)q\left(y \mid \theta^{(t)}\right)}, 1\right)$$

ekkor $\theta^{(t+1)}=y$ és minden egyéb esetben $\theta^{(t+1)}= \theta^{(t)}$.

# A priorok megválasztása

- $\alpha \sim \Gamma(\alpha_u,\lambda_u)$
- $\lambda \sim \Gamma(r,\beta)$
- $\eta \sim \Gamma(q,\rho)$
- $a \sim N(\mu,\tau)$
- $c \sim N(\nu,\kappa)$
- $p_{00} \sim \beta(u_1,v_1)$
- $p_{11} \sim \beta(u_2,v_2)$

**Remember the model:**

$$Y_t = \begin{cases} Y_{t-1} + \varepsilon_{1,t}, & if \quad I_t = 0 \\ a(Y_{t-1} - c) + c + \varepsilon_{2,t}, & if \quad I_t = 1 \end{cases}$$

$$P = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$$

$$\varepsilon_{1,t} \sim \Gamma(\alpha,\lambda)$$

$$\varepsilon_{2,t} \sim N(0, \sigma^2)$$

# Teljes feltételesek λ –ra és η-ra

- λ eloszlása:

$$\lambda \mid \left(\{Y_t\},\{I_t\},\alpha\right) \sim \Gamma(n_0\alpha + r, \sum_{I_t=0}(Y_t - Y_{t-1}) + \beta)$$

- η eloszlása:

$$\eta \mid \left(\{Y_t\},\{I_t\},a,c\right) \sim \Gamma\left(\frac{n_1}{2} + q, \sum_{I_t=1}\left(Y_t - c - a\left(Y_{t-1} - c\right)\right)^2 + \rho\right)$$

# Teljes feltételesek *a-ra és c-re*

- C eloszlása:

$$c \mid (\{Y_t\},\{I_t\},a) \sim N\left(\frac{(1-a)\sum_{I_t=1}(Y_t - aY_{t-1})^2 + \upsilon\kappa}{n_1(1-a)^2+\kappa}, \frac{1}{n_1(1-a)^2+\kappa}\right)$$

- 

$$a \mid (\{Y_t\},\{I_t\},c) \sim N\left(\frac{\mu\tau + \eta\sum_{I_t=1}(Y_t - c)(Y_{t-1}-c)}{\tau+\eta\sum_{I_t=1}(Y_{t-1}-c)^2}, \frac{1}{\tau+\eta\sum_{I_t=1}(Y_{t-1}-c)^2}\right)$$

# Teljes feltételesek az átmenetvalószínűségekre

- $P_{00}$ eloszlása:

$$p_{00} \mid (\{I_t\}) \sim \beta(n_{00} + u_1, n_{01} + v_1)$$

- $P_{11}$ eloszlása:

$$p_{11} \mid (\{I_t\}) \sim \beta(n_{11} + u_1, n_{10} + v_1)$$

# A teljes feltételesek a rejtett állapotokra

e.g.:

$$P\left(I_t = 0 \mid I_{t-1} = 0, I_{t+1} = 0, Y_t, Y_{t-1}\right) = \frac{p_{00}^2 \Gamma\left(Y_t - Y_{t-1}\right)}{p_{00}^2 \Gamma\left(Y_t - Y_{t-1}\right) + \left(1 - p_{00}\right)\left(1 - p_{11}\right) N\left(Y_t - c - a(Y_{t-1} - c)\right)}$$

$$P\left(I_t = 1 \mid I_{t-1} = 1, I_{t+1} = 1, Y_t, Y_{t-1}\right) = \frac{p_{11}^2 N\left(Y_t - c - a(Y_{t-1} - c)\right)}{\left(1 - p_{00}\right)\left(1 - p_{11}\right)\Gamma\left(Y_t - Y_{t-1}\right) + p_{11}^2 N\left(Y_t - c - a(Y_{t-1} - c)\right)}$$

# Az $\alpha$ felújítása

- A Metropolis-Hastings lépések során $\alpha$ felújítására normális eloszlást alkalmaztunk: $\alpha^* \sim N(\alpha,\delta^2)$

# Mi történik, ha az állapot-folyamat mégsem Markov?

- A Markovitás következménye, hogy az egy rezsimben eltöltött idő geometriai eloszlású
- Egy lehetséges általánosítás:
  - Legyen a növekvő rezsimek hosszának eloszlása: NegBin($\beta$,$p_1$)
  - A csökkenő rezsimek maradjanak geometriai eloszlásúak: Geom($p_0$)
  - A rezsimhosszak maradjanak függetlenek
- Speciális eset: $\beta$=1 (az előző modellt nyerjük)

