NORMALIZÁLT EMPIRIKUS MÉRTÉKEK SZORZATA SZERINTI INTEGRÁLOK BECSLÉSÉRŐL

Major Péter
Magyar Tudományos Akadémia Rényi Alfréd Matematikai Kutató Intézete

Az előadásban a következő probléma vizsgálatával foglalkoztam. Legyen adva független, egyforma eloszlású $\xi_1,\dots,\xi_n$ valószínűségi változók sorozata, amelyek értékeiket valamilyen $(X,\mathcal{X})$ mértéktéren veszik fel. Jelölje μ_n e valószínűségi változók által elkészített empirikus mértéket, azaz legyen

$$\mu_n(A)=\frac{1}{n}\#\{j\colon\ \xi_j\in A,\ 1\le j\le n\},\quad A\in\mathcal{X}.$$

Tekintsük a μ_n empirikus mérték, ν_n, $\nu_n(A)=\sqrt{n}(\mu_n(A)-\mu(A))$, $A\in\mathcal{X}$, normalizáltját, valamint vegyünk egy k-változós $f(x_1,\dots,x_k)$ k-változós függvényt az $(X^k,\mathcal{X}^k)$ szorzattéren. Ezután definiáljuk az f függvény (véletlen) integrálját a normalizált ν_n mérték szerint, azaz tekintsük az alábbi integrált:

$$J_{n,k}(f)=\frac{1}{k!}\int' f(u_1,\dots,u_k)\sqrt{n}(\mu_n(\,du_1)-\mu(\,du_1))\dots\sqrt{n}(\mu_n(\,du_k)-\mu(\,du_k)), \tag{1}$$

ahol a $\int'$ azt jelenti, hogy az $u_j=u_l,\ 1\le j<l\le k,$ átlókat kihagyjuk az integrálási tartományból.

Célunk az, hogy jó becslést adjunk a $P\left(|J_{n,k}(f)|>x\right)$ valószínűségre minden $x>0$ számra. (Megjegyzem, hogy az átlók kihagyása az integrálási tartományból természetes, ugyanis az érdekes alkalmazásokban ilyen tipusú integrálokat kell vizsgálni. Másrészt az integrál viselkedése lényegesen különböző lenne az átlókban. Az átlóban ugyanis, amikor az egyes mintaelemek hozadékát tekintjük az empirikus mérték normalizáltjának szorzatában, az egyes mintaelemek többszörös multiplicitással jelennek meg.)

A kérdés vizsgálata előtt megjegyzést tettem arról, hogy miért lehet fontos a felvetett kérdés vizsgálata. Ilyen probléma jelenik meg akkor, ha a matematikai statisztika egyik fontos problémájának a *maximum likelihood* módszernek természetes általánosítását az úgynevezett *nem-paraméteres maximum-likelihood* módszert vizsgáljuk, és megpróbáljuk bebizonyítani a paraméteres esetben alapvető eredmények általánosítását ebben az esetben is.

A maximum likelihood módszerről szóló talán legfontosabb eredmény azt állítja, hogy nagyon általános feltételek mellett a maximum likelihood becslés hibája beszorozva $\sqrt{n}$-nel aszimptotikusan normális eloszlású nulla várható értékkel és ismert szórásnégyzettel. Az eredmény bizonyításában egy viszonylag bonyolult egyenlet, az úgynevezett maximum-likelihood egyenlet gyökeinek a viselkedését kell vizsgálnunk. Ez a feladat eredeti formájában meglehetősen nehéz, de leegyszerűsödik, ha az egyenlet linearizáltját tekintjük. Jelen esetben ez azt jelenti, hogy az egyenlet baloldalán szereplő kifejezésnek tekintjük a Taylor sor fejtését az (ismeretlen) becsülendő paraméter körül. Ekkor egy olyan egyenletet kapunk, amelynek a gyökét jól tudjuk vizsgálni. Ahhoz, hogy bebizonyítsuk a kívánt eredményt, be kell látnunk, hogy a linearizálás segítségével kapott közelítő becslés jól közelíti a valódi egyenlet gyökét. Ezt viszonylag egyszerűen megmutathatjuk a Taylor-sorok tulajdonságainak felhasználásával.

