数理解析研究所講究録 1308 巻 2003 年 185-198

核関数のフーリエ級数展開による U-, V-統計量の解析について

武蔵工業大学・工学部　金川 秀也 (Shuya Kanagawa)
Faculty of Engineering,
Musashi Institute of Technology

§0. 序

$\{\xi_j, j\geq 1\}$ を分布 μ に従う実数値確率変数列とする．また $u(x_1,x_2,\cdots,x_k)$ を実対称関数とする． $u(x_1,x_2,\cdots,x_k)$ を核関数とする統計量を対称統計量と呼ぶ．対称統計量の一種であるU-統計量あるいはV-統計量の漸近的な性質を核関数のフーリエ級数展開によって調べる方法について解説する．

Example 1. U-統計量（degree k）: $U_n := \dfrac{1}{\binom{n}{k}} \displaystyle\sum_{1\leq i_1<i_2<\cdots<i_k\leq n} u\left(\xi_{i_1},\xi_{i_2},\cdots,\xi_{i_k}\right)$

<u>Sample mean</u> : $u(x)=x \quad (k=1)$

<u>Sample variance</u> : $u(x_1,x_2)=\dfrac{1}{2}(x_1-x_2)^2 \quad (k=2)$

Example 2. V-統計量（degree k）: $V_n := n^{-k} \displaystyle\sum_{1\leq i_1\leq i_2\leq\cdots\leq i_k\leq n} u\left(\xi_{i_1},\xi_{i_2},\cdots,\xi_{i_k}\right)$

<u>Cramér-von Mises-Smirnov statistic</u> :

$$u(x_1,x_2)=\int_0^1 w(u)\left(I_{\{x_1\leq u\}}-u\right)\left(I_{\{x_2\leq u\}}-u\right)du \quad (k=2)$$

このとき

$$V_n=\frac{1}{n-1}\int_0^1 w(x)\left(F_n(x)-1\right)^2 dx,$$

ただし $w(x)$ は重み関数、$F_n(x)$ は $\{\xi_j, 1\leq j\leq n\}$ の経験分布関数.

任意の $x_2,\cdots,x_k$ に対して

$$\int_{-\infty}^{\infty} u(y,x_2,\cdots,x_k)\mu(dy)=0$$

のとき$u\left(x_1,x_2,\cdots,x_k\right)$は退化しているという．

§１．degree ２の退化した核関数のフーリエ級数展開

$\left\{\xi_j, j\geq 1\right\}$がi.i.d.の場合はHoeffding(1948)によるH-分解によって退化していない$u\left(x_1,x_2,\cdots,x_k\right)$に関するU-統計量、V-統計量について中心極限定理（漸近正規性）、Donsker型不偏原理、概収束型不偏原理、漸近展開、大偏差原理など極めて精密に漸近的な性質が調べられている。またYoshihara(1976)によって$\left\{\xi_j, j\geq 1\right\}$が$\phi$-mixing性や absolutely regular性程度の弱従属性を持つ場合は$\left\{\xi_j, j\geq 1\right\}$をi.i.d.確率変数列で近似することによってH-分解を用いて退化していない$u\left(x_1,x_2,\cdots,x_k\right)$に関するU-統計量、V-統計量について中心極限定理が成り立つことが示されている。弱従属確率変数列の独立確率変数列による近似についてはYoshihara (1993)を参照のこと．

一方、退化している核関数$u\left(x_1,x_2,\cdots,x_k\right)$に対して$\left\{\xi_j, j\geq 1\right\}$がi.i.d.や弱従属の場合は Dynkin-Mandelbaum (1983)、Dehling (1983)、Kanagawa-Yoshihara (1993)、Dehling-Denker-Philipp (1994)、Kanagawa and Yoshihara (1994)、Kanagawa (1999)によって各種の極限定理が調べらている．$\left\{\xi_j, j\geq 1\right\}$が退化している場合や強い従属性を持つ場合は直接H-分解が使えないために何らかの代わりの方法を見つける必要があるが、その一つとして$u\left(x_1,x_2,\cdots,x_k\right)$をフーリエ級数展開し、各種の漸近的な性質を調べることについて考察する．

degree 2 で核関数$u\left(x_1,x_2\right)$が退化している場合はSerfling(1980)によって次のような固有値と固有関数を用いたフーリエ級数展開が示されている。

