

PPKE ITK
Pázmány Péter Katolikus Egyetem
Információs Technológiai és Bionikai Kar

Sztochasztikus folyamatok jegyzet

előadó: Várdainé Kollár Judit

Lektorálta: dr Rásonyi Miklós
Lejegyezte: Jánossy Bálint

Statisztikai minta: : Valamely ξ **valószínűségi változó** a vonatkozó véges számú független, azonos körülmények között végrehajtott megfigyelés: $\xi_1, \xi_2, \dots, \xi_n$ eredményét nevezzük statisztikai mintának. ξ_i a minta i . eleme, n a minta nagysága.

A statisztikai minta tehát véges sok, független, azonos eloszlású **valószínűségi változó** g. yüttese. A valóságban nem feltétlenül valósul meg a függetlenség, (vízállás mérése havi/napi rendszerességgel). Vannak módszerek a függetlenség megállapítására. (erről később)

Mintaátlag: $\bar{\xi} = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n} = \frac{\sum_{i=1}^n \xi_i}{n}$

Tapasztalati szórásnégyzet: $s_n^2 = \frac{(\xi_1 - \bar{\xi})^2 + (\xi_2 - \bar{\xi})^2 + \dots + (\xi_n - \bar{\xi})^2}{n} = \frac{\sum_{i=1}^n (\xi_i - \bar{\xi})^2}{n}$

Determinisztikus folyamat: Ha bizonyos körülmények együttese fennáll, az egyértelműen meghatározza az esemény lefolyását.

Megjegyzés: (pl. 1 bar nyomáson a tiszta víz 100 °C forr)

Sztochasztikus folyamat: Valamelyik szükséges körülmény hiányzik, az esemény lefolyása bizonytalanná válik.

Megjegyzés: (pl. x bar nyomáson($x=?$)) a tiszta víz már nem biztos, hogy 100 °C forr)

Véletlen jelenség: A rendelkezésre álló adatok nem határozzák meg egyértelműen az események lefolyását.

Statisztikai minta: Véges sok, egymástól független, azonos eloszlású valószínűségi változó együttese. $(\xi_1, \xi_2, \dots, \xi_n)$

Minta nagysága: Minta elemeinek száma. *Megjegyzés:* (pl. $n=100$ rúd, $n=70$ év)

Empirikus(tapasztalati) jellemzők:

- mintaközép/mintaátlag: $\hat{\xi} = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n}$

- szórásnégyzet: $s_n^2 = \frac{\sum_{i=1}^n (\xi_i - \hat{\xi})^2}{n}$

- korrigált tapasztalati szórásnégyzet: $s_n^{*2} = \frac{\sum_{i=1}^n (\xi_i - \hat{\xi})^2}{n-1} = \frac{n}{n-1} \cdot s_n^2$

Becslélmélet: ξ valószínűségi változó, ismerjük az eloszlás és a sűrűség függvényének a típusát. Ekkor:

$$f(x) = f(x, \Theta_1, \Theta_2, \dots, \Theta_n)$$

Torzítatlan becslés: $\tilde{\Theta} = \tilde{\Theta}(\xi_1, \xi_2, \dots, \xi_n)$ függvény és $E(\tilde{\Theta}) = \Theta$, azaz a $\tilde{\Theta}$ várható értéke megegyezik a becsült paraméter értékével.

Megjegyzés: (ahol Θ statisztika a mi folyamatunk valószínűségi változója, melynek értékei a Θ a becsült paraméter értéke körül szóródnak. Tudni szeretnénk, hogy $\tilde{\Theta}$ nagy valószínűséggel legfeljebb mekkora távolságra lehet Θ -tól.)

Megbízhatóság(konfidencia) intervallum: Tfh. $\tilde{\Theta}$ torzítatlan becslés, Θ ismeretlen. A $[\alpha_1, \alpha_2)$ intervallumot megbízhatósági intervallumnak nevezzük, ha adott (kicsi) $\varepsilon > 0$ -ra $P(\alpha_1 \leq \Theta < \alpha_2) = 1 - \varepsilon$.

Megjegyzés: ahol $P(\alpha_1 \leq \Theta < \alpha_2)$ annak a valószínűsége, hogy $\Theta \in [\alpha_1, \alpha_2)$

Maximum-likelihood módszer: Tfh. $\xi_1, \xi_2, \dots, \xi_n$ független, azonos eloszlású valószínűségi változó, sűrűség függvénye f_{ξ_i} . Ekkor:

$$f_{\xi_1, \xi_2, \dots, \xi_n}(x_1, x_2, \dots, x_n) \stackrel{\text{független}}{=} \prod_{i=1}^n f_{\xi_i}(x_i) \xrightarrow{jel} L(\Theta, x_1, x_2, \dots, x_n)$$

$\tilde{\Theta} = L(\Theta, x_1, x_2, \dots, x_n)$ -t nevezzük Likelihood fv-nek.

Megjegyzés: Több lehetőség közül azt fogadjuk el egy ismeretlen paraméter valódi értékéül, amely mellett a ténylegesen bekövetkező eseménynek a legnagyobb a valószínűsége.

Statisztikai hipotézisnek nevezzük a valószínűségi változók

- eloszlására vagy
- valamely paraméterére vonatkozó állításokat. (H_0, H_1)

Dönteni akarunk az állítás helyességéről: H_0 hipotézisünk igazságáról.

Statisztikai próba feladata, hogy az X valószínűségi változóra vonatkozó statisztikai minta alapján döntsünk a H_0 hipotézisről. H_0 -t vagy elfogadjuk, vagy elutasítjuk/elvetjük.

Statisztikai próba lényege: K kritikus tartomány megválasztása úgy, hogy H_0 mellett számított $U = u(x_1, \dots, x_n)$ vagy $T = t(x_1, \dots, x_n)$ statisztikai érték kis valószínűséggel essen a K tartományba. (feladattól függ, hogy u -t vagy t -t számolunk)

Annak a valószínűségét, hogy H_0 fennáll és a próbastatisztika a K tartományba esik a **próba terjedelmének** nevezzük. (jel.: $P(u \in K) = \alpha$)

Hipotézis vizsgálat menete:

- Mintavétel
- H_0 hipotézis megalkotása, (ebből a) H_1 ellenhipotézis megfogalmazása
- Megadunk $0 < \alpha < 1$ korlátot, az elsőfajú hiba(H_0 -t elvetjük, bár igaz) valószínűségére ($P(u \in K) = \alpha$)
- Konstruálunk egy statisztikai függvényt (próbafüggvényt) a minta elemeiből, illetve a H_0 hipotézisből
- Meghatározzuk a próbafüggvény eloszlását, ha H_0 teljesül
- Az 5. lépésben kapott eloszlás alapján meghatározzuk E (elfogadási), és K (kritikus) tartományokat ($E \cup K = \mathbb{R}$ és $E \cap K = \emptyset$)

Az elfogadási tartomány tipikusan lehet:

- $E = [-a_\alpha, a_\alpha]$
- $E = [-a_\alpha, \infty)$
- $E = (-\infty, a_\alpha]$

Ha $u \in E$ (vagy $t \in E$) elfogadjuk H_0 -t.

Ha $u \in K$ (vagy $t \in K$) elutasítjuk H_0 -t, és H_1 -et fogadjuk el.

Elsőfajú hiba: H_0 igaz, de $u(t) \in K$, ezért H_0 -t elvetjük.

Másodfajú hiba: H_0 nem igaz (H_1 igaz), mégis $u(t) \in E$, tehát H_0 -t elfogadjuk.

	H_0 -t elfogadjuk	H_0 -t elutasítjuk
H_0 igaz	✓	elsőfajú hiba
H_0 hamis	másodfajú hiba	✓

(✓ = nincs hiba, minden renben)

Az elsőfajú hiba csökkentése a másodfajú hiba növekedésével jár.

u próba:

Adott n elemű minta $\xi_1, \dots, \xi_n$, amely $N(m, \sigma)$ eloszlású és független, σ ismert.
Paraméteres próba eloszlás valamelyik paraméterére vonatkozó hipotézis igazságát vizsgáljuk.

- 1 mintás u próba:

H_0 hipotézis: $E(\xi) = m$.