Ha a maximum likelihood módszert megpróbáljuk általánosítani a nem-paraméteres esetre (aminek a definiciója korántsem egyszerű, és annak megtalálása nem tekinthető lezárt, megoldott problémának), akkor felmerül az a kérdés, hogy lehetséges-e a fent vázolt módszer általánosítása erre az általánosabb esetre. Jelenleg csak részleges eredményeket tudok bizonyítani, amelyek azt mutatják, hogy bizonyos speciális nem-triviális esetekben ez megtehető. Viszont ezekben a vizsgálatokban annak a lépésnek a bizonyítása, hogy a keresett gyöknek a linearizált egyenlet segítségével kapott közelítő megoldása közel van az igazi megoldáshoz nem olyan egyszerű, mint a parametrikus esetben. Annak érdekében, hogy ezt megmutassuk, szükséges az előadás elején megfogalmazott feladatot megoldani. Ugyanis a linearizáció során olyan hibatag jelenik meg, mint amilyet ebben a problémban vizsgálunk, (pontosabban egy ilyen kifejezés beszorozva az $n^{-1/2}$ konstanssal.) Bár egyelőre csak bizonyos speciális esetekben tudtam ezt a módszert alkalmazni, mégis azt sejtem, hogy az sokkal általánosabb körülmények között is alkalmazható. Ezért úgy vélem, hogy az előadásban vizsgált probléma nagyon fontos.

Megjegyzem, hogy valójában egy általánosabb problémát kell vizsgálni, de annak az itt tárgyalt kérdés az első megkerülhetetlen része. Az általános kérdés a következő: Legyen adva k-változós $f(x_1,\dots,x_k)$ függvényeknek egy $\mathcal{F}$ családja, és adjunk jó becslést a $P\left(\sup\limits_{f\in\mathcal{F}}|J_{n,k}(f)|>x\right)$ valószínűségre. Azt várjuk, hogy amennyiben ez a függvénycsalád szép, akkor hasonló becslést kaphatunk, mint akkor, ha csak egy függvényt tekintünk. Természetes jelöltek az ilyen feltételt teljesítő függvénycsaládokra az úgynevezett Vapnik–Ĉervonenkis osztályt alkotó függvénycsaládok. Ezzel a most megfogalmazott kérdéssel azonban ebben az előadásban nem foglalkoztam. Csak megjegyeztem, hogy miért jelenik meg ez a probléma a nem paraméteres maximum-likelihood becslésben.

Tekintsük a következő tipikus a nem-paraméteres maximum likelihood módszerrel megoldandó feladatot: Becsüljünk meg egy ismeretlen $F(\cdot)$ eloszlásfüggvényt bizonyos megfigyelések alapján valamely $F_n(\cdot)$ véletlen becslőfüggvénnyel. A becslés hibájának szuprémumát, azaz a $\sup\limits_{-\infty<x\infty}|F_n(x)-F(x)|$ kifejezést akarjuk vizsgálni. E kifejezés vizsgálata során az $F_n(x)-F(x)$ valószínűségi változót tekintjük minden $-\infty<x<\infty$ számra. Alkalmazzunk alkalmas linearizációt, Ezután egy olyan kifejezést kell vizsgálnunk, mint amilyent az előadás elején említettem, annak érdekében, hogy megbecsüljük a linearizáció során elkövetett hibát. Viszont vizsgálandó integrálban megjelenő f függvény függ attól az x számtól, ahol az $F_n(x)-F(x)$ különbséget vizsgáljuk. Mivel minket az $F_n(x)-F(x)$ kifejezések szuprémuma érdekel, ezért nem egyetlen (véletlen) integrál értékét kell vizsgálnunk, hanem az előző paragrafusban említett problémát kell tekintenünk, azaz alkalmas függvényosztályba eső fïggvények integráljának szuprémumát kell becsülnünk.

