Sem. on Probab.
Vol.29 1968年
P1-95

SEMINAR ON PROBABILITY

Vol. 29

Gaussian Process の ε-エントロピー

馬場良和・加地紀臣男・井原俊輔

1968

確率論セミナー

Sem. on Probab.
Vol.29 1968年
P1-95

はじめに

C.E.Shannon [28] は通信系の、与えられた精度のもとでの情報伝達速度の概念を導入し、さらに関数空間の次元を測るべく *dimension rate* を定義した。A.N.Kolmogorov はこれらを発展,精密化して[13]において確率分布の ε-エントロピーを、また[14](V.M.Tikhomirov との共著)において距離空間の ε-エントロピーを定義した。これらの仕事は Kolmogorov の情報理論に関する一連の(1950年代後半から現在に至る)仕事の一つであった。Kolmogorov による ε-エントロピーの導入は、通信理論を背景とし、情報量を用いて定義されるものであって、Kolmogorov 以後の、A.Rényi [23] による定義、最近の E.C.Ponser、E.R.Rodemich、H.Rumsey [22] らによる定義などがエントロピーを直接用いているのと異っている。ところで,Kolmogorov [13] は,確率分布(有限次元空間または関数空間における)の ε-エントロピーの評価を、有限次元分布、ブラウン運動,拡散過程、正規定常過程などの例について、証明なしに与えた。情報量 $I(\xi,\eta)$ の $\xi=\eta$ の特別な場合としてのエントロピー $H(\xi)=I(\xi,\xi)$ —— 確率変数 ξ またはその分布の複雑さを表わす量と考えられる —— は多くの興味ある場合に値 $+\infty$ をとるが、ε-エントロピーは,ξ を"ε-近似"して ξ の複雑さを有限な量 —— ε と ξ の関数 —— として把握しようという思想にもとずく。Kolmogorov は上のような例でその具体的な関数形を求めているが,確率分布の ε-エントロピーについてはそれ以上の追求はしなかったようである。

Kolmogorov [13](または Gelfand-Kolmogorov-Jaglom [7])による上述の結果は、M.Pinsker [20] によって Gauss 過程の場合に証明され、n 次元の連続分布の場合が、Linikov [16] によって拡張された形で証明された。なお、Gelfand-Kolmogorov-Jaglom [7] にある n 重マルコフ

Sem. on Probab.
Vol.29 1968年
P1-95

Gauss過程の場合は K.Kazi〔11〕によって精密化されて解かれた。また、K.Kazi〔12〕は拡散過程の場合の評価を行った。これと Kolmogorov〔13〕による結果との間には若干の差があるが、これで Kolmogorov〔13〕（または〔7〕）の与えたすべての結果にほぼ証明がついたことになる。

このセミナーノートでは以上の結果の紹介を第一の目的としている（2章§2, §3, §4、3章§5, §6, §7、4章§9, §10）

第二に、多次元パラメーターのブラウン運動のε-エントロピーに関する、Y.Baba〔1〕の結果の紹介（3章§8）と、S.Ihara による、Gauss 過程の絶対連続性とε-エントロピーの関係についての結果（これによって、例えば、ブラウン運動と互に絶対連続な Gauss 過程のε-エントロピーはブラウン運動のε-エントロピーと漸近的に等しいことが示される）の紹介（5章§11）などを行う。

Gaussian でない場合（例えば K.Kazi〔12〕の拡散過程の場合）や、多次元パラメーターの Gauss 過程を取扱う必要と、叙述の一般化、簡明化のために、われわれは、2章§4、3章§5において、Pinsker〔20〕の結果（Gauss 過程のみを扱っているためやや見通しの悪い部分がある）をできるだけ一般の形で与えることにした。

なお、第一章では、このノート全体を通して用いられる情報量の定義とその性質を述べたが、後で使われない性質については述べなかった。また、附録では、このノートで主に取扱う Kolmogorov〔13〕によるε-エントロピーと他に定義されているε-エントロピーとの関係について若干の考察を与えた。

執筆の分担は次の通りである：　1章（井原）；2章§2, §3（加地），§4（馬場）；3章§5, §8（馬場），§6, §7（加地）；4章（井原）；5章（井原）；附録（馬場）。

内容や用語については、統一のとれたものにすべく努力したが、若干の重複や用語、記号の不統一があるかも知れないのでお許し願いたい。

最後に、このノートを作るにあたっては確率論セミナーのたくさん

Sem. on Probab.
Vol.29 1968年
P1-95

の方々に大変お世話になった。これの方々の有形，無形の援助なしには、このノートはできあがらなかったと思われる。われわれは、そのことに心から感謝したい。

1968年9月　　　　著者

Sem. on Probab.
Vol.29 1968年
P1-95

目　次

	はじめに	1
第1章	情報量	5
§1	情報量の定義と性質	5
第2章	ε-エントロピー	17
§2	ε-エントロピーの定義	17
§3	有限次元分布のε-エントロピー	18
§4	Gaussian system $\xi=(\xi_1,\xi_2,\cdots\cdots)$ のε-エントロピー	24
第3章	確率過程のε-エントロピー	36
§5	確率過程のε-エントロピー	36
§6	Diffusion process のε-エントロピー	42
§7	多重 Markov Gaussian process のε-エントロピー	49
§8	多次元パラメーターの Brown 運動のε-エントロピー	57
第4章	Stationary Gaussian process のε-エントロピー	65
§9	Stationary Gaussian process のε-エントロピー	65
§10	Stationary Gaussian process の単位時間当りのε-エントロピー	72
第5章	Gaussian process の絶対連続性, path の連続性との関係	78
§11	Gaussian process の絶対連続性, path の連続性との関係	78
附　録	ε-エントロピーの種々の定義	88
	引用文献	94

Sem. on Probab.
Vol.29 1968年
P1-95

第１章　情報量

§１．情報量の定義と性質

Shannon [28] は確率変数のあいまいさ(複雑さ)を表わす量として，エントロピーという量を定義した。これに対し，情報(他の確率変数)を与えるとあいまいさが減る。この減った量を情報を与えたことによって伝達した情報量と定義した。そしてその後，Kolmogorov 等がこれを数学的にきちんと定式化した。

ここでは情報量のもつ意味等にはあまり立ち入らず，Kolmogorov [13]，Pinsker [19]，Dobrushin [4] 等に従い，天下り的に情報量の定義を与え，そして ε-エントロピーの定義，計算に必要なことを中心に情報量の基本的性質を整理しておく。

$(X, \mathcal{B}_X, \mu_X)$，$(Y, \mathcal{B}_Y, \mu_Y)$ を2つの測度空間とし，ξ，η を確率空間 $(\Omega, \mathcal{B}, P)$ 上で定義され各々 $(X, \mathcal{B}_X)$，$(Y, \mathcal{B}_Y)$ の値をとる確率変数とする。P_ξ，P_η を各々 ξ，η の確率分布，$P_{\xi\eta}$ を ξ と η の結合分布，$P_\xi \times P_\eta$ を P_ξ と P_η の積とする。

$(X \times Y, \mathcal{B}_X \times \mathcal{B}_Y)$ の有限分割 $\{C_1, \cdots, C_n\}$ (i.e. $C_i \in \mathcal{B}_X \times \mathcal{B}_Y$，$C_1, \cdots, C_n$ は互いに disjoint，$\bigcup_{i=1}^{n} C_i = X \times Y$) に対し，

$$(1.1) \qquad I(\{C_i\}) = \sum_{i=1}^{n} P_{\xi\eta}(C_i) \log \frac{P_{\xi\eta}(C_i)}{P_\xi \times P_\eta (C_i)} \ ^{*)}$$

とおく。

確率変数 ξ と η の(間の)情報量 $I(\xi, \eta)$ を次式で定義する。

$$(1.2) \qquad I(\xi, \eta) = \sup I(\{A_i \times B_j \,;\, i=1, \cdots, m,\ j=1, \cdots, n\})$$

ここで上限は $(X, \mathcal{B}_X)$ の有限分割 $\{A_1, \cdots, A_m\}$ 及び $(Y, \mathcal{B}_Y)$ の有限

---

*) $\log$ の底は e とする。又 $0 \log \frac{0}{A} = 0 \ (A \geqq 0)$ とする。

Sem. on Probab.
Vol.29 1968年
P1-95

分割 $\{B_1, \cdots, B_n\}$ の全体についてとる。

この定義では上限をとる範囲を（$X \times Y$, $\mathcal{B}_X \times \mathcal{B}_Y$）の分割のうち特殊な分割，即ち矩形による分割のみに限っているが，次の定理が成り立ちこのことは何ら本質的でないことがわかる。

<u>定理 1. 1.</u>（Dobrushin [4]）

$\mathcal{B}_X \times \mathcal{B}_Y$ の subalgebra $\mathfrak{A}$ を $\mathfrak{A}$ を含む最小の σ-algebra は $\mathcal{B}_X \times \mathcal{B}_Y$ であるようなものとし，（$X \times Y$, $\mathfrak{A}$）の有限分割の全体を $D(\mathfrak{A})$ とする。$\mathcal{R} \subset D(\mathfrak{A})$ を $D(\mathfrak{A})$ に属する任意の分割 $\{D_i\}$ に対して $\{D_i\}$ の細分で $\mathcal{R}$ に属する分割 $\{C_j\}$ が存在するようなものとする。このとき，

(1, 3) $$I(\xi, \eta) = \sup_{\{C_i\} \in \mathcal{R}} I(\{C_i\})$$

<u>Remark</u> この定理で，特に $\mathcal{R} = \mathfrak{A} = \mathcal{B}_X \times \mathcal{B}_Y$ とおけば，情報量の定義を（1. 2）の代りに次のようにしてよいことがわかる。

(1, 2′) $$I(\xi, \eta) = \sup I(\{C_i\})$$

ここで上限は（$X \times Y$, $\mathcal{B}_X \times \mathcal{B}_Y$）の有限分割 $\{C_i\}$ の全体についてとる。

定理 1. 1. を証明するために補題を 2 つ用意する。

<u>補題 1. 1.</u>

$\{C_i\}$, $\{D_i\}$ は（$X \times Y$, $\mathcal{B}_X \times \mathcal{B}_Y$）の有限分割で $\{C_i\}$ は $\{D_i\}$ の細分とすると、

(1. 4) $$I(\{C_i\}) \geqq I(\{D_i\})$$

<u>証明</u> (1 4) は任意の非負数 $r_1, \cdots, r_m$, $s_1, \cdots, s_m$ に対して成り立つ不等式（帰納法で容易に示せる。）

(1. 5) $$\sum_{i=1}^{m} r_i \log \frac{r_i}{s_i} \geqq (r_1 + \cdots + r_m) \log \frac{r_1 + \cdots + r_m}{s_1 + \cdots + s_m}$$

を用いて容易に証明される。

<u>補題 1. 2.</u>

$(Z, \mathcal{B}_Z)$ を可測空間，$\mathcal{B}_Z$ の subalgebra $\mathfrak{A}$ を $\mathfrak{A}$ を含む最小の σ–

Sem. on Probab.
Vol.29 1968年
P1-95

algebra が $\mathcal{B}_Z$ であるようなものとする。μ_1, μ_2 を共に $(Z, \mathcal{B}_Z)$ 上の測度とするとき，$^\forall\varepsilon > 0$ と $^\forall D \in \mathcal{B}_Z$ に対し，次のような $A \in \mathfrak{A}$ が存在する。

(1.6)　　$\mu_1(A \ominus D) \leqq \varepsilon$　かつ　$\mu_2(A \ominus D) \leqq \varepsilon$

($A \ominus D$ は A と D の対称差集合)

<u>証明</u>　$\tilde{\mathfrak{A}} = \{D \in \mathcal{B}_Z ; {}^\forall\varepsilon > 0$ に対し (1.6) をみたす $A \in \mathfrak{A}$ が存在$\}$ とおくと，明らかに $\tilde{\mathfrak{A}} \supset \mathfrak{A}$。従って $\tilde{\mathfrak{A}}$ の単調列の極限が再び $\tilde{\mathfrak{A}}$ に属することを示せば $\tilde{\mathfrak{A}} = \mathcal{B}_Z$ となり、補題の成立がわかる。

$D_1 \supset D_2 \supset \cdots\cdots$，$\bigcap_n D_n = D$，$D_n \in \tilde{\mathfrak{A}}$ とすると十分大きな n に対しては $\mu_1(D_n \ominus D) \leqq \frac{\varepsilon}{2}$，$\mu_2(D_n \ominus D) \leqq \frac{\varepsilon}{2}$，又この D_n に対し次の様な $A \in \mathfrak{A}$ が存在する。$\mu_1(D_n \ominus A) \leqq \frac{\varepsilon}{2}$，$\mu_2(D_n \ominus A) \leqq \frac{\varepsilon}{2}$。従って $\mu_1(D \ominus A) \leqq \varepsilon$ かつ $\mu_2(D \ominus A) \leqq \varepsilon$，即ち $D \in \tilde{\mathfrak{A}}$ である。

<u>定理 1.1 の証明</u>　(1.3) の右辺を $I_{\mathcal{R}}(\xi, \eta)$，(1.2′) の右辺を $\tilde{I}(\xi, \eta)$ とおく。特に $\mathcal{R} = \mathfrak{A}$ = 矩形による有限分割の全体としたとき情報量の定義より $I_{\mathcal{R}}(\xi, \eta) = I(\xi, \eta)$ だから，結局定理の証明のためには $I_{\mathcal{R}}(\xi, \eta) = I(\xi, \eta)$ を示せば十分である。

$P_\xi \times P_\eta(C) = 0$，　$P_{\xi\eta}(C) > 0$ なる $C \in \mathcal{B}_X \times \mathcal{B}_Y$ が存在するときには容易に $\tilde{I}(\xi, \eta) = \infty = I_{\mathcal{R}}(\xi, \eta)$ がわかる。補題 1.1 より常に $\tilde{I}(\xi, \eta) \geqq I_{\mathcal{R}}(\xi, \eta)$ だから，$I(\{C_i\}) < \infty$ なる $(X \times Y, \mathcal{B}_X \times \mathcal{B}_Y)$ の有限分割 $\{C_1, \cdots, C_m\}$ と $^\forall\varepsilon > 0$ に対し，

(1.7)　　$I(\{D_i\}) \geqq I(\{C_i\}) - \varepsilon$

をみたす分割 $\{D_i\} \in \mathcal{R}$ が存在することを言えばよい。補題 1.2 より $^\forall\delta > 0$ に対し次の様な $\tilde{C}_i \in \mathfrak{A}$ $(i = 1, \cdots, m)$ が存在する。

$$P_{\xi\eta}(C_i \ominus \tilde{C}_i) \leqq \delta, \quad P_\xi \times P_\eta(C_i \ominus \tilde{C}_i) \leqq \delta \quad (i = 1, \cdots, m)$$

そこで，$\tilde{D}_1 = \tilde{C}_1$，$\tilde{D}_j = \tilde{C}_j - \bigcup_{i<j} \tilde{D}_i$ $(j = 2, \cdots, m)$，$\tilde{D}_{m+1} = X \times Y - \bigcup_{j=1}^{m} \tilde{D}_j$ とおくと．$\{\tilde{D}_1, \cdots, \tilde{D}_{m+1}\} \in D(\mathfrak{A})$ であり，$P_{\xi\eta}(\tilde{D}_j \ominus C_j) \leqq \delta m$，$P_\xi \times P_\eta(\tilde{D}_j \ominus C_j) \leqq \delta m$ $(j = 1, \cdots, m)$，$P_{\xi\eta}(\tilde{D}_{m+1}) \leqq \delta m$，$P_\xi \times P_\eta(\tilde{D}_{m+1}) \leqq \delta m$ である。δ は任意だから $I(\{C_1, \cdots, C_m\})$ と $I(\{\tilde{D}_1, \cdots, \tilde{D}_m\})$ は任意に近

Sem. on Probab.
Vol.29 1968年
P1-95

くできる。一方

$$P_{\xi\eta}(\tilde{D}_{m+1})\log\frac{P_{\xi\eta}(\tilde{D}_{m+1})}{P_\xi\times P_\eta(\tilde{D}_{m+1})}\geqq P_{\xi\eta}(\tilde{D}_{m+1})\log P_{\xi\eta}(\tilde{D}_{m+1})$$

$$\geqq m\delta\log m\delta$$

だから，δを十分小さくすれば(1.7)をみたすように$\{\tilde{D}_j\}\in D(\mathfrak{A})$がとれる．$\{D_j\}\in\mathcal{R}$ を $\{\tilde{D}_j\}$の細分とすると，補題1.I より$\{D_j\}$は(1.7)をみたす。

確率変数ξのエントロピー$H(\xi)$を次式で定義する。

(1.8) $H(\xi)=I(\xi,\xi)$

エントロピーは，例えばξが連続分布をもてば常に $H(\xi)=\infty$ となってしまう。もし $H(\xi)<\infty$, $H(\eta)<\infty$ ならば

(1.9) $I(\xi,\eta)=H(\xi)+H(\eta)-H((\xi,\eta))$

が成り立つことが容易にわかる。

次に，確率変数ξの *differential entropy* を定義しよう。

P_ξが *density* p_ξ をもつとき

(1.10) $h(\xi)=-\int_X p_\xi(x)\log p_\xi(x)\,d\mu_X(x)$

をξの *differential entropy* という。

実際に情報量を計算するときには次の定理が重要な役割を果す。

<u>定理 1.2</u> (Gelfand, Yaglom, Perez)

もし $P_{\xi\eta}\prec P_\xi\times P_\eta$ *) でなければ，$I(\xi,\eta)=\infty$

もし $P_{\xi\eta}\prec P_\xi\times P_\eta$ ならば

(1.11) $I(\xi,\eta)=\iint_{X\,Y}\log a(x,y)\,dP_{\xi\eta}(x,y)$

ここで

$$a(x,y)=\frac{dP_{\xi\eta}(x,y)}{dP_\xi\times P_\eta(x,y)}$$

特にP_ξ, P_η, $P_{\xi\eta}$ が *density* p_ξ, p_η, $p_{\xi\eta}$ をもつとき (1.11) は

*) $\mu\prec\nu$ はμがνに関し絶対連続であることを示す。

Sem. on Probab.
Vol.29 1968年
P1-95

$$(1.11')\quad I(\xi,\eta) = \int_X\int_Y p_{\xi\eta}(x,y)\log\frac{p_{\xi\eta}(x,y)}{p_\xi(x)\,p_\eta(y)}\,d\mu_X(x)\,d\mu_Y(y)$$

