A KÜLFÖLDI SZAKIRODALOMBÓL

SZTOCHASZTIKUS FOLYAMATOK ÉS STATISZTIKAI KÖVETKEZTETÉSEK (II).*

Írta: ULF GRENANDER

5. fejezet

A BECSLÉS PROBLÉMÁJA

5. 1. Torzítatlan becslések. Tegyük most fel, hogy az $x(t)$ folyamatot megadó valószínűségi mérték az általunk ismert P_α osztályhoz tartozik, ahol α egy valós paraméter, amelynek értékei kitöltik az (a, b) véges intervallumot. Hogy elkerüljük a szinguláris esetet, feltesszük még, hogy semmilyen α_1, α_2 értékpár esetén sem létezik olyan S halmaz, amelyre $P_{\alpha_1}=0$ és $P_{\alpha_2}\neq 0$. A folyamat egy észlelt realizációja alapján el akarjuk dönteni, hogy a hipotézisek melyike igaz (mi az α igazi értéke), vagyis olyan $t(\omega)$ függvényt akarunk szerkeszteni, amely becslést ad α-ra. Tetszőleges S halmazra érvényes, hogy

$$P_\alpha(S) = \int\limits_S f(\omega, \alpha)\, dP_0(\omega),$$

ahol P_0 az α egy rögzített α_0 értékéhez tartozó mérték. Tekintsük most az $f(\omega, \alpha)$ likelihood függvényt olyan sztochasztikus folyamatnak, amely az α paramétertől függ (ez esetben tehát α-t tekintjük időparaméternek) és amelyet a P_0 valószínűségi mérték ad meg. Ennek a folyamatnak az átlaga nyilvánvalóan

$$\mathbf{E}_0 f(\omega, \alpha) = \int\limits_\Omega f(\omega, \alpha)\, dP_0(\omega) = P_\alpha(\Omega) = 1.$$

Induljunk ki abból a természetes feltevésből, hogy az $f(\omega, \alpha)$ folyamat középben folytonos, és véges a szórása. Az $f(\omega, s)=f(s)$ jelöléssel a folyamat kovariancia függvénye

$$\varrho(s, t) = \mathbf{E}_0[f(s)-1][f(t)-1].$$

Képezzük az ehhez a maghoz tartozó integrálegyenletet a szokásos módon, és jelöljük sajátértékeit λ_ν-vel, sajátfüggvényeit $\varphi_\nu(\alpha)$-val. Ekkor minden $\alpha \in A$ esetén

$$f(\omega, \alpha) = \underset{n\to\infty}{\mathrm{l.i.m.}} \sum_1^n \varphi_\nu \frac{\psi_\nu(\omega)}{\sqrt{\lambda_\nu}} + 1,$$

* Stochastic processes and statistical inference, *Arkiv för Matematik*, **1** (1950), 195–277. A fordítás első része, amely az eredeti dolgozat 1–4. fejezetét, a tartalomjegyzéket és a teljes irodalomjegyzéket tartalmazza, az *MTA III. Oszt. Közl.*, **15** (1965), 51–87. oldalán jelent meg.

ahol $\{\psi_\nu(\omega)\}$ egy $L_2(\Omega)$-ban ortonormált rendszer és az átlagban való konvergencia a P_0 mértékre nézve értendő.

Vizsgáljuk most minimális szórású torzítatlan becslések egzisztenciáját, azaz olyan $t(\omega)$ függvényekét, amelyek kielégítik az

$$\begin{cases} \mathbf{E}_\alpha t(\omega) = \alpha \\ \mathbf{E}_\alpha [t(\omega)]^2 < \infty \end{cases}$$

feltételeket minden $\alpha \in A$ esetén. Ha a $\{\psi_\nu\}$ rendszer nem teljes $L_2(\Omega)$-ban P_0-ra nézve, akkor hozzácsatoljuk az ortogonális komplementumát, a

$$\{\psi'_\mu\} = L_2(\Omega) \ominus \{\psi_\nu\}$$

ortonormált rendszert. A fenti feltételeket kielégítő tetszőleges becslést P_0-ra nézve átlagban konvergens sorba lehet fejteni a $\{\psi_\nu\}$ és $\{\psi'_\mu\}$ rendszerek szerint:

$$t(\omega) = \sum_1^\infty t_\nu \psi_\nu(\omega) + \sum_1^\infty t'_\mu \psi'_\mu(\omega).$$

Ebből következik, hogy

$$\alpha = \mathbf{E}_\alpha t(\omega) = \int_\Omega t(\omega) f(\omega, \alpha) dP_0(\omega) = \sum_1^\infty t_\nu \frac{\varphi_\nu(\alpha)}{\sqrt{\lambda_\nu}} + \mathbf{E}_0 t(\omega).$$

A jobb oldali sor egyenletesen konvergens, mert

$$\left| \sum_n^\infty t_\nu \frac{\varphi_\nu(\alpha)}{\sqrt{\lambda_\nu}} \right|^2 \leqq \sum_n^\infty t_\nu^2 \sum_n^\infty \frac{\varphi_\nu(\alpha)^2}{\lambda_\nu} \leqq \sum_n^\infty t_\nu^2 \sum_1^\infty \frac{\varphi_\nu(\alpha)^2}{\lambda_\nu}$$

és mert

$$\sum_1^\infty \frac{\varphi_\nu(\alpha)^2}{\lambda_\nu} = \varrho(\alpha, \alpha)$$

folytonos függvénye α-nak. Így tehát a sort megszorozhatjuk $\varphi_\nu(\alpha)$-val és tagonként integrálhatjuk, aminek eredménye:

$$\gamma_\nu \int_a^b (\alpha - \alpha_0) \varphi_\nu(\alpha) d\alpha = \frac{t_\nu}{\sqrt{\lambda_\nu}}.$$

Ebből látható, hogy a véges szórású torzítatlan becslés létezésének szükséges feltétele $\sum_1^\infty \lambda_\nu \gamma_\nu^2 < \infty$.

Egyszerűség kedvéért tegyük fel, hogy $\{\varphi_\nu(\alpha)\}$ teljes $L_2(A)$-ban. Ekkor a $\sum_1^\infty \lambda_\nu \gamma_\nu^2 < \infty$ feltétel egyszersmind elégséges feltétele is a véges szórású torzítatlan becslés létezésének.

Tekintsük ugyanis a

$$t(\omega) = \sum_1^\infty \sqrt{\lambda_\nu} \gamma_\nu \psi_\nu(\omega) + \sum_1^\infty t'_\mu \psi'_\mu(\omega)$$

kifejezést, ahol a t'_μ valós számok kielégítik a $\sum_1^\infty t'^2_\mu < \infty$ feltételt, de egyébként tetszőlegesek lehetnek. Ebben az esetben a sor nyilvánvalóan P_0-ra nézve átlagban konvergens, és

$$\mathbf{E}_\alpha t(\omega) = \sum_1^\infty t_\nu \varphi_\nu(\alpha) + \mathbf{E}_0 t(\omega).$$

Ugyanazzal a módszerrel, mint fentebb, ki lehet mutatni, hogy $\sum_1^\infty t_\nu \varphi_\nu(\alpha) = \alpha - \alpha_0$ minden $\alpha \in A$ esetén. Így tehát $t(\omega) + \alpha_0 - \mathbf{E}_0 t(\omega)$ torzítatlan becslése α-nak. Az eredményeket összefoglalva:

A véges szórású torzítatlan becslés létezésének szükséges feltétele

$$\sum_1^\infty \lambda_\nu \left\{ \int_a^b (\alpha - \alpha_0) \varphi_\nu(\alpha)\, d\alpha \right\}^2 < \infty.$$

Ha a $\{\varphi_\nu\}$ rendszer teljes, akkor ebből a feltételből következik a torzítatlan becslések olyan seregének a létezése, amelynek dimenziója megegyezik az $L_2(\Omega) \ominus \{\psi_\nu\}$ tér dimenziójával.

Ha létezik egynél több torzítatlan becslés, és ha van alapja olyan feltevésnek, hogy α igazi értéke közel esik α_0-hoz, akkor természetesnek tűnik annak a becslésnek a kiválasztása, amelynek $\mathbf{D}_0(t)$ szórása a legkisebb. Egyes esetekben az is előfordulhat, hogy a torzítatlan becslések között olyat találunk, amelynek a szórása minden α-ra minimális.

A becslés-szerkesztés fenti módszerének vannak elméleti előnyei (többek között az, hogy a módszer még kiterjeszthető), de kevéssé alkalmas gyakorlati használatra, ezért a következő szakaszokban a becslésnek néhány gyakorlatilag előnyösebb módját fogjuk ismertetni. Itt csupán arra szorítkozunk, hogy egy egyszerű példa kapcsán bemutassuk a becslés megszerkesztésének egy olyan módszerét, amely néhány speciális esetben célra vezet. Tekintsük a 4. 10. szakaszban vizsgált Poisson-folyamatot, amelynek β sűrűsége állandó a $(0, t)$ intervallumban. A β sűrűségnek olyan torzítatlan β^* becslését keressük, amelynek a szórása minimális. Ha $\beta^* = \beta^*(n, t_1, t_2, \ldots, t_n)$ akkor

$$\beta = \mathbf{E}_\beta \beta^* = \sum_0^\infty P\{n = \nu\}\, \mathbf{E}_\beta[\beta^* | n = \nu] = \sum_0^\infty \frac{e^{-\beta t}(\beta t)^\nu}{\nu!} \mathbf{E}_\beta[\beta^* | n = \nu],$$

tehát β-nak bizonyos intervallumból vett értékeire az alábbi egyenletnek kell fennállania

$$\beta e^{\beta t} = \sum_0^\infty \frac{(\beta t)^\nu}{\nu!} \mathbf{E}_\beta[\beta^* | n = v].$$

Ismert n esetén a $t_1, t_2, \ldots, t_n$ mennyiségek feltételes sűrűségfüggvénye

$$\frac{e^{-\beta t} \beta^n}{\frac{e^{-\beta t}(\beta t)^n}{n!}} = \frac{n!}{t^n} \qquad (0 < t_1 < t_2 < \ldots < t_n < t),$$

tehát $\mathbf{E}_\beta[\beta^*|n=\nu]$ értéke független β-tól. Mivel $\beta\cdot e^{\beta t}$ a β változónak egész függvénye, a Taylor-együtthatókat egyenlővé tehetjük, és ilyen módon azt kapjuk, hogy

$$\mathbf{E}_\beta[\beta^*|n=\nu]=\frac{\nu}{t}.$$

Azonban

$$\mathbf{E}_\beta(\beta^*-\beta)^2=\sum_0^\infty P_n\mathbf{E}_\beta[(\beta^*-\beta)^2|n]=\sum_0^\infty P_n\left\{\mathbf{E}_\beta\left[\left(\beta^*-\frac{n}{t}\right)^2\middle|\,n\right]+\right.$$

$$\left.+2\mathbf{E}_\beta\left[\left(\beta^*-\frac{n}{t}\right)\left(\frac{n}{t}-\beta\right)\middle|\,n\right]+\mathbf{E}_\beta\left[\left(\frac{n}{t}-\beta\right)^2\middle|\,n\right]\right\}.$$

Itt a jobb oldal második tagja eltűnik, mert

$$\mathbf{E}_\beta\left[\left(\beta^*-\frac{n}{t}\right)^2\middle|\,n\right]=0.$$

Így tehát nyilvánvalóan megkapjuk az egyetlen minimális szórású torzítatlan becslést, ha β^*-ot úgy választjuk meg, hogy az első tag (amelynek értéke egyébként >0 volna) nullával legyen egyenlő:

$$\beta^*=\frac{n}{t}.$$

5. 2. A lineáris becslések egy osztálya. Az előző szakaszban azzal az esettel foglalkoztunk, amikor az α paraméter értéke teljesen meghatározza a P_α valószínűségi eloszlásokat. Gyakran előfordul, hogy nem tudjuk, vagy nem akarjuk megadni a valószínűségi eloszlásokat. Megtörténhet, hogy ennek ellenére mégis lehet találni minimális szórású torzítatlan becsléseket, de ez esetben a becsléseknek csupán valamilyen leszűkített osztályában. Tekintsük pl. a következő feladatot: meg akarjuk becsülni az $x(t)$ sztochasztikus folyamat m átlagát, ha a folyamatról tudjuk, hogy középben folytonos és létezik adott $r(s,t)$ korrelációs függvénye. Tekintsük a lineáris becslések alábbi osztályát

$$m^*=\int_a^b f(t)x(t)\,dt,$$

ahol $f(t)$ négyzetesen integrálható (a,b)-n. Az integrálást akár az 1. 3. szakasz értelmében foghatjuk fel, akár pedig Doob szerint (ez utóbbi esetben olyan mintateret kell választani, amelyben az $x^2(t)$ folyamat D-integrálható). Hogy az ebből az osztályból kapott becslés torzítatlan legyen, meg kell követelni az

$$\int_a^b f(t)\,dt=1$$

feltétel teljesülését. Ebben az esetben a szórás

$$\mathbf{E}(m^*-m)^2=\int_a^b\int_a^b r(s,t)f(s)f(t)\,ds\,dt.$$

Bevezetve a folyamat $r(s, t)$-magú integrálegyenletének λ_ν sajátértékeit és φ_ν sajátfüggvényeit (feltesszük, hogy a $\{\varphi_\nu\}$ függvények teljes rendszert alkotnak $L_2(a, b)$-ben), és felhasználva a korrelációs függvény bilineáris előállítását, következik

$$\mathbf{E}(m^* - m)^2 = \sum_1^\infty \frac{c_\nu^2}{\lambda_\nu},$$

ahol

$$c_\nu = \int_a^b f(t)\varphi_\nu(t)\,dt, \qquad (\nu = 1, 2, \ldots).$$

Keressük a $\frac{c_\nu^2}{\lambda_\nu}$ kifejezés minimumát a $\sum_1^\infty c_\nu a_\nu = 1$ feltétellel, ahol

$$a_\nu = \int_a^b \varphi_\nu(t)\,dt.$$

A SCHWARZ-féle egyenlőtlenség felhasználásával könnyen ki lehet mutatni, hogy ha $\sum_1^\infty a_\nu^2 \lambda_\nu < \infty$, akkor

$$1 = \left[\sum_1^\infty c_\nu a_\nu\right]^2 \leqq \left[\sum_1^\infty |c_\nu a_\nu|\right]^2 \leqq \sum_1^\infty \frac{c_\nu^2}{\lambda_\nu} \sum_1^\infty a_\nu^2 \lambda_\nu,$$

vagyis

$$\mathbf{E}(m^* - m)^2 \geqq \frac{1}{\sum_1^\infty a_\nu^2 \lambda_\nu},$$

ahol egyenlőség csak abban az esetben következik be, ha

$$c_\nu = \frac{a_\nu \lambda_\nu}{\sum_1^\infty a_\nu^2 \lambda_\nu}.$$

Legyen

$$f_N(t) = \frac{\sum_1^N a_\nu \lambda_\nu \varphi_\nu(t)}{\sum_1^N a_\nu^2 \lambda_\nu}$$

és

$$m_N^* = \int_a^b x(t) f_N(t)\,dt = m + \frac{\sum_1^N a_\nu \sqrt{\lambda_\nu}\, x_\nu}{\sum_1^N a_\nu^2 \lambda_\nu},$$

ahol $\{x_\nu\}$ nulla várható értékű és 1 szórású korrelálatlan valószínűségi változók sorozata. Ha N a végtelenhez tart, akkor *a becsléseknek ez a sorozata nyilvánvalóan átlagban konvergál egy olyan m^* becsléshez, amely torzítatlan és amelynek* $\frac{1}{\sum_1^\infty a_\nu^2 \lambda_\nu}$

szórása a lineáris becslések tekintett osztályában minimális (lásd GRENANDER [1]). Ha úgy kívánjuk, ebből a sorozatból ki lehet választani egy majdnem biztosan konvergens részsorozatot is. Meg kell jegyezni, hogy a sorozat határértéke nem mindig tartozik a becslések itt tekintett osztályába. A becslés-osztálynak erre a hátrányos tulajdonságára később még visszatérünk.