<u>定理１．</u> $u\left(x_1,x_2\right)$は$R\times R$上の対称でξ_jの分布μについて２乗可積分実数値関数で退化しているとする、i.e.任意のxに対し$E\left(u(\xi_1,x)\right)=0$．$f\in L^2=L^2(d\mu\times d\mu)$に対して$T_u f(x):=E\left[u(\xi_1,x)f(\xi_1)\right]$によって$T_u: L^2\rightarrow L^2$ (trace class)を定義すると、T_uは固有関数$\left\{g_i\right\}$と固有値$\left\{\lambda_i\right\}$を持つ．これらが次の条件を満足すると仮定する．

$$\begin{cases} E\left(g_i(\xi_1)\right)=0, \quad E\left(g_i^2(\xi_1)\right)=1 \\ E\left(g_i(\xi_1)g_j(\xi_1)\right)=0 \ (i\neq j), \quad E\left(h(\xi_1,x)g_i(\xi_1)\right)=\lambda_i g_i(x) \end{cases}$$

この時$L^2\left(d\mu\times d\mu\right)$収束の意味で$\sum_{k=1}^{N}\lambda_k g_k\left(x_1\right)g_k\left(x_2\right)$は$u\left(x_1,x_2\right)$に収束する、

$$\lim_{N\to\infty}\iint\left(u(x_1,x_2)-\sum_{k=1}^{N}\lambda_k g_k(x_1)g_k(x_2)\right)^2 d\mu(x_1)d\mu(x_2)=0. \tag{1}$$

Mercerの定理より$u\in L^2$が連続な非負定値関数の場合に各点収束の意味で(1)が収束することが知られているが、最近佐藤坦氏によって区分的に連続な関数の場合にも成立することが示された．

<u>定理２．</u> 佐藤(1992)　Xは可分距離空間、μはX上のボレル測度、$u(\cdot,\cdot)$は$X\times X$上の関数で、$\mu(X\setminus X_0)=0$であるμ－可測集合$X_0\subset X$に対して、$X_0\times X_0$上で$u(\cdot,\cdot)$が連続とする．この時

$$\iint_{X\times X}|u(x,y)|^2\mu(dx)\mu(dy)<\infty \tag{2}$$

$$\iint_{X\times X}u(x,y)f(x)\overline{f(x)}\mu(dx)\mu(dy)\geq\infty,\quad f\in L^2(d\mu\times d\mu) \tag{3}$$

このとき積分作用素$T_u f(x):=E[u(\xi_1,x)f(\xi_1)]$がneclearであるための必要十分条件は

$$\int_X u(x,x)\mu(dx)<\infty \tag{4}$$

である．

定理2（Mercer-佐藤の定理）よりμがルベーグ測度について絶対連続である場合は階段関数のような区分的に連続な関数$u(x,y)$に対して(1)が各点連続の意味で収束する．

<u>系１．</u> $\mu(X\setminus X_0)=0$であるμ－可測集合$X_0\subset X$に対して、$X_0\times X_0$上で$u(\cdot,\cdot)$が連続とする．(2), (3), (4) の条件の下で

$$u(x_1,x_2)=\sum_{k=1}^{\infty}\lambda_k g_k(x_1)g_k(x_2)\qquad \text{a.e. } \mu(dx)\mu(dy) \tag{5}$$

またヒルベルト空間Hを次のように定義する

$$H:=\left\{\boldsymbol{x}=(x_1,x_2,\cdots)\in\mathbf{R}^\infty:\ \sum_{k=1}^{\infty}|\lambda_k|\xi_k^2<\infty\right\}$$

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \sum_{k=1}^{\infty} |\lambda_k| x_k y_k \tag{6}$$

$$\|\boldsymbol{x}\| := \left(\sum_{k=1}^{\infty} |\lambda_k| x_k^2 \right)^{1/2}$$

固有値についての条件

$$\sum_{k=1}^{\infty} |\lambda_k| < \infty \tag{7}$$

の下で

$$E\left(\sum_{k=1}^{\infty} |\lambda_k| g_k^2(\xi_i) \right) = \sum_{k=1}^{\infty} |\lambda_k| E\left(g_k^2(\xi_i) \right) = \sum_{k=1}^{\infty} |\lambda_k| < \infty,$$