となる。さらに η に対する ξ の条件つき確率が *density* $p_{\xi|\eta}(x|y)$ をもつとき

$$(1.12)\quad h(\xi|\eta) = -\int_X p_{\xi|\eta}(x|y)\log p_{\xi|\eta}(x|y)\,d\mu_X(x)$$

とおく（これを条件つき *differential entropy* という）と

$$(1.13)\quad I(\xi,\eta) = h(\xi) - E_\eta h(\xi|\eta)$$

である。

定理の証明のために次の補題を用意する。

<u>補題 1.3</u>

$F(x)$ を $[0,\infty)$ 上の分布関数とすると

$$\int_0^\infty x\log x\,dF(x) \geqq \int_0^\infty x\,dF(x)\,\log\left(\int_0^\infty x\,dF(x)\right)$$

<u>証明</u>　$x\log x$ に Jensen の不等式を適用することによって証明される。

<u>定理 1.2 の証明</u>　$P_{\xi\eta} \prec P_\xi\times P_\eta$ でないとすると，$P_{\xi\eta}(A)>0$，$P_\xi\times P_\eta(A)=0$ なる $A\in\mathcal{B}_X\times\mathcal{B}_Y$ が存在する。従ってこのとき，$I(\xi,\eta) \geqq I(\{A,A^c\}) = \infty$。

$P_{\xi\eta}\prec P_\xi\times P_\eta$ のとき，$P_\xi\times P_\eta(B)>0$ なる $B\in\mathcal{B}_X\times\mathcal{B}_Y$ を固定し

$$F_B(u) = P_\xi\times P_\eta(a(x,y)\leqq u\,|\,B) = \frac{P_\xi\times P_\eta(\{a(x,y)\leqq u\}\cap B)}{P_\xi\times P_\eta(B)}$$

とおくと

$$\int_0^\infty u\,dF_B(u) = \frac{1}{P_\xi\times P_\eta(B)}\int_0^\infty u\,dP_\xi\times P_\eta(a^{-1}(u)\cap B)$$

$$= \frac{1}{P_\xi\times P_\eta(B)}\int_B a(x,y)\,dP_\xi\times P_\eta(x,y) = \frac{P_{\xi\eta}(B)}{P_\xi\times P_\eta(B)}$$

同様に

$$\int_0^\infty u\log u\,dF_B(u) = \frac{1}{P_\xi\times P_\eta(B)}\int_B \log a(x,y)\,dP_{\xi\eta}(x,y)$$

従って補題 1.3 より

Sem. on Probab.
Vol.29 1968年
P1-95

$$\int_0^\infty \log a(x,y)\,dP_{\xi\eta}(x,y) \geqq P_{\xi\eta}(B)\log\frac{P_{\xi\eta}(B)}{P_\xi\times P_\eta(B)}$$

上式は $P_\xi\times P_\eta(B)=0$ のときも 右辺＝左辺＝0 で成立。従って（$X\times Y$, $\mathcal{B}_X\times\mathcal{B}_Y$）の任意の分割 $\{C_i\}$ に対して

$$\sum_i \int_{C_i} \log a(x,y)\,dP_{\xi\eta}(x,y) \geqq I(\{C_i\})$$

次に逆向きの不等式を示す。$M\to\infty$ のとき $P_{\xi\eta}((x,y);|\log a(x,y)|>M)\to 0$ だから，$^\forall\varepsilon>0$ に対しある M_0 が存在し，$M>M_0$ ならば

$$0 \geqq P_{\xi\eta}(|\log a(x,y)|>M)\log P_{\xi\eta}(|\log a(x,y)|>M) > -\varepsilon$$

$M(>M_0)$ を固定し，$\{(x,y);|\log a(x,y)|\leqq M\}=\bigcup_{i=1}^n C_i$ (*disjoint*) とし，$g_i=\inf_{(x,y)\in C_i} a(x,y)\geqq e^{-M}$，$h_i=\sup_{(x,y)\in C_i} a(x,y) \leqq e^{-M}$ とおくと，

$$P_{\xi\eta}(C_i)\log h_i \geqq \int_{C_i} \log a(x,y)\,dP_{\xi\eta}(x,y) \geqq P_{\xi\eta}(C_i)\log g_i$$

$$P_{\xi\eta}(C_i)\log h_i \geqq P_{\xi\eta}(C_i)\log\frac{P_{\xi\eta}(C_i)}{P_\xi\times P_\eta(C_i)} \geqq P_{\xi\eta}(C_i)\log g_i$$

が成り立つ。$\{C_i\}$ を $\log h_i - \log g_i < \varepsilon$ をみたすようにとれば

$$\left|\sum_{i=1}^\infty P_{\xi\eta}(C_i)\log\frac{P_{\xi\eta}(C_i)}{P_\xi\times P_\eta(C_i)} - \sum_{i=1}^n \int_{C_i}\log a(x,y)\,dP_{\xi\eta}(x,y)\right| < \varepsilon$$

$C_0=\{(x,y);|\log a(x,y)|>M\}$ とおけば，

$$P_{\xi\eta}(C_0)\log\frac{P_{\xi\eta}(C_0)}{P_\xi\times P_\eta(C_0)} \geqq P_{\xi\eta}(C_0)\log P_{\xi\eta}(C_0) > -\varepsilon$$

以上より

$$I(\xi,\eta) \geqq I(\{C_0,C_1,\cdots,C_n\}) > \int_{(|\log a(x,y)|\leqq M)} \log a(x,y)\,dP_{\xi\eta}(x,y) - 2\varepsilon$$

従って，$\varepsilon\to 0$，$M\to\infty$ として結論を得る。

(1.11′) は (1.11) から (1.13) は (1.11′) から容易に得られる。

ここで情報量のおもな性質をまとめておこう。

<u>定理 1.3.</u>

Sem. on Probab.
Vol.29 1968年
P1-95

(I) $0 \leqq I(\xi, \eta) \leqq \infty$

$I(\xi,\eta) = 0$ となるのは ξ と η が独立のとき，ただそのときのみである。

(II) $I(\xi,\eta) = I(\eta,\xi)$

(III) f を $(Y, \mathcal{B}_Y)$ から $(Z, \mathcal{B}_Z)$ への可測関数とすると，

$$I(\xi, f(\eta)) \leqq I(\xi,\eta)$$

ただし，$f(\eta)(\omega) = f(\eta(\omega))$

さらにもし f が $1:1$ でかつ f^{-1} も可測関数ならば

$$I(\xi, f(\eta)) = I(\xi,\eta)$$

(IV) $I((\xi,\eta),\zeta) + I(\xi,\eta) = I(\xi,(\eta,\zeta)) + I(\eta,\zeta)$

(IV′) $I((\xi_1,\xi_2),(\eta_1,\eta_2)) + I(\xi_1,\xi_2) + I(\eta_1,\eta_2)$

$= I((\xi_1,\eta_1),(\xi_2,\eta_2)) + I(\xi_1,\eta_1) + I(\xi_2,\eta_2)$

(IV″) ξ_1 と ξ_2 とが独立ならば，

$$I((\xi_1,\xi_2),(\eta_1,\eta_2)) \geqq I(\xi_1,\eta_1) + I(\xi_2,\eta_2)$$

上式で等式が成り立つのは (ξ_1,η_1) と (ξ_2,η_2) が独立のとき，ただそのときのみである。

(V) 確率変数の列 ξ, η, ζ が Markov Chain なすならば，

$$I((\xi,\eta),\zeta) = I(\eta,\zeta),\quad I(\xi,(\eta,\zeta)) = I(\xi,\eta)$$

逆に $I((\xi,\eta),\zeta) = I(\eta,\zeta) < \infty$ (又は $I(\xi,(\eta,\zeta)) = I(\xi,\eta) < \infty$) ならば，$\xi, \eta, \zeta$ は Markov chain をなす。

(V′) ξ と ζ が独立ならば

$$I(\xi,(\eta,\zeta)) = I(\xi,\eta)$$

(V″) $\xi_1, \xi_2, \xi_3, \xi_4$ が Markov chain をなすならば

$$I(\xi_1,\xi_4) \leqq I(\xi_2,\xi_3)$$

(VI) $\xi = (\xi_1, \xi_2, \cdots)$, $\eta = (\eta_1, \eta_2, \cdots)$ のとき

$$I(\xi,\eta) = \lim_{n,m\to\infty} I((\xi_1,\cdots,\xi_n),(\eta_1,\cdots,\eta_m))$$

(VII) $\xi = (\xi_1, \xi_2, \cdots)$, ξ_i $(i = 1, 2, \cdots)$ は互いに独立，$\eta = (\eta_1, \eta_2, \cdots)$ のとき

$$I(\xi,\eta) \geqq \sum_{i=1}^{\infty} I(\xi_i,\eta_i)$$

Sem. on Probab.
Vol.29 1968年
P1-95

上式で等号が成り立つのは (ξ_i, η_i) $(i=1,2,\cdots)$ が互いに独立のとき，ただそのときのみである。

(VIII) $\xi, \xi_1, \xi_2, \cdots$ は各々 $(X, \mathcal{B}_X)$, $(X_1, \mathcal{B}_{X_1})$, $(X_2, \mathcal{B}_{X_2})\cdots$ の値をとり $(\xi_n, \eta_n) \to (\xi, \eta)$ (in law) ならば

$$I(\xi, \eta) \leqq \varliminf_{n\to\infty} I(\xi_n, \eta_n)$$

証明 (I) $I(\xi,\eta) \geqq 0$ は，定義と (1.5) より明らか。

(II), (III) 定義より明らか。

(IV) 定義から結局 ξ, η, ζ が有限個の値をとる確率変数の場合に証明すれば十分である。このとき (1.9) が適用でき，

$$I((\xi,\eta),\zeta) + I(\xi,\eta) = H((\xi,\eta)) + H(\zeta) - H((\xi,\eta,\zeta))$$
$$+ H(\xi) + H(\eta) - H((\xi,\eta))$$
$$I(\xi,(\eta,\zeta)) + I(\eta,\zeta) = H(\xi) + H((\eta,\zeta)) - H((\xi,\eta,\zeta))$$
$$+ H(\eta) + H(\zeta) - H((\eta,\zeta))$$

上の2式を較べて (IV) を得る。

(IV′) は (IV) から，(IV″) は (IV′) から容易に導ける。

(V) 前半，ξ, η, ζ は各々可測空間 $(X, \mathcal{B}_X)$, $(Y, \mathcal{B}_Y)$ $(Z, \mathcal{B}_Z)$ の値をとるものとする。ξ, η, ζ が Markov chain をなすことから，$\forall A \in \mathcal{B}_X$, $\forall B \in \mathcal{B}_Y$, $\forall C \in \mathcal{B}_Z$ に対し

$$P(\xi \in A, \zeta \in C \mid \eta) = P(\xi \in A \mid \eta) P(\zeta \in C \mid \eta)$$

従って $\forall A \in \mathcal{B}_X$, $\forall B \in \mathcal{B}_Y$, $\forall C \in \mathcal{B}_Z$ に対し

$$(1.14) \qquad P_{\xi\eta\zeta}(A\times B\times C) = \int_B P(\xi\in A, \zeta\in C \mid \eta = y)\, dP_\eta(y)$$
$$= \int_B P(\xi \in A \mid \eta = y) P(\zeta \in C \mid \eta = y)\, dP_\eta(y)$$

常に，$I((\xi,\eta),\zeta) \geqq I(\eta,\zeta)$ だから，$I(\eta,\zeta) < \infty$ のときに証明すればよい。このとき定理 1.2 より，$P_{\xi\eta} \prec P_\xi \times P_\eta$，そこで

$$a_{\eta\zeta}(y,z) = \frac{dP_{\eta\zeta}(y,z)}{dP_\eta\times P_\zeta(y,z)}$$

とおく。ここで一般に次のことが成り立つことに注意する。任意の $A \in \mathcal{B}_X$

Sem. on Probab.
Vol.29 1968年
P1-95

と任意の$\mathcal{B}_X \times \mathcal{B}_Z$-可測関数$f(x,z)$に対し

(1.15) $$\int_{A\times Y\times Z} f(y,z)\, dP_{\xi\eta}\times P_\zeta(x,y,z)$$
$$= \int_{Y\times Z} f(y,z)\, P(\xi\in A\mid \eta=y)\, dP_\eta\times P_\zeta(y,z)$$

(1.15)で特に

$$f(y,z)=\begin{cases} a_{\eta\zeta}(y,z) & \text{if } (y,z)\in B\times C \\ 0 & \text{if } (y,z)\notin B\times C \end{cases}$$
$$(B\in\mathcal{B}_Y,\quad C\in\mathcal{B}_Z)$$

とおくと

(1.16) $$\int_{A\times B\times C} a_{\eta\zeta}(y,z)\, dP_{\xi\eta}\times P_\zeta(x,y,z)$$
$$= \int_{B\times C} a_{\eta\zeta}(y,z)\, P(\xi\in A\mid\eta=y)\, dP_\eta\times P_\zeta(y,z)$$
$$= \int_{B\times C} P(\xi\in A\mid\eta=y)\, dP_{\eta\zeta}(y,z)$$

又(1.15)でξとζを入れかえ

$$f(y,z)\equiv f(y)=\begin{cases} P(\xi\in A\mid\eta=y) & \text{if } y\in B \\ 0 & \text{if } y\notin B \end{cases}$$
$$(B\in\mathcal{B}_Y)$$

とおくと

(1.17) $$\int_{B\times C} P(\xi\in A\mid\eta=y)\, dP_{\eta\zeta}(y,z)$$
$$= \int_B P(\xi\in A\mid\eta=y)\, P(\zeta\in C\mid\eta=y)\, dP_\eta(y)$$

従って(1.16), (1.17), (1.14)より

$$\int_{A\times B\times C} a_{\eta\zeta}(y,z)\, dP_{\xi\eta}\times P_\zeta(x,y,z) = P_{\xi\eta\zeta}(A\times B\times C)$$

故に, $P_{\xi\eta\zeta} \prec P_{\xi\eta}\times P_\zeta$ でかつ

$$\frac{dP_{\xi\eta\zeta}(x,y,z)}{dP_{\xi\eta}\times P_\zeta(x,y,z)} = a_{\eta\zeta}(y,z)$$

である。 従って定理 1.2 より

$$I((\xi,\eta),\zeta) = \int_{X\times Y\times Z} \log a_{\eta\zeta}(y,z)\, dP_{\xi\eta\zeta}(x,y,z)$$
$$= \int_{Y\times Z} \log a_{\eta\zeta}(y,z)\, dP_{\eta\zeta}(y,z) = I(\zeta,\eta)$$

Sem. on Probab.
Vol.29 1968年
P1-95

このことより (IV) より、 $I(\xi,(\eta,\zeta)) = I(\xi,\eta)$

後半は、あとで使わないので証明は省略する。

(V′). ξ と ζ が独立なら, ξ,η,ζ は Markov Chain 故 (V) より明らか。

(V″). (III) より

$$I((\xi_1,\xi_2),(\xi_3,\xi_4)) \geqq I(\xi_1,(\xi_3,\xi_4)) \geqq I(\xi_1,\xi_4)$$

一方 (V) より

$$I((\xi_1,\xi_2),(\xi_3,\xi_4)) = I(\xi_2,(\xi_3,\xi_4)) = I(\xi_2,\xi_3)$$

(VI) 次の式を証明すればよい。

(1.18) $\quad I(\xi,(\eta_1,\cdots,\eta_m)) \uparrow I(\xi,\eta) \quad \text{as } m\to\infty$

(III) より明らかに

(1.19) $\quad I(\xi,\eta) \geqq I(\xi,(\eta_1,\cdots,\eta_{m+1})) \geqq I(\xi,(\eta_1,\cdots,\eta_m))$

$\xi,\eta_1,\eta_2,\cdots$ の値域を各々 $(X,\mathcal{B}_X),(Y_1,\mathcal{B}_{Y_1}),(Y_2,\mathcal{B}_{Y_2}),\cdots$ とする。 $\mathfrak{A}$ を次の形の集合の有限個の和からなる algebra とする。

$$E\times F_1\times F_2\times\cdots\ (E\in\mathcal{B}_X,\ F_j\in\mathcal{B}_{Y_j},\ \text{有限個の } j \text{ を除き } F_j = Y_j)$$

$\mathcal{R}$ を上の集合による有限分割の全体とする。 このとき明らかに $\mathfrak{A}$ は σ-algebra $\mathcal{B}_X\times\mathcal{B}_Y$ ($\mathcal{B}_Y=\mathcal{B}_{Y_1}\times\mathcal{B}_{Y_2}\times\cdots$) を生成し, $\mathcal{R}$ は定理1.1の仮定をみたす。 従って

$$I(\xi,\eta) = \sup_{\{C_i\}\in\mathcal{R}} I(\{C_i\})$$

ところで $\{C_i\}\in\mathcal{R}$ を1つ固定すると, ある m が存在し

$$I(\{C_i\}) \leqq I(\xi,(\eta_1,\cdots,\eta_m))$$

故に

$$I(\xi,\eta) \leqq \varliminf_{m\to\infty} I(\xi,(\eta_1,\cdots,\eta_m))$$

従って (1.19) と合せ (1.18) を得る。

(VII), (VI) と (V″) より明らか。

(VIII), 仮定より ${}^{\forall}C\in\mathcal{B}_X\times\mathcal{B}_Y$ に対して

Sem. on Probab.
Vol.29 1968年
P1-95

$$\lim_{n\to\infty} P_{\xi_n \eta_n}(C) = P_{\xi\eta}(C)$$

従って，$(X\times Y, \mathcal{B}_X \times \mathcal{B}_Y)$ の任意の有限分割 $\{C_i\}$ に対して

$$\lim_{n\to\infty} \sum_i P_{\xi_n\eta_n}(C_i) \log \frac{P_{\xi_n\eta_n}(C_i)}{P_{\xi_n}\times P_{\eta_n}(C_i)}$$

$$= \sum_i P_{\xi\eta}(C_i) \log \frac{P_{\xi\eta}(C_i)}{P_\xi \times P_\eta(C_i)}$$

故に

$$I(\xi, \eta) \leqq \varliminf_{n\to\infty} I(\xi_n, \eta_n)$$

次に，Gaussian random variable の場合に，情報量を具体的に求める式を導く。

定理 1.4

ξ, η は (ξ, η) が Gaussian である，各々 k, l 次元 Gaussian random variable とし，ξ, η, (ξ, η) の covariance matrix を各々 A, B, C とする。$\tilde{A}$, $\tilde{B}$ を各々 $|\tilde{A}| \equiv \det \tilde{A} \neq 0$, $|\tilde{B}| \neq 0$ なる，A, B の小行列式の中で最高次（k', l'次とする）のもの，$\tilde{C}$ を $\tilde{A}$, $\tilde{B}$ を含む $k'+l'$次の C の小行列式とする。