$0 < \alpha < 1$

Kétoldali próba	$H_0 : E(\xi) = m$	$E = [-a_\alpha, a_\alpha]$
	$H_1 : E(\xi) \neq m$	$K = \mathbb{R} \setminus [-a_\alpha, a_\alpha]$
Egyoldali próba	$H_0 : E(\xi) = m$	$E = (-\infty, a_\alpha]$
	$H_1 : E(\xi) > m$	$K = [a_\alpha, \infty)$
	$H_0 : E(\xi) = m$	$E = [-a_\alpha, \infty)$
	$H_1 : E(\xi) < m$	$K = (-\infty, -a_\alpha]$

Próba függvény: $u = \frac{\bar{\xi} - m}{\sigma/\sqrt{n}}, \quad N(0, 1)$

Szignifikancia szint(legtöbbször): $\alpha \cong 0,01; 0,05$ vagy $0,1$ (elsőfajú hiba nagysága)

- 2 mintás u próba:

pl. gyárban 2 futószalag

Függetlennek kell lennie a 2 mintának(pl. nem független, ha 1 futószalag, csak 2 különböző napon)

$\xi_1, \dots, \xi_{n_1}$ és $\eta_1, \dots, \eta_{n_2}$ független minták, $N(m_1, \sigma_1)$ és $N(m_2, \sigma_2)$ eloszlásúak, valamint σ_1, σ_2 ismert.

Kétoldali próba	$H_0 : m_1 = m_2$	$E = [a_{-\alpha}, a_\alpha]$
	$H_1 : m_1 \neq m_2$	$K = \mathbb{R} \setminus [a_{-\alpha}, a_\alpha]$
Egyoldali próba	$H_0 : m_1 = m_2$	$E = (-\infty, a_\alpha]$
	$H_1 : m_1 > m_2$	$K = \mathbb{R} \setminus [a_\alpha, \infty)$
	$H_0 : m_1 = m_2$	$E = [-a_\alpha, \infty)$
	$H_1 : m_1 < m_2$	$K = \mathbb{R} \setminus (-\infty, -a_\alpha]$

Próba függvény: $u = \frac{\bar{\xi} - \bar{\eta}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}, \quad N(0, 1)$

Student féle t próba:

Normális eloszlású valószínűségi változó várható értékére vonatkozó hipotézis, ismeretlen szórás esetén.

A szórás ismeretlen, becsüljük az $s_n^* = \frac{\sqrt{\sum_{i=1}^n (\xi_i - \hat{\xi})^2}}{n-1}$ korrigált tapasztalati szórással. Kétmintás próbát megelőzi az **F** próba. *Megjegyzés:* lásd később

- 1 mintás t próba:

$\xi_1, \dots, \xi_n$ eloszlása ismert $N(m, \sigma)$, $-\infty < m < \infty$; $\sigma > 0$.

H_0 hipotézis: $E(\xi) = m$

próbafüggvény(próbastatisztika): $t = \frac{\bar{\xi} - m}{s_n^*/\sqrt{n}}$

Az n elemű mintát $(n-1)$ szabadság fokúnak nevezzük.

egy oldali: $P(|t| > t_\alpha) = \alpha$

két oldali : $P(|t| > t_{\frac{\alpha}{2}}) = \alpha$ $1 - \alpha$ szignifikancia szinten

A formulából látható, hogy t eloszlása nem függ az ismeretlen σ - tól, ezért az elsőfajú hiba valószínűsége is az. Bármely σ esetén a próba terjedelme α .

- 2 mintás t próba:

$\xi_1, \dots, \xi_{n_1}$ és $\eta_1, \dots, \eta_{n_2}$ független minták, σ_1, σ_2 ismeretlen.

Próba függvény: $t_{(n_1-1+n_2-1)} = \frac{\bar{\xi} - \bar{\eta}}{\sqrt{(n_1-1)s_1^{*2} + (n_2-1)s_2^{*2}}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}$

Ism. u, t próbát alkalmazunk, ha a szórásuk megegyezett.

Fisher féle F próba:

$$\left. \begin{array}{l} \xi \\ \eta \end{array} \right\} \begin{array}{ll} n \text{ db minta} & \text{Normál eloszlású,} \\ m \text{ db minta} & \text{független minták} \end{array}$$

$$H_0 : \sigma_\xi = \sigma_\eta$$

$$H_1 : \sigma_\xi \neq \sigma_\eta$$

$$F_{f_\xi, f_\eta}^* = \frac{\max(s_\xi^{*2}, s_\eta^{*2})}{\min(s_\xi^{*2}, s_\eta^{*2})} > 1$$

F_{f_ξ, f_η}^* szabadságfoka:

- $(s_\xi^{*2} > s_\eta^{*2} \text{ esetén}) : (n - 1)$
- $(s_\eta^{*2} > s_\xi^{*2} \text{ esetén}) : (m - 1)$

$1 - \alpha$ szignifikancia szinten:

$F_{f_\xi, f_\eta}^* < F_{1-\epsilon} \rightarrow H_0$ -t elfogadjuk, azonos a két minta szórása.

$F_{f_\xi, f_\eta}^* \geq F_{1-\epsilon} \rightarrow H_0$ -t elvetjük, a két mintának nem azonos a szórásnégyzete,

Eddig paraméteres próbákat néztünk. u, t próbáknál az m volt azonos, F -nél a σ .

Eddig a paraméterek helyességét, mostantól magát az eloszlást vizsgáljuk.

Nem paraméteres próbák: χ^2 próbák

Legyenek $\xi_1, \xi_2, \dots, \xi_n$, $N(0, 1)$, független valószínűségi változók. Legyen $\eta = \xi_1^2 + \dots + \xi_n^2$. Ekkor η eloszlását n szabadságfokú **χ^2 eloszlásnak** nevezzük. $E(\eta) = n$ és $D^2(\eta) = 2n$

Megjegyzés: A Centrális határeloszlás-tétel alapján nagy n -re a χ^2 eloszlása normális eloszlással közelíthető.

Legyen $A_1, \dots, A_r$ teljes eseményrendszer, azaz: $\sum_{i=1}^r A_i = I$ és $A_i \cap A_j = \emptyset$ ahol $i \neq j$ és I a teljes esemény halmaza.

Vizsgáljuk azt a H_0 hipotézist, hogy: $P(A_i) = p_i$. Ahol p_i $i = 1, \dots, r$ nevezetes eloszlás, és $\sum_{i=1}^r p_i = 1$

Ezt a hipotézis vizsgálatot nevezzük **χ^2 próbának**.

Tfh. n kísérlet során A_i ν_i -szer ($\nu = \text{nű}$) következik be. $\sum_{i=1}^r \nu_i = n$

ν_i valószínűségi változók binomiális eloszlásúak. (A_i vagy bekövetkezik, vagy nem.) $E(A_i) = np_i$

$$P(\nu_1 = k_1, \nu_2 = k_2, \dots, \nu_r = k_r | H_0) \stackrel{\text{f.l.n.}}{=} P(\nu_1 = k_1) \cdot P(\nu_2 = k_2) \cdot \dots \cdot P(\nu_r = k_r) =$$

$$\frac{n!}{k_1! k_2! \dots k_r!} p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}, \text{ ahol } \sum_{i=1}^r k_i = n$$

Ez az un. **Polinomiális eloszlás**.

Megj.: Két **valószínűségi változó** s etén a binomiális eloszlást kapjuk.

Képezzük a következő statisztikát:

$$\chi^2 = \sum_{i=1}^r \frac{(\nu_i - np_i)^2}{np_i}$$

Amennyiben a H_0 hipotézis igaz, akkor $E(\nu_i) \approx np_i$. Azaz, a számláló minden tagjában a ν_i értékének a saját várható értékétől vett távolságának négyzete szerepel, azaz ha a χ^2 kicsi, akkor jó a próba.

χ^2 : r db standardizált binomiális eloszlású valószínűségi változó négyzetének összege.

Megj.: A binomiális eloszlás nagy n és kicsi p esetén közelíthető normál eloszlással. ($np > 10$)

Tetszőleges $\alpha > 0$ -hoz meghatározható (χ^2 táblázatból) olyan χ^2_{r-1} kritikus érték, melyre: $P(\chi^2 < \chi^2_{r-1}(\alpha)) = 1 - \alpha$

Ezt nevezzük $r - 1$ **szabadságfokú χ^2 eloszlásnak**.

Ahány paraméter becsült az eloszlásból, az annyival csökkenti a szabadságfokát (az $(r - 1)$ -et).