Tegyük fel viszont, hogy $\sum_{1}^{\infty} a_\nu^2 \lambda_\nu = \infty$. Könnyen belátható, hogy ebben az esetben létezik olyan $m_{N\nu}^*$ részsorozat, amely majdnem biztosan az igazi m-értékhez tart, ha $\nu \to \infty$. Így tehát — annak ellenére, hogy a folyamatnak csupán a korrelációs függvénye ismeretes — *megállapítható, hogy ha* $\sum_{1}^{\infty} a_\nu^2 \lambda_\nu = \infty$, *akkor a szinguláris esettel van dolgunk.* A 4. 4. szakaszban láttuk, hogy a normális folyamatoknál ez szükséges feltétele is a szinguláris esetnek (ha csak pozitív definit korrelációs függvényekre szorítkozunk).

Amint ezt a 4. 6. szakaszban láttuk, az $f_n(t)$ függvénysorozat egy $L_2(T)$-beli függvényhez való konvergálása maga után vonja az

$$\int_a^b r(s,t) f(t)\,dt = 1$$

integrálegyenlet egy négyzetesen integrálható megoldásának a létezését. Mivel ez ritkán fordul elő, kézenfekvő a lineáris becslések

$$m^* = \int_a^b x(t)\,dF(t)$$

tágabb osztályát tekinteni, ahol $F(t)$ korlátos variációjú függvény, és az integrált valamilyen megfelelő módon (pl. KARHUNEN szerint) értelmezzük.

Az előzőkhöz hasonlóan megköveteljük az

$$\begin{cases} \mathbf{E}m^* = m \int_a^b dF(t) = m \\ \mathbf{E}(m^* - m)^2 = \int_a^b \int_a^b r(s,t)\,dF(s)\,dF(t) = \min. \end{cases}$$

feltételek teljesülését. Tegyük fel, hogy $F(t)$ kielégíti ezeket a feltételeket és legyen α és β két pont az (a, b) intervallumon. Ha

$$G = \varepsilon(t-\alpha) - \varepsilon(t-\beta)$$

és δ tetszőleges valós szám, akkor az $F(t) + \delta G(t)$ súlyfüggvény szintén torzítatlan becslést ad, mivel

$$\int_a^b d[F(t) + \delta G(t)] = 1.$$

Továbbá az

$$\int_a^b r(s,t)\,dF(t) = R(s)$$

jelöléssel

$$\int_a^b\int_a^b r(s,t)d[F(s)+\delta G(s)]d[F(t)+\delta G(t)]=\int_a^b\int_a^b r(s,t)dF(s)dF(t)+2\delta\int_a^b R(t)dG(t)+$$

$$+\delta^2\int_a^b\int_a^b r(s,t)dG(s)dG(t).$$

Mivel az utóbbi kifejezésnek δ minden értékére nagyobbnak kell lennie az

$$\int_a^b\int_a^b r(s,t)dF(s)dF(t)$$

integrálnál, szükségszerűen teljesülnie kell az

$$\int_a^b R(t)dG(t)=R(\alpha)-R(\beta)$$

feltételnek, tehát

$$R(s)=\int_a^b r(s,t)dF(t)\equiv c$$

minden $s\in T$-re. Nyilvánvaló, hogy az utóbbi egyenlet jobb oldalán levő állandó éppen becslésünk szórásával, vagyis az m^* lineáris becslés minimális szórásával egyenlő[1].

Tegyük fel másrészt, hogy $F(t)$ kielégíti a fenti integrálegyenletet és az $\int_a^b dF(t)=1$ feltételt. Legyen $H(t)$ egy másik korlátos variációjú függvény (a,b)-n, amelyre $\int_a^b dH(t)=1$, és legyen $H=F+G$, akkor

$$\int_a^b\int_a^b r(s,t)dH(s)dH(t)=\int_a^b\int_a^b r(s,t)dF(s)dF(t)+2\int_a^b\int_a^b r(s,t)dF(s)dG(t)+$$

$$+\int_a^b\int_a^b r(s,t)dG(s)dG(t).$$

A jobb oldal utolsó tagja nyilvánvalóan nem negatív, továbbá

$$\int_a^b\int_a^b r(s,t)dF(s)dG(t)=c\int_a^b dG(t)=0.$$

Így tehát bebizonyítottuk, hogy

$$\mathbf{D}^2(m_H^*)\geqq\mathbf{D}^2(m_F^*).$$

[1] Lásd e tekintetben még az 5.4. szakaszt, ahol ugyanezek az eredmények általánosabb alakban tetszőleges lineáris becslések esetére adódnak.

Már korábban láttuk, hogy az $e^{-\beta|t-s|}$ korrelációs függvény esetében (ez a stacionárius Markov-folyamatnak felel meg, ha teljesül az a járulékos feltétel, hogy minden valószínűségi eloszlás normális) nem lehet leghatékonyabb $f(t)$ függvényt szerkeszteni $L_2(t)$-ben a folyamat átlagára vonatkozó hipotézis ellenőrzésére.

Ha viszont az

$$\int_0^T e^{-\beta|t-s|}dF(t)=\frac{1}{2+\beta T}$$

egyenletet tekintjük, könnyen belátható, hogy az

$$F(t)=\frac{\varepsilon(t)+\varepsilon(t-T)+\beta t}{2+\beta T}$$

korlátos variációjú függvény kielégíti ezt az egyenletet, és hogy

$$\int_0^T dF(t)=1.$$

Ebből következik, hogy az

$$m^*=\frac{x(0)+x(T)+\beta\int_0^T x(t)dt}{2+\beta T}$$

kifejezés m torzítatlan becslését adja, és szórása minimális a becslések tekintett osztályában. Később látni fogjuk, hogy normális folyamat esetében ez a becslés a becsléseknek egy sokkal tágabb osztályában is legjobb.

Második példaként tekintsünk egy az (a, b) intervallumon megadott, időben homogén ortogonális (korrelálatlan növekményű) folyamatot, amelynek létezzék az általunk nem ismert m átlaga és amelynek a korrelációs függvénye $r(s, t)=$ $=\min(s, t)$ legyen. Ebben az esetben az

$$\int_a^b \min(s, t)dF(t)=c$$

integrálegyenletnek létezik az

$$F(t)=\varepsilon(t-a)\cdot\frac{c}{a}$$

megoldása, tehát a leghatékonyabb becslés

$$m^*=x(a).$$

5. 3. A számtani közép szerinti becslés. A stacionárius folyamatok esetében az 5. 2. szakaszban vizsgált feladatnak van néhány speciális sajátsága. Ha a spektrál függvénynek nincs szakadása a $\lambda=0$ pontban (azaz a spektrális energia a $\lambda=0$ pontban zérus), akkor az

$$m^*=\frac{1}{2T}\int_{-T}^{T} x(t)dt$$

becslés, amely nyilvánvalóan torzítatlan, $T\to\infty$ esetén sztochasztikusan m-hez tart (lásd az 1. 3. szakaszt), vagyis m^* itt konzisztens becslés. Ennek a becslésnek, amelyet *számtani közép szerinti becslésnek* fogunk nevezni, van egy másik optimális tulajdonsága is, amellyel ebben a szakaszban foglalkozunk. Lehetségesnek látszik a kapott eredmény általánosítása is; ezzel a szerző egy későbbi munkájában kíván foglalkozni.

Tegyük fel, hogy a spektrum abszolút folytonos és létezik a $h(\lambda)$ spektrális sűrűségfüggvény, amely folytonos a $\lambda=0$ helyen és korlátos. Tekintsük az alábbi típusú torzítatlan becsléseket:

$$m_T^*=\int_{-T}^{T} f(t)x(t)\,dt;\quad \int_{-T}^{T} f(t)\,dt=1.$$

Az

$$f(t)=\frac{1}{T}g\left(\frac{t}{T}\right)$$

helyettesítéssel ekkor

$$m_T^*=\int_{-1}^{1} g(u)x(uT)\,du;\quad \int_{-1}^{1} g(u)\,du=1.$$

A $g(u)$ függvény itt tehát azt a viszonylagos súlyt képviseli, amelyet a folyamat különböző t-khez tartozó értékeinek tulajdonítunk. A torzítatlan becsléseknek arra a reguláris osztályára szorítkozva, amely a $(-1, 1)$ intervallumon egyenletesen folytonos és egyenletesen korlátos C függvényosztályhoz tartozó $g(u)$ függvényeknek felel meg, ki fogjuk mutatni, hogy *ebben az osztályban a számtani közép szerinti becslés aszimptotikusan minimális szórású, ha T a végtelenhez tart.* Pontosabban: definiáljuk a

$$\mu_T^*=\frac{1}{2T}\int_{-T}^{T} x(t)\,dt$$

becslés hatékonyságát („efficienciáját") az

$$e_T=\frac{\inf\limits_{m^*\in C} \mathbf{D}^2 m_T^*}{\mathbf{D}^2\mu_T^*};\quad 0\leqq e_T\leqq 1$$

kifejezéssel (a hatékonyságnak ez a definíciója eltér a klasszikus elméletben használt meghatározástól, amennyiben a folyamatnak csupán a lineáris tulajdonságait veszi tekintetbe). Ki fogjuk mutatni, hogy $e_T\to 1$, ha T a végtelenhez tart. Legyen ugyanis $\lim\limits_{T\to\infty} e_T=e$. Ekkor létezik olyan $T_\nu\to\infty$ sorozat és olyan $g_\nu(\mu)\in C$ függvénysorozat, hogy a hozzátartozó $m_{T_\nu}^*$ becslés-sorozatra érvényes a

$$\frac{\mathbf{D}^2 m_{T_\nu}^*}{\mathbf{D}^2\mu_{T_\nu}^*}\to e$$

aszimptotikus összefüggés $\nu\to\infty$ esetén. Vezessük be a

$$\gamma_\nu(\lambda)=\int_{-1}^{1} e^{iu\lambda} g_\nu(u)\,du$$

függvényeket. Ekkor

$$\mathbf{D}^2(m_{T_\nu}^*) = \frac{1}{T_\nu^2} \int_{-T_\nu}^{T_\nu} \int_{-T_\nu}^{T_\nu} r(s,t) g_\nu\left(\frac{t}{T_\nu}\right) ds\, dt = \int_{-\infty}^{\infty} |\gamma_\nu(T_\nu \lambda)|^2 h(\lambda)\, d\lambda =$$

$$= \frac{1}{T_\nu} \int_{-\infty}^{\infty} |\gamma_\nu(\mu)|^2 h\left(\frac{\mu}{T_\nu}\right) d\mu.$$

A számtani közép szerinti becslésre ugyanilyen módon a

$$\mathbf{D}^2 \mu_{T_\nu}^* = \frac{1}{T_\nu} \int_{-\infty}^{\infty} \frac{\sin^2 \mu}{\mu^2} h\left(\frac{\mu}{T_\nu}\right) d\mu$$

egyenletet kapjuk, amelyből a Fejér-féle magok egy tulajdonságánál fogva következik, hogy

$$T_\nu \mathbf{D}^2 \mu_{T_\nu}^* \to \pi h(0)$$

ha $\nu \to \infty$. A $g_\nu(u)$ függvények egyenletes folytonossága következtében ki lehet választani olyan részsorozatot, amely a folytonos $g(u)$ függvényhez tart. Ha feltesszük, hogy ez már megtörtént, akkor a korlátos konvergenciára vonatkozó LEBESGUE-tétel szerint

$$\int_{-1}^{1} |g(u) - g_\nu(u)|^2 du \to 0$$

ha $\nu \to \infty$. PLANCHEREL tételének felhasználásával továbbá azt kapjuk, hogy

$$\int_{-\infty}^{\infty} |\gamma_\nu(\mu) - \gamma(\mu)|^2 d\mu \to 0,$$

és

$$\int_{-\infty}^{\infty} \left| \gamma_\nu(\mu) \sqrt{h\left(\frac{\mu}{T_\nu}\right)} - \gamma(\mu) \sqrt{h\left(\frac{\mu}{T_\nu}\right)} \right|^2 d\mu \leqq H \int_{-\infty}^{\infty} |\gamma_\nu(\mu) - \gamma(\mu)|^2 d\mu \to 0$$

ha $\nu \to \infty$. Ennélfogva

$$\int_{-\infty}^{\infty} |\gamma_\nu(\mu)|^2 h\left(\frac{\mu}{T_\nu}\right) d\mu - \int_{-\infty}^{\infty} |\gamma(\mu)|^2 h\left(\frac{\mu}{T_\nu}\right) d\mu \to 0.$$

Azonban

$$\left| \int_{-\infty}^{\infty} |\gamma(\mu)|^2 h\left(\frac{\mu}{T_\nu}\right) d\mu - h(0) \int_{-\infty}^{\infty} |\gamma(\mu)|^2 d\mu \right| \leqq \left| \int_{|\mu| < \varepsilon T_\nu} |\gamma(\mu)|^2 \left[h\left(\frac{\mu}{T_\nu}\right) - h(0) \right] d\mu \right| +$$

$$+ \left| \int_{|\mu| \geqq \varepsilon T_\nu} |\gamma(\mu)|^2 \left[h\left(\frac{\mu}{T_\nu}\right) - h(0) \right] d\mu \right.$$

és ha ε értékét olyan kicsire választjuk, hogy $|h(\mu)-h(0)|<\delta$ legyen ha $|\mu|<\varepsilon$, és azután T_ν-t olyan nagyra választjuk, hogy $\int\limits_{|\mu|\geqq\varepsilon T\mu} |\gamma(\mu)^2 d\mu<\delta$ legyen, akkor

$$\int_{-\infty}^{\infty} |\gamma_\nu(\mu)|^2 h\left(\frac{\mu}{T_\nu}\right) d\mu \to h(0) \int_{-\infty}^{\infty} |\gamma(\mu)|^2 d\mu$$

ha $\nu\to\infty$. Azonban PLANCHEREL tétele szerint

$$\int_{-\infty}^{\infty} |\gamma(\mu)|^2 d\mu = 2\pi \int_{-1}^{1} |g(u)|^2 du,$$

tehát a SCHWARZ-féle egyenlőtlenség értelmében

$$e=2\int_{-1}^{1} |g(u)|^2 du \geqq \left\{\int_{-1}^{1} |g(u)|du\right\}^2 \geqq \left\{\int_{-1}^{1} g(u)du\right\}^2 = 1.$$

Az e mennyiség definíciójánál fogva azonban csupán az egyenlőség lehetséges, tehát állításunkat bebizonyítottuk.

Ennek az eredménynek a legfontosabb következménye az alábbi

KOROLLÁRIUM: *Az*

$$m^* = \frac{1}{T}\int_{-T}^{T} g\left(\frac{t}{T}\right) x(t)dt$$

típusú becslések között — ahol $g(u)$ a $(-1, 1)$-ben definiált olyan függvény, amelyre $\int_{-1}^{1} g(u)du = 1$, és amely független T-től — nem lehet találni a számtani közép szerinti becslésnél aszimptotikusan hatékonyabb becslést.

A fenti bizonyítás továbbá azt is kimutatja, hogy e becslések közül a leghatékonyabb ténylegesen azonos a számtani közép szerinti becsléssel.

Megjegyezzük, hogy a $g_\nu(u)\in C$ feltételt nem lehet elhagyni anélkül, hogy ugyanakkor ne szűkítenénk valamiképpen a tekintett folyamatok osztályát. Vegyük ugyanis példaképpen azt a folyamatot, amelynek korrelációs függvénye $e^{-\frac{t^2}{2}}$. Ha ennek a folyamatnak észleljük egy realizációját tetszőleges nem elfajult intervallumon, akkor ismerjük a realizációt t minden értékére, mert ez a folyamat analitikus minden t-re. Ezért a realizáció tetszőlegesen kicsi szakaszon ismert értékeiből kiindulva képezhetjük az $\frac{1}{2A}\int_{-A}^{A} x(t)dt$ kifejezést, amely a folyamat spektrumának folytonossága miatt $A\to\infty$ esetén átlagban az m igazi értékéhez tart. A számtani közép szerinti becslés efficienciája tehát ebben az esetben nulla az összes lineáris becslések osztályában.

5. 4. Doob-féle elemi folyamatok. Láttuk, hogy a folyamatok átlagára vonatkozó becslések vizsgálatánál sokszor nem volt elegendő csupán $\int x(t)f(t)dt$ típusú becslésekre szorítkozni, ahol $f(t)$ négyzetesen integrálható függvény, hanem az előbbiekben Stieltjes-integrálokat is be kellett vezetni. Ezért szükséges, hogy ezt a kérdést más szempontból is megvizsgáljuk.