このとき

$$(1.20) \qquad I(\xi, \eta) = \begin{cases} \dfrac{1}{2} \log \dfrac{|\tilde{A}||\tilde{B}|}{\tilde{C}} & \text{if } |\tilde{C}| \neq 0 \\ \infty & \text{if } |\tilde{C}| = 0 \end{cases}$$

特に，$k = l = 1$ のときには

$$(1.21) \qquad I(\xi, \eta) = -\frac{1}{2} \log(1 - r^2(\xi, \eta))$$

ここで $r(\xi, \eta)$ は ξ, と η の相関係数。

証明　一般性を失うことなく $E\xi = E\eta = 0$ としてよい。

(I) $|A||B| \neq 0$ (i.e $\tilde{A} = A$, $\tilde{B} = B$) の場合

$|C| \neq 0$ ならば，P_ξ, P_η, $P_{\xi\eta}$ は各々 density

$$p_\xi(x) = (2\pi)^{-\frac{k}{2}} |A|^{-\frac{1}{2}} \exp\{-\tfrac{1}{2}(A^{-1}x, x)\}$$

$$p_\eta(y) = (2\pi)^{-\frac{l}{2}} |B|^{-\frac{1}{2}} \exp\{-\tfrac{1}{2}(B^{-1}y, y)\}$$

Sem. on Probab.
Vol.29 1968年
P1-95

$$p_{\xi\eta}(z) = (2\pi)^{-\frac{k+l}{2}} |C|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(C^{-1}z, z)\right\} \quad (z = (x, y))$$

をもつ。従って定理 1.2 より

$$I(\xi, \eta) = \int_{R^k}\int_{R^l} \left[\frac{1}{2}\log\frac{|A||B|}{|C|} - \frac{1}{2}\left\{(C^{-1}z, z) - (A^{-1}x, x) - (B^{-1}y, y)\right\}\right] dP_{\xi\eta}(x, y)$$

$$= \frac{1}{2}\log\frac{|A||B|}{|C|} - \frac{1}{2}(k+l) + \frac{k}{2} + \frac{l}{2} = \frac{1}{2}\frac{|A||B|}{|C|}$$

$|C| = 0$ のときは，$P_{\xi\eta} \prec P_\xi \times P_\eta$ でないから定理 1.2 より $I(\xi, \eta) = \infty$

(II) $|A||B| = 0$ の場合

このときは，容易に (I) の場合に帰着できる。

次に $k = l = 1$ のときには，$E\xi^2 = \sigma_\xi^2$，$E\eta^2 = \sigma_\eta^2$ とおくと，

$$A = \sigma_\xi^2, \quad B = \sigma_\eta^2, \quad C = \begin{pmatrix} \sigma_\xi^2 & \sigma_\xi\sigma_\eta r(\xi, \eta) \\ \sigma_\xi\sigma_\eta r(\xi, \eta) & \sigma_\eta^2 \end{pmatrix}$$

従って，(1.20) より

$$I(\xi, \eta) = -\frac{1}{2}\log\left(1 - r^2(\xi, \eta)\right)$$

Sem. on Probab.
Vol.29 1968年
P1-95

第 2 章 ε-エントロピー

§2 ε-エントロピーの定義

以下では距離 $\rho(x, y)$, $(x, y \in X)$ の入った距離空間 $(X, \mathcal{B}_X)$ の値をとる確率変数 $\xi = \xi(\omega)$, $\omega \in \Omega$ を問題とする。

<u>定義</u> 正数 ε に対して

$$H_\varepsilon(\xi) = \inf_{\dot{\xi}} \left\{ I(\xi, \dot{\xi}) ; \begin{array}{l} \dot{\xi} = \dot{\xi}(\omega) \text{ は } X \text{ の値をとる確率変数} \\ E\rho(\xi, \dot{\xi})^2 \leqq \varepsilon^2 \end{array} \right\}$$

で定義される量を"確率変数 ξ の ε-エントロピー"と呼ぶ。

<u>*Remark*</u> 例えば確率過程 $\xi = \xi(t, \omega)$ の場合には、定理 5.1、の脚註にも述べてあるように、定義式の $\inf$ をとる範囲に、「$\xi = \xi(t, \omega)$ が (t, ω)-可測である」というような、自然な制限を加えることもある。

ε-エントロピーの概念は、すでに *Shannon* [28] に *transmission rate* という言葉で基本的なアイデアが述べられていたが、これをいろいろな $\xi = \xi(\omega)$ に対して計算を実行して、ξ の特性量としての意義を強調したのは *Kolmogorov* である。ここでわれわれは $\varepsilon \to 0$ の時の $H_\varepsilon(\xi)$ の漸近的動向に興味を持つ。

距離 ρ のとり方はそれぞれの X に対していろいろ考えられるが、われわれは $\inf$ をとる条件 $E\rho(\xi, \dot{\xi})^2 \leqq \varepsilon^2$ が二次のモーメントに関する制約となるように定めるのが普通である。このノートで主として取り扱うのは X と ρ が以下のような場合である。

Sem. on Probab.
Vol.29 1968年
P1-95

(1) n次元確率変数 : $\xi=(\xi_1,\cdots,\xi_n)$ (2章, §3)

$$X=R^n,\quad \rho(x,y)=\left\{\sum_{i=1}^{n}(x_i-y_i)^2\right\}^{\frac{1}{2}}$$

$$x=(x_1,\cdots,x_n),\quad y=(y_1,\cdots,y_n)$$

(2) l^2の確率変数 : $\xi=(\xi_1,\xi_2,\cdots)$, $\sum_{i=1}^{\infty}E\xi_i^2<\infty$ (2章, §4)

$$X=l^2,\quad \rho(x,y)=\left\{\sum_{i=1}^{\infty}(x_i-y_i)^2\right\}^{\frac{1}{2}}$$

$$x=(x_1,x_2,\cdots),\quad y=(y_1,y_2,\cdots)$$

(3) 確率過程 : $\xi=\xi(t,\omega)$, $0\leqq t\leqq T$, $E\int_0^T|\xi(t,\omega)|^2dt<\infty$

(3章, §5)

$$X=L^2[0,T],\quad \rho(x,y)=\left\{\int_0^T|x(t)-y(t)|^2dt\right\}^{\frac{1}{2}}$$

$$x=x(t),\quad y=y(t),\quad 0\leqq t\leqq T$$

§ 3、 n次元確率変数の ε-エントロピー

ここでは、n次元実数値確率変数 $\xi=(\xi_1,\cdots,\xi_n)\in R^n$ の ε-エントロピーをユークリッド距離 $\rho(x,y)=|x-y|=\left\{\sum_{i=1}^{n}(x_i-y_i)^2\right\}^{\frac{1}{2}}$ に基づいて計算する。これを最初に厳密に証明したのは *Linikov* [16] である。そこには以下のような仮定よりもっと緩かな条件の下で一般的な距離*) に対して結果を導いているが、ここでは計算の簡明さの故に、*Geerish and*

---

*) 例えば $\rho(x,y)=\left\{\sum_{i=1}^{n}(x_i-y_i)^{\alpha}\right\}^{\beta}$ $(\alpha,\beta>0)$ など.

Sem. on Probab.
Vol.29 1968年
P1-95

Schultheiss [6] にならって見やすく書きかえることにする。

定理 3.1.

$\xi=(\xi_1,\cdots,\xi_n)$ は有界で連続な密度関数 $p_\xi(x)=p_\xi(x_1,\cdots,x_n)$ をもち（従ってその differential entropy は

$$h(\xi)=-\int_{R^n}p_\xi(x)\log p_\xi(x)\,dx>-\infty$$

となる），二次のモーメント $\sigma^2=E|\xi|^2=E\sum_{i=1}^{n}|\xi_i|^2<\infty$ とする。

このとき、

$$H_\varepsilon(\xi)=n\log\frac{1}{\varepsilon}+\frac{n}{2}\log n-n\log\sqrt{2\pi e}+h(\xi)+o(1)$$

証明　$H_\varepsilon(\xi)$ の評価を上と下から行なう。

$H_\varepsilon(\xi)=\inf_{\dot\xi}\{I(\xi,\dot\xi)\,;\,E|\xi-\dot\xi|\leqq\varepsilon^2\}$ の定義式で、確率変数の対 $(\xi,\dot\xi)$ に対して情報量 $I(\xi,\dot\xi)$ も平均距離 $E|\xi-\dot\xi|^2$ も同時分布が密度関数をもつ n 次元確率変数の対 $(\xi,\dot\xi)$ によって、任意の精度で近似できる.（2章, §4、補題 4.1 の証明参照）から、$H_\varepsilon(\xi)$ の計算には $(\xi,\dot\xi)$ の密度関数が存在するような $\dot\xi$ について inf を考えればよいことに注意する。

（第一段）下からの評価

$(\xi,\dot\xi)$ の密度関数が存在するような場合は、$\dot\xi$ に対する ξ の条件つき確率の密度関数 $p_{\xi|\dot\xi}(x|y)$ が存在するから，今のべた注意によって明らかなように

$$(3.1)\qquad H_\varepsilon(\xi)\geqq\inf_{\dot\xi}\left\{I(\xi,\dot\xi)\,;\ \begin{array}{l}\dot\xi\text{ に対する }\xi\text{ の条件つき確率の密}\\ \text{度関数 }p_{\xi|\dot\xi}(x|y)\text{ が存在し、}\\ E|\xi-\dot\xi|^2\leqq\varepsilon^2\end{array}\right\}$$

(3.1) の右辺の値（$\underline{H}$ とおく）を評価しよう．条件をみたすように $(\xi,\dot\xi)$ の対に対して

$$I(\xi,\dot\xi)=h(\xi)-E\,h(\xi|\dot\xi)$$

但し、

$$h(\xi,\dot\xi)=-\int_{R^n}p_{\xi|\dot\xi}(x|y)\log p_{\xi|\dot\xi}(x|y)\,dx$$

が成り立つから、

Sem. on Probab.
Vol.29 1968年
P1-95

(3.2) $\quad \underline{H} \geq h(\xi) - \sup_{\dot{\xi}} E h(\xi|\dot{\xi})$

よって $\sup_{\dot{\xi}} E h(\xi|\dot{\xi})$ を求める。まず

(3.3) $\quad E|\xi-\dot{\xi}|^2 = \int_{R^n}\int_{R^n}|x-y|^2 p_{\xi|\dot{\xi}}(x|y)\, p_{\dot{\xi}}(y)\,dx\,dy = \varepsilon^2$

の範囲で $\sup_{\dot{\xi}} E h(\xi|\dot{\xi})$ を求めてみよう。

<u>補題 3.1</u> *)

$$\sup_{\dot{\xi}}\{E h(\xi|\dot{\xi})\,;\,E|\xi-\dot{\xi}|^2=\varepsilon^2\} = -n\log\frac{1}{\varepsilon} - \frac{n}{2}\log n + n\log\sqrt{2\pi e},$$

ここで右辺の値は $\dot{\xi}$ に対する ξ の条件つき確率の密度関数が特に

$$g(x|y) = \frac{1}{\left(\frac{2\pi\varepsilon^2}{n}\right)^{\frac{n}{2}}}\exp\left\{-\frac{n}{2\varepsilon^2}|x-y|^2\right\}$$

で与えられる場合に達成される。

<u>証明</u> 条件つき確率の密度関数を上式の $g(x|y)$ で与えるような確率変数を ξ' とすると，明かに

$$E|\xi-\xi'|^2 = \varepsilon^2,$$

しかも (3.3) を満すような任意の $\dot{\xi}$ に対して

$$\int_{R^n} p_{\dot{\xi}}(x)\left[\int_{R^n} p_{\xi|\dot{\xi}}(x|y)\log g(x|y)\,dx\right]dy$$

$$= n\log\frac{1}{\varepsilon} + \frac{n}{2}\log n - n\log\sqrt{2\pi e} \quad (= E h(\xi|\xi'))$$

と計算できることに注意すれば，(3.3) をみたす任意の $\dot{\xi}$ に対して

$$E h(\xi|\dot{\xi}) + n\log\frac{1}{\varepsilon} + \frac{n}{2}\log n - n\log\sqrt{2\pi e}$$

$$= \int_{R^n} p_{\dot{\xi}}(y)\left[\int_{R^n} p_{\xi|\dot{\xi}}(x|y)\log\frac{g(x|y)}{p_{\xi|\dot{\xi}}(x|y)}\,dx\right]dy$$

$$\leq \int_{R^n} p_{\dot{\xi}}(y)\left[\int_{R^n}\left\{\frac{g(x|y)}{p_{\xi|\dot{\xi}}(x|y)}-1\right\}p_{\xi|\dot{\xi}}(x|y)\,dx\right]dy^{**)}$$

---

*) 以下の計算は「分散一定という制限の下で *differential entropy* が最大となるのは *Gaussian random variable* である」という事実を証明するのと本質的に同じである。

**) 不等式 $\log x \leq x-1$ $(x>0)$ を用いた。$x=1$ の時にのみ等号が成り立つ。

Sem. on Probab.
Vol.29 1968年
P1-95

$$= \iint p_{\xi'}(y)\, g(x|y)\, dx\, dy - \iint p_{\xi,\xi'}(x,y)\, dx\, dy$$

$$= 1 - 1 = 0$$

が成り立つ。不等号の部分で等号が成り立つのは $p_{\xi|\xi'}(x|y) = g(x|y)$ の時に限る。

補題 3.1 では (3.3) の条件の下で $\sup_{\xi'} E\,h(\xi|\xi')$ を求めたのであるが，その結果を見ると右辺の値は $\varepsilon \to 0$ の時に単調に減少してゆくから

$$\sup_{\xi'} \left\{ E\,h(\xi|\xi');\ \begin{array}{l} E|\xi-\xi'|^2 \leqq \varepsilon^2 \\ p_{\xi|\xi'}(x|y) \text{ が存在する} \end{array} \right\}$$

$$= -n\log\frac{1}{\varepsilon} - \frac{n}{2}\log n + n\log\sqrt{2\pi e}$$

としてよい。この結果と (3.1), (3.2) をあわせて

(3.4) $$H_\varepsilon(\xi) \geqq \underline{H} = n\log\frac{1}{\varepsilon} + \frac{n}{2}\log n - n\log\sqrt{2\pi e} + h(\xi)$$

の評価を得る。

(第二段) <u>上からの評価</u>

評価の方針は $E|\xi-\xi'|^2 = \varepsilon^2$ を満すような一つの ξ' をうまく選んで，$I(\xi,\xi')$ を計算することである。天下り的であるが ξ に対する条件つき確率の密度関数が

(3.5) $$p_{\xi'|\xi}(y|x) = \frac{1}{\left(\frac{2\pi\alpha\varepsilon^2}{n}\right)^{n/2}} \exp\left\{-\frac{n}{2\alpha\varepsilon^2}|y-\alpha x|^2\right\}$$

(但し $\alpha = 1-\frac{\varepsilon^2}{\sigma^2}$, $\sigma^2 = E|\xi|^2$ とする)

で与えられる n 次元確率変数 $\xi' = (\xi'_1, \cdots, \xi'_n)$ を考える。*)

---

*) この密度関数は実は $\xi = (\xi_1, \cdots, \xi_n)$ が Gaussian である場合に第一段で示した $g(x|y)$ (ξ' に対する ξ の条件つき密度) から逆 Fourier 変換によって求めた「ξ に対する ξ' の条件つき密度関数」を流用した。

Sem. on Probab.
Vol.29 1968年
P1-95

補題 3.2.