# Az általánosított modell becslése: 1. megközelítés

- A strukturális paraméterek: $\underline{\theta} = (\alpha, \lambda, \eta, a, c, p_0, \beta, p_1)$
- A posterior a megfelelő módon változtatható
- $\beta$-ra gamma eloszlású priort feltételézünk
- Minden lépés ugyanaz, mint az előző modellben, kivéve $\beta$-t, $p_{1-t}$ $I_t$-ket
- $\beta$ Metropolis-Hastings lépésekkel újítjuk fel
- Az $I_t$-k eloszlása már nem csak $Y_{t-1}$, $Y_t$, $I_{t-1}$ és $I_{t+1}$-től függ, hanem a t utáni 0 és 1 futamok hosszától is.

# Az általánosított modell becslése: A „reversible jump” megközelítés

- Az $I_t$-k helyett vezessük be a rezsimek váltáspontjait, mint látens változókat:
  - A csökkenő periódusok végpontjai: $\underline{\mathbf{s}}=(s_1,...,s_k)$
  - A növekvő periódusok végpontjai: $\underline{\mathbf{t}}=(t_1,...,t_k)$
- A kezdőértékeket és a váltáspontok számát az előző modellből becsültük meg
- Ha ismerjük a váltáspontokat, $\{I_t\}$ meghatározható, így az előző algoritmus nagy része alkalmazható
- Az új probléma és feladat: a váltáspontok felújítása

# És ha a váltáspontok száma is paraméter?

- A valóságban a váltáspontok számát nem tudjuk előre meghatározni
- A megoldás: legyen ez is paraméter és a strukturális változókhoz hasonlóan újítsuk fel ezt is
- A standard MCMC technikák itt már nem alkalmazhatóak, itt „reversible jump” MCMC-t kell felépíteni (see Green, 1995)

# A „reversible jump” algoritmus

- 3 lépést definiáltunk:
  - Step 1. : Minden paramétert felújítunk, kivéve a váltáspontok számát
  - Step 2. : „Születési” lépés: egy új váltáspontot szúrunk be két véletlenszerűen kiválasztott már meglévő váltáspont közé
  - Step 3. : „Halál” lépés : egy véletlenszerűen kiválasztott váltáspontot törlünk
- A lépéseket a következő valószínűségekkel választjuk: $q_1,q_B,q_D$ , $q_1+q_B+q_D=1$.

# A váltáspontok felújítása az első lépésben

- Teljes feltételes az $\underline{s}$ -re:

$$P\left(s_i = t_{i-1} + x \mid t_{i-1}, t_i, \{Y_t\}\right) =$$

$$\frac{P\left(\{Y_t\}_{t_{i-1} \le t \le t_i} \mid t_{i-1}, t_i, s_i = t_{i-1} + x\right) P\left(s_i = t_{i-1} + x \mid t_{i-1}, t_i\right)}{\sum_y P\left(\{Y_t\}_{t_{i-1} \le t \le t_i} \mid t_{i-1}, t_i, s_i = t_{i-1} + y\right) P\left(s_i = t_{i-1} + y \mid t_{i-1}, t_i\right)}$$

- Ahol a második tag : ($N_1$ negatív binomiális, $N_2$ geometriai eloszlású)

$$P\left(s_i = t_{i-1} + x \mid t_{i-1}, t_i\right) = P\left(N_1 = x\right) P\left(N_1 + N_2 = t_i - t_{i-1}\right) =$$

$$\frac{NegBin\left(x, b, 1 - p_{00}\right) Geom\left(t_i - t_{i-1} - x, 1 - p_{11}\right)}{\sum_y NegBin\left(y, b, 1 - p_{00}\right) Geom\left(t_i - t_{i-1} - y, 1 - p_{11}\right)}$$

# Elfogadási valószínűségek a második és harmadik lépésben

- A születés lépésében az alábbi mozgási típus elfogadásának valószínűsége: $(t,s) \rightarrow (t^*,s^*)$