Tegyük fel, hogy ismerjük az $y(t) = m+x(t)$ folyamatnak az (a, b) időintervallumban észlelt értékeit, és legyen $\mathbf{E}x(t)=0$. Keressük m-nek egy minimális szórású torzítatlan becslését. Ha m^* torzítatlan becslés és

$$m^* = \operatorname*{l.i.m.}_{n\to\infty} m_n^*; \quad m_n^* = \sum_1^n e_\nu^{(n)} y(t_\nu^{(n)}); \quad t_\nu^{(n)} \in (a, b),$$

akkor

$$\mathbf{E}m_n^* = m\sum_1^n c_\nu^{(n)} \to m, \quad \text{ha} \quad n\to\infty.$$

Ebben az esetben tehát

$$m^* = \operatorname*{l.i.m.}_{n\to\infty} \left\{\sum_1^n c_\nu^{(n)} y(t_\nu^{(n)}) + \left[1-\sum_1^n c_\nu^{(n)}\right] y(a)\right\},$$

mert a szögletes zárójelben levő kifejezés zérushoz tart, ha $n\to\infty$. Ilyenképpen minden torzítatlan lineáris becslés

$$\sum_1^n c_\nu^{(n)} y(t_\nu^{(n)})$$

alakú véges összegek határértékének tekinthető, ahol $\sum_1^n c_\nu^{(n)} = 1$.

Tekintsük most a

$$\sum_1^n c_\nu x(t_\nu)$$

alakú valószínűségi változók által alkotott $M_0 \subset L_2(X; a, b)$ halmazt, ahol

$$\sum_1^n c_\nu = 1; \quad t_\nu \in (a, b); \quad \nu = 1, 2, \ldots, n.$$

Ennek a halmaznak átlag szerinti konvergenciában való lezárásával egy új $M \subset \subset L_2(X; a, b)$ halmazt kapunk. Mivel az M halmaz nyilvánvalóan zárt és konvex, létezik legalább egy olyan μ^* eleme, amelyre

$$\|\mu^*\| = \inf_{z\in M} \|z\|.$$

Könnyen ki lehet mutatni, hogy valójában csupán egyetlen ilyen elem létezhetik. Legyen ugyanis μ_1^* egy másik ilyen elem. Ekkor $\frac{\mu^*+\mu_1^*}{2} \in M$, és

$$\left\|\frac{\mu^*+\mu_1^*}{2}\right\|^2 = \frac{\|\mu^*\|^2}{4} + \frac{\|\mu_1^*\|^2}{4} + 2Re\frac{\mathbf{E}\mu^*\overline{\mu}_1^*}{4} < \|\mu^*\|^2,$$

ami ellentmond μ^* definíciójának. Így tehát *a torzítatlan lineáris becslések osztályában létezik egyetlen minimális szórású becslés.*

Jelöljük ezt a becslést m^*-gal és tekintsük az $\mathbf{E}m^*\overline{x(t)}$, $a \leqq t \leqq b$, kifejezést. Ha

$$\mathbf{E}m^*\overline{x(\alpha)} \neq \mathbf{E}m^*\overline{x(\beta)}; \quad a \leqq \alpha, \beta \leqq b,$$

akkor bevezetjük az

$$m_1^* = m^* + \varepsilon[x(\alpha) - x(\beta)]$$

becslést, amely szintén torzítatlan. Ebben az esetben

$$\|m_1^* - m\|^2 = \|m^* - m\|^2 + 2Re\{\bar{\varepsilon}\mathbf{E}[m^* - m]\overline{[x(\alpha) - x(\beta)]}\} + |\varepsilon|^2 \|x(\alpha) - x(\beta)\|^2,$$

tehát lehet választani olyan ε értéket, hogy $\mathbf{D}m_1^* < \mathbf{D}m^*$ legyen, ellentmond m^* definíciójának. *Következésképpen az* $\mathbf{E}m^*\overline{x(t)}$ *függvénynek állandó értéket kell felvennie minden az* (a, b) *intervallumhoz tartozó* t *esetében:*

$$\mathbf{E}m^*\overline{x(t)} = c; \quad t \in (a, b).$$

Mivel azonban $m^* = \underset{n\to\infty}{\text{l.i.m.}}\, m_n^*$, ahol

$$m_n^* = \sum_1^n c_\nu^{(n)} y(t_\nu^{(n)}),$$

következik

$$\mathbf{D}^2 m^* = \|m^* - m\|^2 = \lim_{n\to\infty} \mathbf{E}(m^* - m)\overline{(m_n^* - m)} = \lim_{n\to\infty} \mathbf{E}\, m^*\overline{(m_n^* - m)} = c \sum_1^n \overline{c_\nu^{(n)}} = c,$$

tehát *a c állandó egyenlő az* m^* *becslés szórásával.*

Megjegyezzük még, hogy *ennek az egyenletnek a megoldása mindig megadja az egyértelműen meghatározott legkisebb szórású torzítatlan lineáris becslést.* Ugyanis ha

$$\left.\begin{aligned} \mathbf{E}m^*\overline{x(t)} &= c \\ \mathbf{E}m_1^*\overline{x(t)} &= c_1 \end{aligned}\right\}, \quad a \leqq t \leqq b,$$

akkor a $c = c_1$ egyenlőségből következik, hogy $\|m^* - m_1\|^2 = 0$, vagyis hogy a két becslés egybeesik. Ha viszont $c \neq c_1$, akkor mindig feltehetjük, hogy $c \neq 0$, ekkor azonban nyilvánvalóan

$$m_1^* = \frac{c_1}{c} m^*,$$

és abból, hogy mindkét becslés torzítatlan, közvetlenül következik $c_1 = c$.

Alkalmazzuk az itt kapott eredményeket a stacionárius folyamatoknak egy Doob által bevezetett fontos osztályára [5]. Itt csupán ennek az osztálynak a reguláris (nem-determinisztikus) típusú folyamataival kívánunk foglalkozni, vagyis feltesszük, hogy a folyamat korrelációs függvényét elő lehet állítani az

$$r(t) = \int_{-\infty}^{\infty} e^{it\lambda} \frac{d\lambda}{|a_n(i\lambda)^n + a_{n-1}(i\lambda)^{n-1} + \ldots + a_0|^2}$$

alakban, ahol a nevezőben levő polinom együtthatói valósak és a λ_ν zérushelyek mind a felső félsíkban helyezkednek el:

$$a_n(i\lambda)^n+\ldots+a_0=a_n i^n \prod_1^n (\lambda-\lambda_\nu).$$

Ugyanezt a folyamatot elő lehet állítani egy állandó együtthatójú lineáris differenciálegyenlet megoldásaként is, (KARHUNEN [2]). Ha $n=1$, akkor olyan folyamatot kapunk, amelynek a korrelációs függvénye $e^{-\beta|t|}$.

Közvetlenül belátható, hogy a tekintett folyamatnak léteznek a deriváltjai az $(n-1)$-edrendű deriváltig bezárólag (az átlagban való konvergencia értelmében). Tekintsük az

$$m^*=\frac{\sum\limits_0^{n-1}\{\alpha_\nu x^{(\nu)}(0)+\beta_\nu x^{(\nu)}(T)\}+a_0\int\limits_0^T x(s)\,ds}{2a_1+a_0T}$$

becslést, ahol $x(t)$ a megfigyelt folyamat, és

$$\left.\begin{aligned}\alpha_\nu&=(-1)^\nu a_{\nu+1}\\ \beta_\nu&=a_{\nu+1}\end{aligned}\right\},\quad \nu=0,1,\ldots,(n-1).$$

Ilyen becslés mindig lehetséges, mert

$$ia_1=(-1)^{n-1}i^n a_n\sum_1^n\frac{\lambda_1\ldots\lambda_n}{\lambda_\nu}=-a_0\sum_1^n\frac{1}{\lambda_\nu}=-a_0\sum_1^n\frac{\bar\lambda_\nu}{|\lambda_\nu|^2},$$

ahonnan kitűnik, hogy $i\dfrac{a_1}{a_0}$ $(a_0\neq 0)$ képzetes része pozitív, tehát a_1 és a_0 azonos előjelűek, és így a becslés nevezője nem lehet zérus. A becslés nyilvánvalóan torzítatlan, és ki fogjuk mutatni, hogy egyszersmind a szórása is minimális. Tekintsük az

$$\mathbf{E}m^*\overline{[x(t)-m]}=\frac{\sum\limits_0^{n-1}\{\alpha_\nu r^{(\nu)}(-t)+\beta_\nu r^{(\nu)}(T-t)\}+a_0\int\limits_0^T r(s-t)\,dt}{2a_1+a_0T}$$

kifejezést. Azonban

$$a_n\frac{d^n r(t)}{dt^n}+a_{n-1}\frac{d^{n-1}r(t)}{dt^{n-1}}+\ldots+a_0=\int\limits_{-\infty}^{\infty}e^{it\lambda}\frac{a_n(i\lambda)^n+\ldots+a_0}{[a_n(i\lambda)^n+\ldots+a_0]^2}\,d\lambda=$$

$$=\int\limits_{-\infty}^{\infty}e^{it\lambda}\frac{d\lambda}{a_n(-i)^n\prod\limits_1^n(\lambda-\bar\lambda_\nu)},$$

ahol a jobb oldal $t>0$ esetében CAUCHY tétele szerint azonosan zérus. Hasonlóképpen kapjuk $t<0$ esetére, hogy

$$a_n\frac{d^n r(t)}{dt^n}-a_{n-1}\frac{d^{n-1}r(t)}{dt^{n-1}}+\ldots+(-1)^n a_0 r(t)=0.$$

Így tehát

$$a_0\int_0^T r(s-t)\,ds = a_0\lim_{\varepsilon\to 0}\left\{\int_0^{t-\varepsilon} r(s-t)\,ds+\int_{t+\varepsilon}^{T} r(s-t)\,ds = \lim_{\varepsilon\to 0}\{a_1[r(-\varepsilon)-r(-t)]-\right.$$

$$-a_2[r'(-\varepsilon)-r'(-t)]+\ldots+(-1)^{n+1}a_n[r^{(n-1)}(-\varepsilon)-r^{(n-1)}(-t)]-a_1[r(T-t)-r(\varepsilon)]-$$

$$\left.-a_2[r'(T-t)-r'(\varepsilon)]-\ldots-a_n[r^{(n-1)}(T-t)-r^{(n-1)}(\varepsilon)]\right\}=2a_1r(0)+2a_3r''(0)+\ldots$$

$$\ldots+2a_\mu r^{(\mu-1)}(0)-a_1r(-t)-a_1r(T-t)+a_2r'(-t)-a_2r'(T-t)+\ldots$$

$$\ldots+(-1)^n a_n r^{(n-1)}(-t)-a_n r^{(n-1)}(T-t),$$

ahol $\mu=n$, ha n páratlan és $\mu=n-1$, ha n páros. A fentiekből következik, hogy

$$\mathrm{E}m^*\overline{[x(t)-m]}=\frac{2a_1r(0)+2a_2r''(0)+\ldots+2a_\mu r^{(\mu-1)}(0)}{2a_1+a_0T}=\text{konst.}$$

A jobb oldal itt nem függ t-től, amiből következik, hogy m^* *valóban megadja az egyetlen minimális szórású torzítatlan becslést.* A szórás meghatározására csupán az utóbbi egyenlet jobb oldalán szereplő állandó értékét kell kiszámítani. Az $r(t)$ számára megadott két differenciálegyenletet $-A$ és A határok között integrálva, és tekintetbe véve, hogy $r(\pm A)$, $r'(\pm A)$, ..., $r^{(n-1)}(\pm A)$ a Fourier-együtthatókra vonatkozó LEBESGUE-tétel szerint nullához tart, ha $A\to\infty$, azt kapjuk, hogy

$$2a_1r(0)+2a_3r''(0)+\ldots+2a_\mu r^{(\mu-1)}(0)=\lim_{A\to\infty}a_0\int_{-A}^{A}r(t)\,dt.$$

Így tehát a fenti leghatékonyabb becslés szórása

$$\frac{a_0 2\pi\frac{1}{a_0^2}}{2a_1+a_0T}=\frac{2\pi}{a_0(2a_1+a_0T)}.$$

5. 5. Tisztán nem-determinisztikus folyamatok. Láttuk, hogyha a folyamat egy realizációjának egy A hosszúságú szakaszon való ismerete alapján extrapolálni tudjuk a folyamat értékeit minden más időpontra, akkor a folyamat A-nál hosszabb időszakaszon végzett megfigyelése alapján az átlag számára szerkesztett, számtani közép szerinti becslés hatékonysága nulla. Ennek az eshetőségnek az elkerülése végett közelfekvő a vizsgálatot olyan sztochasztikus folyamatokra korlátozni, amelyek regulárisak, más szóval „tisztán nem-determinisztikusak" (ilyen folyamatokról lásd pl. HANNER [1] és KARHUNEN [4] [2]). Ebben az esetben a folyamat spektrál eloszlásfüggvénye abszolút folytonos és azzal a tulajdonsággal bír, hogy létezik olyan abszolút értékben négyzetesen integrálható $f(\lambda)$ függvény, amelyre

$$|f(\lambda)|^2=F'(\lambda)$$

[2]Magának a reguláris folyamatnak a fogalmát KOLMOGOROV vezette be először [1*]. Lásd még DOOB [1*], XII. fej.

és ahol a

$$g(a) = \underset{A\to\infty}{\text{l. i. m.}} \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} e^{ia\lambda} f(\lambda)\, d\lambda$$

függvény a negatív féltengelyen mindenütt nulla. Feltesszük még ezenkívül, hogy λ kis értékeire

$$f(\lambda) = f_0 + f_1\lambda(1 + 0(\lambda)),$$

ahol f_0 valós és zérustól különböző.

Tekintsük λ függvényeinek azt a Hilbert-terét, amelyben a belső szorzatot a következő kifejezés adja meg:

$$(\varphi, \psi) = \int_{-\infty}^{\infty} \varphi(\lambda)\overline{\psi(\lambda)}\, dG(\lambda),$$

ahol

$$G(\lambda) = F(\lambda) + \varepsilon(\lambda).$$

Ebben a térben a

$$H(\lambda) = \frac{e^{-iT\lambda}\dfrac{1}{f(\lambda)} - e^{-iT\lambda}\dfrac{1}{f(\lambda)}}{\lambda}$$

függvény, úgyszintén az $e^{it\lambda}$, $-T \leqq t \leqq T$, függvények nyilvánvalóan véges normájúak. Hilbert-terünknek az $e^{it\lambda}$, $-T = t = T$ függvényeken kifeszített alterét jelöljük $\lambda_2(T)$-vel. Legyen

$$H(\lambda) = e^{-iT\lambda}\frac{\dfrac{1}{f(\lambda)} - \dfrac{1}{f(0)}}{\lambda} + \frac{e^{-iT\lambda} - e^{iT\lambda}}{f(0)\lambda} + e^{iT\lambda}\frac{\dfrac{1}{f(0)} - \dfrac{1}{f(\lambda)}}{\lambda} = H_1(\lambda) + H_2(\lambda) + H_3(\lambda).$$

Ekkor a H_1, H_2, H_3 függvények mindegyike véges normájú, azonkívül

$$H_2 = \frac{1}{if(0)} \int_{-T}^{T} e^{i\lambda t}\, dt \in \lambda_2(T).$$

Legyen

$$\begin{cases} H^*_{1,T} = P_{\lambda_2(T)} H_1, \\ H^*_{3,T} = P_{\lambda_2(T)} H_3 \end{cases}$$

és

$$H^*_T(\lambda) = H^*_{1,T}(\lambda) + H_2(\lambda) + H^*_{3,T}(\lambda) \in \lambda_2(T).$$