(3.5) で与えられる $\mathring{\xi}$ について

$$E|\xi-\mathring{\xi}|^2 = \varepsilon^2$$

証明 $E|\xi-\mathring{\xi}|^2 = \sum_{i=1}^{n}\int_{R^n} p_\xi(x)\left[\int_{R^n}(x_i-y_i)^2 p_{\mathring{\xi}|\xi}(y|x)\,dy\right]dx$

$$= \sum_{i=1}^{n}\int_{R^n} p_\xi(x)\left[\int_{R^n}\frac{1}{\left(\frac{2\pi\alpha\varepsilon^2}{n}\right)^{\frac{1}{2}}}(x_i-y_i)^2 \exp\left\{-\frac{n}{2\alpha\varepsilon^2}(y_i-\alpha x_i)^2\right\}dy_i\right]dx$$

$$= \sum_{i=1}^{n}\int_{R^n} p_\xi(x)\left[\frac{\alpha\varepsilon^2}{n} + (\alpha-1)^2 x_i^2\right]dx$$

$$= n\cdot\frac{\alpha\varepsilon^2}{n} + (\alpha-1)^2\sigma^2 = \varepsilon^2 .$$

補題 3.3

(3.5) で与えられる $\mathring{\xi}$ について

$$\lim_{\varepsilon\to 0} h(\mathring{\xi}) = h(\xi).$$

証明 $\mathring{\xi}$ の特性関数 $\varphi_{\mathring{\xi}}(z) = E(e^{i(z,\mathring{\xi})})$, $z=(z_1,\cdots,z_n)$

$$\varphi_{\mathring{\xi}}(z) = \int_{R^n} e^{i(z,y)}\left[\int_{R^n} p_\xi(x)\frac{1}{\left(\frac{2\pi\alpha\varepsilon^2}{n}\right)^{n/2}}\exp\left\{-\frac{n}{2\alpha\varepsilon^2}|y-\alpha x|^2\right\}dx\right]dy$$

に於いて変数変換 $u = y-\alpha x$ を行なうと

$$= \left[\int_{R^n} e^{i\alpha(z,x)} p_\xi(x)\,dx\right]\cdot\left[\int_{R^n} e^{i(z,u)}\frac{1}{\left(\frac{2\pi\alpha\varepsilon^2}{n}\right)^{n/2}}\cdot\exp\left\{-\frac{n}{2\alpha\varepsilon^2}|u|^2\right\}du\right]$$

$$= \varphi_\xi(\alpha z)\cdot\exp\left\{-\frac{\alpha\varepsilon^2}{2n}|z|^2\right\}$$

($\varphi_\xi(\cdot)$ は ξ の特性関数)

となる。 $\varepsilon\to 0$ とすれば $\alpha\to 1$, 従って $\varphi_{\mathring{\xi}}(z)\to\varphi_\xi(z)$。

$$\lim_{\varepsilon\to 0} p_{\mathring{\xi}}(x) = p_\xi(x).$$

Sem. on Probab.
Vol.29 1968年
P1-95

そこで正数列 $\varepsilon_1 \geqq \varepsilon_2 \geqq \cdots \to 0$ を選べば，関数列 $p_{\dot{\xi}}(x;\varepsilon_i)\log p_{\dot{\xi}}(x;\varepsilon_i)$ は一様有界で $p_\xi(x)\log p_\xi(x)$ に収束するから

$$\lim_{\varepsilon\to 0} h(\dot{\xi}) = -\lim_{\varepsilon\to 0}\int_{R^n} p_{\dot{\xi}}(x;\varepsilon)\log p_{\dot{\xi}}(x;\varepsilon)\,dx$$

$$= -\int_{R^n} p_\xi(x)\log p_\xi(x)\,dx = h(\xi).$$

さて，補題 3.2 によれば $E|\xi-\dot{\xi}|^2 = \varepsilon^2$ であるから，(3.5)で与えられる $\dot{\xi}$ に対して $\bar{H} = I(\xi,\dot{\xi})$ を計算すれば $H_\varepsilon(\xi)$ の上からの評価が得られる。この $\dot{\xi}$ に対して容易に計算されるように

$$E h(\dot{\xi}|\xi) = -n\log\frac{1}{\varepsilon} - \frac{n}{2}\log n + n\log\sqrt{2\pi e\alpha}$$

これを

$$\bar{H} = I(\xi,\dot{\xi}) = h(\dot{\xi}) - E h(\dot{\xi}|\xi)$$

に代入して補題 3.3 を適用すれば，十分小さい $\varepsilon>0$ について

$$(3.6)\qquad H_\varepsilon(\xi) \leqq \bar{H} = n\log\frac{1}{\varepsilon} + \frac{n}{2}\log n - n\log\sqrt{2\pi e\left(1-\frac{\varepsilon^2}{\sigma^2}\right)} + h(\xi) + o(1)$$

故に $\varepsilon\to 0$ の時

$$H_\varepsilon(\xi) \leqq n\log\frac{1}{\varepsilon} + \frac{n}{2}\log n - n\log\sqrt{2\pi e} + h(\xi) + o(1)$$

オ一段とオ二段を合せて定理 3.1 が証明された。

Sem. on Probab.
Vol.29 1968年
P1-95

§4. Gaussian system $\xi=(\xi_1,\xi_2,\cdots)$ の ε-エントロピー

この節では，はじめに一般の無限次元確率ベクトル $\xi=(\xi_1,\xi_2,\cdots)$ の ε-エントロピーについて成り立つ結果を述べ，次に Gaussian の場合に成り立つ特殊の事情を反映した精密な結果を述べる。

次のような無限次元確率ベクトル $\xi=(\xi_1,\xi_2,\cdots)$ を考える：$\{\xi_k\}_{k\geq 1}$ は平均 0，分散 ρ_{kk} の確率変数列で，$\sum_{k\geq 1}\rho_{kk}<\infty$ をみたすものとする。このとき，$\sum_{k\geq 1}E\xi_k^2=E(\sum_{k\geq 1}\xi_k^2)<\infty$ だから，$\sum_{k\geq 1}\xi_k^2<\infty$ が確率 1 で成り立つ。すなわち，$\xi\in\ell^2$ (a. e)。
$E(\xi_k\xi_\ell)\equiv\rho_{k\ell}$ ($1\leqslant k,\ \ell<\infty$) を要素とする行列は，ℓ^2 での正値，対称，完全連続（$\because$）$\sum_{k,\ell\geq 1}|\rho_{k\ell}|\leqq(\sum_{k\geq 1}\rho_{kk})^{\frac{1}{2}}(\sum_{\ell\geq 1}\rho_{\ell\ell})^{\frac{1}{2}}<\infty$，これから，$\sum_{k,\ell\geq 1}|\rho_{k\ell}|^2<\infty$ がでる。）な作用素である。その固有値とそれに対応する完全正規直交系をなす固有ベクトルを $\lambda_1\geqq\lambda_2\geqq\cdots\geqq 0$; $(a_{11},a_{12},\cdots),(a_{21},a_{22},\cdots),\cdots$, とする。行列 $A=(a_{k,\ell})_{1\leqq k,\ \ell<\infty}$ は ℓ^2 での等距離作用素となる。

ξ を A で変換したものを $\eta=(\eta_1,\eta_2,\cdots)$ としよう。そうすると $\{\eta_k\}_{k\geq 1}$ は平均 0，分散 λ_k，互いに correlation のない確率変数列となる。なぜなら

$$
\begin{aligned}
E\eta_k\eta_\ell &= E(\sum_{i\geq 1}a_{ki}\xi_i)(\sum_{j\geq 1}a_{\ell j}\xi_j)\\
&=\sum_{i\geq 1}a_{ki}\sum_{j\geq 1}a_{\ell j}\rho_{ij}\\
&=\sum_{i\geq 1}a_{ki}\lambda_\ell a_{\ell i}\ =\delta_{k\ell}\lambda_k.
\end{aligned}
$$

このとき次の定理がなりたつ。

<u>定理 4.1</u>

$$H_\varepsilon(\xi)\ =\ H_\varepsilon(\eta)$$

<u>証明</u>　$\dot{\xi}=(\dot{\xi}_1,\dot{\xi}_2,\cdots)$ を $E(\sum_{k\geq 1}|\xi_k-\dot{\xi}_k|^2)\leqq\varepsilon^2$ をみたす任意のベクトル（このことから，$\dot{\xi}$ は確率 1 で ℓ^2 の元となる）とし，$\dot{\xi}$ を A で

Sem. on Probab.
Vol.29 1968年
P1-95

変換したものを $\dot{\eta}$ とする。そのとき，A が逆 A^{-1} をもち，これらが連続したがって l^2 の topological Borel field に関して可測であることから，

(4.1) $\quad I(\xi, \dot{\xi}) = I(\eta, \dot{\eta})$ *)

またあきらかに 確率 1 で

(4.2) $\quad \sum_{k\geq 1} |\xi_k - \dot{\xi}_k|^2 = \sum_{k\geq 1} |\eta_k - \dot{\eta}_k|^2$

がなりたつ。このことから $E(\sum_{k\geq 1} |\eta_k - \dot{\eta}_k|^2) \leqq \varepsilon^2$，したがって ε-エントロピーの定義によって，

(4.3) $\quad H_\varepsilon(\xi) \geqq H_\varepsilon(\eta)$

がなりたつが，η に対する $\dot{\eta}$ を $E(\sum_{k\geq 1}|\eta_k - \dot{\eta}_k|^2) \leqq \varepsilon^2$ をみたすようにとったと考えて同様な論法を用いれば

(4.4) $\quad H_\varepsilon(\xi) \leqq H_\varepsilon(\eta)$

がなりたって定理が示される。

Remark. この定理によって，われわれは corelation のない場合のみを（したがって Gauss の場合には独立な場合のみを）考えればよいことがわかる。しかし，一般に $H_\varepsilon(\eta)$ の評価は困難である。一般的には，上または下からの（かなりあらっぽい）評価が与えられるだけである。

定理 4.2

$\xi = (\xi_1, \xi_2, \cdots)$ は平均 0，分散 ρ_{kk}，$\sum_{k\geq 1} \rho_{kk} < \infty$ の無限次元確率ベクトルとする。そのとき，$\forall n \geqq 1$ に対して，

$$H_\varepsilon(\xi) \geqq H_\varepsilon((\xi_1, \xi_2, \cdots, \xi_n))$$

証明 情報量の性質 §1. 定理 1.3 (III) **) により，$\forall n \geqq 1$ に対して

$$I((\xi_1, \xi_2, \cdots), (\dot{\xi}_1, \dot{\xi}_2, \cdots)) \geqq I((\xi_1, \cdots, \xi_n), (\dot{\xi}_1, \cdots, \dot{\xi}_n))$$

がなりたつ。このことを用いれば

*) ここで，情報量の性質 §1. 定理 1.3 (III) を用いたが，そのさい もとになる 確率空間 $(\Omega, \mathcal{B}, P)$ を $(\hat{\Omega}, \hat{\mathcal{B}}, \hat{P})$ に制限して考えればよい。
ここで $\hat{\Omega} = \{\omega, \xi(\omega) \in l^2\}$，$P(\hat{\Omega}) = 1$，$\hat{\mathcal{B}} = \{E \cap \hat{\Omega} ; E \in \mathcal{B}\}$，$P$ は $\hat{P}$ を $\hat{\mathcal{B}}$ に制限したもの。

**) 以下で 情報量の性質（・）という場合には，すべて §1. 定理 1.3 でのものを指すことにする。

Sem. on Probab.
Vol.29 1968年
P1-95

$$H_\varepsilon(\xi) = \inf_{(\mathring{\xi}_1, \mathring{\xi}_2, \cdots)} \{ I((\xi_1, \xi_2, \cdots), (\mathring{\xi}_1, \mathring{\xi}_2, \cdots)) ; \sum_{k \geq 1} E|\xi_k - \mathring{\xi}_k|^2 \leqq \varepsilon^2 \}$$

$$\geqq \inf_{(\mathring{\xi}_1, \mathring{\xi}_2, \cdots)} \{ I((\xi_1, \cdots, \xi_n), (\mathring{\xi}_1, \cdots, \mathring{\xi}_n)) ; \sum_{k \geq 1} E|\xi_k - \mathring{\xi}_k|^2 \leqq \varepsilon^2 \}$$

$$\geqq \inf_{(\mathring{\xi}_1, \cdots, \mathring{\xi}_n)} \{ I((\xi_1, \cdots, \xi_n), (\mathring{\xi}_1, \cdots, \mathring{\xi}_n)) ; \sum_{k=1}^{n} E|\xi_k - \mathring{\xi}_k|^2 \leqq \varepsilon^2 \}$$

$$= H_\varepsilon((\xi_1, \cdots, \xi_n))$$

ここで，$H_\varepsilon(\xi)$ は $\varepsilon \downarrow 0$ のときに $H_\varepsilon(\xi)$ の定義から単調に増大することを用いた。

Remark　この証明からわかるように，$H_\varepsilon((\xi_1, \cdots, \xi_n))$ は n の単調増加関数である。

次に $H_\varepsilon(\xi)$ の上からの評価を与える。

定理 4.3

定理 4.2 と同じ仮定のもとに，

$$H_\varepsilon(\xi) \leqq H_\theta((\xi_1, \xi_2, \cdots, \xi_n))$$

がなりたつ。ここで n は $\varepsilon^2 - \sum_{k=n+1}^{\infty} \rho_{kk} \equiv \theta^2 > 0$ をみたす数で $\theta^2 = \theta^2(n)$。

証明　ε − エントロピーの定義から

$$H_\varepsilon(\xi) \leqq \inf_{(\mathring{\xi}_1, \cdots, \mathring{\xi}_n)} \{ I((\xi_1, \xi_2, \cdots), (\mathring{\xi}_1, \cdots, \mathring{\xi}_n, 0, 0, \cdots)) ; (\mathring{\xi}_1, \cdots, \mathring{\xi}_n) \text{は} (\xi_{n+1}, \xi_{n+2}, \cdots) \text{と独立で，} \sum_{k=1}^{n} E|\xi_k - \mathring{\xi}_k|^2 + \sum_{k \geq n+1} \rho_{kk} \leqslant \varepsilon^2 \text{をみたす.} \}$$

$$= \inf \{ I((\xi_1, \cdots, \xi_n), (\mathring{\xi}_1, \cdots, \mathring{\xi}_n)) ; \sum_{k=1}^{n} E|\xi_k - \mathring{\xi}_k|^2 \leqslant \theta^2 \}$$

$$= H_\theta((\xi_1, \cdots, \xi_n))$$

ここで，不等式から等式に移行する部分では，情報量の性質 (V') を用いることによって，$I((\xi_1, \xi_2, \cdots), (\mathring{\xi}_1, \cdots, \mathring{\xi}_n, 0, 0, \cdots)) = I((\xi_1, \cdots, \xi_n), (\mathring{\xi}_1, \cdots, \mathring{\xi}_n))$ となることを用いた。

Remark　K. Kagi [12] は拡散過程の ε-エントロピーの評価をこれら

Sem. on Probab.
Vol.29 1968年
P1-95

の評価を用いて行った(3章. §6)。その際，問題は有限次元の絶対連続分布の ε-エントロピーの評価に帰着することになる。

さて，Gaussian system $\xi=(\xi_1,\xi_2,\cdots):\{\xi_k\}_{k\geq 1}$ は平均 0，分散 $\rho_k=\rho_{kk}$，$\sum_{k\geq 1}\rho_k<\infty$，$\rho_1\geqq\rho_2\geqq\cdots\geqq 0$ をみたす、互に独立な Gaussian random variable の列，のε-エントロピーの評価に移ろう。このときには，きっちりした評価ができて次の定理がなりたつ。

<u>定理 4.4</u> (Pinsker [20])

(4.5) $\quad H_\varepsilon(\xi)=\frac{1}{2}\sum_{k\geq 1}\log\left(\frac{\rho_k}{\theta^2}\vee 1\right)$

ここで，θ^2 は方程式

(4.6) $\quad \sum_{k\geq 1}\min(\rho_k,\theta^2)=\varepsilon^2$

から一意に定まる数である。

<u>Remark 1.</u> ε^2 を与えたとき(4.6)から θ^2 が一意的に定まることは，次のようにしてわかる: $\rho_k=\rho_{kk}$ とおき，$f(\rho)=\sum_{k\geq 1}\min(\rho_k,\rho)$ が ρ の連続かつ真に増加($0<\rho<\rho_1$)な関数であることが示されれば，ε^2 は $\sum_{k\geq 1}\rho_k$ より小であるとして一般性を失なわないから $f(\rho)=\varepsilon^2$ をみたす $\rho=\theta^2$ が一意に定まる。

$\rho'>\rho$ としたとき，$f(\rho')=n'\rho'+\sum_{k=n'+1}^{\infty}\rho_k$，$f(\rho)=n\rho+\sum_{k=n+1}^{\infty}\rho_k$，ここで $n\geqq n'$ となっているから，

$$f(\rho')-f(\rho)=\sum_{k=1}^{n'}(\rho'-\rho)+\sum_{k=n'+1}^{n}(\rho_k-\rho)\quad>0$$

また，$(0,\rho_1)$ に属する任意の ρ,ρ' に対して，

$$|f(\rho')-f(\rho)|\leqq\max(n,n')\cdot|\rho'-\rho|+|n-n'|\cdot|\rho'-\rho|$$

となるから，$f(\rho)$ は連続である。

定理の証明は，Gauss の場合の特殊性: 1) 1次元 Gauss 分布の ε-エントロピーがきっちり求まる。 2) 同次分布も Gaussian のときに ε-エントロピーの下限が到達される。 3) $\{(\xi_k,\tilde{\xi}_k)\}_{k\geq 1}$ が独立な対の列のときに ε-エントロピーの下限が到達される。を用いてなされる。

Sem. on Probab.
Vol.29 1968年
P1-95

Gauss以外のときには，こうした事情の成立の保証がない。なお，以下で考える確率変数の平均はすべて 0 であると仮定する。次の 1)～3) に対応する 3つの補題を用意する。

補題 4.1

$\xi=(\xi_1,\cdots,\xi_m)$ を与えられた m 次元 Gaussian random variable，(a_{ij}) を与えられた $(m+n)$ 行－$(m+n)$ 列の行列とする。また，n 次元 random variable $\eta=(\eta_1,\cdots,\eta_n)=(\xi_{m+1},\cdots,\xi_{m+n})$ で，$E\xi_i\xi_j=a_{ij}\ (1\leqq i\cdot j\leqq m+n)$ をみたすものと考える。そのとき，そのような η の中で $I(\xi,\eta)$ を最小にするものは (ξ,η) が $(m+n)$ 次元 Gaussian random variable になるものである。

証明　(ξ,η) を補題の条件をみたす任意の $(m+n)$ 次元 random variable とする。最初に (ξ,η) の同次分布 $P_{\xi\eta}$ が密度関数 $p_{\xi\eta}$ をもつ場合を考える。そのとき，p_ξ, p_η, $p_{\xi\eta_g}$, p_{η_g} をそれぞれ対応する分布 P_ξ, P_η, $P_{\xi\eta_g}$, P_{η_g} の密度関数とする。ここで $P_{\xi\eta_g}$ は $(m+n)$ 次元 Gauss 分布，P_{η_g} はその n 次元の周辺分布になっている場合とする。

$$(4.7)\qquad I(\xi,\eta)-I(\xi,\eta_g)=\int_{R^m}\int_{R^n}\log\frac{p_{\xi\eta}(x,y)}{p_\xi(x)\,p_\eta(y)}\,dx\,dy$$
$$-\int_{R^m}\int_{R^n}\log\frac{p_{\xi\eta_g}(x,y)}{p_\xi(x)\,p_{\eta_g}(y)}\,p_{\xi\eta_g}(x,y)\,dx\,dy$$

が非負であることを示せばよい。ここで p_{η_g} は n 次元 Gauss 分布の密度関数だから，$\log(p_{\xi\eta_g}(x,y)/p_\xi(x)\,p_{\eta_g}(y))$ は $(m+n)$ 個の変数の 2次形式であり，$E\xi_i\xi_j=a_{ij}$ を (ξ,η), (ξ,η_g) が共にみたしていることから (4.7) の第2項は

$$-\int_{R^m}\int_{R^n}\log\frac{p_{\xi\eta_g}(x,y)}{p_\xi(x)\,p_{\eta_g}(y)}\,p_{\xi\eta}(x,y)\,dx\,dy$$

に等しい。これから

$$(4.8)\qquad I(\xi,\eta)-I(\xi,\eta_g)=\int_{R^m}\int_{R^n}\log\frac{p_{\xi\eta}(x,y)\,p_{\eta_g}(y)}{p_{\xi\eta_g}(x,y)\,p_{\eta_g}(y)}\,p_{\xi\eta}(x,y)\,dx\,dy$$

Sem. on Probab.
Vol.29 1968年
P1-95

$$\geq \int_{R^m}\int_{R^n} \left(1 - \frac{p_{\xi\eta_g}(x,y)\, p_\eta(y)}{p_{\xi\eta}(x,y)\, p_{\eta_g}(y)}\right) p_{\xi\eta}(x,y)\, dx\, dy$$