Először megfogalmazom azt az eredményt, amely megadja a felvetett kérdésre a választ, azután összehasonlítom ezt az eredményt néhány más eredménnyel, annak érdekében, hogy jobban megértsük a kapott becslés jelentését.

Tétel A. *Tekintsünk valamilyen független $(X, \mathcal{X})$ térbeli értékeket felvevő független, egyforma eloszlású $\xi_1, \dots, \xi_n$ valószínűségi változókat, amelyek μ eloszlása nem atomos. Legyen $f = f(x_1, \dots, x_k)$ olyan mérhető függvény a $(X^k, \mathcal{X}^k, \mu^k)$ szorzattéren, $k \ge 1$, amelyre*

$$\|f\|_\infty = \sup_{x_j \in X,\ 1 \le j \le k} |f(x_1, \dots, x_k)| \le 1, \tag{2a}$$

és

$$\frac{1}{k!}\|f\|_2^2 = Ef^2(\xi_1, \dots, \xi_k) = \frac{1}{k!}\int f^2(x_1, \dots, x_k)\mu(\,dx_1)\dots\mu(\,dx_k) \le \sigma^2 \tag{2b}$$

valamilyen $\sigma > 0$ számmal. Ekkor léteznek olyan $C = C_k > 0$ és $\alpha = \alpha_k > 0$, $k = 1, 2, \dots,$ konstansok, amelyekre az (1) formulában definiált $J_{n,k}(f)$ integrál teljesíti a

$$P\left(|J_{n,k}(f)| > x\right) \le C \max\left(\exp\left\{-\alpha\left(\frac{x}{\sigma}\right)^{2/k}\right\}, \exp\left\{-\alpha(nx^2)^{1/(k+1)}\right\}\right) \tag{3}$$

egyenlőtlenséget minden $x > 0$ számra. A $C = C_k > 0$ és $\alpha = \alpha_k > 0$ konstansok csak a k paramétertől függnek.

Jegyezzük meg, hogy nagy n számra a $\sqrt{n}(\mu_n(\cdot) - \mu(\cdot))$ normalizált mérték közel van egy (véletlen) Gauss eloszlású mértékhez. Ez azt sugallja, hogy egy k-változós függvénynek a Tétel A-ban vizsgált normalizált empirikus mérték szorzatmértéke szerinti integrálja hasonlóan viselkedik, mint egy k-változós függvénynek egy Wiener-folyamat szerinti többváltozós úgynevezett Wiener–Itô integrálja. Az alábbi Tétel B-ben megfogalmazzuk a Tétel A egy olyan analogonját, ahol többváltozós függvények Wiener–Itô integrálját tekintjük.

Tétel B. *Legyen $f(x_1, \dots, x_k)$ k-változós valós értékű függvény, amelyre*

$$\frac{1}{k!}\int_0^\infty \cdots \int_0^\infty f^2(x_1, \dots, x_k)\,dx_1 \dots\, dx_k \le \sigma^2 \tag{4}$$

valamely $\sigma > 0$ számmal. Legyen $W(x)$ Wiener folyamat a $x \ge 0$ félegyenesen, és tekintsük az

$$I_{n,k}(f) = \int_0^\infty \cdot \int_0^\infty f(x_1, \dots, x_k) W(\,dx_1)\dots W(\,dx_k) \tag{5}$$

k-változós Wiener–Itô integrált. Ez a Wiener–Itô integrál teljesíti a

$$P(|I_{n,k}(f)| > x) \le C_k \exp\left\{-\alpha_k\left(\frac{x}{\sigma}\right)^{2/k}\right\} \tag{6}$$

egyenlőtlenséget valamilyen $\alpha_k > 0$ és $C_k > 0$ konstansokkal, amelyek csak a k paramétertől függnek.