が成り立つのでH-値確率変数 $G_i := \left(g_1(\xi_i), g_2(\xi_i), g_3(\xi_i), \cdots \right)$ $i \geq 1$を用いて$\{V_n, n \geq 1\}$を書き直すことができる.

$$\boldsymbol{x} \in H := \left\{ \boldsymbol{x} = (x_1, x_2, \cdots) \in \mathbf{R}^{\infty} : \sum_{k=1}^{\infty} |\lambda_k| x_k^2 < \infty \right\}$$

に対して

$$h(\boldsymbol{x}) := \sum_{k=1}^{\infty} \lambda_k x_k^2$$

とおくと、

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} G_i = \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} g_1(\xi_i), \frac{1}{\sqrt{n}} \sum_{i=1}^{n} g_2(\xi_i), \frac{1}{\sqrt{n}} \sum_{i=1}^{n} g_3(\xi_i), \cdots \right) \in H. \tag{8}$$

$$nV_n = \frac{1}{n} \sum_{1 \leq i, j \leq n} u\left(\xi_i, \xi_j\right) = \frac{1}{n} \sum_{1 \leq i, j \leq n} \sum_{k=1}^{\infty} \lambda_k g_k(\xi_i) g_k(\xi_j) \tag{9}$$

$$= \frac{1}{n} \sum_{k=1}^{\infty} \lambda_k \sum_{1 \leq i, j \leq n} g_k(\xi_i) g_k(\xi_j) = \frac{1}{n} \sum_{k=1}^{\infty} \lambda_k \left\{ \sum_{i=1}^{n} g_k(\xi_i) \right\}^2 = h\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} G_i \right).$$

ここでH-値確率変数 $G_i := \left(g_1(\xi_i), g_2(\xi_i), g_3(\xi_i), \cdots \right)$に関する極限定理を用いて$\{V_n, n \geq 1\}$の漸近的な性質を調べることができる.

§２．degree ２のU-統計量、V-統計量の表現

$$U_n := \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} u\left(\xi_i, \xi_j\right),\ V_n := \frac{1}{n^2} \sum_{1 \le i, j \le n} u\left(\xi_i, \xi_j\right)$$ について

$$nV_n = \frac{1}{n} \sum_{1 \le i, j \le n} u\left(\xi_i, \xi_j\right) = \frac{1}{n} \sum_{1 \le i, j \le n} \sum_{k=1}^{\infty} \lambda_k g_k\left(\xi_i\right) g_k\left(\xi_j\right) \tag{10}$$

$$= \frac{1}{n} \sum_{k=1}^{\infty} \lambda_k \sum_{1 \le i, j \le n} g_k\left(\xi_i\right) g_k\left(\xi_j\right) = \frac{1}{n} \sum_{k=1}^{\infty} \lambda_k \left\{ \sum_{i=1}^{n} g_k\left(\xi_i\right) \right\}^2$$

$$= h\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} G_i \right).$$

同様に

$$\sqrt{n(n-1)}\, U_n = \frac{2}{\sqrt{n(n-1)}} \sum_{1 \le i < j \le n} u\left(\xi_i, \xi_j\right) \tag{11}$$

$$= \frac{2}{\sqrt{n(n-1)}} \sum_{1 \le i < j \le n} \sum_{k=1}^{\infty} \lambda_k g_k\left(\xi_i\right) g_k\left(\xi_j\right)$$

$$= \frac{1}{\sqrt{n(n-1)}} \sum_{k=1}^{\infty} \lambda_k \sum_{1 \le i < j \le n} 2 g_k\left(\xi_i\right) g_k\left(\xi_j\right)$$

$$= \frac{1}{\sqrt{n(n-1)}} \sum_{k=1}^{\infty} \lambda_k \left[\left\{ \sum_{i=1}^{n} g_k\left(\xi_i\right) \right\}^2 - \sum_{i=1}^{n} g_k^2\left(\xi_i\right) \right]$$

$$= \frac{n}{\sqrt{n(n-1)}} \sum_{k=1}^{\infty} \lambda_k \left[\left\{ \sum_{i=1}^{n} \frac{g_k\left(\xi_i\right)}{\sqrt{n}} \right\}^2 - \frac{1}{n} \sum_{i=1}^{n} \left(\frac{g_k\left(\xi_i\right)}{\sqrt{n}} \right)^2 \right]$$