（ここで不等式 $\log x \geqq 1 - \frac{1}{x}$ $(x \geqq 0)$ *) を用いた）

$$= 1 - \int_{R^n} \frac{p_{\eta_g}(y)\, p_\eta(y)}{p_{\eta_g}(y)}\, dy = 0$$

次に $P_{\xi\eta}$ が密度関数を持たない場合を考える。このときには，任意に固定した $\delta > 0$ に対して，互いに独立かつ (ξ, η) とも独立な $(m+n)$ 個の $N(0, \delta^2)$ にしたがう *Gaussian random variable* $\zeta_1, \cdots, \zeta_{m+n}$ をとり，

$$\overline{\xi}_i = \xi_i + \zeta_i \ (1 \leqq i \leqq m+n), \quad \overline{\xi} = (\overline{\xi}_1, \cdots, \xi_n),$$
$$\overline{\eta} = (\overline{\xi}_{m+1}, \cdots, \overline{\xi}_{m+n})$$

とおく。そのとき，4個の *random variable* の列 $(\overline{\xi}, \xi, \eta, \overline{\eta})$ と $(\overline{\xi}, \xi, \eta_g, \overline{\eta}_g)$ は *Markov chain* をなすから情報量の性質(V″)によって，

(4.9) $\quad I(\overline{\xi}, \overline{\eta}) \leqq I(\xi, \eta), \quad I(\overline{\xi}, \overline{\eta}_g) \leqq I(\xi, \eta_g)$

がなりたつ。またここで $\delta \to 0$ とすると，$(\overline{\xi}, \overline{\eta})$；$(\overline{\xi}, \overline{\eta}_g)$ はそれぞれ (ξ, η)，(ξ, η_g) に法則収束するから，情報量の性質(VIII)によって

(4.10) $\quad \varliminf_{\delta \to 0} I(\overline{\xi}, \overline{\eta}) \geqq I(\xi, \eta), \quad \varliminf_{\delta \to 0} I(\overline{\xi}, \overline{\eta}_g) \geqq I(\xi, \eta_g)$

がなりたつ。また $P_{\overline{\xi}\overline{\eta}}$，$P_{\overline{\xi}\eta_g}$ は密度関数をもつから，はじめに述べたことから，

(4.11) $\quad I(\overline{\xi}, \overline{\eta}) \geqq I(\overline{\xi}, \overline{\eta}_g)$

がなりたっている。(4.9), (4.10), (4.11) から，$I(\xi, \eta) = \infty$ または

$$I(\xi, \eta) \geqq I(\xi, \eta_g)$$

がなりたつことが示された。

<u>系.1.</u>

$\xi = (\xi_1, \cdots, \xi_m)$ を与えられた m 次元 *Gaussian random variable*，A を与えられた $(m+n)$ 行 - $(m+n)$ 列の行列 (a_{ij}) の集合とする。

*) $x = 0$ のときには $-\infty = -\infty$ と考える。

Sem. on Probab.
Vol.29 1968年
P1-95

そのとき $E\xi_i\xi_j = a_{ij}$ $(1 \leq i, j \leq m+n)$, $(a_{ij}) \in A$ をみたす $\eta = (\eta_1, \cdots, \eta_n) \equiv (\xi_{m+1}, \cdots, \xi_{m+n})$ の集合を W_A とすると,

$$H(W_A) \equiv \inf_{\eta \in W_A} I(\xi, \eta) = \inf\{I(\xi, \eta);\ \eta \in W_A,\ P_{\xi\eta} \text{ は Gauss 分布にしたがう}\}$$

証明　$(a_{ij}) \in A$ を任意に固定する度に, 補題から $P_{\xi\eta}$ が Gauss 分布のときに $I(\xi, \eta)$ が最小となるから系がなりたつ。

系 2.

$\xi = (\xi_1, \xi_2, \cdots)$ を与えられた Gaussian system ; A と B を与えられた無限行列 (a_{ij}), (b_{ij}) $(1 \leq i,\ j < \infty)$ の集合とする。そのとき, $E\xi_i\eta_j = a_{ij}$, $E\eta_i\eta_j = b_{ij}$ $(1 \leq i,\ j < \infty)$, $(a_{ij}) \in A$, $(b_{ij}) \in B$ をみたす無限次元 random variable $\eta = (\eta_1, \eta_2, \cdots)$ の集合を $W_{A.B}$ とすると,

$$H(W_{A.B}) \equiv \inf I(\xi, \eta) = \inf\{I(\xi, \eta);\ \eta \in W_{(A.B)},\ (\xi, \eta) \text{ Gaussian system}\}$$

証明　情報量の性質(VI)と系1により系2の成立が示される。

系1, 系2 は Gaussian random variable (有限次元または l^2 の値をとる) の ε-エントロピーの評価の際に, Gauss の範囲で議論ができることを示している. 何故なら, ユークリッド空間 R^n または l^2 の場合 ε-エントロピーを定義している条件 $E\rho(\xi, \eta)^2 \leq \varepsilon^2$ は ξ, η の座標に関する2次のモーメントに関する条件になるからである。以下で考える ε-エントロピーを定める距離はこれらの距離である.

補題 4. 2

ξ を正規分布 $N(0, \rho_{\xi\xi})$ にしたがう random variable としたとき,

$$H_\varepsilon(\xi) = \frac{1}{2}\log\left(\frac{\rho_{\xi\xi}}{\varepsilon^2} \vee 1\right) \tag{4.12}$$

証明　$\rho_{\xi\xi} \leq \varepsilon^2$ のときには, $H_\varepsilon(\xi) = \inf_{E|\xi-\eta|^2 \leq \varepsilon^2} I(\xi, \eta) = I(\xi, 0) = 0$ となるから (4.12) がなりたっている.

したがって以下では $\rho_{\xi\xi} > \varepsilon^2$ と仮定する. 補題 4.1 の系 1. から $P_{\xi\eta}$ が 2次元分布にしたがう random variable η の　を考えれば十分である.

Sem. on Probab.
Vol.29 1968年
P1-95

したがって，η を $E|\xi-\eta|^2 \leqq \varepsilon^2$ をみたし，(ξ,η) が2次元 Gaussian random variable となるものとする。$\xi-\eta=\zeta$ とおくと，ζ が $E\zeta^2 \equiv \rho_{\zeta\zeta}=\varepsilon^2$ かつ $E_{\xi\zeta} \equiv \rho_{\xi\zeta}=\rho_{\zeta\zeta}=\varepsilon^2$（このとき $E\eta\zeta=E(\xi-\zeta)\zeta=0$，したがって η と ζ は独立）をみたす Gaussian random variable のときに $I(\xi,\eta)=I(\xi,\xi-\zeta)$ の下限が到達されることが示される。なぜなら，定理 1.4 によって

$$I(\xi,\eta) = \frac{1}{2}\log\frac{\rho_{\xi\xi}\rho_{\eta\eta}}{\rho_{\xi\xi}\rho_{\eta\eta}-\rho_{\xi\eta}^2}$$

となるが，これを ξ と ζ で表現すると，$\rho_{\eta\eta}=\rho_{\xi\xi}-2\rho_{\xi\zeta}+\rho_{\zeta\zeta}$，$\rho_{\xi\eta}=\rho_{\xi\xi}-\rho_{\xi\zeta}$ を代入して，

$$I(\xi,\eta) = \frac{1}{2}\log\frac{\rho_{\xi\xi}(\rho_{\xi\xi}-2\rho_{\xi\zeta}+\rho_{\zeta\zeta})}{\rho_{\xi\xi}\rho_{\zeta\zeta}-\rho_{\xi\zeta}^2}$$

となるがこれを $\rho_{\xi\zeta}$ の関数と考えると，この右辺を最小にする $\rho_{\xi\zeta}$ の値は容易に $\rho_{\zeta\zeta}$ であることがわかる。そしてそのとき

$$I(\xi,\eta)=\frac{1}{2}\log\frac{\rho_{\xi\xi}(\rho_{\xi\xi}-\rho_{\zeta\zeta})}{\rho_{\zeta\zeta}(\rho_{\xi\xi}-\rho_{\zeta\zeta})} = \frac{1}{2}\log\frac{\rho_{\xi\xi}}{\rho_{\zeta\zeta}} \geqq \frac{1}{2}\log\frac{\rho_{\xi\xi}}{\varepsilon^2}$$

となるから，$\rho_{\zeta\zeta}=\varepsilon^2$ のとき $I(\xi,\eta)$ が最小となり，その値は $\frac{1}{2}\log\frac{\rho_{\xi\xi}}{\varepsilon^2}$ である。

次に Gaussian system $\xi=(\xi_1,\xi_2,\cdots)$ の場合に戻って考える。補題 4.1 とその系で $H_\varepsilon(\xi)=\inf\{I(\xi,\eta);\, E\rho(\xi,\eta)^2 \leqq \varepsilon^2,\ (\xi,\eta)$ Gaussian system$\}$ としてよいことがわかったが，さらに次の補題がなりたつ。

<u>補題 4.3</u>

Gaussian system $\xi=(\xi_1,\xi_2,\cdots)$；$\{\xi_k\}_{k\geq 1}$ は $N(0,\rho_k)$ にしたがい，$\sum_{k\geq 1}\rho_k<\infty$，互に独立な Gaussian random variable の ε-エントロピーの評価において

$$H_\varepsilon(\xi)=\inf\{I(\xi,\eta);\, E\rho(\xi,\eta)^2 \leq \varepsilon^2,\ (\xi,\eta)\ \text{Gaussian system},\ \{(\xi_k,\eta_k)\}_{k\geq 1}\ \text{は対として互に独立}\}$$

としてよい。

Sem. on Probab.
Vol.29 1968年
P1-95

証明　$E\rho(\xi,\eta)^2=\sum_{k\geq 1}E|\xi_k-\eta_k|^2\leqq\varepsilon^2$, (ξ,η) Gaussian system をみたす任意の η に対して

$E\xi_k\eta_k=E\xi_k\tilde{\eta}_k$, $E\eta_k^2=E\tilde{\eta}_k^2$, $\forall k\geqq 1$ かつ
$\{(\xi_k,\eta_k)\}_{k\geq 1}$ は互いに独立

となるような Gaussian system $\tilde{\eta}=(\tilde{\eta}_1,\tilde{\eta}_2,\cdots)$ で $(\xi,\tilde{\eta})$ も Gaussian system となっているものをとったときに，$E(\xi_k-\eta_k)^2=E(\xi_k-\tilde{\eta}_k)^2$, $\forall k\geqq 1$ であるから $\tilde{\eta}$ は補題の条件をみたしている。したがって，

(4.13)　$I(\xi,\eta)\geqq I(\xi,\tilde{\eta})$

を示せばよい。それにはまず Gaussian random variable の情報量が2次のモーメントのみで定まることから

(4.14)　$I(\xi_k,\eta_k)=I(\xi_k,\tilde{\eta}_k)$,　$k\geq 1$

がなりたつことにまず注意する。また情報量の性質(VII)から $\xi=(\xi_1,\xi_2,\cdots)$, $\{\xi_k\}_{k\geq 1}$ は互いに独立, $\eta=(\eta_1,\eta_2,\cdots)$ のときに

(4.15)　$I(\xi,\eta)\geqq\sum_{k\geq 1}I(\xi_k,\eta_k)$

がなりたっているから，(4.14) と合わせて

$$I(\xi,\eta)\geqq\sum_{k\geq 1}I(\xi_k,\eta_k)=\sum_{k\geq 1}I(\xi_k,\tilde{\eta}_k)=I(\xi,\tilde{\eta})$$

がなりたつ。ここで最後の等式は情報量の性質(VII)による。

定理4.4.の証明

(4.16)　$E\rho(\xi,\eta)^2=\sum_{k\geq 1}E|\xi_k-\eta_k|^2=\sum_{k\geq 1}\varepsilon_k^2\leqq\varepsilon^2$

をみたす η についての $I(\xi,\eta)$ の下限を求めればよいが補題1.とその系また補題3.によって (ξ,η) が Gaussian system であって，$\{(\xi_k,\eta_k)\}_{k\geq 1}$ が互いに独立な対の場合を考えれば十分である。$E|\xi_k-\eta_k|^2=\varepsilon_k^2$, $H_{\varepsilon_k}(\xi_k)=I(\xi_k,\eta_k)$ をみたす $\{\eta_k\}_{k\geq 1}$ がすでにとられたとしよう。そうすると，

(4.17)　$H_\varepsilon(\xi)=\inf I(\xi,\eta)$

$$=\inf\sum_{k\geq 1}H_{\varepsilon_k}(\xi_k)=\inf\left\{\frac{1}{2}\sum_{k\geq 1}\log\left(\frac{\rho_k}{\varepsilon_k^2}\vee 1\right)\right\}$$

Sem. on Probab.
Vol.29 1968年
P1-95

となる。ここでの2行目の下限は $\sum_{k\geq1}\varepsilon_k^2 \leqq \varepsilon^2$ をみたすような $\{\varepsilon_k\}_{k\geq1}$ のすべてのとり方についてとる。したがって (4.17) が (4.5) (と (4.6)) に一致することを示せばよい。

$\sum_{k\geq1}\log(\frac{\rho_k}{\varepsilon_k^2}\vee1)$ をできるだけ小さくするための $\{\varepsilon_k\}_{k\geq1}$ の取り方としては，1) $\sum_{k\geq1}\varepsilon_k^2=\varepsilon^2$ とする. 2) $\varepsilon_k^2\leqq\rho_k$, $\forall k\geq1$ とする。 3) ある N_0 に対して，$\varepsilon_1^2=\varepsilon_2^2=\cdots=\varepsilon_{N_0}^2=\theta_0^2$ とする．とすればよいことがすぐわかる。したがって，この 1), 2), 3) をみたすように $\{\varepsilon_k\}_{k\geq1}$ を任意にとったときに

(4.18) $$\sum_{k\geq1}\log(\frac{\rho_k}{\varepsilon_k^2}\vee1)-\sum_{k\geq1}\log(\frac{\rho_k}{\theta^2}\vee1)\geqq0$$

を示せばよい。（もちろん，等号をなりたたせることはできる。）

そこで，$\rho_k>\theta^2$ をみたす k の最大値を N とする。そのとき次の3つの場合が考えられる.

i) $N>N_0$, ii) $N=N_0$, iii) $N<N_0$

しかし，i)の場合は 1)～3) を考慮すれば ii) に帰着されることがわかる。そこで ii) と iii) の場合に (4.18) を示す.

ii) $N=N_0$ のとき，このとき $\theta^2\geqq\theta_0^2$ の場合だけ考えれば十分である。($\because$) (4.6) と 1), 2) による)

したがって

$$(4.18)\text{の左辺}=\sum_{k=1}^{N_0}\log\frac{\rho_k}{\theta_0^2}+\sum_{k\geq N_0+1}\log\frac{\rho_k}{\varepsilon_k^2}-\sum_{k=1}^{N}\log\frac{\rho_k}{\theta^2}$$
$$=\sum_{k\geq N+1}\log\frac{\rho_k}{\varepsilon_k^2}-N\log\frac{\theta_0^2}{\theta^2}\geqq0.$$

iii) $N<N_0$ のとき， このとき 2)によって $\theta^2\geqq\theta_0^2$ としてよいことがわかる。そして，

$$(4.18)\text{の左辺}=\sum_{k=1}^{N_0}\log\frac{\rho_k}{\theta_0^2}+\sum_{k\geq N_0+1}\log\frac{\rho_k}{\varepsilon_k^2}-\sum_{k=1}^{N}\log\frac{\rho_k}{\theta^2}$$
$$=\sum_{k=N+1}^{N_0}\log\frac{\rho_k}{\theta_0^2}+\sum_{k\geq N_0+1}\log\frac{\rho_k}{\varepsilon_k^2}+N\log\frac{\theta^2}{\theta_0^2}\geqq0$$

系.

Gaussian system $\xi=(\xi_1,\xi_2,\cdots)$; $\{\xi_k\}_{k\geq1}$ は互いに独立，$N(0,\rho_k)(\sum_{k\geq1}\rho_k<\infty)$ にしたがう random variable の列の ε-エントロピー

Sem. on Probab.
Vol.29 1968年
P1-95

$H_\varepsilon(\xi)$ は ε の連続関数である。

証明　$\varepsilon_1 > \varepsilon_2 > 0$ とする。 $\varepsilon_i^2 = N_i \theta_i^2 + \sum_{k \geq N_i+1} \rho_k$ $(i = 1, 2)$ とする。そのとき明らかに $\theta_1^2 > \theta_2^2$ であるから $N_1 \leqq N_2$ となることがわかる。

$$2(H_{\varepsilon_2}(\xi) - H_{\varepsilon_1}(\xi)) = \sum_{k=1}^{N_2} \log \frac{\rho_k}{\theta_2^2} - \sum_{k=1}^{N_1} \log \frac{\rho_k}{\theta_1^2}$$
$$= \sum_{k=N_1+1}^{N_2} \log \frac{\rho_k}{\theta_2^2} + \sum_{k=1}^{N_1} \log \frac{\theta_1^2}{\theta_2^2}$$

ここで，$\theta_1^2 \geqq \rho_{N_1+1} \geqq \cdots \geqq \rho_{N_2} \geqq \theta_2^2$．そこで $\varepsilon_1 \to \varepsilon_2$ (又は $\varepsilon_2 \to \varepsilon_1$) とすれば $\theta_1^2 \to \theta_2^2$ また $\rho_k / \theta_2^2 \to 1$ $(N_1 + 1 \leqq k \leqq N_2)$

したがって

$$H_{\varepsilon_2}(\xi) - H_{\varepsilon_1}(\xi) \longrightarrow 0$$

Remark 2.　この定理の証明の過程でわかるが Gaussian system $\xi = (\xi_1, \xi_2, \cdots)$ において，$\xi_k = \eta_k + \zeta_k$，$k \geqq 1$ としたとき，$\{\zeta_k\}_{k \geq 1}$ は互いに独立であり，かつ $\eta = (\eta_1, \eta_2, \cdots)$ とも独立な Gaussian random variable で

$$E\zeta_k^2 = \varepsilon_k^2 = \begin{cases} \theta^2, & \rho_k \geqq \theta^2 \text{ のとき} \\ \rho_k, & \rho_k < \theta^2 \text{ のとき} \end{cases}$$

となっているときに $H_\varepsilon(\xi)$ の下限が到達されている。このことは Kolmogorov [13] に述べてある。またこの $\eta = \xi - \zeta$ をとったとき

$$H_\varepsilon(\xi) = \sum_{\rho_k > \theta^2} I(\xi_k, \eta_k) = I((\xi_1, \cdots, \xi_k), (\eta_1, \cdots, \eta_k))$$
$$= I((\xi_1, \xi_2, \cdots), (\eta_1, \eta_2, \cdots, \eta_k, 0, 0, \cdots))$$
$$= I(\xi, (\eta_1, \cdots, \eta_k))$$

したがって ξ は有限次元の random variable $(\eta_1, \cdots, \eta_k)$ で近似されていることになる。

Remark 3.　有限次元 Gaussian random variable $\xi = (\xi_1, \cdots, \xi_N)$; $\{\xi_k\}_{k \geq 1}$ は互いに独立，$N(0, \rho_k)$ にしたがう。に対しては，定理は次の形になる。

$\varepsilon^2 \leqq N \min_{1 \leq k \leq N} \rho_k$ ならば

Sem. on Probab.
Vol.29 1968年
P1-95

$$H_\varepsilon(\xi) = \frac{1}{2}\sum_{k=1}^{N} \log \frac{\rho_k}{\varepsilon^2/N} = N \log \frac{1}{\varepsilon} + \frac{1}{2}\left(N \log N + \sum_{k=1}^{N} \log \rho_k\right)$$

例. $\rho_k = C k^{-s}$ ($C>0$, $s>1$ は定数) $k \geqq 1$ の場合には, 定理から

(4.19) $$H_\varepsilon(\xi) = \frac{1}{2} C^{\frac{1}{s-1}} s^{\frac{s}{s-1}} (s-1)^{-\frac{1}{s-1}} \varepsilon^{-\frac{2}{s-1}} + o\left(\varepsilon^{-\frac{2}{s-1}}\right)$$

となることがわかる。

追補

今迄 $H_\varepsilon(\xi) < \infty$ となることに触れなかったが, そのことは定理 4.3、と、2次のモーメントが与えられた時にはその中で Gaussian random variable の ε-エントロピーが最大である(そして、それは有限である)こと —— Kolmogorov 〔13〕 p. 310 にある Pinsker の定理による —— を用いて示される.