Annak érdekében, hogy a Tétel B tartalmát jobban megértsük, a következő megjegyzést teszem.

Megjegyzés: Ha ξ standard normális eloszlású valószínűségi változó, akkor $P(\sigma|\xi^k| > x) \le \text{const.} \exp\left\{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^{2/k}\right\}$, és ez az egyenlőtlenség lényegében éles.

Valóban,

$$P(\sigma|\xi^k| > x) = P\left(|\xi| > \left(\frac{x}{\sigma}\right)^{1/k}\right) \le \text{const.} \exp\left\{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^{2/k}\right\},$$

mivel egy ξ standard normális eloszlású valószínűségi változó esetén $P(|\xi| > u) < \text{const.}\, e^{-u^2/2}$, és ez a becslés lényegében éles.

Összehasonlítva a Tétel B eredményét a Megjegyzés állításával azt látjuk, hogy egy k-változós Wiener–Itô integrál eloszlására hasonló becslést tudunk adni, mint egy $\text{const.}\, \sigma\xi^k$ valószínűségi változó eloszlására, ahol σ^2 a Wiener–Itô integrál szórásnégyzete. (Jegyezzük meg, hogy a (4) formulában megfogalmazott feltételnek a baloldalán szereplő kifejezés a Wiener-Itô integrál szórásnégyzetével egyenlő.) Érdemes a Tétel A és Tétel B eredményét is összehasonlítani. Azt láthatjuk, hogy a Tétel A-ban és Tétel B-ben szereplő két becslés hasonló. Két különbség van. Az egyik az, hogy a Tétel A-ban szereplő becslésben, a (3) formulában szereplő maximumban van egy másik tag, aminek a megfelelője a Tétel B-ben megfogalmazott (6) formulában nem szerepel. Jegyezzük meg, hogy abban az esetben, ha a tekintett f függvény négyzetintegrálja vagy másképp megfogalmazva a tekintett véletlen integrál szórásnégyzete nem túl kicsi, akkor a (3) képlet jobboldalán szereplő kifejezésben az első tag adja meg a maximum értékét. Ezért ebben az esetben a normalizált empirikus mérték k-szoros szorzata szerinti sztochasztikus integrál eloszlásfüggvényére hasonló becslés érvényes, mint egy Wiener folyamat szerinti k-szoros integrálra. A másik különbség a két eredmény körött az, hogy a Tétel A feltételei között szerepel egy a (2a) formulában megfogalmazott feltétel, aminek nincs megfelelője a Tétel B-ben.

Azt mondhatjuk, hogy a normalizált empirikus mérték szerinti k-változós integrálok hasonlóan viselkednek a k-változós Wiener-Itô integrálokhoz, feltéve, hogy ezen integrálok szórásnégyzete nem túl kicsi. (Az, hogy mely σ^2 érték számít kicsinek az egyrészt a minta n elemszámtól, másrészt attól az x számtól függ, amely meghatározza, hogy mely pontban akarjuk a véletlen integrál eloszlásfüggvényét becsülni.) Felmerül az a kérdés, hogy ez a szórásnégyzetről szóló megszorítás lényeges-e vagy esetleg van remény arra, hogy egy finomabb érvelés segítségével meg lehet tőle szabadulni. A következő példában megmutatjuk, hogy egy ilyen megszorítás valóban szükséges. Az egyszerűség kedvéért csak $k = 1$ esetben mutatok példát, de némi plusz munkával hasonló példát lehet konstruálni tetszőleges $k \ge 1$ multiplicitású integrálokra is.