$$= \frac{n}{\sqrt{n(n-1)}} \sum_{k=1}^{\infty} \lambda_k \left\{ \sum_{i=1}^{n} \frac{g_k\left(\xi_i\right)}{\sqrt{n}} \right\}^2 - \frac{1}{\sqrt{n(n-1)}} \sum_{k=1}^{\infty} \lambda_k \sum_{i=1}^{n} \frac{g_k^2\left(\xi_i\right)}{n}$$

$$= \frac{n}{\sqrt{n(n-1)}} h\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} G_i \right) - \frac{1}{\sqrt{n(n-1)}} \sum_{k=1}^{\infty} \lambda_k \sum_{i=1}^{n} \frac{g_k^2\left(\xi_i\right)}{n}.$$

§ 2. degree $k\ (k \ge 3)$ の退化した核関数のフーリエ級数展開

次にdegreeが３以上の核関数$u\left(x_1, x_2, \cdots, x_k\right)$に関するU-統計量やV-統計量につ

いても、核関数のフーリエ級数が各点収束すれば、degree 2の場合と同様にそれらの統計量が無限次元確率変数列の和を用いて表現できることを示す．まず$u(x_1,x_2,\cdots,x_k)$のフーリエ級数展開について考察する．以下の議論は、高橋陽一郎「実解析とFourier解析１」、岩波講座、第二章を参考とした．

$T=[0,1]$、$C(T)$をT上の連続関数全体、$C(T)$上の内積を$\langle f,g\rangle=\int_T f(t)\overline{g(t)}dt$, $f,g\in C(T)$、$e_1,e_2,\cdots$を$C(T)$上の正規直行系とする。$u:T^m\to\mathbf{R}$に関するフーリエ係数は、$k=(k_1,k_2,\cdots,k_m)\in Z^m$（$Z$は整数全体）として

$$\hat{u}(k)=\int_{T^m}u(t)\overline{e_k(t)}dt$$

によって与えられる．ただし、$t=(t_1,t_2,\cdots,t_m)\in T^m$に対して

$$e_k(t)=e_{k_1}(t_1)e_{k_2}(t_2)\cdots e_{k_m}(t_m)=\exp\left(2\pi i\sum_{j=1}^{m}k_jt_j\right).$$

もし$u\in L^1(T^m)$ならばフーリエ係数はFubini の定理によって累次積分で表される．

$$S_N(u,t)=\sum_{k_1=-N}^{N}\cdots\sum_{k_m=-N}^{N}\hat{u}(k)e_k(t),$$

$$\langle f,g\rangle=\int_{T^m}f(t)\overline{g(t)}dt,\quad f,g\in C(T^m),$$

$$\|f\|_2=(\langle f,f\rangle)^{1/2},$$

とおくと$u\in L^1(T^m)$のフーリエ級数展開は次のように表される．

<u>定理３.</u>　$u\in C(T^m)$のとき

$$\min_{\{c_k\}}\left\|u-\sum_{k_1=-N}^{N}\cdots\sum_{k_m=-N}^{N}c_ke_k\right\|_2^2=\|u\|_2^2-\sum_{k_1=-N}^{N}\cdots\sum_{k_m=-N}^{N}|\hat{u}(k)|^2$$

$$\|u\|_2=\left(\sum_{k\in Z^m}|\hat{u}(k)|^2\right)^{1/2}.$$

<u>定理４.</u>　$u \in C^m(T^m)$のとき、$t=(t_1,t_2,\cdots,t_m)\in T^m$に関して一様に

$$S_N(u,t)=\sum_{k_1=-N}^{N}\cdots\sum_{k_m=-N}^{N}\hat{u}(k)e_k(t)\to u(t)\quad(N\to\infty). \tag{12}$$

定理３では、$u \in C^m(T^m)$に対して各点一様収束が成り立つので従属確率変数列$\{\xi_j, j\geq 1\}$に対し、$\xi=(\xi_{i_1},\xi_{i_2},\cdots,\xi_{i_m})$と置いたとき、確率１で

$$S_N(u,\xi)=\sum_{k_1=-N}^{N}\cdots\sum_{k_m=-N}^{N}\hat{u}(k)e_{k_1}(\xi_{i_1})e_{k_2}(\xi_{i_2})\cdots e_{k_m}(\xi_{i_m})\to u(\xi_{i_1},\xi_{i_2},\cdots,\xi_{i_m})\quad(N\to\infty)$$