Sem. on Probab.
Vol.29 1968年
P1-95

第三章 確率過程の ε-エントロピー

この章では，はじめに (ω, t)-可測，$E\int_0^T \xi(t,\omega)^2 dt < \infty$ をみたす確率過程 $\xi = \xi(t)$, $0 \leqq t \leqq T$ の ε-エントロピーについて一般になりたつ結果を述べ，次に平均連続な Gauss 過程の場合の特別な事情を述べる.(§5) §6. では，§5 の結果と §4 の評価を拡散過程の場合に適用して ε-エントロピーの計算を行う。§7 では n-重マルコフ Gauss 過程の場合の評価を行い，n-重マルコフの n が結果にどう反映するかを見る。また，§8 では，§5 の結果がパラメーター t を多次元にした場合にそのまま拡張されることを用いて，パラメーターの次元の ε-エントロピーへの反映をしらべる。

§5. 確率過程の ε-エントロピー

(ω, t)-可測*)，$E\xi(t) = 0$ $(0 \leqq t \leqq T)$**)，$E\int_0^T \xi(t,\omega)^2 dt < \infty$ をみたす確率過程 $\xi = \xi(t,\omega)$, $0 \leqq t \leqq T$ を考える。そのとき，ξ は $L^2[0,T]$ の値をとる確率変数とみなせることにまず注意する。そのことは次のようにして示すことができる***)：1) $\forall \varphi(t) \in L^2[0,T]$ に対して，$\varphi(t)\xi(t,\omega)$ は (ω,t)-可測であり，

$$E\int_0^T |\varphi(t)\xi(t,\omega)|\, dt \leqq \left(\int_0^T |\varphi(t)|^2 dt\right)^{1/2} \cdot E\left(\int_0^T \xi(t,\omega)^2 dt\right)^{1/2}$$

$$\leqq \left(\int_0^T |\varphi(t)|^2 dt\right)^{1/2} \cdot \left(E\int_0^T \xi(t,\omega)^2 dt\right)^{1/2} < \infty$$

となるから，Fubini の定理により，$\int_0^T \varphi(t)\xi(t,\omega)\,dt \equiv (\varphi, \xi(\cdot,\omega))$ は ω-可測関数。2) $\{\psi_n\}_{n\geqq 1}$ を $L^2[0,T]$ での 0 でない dense set とすると $\varphi_n \equiv \psi_n/\|\psi_n\|$, $n \geqq 1$ は半径 1 の球面上での dense である。そして $\forall \xi \in L^2[0,T]$ に対して $\|\xi\| = \sup_n |(\varphi_n, \xi)|$。3) $\{\omega; \|\xi(\cdot,\omega)\| < 1\}$ は 1), 2) により ω-可測集合，したがってこのことから ξ が可測であることがでる。

---

*) このことは例えば ξ が確率連続なら保証される。

**) この仮定は本質的でない。これがなりたたない場合には，以下の議論に若干の修正をすれば同じような結果がなりたつ。

***) このことは静岡のセミナーでの討論に負う。

Sem. on Probab.
Vol.29 1968年
P1-95

次に，$E\xi(s)\xi(t) = r(s.t)$[*)] を核とする $L^2[0,T]$ での積分作用素 R を考える。これは，対称，正値，完全連続である。R の固有値と対応する完全正規直交系をなしている固有関数系を $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq 0$，$\{\varphi_k(t)\}_{k\geqq 1}$ とする。次の積分によって確率変数 ξ_k $(k\geqq 1)$ を定義する．

$$(5.1)\qquad \xi_k = \int_0^T \xi(t,\omega)\,\varphi_k(t)\,dt$$

$\{\xi_k\}_{k\geqq 1}$ は平均 0，分散 λ_k で互いに correlation のない確率変数列である。何故なら前者は Fubini の定理から直ちに従い，後者は

$$\begin{aligned}
E\,\xi_k\xi_l &= E\left(\int_0^T \xi(t,\omega)\varphi_k(t)\,dt\right)\left(\int_0^T \xi(s,\omega)\varphi_l(s)\,ds\right)\\
&= E\left(\int_0^T\!\!\int_0^T \xi(t,\omega)\xi(s,\omega)\varphi_k(t)\varphi_l(s)\,dt\,ds\right)\\
&= \int_0^T\!\!\int_0^T r(s,t)\,\varphi_k(t)\varphi_l(s)\,dt\,ds\\
&= \lambda_k\int_0^T \varphi_k(t)\varphi_l(t)\,dt = \lambda_k\,\delta_{kl}
\end{aligned}$$

となるから。ここで，Parceval の等式

$$(5.2)\qquad \sum_{k\geqq 1}\xi_k^2 = \int_0^T \xi(t,\omega)^2\,dt \quad ,\ \forall\omega$$

がなりたつ。[**)] これから

$$(5.3)\qquad \int_0^T E\,\xi(t,\omega)^2\,dt = \sum_{k\geqq 1} E\,\xi_k^2 = \sum_{k\geqq 1}\lambda_k < \infty$$

すなわち，$\sum_{k\geqq 1}\lambda_k$ は収束する。

定理 5.1

$$(5.4)\qquad H_\varepsilon(\xi(t)) = H_\varepsilon((\xi_1,\xi_2,\cdots))$$

証明　$\dot{\xi}(t)$ を

$$(5.5)\qquad E\left[\int_0^T (\xi(t)-\dot{\xi}(t))^2\,dt\right] \leqq \varepsilon^2$$

をみたす確率過程とする。$\dot{\xi}(t,\omega)$ は (ω,t)-可測[***)] かつ $E\int_0^T \dot{\xi}(t)^2\,dt < \infty$ をみたすと考えてよい。したがって，$\{\xi_k\}_{k\geqq 1}$ に対応する $\{\dot{\xi}_k\}_{k\geqq 1}$ が

---

*) $r(s,t)$ が (s,t)-可積分なことは上の 1) と同様にして示せる。

**) ここで，また以下で §4 の最初の脚注のような注意をする。

***) (5.5) をみたす $\dot{\xi}(t)$ というだけでは，(ω,t)-可測性が出て来ないかも知れないが，積分の順序交換などを保証するためにも，(ω,t)-可測でないもので近似することは考えないことにする。必要ならば確率過程の ε-エントロピーの定義を変えればよい。

Sem. on Probab.
Vol.29 1968年
P1-95

(5.1′) $$\dot{\xi}_k = \int_0^T \dot{\xi}(t,\omega)\varphi_k(t)\,dt$$

で定義される。そして

(5.2′) $$\sum_{k\geq 1}|\xi_k - \dot{\xi}_k|^2 = \int_0^T |\xi(t,\omega) - \dot{\xi}(t,\omega)|^2\,dt\ , \quad {}^\forall\omega$$

がなりたつ。これから

(5.6) $$\sum_{k\geq 1} E|\xi_k - \dot{\xi}_k|^2 \leqq \varepsilon^2$$

また ω を固定したとき $L^2[0,T]$ から ℓ^2 への写像（5, 1）（または(5.1′)）

$$K : \xi(\cdot,\omega) \longrightarrow \{\xi_1, \xi_2, \cdots\}$$

は等距離作用素であるから K^{-1} が存在して，K, K^{-1} は連続，したがって，*topological Borel frield* に関して可測である。したがって情報量の性質（定理1, 3）(III) によって，

(5.7) $$I(\xi(t), \dot{\xi}(t)) = I((\xi_1, \xi_2, \cdots), (\dot{\xi}_1, \dot{\xi}_2, \cdots))$$

がなりたつ。以上から

(5.8) $$H_\varepsilon(\xi(t)) \geqq H_\varepsilon((\xi_1, \xi_2, \cdots))$$

の成立がわかるが $\{\xi_1, \xi_2, \cdots\}$ から出発して上の推論を逆にすれば (5.8) と逆の不等式がなりたつことがわかる。

この定理によって，確率過程の ε-エントロピーの評価は対応する無限次元確率ベクトルの ε-エントロピーの評価に帰着されることがわかった。また以上の結果においてはパラメーターの次元は何等本質的でないから，まったく同様の結果が多次元パラメーターの確率過程においても成立する。

最後に平均連続な確率過程の場合を考察する。まず，はじめにこの場合 Riemann の意味で考えた積分 $\int_0^T \xi(t,\omega)\varphi_k(t)\,dt$ が定義できる。ここで $\varphi_k(t)$, $(k\geq 1)$ は連続関数である。Δ を区間 $[0,T]$ の分割，

$$\Delta : 0 = t_0 < t_1 < \cdots < t_n = T$$

δ をこれらの細区間の最大の巾，$\delta = \max_i (t_i - t_{i-1})$ とする。
そのとき，

$$\underset{\substack{n\to\infty \\ \delta\to 0}}{\mathrm{l.i.m}} \sum_{\Delta} \xi(t_i)\varphi_k(t_i)\cdot(t_i - t_{i-1}) \quad \text{in } L^2[\Omega]$$

Sem. on Probab.
Vol.29 1968年
P1-95

が存在することは普通の Riemann 積分と同様に示せる。そこで

(5.9) $\quad \mathrm{l.i.m.}\sum_{\Delta}\xi(t_i)\varphi_k(t_i)(t_i-t_{i-1}) = R\text{-}\int_0^T \xi(t,\omega)\varphi_k(t)\,dt$

と書くことにする．そのとき，次の定理がなりたつ。

定理 5.2 *)

(5.1) で定義された Lebesgue 積分による ξ_k, $(k\geq 1)$ と $R\text{-}\int_0^T\xi(t,\omega)\varphi_k(t)\,dt$ は $L^2[\Omega]$ の元として等しい。

証明　簡単のために，$\eta_k = R\text{-}\int_0^T\xi(t,\omega)\varphi_k(t)\,dt$ とかこう $\forall f\in L^2[\Omega]$ に対して $(\xi_k, f) = (\eta_k, f)$ が成立することをまず示そう。ここで $(\ ,\ )$ は $L^2[\Omega]$ での積である。

$$
\begin{aligned}
(\eta_k, f) &= \lim\Big(\sum_{\Delta}\xi(t_i,\omega)\varphi_k(t_i)(t_i-t_{i-1}),\ f\Big)\\
&= \lim E\Big(\sum_{\Delta} f(\omega)\xi(t_i,\omega)\varphi_k(t_i)(t_i-t_{i-1})\Big)\\
&= \lim\sum E f(\omega)\xi(t_i,\omega)\varphi_k(t_i)(t_i-t_{i-1})\\
&= \int_0^T E(\xi(t)f(\omega))\varphi_k(t)\,dt
\end{aligned}
$$

ここで，$E(\xi(t,\omega)f(\omega))$ が t の連続関数になることを用いた。一方，

$$
\begin{aligned}
(\xi_k, f) &= E f(\omega)\int_0^T\xi(t,\omega)\varphi_k(t)\,dt\\
&= \int_0^T E(\xi(t,\omega)f(\omega))\varphi_k(t)\,dt\\
&= (\eta_k, f)
\end{aligned}
$$

次に，fとして特に集合 $\{\omega;\ \xi_k>\eta_k\}$, $\{\omega;\ \xi_k<\eta_k\}$ の特性関数を代入すれば容易に $P(\xi_k>\eta_k) = P(\xi_k<\eta_k) = 0$，したがって

$$P(\xi_k \neq \eta_k) = 0$$

を得る。

さて，$\xi(t,\omega)$ が平均連続な Gauss 過程のときには $R\text{-}\int_0^T\xi(t,\omega)\varphi_k(t)\,dt$ は Gauss 分布にしたがう。したがって今の定理によって，$\xi_k = \int_0^T\xi(t,\omega)\varphi_k(t)\,dt$ も Gauss 分布にしたがう。そしてそのとき，$\{\xi_k\}_{k\geq 1}$ は Gauss 分布 $N(0,\lambda_k)$ にしたがう独立な確率変数列であるから，$\{\xi_k\}_{k\geq 1}$ は Gaussian

*) 近藤亮司氏による。

Sem. on Probab.
Vol.29 1968年
P1-95

system である。この ε-エントロピー は 定理 4.4 で 求まっているから，結局 定理 5.1 と 合わせれば 次の 定理 がなりたつことがわかった。

<u>定理 5.3</u> (Pinsker [20])

$\xi = \xi(t,\omega)$, $0 \leqq t \leqq T$ を平均連続な Gauss 過程，$\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq 0$ を $E\xi(s)\xi(t) = \gamma(s,t)$ を核とする $L^2[0,T]$ での積分作用素の固有値とすれば，

$$(5.10) \qquad H_\varepsilon(\xi) = \frac{1}{2}\sum_{k\geqq 1} \log\left(\frac{\lambda_k}{\theta^2} \vee 1\right)$$

がなりたつ。ここで θ^2 は方程式 (4.6) からきまる数である。

<u>例.</u> ξ が ブラウン運動 のときには

$$(5.11) \qquad H_\varepsilon(\xi) = \frac{2T^2}{\pi^2}\cdot\frac{1}{\varepsilon^2} + o\left(\frac{1}{\varepsilon^2}\right) \qquad (\text{Pinsker})$$

このとき，$E\xi(s)\xi(t) = \min(t,s)$，これから容易に

$$\lambda_n = \frac{4T^2}{\pi^2(1+2n)^2} = \frac{T^2}{\pi^2}\cdot\frac{1}{n^2} + o\left(\frac{1}{n^2}\right)$$

を得るから，定理 4.4 の Example (4.19) 式 によって (5.11) を得る。
この結果は Kolmogorov [13] に A.M. Jaglom が $\frac{1}{\varepsilon^2}$ の係数の評価をしたとして述べてあるが，その評価には若干誤りがある。

以上においては，パラメーター t の動く範囲 $0 \leqq t \leqq T$ は固定して考えてきたが，$T \to \infty$ のときの ε-エントロピーの様子を調べるものとして，単位時間当りの ε-エントロピー がある。(Kolomogorov [13] が導入)

$\xi = \xi(t)$ $0 \leqq t < \infty$ or $-\infty < t < \infty$ を (ω, t)-可測で ${}^\forall S, T$ に対し，$E\int_0^T \xi(t)^2 dt < \infty$ なる確率過程 とする。$\xi_S^T \equiv \{\xi(t),\ S \leqq t \leqq T\}$ とし，もし 極限

$$(5.12) \qquad \bar{H}_\varepsilon(\xi) = \lim_{T-S\to\infty} \frac{1}{T-S}\, H_{\varepsilon\sqrt{T-S}}(\xi_S^T)$$

が存在するとき，この極限を単位時間当りの ε-エントロピー という。

単位時間当りの ε-エントロピー が計算されている例は 次のものがある。

Sem. on Probab.
Vol.29 1968年
P1-95

例 1. ブラウン運動

ξ がブラウン運動のとき

(5.13) $\qquad \overline{H}_\varepsilon(\xi) = \frac{2}{\pi^2}\frac{1}{\varepsilon^2} + o\left(\frac{1}{\varepsilon^2}\right)$

実は, $\forall S\ \forall T$ に対して, $\frac{1}{T} H_{\varepsilon\sqrt{T}}(\xi^{T+S}) = \overline{H}_\varepsilon(\xi) + o(\overline{H}_\varepsilon(\xi))$ が成り立っている。

例 2. *Stationary Gaussian Process*

この *Process* の単位時間当りの ε-エントロピーについては, §10. で詳しくのべる。

Sem. on Probab.
Vol.29 1968年
P1-95

§6. *Diffusion process* の ε-エントロピー

この§では確率積分方程式

$$(6.1) \qquad \xi(t) = \int_0^t b(u, \xi(u))\, du + \int_0^t a(u, \xi(u))\, dB(u)$$

から構成される一次元 *diffusion process* $\xi = \xi(t)$, $0 \leqq t \leqq T$ の ε-エントロピーの計算を行なう。ここに係数 $a(t, x)$, $b(t, x)$ に関して以下の仮定を設ける。

$$\begin{cases} |a(t,x) - a(t,y)| + |b(t,x) - b(t,y)| \leqq L|x-y|, \\ b(t,x)^2 \leqq L(1+x^2), \\ 0 < k \leqq a(t,x)^2 \leqq K \end{cases}$$

<u>定理 6.1</u>

以上の仮定の下で，十分小さい $\varepsilon > 0$ に対して

$$H_\varepsilon(\xi) \leqq \frac{2}{\pi^2} \left\{ \int_0^T \sqrt{E\, a(u, \xi(u))^2}\, du \right\}^2 \frac{1}{\varepsilon^2} + o\left(\frac{1}{\varepsilon^2}\right),$$

$$H_\varepsilon(\xi) \geqq \frac{k}{eK} \cdot \frac{1}{\pi^2} \left\{ \int^T \sqrt{E\, a(u, \xi(u))^2}\, du \right\}^2 \frac{1}{\varepsilon^2} + o\left(\frac{1}{\varepsilon^2}\right)$$