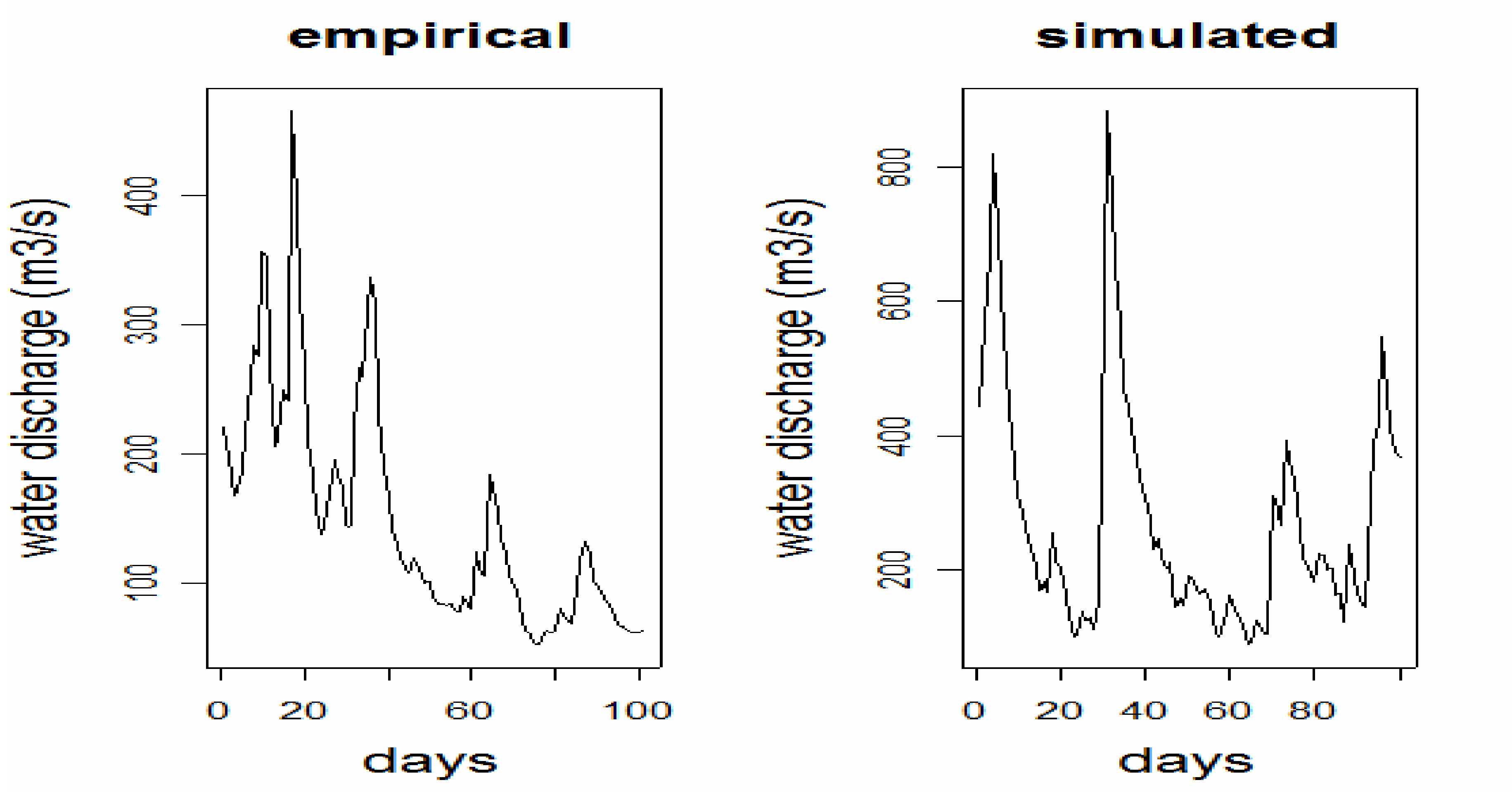
$$\alpha_{birth} = \min\left(1, \frac{f\left(\underline{\theta}, k+1, \underline{t}^*, \underline{s}^* \mid \{Y_t\}\right)}{f\left(\underline{\theta}, k, \underline{t}, \underline{s} \mid \{Y_t\}\right)} \frac{\frac{q_D}{2k+1}}{\frac{q_B}{l}} \frac{2}{(m-2)(m-1)}\right)$$