が成り立つ．しかし、Cramér-von Mises-Smirnov statisticでは核関数

$$u(x_1,x_2)=\int_0^1 w(u)\left(I_{\{x_1\leq u\}}-u\right)\left(I_{\{x_2\leq u\}}-u\right)du \tag{13}$$

が連続（区分的連続）ではないのでこのままでは定理２を適用できない．そのため不連続な関数に対するフーリエ級数展開についても調べる必要がある．

以上のように核関数$u(x_1,x_2,\cdots,x_k)$のフーリエ級数が各点収束すれば、degree 2の場合と同様にU-統計量やV-統計量が無限次元確率変数列の和を用いて表現できることを示す．

可分なバナッハ空間Xとノルム$\|\cdot\|$を次の様の定義する．

$$X:=\left\{\boldsymbol{x}=(x_1,x_2,\cdots)\in\mathbf{R}^\infty : \sum_{-\infty\leq k_1,\cdots,k_m<\infty}|\hat{u}(k)||x_{k_1}|\cdots|x_{k_m}|<\infty\right\},$$

$$\|\boldsymbol{x}\|:=\left\{\sum_{-\infty\leq k_1,\cdots,k_m<\infty}|\hat{u}(k)||x_{k_1}|\cdots|x_{k_m}|\right\}^{1/m}.$$

X上の連続関数$\varphi(\boldsymbol{x})$を

$$\varphi(\boldsymbol{x}):=\sum_{-\infty\leq k_1,\cdots,k_m<\infty}\hat{u}(k)x_{k_1}\cdots x_{k_m} \tag{14}$$

とおき、X-値確率変数$\mathbf{G}_k := \left(e_1(\xi_k), e_2(\xi_k), \cdots\right)$, $k \geq 1$とすると、

$$\sum_{k=1}^{n} \mathbf{G}_k = \left(\sum_{k=1}^{n} e_1(\xi_k), \sum_{k=1}^{n} e_2(\xi_k), \cdots \right) \tag{15}$$

より、

$$n^m V_n = \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_m \leq n} u\left(\xi_{i_1}, \xi_{i_2}, \cdots, \xi_{i_m}\right) \tag{16}$$

$$= \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_m=-\infty}^{\infty} \hat{u}(k) \left\{ \sum_{j=1}^{n} e_{k_1}(\xi_j) \sum_{j=1}^{n} e_{k_2}(\xi_j) \cdots \sum_{j=1}^{n} e_{k_m}(\xi_j) \right\}$$

$$= \varphi\left(\sum_{k=1}^{n} \mathbf{G}_k \right).$$

ゆえに$n^m V_n$がX-値確率変数列$\left\{\mathbf{G}_k,\, k \geq 1\right\}$の部分和の汎関数として表すことが出来るので、$\left\{\mathbf{G}_k,\, k \geq 1\right\}$に関する極限定理を応用することでV-統計量$V_n$の漸近的な性質を調べることが出来る。通常のフーリエ級数展開は固有関数による展開のように核関数$u\left(x_1, x_2, \cdots, x_k\right)$に対称性を仮定する必要がないので、U-, V-統計量含む、より一般の対称統計量にも面倒な核関数の対称化を行わずに応用できる.

§3. 大偏差原理への応用

核関数のフーリエ級数展開にヒルベルト空間値確率変数列に関する極限定理を応用することでU-, V-統計量の漸近的な性質を従来の方法よりも明解に考察することが出来る. Cramér-von Mises-Smirnov 統計量のような退化した不連続な核関数を持つ対称統計量に関する重複対数の法則や大偏差理論のような概収束性を論じる場合に、定理1が2乗平均の意味での収束であることから確率変数列に対していくつかのテクニカルな条件を仮定する必要があった. しかし定理2（Mercer-佐藤の定理）を用いる事によって、そのような条件を省略することが出来る. まず大偏差原理について解説する.