Példa: Legyen $\xi_1, \dots, \xi_n$ független a $[0, 1]$ intervallumban egyenletes eloszlású valószínűségi változók sorozata, és tekintsük a következő valamely p paramétertől, $0 < p < \frac{1}{2}$, függő $f(x) = f_p(x)$ függvényt a $[0, 1]$ intervallumon:

$$f(x) = f_p(x) = \begin{cases} 1 & \text{ha } 0 \le x < p, \\ -1 & \text{ha } p \le x < 2p, \\ 0 & \text{ha } 2p \le x \le 1. \end{cases}$$

Tekintsük a most definiált $f(\cdot)$ függvény integrálját a $\xi_1, \dots, \xi_n$ valószínűségi változók μ_n empirikus mértékének normalizáltja szerint. Vegyük észre, hog jelen esetben $\sigma^2 = \int_0^1 f^2(u)\,du = 2p$. Alsó becslést adunk az $\int f_n(u)\sqrt{n}(\mu_n(\,du) - du)$ integrálra, és ebből a becslésből látható lesz, hogy kis p, azaz kis σ^2 esetén a (3) formulában az első tag, azaz a Gauss közelítés által adott becslés nem teljesül a vizsgált valószínűségre. Ezért kellett egy második korrekciós tagot is megadni ebben a becslésben.

Vezessük be az

$$\eta_k = \begin{cases} 1 & \text{ha } 0 \le \xi_k < p, \\ -1 & \text{ha } p \le \xi_k < 2p. \\ 0 & \text{ha } 2p \le \xi_k \le 1, \end{cases} \qquad 1 \le k \le n$$

valószínűségi változókat. Ekkor

$$J_{n,1}(f) = \sqrt{n} \int_0^1 f(x)[\mu_n(\,dx) - \,dx] = \sqrt{n} \int_0^1 f(x)\mu_n(\,dx) = \frac{1}{\sqrt{n}} \sum_{k=1}^n \eta_k.$$

Továbbá, mivel $P(\eta_k = 1) = P(\eta_k = -1) = p$, $P(\eta_k = 0) = 1 - 2p$, azért

$$\begin{aligned} &P(J_{n,1}(f) > x) \\ &\quad = P\left(\sum_{k=1}^n \eta_k > \sqrt{n}x\right) \ge P\left(\eta_k = 1, \text{ ha } 0 \le k \le \sqrt{n}x, \sum_{k=\sqrt{n}x+1}^n \eta_k \ge 0\right) \\ &\quad = \prod_{k=1}^{\sqrt{n}x} P(\eta_k = 1) \cdot P\left(\sum_{k=\sqrt{n}x+1}^n \eta_k \ge 0\right) = p^{\sqrt{n}x} \cdot \frac{1}{2}. \end{aligned}$$

Kis $p = \frac{\sigma^2}{2}$ paraméter esetén ez a valószínűség viszonylag nagy (összehasonlítva a Gauss approximáció által sugallt becsléssel). Ugyanis

$$P(J_{n,1}(f) > x) \ge p^{\sqrt{n}x} \cdot \frac{1}{2} \ge C \exp\left\{-\sqrt{n}x \log \frac{2}{\sigma^2}\right\}$$

alkalmas (univerzális C konstanssal), míg a Gauss közelítés azt sugallná, hogy

$$P(J_{n,1}(f) > x) \le C \exp\left\{-\frac{x^2}{2\sigma^2} + \log \sigma\right\}.$$

Ha $\sqrt{n}x \log \frac{2}{\sigma^2} \ll \frac{x^2}{2\sigma^2}$, azaz $x \gg \sqrt{n}\sigma^2 \log \frac{1}{\sigma^2}$, akkor a Gauss approximáció által sugallt becslés nem érvényes a Tétel A-ban vizsgált kifejezésre. Ez példát mutat arra, hogy a nrmális közelítés által sugallt kifejezés mellett szükség van még egy korrekciós tagra a (3) formulában megadott becslésben.