の評価が成り立つ。*)

<u>証明</u> (6.1)で与えられる $\xi(t)$, $0 \leqq t \leqq T$ の *covariance function* $r(t,s) = E(\xi(t) - E\xi(t))(\xi(s) - E\xi(s))$ は

$$(6.2) \qquad r(t,s) = \int_0^{t \wedge s} E\, a(u, \xi(u))^2\, du$$

であるから，$\xi = \xi(t)$ は平均連続な確率過程である。定理 5.1 によれば $\xi = \xi(t)$ の ε-エントロピーは $H_\varepsilon((\xi_1, \xi_2, \cdots))$ に等しい。但し

$$(6.3) \qquad \xi_j = \int_0^T \{\xi(u) - E\,\xi(u)\}\, \varphi_j(u)\, du,$$

ここで $\varphi_j(u)$ を積分方程式

---

*) Kolmogorov [13] は証明なしに $H_\varepsilon(\xi) = \frac{4}{\pi}\left\{\int_0^T E\, a(u, \xi(u))^2\, du\right\}\frac{1}{\varepsilon^2} + o\left(\frac{1}{\varepsilon^2}\right)$ という評価を示しているが，この $\frac{1}{\varepsilon^2}$ の係数はいろいろな結果を考え合せると [illegible] ないように思われる。

Sem. on Probab.
Vol.29 1968年
P1-95

(6.4) $\displaystyle\int_0^T r(t,s)\varphi(s)\,ds = \lambda\varphi(t)$

の j 番目の固有関数とする。

(6.3) で定義される ξ_j の分散は方程式 (6.4) の j 番目の固有値 λ_j に等しいが，λ_j の漸近的な値をまず求めておこう。

補題 6.1

λ_j は漸近的に

$$\lambda_j = Cj^{-2} + o(j^{-2}), \quad C = \frac{1}{\pi^2}\left\{\int_0^T \sqrt{E\,a(u,\xi(u))^2}\,du\right\}^2$$

と表わされる。

証明　(6.2) を方程式 (6.4) に代入して，これを微分方程式に直すと

$$\begin{cases} -\dfrac{d}{dt}\left\{\dfrac{1}{E\,a(t,\xi(t))^2}\dfrac{d\varphi}{dt}\right\} = \dfrac{1}{\lambda}\varphi \\ \varphi(0) = \varphi'(T) = 0 \end{cases}$$

という Sturm-Liouville type の境界値問題となる。（ここで $a(t,x)^2 \geqq k > 0$ の条件が効いている）。ところがこの問題で λ_j の値は §7. 例 2 と全く同様に

$$\lambda_j = \frac{1}{\pi^2}\left\{\int_0^T \sqrt{E\,a(u,\xi(u))^2}\,du\right\}^2 \cdot j^{-2} + o(j^{-2})$$

と評価できる。

定理 4.2，定理 4.3 を用いて $H_\varepsilon((\xi_1, \xi_2, \cdots))$ を有限次元確率変数の ε-エントロピーで上と下から評価する方針で，有限次元確率変数に対する定理 3.1 の結果を適用するために (6.3) で定義される n 次元確率変数 $\xi = (\xi_1, \cdots, \xi_n)$ の密度関数が存在して定理 3.1 の条件を満すことをいう。

まず $(\xi_1, \xi_2, \cdots \xi_n)$ の特性関数 $\chi(z) = E e^{i\sum_{j=1}^n z_j \xi_j}$ の絶対値を考えてみよう。ξ_j は (6.3) の形に書けるから

(6.5) $\displaystyle |\chi(z)| \leqq \left| E \exp\left\{ i\sum_{j=1}^n z_j \int_0^T \xi(u)\varphi_j(u)\,du \right\}\right|.$

そこで右辺の値を調べるために確率過程

Sem. on Probab.
Vol.29 1968年
P1-95

$$\eta(t;z) = \exp\{i\sum_{j=1}^{n} z_j \int_0^T \xi(u)\varphi_j(u)\,du\}$$

を考える。この式に，$u \geqq t$ に対して確率1で成り立つ

$$\xi(u) = \xi(t) + \int_t^u b(s,\xi(s))\,ds + \int_t^u a(s,\xi(s))\,dB(s)$$

を代入して積分順序の変換を行なえば

$$\eta(t;z) = \exp[i\,\xi(t)\sum_{j=1}^{n} z_j f_j(t)]\,\tilde{\eta}(t;z),$$

但し，

$$\begin{cases} \tilde{\eta}(t;z) = \exp[i\sum_{j=1}^{n}\lambda_j\{\int_t^T b(s,\xi(s))f_j(s)\,ds + \int_t^T a(s,\xi(s))f_j(s)\,dB(s)\}] \\ f_j(s) = \int_s^T \varphi_j(u)\,du \end{cases}$$

を得る。但し，ここで普通の積分と確率積分の積分順序が交換できることを保証する次の補題が必要である。

<u>補題 6.2</u>

$$\int_t^T \varphi(u)\{\int_t^u a(s,\xi(s))\,dB(s)\}\,du = \int_t^T a(s,\xi(s))\{\int_s^T \varphi(u)\,du\}\,dB(s)$$

この補題の証明は確率積分の変換公式を証明するのと同様の方法でできるから，ここでは省略する。

次に

$$\begin{aligned} v(t,x;z) &= E_{t,x}[\eta(t;z)]^{*)} \\ &= \exp\{ix\sum_{j=1}^{n}\lambda_j f_j(t)\}\cdot\tilde{v}(t,x;z) \end{aligned}$$

但し

$$\tilde{v}(t,x;z) = E_{t,x}[\tilde{\eta}(t;z)]$$

とおけば，(6.5) からわかるように

(6 6) $\quad |\chi(z)| \leqq |v(0,0;z)| = |\tilde{v}(0,0;z)|$。

$|\tilde{v}(0,0;z)|$ の値を評価するために $\tilde{v}(t,x;z)$ の満すべき方程式を導き，それを解いてみよう。

<u>補題 6.3</u>

$\tilde{v}(t,x;z)$ は方程式

---

*) $E_{t,x}$ は測度 $P_{t,x}(\cdot) = P(\cdot\,|\,\xi(t)=x)$ による平均値を表わす。

Sem. on Probab.
Vol.29 1968年
P1-95

$$(6.7)\qquad \tilde{v}(t,x;z) = 1 + E_{t,x}\int_t^T A(s,\xi(s))\,\tilde{v}(s,\xi(s);z)\,ds,$$

但し,

$$A(t,x) = i\,b(t,x)\sum_j z_j f_j(t) - \frac{1}{2}a(t,x)^2\sum_j z_j^2 f_j^2(t)$$

を満す。

証明　$\tilde{\eta}(t;z)$ に確率積分の変換公式を適用すれば

$$\tilde{\eta}(t) = \tilde{\eta}(T) + \int_t^T A(s,\xi(s))\,\tilde{\eta}(s)\,ds + i\sum_j \lambda_j\int_t^T b(s,\xi(s))f_j(s)\,\tilde{\eta}(s)\,dB(s)$$

この両辺の平均値 $E_{t,x}$ をとって, マルコフ性と確率積分の性質を使えば (6.7)の方程式を得る。

補題 6.4

方程式 (6.7) は唯一つの解をもち, それは

$$(6.8)\qquad \tilde{v}(t,x;z) = E_{t,x}\exp\left\{\int_t^T A(s,\xi(s))\,ds\right\}$$

と表わせる。

証明.　逐次代入法で解を求める。

$$\tilde{v}_0(t,x;z) = 1$$

$$\tilde{v}_{p+1}(t,x;z) = E_{t,x}\int_t^T A(s,\xi(s))\,\tilde{v}_p(s,\xi(s);z)\,ds \qquad (p = 0,1,2,\cdots)$$

とおけば数学的帰納法によって,

$$\tilde{v}_p(t,x;\lambda) = E_{t,x}\frac{1}{p!}\left\{\int_t^T A(s,\xi(s))\,ds\right\}^p$$

となることが容易に示される。 $|A(t,x)|$ は有界であることから級数 $\sum_{p=0}^{\infty}$ $\tilde{v}_p(t,x;z)$ は $E_{t,x}\left\{\exp\int_t^T A(s,\xi(s))\,ds\right\}$ に絶対収束し, 明らかにこの級数は方程式(6.7)を満す。解が unique であることは通常の方法によって容易に証明される。

従って (6.6) に注意すれば $|\chi(z)|$ について次のような評価を得る。

補題 6.5

$$|\chi(z_1,\cdots,z_n)| \leqq \exp\left\{-\frac{1}{2}\cdot\frac{k}{K}\sum_{j=1}^{n}\lambda_j z_j^2\right\}$$

Sem. on Probab.
Vol.29 1968年
P1-95

証明　補題 6.4 の (6.8) から

$$|\chi(z_1,\cdots,z_n)| \leqq |\tilde{v}(0,0;z)|$$
$$\leqq E\left|\exp\left\{\int_0^T A(s,\xi(s))ds\right\}\right|$$
$$\leqq E\exp\left\{-\frac{1}{2}\int_0^T a(s,\xi(s))^2\sum_j z_j^2 f_j(s)^2 ds\right\}$$
$$\leqq \exp\left\{-\frac{k}{2}\sum_j z_j^2\int_0^T f_j(s)^2 ds\right\}$$

一方，積分方程式 (6.4) を

$$\lambda_j\varphi_j(t) = \int_0^t r(s)\varphi_j(s)ds + r(t)\int_t^T \varphi_j(s)ds$$
$$\left(r(s) = \int_0^s E\,a(u,\xi(u))^2 du \right)$$

と書きかえて，両辺に $\varphi_j(t)$ を乗じた後 t について 0 から T まで積分すれば

$$\lambda_j = 2\int_0^T r(t)\varphi_j(t)\left\{\int_t^T \varphi_j(s)ds\right\}dt$$
$$= -\int_0^T r(t)\frac{d}{dt}\left\{\int_t^T \varphi_j(s)ds\right\}^2 dt$$
$$= \int_0^T r'(t)\left\{\int_t^T \varphi_j(s)ds\right\}^2 dt$$
$$= \int_0^T E\,a(t,\xi(t))^2 f_j(t)^2 dt \text{。}$$

$$\therefore \quad \int_0^T f_j(t)^2 dt \geqq \frac{\lambda_j}{K}$$

この関係と上記の不等式関係から補題 6.5 が得られる。補題 6.5 をもとにして n 次元確率変数 $(\xi_1,\cdots,\xi_n)$ の密度関数について以下のことがわかる。

補題 6.6. $(\xi_1,\cdots,\xi_n)$ の密度関数が存在して，有界連続である。またその *differential entropy* は

$$(6.9) \qquad h((\xi_1,\cdots,\xi_n)) \leqq \log\prod_{j=1}^n (2\pi e\lambda_j)^{\frac{1}{2}},$$

$$(6.10) \qquad h((\xi_1,\cdots,\xi_n)) \geqq -\left|\log\prod_{j=1}^n\left(\frac{K}{2\pi k\lambda_j}\right)^{\frac{1}{2}}\right| \text{。}$$

証明. 補題 6.5 で見る通り $(\xi_1,\cdots,\xi_n)$ の特性関数 $\chi(z_1,\cdots,z_n)$ は $(z_1,\cdots,z_n)$ の L^1-関数であるから，密度関数 $p(x_1,\cdots,x_n)$ が存在して

Sem. on Probab.
Vol.29 1968年
P1-95

$$p(x_1,\cdots,x_n)=\frac{1}{(2\pi)^n}\int \exp(-i\sum_j z_j x_j)\chi(z_1,\cdots,z_n)\,dz_1\cdots dz_n$$

と表わせる。補題6.5からわかる通りこれは連続で有界である：

$$\begin{aligned}M&=\sup p(x_1,\cdots,x_n)\\&\leqq\frac{1}{(2\pi)^n}\int|\chi(z_1,\cdots,z_n)|\,dz_1\cdots dz_n\\&=\prod_{j=1}^{n}\left(\frac{K}{2\pi k\lambda_j}\right)^{\frac{1}{2}}\end{aligned}$$

differential entropy に関する第一の不等式は「$E\xi_j=0$, $E\xi_j\xi_k=\lambda_j\delta_{jk}$ の条件の下で differential entropy の最大値は $(\xi_1,\cdots,\xi_n)$ が Gaussian の時に達成される」という事実から導かれ，又第二の不等式は自明な関係式 $h((\xi_1,\cdots,\xi_n))\geqq-|\log M|$ から得られる。

以上の準備のもとに定理の評価を導こう。

(i) 上からの評価

任意の n に対して

$$\begin{aligned}H_\varepsilon(\xi)&=H_\varepsilon((\xi_1,\xi_2,\cdots))\\&\leqq H_\theta((\xi_1,\cdots,\xi_n))\qquad(\text{定理}4.3)\\&\qquad(\text{ここで }\theta^2=\varepsilon^2-\sum_{j=n+1}^{\infty}\lambda_j)\\&\leqq n\log\frac{1}{\theta}+\frac{n}{2}\log n-n\log\sqrt{2\pi e\alpha}\\&\qquad+h((\xi_1,\cdots,\xi_n))+o(1)\quad((3.6)\text{から})\\&\qquad(\text{但し }\alpha=1-\frac{\varepsilon^2}{\sum_{j=1}^{n}\lambda_j})\\&\leqq\frac{1}{2}\log\prod_{j=1}^{n}\frac{\lambda_j}{\frac{\theta^2}{n}}-n\log\sqrt{\alpha}\quad((6.9)\text{を代入して})\end{aligned}$$

の不等式が成り立つ。

そこで $\lambda_j=Cj^{-2}+o(j^{-2})$（補題6.1）と $\theta^2=\varepsilon^2-\sum_{j=n+1}^{\infty}\lambda_j$ から $\frac{\theta^2}{n}\sim\lambda_n$ となるように n を選ぶと，$n=\left[\frac{2C}{\varepsilon^2}\right]$。

$$H_\varepsilon(\xi)\leqq n+o(n)=2C\cdot\frac{1}{\varepsilon^2}+o\left(\frac{1}{\varepsilon^2}\right).$$

ii) 下からの評価

任意の n に対して

Sem. on Probab.
Vol.29 1968年
P1-95

$$
\begin{aligned}
H_\varepsilon(\xi) &= H_\varepsilon((\xi_1, \xi_2, \cdots)) \\
&\geqq H_\varepsilon((\xi_1, \cdots, \xi_n)) \qquad \text{(定理 4.2)} \\
&\geqq n \log \frac{1}{\varepsilon} + \frac{n}{2} \log n - n \log \sqrt{2\pi e} + h((\xi_1, \cdots, \xi_n)) \\
&\qquad \text{((3.4) から)} \\
&\geqq n \log \frac{1}{\varepsilon} + \frac{n}{2} \log n - n \log \sqrt{2\pi e} - \left| \log \prod_{j=1}^{n} \left(\frac{K}{2\pi k \lambda_j} \right)^{\frac{1}{2}} \right| \\
&\qquad \text{((6.10) から)} \\
&= \frac{n}{2} \log \frac{kC}{\varepsilon^2 e K} - \frac{n}{2} \log n + n + o(n)
\end{aligned}
$$

がなりたつ。ここで 第一項と第二項が打消しあうように $n = \left[\frac{kC}{ek} \cdot \frac{1}{\varepsilon^2} \right]$ととれば,

$$= n + o(n) = \frac{kC}{eK} \cdot \frac{1}{\varepsilon^2} + o\left(\frac{1}{\varepsilon^2}\right) 。$$

$C = \frac{1}{\pi^2} \left\{ \int_0^T \sqrt{E a(u, \xi(u))^2}\, du \right\}^{\frac{1}{2}}$ であったから (i),(ii) をあわせて定理 6.1 が証明された。

<u>Remark</u> 定理 6.1 の上からの評価 (i) は $\lambda_k = Ck^{-s}$ $(c>0,\ s>1)$ の時の *Gaussian random variable* に対する評価と全く同じである。(§4. 定理 4.4 及び *Example* の (4.19) の式)。一方, 下からの評価(ii) では情報量と *differential entropy* について相当に粗っぽい評価を行なっている。(定理 4.2 及び (6.10))。最終的には $H_\varepsilon(\xi)$ の十分小さい $\varepsilon > 0$ に対する漸近式を出来るだけきっちりと表式したいのであるが, ここに述べた方法ではこれ以上に余り改良が図れないように思える。$\xi = \xi(t)$ を可算個の確率変数 $(\xi_1, \xi_2, \cdots)$ に変法する種々の方法の中で $\xi = \xi(t)$ のもつ情報量を保存して かつ $\{\xi_j\}$ が独立となるような変換を巧く見出してきて計算する必要があろう。

Sem. on Probab.
Vol.29 1968年
P1-95

§7 多重マルコフ Gaussian process の ε-エントロピー

ここでは、

$$\xi(t) = \int_0^t \sum_{i=1}^N f_i(t) g_i(u) dB(u) \qquad (0 \leqq t \leqq T)$$

ただし $B(t)$ は $EB(t)=0$, $EB(t)^2 = t$ のブラウン運動、の形で表現される gaussian process の ε-エントロピーを評価する。gaussian process の ε-エントロピーを求める pinsker の結果（§5）によって、この gaussian process の covariance function を核とする積分方程式の固有値の収束の速さを計算すればよい。

$\xi(t)$ についての仮定

1、 f_i, g_i $(i=1,\cdots,N)$ は $[0,T]$ で $2N$ 回微分可能で、 $t \in [0,T]$ で $f_i(t) \neq 0$, $W(g_1,\cdots,g_i)(t) \neq 0$ *) $(i=1,\cdots,N)$

2、 $\sum_{i=1}^{N-1} f_i^{(k)}(t) g_i(t) \equiv 0 \quad (k=0,1,\cdots,N-2)$

$\sum_{i=1}^{N} f_i^{(N-1)}(t) g_i(t) \neq 0 \qquad t \in [0,T]$

このとき、$\xi(t)$は非定常で、Lévy の狭義N重マルコフ過程となっている。このような確率過程については Hida [9] でくわしく取り扱われているから、その詳細はそれを参照せられたい。