(pl. ha becsüljük a λ -t a Poisson eloszlásnál, a szabadságfok $r - 2$ lesz)

Illeszkedés vizsgálat

Van egy megfigyelt adatunk, és ehhez feltételezünk egy eloszlást.

Megfigyelt valószínűségi változók eloszlása megegyezik-e az előre megadott eloszlással.

Lehetnek diszkrét vagy folytonos **valószínűségi változó** .

Diszkrét v.v. esetén

Adott (x_i, p_i) elméleti $i = 1, \dots, r$; valamint $q_i = P(\xi_i = x_i)$ tapasztalati eloszlás

Megfigyelés: $\xi_1, \dots, \xi_r$ valószínűségi változók függetlenek és azonos eloszlásúak.

Jelöljük $q_1, \dots, q_r$ -val az eloszlásukat. (Azt adják meg, hogy x_i -t milyen valószínűséggel veszik fel.)

$\nu_i = \{m : \xi_i = x_i\}$ (Megadja annak a gyakoriságát, hogy a megfigyelés során hányszor vette fel az elméleti értéket)

$H_0 : q_i = p_i$ azaz a két eloszlás megegyező $\forall i = 1, \dots, r$ -re.

$H_1 : q_i \neq p_i$ legalább egy i -re.

Ha $\nu_i \approx np_i$, alkalmazhatjuk a χ^2 próbát.

$$q(\xi_1, \dots, \xi_r) = \frac{\sum_{i=1}^r (\nu_i - np_i)^2}{np_i}$$

(Ahol ν_i azt adja meg, hogy gyakorlatilag hányszor következett be. Az np_i pedig, hogy az elméleti eloszlás szerint hányszor következett be.)

Ez egy $(r - 1)$ szabadságfokú χ^2 eloszlás.

$P(H_0 \text{ igaz}; g > c_r) = \alpha$, ahol g a kiszámolt, c_r az elméleti(táblázatból) (első-fajú hiba)

Az α -t adjuk meg, így minimalizáljuk az elsőfajú hibát.

Ha $g > c_r$ elvetjük H_0 -t
 Ha $g < c_r$ elfogadjuk H_0 -t

Folytonos v.v. esetén:

Ha az eseményeket $(-\beta, \beta)$ intervallumon vizsgáljuk, akkor ezt felosztjuk r db részintervallumra, és ezeken nézem a ν_i -t és az np_i -t. (r részintervallum esetén $i = 1, \dots, r$) Majd erre alkalmazom a χ^2 próbát.

Homogenitás vizsgálat

Célja: annak eldöntése, hogy két valószínűségi változó: ξ, η egyforma eloszlású-e. Vagyis az ξ -re vonatkozó $\xi_1, \dots, \xi_n$, valamint az η -ra vonatkozó $\eta_1, \dots, \eta_m$ statisztikai minta ugyanazon eloszlású sokaságból származik-e.

$H_0 : P(\xi < x) = P(\eta < x), \quad x \in \mathbb{R},$ azaz: $p_i = q_i, \forall i$ -re.

$H_1 : \exists i, \text{ hogy } p_i \neq q_i$

$\left. \begin{matrix} \xi_1, \dots, \xi_n \\ \eta_1, \dots, \eta_m \end{matrix} \right\}$ minták függetlenek és azonos eloszlásúak $\begin{matrix} \nu_i \\ \mu_i \end{matrix} \quad \left\{ \begin{matrix} \text{az } i \text{ intervallumba eső} \\ \text{minta elemek száma/gyakorisága} \end{matrix} \right.$

$$\chi^2 = n \cdot m \sum_{i=1}^r \left(\frac{\left(\frac{\nu_i}{n} - \frac{\mu_i}{m} \right)^2}{\frac{\nu_i}{n} + \frac{\mu_i}{m}} \right) \quad r - 1 \text{ szabadságfokú } \chi^2 \text{ eloszlást ad.}$$

$$\sum_{i=1}^r \nu_i = n, \quad \sum_{i=1}^r \mu_i = m, \quad I\text{-t } r \text{ részre osztjuk.}$$

Függetlenség vizsgálata

Célja: Két valószínűségi változó, ξ, η között van-e kapcsolat, vagy függetlenek egymástól.

azaz: $A_1, \dots, A_r$ és $B_1, \dots, B_s$ teljes eseményrendszer

H_0 : $P(A_i \cap B_j) = P(A_i)P(B_j)$ azaz A_i, B_j független $\forall i = 1, \dots, r, j = 1, \dots, s$ -re.

H_1 : nem függetlenek.

$\nu_{ij} = A_i \cap B_j$ gyakorisága n független kísérletben.

Próbastatisztika:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(\nu_{ij} - np_i q_j)^2}{np_i q_j} = \sum_{i=1}^r \sum_{j=1}^s \frac{\left(\nu_{ij} - \frac{\nu_i \nu_j}{n}\right)^2}{\frac{\nu_i \nu_j}{n}}$$

Ha $\chi^2_{(r-1)(s-1)} < c_k$, elfogadjuk $1 - \alpha$ szinten H_0 -t

Hipotézis vizsgálat vége, sztochasztikus folyamatok/Idősorok kezdete

Sztochasztikus folyamatok

Def: Legyen $x_t; t \in N$ valószínűségi változók halmaza. Ezt **diszkrét idejű** sztochasztikus folyamatnak nevezzük.

Megjegyzés: $t = 0 : x_0$ - most mérem, vizsgálom, észlelem

$t > 0 : x_t$ - jövő

Def: Legyen $x_t; t \in Z$ valószínűségi változók halmaza, diszkrét idejű sztochasztikus folyamat.

$$x_t \begin{cases} t \leq 0 : \text{múlt és jelen} & \text{Ezt vizsgálom, mérem, megfigyelem..} \\ t > 0 : \text{jövő} & \text{Ezt akarjuk becsülni, sejteni.} \end{cases}$$

t bármi lehet (pl. év, hónap, perc, stb.)

Def: $x_t : \Omega \rightarrow R$ függvény valamilyen véletlen esemény értéke a t időpontban.

$\omega \in \Omega$ - az eseménytér egy fix állapota.

$x_0(\omega), x_1(\omega), \dots, x_n(\omega)$ számsorozat a folyamat $\omega \in \Omega$ -hoz tartozó trajektóriája/megvalósulása.

Példa: ε_t - fehér zaj, ha független, azonos eloszlású valószínűségi változót jelöl.

Általában: $E(\varepsilon_t) = 0$

$D^2(\varepsilon_t) = c, c \in R^+$; gyakran $c = 1$

Példa: $x_t; t \in N$ kockadobás értéke $P(x_t = l) = \frac{1}{6}, l = 1, \dots, 6$ esetén.

Példa: Szimmetrikus véletlen bolyongás

"kis átmérőjű csőben bolyongtatunk egy részecskét"

Legyen x_j egy lépés jelölése.

$$\left. \begin{aligned} P(x_j = -1) &= \frac{1}{2} & (\text{balra léptem}) \\ P(x_j = 1) &= \frac{1}{2} & (\text{jobbra léptem}) \end{aligned} \right\} \Rightarrow x_j\text{-k azonos eloszlásúak, függetlenek}$$

Általános:

$$\left. \begin{aligned} P(x_j = x_t - 1) &= \frac{1}{2} & (\text{balra léptem}) \\ P(x_j = x_t + 1) &= \frac{1}{2} & (\text{jobbra léptem}) \end{aligned} \right\} \Rightarrow x_j\text{-k azonos eloszlásúak, függetlenek}$$

Ennek módosítása:

- amikor 2-t léphetek balra vagy jobbra
- nem szimmetrikus a bolyongás, azaz a balra ill. jobbra lépés valószínűsége nem egyforma.

Def: Az időben változatlan viselkedésű, egyensúlyban lévő rendszereket nevezzük **stacionáriusnak**.

(Pl. eső esésének valószínűsége + eső mennyiség várhatóérték)

Def: Egy x_t ; $t \in Z$ folyamatot **erősen stacionáriusnak** nevezünk, ha $\forall k \in Z$ -re az $(x_0, \dots, x_n)$ eloszlása megegyezik $(x_{0+k}, \dots, x_{n+k})$ eloszlásával. Azaz az eloszlására teljesül az *időinvariáns tulajdonság*.