Láttuk, hogy az empirikus mérték szerinti többszörös integrálokra a Tétel A-ban adott becslés hasonló a Tétel B-ben megfogalmazott, és Wiener–Itô integrálokról szóló becsléshez. Természetes azt várni, hogy a Tétel B bizonyításában alkalmazott módszer alkalmas adaptációja segítségével be lehet bizonyítamo a Tétel A-t. Az egyik lehetséges ilyen hozzáálláson alapuló bizonyítás a következő: Írjuk fel a

$$P\left(|J_{n,k})(f)| > x\right) \le \frac{EJ_{n,k}(f)^{2M}}{x^{2M}}$$

Markov egyenlőtlenséget, amely érvényes minden $M = 1, 2, \dots$ számra, és alkalmazzuk azt az M paramétert, amelyik a lehető legjobb becslést adja. Ez a módszer alkalmazható, de ehhez szükséges jó becslést adni a $J_{n,k}(f)$ véletlen integrálok momentumaira. Az alább megfogalmazott eredmény olyan becslést ad, amelyikből viszonylag egyszerűen levezethető a Tétel A eredménye.

Segédtétel. *Ha adottak $\xi_1, \dots, \xi_n$ független μ eloszlású valószínűségi változók, valamint egy $f = f(x_1, \dots, x_k)$ k-változós függvény, amely teljesíti az*

$$\|f\|_\infty = \sup_{x_j \in X,\ 1 \le j \le k} |f(x_1, \dots, x_k)| \le 1,$$

és

$$\|f\|_2^2 = \int f^2(x_1, \dots, x_k)\mu(\,dx_1) \dots \mu(\,dx_k) \le \sigma^2,$$

feltételeket, és a μ mérték nem atomos, akkor léteznek olyan C_k konstansok, $k = 1, 2, \dots$, amelyekre minden $M = 2^m$ alakú számra, ahol m nem negatív egész, valamint minden $n = 1, 2, \dots$ paraméterre

$$EJ_{n,k}(f)^{2M} \le \left(C_k \sigma^2 M^k\right)^M \cdot \max\left(1, \left(\frac{M}{n\sigma^2}\right)^M\right). \tag{7}$$

A C_k számok csak a k paramétertől függnek, azaz nem függnek az f függvénytől.

Jegyezzük meg, hogy a (7) formulában szereplő kifejezés fő tagja, a $\left(C_k \sigma^2 M^k\right)^M$ kifejezés, olyan nagyságrendű, mint egy const. $\sigma\xi^k$ valószínűségi változó $2M$-ik momentuma, ahol ξ standard normális eloszlású valószínűségi változó. Ez azt jelenti, hogy a $J_{n,k}$ véletlen integrálok momentumai úgy viselkednek, mint ahogy azt a normális közelítés sugallja, feltéve, hogy azok σ^2 szórásnégyzete nem túl kicsi. De annak érdekében, hogy minden $\sigma > 0$ paraméterre érvényes becslést kapjunk szükség volt egy korrekciós tagot is megjelentetni a (7) formulában.

Felmerül az a kérdés, hogy hogyan lehet a most megfogalmazott Segédtételt bebizonyítani. Vizsgáljuk meg, hogyan lehet az $I_{n,k}(f)$ Wiener–Itô integrálok momentumainak becsléséről szóló analog eredményt bebizonyítani.

A momentumok becslését megtehetjük a sztochasztikus integrálokról szóló egyik alapvető eredmény, az úgynevezett diagram formula segítségével. Ezen eredmény megmutatja, hogy hogyen lehet két $I_{n,k}(f)$ és $I_{n,l}(g)$ k illetve l multiplicitású Wiener–Itô integrál szorzatát felírni alkalmas (különböző multiplicitású) Wiener–Itô integrálok összegeként. (Az integrálokban szereplő integrandusokat bizonyos diagramok segítségével definiálják, innen származik az eredmény diagram formula elnevezése.)