Ha

$$x(t) = \int_{-\infty}^{\infty} e^{it\lambda}\, dZ(\lambda)$$

az $x(t)$ folyamat spektrális felbontása, akkor

$$y(t) = m + x(t) = \int_{-\infty}^{\infty} e^{it\lambda}\, dZ'(\lambda),$$

ahol

$$Z'(\lambda) = Z(\lambda)+m\varepsilon(\lambda).$$

Az $y(t)$ folyamatnak a $(-T, T)$ intervallumon történt megfigyelése alapján meg lehet szerkeszteni az

$$m_0^* = \int_{-\infty}^{\infty} H_T^*(\lambda)\,dZ'(\lambda)$$

becslést. Ha ugyanis

$$H_T^*(\lambda)\ \underset{n\to\infty}{\mathrm{l.\,i.\,m.}} \sum_1^n c_\nu^{(n)} e^{it_\nu^{(n)}}\lambda, \quad -T \leqq t_\nu^{(n)} \leqq T,$$

akkor nyilvánvalóan

$$m_0 = \underset{n\to\infty}{\mathrm{l.\,i.\,m.}} \sum_1^n c_\nu^{(n)} y\big(t_\nu^{(n)}\big).$$

Ebben az esetben

$$\mathbf{E}m_0^* = mH_T^*(0) = m[H_{1,T}^*(0)+H_2(0)+H_{3,T}^*(0)].$$

Azonban az

$$\frac{1}{\lambda}\left[\frac{1}{\overline{f(\lambda)}}-\frac{1}{f(0)}\right]$$

függvénynek $H_{1,T}'(\lambda)$ projekciója az $\{e^{i(T+t)\lambda};\ -T \leqq t \leqq T\}$ elemek által kifeszített altérre nyilvánvalóan egyenlő az $e^{iT\lambda}H_{1,T}^*(\lambda)$ függvénnyel, ezért $T\to\infty$ esetén egy $H_{1,\infty}'(\lambda)$ határértékhez tart. Így tehát

$$(H_1, 1) = (H_1^*, 1) = e^{-iT\lambda}H_{1,T}'(\lambda) - e^{-iT\lambda}H_{1,\infty}'(\lambda), 1) +$$
$$+ (e^{-eT\lambda}H_{1,\infty}', 1).$$

$T\to\infty$ esetén az egyenlet bal oldalának a határértékét közvetlenül ki lehet számítani:

$$(H_1', 1) = \int_{-\infty}^{\infty} H_1(\lambda)\,dG(\lambda) =$$

$$= \int_{-\infty}^{\infty} \frac{f(\lambda) - \dfrac{|f(\lambda)|^2}{f(0)}}{\lambda} e^{-iT\lambda}d\lambda + H_1(0) \to H_1(0) = \frac{\bar{f}_1}{f_0^2}.$$

De a jobb oldal első tagja $T\to\infty$ esetén zérushoz tart:

$$|(e^{-iT\lambda}H_{1,T} - e^{-iT\lambda}H_{1,\infty}, 1)| \leqq \|1\|\cdot\|H_{1,T}' - H_{1,\infty}'\| \to 0,$$

a második tag határértéke pedig:

$$(e^{-iT\lambda}H_{1,\infty}, 1) = \int_{-\infty}^{\infty} H_{1,\infty}'(\lambda)e^{-iT\lambda}|f(\lambda)|^2 d\lambda + H_{1,\infty}'(0) \to H_{1,\infty}'(0).$$

Így tehát

$$\lim_{T\to\infty} H_{1,T}^*(0) = \lim_{T\to\infty} H_{1,T}'(0) = H_{1,\infty}' = H_1(0).$$

Hasonlóképpen lehet bizonyítani, hogy

$$\lim_{T\to\infty} H_{3,T}^*(0) = H_3(0).$$

A fentiekből következik, hogy az alábbi kifejezés a folyamat átlagának torzítatlan becslése:

$$m^* = \frac{\int\limits_{-\infty}^{\infty} H_T^*(\lambda)\, dZ'(\lambda)}{H_{1,T}^*(0) + \dfrac{2T}{if(0)} + H_{3,T}^*(0)},$$

(a nevező különbözik nullától, ha T elég nagy). Ki fogjuk mutatni, hogy ennek a becslésnek a szórása minimális.

Ugyanis
$$\mathbf{E} m^* \overline{x(t)} = \mathbf{E}[m^* - m]\overline{x(t)},$$

és

$$\left[H_{1,T}^*(0) + \frac{2T}{if(0)} + H_{3,T}^*(0)\right] \mathbf{E}[m^* - m]\overline{x(t)} = \int\limits_{-\infty}^{\infty} H_T^*(\lambda) e^{-it\lambda} F'(\lambda)\, d\lambda =$$

$$= \int\limits_{-\infty}^{\infty} H_T^*(\lambda) e^{-it\lambda} dG(\lambda) - H_T(0) = \int\limits_{-\infty}^{\infty} H(\lambda) e^{-it\lambda} dG(\lambda) - H_T^*(0) =$$

$$= \int\limits_{-\infty}^{\infty} e^{-it\lambda} H(\lambda) F'(\lambda)\, d\lambda + H(0) - H_T^*(0).$$

Legyen most

$$n(t) = \int\limits_{-\infty}^{\infty} H(\lambda) e^{-it\lambda} F'(\lambda)\, d\lambda.$$

Ekkor $-T \leqq t_1,\ t_2 \leqq T$ esetén

$$n(t_2) - n(t_1) = \int\limits_{-\infty}^{\infty} \frac{e^{i(-T-t_2)\lambda} - e^{i(-T-t_1)\lambda}}{\lambda} f(\lambda)\, d\lambda - \int\limits_{-\infty}^{\infty} \frac{e^{i(T-t_2)\lambda} - e^{i(T-t_1)\lambda}}{\lambda} f(\lambda)\, d\lambda.$$

A jobb oldal első tagja itt PLANCHEREL tétele szerint:

$$\frac{i}{\sqrt{2\pi}} \int\limits_{-T-t_1}^{-T-t_2} g(a)\, da = 0,$$

mert a fenti feltétel értelmében

$$\begin{cases} -T \leqq t_2 \\ -T \leqq t_1. \end{cases}$$

Ugyanígy lehet kimutatni, hogy a második tag is nulla, mert

$$\begin{cases} t_2 \leqq T \\ t_1 \leqq T. \end{cases}$$

Ezért

$$n(t) = C(T) \quad \text{ha} \quad -T \leqq t \leqq T,$$

tehát m^* csakugyan minimális szórású becslés. Hasonlóképpen lehet kimutatni, hogy $C(T)$ valóban független T-től.

Számítsuk most ki, mekkora az m^* becslés szórása. Tekintsük a

$$C = \int_{-\infty}^{\infty} H(\lambda) F'(\lambda) d\lambda = 2i \int_{-\infty}^{\infty} \cos T\lambda \frac{Im f(\lambda)}{\lambda} d\lambda - 2i \int_{-\infty}^{\infty} \frac{\sin T\lambda}{\lambda} \operatorname{Re} f(\lambda) d\lambda$$

kifejezést. A jobb oldalon az első integrál nullához tart, ha $T \to \infty$ (mert integrálható függvénynek a Fourier-transzformáltja), a második tag határértéke pedig $-2i\pi f(0)$. Így tehát az m mennyiség legjobb torzítatlan becslésének a szórása

$$\mathbf{D}^2 m^* = \frac{-2if(0)\pi + H(0) - H_1^*(0)}{H_{1,T}^*(0) + \dfrac{2T}{if(0)} + H_{3,T}^*(0)} \sim \frac{\pi f(0)^2}{T} = \frac{\pi F'(0)}{T}.$$

Ezzel bebizonyítottuk, hogy a reguláris (tisztán nem-determinisztikus) sztochasztikus folyamatok itt vizsgált osztályának az esetében az aritmetikai közép szerinti becslés aszimptotikusan efficiens a lineáris torzítatlan becslések osztályában (és nem csupán az 5. 3. szakaszban vizsgált szűkebb osztályban).

5. 6. A becslések efficienciája. Ebben a szakaszban és a következő szakaszokban a maximum likelihood módszerrel fogunk foglalkozni. Habár általában ennek a módszernek az átvitele lehetséges a folytonos idő-paraméterű sztochasztikus folyamatokra, néhány igen érdekes és mélyreható bonyodalom vetődik fel. Ezeknek a problémáknak a megoldását az 5. 7—5. 9. szakaszokban fogjuk először megkísérelni.

Itt is feltesszük, hogy az α paraméterre (amely az A intervallumból veszi értékeit) a reguláris eset áll fenn, továbbá feltesszük, hogy az $f(\omega, \alpha)$ likelihood függvény majdnem biztosan differenciálható, és a derivált eleget tesz a

$$\left| \frac{\partial f(\omega, \alpha)}{\partial \alpha} \right| < F(\omega)$$

feltételnek, ahol $F(\omega)$ egy véges szórású valószínűségi változó (a P_0 valószínűségi mértékre nézve), és ezenkívül

$$E_\alpha \left(\frac{\partial \log f(\omega, \alpha)}{\partial \alpha} \right)^2 < \infty.$$

Vizsgáljuk az α paraméter véges szórású $\alpha^*(\omega)$ becsléseit; ebben az esetben SAKS [1] könyve 15. 1. szakaszában található tétel alkalmazásával *ennek a szórásnak a minimális értékére egy olyan kifejezést lehet levezetni, amely hasonló a véges dimenziószámú esetben kapott kifejezéshez* (lásd CRAMÉR [4]).

Legyen a becslésünk torzított és a torzítás nagysága legyen $b(\alpha)$, akkor

$$\alpha + b(\alpha) = \mathbf{E}_\alpha a^* = \mathbf{E}_0[\alpha^*(\omega) f(\omega, \alpha)]$$

és

$$1 + \frac{db(\alpha)}{d\alpha} = E_0 \left[\alpha^*(\omega) \frac{\partial f(\omega, \alpha)}{\partial \alpha} \right]$$

(ez utóbbi egyenlet jobb oldalán az integrál alatti kifejezés abszolút értékben nem nagyobb az $|\alpha^*(\omega)|F(\omega)$ mennyiségnél, ez utóbbi kifejezés viszont a SCHWARZ-féle egyenlőtlenség értelmében integrálható a P_0 mértékre nézve). Hasonlóképpen kapjuk az

$$1 = \mathbf{E}_\alpha 1 = \mathbf{E}_0 f(\omega, \alpha)$$

és

$$0 = \mathbf{E}_0 \left(\frac{\partial f(\omega, \alpha)}{\partial \alpha} \right)$$

egyenleteket. Így tehát

$$\left(1 + \frac{db}{d\alpha}\right)^2 = \left[\mathbf{E}_0 (\alpha^* - \alpha) \frac{\partial f(\omega, \alpha)}{\partial \alpha}\right]^2 \leqq \mathbf{E}_0[(\alpha^* - \alpha)^2 f(\omega, \alpha)] \mathbf{E}_0 \left[\left(\frac{\partial \log f(\omega, \alpha)}{\partial \alpha}\right)^2 f(\omega, \alpha)\right] =$$

$$= \mathbf{E}_\alpha (\alpha^* - \alpha)^2 E_\alpha \left(\frac{\partial \log f}{\partial \alpha}\right)^2,$$

amiből következik, hogy

$$\mathbf{E}_\alpha (\alpha^* - \alpha)^2 \geqq \frac{\left(1 + \frac{db}{d\alpha}\right)^2}{\mathbf{E}_\alpha \left(\frac{\partial \log f(\omega, \alpha)}{\partial \alpha}\right)^2}.$$

Mellékesen megjegyezzük, hogy ha ezt az eredményt hozzárendelt valószínűségi változókkal bíró pontfolyamatokra alkalmazzuk (amelyeket a 3. 1. szakaszban ismertetett módon koordinátáikkal adunk meg), akkor olyan eredményt kapunk, amely formálisan analóg WOLFOWITZ-nak [1] a szekvenciális becslések minimális szórására vonatkozó tételével.

Könnyen belátható, hogy a legutóbbi összefüggésben az egyenlőség esete akkor és csak akkor következik be, ha

$$\frac{\partial \log f(\omega, \alpha)}{\partial \alpha} = k(\alpha)[\alpha^*(\omega) - \alpha].$$

Ebből, ugyanúgy mint a klasszikus esetben, következik, hogy ha egyáltalában létezik efficiens becslés, akkor elő lehet állítani a maximum likelihood egyenlet egyetlen, nem azonosan állandó megoldásaként.

Példaképpen vizsgáljuk az 5. 2. szakaszban tárgyalt becsléselméleti feladatot és tegyük fel, hogy normális folyamattal és a reguláris esettel van dolgunk. Ez esetben a likelihood függvény

$$f(\omega, m) = e^{-\frac{m^2}{2} \sum_1^\infty a_\nu^2 \lambda_\nu + m \sum_1^\infty a_\nu \lambda_\nu x_\nu},$$

amely kielégíti a regularitás minden feltételét. Ha m^* az m paraméternek véges szórású torzítatlan becslése, akkor a fenti eredménynek megfelelően

$$\mathbf{D}_m^2 m^* \geqq \frac{1}{\mathbf{E}_m \left[\sum_1^\infty a_\nu \lambda_\nu x_\nu - m \sum_1^\infty \lambda_\nu a_\nu^2\right]^2} = \frac{1}{\sum_1^\infty a_\nu^2 \lambda_\nu}.$$

Mivel azonban a jelen esetben

$$\frac{\partial \log f(\omega, m)}{\partial m} = m \left\{ \frac{\sum_1^\infty a_\nu \lambda_\nu x_\nu}{\sum_1^\infty a_\nu^2 \lambda_\nu} - m \right\} \sum_1^\infty a_\nu^2 \lambda_\nu ,$$

tehát az alábbi efficiens becslést kapjuk:

$$m^* = \frac{\sum_1^\infty a_\nu \lambda_\nu x_\nu}{\sum_1^\infty a_\nu^2 \lambda_\nu},$$

amely eredményt közvetlen számítással sem nehéz igazolni.

Különösen érdekesek számunkra a normális stacionárius Markov-folyamatok. Ez esetben

$$f(\omega, m) = e^{\frac{m}{2}\left[x(0)+x(T)+\beta \int_0^T x(t)dt\right] - \frac{m^2}{2}\left(1+\frac{T}{2}\right)}$$

és a fenti módszerrel az alábbi alakú minimális szórású torzítatlan becslést kapjuk:

$$m^* = \frac{x(0)+x(T)+\beta \int_0^T x(t)\,dt}{2+\beta T}.$$

Így tehát ez a becslés ilyen értelemben a legjobb az összes véges szórású becslések közül.

5. 7. A maximum likelihood módszer. Majdnem ugyanolyan módszerekkel, mint a véges dimenziójú esetben, bizonyítani lehet a becslések számos általános tulajdonságát, pl. azt, hogy ugyanannak a paraméternek két efficiens becslése feltétlenül 1 valószínűséggel azonos. Hasonló módszereket lehet alkalmazni több paraméter egyidejű becslésének az esetére is. A maximum likelihood módszernek a sztochasztikus folyamatokra való alkalmazásakor azonban néhány új nehézség merül fel.

Kezdjük az alábbi eredménnyel, amely viszont még nagyon könnyen bizonyítható. Tegyük fel, hogy teljesül az alábbi három feltétel:

1. A $\dfrac{\partial^\nu \log f(\omega, \alpha)}{\partial \alpha^\nu}$ deriváltak, $\nu = 1, 2, 3$, majdnem biztosan léteznek.

2. Minden $\alpha \in A$ esetén érvényesek a

$$\left|\frac{\partial f}{\partial a}\right| < F_1(\omega), \quad \left|\frac{\partial^2 f}{\partial a^2}\right| < F_2(\omega), \quad \left|\frac{\partial^3 \log f}{\partial \alpha^3}\right| < H(\omega)$$

egyenlőtlenségek, ahol $\mathbf{E}_0 F_1 < \infty$, $E_0 F_2 < \infty$ és $\mathbf{E}_\alpha H < k$.

3. Minden $\alpha \in A$ esetén $\mathbf{E}_\alpha \left(\dfrac{\partial \log f}{\partial \alpha}\right)^2$ pozitív és véges.