Def: Az x_t ; $t \in Z$ **gyengén stacionárius** folyamat, ha

- $E(x_t^2) < \infty \quad \forall t \in Z$
- $E(x_t) = E(x_0) \quad \forall t \in Z$
- $cov(x_0, x_m) = cov(x_{0+k}, x_{m+k}) \quad \forall k, m \in Z$ (időinvaráns)

A $cov(x_k, x_m)$ a k . és m . időpontban lévő folyamatot hasonlítja össze, ez csak k és az m távolságától függ (pl. $t-s=1$).

Def: Legyen x_t ; $t \in Z$ gyengén stacionáris folyamat, ekkor az $R_x(k) = cov(x_0, x_k)$, $k \in Z$ **kovariancia függvénynek** nevezzük.

Állítás: Az így definiált kovariancia függvény tulajdonságai:

- $R_x(0) = D^2(x_t)$ bármely $t \in Z$
- $R_x(k) = R_x(-k)$ (páros függvény)
- $cov(x_k, x_m) = R_x(m - k)$

Biz:

- $R_x(0) = cov(x_0, x_0) \xrightarrow{\text{gyengén stac.}} cov(x_{0+t}, x_{0+t}) = cov(x_t, x_t) = D^2(x_t)$
- $R_x(k) = cov(x_0, x_k) \xrightarrow{\text{idő inv.}} cov(x_{0-k}, x_{k-k}) = cov(x_{-k}, x_0) = R_x(-k)$
- $cov(x_k, x_m) \xrightarrow{-k} cov(x_{k-k}, x_{m-k}) = cov(x_0, x_{m-k}) = R_x(m - k)$

Ezen tulajdonságokból adódóan:

(Emlékeztető: ε_t fehér zaj, ha $E(\varepsilon_t) = 0$, $D^2(\varepsilon_t) = \sigma^2 > 0$)

fehér zaj tulajdonságai:

- $cov(x_0, x_0) = \sigma^2$

- $cov(x_i, x_j) = 0 \quad i \neq j$
- Az előző átalakítva: $R_x(i - j) = 0$

Fehér zaj pl. $\varepsilon_t \rightarrow N(0, 1)$

Mivel függetlenek $\rightarrow cov(\varepsilon_i, \varepsilon_j) = 0$

Mozgó átlag - MA folyamat

Legyen $x_t = \sum_{j=-m}^n \alpha_j \varepsilon_{t-j} \quad t \in \mathbb{Z}; \varepsilon_t$ fehér zaj, $\alpha_j \in \mathbb{R}$ együtthatók.

Példa₁: $x_t = 3\varepsilon_{t-2} + 4\varepsilon_{t-1} + 3\varepsilon_t + 2\varepsilon_{t+1}$

Példa₂: Legyen $n = m$, $\alpha_j = \frac{1}{2n+1} \quad (\forall -m \leq j \leq m)$

Ekkor: $x_t = \frac{1}{2n+1} \sum_{j=-m}^m \varepsilon_{t-j}$.

Általánosan:

$$\sum_{j \in \mathbb{Z}} \alpha_j^2 < \infty$$

$x_t = \sum_{j=-\infty}^{\infty} \alpha_j \varepsilon_{t-j}$ **végtelen MA** folyamat.

Kauzális MA folyamat:

ha $\forall \alpha_j = 0 \quad j < 0$ esetén, akkor $x_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}$

A jövő csak a jelentől és a múltól függ.

$$R_x(k) = \sum_{j=0}^{\infty} \alpha_j \alpha_{j+k} \sigma^2$$

Def: Azt mondjuk, hogy x_t **autoregresszív (AR)** folyamat, ha x_t gyengén stacionárius és előáll $x_{t+1} = \Phi x_t + \varepsilon_{t+1}$ rekurzióval, ahol $\Phi \in \mathbb{R}$, ε_{t+1} fehér zaj.

Megj. Ha $\Phi > 1$ x_t Φ ütemben nő,

ha $\Phi < 1$ x_t Φ ütemben csökken.

Példa₃

x_t - a Balaton vízállása t időpontban

$x_{t+1} = 0,9x_t + \varepsilon_{t+1}$, ahol ε_{t+1} a csapadék ($\Phi < 1$ mert a tó vize párolog).

Def: Legyen Y_n valószínűségi változók sorozata és Y valószínűségi változó.

$$E(Y_n^2) < \infty; E(Y^2) < \infty$$

Azt mondjuk, hogy Y_n L^2 értelemben tart Y -hoz ($Y_n \rightarrow Y$), ha $E((Y_n - Y)^2) \rightarrow 0, \quad (n \rightarrow \infty)$.

2 fajta idősor:

- determinisztikus
 - trend
 - szezonális
 - ciklus
- sztochasztikus
 - független valószínűségi változó(ka)t vetünk össze az elmúlt időszak folyamatai

Nincsenek előre adataim, a jövőt akarom jelezni/jósolni

Például napi hozam

Időtől független E, σ^2

Ötös lottó átlagos húzás

Fehér zaj $cov = 0$ E ált.

Azt feltételezzük, hogy az idősor stacionárius, nem tartalmaz trend hatást, és az időbeli sztochasztikus kapcsolatok állandóak.

Def: Y_t, Y valószínűségi változók $E(Y_t^2) < \infty$ és $E(Y^2) < \infty$.

Y_n L^2 értelemben tart Y -hoz, ha $\lim_{n \rightarrow \infty} E((Y_n - Y)^2) = 0$

Tétel: Legyen $|\phi| < 1$, ε_t fehér zaj, ekkor $\exists x_t, t \in \mathbb{Z}$ AR-nak MA előállítás, melyre $\lim_{k \rightarrow \infty} E((x_t - \sum_{k=0}^{\infty} \Phi^k \varepsilon_{t-k})^2) = 0$, azaz x_t előáll ε_k ∞ soraként (lineáris kombinációjaként).

Biz:

$$x_t = \phi x_{t-1} + \varepsilon_t = \phi(\phi x_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \phi(\phi(\phi x_{t-3} + \varepsilon_{t-2}) + \varepsilon_{t-1}) + \varepsilon_t = \dots =$$

$$= \phi^{n+1} x_{t-(n+1)} + \phi^n \varepsilon_{t-n} + \phi^{n-1} \varepsilon_{t-(n-1)} + \dots + \varepsilon_t =$$

$$= \phi^{n+1} x_{t-n-1} + \sum_{k=0}^n \phi^k \varepsilon_{t-k}$$

$$E[(x_t - \sum_{n=0}^{\infty} \phi^n \varepsilon_{t-n})^2] = E[(\phi^{n+1} x_{t-(n+1)})^2] = \lim_{n \rightarrow \infty} \phi^{2n+2} E[(x_{t-(n+1)})^2] = \lim_{n \rightarrow \infty} \phi^{2n+2} E[(x_0)^2] = 0$$

Állítás: Az $x_t = \Phi x_{t-1} + \varepsilon_t$ AR folyamat csak ε_t jelen és múlt értékeitől függ, így x_t és ε_{t+k} korrelálatlanok, függetlenek, $\forall k \geq 1$ -re.

x_t és az ε_{t+k} korrelálatlanok $k \geq 1$.

$$cov(x_t, \varepsilon_{t+k}) = E[x_t \varepsilon_{t+k}] - E[x_t] E[\varepsilon_{t+k}] = 0 \text{ Tehát } cov(x_t, \varepsilon_{t+k}) = 0$$

Megjegyzés: Ez azt jelenti, hogy $cov(x_t, \varepsilon_{t+k}) = 0, \forall k \geq 1$

$$D^2(x_t) = D^2(\phi x_{t-1} + \varepsilon_t) = \phi^2 D^2(x_{t-1}) + D^2(\varepsilon_t) + \underbrace{2cov(\phi x_{t-1}, \varepsilon_t)}_0 = x_t \text{ staci-}$$

onárius, $\Rightarrow D^2(x_t) = D^2(x_{t-1})$,

valamint függetlenek, ezért a kovarianciájuk 0.

$$\text{Tehát: } D^2(x_t)[1 - \phi^2] = \sigma^2 \Rightarrow D^2(x_t) = \frac{\sigma^2}{1 - \phi^2}$$

Ez $|\phi| < 1$ -re érvényes ($D^2 > 0$)

$$E(x_t) = E(\phi x_{t-1} + \varepsilon_t) = \phi E(x_{t-1}) + \underbrace{E(\varepsilon_t)}_0$$

ε_t fehér zaj, tehát $E(\varepsilon_t) = 0$.

$$E(x_t) = E(x_{t-1})$$

Mivel $\phi < 1 \Rightarrow E(x_t) = 0$

$R_x(k)$:

Def: Legyen x_t , $t \in \mathbb{Z}$ -re tetszőleges sztochasztikai folyamat.