E formula azért hasznos, mert a Wiener–Itô integrálok első és második momentumát egyszerűen ki tudjuk fejezni. Nevezetesen,

$$\begin{aligned} &EI_{n,k}(f)=0, \\ &EI_{n,k}(f)^2=\frac{1}{k!}\int_0^\infty\cdots\int_0^\infty f^2(x_1,\dots,x_k)\,dx_1\dots\,dx_k=\sigma^2. \end{aligned} \tag{8}$$

Ezután az $EI_{n,}(f)^4$ negyedik momentumot a következő módon becsülhetjük meg a diagram formula segítségével: Írjuk fel az $I_{n,k}(f)^4=(I_{n,k}(f)\cdot I_{n,k}(f))^2$ azonosságot, és fejezzük ki az $I_{n,k}(f)\cdot I_{n,k}(f)$ szorzatot a diagram formula segítségével, mint Wiener–Itô integrálok összegét. Ez azt jelenti, hogy elég Wiener–Itô integrálok összegének négyzetét becsülni, ami lehetséges a (8) formula segítségével. Ezután felírva az

$$I_{n,k}(f)^8=(I_{n,k}(f)\cdot I_{n,k}(f))^4$$

azonosságot, hasonlóan megbecsülhetjük az $EI_{n,k}(f)^8$ momentumot a negyedik momentumra adott becslések segítségével, majd indukcióval becsülhetjük az $EI_{n,k}(f)^{2M}$ momentumokat minden $M=2^m$ alakú kitevőre. Az itt vázolt módszer kidolgozásával be lehet látni, hogy

$$EI_{n,k}(f)^{2M}\le\left(C_k\sigma^2M^k\right)^M$$

minden $M=2^m$ alakú számra, ahol $\sigma^2=\frac{1}{k!}\int_0^\infty\cdots\int_0^\infty f(x_1,\dots,x_k)\,dx_1\dots\,dx_k$. Ezen becslés segítségével be lehet bizonyítani a Tétel B-t.

Lehet-e ezt a módszert adaptálni úgy, hogy annak segítségével be tudjuk bizonyítani a Tétel A eredményét? A válasz erre a kérdésre igenlő, de a program végrehajtása sok technikai nehézség legyőzését igényli. Igaz a (8) formula következő módosított változata:

$$\begin{aligned} |EJ_{n,k}(f)|&\le\frac{C^k}{k^{k/2}}\int|f(u_1,\dots,u_k)|\mu(\,du_1)\dots\mu(\,du_k), \\ EJ_{n,k}(f)^2&\le\frac{C^k}{k^k}\int f^2(x_1,\dots,x_k)\mu(\,du_1)\dots\mu(\,du_k). \end{aligned}$$

Ez a formula némileg eltér a Wiener–Itô integrál első két momentumát kifejező (8) formulától. Emiatt megjelenik néhány technikai probléma, de igazi nehézséget ez a különbség nem okoz. Megjegyzem, hogy az eltérés oka az, hogy míg a Wiener–Itô integrálokban szereplő Wiener mértékre igaz, hogy diszjunkt halmazok Wiener mértéke független, addig diszjunkt halmazok normalizált empirikus mértéke között van bizonyos enyhe függés. A fontos kérdés azonban az, hogy be tudjuk-e bizonyítani a diagram formula megfelelőjét abban az esetben, ha többváltozós Wiener-Itô integrálok helyett a normalizált empirikus mérték szerinti többváltozós integrálokat tekinünk, és be tudjuk-e bizonyítani egy ilyen formula segítségével a segédtételt. A válasz ezekre a kérdésekre igenlő. A lényeges újdonság az, hogy a normalizált empirikus mértékek szrinti integrálok szorzatát kifejező diagram formulában új diagramok és új e diagramoknak megfelelő integrálok is megjelennek. Ezeket az új integrálokat is meg kell becsülni. Ez megoldható, bár nem egyszerű feladat. Ezek az új diagramok, illetve a nekik megfelelő íntegrálok adják azt a hozadékot, amelynek becslése miatt megjelenik a korrekciós tag a (8) formulában.