Tegyük fel, hogy ismerjük a folyamat N független realizációját, és ezeknek a realizációknak az alapján keresünk becslést α-ra. A megfigyelt realizációkat az $\omega_1, \omega_2, \dots, \omega_N$ szimbólumokkal jelölve, vizsgáljuk a megfelelő

$$f(\omega_1, \omega_2, \dots, \omega_N; \alpha) = f(\omega_1; \alpha) f(\omega_2; \alpha) \dots f(\omega_N; \alpha)$$

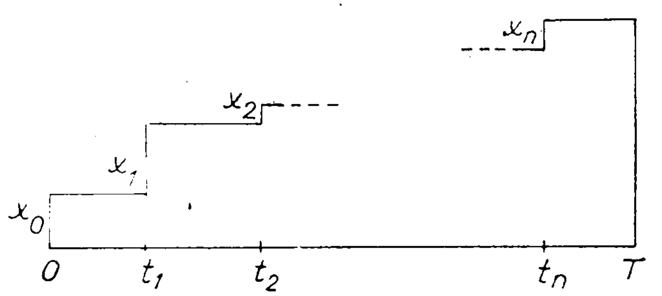
likelihood függvényt. A jelen esetben ki lehet mutatni, hogy a

$$\frac{\partial \log f(\omega_1, \omega_2, \ldots, \omega_N; \alpha)}{\partial \alpha} = 0$$

likelihood egyenletnek létezik $\alpha^*(\omega_1, \omega_2, \ldots, \omega_N)$ megoldása, amely az α paraméternek $N \to \infty$ esetére konzisztens, aszimptotikusan normális és aszimptotikusan efficiens becslése.

Ennek az állításnak a bizonyítása pontosan ugyanolyan, mint a CRAMÉR [4] könyvben az 500—503. oldalakon található megfelelő bizonyítás.

Az eredmény szemléltetésére tekintsük az EINSTEIN [1] által kavicsoknak folyómederben történő mozgására alkalmazott sztochasztikus folyamatot. Figyeljünk meg a $(0, T)$ időintervallum folyamán egy bizonyos kavicsot; minden egyes időpontban kétféle lehetséges állapot közül az egyik állapotban fogjuk találni: vagy mozog, vagy pedig mozdulatlanul fekszik a meder fenekén. Tegyük fel az egyszerűség kedvéért, hogy a mozgás egyes ugrásokból tevődik össze, éspedig úgy, hogy minden egyes ugrás elhanyagolhatóan kicsi idő alatt történik (azaz pillanatnyinak tekinthető). Tegyük fel továbbá, hogy a $t=0$ időpontban a kavics mozog. A többi olyan $t_1, t_2, \ldots, t_n$ időpont, amikor a kavics mozog, legyen rögzített $1/\vartheta$ sűrűségű Poisson eloszlású a $(0, T)$ intervallumon. Könnyű belátni, hogy ϑ az átlagos nyugalmi idő. A $0, t_1, \ldots, t_n$ időpontokban történő ugrások alkalmával megtett távolságok legyenek $x_0, x_1, \ldots, x_n$ (lásd az 1. ábrát). Legyenek itt az x_i távolságok független valószínűségi változók, amelyek csak pozitív értékeket vehetnek fel $1/\xi \cdot e^{-x/\xi}$ valószínűségsűrűséggel, ahol a ξ paraméter egy ugrás átlagos távolságával egyenlő. Tegyük fel, hogy ismerjük N különböző (azonos méretű) kavicson végzett N megfigyelés eredményeit, és ezeknek az adatoknak az alapján akarjuk a ϑ paramétert megbecsülni.



1. ábra

Az $x_0, t_1, x_1, \ldots, t_n, x_n$ koordinátatérben a valószínűség-sűrűséget a $dx_0 \ldots dx_n$ $dt_1 \ldots dt_n$ térfogatelembe való tartozás valószínűségével adjuk meg:

$$\frac{1}{\xi} e^{-\frac{x_0}{\xi}} dx_0 \frac{1}{\vartheta} e^{-\frac{t_1}{\vartheta}} dt_1 \ldots \frac{1}{\vartheta} e^{-\frac{t_n - t_{n-1}}{\vartheta}} dt_n \frac{1}{\xi} e^{-\frac{x_n}{\xi}} dx_n e^{-\frac{T-t_n}{\vartheta}} =$$

$$= \xi^{-(n+1)} e^{-\frac{X}{\xi}} \vartheta^{-n} e^{-\frac{T}{\vartheta}} dx_0 \ldots dx_n dt_1 \ldots dt_n,$$

ahol az $X = x_0 + x_1 + \ldots + x_n$ jelölést alkalmaztuk. A különböző realizációknak megfelelő koordinátákat az $i = 1, 2, \ldots, N$ indexekkel számozva az összes realizációkat jellemző koordinátatérben a valószínűségsűrűséget az alábbi alakban kapjuk meg:

$$\xi^{-\sum\limits_1^N (n_i+1)} e^{-\sum\limits_1^N \frac{x_i}{\xi}} \vartheta^{-\sum\limits_1^N n_i} e^{-\frac{NT}{\vartheta}} .$$

Ennek alapján könnyen meg tudjuk határozni az X mennyiségnek (a kavics összelmozdulásának) a valószínűségeloszlását. Itt csupán az eloszlás első négy momentumára van szükségünk, amelyek:

$$\alpha_1 = \xi \left(1 + \frac{T}{\vartheta}\right)$$

$$\alpha_2 = \xi^2 \left(2 + 4\frac{T}{\vartheta} + \frac{T^2}{\vartheta^2}\right)$$

$$\alpha_3 = \xi^3 \left(6 + 18\frac{T}{\vartheta} + 9\frac{T^2}{\vartheta^2} + \frac{T^3}{\vartheta^3}\right)$$

$$\alpha_4 = \xi^4 \left(24 + 96\frac{T}{\vartheta} + 72\frac{T^2}{\vartheta^2} + 16\frac{T^3}{\vartheta^3} + \frac{T^4}{\vartheta^4}\right)$$

$$\mu_2 = \xi^2 \left(1 + 2\frac{T}{\vartheta}\right)$$

$$\mu_3 = \xi^3 \left(2 + 6\frac{T}{\vartheta}\right)$$

$$\mu_4 = \xi^4 \left(9 + 36\frac{T}{\vartheta} + 12\frac{T^2}{\vartheta^2}\right)$$

ahol a szokásos módon a közönséges momentumokat α_i-vel és a centrális momentumokat μ_i-vel jelöltük. A ϑ paraméter maximum likelihood becslése igen egyszerű:

$$\vartheta^* = \frac{NT}{\sum\limits_1^N n_i} .$$

Ha $\sum\limits_1^N n_i = 0$, akkor nyilvánvalóan $\vartheta^* = \infty$, azonban ennek az esetnek a valószínűsége nullához tart, amikor $N \to \infty$, tehát ennek a lehetőségnek csekély a gyakorlati jelentősége. A ϑ^*-ra kapott kifejezésből következik, hogy

$$\sqrt{N}(\vartheta^* - \vartheta) = -\vartheta \frac{\frac{1}{\sqrt{N}} \sum\limits_1^N \left(n_i - \frac{T}{\vartheta}\right)}{\frac{1}{N} \sum\limits_1^N n_i} ,$$

ahol az n_i mennyiségek Poisson-eloszlású független valószínűségi változók T/ϑ átlaggal. A centrális határeloszlástétel értelmében az $\frac{1}{\sqrt{N}}\sum_1^N\left(\eta_i-\frac{T}{\vartheta}\right)$ mennyiség eloszlása aszimptotikusan normális 0 átlaggal és T/ϑ szórással. Mivel $N\to\infty$ esetén az $\frac{1}{N}\sum_1^N n_i$ mennyiség valószínűségben T/ϑ-hoz tart, ezért (a CRAMÉR [4] könyv 20. 6. szakaszában megadott tétel szerint) a ϑ^* becslés aszimptotikusan normális, ϑ átlaggal és nagy N értékek esetén közelítőleg $\frac{1}{N}\frac{\vartheta^3}{T}$ szórással. A becslés efficienciájának a kiszámítására megjegyezzük, hogy

$$\mathbf{E}_\vartheta\left(\frac{\partial \log f(\omega,\vartheta)}{\partial\vartheta}\right)=\mathbf{E}_\vartheta\left(\frac{n}{\vartheta}-\frac{T}{\vartheta^2}\right)^2=\frac{T}{\vartheta^3},$$

amiből közvetlenül belátható, hogy ϑ^* aszimptotikusan efficiens becslés. A folyamat koordinátáiként itt az ugrások időpontjait és hosszúságait használtuk; de a különböző realizációk adataiból megszerkesztett legjobb becslésben csupán a megfigyelt ugrások száma szerepelt. EINSTEIN [1] munkája szerint csak az $X_1, X_2, \ldots, X_N$ változókat figyelték meg; a hordalékok vizsgálatánál az n_i mennyiségeket feltehetőleg nem is tudják megfigyelni. Amennyiben ugyanennek a folyamatnak másféle alkalmazásainál ez a megfigyelés lehetséges volna (habár esetleg nehéz), érdekes megvizsgálni, hogy milyen mértékben csökken a ϑ paraméter becslésének az efficienciája, ha csupán az X_i értékeket figyelik meg. Vizsgáljuk a ϑ paraméternek az X_i értékek alapján szerkesztett következő becslését, amelyet EINSTEIN a momentumok módszerével állított elő, ([1] p. 38):

$$\vartheta_0=T\left\{\frac{1}{\sqrt{1-\frac{m_2}{a_2^2}}}-1\right\},$$

ahol

$$\begin{cases} a_1=\frac{1}{N}\sum_1^N X_i \\ m_2=\frac{1}{N}\sum_1^N X_i^2-a_i^2. \end{cases}$$

A $\vartheta_0^*-\vartheta$ különbséget a minta-momentumok $H(a_1, m_2)$ függvényeként tekintve, azt kapjuk, hogy

$$H(\alpha_1\mu_2)=0 \qquad \mu_2(a_1)\sim\frac{\xi^2(1+2\theta)}{N}$$

$$\left(\frac{\partial H}{\partial a_1}\right)_0=-\frac{\vartheta^3}{\xi T^2}(1+2\theta) \qquad \mu_{11}(a_1, m_2)\sim\frac{\xi^3(2+6\theta)}{N}$$

$$\left(\frac{\partial H}{\partial m_2}\right)_0=\frac{\vartheta^3}{2\xi^2T^2}(1+\theta) \qquad \mu_2(m_2)\sim\frac{\xi^4(8+32\theta+8\theta^2)}{N},$$

ahol $\theta=\frac{T}{\vartheta}$ és a 0-indexszel azokat a mennyiségeket jelöltük, amelyeket az $a_1=\alpha_1$

és $m_2 = \mu_2$ értékeknek a zárójeles kifejezésekbe való helyettesítésekor kapunk. A jobb oldali oszlopban a μ kifejezések a zárójelbe írt mennyiségek centrális momentumait jelentik. A CRAMÉR [4] könyv 28. 4. szakaszában megadott tétel alkalmazásával könnyen ki lehet mutatni, hogy a $\vartheta^* - \vartheta$ mennyiség eloszlása ebben az esetben aszimptotikusan normális 0 átlaggal és

$$\frac{\vartheta^6}{NT^4}(1+6\theta+10\theta^2+8\theta^3+2\theta^4) \qquad \text{szórással.}$$

Így tehát az EINSTEIN-féle becslés efficienciája

$$e(\vartheta_0^*) = \frac{\theta^3}{1+6\theta+10\theta^2+8\theta^3+2\theta^4}.$$

Tekintsük még a következő, az előbbivel rokon sztochasztikus folyamatot, amely az orvosi statisztikának egy problémájával kapcsolatban fordul elő.

Vizsgáljuk valamilyen A jelenség előfordulását vagy hiányát egy intervallumon, amelyet szokás szerint $(0, T)$-vel jelölünk, noha az adott esetben a paraméterünk nem időt jelent, hanem egy egyenesen mért távolság-koordinátát. Tegyük fel, hogy $t=0$ esetén az A^* eseményt észleljük (vagyis az A eseményt nem észleljük). Legyen a mintaterünk olyan függvények tere, amelyeknek az értéke általában csak 0 vagy 1 lehet (éspedig 0 az A^* és 1 az A észlelésekor), $t=0$ esetén pedig csakis 0 lehet. A folyamat koordinátái legyenek az $n_1, n_2, t_1, t_2, \ldots, t_{n_1+n_2}$ számok, ahol t_i az állapotváltozások időpontjait jelenti, n_1 az A állapotok kezdeti időpontjainak a számát, n_2 pedig ugyanezen állapotok vég-időpontjainak a számát jelenti. Nyilvánvaló, hogy vagy $n_1 = n_2$, vagy pedig $n_1 = n_2 + 1$, attól függően, hogy a $t = T$ időpontban az A^* állapotot, illetve az A állapotot észleljük-e. Tegyük fel, hogy az A^* (illetve A) intervallumok hosszának a sűrűségfüggvénye $\beta e^{-\beta t}$, $t>0$ (illetőleg $\alpha e^{-\alpha t}$, $t>0$). Ekkor a koordinátatér egy térfogatelemébe való tartozás valószínűségét a következő módon kapjuk meg:

$$e^{-\beta t_1}\beta dt_1 e^{-\alpha(t_2-t_1)}\alpha dt_2 \ldots e^{-\alpha(t_{n_1+n_2}-t_{n_1+n_2-1})}\alpha dt_{n_1+n_2} e^{-\beta(T-t_{n_1+n_2})} =$$
$$= e^{-\beta l'}\beta^{n_1} e^{-\alpha\lambda'}\alpha^{n_2}\, dt_1 \ldots dt_{n_1+n_2}$$

az $n_1 = n_2$ esetben, és

$$e^{-\beta t_1}\beta dt_1 \ldots e^{-\alpha(T-t_{n_1+n_2})} = e^{-\beta l''}\beta^{n_1} e^{-\alpha\lambda''}\alpha^{n_2} dt_1 \ldots dt_{n_1+n_2}$$

az $n_1 = n_2 + 1$ esetben.

Itt az alábbi jelöléseket vezettük be:

$$\begin{aligned} l' &= t_1 + t_3 - t_2 + \ldots + T - t_{n_1+n_2}, \\ l'' &= t_1 + t_3 - t_2 + \ldots + t_{n_1+n_2} - t_{n_1+n_2-1}, \\ \lambda' &= t_2 - t_1 + \ldots + t_{n_1+n_2} - t_{n_1+n_2-1}, \\ \lambda'' &= t_2 - t_1 + \ldots + T - t_{n_1+n_2}. \end{aligned}$$

Bevezetve az összes A^* intervallumok L összesített hosszúságát, látjuk, hogy az első esetben $L = l'$, és a második esetben $L = l''$. Hasonló összefüggések érvényesek az A intervallumok Λ összesített hosszúságára. Így tehát mindkét esetben

$$\beta^{n_1} e^{-\beta L}\alpha^{n_2} e^{-\alpha\Lambda}\, dt_1 \ldots dt_{n_1+n_2}$$

alakban lehet a valószínűség-eloszlást felírni. A kísérletet N-szer, egymástól függetlenül elvégezve, a következő maximum likelihood becsléseket kapjuk:

$$\beta^* = \frac{\sum_1^N n_{1,i}}{\sum_1^N L_i}, \qquad \alpha^* = \frac{\sum_1^N n_{2,i}}{\sum_1^N \Lambda_i}.$$

A becslések nevezőjében szereplő mennyiségek azt az összesített időt adják meg, amelynek folyamán az éppen meglevő állapotnak (A^* az első esetben, A a második esetben) az ellenkező állapotba való átmenete várható, a számlálókban levő kifejezések pedig megadják azoknak az eseteknek a számát, amikor ilyen változás történt.

5. 8. Metrikus tranzitivitás — konzisztens becslések. Eddig azt az esetet vizsgáltuk, amikor a folyamatnak N független realizációja volt megfigyelhető. Kimutattuk, hogy ekkor a maximum likelihood módszer segítségével olyan becslést lehet szerkeszteni, amely aszimptotikusan konzisztens és asszimptotikusan efficiens, ha N végtelenhez tart. Abban a fontos esetben, amikor a folyamat stacionárius, azt lehetne remélni, hogy a maximum likelihood módszerrel a folyamat egyetlen T hosszúságú realizációjából is kaphatunk ilyen becslést, feltéve hogy T a végtelenhez tart. Ez a gondolat közelfekvő, mert ha T nagy, akkor a $(0, T)$ intervallumot fel lehet osztani nagy számú I_n intervallumra, amelyeket más I_n' intervallumok választanak el egymástól, éspedig úgy, hogy az utóbbiak összes hossza elhanyagolhatóan kicsi az egész $(0, T)$ intervallum hosszához képest, de mindegyik I_n' intervallum mégis elég hosszú ahhoz, hogy két különböző I_n intervallumon a folyamat értékei egymástól gyakorlatilag függetlenek legyenek. Nyilvánvaló azonban, hogy ennek az állításnak az érvényessége valamilyen járulékos feltételhez van kötve, amely biztosítja a folyamat értékeinek az aszimptotikus függetlenségét olyan időpontokban, amelyeket elég hosszú időköz választ el egymástól. Ilyen feltétel szükséges voltát bizonyítja a következő egyszerű példa.