Ekkor $zx_t = x_{t-1}$ kifejezésben z a **visszaléptetési operátor**.

Pl.

$$(3 - z^2 - 4z + 5)x_t = 3x_{t-2} - 4x_{t-1} + 5x_t$$

$$x_t = 4\varepsilon_{t-3} + 5\varepsilon_{t-2} - 2\varepsilon_t \Rightarrow x_t = (4z^3 + 5z^2 - 2)\varepsilon_t$$

$$x_t = \phi x_{t-1} + \varepsilon_t \Rightarrow x_t(1 - \phi z) = \varepsilon_t$$

Def: *ARMA*

x_t , $t \in \mathbb{Z}$ stacionárius folyamat (p, q) -ad rendű *ARMA* folyamatnak nevezük, ha:

$$x_t + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_p x_{t-p} = \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q}$$

ahol α_0 mindig 1, $\alpha_1, \dots, \alpha_p$ és $\beta_0, \dots, \beta_q$ rögzített konstansok, valamint ε_t fehér zaj.

$$x_t + \alpha_1 z x_t + \alpha_2 z^2 x_t + \dots + \alpha_p z^p x_t = \beta_0 \varepsilon_t + \beta_1 z \varepsilon_t + \beta_2 z^2 \varepsilon_t + \dots + \beta_q z^q \varepsilon_t$$

$$x_t(1 + \alpha_1 z + \alpha_2 z^2 + \dots + \alpha_p z^p) = \varepsilon_t(\beta_0 + \beta_1 z + \beta_2 z^2 + \dots + \beta_q z^q)$$

$$A(z)x_t = B(z)\varepsilon_t$$

Tétel 1. Ha $A(z)$ polinom $\forall$ gyöke a komplex sík egységkörén kívül van, azaz $\forall |z_i| > 1$, ha $A(z_i) = 0$, akkor az x_t *ARMA* folyamatnak **IMA**(∞) előállítása:

$$x_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j}$$

Biz. Legyenek $A(z)$ gyökei $z_1, \dots, z_p \in \mathbb{Z}$ vagy $\in \mathbb{C}$, mind különböző.

Ekkor $A(z) = \alpha_p(z - z_1)(z - z_2) \dots (z - z_p)$

$$x_t = \frac{B(z)}{A(z)} \varepsilon_t = \frac{B(z)}{\alpha_p} \frac{1}{(z - z_1)(z - z_2) \dots (z - z_p)} \varepsilon_t =$$

Parciális törtekre bontható:

$$\frac{B(z)}{\alpha_p} \left[\frac{L_1}{z - z_1} + \frac{L_2}{z - z_2} + \dots + \frac{L_p}{z - z_p} \right] \varepsilon_t =$$

Minden elem mértani sorra bontható:

$$\frac{L_j}{z - z_j} = -\frac{L_j}{z_j} \frac{1}{1 - \frac{z}{z_j}} = -\frac{L_j}{z_j} \sum_{i=0}^{\infty} \left(\frac{z}{z_j} \right)^i$$

$$= \frac{B(z)}{\alpha_p} \left[\sum_{j=1}^p \left(-\frac{L_j}{z_j} \right) \sum_{i=0}^{\infty} \left(\frac{z}{z_j} \right)^i \right] \varepsilon_t = \sum_{i=0}^{\infty} \varphi_i z^i \varepsilon_t = \sum_{i=0}^{\infty} \varphi_i \varepsilon_{t-i},$$

ahol φ_i konstansok (z^i -n kívüli tagok a szummákból)

Def: $A(z)$ $x_t = B(z)\varepsilon_t$ *ARMA* folyamat **invertálható**, ha $\exists \varepsilon_t = \sum_{j=0}^{\infty} \lambda_j x_{t-j}$

előállítása.

Tétel 2. Ha $B(z)$ polinom olyan, hogy $\forall$ gyöke kívül van az egységkörön ($\forall |z_i| > 1$, ha $B(z_i) = 0$), akkor **invertálható**, és:

$$\varepsilon_t = \frac{A(z)}{B(z)} x_t$$

Pl.

$$x_t - \frac{1}{4}x_{t-1} = 3\varepsilon_t + \varepsilon_{t-1} \quad (2, 1) \text{ ARMA folyamat.}$$

1. Feladat: Létezik-e az ARMA folyamatnak végtelen MA előállítása? Ha igen, adja meg!

$$x_t(1 - \frac{1}{4}z^2) = \varepsilon_t(3 + z)$$

$$A(z) = 1 - \frac{1}{4}z^2 = 0 \Rightarrow z_{1,2} = \pm 2$$

A Tétel1. szerint előáll $x_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j}$. Most:

$$x_t = \frac{3+z}{1-\frac{1}{4}z^2} \varepsilon_t$$

$$\frac{3+z}{\left(1-\frac{1}{2}z\right)\left(1+\frac{1}{2}z\right)} = \frac{A}{1-\frac{z}{2}} + \frac{B}{1+\frac{z}{2}} = \frac{A + \frac{1}{2}Az + B - \frac{1}{2}Bz}{1-\frac{1}{4}z^2}$$

$$\left. \begin{array}{l} 3 = A + B \\ 1 = \frac{1}{2}A - \frac{1}{2}B \end{array} \right\} \cdot 2 \Bigg\} \oplus$$

$$5 = 2A \Rightarrow A = 2,5$$

$$3 = 2,5 + b \Rightarrow B = 0,5$$

$$\begin{aligned} x_t &= \left[\frac{2,5}{1-0,5z} + \frac{0,5}{1+0,5z} \right] \varepsilon_t = \left[2,5 \sum_{j=0}^{\infty} 0,5^j z^j + 0,5 \sum_{j=0}^{\infty} (-0,5)^j z^j \right] \varepsilon_t = \\ &= \left[2,5 \sum_{j=0}^{\infty} 0,5^j \varepsilon_{t-j} + 0,5 \sum_{j=0}^{\infty} (-0,5)^j \varepsilon_{t-j} \right] = \\ &= \left[\sum_{j=0}^{\infty} 2,5 \cdot 0,5^j + 0,5 \cdot (-0,5)^j \right] \varepsilon_{t-j} \end{aligned}$$

2. Feladat: Invertálható-e az ARMA folyamat? Ha igen, adja meg az előállítását!

$$B(z) = 3 + z \Rightarrow z_{1,2} = \pm 3$$

A Tétel2. szerint invertálható. Tehát:

$$\varepsilon_t = \frac{1 - \frac{1}{4}z^2}{3 + z} x_t$$

Mivel a számlálóban is van z , és nehéz egyszerűsíteni, ezért polinom osztást végzek:

$$\begin{array}{r}
 -\frac{1}{4}z^2 + 1 : z + 3 = -\frac{1}{4}z + \frac{3}{4} - \frac{5}{4(z+3)} \\
 \ominus -\frac{1}{4}z^2 - \frac{3}{4}z \\
 \hline
 \frac{3}{4}z + 1 \\
 \ominus \frac{3}{4}z + \frac{9}{4} \\
 \hline
 -\frac{5}{4}
 \end{array}$$

A polinom osztás után:

$$\varepsilon_t = \left[-\frac{1}{4}z + \frac{3}{4} - \frac{5}{4(z+3)} \right] x_t = \frac{3}{4}x_t - \frac{1}{4}x_{t-1} - \frac{5}{4 \cdot 3} \underbrace{\frac{1}{1 + \frac{z}{3}}}_{\text{mérteni sor}} x_t$$

Amből a mértani sor miatt:

$$\varepsilon_t = \frac{3}{4}x_t - \frac{1}{4}x_{t-1} - \frac{5}{12} \sum_{j=0}^{\infty} \left(-\frac{1}{3} \right)^j z^j x_t, \quad \text{ahol } z^j x_t = x_{t-j}$$

Sztochasztikus folyamatok - ELŐREJELZÉS

Adottak $x_1, \dots, x_T$ minták. Ekkor x_{T+1} megközelítése $\hat{x}_{T+1}$.