<u>定理5.</u>（Cramérの定理） $\left\{\xi_j, j \geq 1\right\}$ i.i.d.実数値確率変数列で平均0，分散1とする. また$E\left\{\exp\left(t|\xi_1|\right)\right\} < \infty$、$t > 0$とする. このとき$S_n = \sum_{i=1}^{n} \xi_i$とおくと、任意の閉集合 $F \subset \mathbf{R}$に対して

$$\overline{\lim_{n\to\infty}}\frac{1}{n}\log P\{S_n\in F\}\leq -\inf_{x\in F}I(x), \tag{17}$$

また任意の開集合$G\subset\mathbf{R}$に対して

$$\underline{\lim_{n\to\infty}}\frac{1}{n}\log P\{S_n\in G\}\geq -\inf_{x\in G}I(x), \tag{18}$$

ただし

$$I(x):=\sup_{\theta}[\theta x-\log M(\theta)],\ M(\theta):=E\{\exp(\theta\xi_1)\}. \tag{19}$$

<u>定理６</u>．(Donsker-Varadhan (1976)) B をバナッハ空間で$\|\cdot\|$をノルムとする．B-値i.i.d.確率変数列$\{\xi_j,j\geq 1\}$は平均０で$E\{\exp(t\|\xi_1\|)\}<\infty,\ t>0$を仮定する．さらに $\Phi:B\to\mathbf{R}$ は連続関数で、定数 C, D 及び任意の$x\in B$に対して

$$\Phi(x)\leq C+D\|x\|. \tag{20}$$

このとき

$$\lim_{n\to\infty}\frac{1}{n}\log E\left\{\exp\left(n\Phi\left(\frac{S_n}{n}\right)\right)\right\}=\sup_{x\in B}[\Phi(x)-h(x)], \tag{21}$$

ただし、$h(x):=\sup_{\varphi\in B^*}[\varphi(x)-\log M(\varphi)]$, B^*はBのdualで$M(\theta):=E\{\exp(\varphi(\xi_1))\}$.

<u>定理７</u>．(Bolthausen (1986))

定理８の条件をすべて仮定する．また$\Phi(\boldsymbol{x}')-h(\boldsymbol{x}')=\sup_{\boldsymbol{x}\in B}[\Phi(\boldsymbol{x})-h(\boldsymbol{x})]$となる$\mathrm{x}'\in B$がただ一つ定まり、$\Phi(\cdot)$ はB上で３次Fréchet微分可能とする．さらにいくつかの仮定の下で

$$\lim_{n\to\infty}\exp\{-n(\Phi(x')-h(x'))\}E\left\{\exp\left(n\Phi\left(\frac{S_n}{n}\right)\right)\right\}=\int\exp\left\{\frac{1}{2}D^2\Phi(x')[y,y]\right\}\gamma(dy),$$

ただし、$\gamma(dy)$はあるガウス測度、D^2は２階のFréchet微分．

<u>定理８</u>．(Kanagawa (1999)) $\{\xi_j,j\geq 1\}$ は i.i.d. で 分布を μ とする. またdegre 2の

対称核関数$u(x,y)$に対して(2), (3), (7)及びCram ér条件

$$E\left(\exp\left(t\|G_1\|\right)\right) < \infty$$

を仮定する．$\boldsymbol{x} = (x_1, x_2, \cdots) \in H$に対して$\Phi(\boldsymbol{x}) := \left|\sum_{k=1}^{\infty} \lambda_k \boldsymbol{x}_k^2\right|^{1/2}$とおく．$G_1$の分布に関するエントロピー関数$h$を

$$h(\boldsymbol{x}) := \sup_{\varphi \in H^*} \left(\varphi(\boldsymbol{x}) - \log M(\varphi)\right),$$

ただしH^*はHのtopological dualで$M(\varphi) := E\left(\exp\left(\varphi(G_1)\right)\right)$. このとき

$$\lim_{n \to \infty} \frac{1}{n} \log E\left(\exp\left(n|V_n|^{1/2}\right)\right) = \sup_{x \in H} \left(\Phi(x) - h(x)\right).$$

さらにμが絶対連続であるとき

$$\lim_{n \to \infty} \frac{1}{n} \log E\left(\exp\left(n|U_n|^{1/2}\right)\right) = \sup_{x \in H} \left(\Phi(x) - h(x)\right).$$