Legyen $y(t)$ az előzőkben már többször tárgyalt normális sztochasztikus folyamat nulla átlaggal és adott $\varrho(s, t)$ korrelációs függvénnyel, és legyen x egy normális eloszlású valószínűségi változó, amely független $y(t)$-től minden t-re. Legyen x átlaga 0 és szórása σ. A megfigyelt folyamat legyen

$$x(t) = m + x + y(t); \quad 0 = t = T,$$

ahol m egy ismeretlen valós paraméter. Az előzőkben kimutattuk, hogy az m paraméter maximum likelihood becslésének, m^*-nak a szórása

$$\begin{cases} \mathbf{D}^2 m_T^* = \inf \int_0^T \int_0^T r(s, t) f(s) f(t)\, ds\, dt, \\ \int_0^T f(s)\, ds = 1, \end{cases}$$

ahol $r(s, t)$ az $x(t)$ folyamat korrelációs függvénye, vagyis

$$r(s, t) = \varrho(s, t) + \sigma^2.$$

Ebből közvetlenül belátható, hogy

$$\mathbf{D}^2 m_T^* \geqq \sigma^2,$$

és így m_T nem lehet m konzisztens becslése $T \to \infty$ esetén. Ez az eredmény nyilván annak a következménye, hogy az $r(s, t)$ korrelációs függvény kifejezésében fellépő σ^2 tag miatt a korrelációs kapcsolat nagyon erős marad $|t-s|$ nagy értékei esetében is. Ezért, ha biztosítani akarjuk konzisztens becslés létezését, akkor a *folyamatra olyan feltételt kell szabni, amely kizárja a túl erős korrelációs kapcsolat lehetőségét* $|t-s|$ *nagy értékeinél.* E célra a következőkben a *metrikus tranzitivitás* feltételét fogjuk használni (lásd az 1. 4. szakaszt).

Hogy a vizsgálatot ne bonyolítsák lényegbe nem vágó nehézségek, a 4. 1. szakaszban tárgyalt egyszerű esettel fogunk foglalkozni: legyen adva két különböző egyszerű hipotézis a P_{α_1}, illetőleg P_{α_2} valószínűség-eloszlásoknak megfelelően, ahol $\alpha_1 \neq \alpha_2$. Bizonyítani fogjuk, hogy a metrikus tranzitivitás feltételének következményeképpen létezik az α paraméternek konzisztens becslése, ha a folyamat T megfigyelési ideje végtelenhez tart (illetve létezik konzisztens próba a két hipotézis összehasonlítására).

Tekintsük a mintatér összes véges dimenziójú I_n intervallumainak $\{I_n\}$ összességét, ahol n a dimenziószám. Ha minden $I \in \{I_n\}$ esetén

$$P_{\alpha_1}(I) = P_{\alpha_2}(I),$$

akkor az eloszlások ekvivalensek, és a feladat triviális. Az ellenkező esetben létezik olyan I intervallum, hogy

$$P_{\alpha_1}(I) \neq P_{\alpha_2}(I) \quad (\text{legyen pl. } P_{\alpha_1}(I) > P_{\alpha_2}(I)).$$

Az 5. 14. szakaszban ki fogjuk mutatni, hogy a metrikus tranzitivitás feltételének teljesülése esetén a $P(I)$ valószínűségre lehet szerkeszteni egy $\pi_T(I)$ konzisztens becslést. Legyen most $f(x)$ olyan valós függvény, hogy

$$f(x) = \begin{cases} \alpha_1, & \text{ha} \quad x \leqq \dfrac{P_{\alpha_1}(I) + P_{\alpha_2}(I)}{2} \\ \alpha_2, & \text{ha} \quad x > \dfrac{P_{\alpha_1}(I) + P_{\alpha_2}(I)}{2}. \end{cases}$$

Ekkor $f[\pi_T(I)]$ valóban az α paraméter konzisztens becslése.

5. 9. A maximum likelihood módszer (folytatás). Ahhoz, hogy a maximum likelihood módszerrel akkor is optimális tulajdonságú becslést állíthassunk elő, ha a folyamatnak csupán egyetlen, de nagyon hosszú idejű realizációja ismeretes, a metrikus tranzitivitáson kívül még egy feltételnek kell teljesülnie. Ez a feltétel nem a folyamat múltbeli és jövőbeli értékei közötti függőségi kapcsolat mértékét, hanem annak jellegét korlátozza. Tekintsük a folyamat értékeit a t időpontot követő időpontokban. Legyen ismeretes a folyamat realizációja az $a \leqq s \leqq b$ $(b < t)$ intervallumban; ekkor be lehet vezetni az $x(s)$ folyamat feltételes valószínűség-eloszlását $s \geqq t > b$ esetére. Ha létezik olyan $(b-a)$-nál kisebb T érték, amelyre ez a feltételes valószínűség csupán a folyamatnak a $b-T \leqq s \leqq b$ intervallumból vett értékeitől függ, akkor a folyamatot általánosított Markov-folyamatnak nevezzük. A folyamatoknak ehhez a típusához tartoznak a közönséges Markov-folyamatok, úgyszintén az olyan folyamatok, amelyeknél a feltételes valószínűségeloszlást teljesen meghatározza magának a folyamatnak az értéke és az összes deriváltak értéke (adott rendű deriváltig) az utolsó megfigyelt időpontban. Diszkrét idő esetén az általánosított Markov-folyamat a közönséges, illetve a véges rendű Markov-lánc.

A továbbiakban feltesszük, hogy a feltételes eloszlásokat úgy lehet megadni, hogy majdnem biztosan valószínűségeloszlások legyenek, és hogy a likelihood függvények az 5. 7. szakaszban megadottakkal analóg feltételeknek tegyenek eleget (itt is csak a reguláris esettel foglalkozunk). Figyeljük meg a folyamatot az $(0, NT)$ időintervallumban, ahol N pozitív egész szám, és jelöljük a $((\nu-1)T, \nu T)$ intervallumnak megfelelő realizációt ω_ν-vel, $\nu=1, 2, \ldots, N$.

Használjuk minden ν esetén ugyanazokat a koordinátákat, és legyen Ω_ν az ω_ν realizáció mintatere.

Tekintsünk egy tetszőleges $A\subset\Omega_N$ halmazt és egy másik $S\subset\Omega_1\times\ldots\times\Omega_N$ halmazt. Legyen $f(\omega_1, \ldots, \omega_N; \alpha)$ a likelihood függvény; ekkor magából a feltételes valószínűségeloszlás definíciójából adódik, hogy

$$\begin{aligned} P_\alpha(SA) &= \int_S P_\alpha\{\omega_N\in A|\omega_1, \ldots, \omega_{N-1}\}\, dP_\alpha(\omega_1, \ldots, \omega_{N-1}) = \\ &= \int_S P_\alpha\{\omega_N\in A|\omega_{N-1}\}\, dP_\alpha(\omega_1, \ldots, \omega_{N-1}) = \\ &= \int_S P_\alpha\{\omega_N\in A|\omega_{N-1}\} f(\omega_1, \ldots, \omega_{N-1}; \alpha)\, dP_0(\omega_1, \ldots, \omega_{N-1}) = \\ &= \int_{SA} f(\omega_1, \ldots, \omega_N; \alpha)\, dP_0(\omega_1, \ldots, \omega_N). \end{aligned}$$

Az

$$\begin{aligned} \int_A f(\omega_1, \ldots, \omega_N; \alpha)\, dP_0(\omega_N|\omega_1, \ldots, \omega_{N-1}) &= \int_A f(\omega_1, \ldots, \omega_N; \alpha)\, dP_0(\omega_N|\omega_{N-1}) = \\ &= g(\omega_1, \ldots, \omega_{N-1}, A; \alpha) \end{aligned}$$

jelöléssel azt kapjuk, hogy (lásd Doob [2])

$$\begin{aligned} \int_S P_\alpha\{\omega_N\in A\,|\,\omega_{N-1}\} f(\omega_1, \ldots, \omega_{N-1}; \alpha)\, dP_0(\omega_1, \ldots, \omega_{N-1}) &= \\ = \int_S g(\omega_1, \ldots, \omega_{N-1}, A; \alpha)\, dP_0(\omega_1, \ldots, \omega_{N-1}) \end{aligned}$$

minden $S\subset\Omega_1\times\ldots\times\Omega_{N-1}$ halmazra. Így tehát majdnem biztosan

$$P_\alpha(\omega_N\in A\,|\,\omega_{N-1}) = \int_A \frac{f(\omega_1, \ldots, \omega_N; \alpha)}{f(\omega_1, \ldots, \omega_{N-1}; \alpha)}\, dP_0(\omega_N\,|\,\omega_{N-1}),$$

amiből következik, hogy az

$$\frac{f(\omega_1, \ldots, \omega_N; \alpha)}{f(\omega_1, \ldots, \omega_{N-1}; \alpha)}$$

viszony majdnem biztosan független az $\omega_1, \omega_2, \ldots, \omega_{N-2}$ realizációktól. A nevező itt 1 valószínűséggel különbözik nullától, mert feltételeztük a reguláris esetet. Használjuk most fel, hogy

$$f(\omega_1, \ldots, \omega_N; \alpha) = f(\omega_1; \alpha)\frac{f(\omega_1, \omega_2; \alpha)}{f(\omega_1; \alpha)} \cdots \frac{f(\omega_1, \ldots, \omega_N; \alpha)}{f(\omega_1, \ldots, \omega_{N-1}; \alpha)}.$$

Ugyanúgy, mint az előzőkben, ki lehet mutatni, hogy a jobb oldali hányadosok mindegyike csupán a benne előforduló utolsó két ω_i szimbólumtól függ. A legutóbbi egyenletet tehát az alábbi alakba lehet írni:

$$f(\omega_1, \ldots, \omega_N; \alpha) = f_1(\omega_1; \alpha) f_2(\omega_2|\omega_1; \alpha) \ldots f_N(\omega_N|\omega_{N-1}; \alpha).$$

Az f-ek indexeit itt el lehet hagyni, mert a folyamat stacionárius:

$$f(\omega_1, \ldots, \omega_N; \alpha) = f(\omega_1; \alpha) f(\omega_2|\omega_1; \alpha) \ldots f(\omega_N|\omega_{N-1}; \alpha).$$

Így tehát

$$\frac{1}{N}\frac{\partial \log f(\omega_1, \ldots, \omega_N; \alpha)}{\partial \alpha} = \frac{1}{N}\frac{\partial \log f(\omega_1; \alpha)}{\partial \alpha} + \frac{1}{N}\sum_2^N \frac{\partial \log f(\omega_\nu|\omega_{\nu-1}; \alpha)}{\partial \alpha},$$

és most a CRAMÉR [4] könyv 501—503. oldalán leírt módszer alkalmazásával a maximum likelihood egyenletet

$$B_0 + B_1(\alpha - \alpha_0) + \tfrac{1}{2}\theta B_2(\alpha - \alpha_0)^2 = 0$$

alakban kapjuk, ahol

$$B_0 = \frac{1}{N}\left(\frac{\partial \log f(\omega_1; \alpha)}{\partial \alpha}\right)_0 + \frac{1}{N}\sum_2^N \left(\frac{\partial \log f(\omega_\nu|\omega_{\nu-1}; \alpha)}{\partial \alpha}\right)_0$$

$$B_1 = \frac{1}{N}\left(\frac{\partial^2 \log f(\omega_1; \alpha)}{\partial \alpha^2}\right)_0 + \frac{1}{N}\sum_2^N \left(\frac{\partial^2 \log f(\omega_\nu|\omega_{\nu-1}; \alpha)}{\partial \alpha^2}\right)_0$$

$$B_2 = \frac{1}{N} H(\omega_1) + \frac{1}{N}\sum_2^N H(\omega_\nu|\omega_{-1}).$$

A metrikusan tranzitív folyamatokra $N \to \infty$ esetén érvényes ergod tétel értelmében ezek a kifejezések valószínűségben az átlagértékeikhez tartanak. Azonban

$$\mathbf{E}_0\left(\frac{\partial \log f(\omega_\nu|\omega_{\nu-1}; \alpha)}{\partial \alpha}\right)_0 = \mathbf{E}_0\left(\frac{\partial \log f(\omega_1, \ldots, \omega_\nu; \alpha)}{\partial \alpha}\right)_0 -$$

$$-\mathbf{E}_0\left(\frac{\partial \log f(\omega_1, \ldots, \omega_{\nu-1}; \alpha)}{\partial \alpha}\right)_0 = 0.$$

Az

$$\mathbf{E}_0\left(\frac{\partial^2 \log f(\omega_2|\omega_1; \alpha)}{\partial \alpha^2}\right)_0 = k$$

jelöléssel (ahol feltehetjük, hogy $k \neq 0$, mert különben ez a triviális eset volna) azt kapjuk, hogy

$$\mathbf{E}_0\left(\frac{\partial \log f(\omega_1, \ldots, \omega_N; \alpha)}{\partial \alpha}\right)_0^2 = -\mathbf{E}_0\left(\frac{\partial^2 \log f(\omega_1, \ldots, \omega_N; \alpha)}{\partial \alpha^2}\right)_0 =$$

$$= -\mathbf{E}_0\left(\frac{\partial^2 \log f(\omega_1; \alpha)}{\partial \alpha^2}\right)_0 + (N-1)k.$$

Ebből következik, hogy $N \to \infty$ esetén

$$B_0 \to 0, \quad B_1 \to -k, \quad B_2 \to M < \infty$$

(valószínűségben való konvergenciával). Most ugyanolyan módon, mint CRAMÉR [4]-ben, ki lehet mutatni, hogy létezik konzisztens maximum likelihood becslés, és hogy

$$\sqrt{Nk}(\alpha^* - \alpha) = \frac{\sqrt{\frac{N}{k}}\, B_0}{u_N},$$

ahol $N \to \infty$ esetén u_N valószínűségben egyhez tart. B_0 átlaga azonban 0 és szórása

$$-\frac{1}{N^2}\mathbf{E}_0\left(\frac{\partial^2 \log f(\omega_1; \alpha)}{\partial \alpha^2}\right)_0 + \frac{N-1}{N^2}k.$$

Tehát bebizonyítottuk, hogy az adott esetben *a maximum likelihood becslés konzisztens és aszimptotikusan efficiens* az aszimptotikus efficiencia WALD [1] által definiált értelmében.

5. 10. A metrikus tranzitivitás feltételei. A metrikus tranzitivitás fogalma nagyon fontos a stacionárius sztochasztikus folyamatok becsléselméletének a felépítésében. Ezzel kapcsolatban hasznosnak bizonyulnak DOOB [2] eredményei, amelyek a Markov-folyamatok esetére vonatkoznak. Itt a metrikus tranzitivitásra két másik feltételt fogunk adni.