Feltevés: x_t (p, q) ARMA, stabil, invertálható.

Def: Oyan $\hat{x}_t = \sum_{j=1}^{\infty} \delta_j x_{t-j}$ becslést keresünk, ahol x_t $(x_{t-1}, x_{t-2}, \dots)$ múltból állítjuk elő, melyre $E(x_t - \hat{x}_t)^2$ a legkisebb négyzetes közelítés.

$$x_t = \frac{B(z)}{A(z)} \varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad \psi_k\text{-k számolhatók.}$$

$$\hat{x}_t = \sum_{j=1}^{\infty} \gamma_j \varepsilon_{t-j} \quad (j = 1\text{-től megy, mert előre jelzést akarunk)}$$

Állítás: $x_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ ez a legjobb négyzetes közelítés. (azaz akkor a legjobb, ha $\gamma_j = \psi_j \quad \forall j = 1, \dots$)

Biz. Tfh. $y_t = \sum_{j=1}^{\infty} \gamma_j \varepsilon_{t-j}$ x_t -nek egy becslése/előrejelzése.

A becslés, akkor a legjobb, ha az $E((x_t - y_t)^2)$ a lehető legkisebb.

$$E((x_t - y_t)^2) = E((\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} - \sum_{j=1}^{\infty} \gamma_j \varepsilon_{t-j})^2) = E((\psi_0 \varepsilon_t + \sum_{j=0}^{\infty} (\psi_j - \gamma_j) \varepsilon_{t-j})^2) = \psi_0^2 \underbrace{E((\varepsilon_t)^2)}_{\sigma^2} + \sum_{j=0}^{\infty} (\psi_j - \gamma_j)^2 \underbrace{E((\varepsilon_{t-j})^2)}_{\sigma^2}$$

A köztes tag (2nem tudom mi) 0, mert ε_t fehér zaj $\Rightarrow E(\varepsilon_t) = 0 \quad \forall t \in \mathbb{Z}$

Akkor a legkisebb, ha $\psi_j = \gamma_j$

$$\left. \begin{aligned} \hat{x}_t &= \sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j} \text{ alakban a legjobb} \\ x_t &= \frac{B(z)}{A(z)} \varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \end{aligned} \right\} \hat{x}_t = x_t - \psi_0 \varepsilon_t = \frac{B(z)}{A(z)} \varepsilon_t - \psi_0 \varepsilon_t = \frac{B(z) - A(z)\psi_0}{A(z)} \varepsilon_t = \hat{x}_t$$

$$\hat{x}_t = \sum_{j=1}^{\infty} \psi_j z^j \varepsilon_t$$

Mivel $j = 1$ -től megy a $j = 0$ tag az 0.

Azaz a számláló első tagja: $\beta_0 - 1\psi_0 = 0$

Amiből adódik, hogy $\psi_0 = \beta_0$

Tehát:

$$\hat{x}_t = \frac{B(z) - A(z)\beta_0}{A(z)} \varepsilon_t$$

Mivel x_t -kel akarjuk előállítani, felhasználjuk, hogy $\varepsilon_t = \frac{A(z)}{B(z)} x_t$.

$$\hat{x}_t = \frac{B(z) - A(z)\beta_0}{A(z)} \frac{A(z)}{B(z)} x_t = \frac{B(z) - A(z)\beta_0}{B(z)} x_t$$

Pl.: Van egy x_t mérési mintánk $A(z)$, $B(z)$ ismert.

$$25x_t - x_{t-2} = 8\varepsilon_t - 2\varepsilon_{t-1} - \varepsilon_{t-2}$$

$$A(z) = 25z^2 - 1$$

$$B(z) = 8z^2 - 2z - 1$$

$$\beta_0 = 8$$

Statisztikai folyamat paramétereinek becslése x_t stacionárius folyamat
 $m = E(x_t) \quad \forall t \in \mathbb{Z}, \quad R_x(k) = \text{cov}(x_0, x_t) \quad \forall k \in \mathbb{Z}$

Például ε_t független azonos eloszlású, és $E(\varepsilon_0) = 0, \quad D^2(\varepsilon_t) < \infty$, konstans
 (tehát fehér zaj)

$$\varepsilon_t \quad t = 0, \dots, T$$

$$\bar{\varepsilon} = \frac{\varepsilon_0 + \varepsilon_1 + \dots + \varepsilon_T}{T+1}, \text{ ami a nagyyszámok törvénye miatt: } E\varepsilon_0$$

Ergodikusság: $x_t, t \in \mathbb{Z}$ gyengén stacionárius folyamatot **ergodikusnak** nevezzük, ha $\frac{x_1 + x_2 + \dots + x_t}{T}$ tart az $E(x_0)$ -hoz L^2 értelemben.
 $(\lim_{T \rightarrow \infty} (E(\frac{x_1 + \dots + x_T}{T}) - E(x_0))^2) = 0)$

Tétel: Ha $R_x(k)$ kovariancia függvény olyan, hogy az $\lim_{k \rightarrow \infty} R_x(k) = 0$, akkor x_t folyamat **ergodikus**.

$$x_t \text{ ARMA, tehát } x_t = \sum_{i=1}^n \phi_i \varepsilon_{t-i}, \text{ ekkor } R_x(k) = \sum_{j=-\infty}^{\infty} \phi_j \phi_{j+k} \sigma^2, \quad \sum \phi_i^2 < \infty$$

Mert:

$$R_x(k) \text{ számolásához felhasználjuk, hogy } x_t = \sum_{i=1}^n \phi_i \varepsilon_{t-i}$$

$$R_x(k) = \text{cov}(x_t, x_{t+k}) = \text{cov}\left(\sum_{j=-\infty}^{\infty} \phi_j \varepsilon_{t-j}, \sum_{i=-\infty}^{\infty} \phi_i \varepsilon_{t+k-i}\right) = \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \phi_j \phi_i \text{cov}(\varepsilon_{t-j}, \varepsilon_{t+k-i})$$

$$R_x(k) = \begin{cases} 0 & t-j \neq t+k-i \\ \phi_j \phi_i \sigma^2 & t-j = t+k-i \end{cases} \text{ akkor nem 0, ha } i = j+k$$

$$\text{Tehát valóban } R_x(k) = \sum_{j=-\infty}^{\infty} \phi_j \phi_{j+k} \sigma^2.$$

Tétel: Ha $\lim_{k \rightarrow \infty} R_x(k) = 0$, és x_t Gauss folyamat (azaz normál eloszlású folyamat), ekkor $\lim_{k \rightarrow \infty} \hat{R}_x^n(k) = R_x(k)$

Konklúzió $x_1, \dots, x_n$, ekkor $\underbrace{E(x_t) \approx \bar{x}_n \text{ és } r_x(k) \approx \hat{R}_x^n(k)}_{\text{a legjobb becslések}} \quad (n > 50, n > 4k$

esetén)

Wiener szűrő: t

émaköre

Spektrális sg. fv.: Legyen $x_t \quad t \in \mathbb{Z}$ gyengén stacionárius folyamat,
 $R_x(k) \quad k \in \mathbb{Z}$ kovariancia függvény. Ekkor: $\Phi_x(u) = \sum_{u \in \mathbb{Z}} R_x(k) e^{-iuk}$ végtelen

sort, ha konvergens, $u \in [-\pi, \pi]$ a folyamat **spektrális sűrűség függvényének** nevezzük

Tulajdonságai:

$$1. \quad \Phi_x(u) = \Phi_x(-u)$$

$$\text{Bizonyítás: } \Phi_x(u) = \sum_{k \in \mathbb{Z}} R_x(k) e^{-iuk} \xrightarrow[\text{(páros)}]{R_x(k) = R_x(-k)} \sum_{k \in \mathbb{Z}} R_x(-k) e^{-iuk} \xrightarrow[-k=j]{\text{új jelölés:}} \sum_{j \in \mathbb{Z}} R_x(j) e^{iuj} = \sum_{j \in \mathbb{Z}} R_x(j) e^{-i(-u)j} = \Phi_x(-u)$$

$$2. \quad \Phi_x(u) \geq 0$$

Bizonyítás: $R_x(k) \sim \sigma^2 \geq 0 \quad e^{-ix} \geq 0$ (e^{-ix} -ből cos-os alak lesz)

3. **Tétel:** Ha $\Phi_x(u)$ létezik, akkor páros, nem negatív **inverze:**
 $R_x(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_x(t) e^{itk} dt$

Markov lánc: $x_1, x_2, \dots, x_t$ valószínűségi változók sorozata **Markov láncot** alkot, ha t kezdőpillanatot követő x_{t+1} valószínűségi változó értéke csak x_t értékétől függ.

$$P(x_{t+1} = i_{t+1} | x_t = i_t, \underbrace{x_{t-1} = i_{t-1}, \dots, x_0 = i_0}_{\text{ezek nem befolyásolják}}) = P(x_{t+1} = i_{t+1} | x_t = i_t)$$

egylépéses átmenet: Rögzített $n \in \mathbb{N}$ mellett $p_{i,j}(n)$ $i, j \in I$ valószínűségeket a Markov lánc **n . lépéséhez** tartozó **egylépéses átmenet valószínűségének** nevezzük.