証明．$H = \left\{\boldsymbol{x} = (x_1, x_2, \cdots) \in \mathbf{R}^\infty : \sum_{k=1}^{\infty} |\lambda_k| x_k^2 < \infty\right\}$、
$G_i = \left(g_1(\xi_i), g_2(\xi_i), g_3(\xi_i), \cdots\right) \in H$及び、定理２（Mercer-佐藤の定理）によって

$$u(\xi_i, \xi_j) = \sum_{k=1}^{\infty} \lambda_k g_k(\xi_i) g_k(\xi_j) \quad \text{a.s.}$$

より

$$\lim_{n \to \infty} \frac{1}{n} \log E\left\{\exp\left\{n|V_n|^{1/2}\right\}\right\} = \lim_{n \to \infty} \frac{1}{n} \log E\left\{\exp\left(\left|\sum_{1 \le i, j \le n} u(\xi_i, \xi_j)\right|^{1/2}\right)\right\}$$

$$= \lim_{n \to \infty} \frac{1}{n} \log E\left\{\exp\left(\left|\sum_{1 \le i, j \le n} \sum_{k=1}^{\infty} \lambda_k g_k(\xi_i) g_k(\xi_j)\right|^{1/2}\right)\right\}$$

$$= \lim_{n\to\infty}\frac{1}{n}\log E\left\{\exp\left(\left|\sum_{k=1}^{\infty}\lambda_k\sum_{1\le i,j\le n}g_k(\xi_i)g_k(\xi_j)\right|^{1/2}\right)\right\}$$

$$= \lim_{n\to\infty}\frac{1}{n}\log E\left\{\exp\left(\left|\sum_{k=1}^{\infty}\lambda_k\left(\sum_{i=1}^{n}g_k(\xi_i)\right)^2\right|^{1/2}\right)\right\}$$

$$= \lim_{n\to\infty}\frac{1}{n}\log E\left\{\exp\left(n\left|\sum_{k=1}^{\infty}\lambda_k\left(\frac{1}{n}\sum_{i=1}^{n}g_k(\xi_i)\right)^2\right|^{1/2}\right)\right\}$$

$$= \lim_{n\to\infty}\frac{1}{n}\log E\left\{\exp\left(n\Phi\left(\frac{S_n}{n}\right)\right)\right\},$$

ただし $S_n := \sum_{i=1}^{n}G_i = \sum_{i=1}^{n}(g_1(\xi_i), g_2(\xi_i), g_3(\xi_i), \cdots)$. 一方

$$\Phi(\boldsymbol{x}) = \left|\sum_{k=1}^{\infty}\lambda_k x_k^2\right|^{1/2} \le \left(\sum_{k=1}^{\infty}|\lambda_k|x_k^2\right)^{1/2} = \|\boldsymbol{x}\|$$

より(20)が成り立つため定理８（Donsker-Varadhan大偏差原理）を用いれば定理が証明される.

<u>系２.</u>　μを[0,1]上の一様分布とする．Cramér-von Mises-Smirnov統計量の核関数

$$u(x_1,x_2) = \int_0^1 w(u)\left(I_{\{x_1\le u\}} - u\right)\left(I_{\{x_2\le u\}} - u\right)du$$

に対して定理１０が成り立つ.

重み関数$w(x) = \dfrac{1}{x(x-1)}$に対してCramér-von Mises-Smirnov統計量の固有値は

$$\lambda_k = \frac{1}{k(k+1)},\quad k\ge 1$$

で与えられる（Csáki(1980), Borovskikh(1985)）．ゆえに

$$V_n = \frac{1}{n-1}\int_0^1 w(x)\left(F_n(x) - 1\right)^2 dx$$

に関するDonsker-Varadhan大偏差原理のエントロピー関数$h(\boldsymbol{x})$は

$$H = \left\{\boldsymbol{x} = (x_1, x_2, \cdots) \in \mathbf{R}^\infty : \sum_{k=1}^{\infty} \frac{x_k^2}{k(k+1)} < \infty\right\}$$

の topological dual H^* に対して

$$h(\boldsymbol{x}) = \sup_{\varphi \in H^*}\left(\varphi(\boldsymbol{x}) - \log M(\varphi)\right)$$