TÉTEL: *Egy folytonos $r(t)$ korrelációs függvényű stacionárius normális folyamat metrikus tranzitivitásának szükséges és elegendő feltétele, hogy a folyamat spektrál eloszlásfüggvénye folytonos legyen.*

A tétel bizonyítására DOOB [1] és ITÔ [1] dolgozataiban található gondolatokat alkalmazunk. Feltesszük, hogy a folyamat D-integrálható, és legyen

$$r(t) = \int_{-\infty}^{\infty} e^{it\lambda}\, dF(\lambda), \quad r(0) = 1,$$

ahol az $F(\lambda)$ spektrál eloszlásfüggvényről feltettük, hogy folytonos. Ha a folyamat nem volna metrikusan tranzitív, akkor léteznék egy olyan minden T_t eltolással szemben invariáns S halmaz, amelyre $P(S) = \varrho$, $0 < \varrho < 1$. Approximáljuk az S halmazt véges dimenziójú intervallumok véges I összegével oly módon, hogy kielégüljenek a

$$P(I) < \varrho + \varepsilon$$
$$P(SI^*) < \varepsilon$$

feltételek, ahol ε egy előre megadott pozitív szám. Nyilvánvalóan feltehetjük, hogy az I összeget képező intervallumok mindegyikének véges hosszúságúak az oldalai. Legyen $T_t I = I_t$. Legyenek $\tau_1 + t, \ldots, \tau_n + t$ olyan időpontok, hogy minden az I_t-hez tartozó véges dimenziójú intervallumot a folyamatnak ezekhez az időpontokhoz tartozó értékei határoznak meg, és vezessük be az

$$x_i = x(\tau_i); \qquad i = 1, 2, \ldots, n,$$
$$x_{n+j} = x(\tau_j + t); \qquad j = 1, 2, \ldots, n$$

jelöléseket. Ezeknek a valószínűségi változóknak az eloszlása nyilvánvalóan $2n$-dimenziós normális eloszlás, amelyet a második momentumok matrixa, $\Lambda(t)$ határoz meg:

$$\Lambda(t) = \left[\begin{array}{ccc|ccc} 1 & r(\tau_1 - \tau_2) & \ldots & r(t) & r(\tau_1 - \tau_2 - t) & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ r(\tau_n - \tau_1) & r(\tau_n - \tau_2) & \ldots 1 & & & \\ \hline r(t) & r(\tau_1 + t - \tau_2) & \ldots & 1 \ldots & & \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ r(\tau_n + t - \tau_1) & r(\tau_n + t - \tau_2) & \ldots & \ldots & & \ldots 1 \end{array}\right] =$$

$$= \begin{bmatrix} A & B(t) \\ B'(t) & A \end{bmatrix},$$

ahol az A matrix nem függ t-től. A spektrum folytonosságából — amint ez könnyen belátható — következik, hogy a második momentumok semmilyen matrixa nem lehet szinguláris. Eléggé nagy t értékek esetén

$$P(II_t) = \frac{[\Lambda(t)]^{-\frac{1}{2}}}{(2\pi)^n} \underset{(x_1, \ldots x_n) \in I}{\int \ldots \int} \underset{(x_{n+1}, \ldots x_{2n}) \in I_t}{\int \ldots \int} e^{-\frac{1}{2} Q(x)} dx_1 \ldots dx_{2n},$$

ahol $Q(x)$ az $x_1, x_2, \ldots, x_{2n}$ változók olyan kvadratikus formája, amelynek a matrixa $\Lambda(t)$ inverze. Vezessük be az összes $\tau_i - \tau_j$ alakú számok részére a $t_1, t_2, \ldots, t_N$ jelölést. Ekkor az abszolút integrálhatóság következtében

$$\frac{1}{2T} \int_{-T}^{T} \sum_1^N |r(t_i + t)|^2 dt = \frac{1}{2T} \int_{-T}^{T} \sum_1^N \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(t_i + t)\lambda - i(t_i + t)\mu} dF(\lambda) dF(\mu) dt =$$

$$= \sum_1^N \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{e^{i(t_i + T)(\lambda - \mu)} - e^{i(t_i - T)(\lambda - \mu)}}{2Ti(\lambda - \mu)} dF(\lambda) dF(\mu).$$

Ismert meggondolások alapján (lásd pl. HOPF [1] p. 16)

$$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \sum_1^N |r(t_i + t)|^2 dt = 0,$$

tehát

$$\lim_{T\to\infty} \sum_{1}^{N} |r(t_i+t)|^2 = 0.$$

Következésképpen létezik olyan, $\nu\to\infty$ esetén végtelenhez tartó t_ν sorozat, hogy $B(t_\nu)\to 0$ és

$$A(t_\nu)\to\begin{Bmatrix} A & 0 \\ 0 & A \end{Bmatrix}.$$

A korlátos konvergenciára vonatkozó LEBESGUE-tétel alkalmazásával kapjuk, hogy

$$P(II_t)\to P(I)^2.$$

Így tehát ν nagy értékeire

$$(\varrho+\varepsilon)^2 > P(II_t) \geqq P(SII_{t\nu}) \geqq P(S)-P(SI^*)-P(SI^*_{t\nu}) > \varrho - 2\varepsilon.$$

Ez az egyenlőtlenség azonban csak úgy lehet érvényes tetszőlegesen kicsi ε-ra, ha vagy $\varrho=1$, vagy $\varrho=0$, ami viszont kiinduló feltevésünknek mond ellent. Ezzel bebizonyítottuk, hogy feltételünk elégséges.

Hogy bizonyíthassuk a feltétel szükséges voltát is, tekintsük az $x^2(t)$ folyamatot. Ez a folyamat véges szórású, mert a normális eloszlásnak léteznek véges negyedik momentumai, és könnyen belátható, hogy

$$\varrho(t) = \mathbf{E}[x^2(s)-\mathbf{E}x^2(s)][x^2(s+t)-\mathbf{E}x^2(s+t)] = 2r^2(t).$$

Tudjuk, hogy a

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x^2(t)\,dt = y$$

határérték majdnem biztosan létezik és szórása

$$\mathbf{D}^2(y) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\varrho(t)\,dt = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} 2r^2(t)\,dt.$$

Ez az utóbbi kifejezés azonban CRAMÉR [1] szerint a

$$\mathbf{D}^2 y = 2\sum_{1}^{N}\Delta_\nu F$$

alakra hozható, ahol $\Delta_\nu F$ az $F(\lambda)$ spektrál eloszlásfüggvény ugrásait jelenti a szakadási pontokon. Az $x(t)$ folyamat metrikus tranzitivitásának tehát szükséges feltétele, hogy $F(\lambda)$ folytonos legyen. Ezzel tételünk bizonyítása teljessé vált.

Ha nem feltételezzük, hogy folyamatunk normális, akkor a korrelációs függvény ismerete nem elégséges a folyamat valószínűségeloszlásainak a meghatározására. Ennek ellenére ebben az esetben is lehet a metrikus tranzitivitás számára egy kritériumot megadni, amely lényegében a fenti kritérium általánosítása.

Az $x(t)$ folyamatra azt mondjuk, hogy *keverő*, ha minden A, B mérhető halmazra érvényes a

$$\lim_{t\to\infty} P(AB_t) = P(A)\cdot P(B)$$

összefüggés, ahol $B_t = T_t B$.

Ismeretes, hogy a keverő tulajdonság maga után vonja a metrikus tranzitivitást, de ennek a fordítottja nem érvényes (lásd HOPF [1]). A fentebb bizonyított tétel szerint a folyamat metrikusan tranzitív, ha nincs pont-spektruma. Ebben az esetben a spektrál eloszlásfüggvény egy abszolút folytonos részből és egy szinguláris részből lehet összetéve. ITÔ kimutatta, hogy ha a szinguláris összetevő is hiányzik, akkor normális folyamat esetén a folyamat keverő. De a normális folyamat esetében már láttuk, hogy a folyamat metrikusan tranzitív már akkor is, ha a spektrumának van szinguláris összetevője. Megjegyezzük azonban, hogy ilyen esetben léteznie kell olyan a végtelenhez tartó t_ν sorozatnak, hogy $\lim_{\nu\to\infty} r(t_\nu)=0$. Ez a körülmény természetes módon rávezet a keverő folyamatok alábbi gyengített fogalmára:

Az $x(t)$ folyamat *részben keverő*, ha minden mérhető A halmazhoz létezik olyan $t_\nu \to \infty$ sorozat (amely A-tól függ), hogy

$$\lim_{\nu\to\infty} P(AA_{t_\nu}) = P(A)^2.$$

TÉTEL: *A stacionárius sztochasztikus folyamat metrikus tranzitivitásának szükséges és elegendő feltétele, hogy a folyamat részben keverő legyen.*

Ha az $x(t)$ részben keverő folyamat, akkor be lehet bizonyítani — pontosan ugyanolyan módon, mint fent —, hogy metrikusan tranzitív. Így tehát itt csak a feltétel szükséges voltát kell külön bizonyítanunk. Válasszunk egy tetszőleges A halmazt és legyen $c(t, \omega)$ az A_t halmaz karakterisztikus függvénye, vagyis az a függvény, amelynek értéke 1, ha $\omega \in A_t$, és 0, ha $\omega \notin A_t$. Ekkor $c(t, \omega)$ stacionárius sztochasztikus folyamat, amelynek korrelációs függvénye

$$\begin{aligned} r_A(t) &= Ec(s, \omega)\, c(s+t, \omega) - Ec(s, \omega)\, Ec(s+t, \omega) = \\ &= P(A_s, A_{s+t}) - P(A_s)\, P(A_{s+t}) = \\ &= P(AA_t) - P(A)^2. \end{aligned}$$

Az ergodikus tulajdonság értelmében

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} r_A(t)\,dt = \lim_{T\to\infty} \frac{1}{T} \int_{-0}^{T} r_A(t)\,dt = 0.$$

Mivel a folyamatot D-integrálhatónak és D-mérhetőnek tételeztük fel, $P(AA_t)$ folytonos függvénye t-nek (lásd pl. HOPF [1]). Az $r_A(t)$ függvénynek vagy végtelen sok zérushelye van és ezeknek a sorozata végtelenhez tart, vagy pedig a függvény értékeinek az előjele ugyanaz marad minden $t > t_0$ időpontra. Mindkét esetben lehet olyan $t_\nu(A)$ sorozatot találni, amelyre

$$\lim_{\nu\to\infty} P(AA_{t_\nu}) = P(A)^2,$$

és ez bizonyítja tételünket.

Megjegyzés: A keverő tulajdonság definíciójában egy bizonyos feltételt szabtunk minden mérhető A halmazra. Az alkalmazások számára ez a megfogalmazás nem nagyon célszerű. Megmutatjuk, hogy teljesen elegendő, ha csupán véges dimenziójú intervallumokat veszünk tekintetbe. Tegyük fel, hogy a

$$\lim_{t\to\infty} P(IJ_t) = P(I)\,P(J)$$

feltétel bármely véges dimenziójú I, J intervallumpárra teljesül. Ha most A tetszőleges mérhető halmaz, akkor approximálni lehet diszjunkt intervallumok $\Sigma=\Sigma I$ véges összegével úgy, hogy

$$P(A^*\Sigma)+P(A\Sigma^*)<\varepsilon.$$

Ekkor

$$|P(A)-P(\Sigma)|<\varepsilon,$$

és

$$P\{AA_t(\Sigma\Sigma_t)^*\}+P\{(AA_t)^*\Sigma\Sigma_t\}\leqq$$
$$\leqq P(AA_t\Sigma^*)+P(A^*\Sigma\Sigma_t)+P(A_t^*\Sigma\Sigma_t)+P(AA_t\Sigma_t^*)\leqq$$
$$\leqq P(A\Sigma^*)+P(A^*\Sigma)+P(A_t\Sigma_t^*)+P(A_t^*\Sigma_t)<2\varepsilon.$$

Azonban

$$P(\sum\sum_t)=P(\sum_\nu I_\nu\sum_\mu I_\mu^t)=\sum_{\nu,\mu}P(I_\nu I_\mu^t).$$

Itt a jobb oldal határértéke $t\to\infty$ esetén

$$\sum_{\nu,\mu}P(I_\nu)\,P(I_\mu),$$

tehát

$$\lim_{t\to\infty}P(\Sigma\Sigma_t)=P(\Sigma)^2.$$

Így

$$P(AA_t)-P(A)^2\leqq$$
$$\leqq|P(AA_t)-P(\Sigma\Sigma_t)|+|P(\Sigma\Sigma_t)-P(\Sigma)^2|+$$
$$+|P(\Sigma)^2-P(A)^2|\leqq 4\varepsilon+|P(\Sigma\Sigma_t)-P(\Sigma)^2|,$$

ezért végül

$$\lim_{t\to\infty}P(AA_t)=P(A)^2.$$

5. 11. Alkalmazások. Alkalmazzuk a maximum likelihood módszert két speciális, egyszerű típusú stacionárius sztochasztikus folyamatra.

Legyen $x(t)$ egy stacionárius normális Markov-folyamat m átlaggal és $e^{-\beta|t-s|}$ korrelációs függvénnyel. Tudjuk, hogy itt a spektrális függvény abszolút folytonos, tehát a folyamat metrikusan tranzitív. A likelihood függvény

$$f(\omega,m)=e^{-\frac{m^2}{2}\left(1+\frac{\beta T}{2}\right)+\frac{m}{2}\left\{x(0)+x(T)+\beta\int\limits_0^T x(t)dt\right\}}$$

és így a maximum likelihood becslés

$$m^*=\frac{x(0)+x(T)+\beta\int\limits_0^T x(t)dt}{2+\beta T},$$

amely tehát az m paraméter konzisztens és aszimptotikusan efficiens becslése. Az adott esetben ez az eredmény nem mond újat, ugyanis az 5. 6. szakaszban láttuk, hogy az m^* becslés már véges T esetén is efficiens.

Tekintsük most a 4. 9. szakaszban vizsgált folyamatot. Ez is stacionárius Markov-folyamat, és korrelációs függvénye szintén $e^{-\beta|t-s|}$, de mivel a folyamat nem normális, nem lehet az előbbivel azonos módon kimutatni, hogy metrikusan tranzitív.

Ehelyett tekintsük azt az I intervallumot, amelyet a folyamat $t_1, t_2, \ldots, t_n$ időpontokban vett értékei határoznak meg, és tekintsünk egy másik J intervallumot, amelyet a $t_1', t_2', \ldots, t_m'$ pontokban vett értékei határoznak meg. Legyen t egy nagy pozitív szám. Ekkor

$$P(IJ_t) = P_0(t)\, P(IJ_t|0) + P_1(t)\, P(IJ_t|1),$$

ahol a 0 index azt a feltételt jelenti, hogy a $(t_n, t_1' + t)$ időintervallumban a folyamat értékeiben egyetlen változás sem következett be, az 1 index pedig annak az alternatív feltételnek felel meg, hogy legalább egy változás történt. Ekkor

$$P_0(t) = e^{-\beta(t_1'+t-t_n)} \to 0, \quad \text{ha } t \to \infty.$$

Azonban

$$P(IJ_t|1) = P(I)\, P(J_t),$$

amiből az előző szakasz végén közölt megjegyzés értelmében közvetlenül adódik, hogy a folyamat metrikusan tranzitív. A maximum likelihood becslés itt igen egyszerű alakú:

$$m^* = \frac{1}{m+1} \sum_0^n x_\nu,$$

amit úgy lehet felfogni, mint a folyamat értékeinek a realizációtól függő súlyfüggvénnyel vett integrálját. A becslés torzítatlan, mert

$$\mathbf{E} m^* = \sum_0^\infty P_\nu \mathbf{E}[m^*|\nu] = m \sum_0^\infty P_\nu = m.$$

A szórást is könnyen lehet kiszámítani:

$$\mathbf{E}(m^* - m)^2 = \sum_0^\infty P_\nu \mathbf{E}[(m^*-m)^2|\nu] = \sum_0^\infty \frac{(\beta T)^\nu}{\nu!} e^{-\beta T} \frac{1}{\nu+1} = \frac{1-e^{-\beta T}}{\beta T}.$$

Mivel azonban

$$\mathbf{E}\left(\frac{\partial \log f(\omega, m)}{\partial m}\right)^2 = \sum_0^\infty P_\nu \mathbf{E}\left\{\left[\sum_0^\nu x_n - (\nu+1)m\right]^2 = e^{-\beta T} \sum_1^\infty \frac{(\beta T)^\nu}{\nu!}(\nu+1) = 1 + \beta T,\right.$$

tehát becslésünk efficienciája

$$e(m^*) = \frac{\beta T}{(1+\beta T)(1-e^{-\beta T})}.$$

$T=0$ esetén ez az efficiencia 1, és ha T nő, akkor $e(m^*)$ először csökken, de nagy T értékek esetén megint 1-hez tart. Ha viszont az m paraméter legjobb lineáris becslését kívánjuk alkalmazni, akkor az alábbi értéket kapjuk:

$$m_L^* = \frac{x(0) + x(T) + \beta \int_0^T x_0(t)\,dt}{2 + \beta T}.$$

Ennek a becslésnek a szórása

$$\mathrm{D}^2 m_L^* = \frac{2}{2+\beta T},$$

az efficienciát tehát ekkor az

$$e(m_L^*) = \frac{1+\frac{\beta T}{2}}{1+\beta T}$$

képlet adja meg. Látjuk, hogy $T=0$ esetén $e(m_L^*)=e(m^*)=1$, a kétféle becslésnek ezért egybe kell esnie. A $T=0$ esetben csakugyan 1 valószínűséggel $n=0$, tehát $m^*=x_0$, és ugyancsak az $m_L^*=x_0$ érték adódik közvetlenül az m_L^* kifejezéséből is $T=0$ esetére. Ha viszont $T\to\infty$, akkor nyilvánvalóan $e(m_L^*)\to 1/2$.