$\underline{\Pi}(n)$ az átmenet valószínűségekhez tartozó **átmenetivalószínűségi mátrix**.

$$\underline{\Pi}(n) = [p_{i,j}(n)]_{i,j \in I}$$

eloszlás: Egy $\underline{\alpha}$ sorvektorról azt mondjuk, hogy **eloszlás**, ha $\underline{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_k]$ $\alpha_i \geq 0 \quad \forall i = 1 \dots k$, $\sum_{i=1}^k \alpha_i = 1$

Az $\underline{A}$ mátrix **sztochasztikus mátrix**, ha minden sorvektora eloszlás.

Markov lánc kezdeti eloszlása megadható $\underline{q} = [q_i]$ sorvektorával,

ahol $\forall i \quad q_i = P(x_0 = i)$. Ebből következik, hogy $\sum_{i=1}^k q_i = 1$, mivel eloszlás. (k a realizációk száma)

$\underline{\Pi}$ a folyamat dinamikáját adja meg. Tehát azt adja meg, hogy milyen valószínűséggel jut egy lépés alatt i -ből j -be.

Markov lác homogenitása: Az x_t , $t \in \mathbb{N}$ Markov lánc **homogén**, ha $P(x_t = i_t | x_{t-1} = i_{t-1})$ független t -től.

Megjegyzés: Innentől kezdve csak homogén Markov láncokkal fogunk foglalkozni.

Tétel: $p_{i,j} = P(x_1 = j | x_0 = i) \geq 0 \quad \forall i, j$ -re és $\sum_{j=0}^n p_{i,j} = 1$

Tétel: Legyen x_t valószínűségi változók sorozatának eloszlása $\underline{q}(x_0)$. Ekkor $\underline{q}(x_1)$ a következőképpen számolható:

$$\underline{q}(x_1) = \underline{q}(x_0) \underline{\Pi} \quad \text{ahol } \underline{q}(x_1), \underline{q}(x_0) \in \mathbb{R}^n, \quad \underline{\Pi} \in \mathbb{R}^N \times \mathbb{R}^N$$

n lépéses átmenet: Az **n lépéses átmenetvalószínűség** annak a valószínűsége, hogy i -ből kiindulva n lépés alatt eljutok k -ba.

$$p_{i,k}^{(n)} = P(x_n = k | x_0 = i)$$

Tétel: Chapman-Kolmogorov egyenlet:

$$\left[p_{i,j}^{(n)} \right] = \underline{\Pi}^n, \quad n \geq 2$$

irreducibilis: Az x_t $t \in \mathbb{N}$ Markov láncot **irreducibilis**nek nevezzük, ha $\forall i, j = 1, 2, \dots, N$ esetén $\exists n = n(i, j)$, hogy $p_{ij}^{(n)} > 0$.

Bármelyik állapotból bármelyik állapot elérhető $k \in \mathbb{Z}$ lépés alatt, pozitív valószínűséggel.

zárttság: A Markov lánc állapotainak S halmaza **zárt**, ha az S halmazon kívül egyetlen állapotot sem érhető el S állapotaiból.

elnyelő: Az i . állapot **elnyelő** állapot, ha $p_{ii} = 1$. Minden elnyelő állapot zárt halmaz.

tranziens: Az i . állapot **tranziens**, ha $\exists$ olyan j állapot, amely elérhető i -ből, de j -ből i nem érhető el. $p_{ij}^{(n)} > 0 \quad \nexists m : p_{ji}^{(m)} > 0$
Ha nem tranziens, akkor **visszatérő**.

kommunikál: Azt mondjuk, hogy i kommunikál j -vel, ha $\exists n, m$ $p_{ij}^{(n)} > 0$ és $p_{ji}^{(m)} > 0$

periodikus: Az i . állapot **periodikus** $k > 1$ periodussal, ha k az a legkisebb szám, amire igaz, hogy az i -ből i -be visszatérő lánc hossza k , illetve k egészszáma többszöröse.

$$\underline{\Pi}^{(k)} = \begin{cases} p_{ij} = 0 & \text{ha } i \neq j \\ p_{ij} = 1 & \text{ha } i = j \end{cases}$$

A nem periodikus visszatérő állapotot **aperiodikus**nak nevezzük.

$\{n > 0, p_{ij}^{(n)}\}$ halmaz legnagyobb közös osztója: $d(i) > 1$, akkor i periodikus. Ha $d(i) = 1$, akkor aperiodikus.

aperiodikus: A Markov lánc **aperiodikus**, ha $\forall i = 1, \dots, N$ állapota aperiodikus.

ergodik: A Markov láncot **ergodik**nek nevezzük, ha aperiodikus, és $\forall p_{ij} > 0 \quad \forall i, j = 1, \dots, N$.

Megjegyzés: (azaz ergodik, ha aperiodikus és irreducibilis)

Tétel: Legyen a Markov lánc ergodik, $(N \times N)$ -s $\underline{\Pi}$ mátrix-szal adott.

$$\text{Ekkor } \exists \underline{x} = (x_1, \dots, x_N), \text{ hogy } \lim_{n \rightarrow \infty} \underline{\Pi}^n = \begin{pmatrix} x_1 & x_2 & \dots & x_N \\ x_1 & x_2 & \dots & x_N \\ \vdots & & & \vdots \\ x_1 & x_2 & \dots & x_N \end{pmatrix}$$

x_j valószínűség független a kiindulástól.

Megjegyzés: ("elfelejti a múltját")

stacionárius állapot: A fenti $\underline{x} = (x_1, \dots, x_N)$ a Markov lánc **stacionárius állapota**: $\underline{x} = \underline{x}\underline{\Pi}$ másképp jelölve: $\underline{q}_{\infty} = \underline{q}_{\infty}\underline{\Pi}$ amit átalakítva: $(\underline{\Pi} - \underline{E})^T \underline{q}_{\infty}^T = \underline{0}$, ahol $q_{\infty_1} + q_{\infty_2} + \dots + q_{\infty_N} = 1$.

Tétel: Ha a Markov lánc irreducibilis és aperiodikus, akkor stabil. Azaz $\underline{q}(x_n) \rightarrow \underline{q}_{\infty}$, $\underline{q}_0$ -tól függetlenül

$\Rightarrow \exists$ stacionárius állapota.
Tömegkiszolgálási rendszerek

Érkeznek igények, amikből mindig egyet kiszolgálunk, majd kiszolgálás után

távozik.

Fontos tulajdonságai:

- várható sorhossz
- átlagos várakozási idő
- optimális alkalmazott (kiszolgáló egységek) száma

Diszkrét idejű rendszert tekintünk.

$$\left. \begin{array}{l} Y_n := (n-1, n] \text{ között érkező igények száma} \\ V_n := (n-1, n] \text{ között kiszolgált igények száma} \\ X_n := n. \text{ időpontban várakozó igények száma} \end{array} \right\} n = 1, 2, \dots \text{ (ez lehet perc, nap, év, stb)}$$

A sorhossz érdekel minket:

$$X_{n+1} = (X_n - V_{n+1})^+ + Y_{n+1}$$

$$\text{Ahol } (X_n - V_{n+1})^+ = \max(X_n - V_{n+1}, 0) = \begin{cases} X_n - V_{n+1} & \text{ha } X_n - V_{n+1} > 0 \\ 0 & \text{különben} \end{cases}$$

Feltevések

- $Y_n \quad n = 1, 2, \dots$ független azonos eloszlású valószínűségi változó
- $V_n \quad n = 1, 2, \dots$ független azonos eloszlású valószínűségi változó
- $\{Y_n\}$ és $\{V_n\}$ sorozatok is függetlenek egymástól

Állítás: Legyen x_0 független (Y_n, V_n) -től. Ha a fenti feltevések igazak, akkor x_t homogén Markov lánc.