によって与えられる．

同様に退化した核関数に対するU-統計量についてBolthausenの大偏差原理も成り立つ．

定理９．(Kanagawa (1999)) 定理８と同じ条件を仮定する．また $\Phi(\boldsymbol{x}') - h(\boldsymbol{x}') = \sup_{\boldsymbol{x} \in B}\left[\Phi(\boldsymbol{x}) - h(\boldsymbol{x})\right]$ となる $\boldsymbol{x}' \in H$ がただ一つ定まり、 $\Phi(\cdot)$ は H 上で３次Fréchet微分可能とする．さらにいくつかの仮定の下で

$$\lim_{n \to \infty} \exp\left\{-n\left(\Phi(x') - h(x')\right)\right\} E\left\{\exp\left(n|U_n|^{1/2}\right)\right\} = \int \exp\left\{\frac{1}{2} D^2\Phi(x')[y,y]\right\}\gamma(dy),$$

ただし、$\gamma(dy)$はあるガウス測度、D^2は２階のFréchet微分．この定理は V_n についても全く同様に成り立つ．

その他の退化した対称核関数についてのU-、V-統計量に関する極限定理について[4]～[7], [9]～[11]を参照されたい．

REFERENCES

[1] Yu. V. Borovskikh, Estimates of characteristic functions of some random variables with applications to ω^2-statistics. II, Theory of Prob. Appl., **30** (1985) 117-127.

[2] E. Csáki, On the standardized empirical distribution function, in B.V. Gnedenko et al. eds., Non-parametric Statistical Inference, Coll. Mathem. Soc. János Bolyai **32** (North-Holland, Amsterdam, 1980) pp. 123-138.

[3] H. Dehling, Limit theorems for sums of weakly dependent Banach space valued

random variables, Wahrscheinlichkeitstheorie verw. Gebiete, **63** (1983), 393-432.

[4] H. Dehling, M. Denker and W. Philipp, Invariance principles for von Mises and U-statistics, Z. Wahrscheinlichkeitstheorie verw. Gebiete, **67** (1984), 139-167.

[5] E. B. Dynkin and A. Mandelbaum, Symmetric statistics, Poisson point processes and multiple Wiener integrals, Ann. Statist., **11** (1983) 739-745.

[6] F. Götze, On Edgeworth expansions in Banach spaces, Ann. Probability, **9** (1981) 852-859.

[7] M. Harel and M. L. Puri, Limiting behaviour of U-statistics, V-statistics, and one sample rank order statistics for non stationary absolutely regular processes, J. Multivariate Anal., **30** (1989), 181-204.

[8] W. Hoeffding, A non-parametric test of independence, Ann. Math. Statist., **19** (1948), 546-557.

[9] S. Kanagawa and K. Yoshihara; The almost sure invariance principles of degenerate U-statistics of degree two for stationary random variables, Stochastic Processes and their Applications, **49** (1994), pp347 - pp356

[10] S. Kanagawa and K. Yoshihara, Almost sure invariance principle for V- and U-statistics based on weakly dependent random variables, Asymptotic Statistics (Proceedings of the Fifth Prague Symposium, Prague, Czech Republic, 1993), Physica-Verlag, (1994) 318 - 327.

[11] S. Kanagawa, A representation of the rate functions in large deviation principles for U-statistics with degenerate kernels, Trends in Probability and Related Analysis (Proceedings of Symposium on Analysis and Probability 1998, Tiwan University, Taipei), 254-264, (1999).

[12] H. Sato, Nuclearity of a nonnegative definite integral kernel on a separable metric space, J. Theoritical Probab., **5** (1992), 349-353.

[13] V. V. Sazonov and V. V. Ul'yanov, Asymptotic expansions of the probability that the sum of independent random variables hits a ball in a Hilbert space, Russian Maht. Surveys, **50:5** (1995) 1045-1063.

[14] R. J. Serfling, Approximation Theorems of Mathematical Statistics, John Wiley and Sons, New York, 1980.

[15] K. Yoshihara, Limiting behaviour of U-statistics for stationary, absolutely regular

processes, Z. Wahrscheinlichkeitstheorie und verw. Gebiete, **35** (1976), 237-252.

[16] K. Yoshihara, Weakly Dependent Stochastic Sequences and Their Applications, Vol. 2, Asymptotic Statistics based on Weakly Dependent Date (Sanseido Co. Ltd., Tokyo, 1993).K. Yoshihara, *Weakly Dependent Stochastic Sequences and Their Applications, Vol. 2, Asymptotic Statistics based on Weakly Dependent Date* (Sanseido Co. Ltd., Tokyo, 1993).

[17] 高橋陽一郎、実解析とFourier解析１、岩波講座、第二章