A számtani közép szerinti becslés

$$m_E^* = \frac{1}{T}\int_0^T x(t)dt.$$

S zórására T nagy értékei esetén aszimptotikusan a

$$D^2 m_E^* \sim \frac{2}{\beta T}$$

összefüggést kapjuk, következésképpen az aszimptotikus efficiencia

$$\lim_{T\to\infty} e(m_E^*) = \frac{1}{2}.$$

Így tehát a számtani közép szerinti becslés esetén, sőt a legjobb lineáris becslésnél is elveszítjük az efficiencia 50%-át, nagy T értékek esetén.

5. 12. Valószínűségi eloszlás a becslések egy típusának az esetére. Tekintsünk pontfolyamatokat hozzárendelt valószínűségi változókkal; ezeknél gyakran jutunk olyan becslésekhez, amelyek $\sum_1^n x_i$ alakú kifejezéseket foglalnak magukban. Vizsgáljuk ilyen kifejezések aszimptotikus eloszlását $T\to\infty$ esetére. Nem kívánunk a legáltalánosabb esettel foglalkozni, ezért feltesszük, hogy az x_i valószínűségi változók függetlenek, átlaguk nulla és szórásuk azonos σ értékű. Feltesszük továbbá, hogy $n\to\infty$, ha $T\to\infty$ (valószínűségben vett konvergencia értelmében). Ekkor a

$$\frac{\sum_1^n x_i}{\sigma\sqrt{n}}$$

mennyiség $T\to\infty$ esetén aszimptotikusan normális eloszlású (0, 1) paraméterekkel, mert

$$P\left\{\frac{\sum_1^n x_i}{\sigma\sqrt{n}} \leq a\right\} = \sum_1^\infty P_T(\nu) P\left\{\frac{\sum_1^n x_i}{\sigma\sqrt{n}} \leq a_a \,|\, \nu = n\right\}$$

és a centrális határeloszlástételből következik, hogy

$$P\left\{\frac{\sum_{1}^{n} x_i}{\sigma\sqrt{n}} \leqq a \mid \nu = n\right\} \to \Phi(a); \quad \nu \to \infty,$$

ahol $\Phi(x)$ a normális eloszlásfüggvény. A tétel szerint ugyanis tetszőleges $\varepsilon > 0$ esetén létezik olyan $\nu(\varepsilon)$ szám, hogy

$$\left| P\left\{\frac{\sum_{1}^{n} x_i}{\sigma\sqrt{n}} \leqq a \mid \nu = n\right\} - \Phi(a)\right| < \varepsilon, \quad \text{ha} \quad \nu > \nu(\varepsilon).$$

Válasszuk most T értékét olyan nagyra, hogy $\sum_{1}^{\nu(\varepsilon)} P_T(\nu) < \varepsilon$. Ekkor

$$\left| P\left\{\frac{\sum_{1}^{n} x_i}{\sigma\sqrt{n}}\right\} - \Phi(a)\right| \leqq \sum_{1}^{\infty} P_T(\nu)\left| P\left\{\frac{\sum_{1}^{n} x_i}{\sigma\sqrt{n}} \leqq a \mid n = \nu\right\} - \Phi(a)\right| \leqq 2\varepsilon + \varepsilon \sum_{\nu(\varepsilon)+1}^{\infty} P_T(\nu) < 3\varepsilon,$$

ami bizonyítja állításunkat.

Ha a $\frac{\mathbf{D}n(T)}{\mathbf{E}n(T)}$ viszonyszám $T \to \infty$ esetén nullához tart, akkor a $\sum_{1}^{n} x_i$ összeg eloszlása aszimptotikusan normális $\{0, \sigma\sqrt{\mathbf{E}n(T)}\}$ paraméterekkel, mert

$$\frac{\sum_{1}^{n} x_i}{\sigma\sqrt{\mathbf{E}n(T)}} = \frac{\sum_{1}^{n} x_i}{\sigma\sqrt{n}}\sqrt{\frac{n}{\mathbf{E}n(T)}}.$$

Az $\frac{n}{\mathbf{E}n(T)}$ valószínűségi változó átlaga ugyanis 1, és szórása $\frac{\mathbf{D}n(T)}{\mathbf{E}n(T)}$, tehát valószínűségben 1-hez tart, ha $T \to \infty$. A CRAMÉR [4] könyv 20.6. szakaszának tétele szerint ebből közvetlenül a fenti eredmény adódik.

Végül $\frac{\sum_{1}^{n} x_i}{n}$ is aszimptotikusan normális, $\left\{0, \frac{\sigma}{\sqrt{\mathbf{E}n(T)}}\right\}$ paraméterekkel, amit az előzőhöz hasonló módon lehet belátni az

$$\sqrt{\mathbf{E}n(T)}\,\frac{\sum_{1}^{n} x_i}{n\sigma} = \frac{\sum_{1}^{n} x_i}{\sigma\sqrt{n}}\sqrt{\frac{\mathbf{E}n(T)}{n}}$$

egyenlet alapján.

5. 13. Becslések approximációja. A következőkben a becsléselmélet egy eredményét ismertetjük, amely a 4. 12. szakaszban tárgyalt becsléssel analóg. Ott a reguláris eset biztosítása érdekében feltételeztük, hogy minden P_α eloszlás abszolút folytonos P_0-ra nézve. Most azt is feltesszük, hogy ez a folytonosság egyenletes $\alpha \in A$ esetén. Ekkor, ha $\alpha^*(\omega)$ az α paraméter becslése, és

$$\mathbf{E}_\alpha[\alpha^*(\omega)]^2 = v(\alpha),$$

— ahol a jobb oldalról feltesszük, hogy α-nak folytonos függvénye —, lehetséges az α^* becslés approximálása olyan $\alpha^*(x_1, \dots, x_n)$ becslésekkel, amelyek csak véges számú koordinátától függnek. Ez az approximáció egyenletes $\alpha \in A$-ra. Képezzük e célból a

$$t_N(\omega) = \begin{cases} \alpha^*(\omega), & \text{ha} \quad |\alpha^*(\omega)| < N \\ 0 & \text{ha} \quad |\alpha^*(\omega)| \geqq N \end{cases}$$

valószínűségi változót. Ekkor

$$\mathbf{E}_\alpha(t_N - \alpha^*)^2 = \int_\Omega (t_N - \alpha^*)^2 \, dP_\alpha = \int_{\alpha^* \geqq N} (\alpha^*)^2 \, dP_\alpha \to 0,$$

ha $N \to \infty$. Azonban

$$\left| \int_{\alpha^* < N} (\alpha^*)^2 \{f(\omega, \alpha_0) - f(\omega, \alpha)\} \, dP_0(\omega) \right| = N^2 \sqrt{\int_\Omega [f(\omega, \alpha_0) - f(\omega, \alpha)]^2 \, dP_0(\omega)} \,,$$

ahol az 5. 1. szakaszban az $f(\omega, \alpha)$ számára megszabott feltételeknek megfelelően a jobb oldal nullához tart, ha $\alpha \to \alpha_0$. Ezért tehát

$$\int_{|\alpha^*| \geqq N} [\alpha^*(\omega)]^2 f(\omega, \alpha) dP_0(\omega) = v(\alpha) - \int_{|\alpha^*| < N} [\alpha^*(\omega)]^2 f(\omega, \alpha) dP_0(\omega)$$

α-nak folytonos függvénye. DINI tétele szerint a fenti konvergencia egyenletes, vagyis tetszőleges $\varepsilon > 0$ esetén létezik olyan $N_0 = N_0(\varepsilon)$, hogy

$$\mathbf{E}_\alpha[t_N - \alpha^*]^2 < \varepsilon, \quad \text{ha} \quad N > N_0.$$

Tekintsük most az

$$\alpha_N^*(x_1, \dots, x_n) = \mathbf{E}_0[t_N | x_1, \dots, x_n]$$

valószínűségi változót. Ha n végtelenhez tart, akkor az $\alpha_N^*(x_1, \dots, x_n)$ mennyiség 1 valószínűséggel $t_N(\omega)$-hoz tart. Vezessük be az

$$\{|\alpha_N^*(x_1, \dots, x_n) - t_N(\omega)| < \varepsilon\} = \mathbf{E}_n \subset \Omega$$

halmazt. Ekkor

$$\mathbf{E}_\alpha[\alpha_N^*(x_1, \dots, x_n) - t_N(\omega)]^2 = \int_\Omega [\alpha_N^*(x_1, \dots, x_n) - t_N(\omega)]^2 \, dP_\alpha(\omega) \leqq$$

$$\leqq \varepsilon^2 P_\alpha(\mathbf{E}_n) + 4N^2 P_\alpha(\mathbf{E}_n^*).$$

Azonban $P_0(\mathbf{E}_n^*) \to 0$, ha $n \to \infty$, következőleg P-nak P_0-ra nézve egyenletes abszolút folytonossága miatt

$$\mathbf{E}_\alpha[\alpha_N^*(x_1, \dots, x_n) - t_N(\omega)]^2 < \delta, \quad \text{ha} \quad n > n_0(N, \delta).$$

A háromszög egyenlőtlenség felhasználásával ekkor megkapjuk a kívánt eredményt:

$$\mathbf{E}_\alpha[\alpha_N^*(x_1, \dots, x_n) - \alpha^*(\omega)]^2 < \varepsilon$$

minden $\alpha \in A$ esetén, ha N és n értékeit elég nagyra választjuk. Így tehát véges számú koordinátától függő olyan becslést kaptunk, amelynek átlaga és szórása tetszőleges pontossággal megközelítheti az $\alpha^*(\omega)$ becslés átlagát és szórását, éspedig egyenletesen $\alpha(\in A)$-ban.

Ebből következik, hogy ha csupán olyan becslésekre szorítkozunk, amelyek véges számú koordinátától függnek, és minden esetben kiválasztjuk ezek közül a legjobbat, (amelyet $\alpha^*(x_1, \ldots, x_n)$-nel jelölünk), akkor eléggé nagy n-re az $\alpha^*(x_1, \ldots, x_n)$ becslés gyakorlatilag éppen olyan jó lesz, mint bármelyik, a folyamat minden koordinátájától függő becslés.

5. 14. Függvények becslése. Eddig főleg azt az esetet vizsgáltuk, amikor az $x(t)$ valószínűségeloszlása néhány benne szereplő valós paraméter kivételével ismeretes volt. Más típusú problémával foglalkoztunk az 5. 2.—5. 5. szakaszokban, ahol azt tettük fel, hogy a folyamat valószínűségeloszlásairól semmi egyebet nem tudunk, mint a megfelelő korrelációs függvény alakját. Megint más jellegű feladattal állunk szemben, amikor *a folyamathoz tartozó valószínűségeloszlás egy ismeretlen függvénytől függ és a folyamat megfigyelése útján kapott adatokból kell becslést adni erre a függvényre.* A következőkben két példával illusztráljuk az ilyen feladatokat.

Legyen $x(t)$ egy valós, stacionárius, normális és D-mérhető folyamat nulla átlaggal és $r(t)$ korrelációs függvénnyel; ez utóbbit — mint eddig — folytonosnak feltételezzük. Az $r(t)$ függvényre keresünk becslést a folyamat realizációjának a $(0, T)$ időszakaszban történt megfigyelése alapján. Ha a folyamat metrikusan tranzitív — azaz ha spektrális függvénye folytonos (lásd az 5. 10. szakaszt) —, akkor lehetséges konzisztens becslést megadni $r(t)$-re. Ismeretes (lásd HOPF [1], p. 54—55), hogy majdnem biztosan t minden értékére

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T x(s)x(s-t)\,ds = r(-t) = r(t).$$

Mivel folyamatunk valós, a korrelációs függvénye páros, tehát elég csupán $t>0$ esetét vizsgálni. De

$$\frac{1}{T}\int_0^T x(s)x(s-t)\,ds - \frac{1}{T}\int_t^T x(s)x(s-t)\,ds = \frac{1}{T}\int_0^t x(s)x(s-t)\,ds,$$

ahol a jobb oldal $T\to\infty$ esetén majdnem biztosan nullához tart minden t-re. Ebből következik, hogy az

$$r_T^*(t) = \begin{cases} \dfrac{1}{T}\displaystyle\int_t^T x(s)x(s-t)\,ds, & \text{ha} \quad 0 \leqq t < T \\ 0, & \text{ha} \quad t \geqq T \end{cases}$$

kifejezés, amely csupán a folyamatnak a $0 \leqq t < T$ intervallumban megfigyelt értékeitől függ, $r(t)$-nek konzisztens becslése.

Második példaként vizsgáljunk olyan $x(t)$ folyamatot, amely szintén stacionárius, D-integrálható és metrikusan tranzitív. Legyen feladatunk az

$$F(a) = P\{x(t) \leqq a\}, \quad -\infty < a < \infty$$

eloszlásfüggvény konzisztens becslése. E célra segédeszközként bevezetjük az

$$e_t(\omega)=\begin{cases}1, & \text{ha} \quad x(t,\omega)\leqq a,\\ 0, & \text{ha} \quad x(t,\omega)>a\end{cases}$$

sztochasztikus folyamatot. Nyilvánvalóan t tetszőleges rögzített értékére $e_t(\omega)$ valószínűségi változó. Az $e_t(\omega)$ függvény mérhető és integrálható a $T\times\Omega$ szorzattéren, ahol T tetszőleges véges intervallum az időtengelyen. Így tehát 1-valószínűséggel

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} e_t(\omega)dt = Ee_t(\omega) = P\{x(t)\leqq a\} = F(a).$$

De $\int_{-T}^{T} e_t(\omega)dt$ annak a $(-T, T)$ intervallumhoz tartozó időszakasznak a hossza, amelynek folyamán $x(t)\leqq a$. Az

$$\frac{1}{2T}m\{x(t,\omega)\leqq a;\ |t|<T\}_t=F_T^*(a,\omega)$$

jelöléssel (ami megengedett, mert FUBINI tétele szerint az m mérték létezik) ekkor

$$\lim_{T\to\infty}F_T^*(a,\omega)=F(a).$$

Legyen $\{a_\nu;\ \nu=1,2,\ldots\}$ egy a valós tengelyen mindenütt sűrű valós számsorozat. A sorozat megszámlálhatósága miatt

$$\lim_{T\to\infty}F_T^*(a_\nu,\omega) = F(a_\nu)$$

majdnem biztosan minden ν-re. De $F_T^*(a,\omega)$ nem csökkenő függvénye a-nak. Ha a folytonossági pontja $F(x)$-nek és $a'_\nu\to a+0$, $a''_\nu\to a-0$, akkor

$$F_T^*(a''_\nu,\omega)\leqq F_T^*(a,\omega)\leqq F_T^*(a'_\nu,\omega),$$

ahonnan következik, hogy

$$\lim_{T\to\infty}F_T^*(a,\omega)=F(a)$$

majdnem biztosan $F(x)$ minden folytonossági pontjában. Így tehát F_T^* konzisztens becslése F-nek. Ugyanilyen módon lehet konzisztens becslést szerkeszteni a folyamat bármelyik többdimenziós eloszlásfüggvénye számára.

Könnyen belátható, hogy a második példánkban a becslés torzítatlan, és hogy az első példában egyszerű módon torzítatlanná lehet tenni a $\dfrac{T}{T-t}$ tényezővel való szorzás útján. Kívánatosnak véljük az efficienciához hasonló valamilyen fogalom bevezetését ilyen természetű feladatok esetére és ennek alapján vizsgálni a függvények becsléseinek a sajátságait.

Mielőtt a becsléselmélet fejezetét lezárnók, utalni kívánunk arra, hogy a konfidenciatartományok definíciója a klasszikus elméletből majdnem szóról szóra átvihető a sztochasztikus folyamatok becsléselméletébe.

Fordította: *dr. Korodi Albert*,
a műszaki tudományok kandidátusa.