- A „halál” lépésben a fenti valószínűség inverze lesz az elfogadási esély

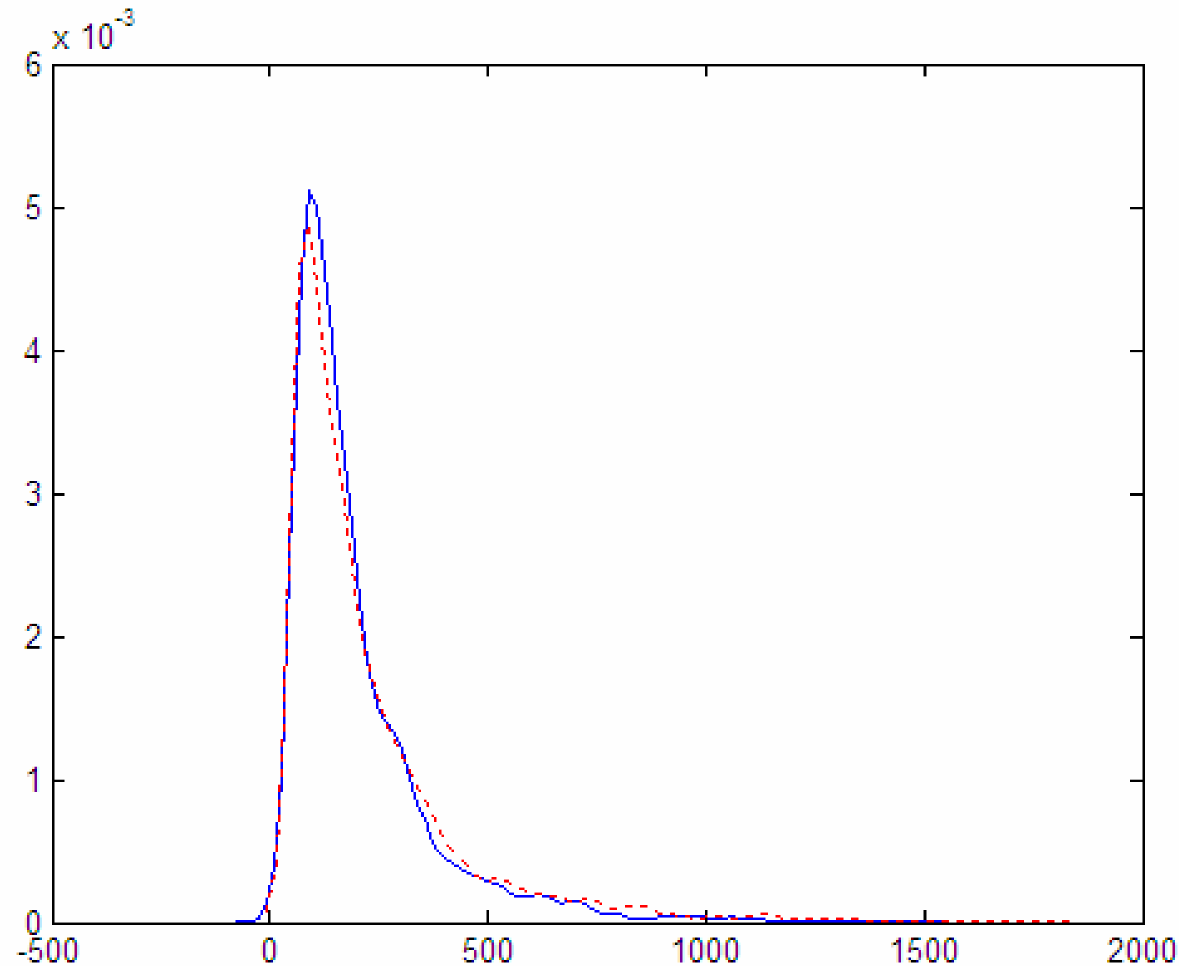
# Eredmények Tivadar esetében

- Mindkét megközelítést kipróbáltuk, a fix dimenziós jobbnak bizonyult
- $\beta = 4.77$ (s.e. 1.38); szignifikánsan különbözik 1-től
- A növekvő rezsimek átlagos hossza 2.5 nap
- A csökkenő rezsimek átlagos hossza 14.4 nap
- $a = 0.815$ (s.e. 0.0023); magas perzisztencia a csökkenő rezsimben
- $\alpha = 0.974$ (s.e. 0.072); a növekvő rezsime majdnem ecponenciális

# A modell visszaadja a hidrográf aszimmetrikus alakját



# Az eredeti (red) és a szimulált(kék) idősorok sűrűségfüggvénye



Köszönöm a figyelmet!