Tétel: Tekintsük a fenti tömeg kiszolgálási modelt:

Tfh. Π főátlója alatt és felett csupa > 0 elem van, valamint $E(Y_i) < E(V_i)$
 $\forall i = 1, 2, \dots$. Ekkor $x_t \quad t \in \mathbb{N}$ stabil Markov lánc. ($\exists$ stacionáris állapota)

Az $x_t \quad t \in \mathbb{N}$ egy $\chi \infty$ állapotterben értelmezett.

Az irreducibilis és az aperiodikus tulajdonságok hasonlóak. $\Pi = [p_{ij}]_{i,j \in \mathbb{N}}$, ahol $p_{ij} = P(x_1 = j | x_0 = i)$

Tétel: Ha a főátló alatt és felett minden elem pozitív ($p_{i,i-1} > 0; p_{i,i+1} > 0$) és a főátlóban legalább egy $p_{ii} \neq 0$

Megjegyzés: Ha $x_t \infty$ állapotterben értelmezett, akkor az irreducibilitásból és aperiodikusságból nem következik a satabilitás.

eloszlási operátor: Eloszlási operátornak nevezzük $q(x_t) = (p(x_t = 0), p(x_t = 1), \dots, p(x_t = k), \dots) \in \mathbb{R}^N$.

stabilitás: Az $x_t \quad t \in \mathbb{N}$ Markov lánc **stabil**, ha $q_\infty \in \mathbb{R}^n$ eloszlási operátorra igaz, hogy $\forall q_\infty^i \in [0, 1] \quad i \in \mathbb{N}$ -re $\sum_{i=0}^{\infty} q_\infty^i = 1$ és $\forall x_0$ -ra $\lim_{t \rightarrow \infty} q^i(x_t) = q_\infty^i$.

Folytonos idejű Markov láncok

Az állapotter diszkrét $x_t \in S \quad S = \{1, 2, \dots\} \quad t \geq 0 \quad t \in \mathbb{R}^+ \cup \{0\}$

folytonos idejű ML: $x_t \quad t \in \mathbb{R}^+ \cup \{0\}$ **folytonos idejű** Markov láncnak nevezzük, ha $\forall n \geq 1$ -re $0 \leq t_0 \leq t_1 \leq \dots \leq t_n$ -re az $x_0, x_0, x_1, \dots, x_n \in S$ teljesül.

$P(x(t_n) = x_n | x(t_{n-1}) = x_{n-1}, x(t_{n-2}) = x_{n-2}, \dots, x(t_0) = x_0) = P(x(t_n) = x_n | x(t_{n-1}) = x_{n-1})$ ha ezek a feltételes valószínűségek léteznek.

Például: 2 egymás utáni történés λ paraméterű exponenciális eloszlású.

x_t jelenti $[0, t]$ hány történés volt.

x_t Folytonos idejű Markov lánc eloszlása (λt) paraméterű Poisson eloszlás.

$$p_{ij}(t) = e^{-(\lambda t)} \frac{(\lambda t)^{j-i}}{(j-i)!}$$

Chapman-Kolmogorov: $p_{ij}(s+t) = \sum_{k \in S} P(x(s+t) = j, x(t) = k | x_0 = i) = \sum_{k \in S} P(x(s+t) = j | x(t) = k) P(x(t) = k | x_0 = i) = \sum_{k \in S} p_{ik}(t) p_{kj}(s) = \sum_{k \in S} p_{ik}(s) p_{kj}(t)$

$$\underline{\Pi}(s+t) = \underline{\Pi}(s) \underline{\Pi}(t) = \underline{\Pi}(t) \underline{\Pi}(s)$$

Tfh. $p_{ij}(t) = \delta_{ij} = \begin{cases} 0 & \text{ha } i \neq j \\ 1 & \text{ha } i = j \end{cases}$ kicsi t -re $\begin{cases} p_{ii} = 1 \\ p_{ij} = 0 \end{cases}$

Tétel: Ha a Markov láncra igaz a fenti feltétel, $p_{ij}(t)$ függvények differenciálhatóak $\forall t > 0$ -ra.

Jelölés: $q_{ij} = p'_{ij}(0)^+ = \lim_{t \rightarrow 0^+} \frac{p_{ij}(t) - \delta_{ij}}{t}$

ráta mátrix: $Q = [q_{ij}] = \Pi'[0^+]$ mátrix a folyamat **ráta mátrix**nak nevezzük. $\sum_{j \in \mathbb{N}} q_{ij} = 0$ (tehát van negatív eleme is)

Tétel: Ha a Markov láncra teljesül, hogy $\sum_{j \in \mathbb{N}} q_{ij} = 0 \quad \sum_i p'_i j = \sum_{k \in S} q_{ik} p_{kj}(t)$

$$\underline{\Pi}'(t) = \underline{Q} \underline{\Pi}(t)$$

Tétel: Folytonos idejű, véges állapotterű, irreducibilis Markov lánc, akkor **stabil**, ha $\underline{x}\underline{Q} = \underline{0} \rightarrow \underline{Q}^t \underline{x}^t = \underline{0}^t$. (Ahol $\underline{x}$ a határeloszlás, $\underline{Q}$ a ráta mártix.)

felújítási folyamat: Egy pontfolyamatot **felújítási folyamatnak** nevezük, ha a pontok közötti távolság azonos eloszlású, tetszőleges valószínűségi változók. Paraméterei: A/B/m/K/M

A - szomszédos igények beérkezése között eltelt idő eloszlásának kódja

B - kiszolgálási idő eloszlásának kódja

m - a rendszerben lévő kiszolgáló egységek száma

K - a rendszer befogadóképessége

M - igényforrások száma

A,B kódja lehet:

- M - Markov tulajdonságú, exponenciális eloszlású
- D - Determinisztikus
- G - tetszőleges eloszlású

K,M alapértelmezettként ∞

Például M/M/1, ahol az első M (az A kódja) utal a (λt) paraméterű Poisson eloszlásra, a második M (a B kódja) a (μ) paraméterű Exponenciális eloszlásra, az 1 a kiszolgáló egységek száma, valamint K,M végtelen.

Megjegyzés: Gyakorlatról(05.16.):

Érkezés: λ intenzitású Poisson folytonos eloszlású

Kiszolgálás: μ paraméterű Exponenciális eloszlású

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ \mu & -(\lambda + \mu) & \lambda & 0 & \\ 0 & \mu & -(\lambda + \mu) & \lambda & \\ 0 & 0 & \mu & -(\lambda + \mu) & \\ \vdots & & & & \ddots \end{bmatrix}$$

Ha $\exists$ stacionárius eloszlása:

$$\mathbf{q} = (q_0, q_1, \dots, q_n, \dots)$$

$$\text{Ahol } q_i = q_0 * \left(\frac{\lambda}{\mu}\right)^i \quad i \in S \quad i > 0$$

$$\text{Kihasználtság: } \frac{\lambda}{\mu} = \rho$$

$$p_i = p_0 \left(\frac{\lambda}{\mu}\right)^i \quad p_0 = 1 - \frac{\lambda}{\mu} = \frac{\mu - \lambda}{\mu}$$

Rendszerben tartózkodó igények száma:

$$N(t) = \sum_{i=0}^{\infty} i p_i = \sum_{i=1}^{\infty} i (1 - \rho) \rho^i = \sum_{i=1}^{\infty} i (1 - \rho) \rho^{i-1} \rho = (1 - \rho) \rho \frac{1}{(1 - \rho)^2} = \frac{\rho}{1 - \rho}$$

Sorhossz:

$$Q = \frac{\rho^2}{1 - \rho} = N(t) - \rho$$

Várakozási idő:

$$T = \frac{1}{\mu - \lambda}$$

Ezek a képletek $M/M/1$ esetre vannak. Tipikusan ilyen a születés/halálozás.

Ha bármilyen hibát találnak, kérem írják meg! vardai@itk.ppke.hu