定理 7.1

以上のような、$\xi(t)$ に対し、

$$H_\varepsilon(\xi) \asymp \varepsilon^{-\frac{2}{2N-1}} \quad **)$$

以下、計算の概略を述べよう。

covariance function $r(t,s) = E\xi(t)\xi(s)$ は、

---

*) $W(g_1,\cdots,g_i)$ は $g_1,\cdots,g_i$ の Wronskian

**) $f(\varepsilon) \asymp g(\varepsilon) \iff f(\varepsilon) = O(g(\varepsilon))$ かつ $g(\varepsilon) = O(f(\varepsilon))$

Sem. on Probab.
Vol.29 1968年
P1-95

$$I(t,s) = \sum_{i=1}^{N} f_i(t)\, k_i(s) \qquad t \geqq s$$

$$k_i(s) = \sum_{j=1}^{N} f_j(s) \int_0^s g_i(u)\, g_j(u)\, du \qquad (i=1,\cdots,N)$$

と書ける。

まず

$$W(f_1,\cdots,f_N)\cdot W(g_1,\cdots g_N) = \Big\{\sum_{i=1}^{N} f_i^{(N-1)}(s)\, g_i(s)\Big\}^N \neq 0$$

$$W(k_1,\cdots,k_N) = \det\{A_{ij}\}\; W(f_1,\cdots,f_N) \neq 0$$

$$\Big(A_{ij}(s) = \int_0^s g_i(u) g_j(u)\, du \Big)$$

$$\begin{aligned} W&(f_1,\cdots,f_N,k_1,\cdots,k_N) \\ &= \Big\{\sum_{i=1}^{N} f_i^{(N-1)}(s)\, g_i(s)\Big\}^N\, W(f_1,\cdots,f_N)\, W(g_1,\cdots,g_N) \\ &= \Big\{\sum_{i=1}^{N} f_i^{(N-1)}(s)\, g_i(s)\Big\}^{2N} \neq 0 \end{aligned}$$

から、$f_1,\cdots,f_N,k_1,\cdots,k_N$ は一次独立となること、及び

$$(7.1)\quad \begin{cases} \displaystyle\sum_{i=1}^{N}\{f_i^{(k)}(s)\, k_i^{(l)}(s) - f_i^{(l)}(s)\, k_i^{(k)}(s)\} = 0 & (0 \leqq k+l \leqq 2N-2) \\ \displaystyle\sum_{i=1}^{N}\{f_i^{(2N-1)}(s)\, k_i(s) - f_i(s)\, k_i^{(2N-1)}(s)\} = (-1)^N \Big\{\sum_{i=1}^{N} f_i^{(N-1)}(s) g_i(s)\Big\}^2 \end{cases}$$

に注意する。

そこで積分方程式、$\int_0^T I(t,s)\,\varphi(s)\,ds = \lambda\varphi(t)$:

$$\lambda\varphi(t) = \sum_{i=1}^{N}\{f_i(t)\, A_i(t) + k_i(t) B_i(t)\}$$

$$A_i(t) = \int_0^t k_i(s)\,\varphi(s)\,ds \;,\quad B_i(t) = \int_t^T f_i(s)\,\varphi(s)\,ds$$

を、微分方程式の境界値問題に変換する。(7.1)を利用して(7.2)の両辺を2N回微分し、$A_i(t)$、$B_i(t)$ を消去すれば常微分方程式

$$\begin{vmatrix} \varphi & f_1 & \cdots & f_N & k_1 & \cdots & k_N \\ \varphi^{(1)} & f_1^{(1)} & \cdots & f_N^{(1)} & k_1^{(1)} & \cdots & k_N^{(1)} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \varphi^{(2N-1)} & f_1^{(2N-1)} & \cdots & f_N^{(2N-1)} & k_1^{(2N-1)} & \cdots & k_N^{(2N-1)} \\ \varphi^{(2N)} - \frac{\varphi}{\lambda}\sum_i\{f_i^{(2N-1)}k_i - f_i k_i^{(2N-1)}\} & f_1^{(2N)} & \cdots & f_N^{(2N)} & k_1^{(2N)} & \cdots & k_N^{(2N)} \end{vmatrix} = 0$$

が得られ、$W(f_1,\cdots,f_N,k_1,\cdots,k_N) \neq 0$ より、実際に2N階の常微分方程式となっている。これを

$$(7.3)\qquad L\varphi \equiv \sum_{k=0}^{2N} p_k(t)\,\varphi^{(k)}(t) = \frac{\varphi}{\lambda}$$

と書こう。

境界条件($t=0$ と $t=T$ での)としては、

Sem. on Probab.
Vol.29 1968年
P1-95

$$A_i(0) = B_i(T) = 0 \qquad (i = 1, \cdots, N)$$
$$k_i^{(k)}(0) = 0 \qquad (i = 1, \cdots, N\ ;\ k = 0, 1, \cdots, N-1)$$

に注意して、

$$(7.4)\quad \begin{cases} \Delta_k(\varphi) \equiv \varphi^{(k)}(0) = 0 \qquad (k = 0, 1, \cdots, N-1) \\ \Delta_k(\varphi) \equiv \begin{vmatrix} \varphi^{(k)}(T) & f_1^{(k)}(T) & \cdots & f_N^{(k)}(T) \\ \varphi^{(N-1)}(T) & f_1^{(N-1)}(T) & \cdots & f_N^{(N-1)}(T) \\ \vdots & \vdots & & \vdots \\ \varphi^{(1)}(T) & f_1^{(1)}(T) & \cdots & f_N^{(1)}(T) \\ \varphi(T) & f_1(T) & \cdots & f_N(T) \end{vmatrix} = 0 \end{cases}$$

$$(k = N, N+1, \cdots, 2N-1)$$

を得る。

この境界値問題の固有値は、(7.3) の基本解、$\chi_1 = \chi_1(t;\lambda)$, $\chi_2 = \chi_2(t;\lambda)$, $\cdots$, $\chi_{2N}(t;\lambda)$ に対して、方程式

$$(7.5)\quad \begin{vmatrix} \Delta_0(\chi_1) & \Delta_0(\chi_2) & \cdots & \Delta_0(\chi_{2N}) \\ \Delta_1(\chi_1) & \Delta_1(\chi_2) & \cdots & \Delta_1(\chi_{2N}) \\ \vdots & \vdots & & \vdots \\ \Delta_{2N-1}(\chi_1) & \Delta_{2N-1}(\chi_2) & \cdots & \Delta_{2N-1}(\chi_{2N}) \end{vmatrix} = 0$$

の根として求まる。

後の計算のために、もっとも簡単な Doob のN重マルコフ過程の場合を計算しておく。Doob [5] 参照。

例 1、確率微分方程式

$$\xi^{(N)}(t) + a_1 \xi^{(N-1)}(t) + \cdots + a_{N-1}\xi^{(1)}(t) + a_N = 0$$

ただし、$\nu^N + a_1\nu^{N-1} + \cdots + a_{N-1}\nu + a_N = 0$ が単根 $\nu_1, \cdots, \nu_N$ をもつとき、初期条件

$$\xi(0) = \xi^{(1)}(0) = \cdots \xi^{(N-1)}(0) = 0$$

のみたす解は

$$\xi(t) = \frac{a}{\delta}\int_0^t \sum_{i=1}^N \tilde{\delta}_{Ni}\, e^{\nu_i(t-u)}\, dB(u)$$

Sem. on Probab.
Vol.29 1968年
P1-95

ただし、$\tilde{\delta}_{Ni}$ は、行列式

$$\delta = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \nu_1 & \nu_2 & \cdots & \nu_N \\ \vdots & \vdots & & \vdots \\ \nu_1^{N-1} & \nu_2^{N-1} & \cdots & \nu_N^{N-1} \end{vmatrix} \quad \text{の} \quad \nu_i^{N-1} \text{ の余因子}$$

と表わされる。

前にならって

$$f_i(t) = \frac{a}{\delta} \tilde{\delta}_{Ni} e^{\nu_i t}, \quad g_i(t) = e^{-\nu_i t} \qquad (i = 1, \cdots, N)$$

$$h_i(t) = \frac{a}{\delta} \sum_{j=1}^{N} \frac{\tilde{\delta}_{Nj}}{\nu_i + \nu_j} \left(e^{\nu_j t} - e^{-\nu_j t} \right)$$

$$\sum_{i=1}^{N} \left\{ f_i^{(2N-1)}(t) h_i(t) - f_i(t) h_i^{(2N-1)}(t) \right\} = (-1)^N \left\{ \sum_{i=1}^{N} f_i^{(N-1)}(t) g_i(t) \right\}^2$$

$$= (-1)^N a^2$$

とすれば、このとき方程式は

$$\frac{(-1)^N}{a^2} (D^2 - \nu_1^2) \cdots (D^2 - \nu_N^2) \varphi = \frac{\varphi}{\lambda} \qquad \left(D\varphi = \frac{d\varphi}{dt} \right).$$

ところで代数方程式

$$\frac{(-1)^N}{a^2} (z^2 - \nu_1^2) \cdots (z^2 - \nu_N^2) = \frac{1}{\lambda}$$

は常に二つの純虚数 $\pm\mu_1$ をもつ。他の複素数根を $\pm\mu_2, \cdots \pm\mu_N$ （μ_1 の虚数部が正、$\mu_2, \cdots, \mu_N$ の実数部が正であるようにとっておく）とすれば、大きい $\frac{1}{\lambda}$ に対しては

$$|\mu_j| \asymp \lambda^{-1/2N} \qquad (j = 1, \cdots, N)$$

$$\text{Real part of } \mu_j \asymp \lambda^{-1/2N} \qquad (j = 2, \cdots, N)$$

この場合の方程式の基本解 $x_{\pm j}(t) = e^{\pm\mu_j t}$ $(j = 1, \cdots, N)$ について、これを（7.5）に代入し、上の漸近関係に留意しながら（7.5）を近似的に解くと、

$$e^{2\mu_1 T} = (-1)^N \left\{ \frac{(\mu_1 + \mu_2) \cdots (\mu_1 + \mu_N)}{(\mu_1 - \mu_2) \cdots (\mu_1 - \mu_N)} \right\}^2$$

$\frac{1}{\lambda}$ が十分大きいとき、右辺の値の偏角は 0 と 2π の間で考えて定数に近

Sem. on Probab.
Vol.29 1968年
P1-95

づくから、μ_iは十分大きな整数nに対して

$$2T|\mu_i| = 2n\pi + O(1)$$

各固有値の重複度は高々 $2N$ であるから、この関係式から十分大きいnに対しては、n番目の固有値 λ_n について、

$$\lambda_n \asymp \pi^{-2N} T^{2N} n^{-2n}$$

の漸近評価が得られる。

従って Doob の意味のN重マルコフ過程 $\xi = \{\xi(t)\}_{0 \leq t \leq T}$ の ε - エントロピーは

$$H_\varepsilon(\xi) \asymp \varepsilon^{-\frac{2}{2N-1}}$$

一般の問題へ戻ろう。

<u>補題 7.1</u>

(7.3) の微分作用素 L は $g_0(t), g_1(t), \cdots g_N(t)$ を適当にとれば

$$L\varphi = \sum_{k=0}^{N} (-1)^k \frac{d^k}{dt^k}\left(g_k(t) \frac{d^k\varphi}{dt^k} \right)$$

の形に書ける。

<u>証明</u>

等式 $\displaystyle\sum_{k=0}^{2N} p_k(t) \frac{d^k\varphi}{dt^k} = \sum_{k=0}^{N} (-1)^k \frac{d^k}{dt^k}\left(g_k(t) \frac{d^k\varphi}{dt^k} \right)$ の両辺を比較すれば、

$$p_i = \begin{cases} \displaystyle\sum_{k=(i+1)/2}^{i} (-1)^k \binom{k}{2k-i} g_k^{(2k-i)} & (i = 0, 1, \cdots, N-1) \\ \displaystyle\sum_{k=(i+1)/2}^{N} (-1)^k \binom{k}{2k-i} g_k^{(2k-i)} & (i = N, N+1, \cdots, 2N) \end{cases}$$

で、二項係数の間の関係式

$$\sum_{k=i}^{2i} (-1)^k \binom{k}{l}\binom{i}{2i-k} = 0 \qquad (0 \leq l \leq i-1)$$

$$\sum_{k=l}^{2i} (-1)^k \binom{k}{l}\binom{i}{2i-k} = \binom{i}{2i-l} \qquad (i \leq l \leq 2i)$$

を利用して、上式より $g_0, g_1, \cdots, g_N$ を消去すると、

$$p_l = \sum_{k=l}^{2N} (-1)^k \binom{k}{l} p_k^{(k-l)} \qquad (l = 0, 1, \cdots, 2N)$$

Sem. on Probab.
Vol.29 1968年
P1-95

ところが、これは微分作用素Lの形式的自己共役性（元の積分作用素は自己共役作用素であるから） $L\varphi = L^{*}\varphi$:

$$\sum_{k=0}^{2N} p_k \frac{d^k\varphi}{dt^k} = \sum_{k=0}^{2N} (-1)^k \frac{d^k}{dt^k}(p_k \varphi)$$

より係数の間の関係を書き下して得られるものに他ならない。

固有値の求め方について、次の Courant の方法がある.

補題 7.2 (Courant)

$[0,T]$ で区分的に連続な関数 $\psi_1, \cdots, \psi_n$ に対して

$$d(\psi_1,\cdots,\psi_{n-1}) = \inf\left\{\frac{D[\varphi]}{\|\varphi\|^2} ; \begin{array}{l}(\varphi,\varphi_j)=0 \quad (j=1,\cdots,n-1)\\ \varphi\in C^{2N}[0,T] \text{ で境界条件 (7.4) を満す}\end{array}\right\}$$

$$D[\varphi] = (L\varphi,\varphi) = \int_0^T \sum_{k=0}^{N}(-1)^k \varphi \frac{d^k}{dt^k}\left(g_k \frac{d^k\varphi}{dt^k}\right) dt$$

とおくとき、 $L\varphi = \frac{1}{\lambda}\varphi$ の境界条件 (7.4) による n 番目の固有値 $\frac{1}{\lambda_n}$ は

$$\frac{1}{\lambda_n} = \max\left\{ d(\psi_1,\cdots,\psi_{n-1}) ; \psi_1,\cdots,\psi_n \text{ が区分的連続関数}\right\}$$

で与えられ、右辺の max は $\psi_1=\varphi_1, \cdots, \psi_{n-1}=\varphi_{n-1}$ （$\varphi_1,\cdots,\varphi_n$ は最初の n 個の固有関数）のときに達せられる.

証明　Courant, Hilbert [3] で二階の微分作用素について証明されているのと全く同じである.

$D[\varphi]$ は境界条件を用いて、

(7.6)
$$D[\varphi] = \sum_{k=1}^{N}(-1)^k \sum_{l=0}^{k-1}(-1)^l \varphi^{(l)}(t)\left(g_k(t)\varphi^{(k)}(t)\right)^{(k-l-1)}\Big|_{t=T} + \int_0^T \sum_{k=0}^{N} g_k(t)\left(\frac{d^k\varphi}{dt^k}\right)^2 dt$$

$$= \left\{\varphi^{(k)}(T)\ (k=0,1,\cdots,N-1) \text{ の二次形式の項}\right\} + \int_0^T \sum_{k=0}^{N} g_k(t)\left(\frac{d^k\varphi}{dt^k}\right)^2 dt$$

と計算される. ところが、 $g_1(t),\cdots,g_N(t)$ は $[0,T]$ で連続で、

Sem. on Probab.
Vol.29 1968年
P1-95

$$g_N(t) = (-1)^N p_{2N}(t) = \frac{1}{\{\sum_i f_i^{(N-1)}(t)\, g_i(t)\}^2} > 0$$

であるから、〔0,T〕で

$$|g_k(t)| \leq \tilde{m} \quad (k=0,1,\cdots,N) \qquad g_N(t) \geq \underset{\sim}{m}$$

となるような正数 $\tilde{m}$, $\underset{\sim}{m}$ がとれる。

微分作用素 $\tilde{L}$ を

$$\tilde{L}\varphi \equiv \sum_{k=0}^{N}(-1)^k \tilde{m} \frac{d^{2k}\varphi}{dt^{2k}}$$

で定義し、境界条件（$\tilde{L}$ に付随する）として、

$$(7.7) \qquad \begin{aligned} &\varphi(0) = \varphi^{(1)}(0) = \cdots = \varphi^{(N-1)}(0) \\ &\varphi^{(k)}(T) = a_{k0}\varphi(T) + a_{k1}\varphi^{(1)}(T) + \cdots + a_{k\,N-1}\varphi^{(N-1)}(T) \\ &\qquad\qquad (k = N, N+1, \cdots, 2N-1) \end{aligned}$$

$\int_0^T \tilde{L}(\varphi)\varphi\, dt$ を計算したとき、非積分項の二次形式がちょうど（7.6）の第一項に一致するように a_{kj} を定める。

境界条件（7.4）の下での $L\varphi = \frac{1}{\lambda}\varphi$ の n 番目の固有値 $\frac{1}{\lambda_n}$ と、境界条件（7.7）の下での $\tilde{L}\varphi = \frac{1}{\tilde{\lambda}}\varphi$ の n 番目の固有値 $\frac{1}{\tilde{\lambda}_n}$ とを比較すると、補題2から $\frac{1}{\lambda_n} \leq \frac{1}{\tilde{\lambda}_n}$ 。そこで、例1の結果を用いて

$$\lambda_n \geq \tilde{\lambda}_n \asymp T^{2N} n^{-2N} .$$

これと同じようにして

$$\underset{\sim}{L}\varphi \equiv \sum_{k=0}^{N-1}(-1)^k(-\tilde{m})\frac{d^{2k}\varphi}{dt^{2k}} + (-1)^N \underset{\sim}{m}\frac{d^{2N}\varphi}{dt^{2N}} = \frac{\varphi}{\underset{\sim}{\lambda}}$$

の（7.7）と同じかたちの境界条件の下での、固有値評価から、

$$\lambda_n \leq \underset{\sim}{\lambda}_n \asymp T^{2N} n^{-2N}$$

以上を合せて

$$\lambda_n \asymp T^{2N} n^{-2N}$$

従って、$\xi(t)$ の ε-エントロピーの評価

$$H_\varepsilon(\xi) \asymp \varepsilon^{-\frac{2}{2N-1}}$$

を得る。

例2. $\xi(t) = \int_0^t g(u)\, dB(u) \qquad (0 \leq t \leq T)$

Sem. on Probab.
Vol.29 1968年
P1-95

の場合は、ε－エントロピーの評価の定数まで *Pinsker* の結果（§5）に従って、きっちり求められる。

このとき微分方程式は

$$L\varphi = -\frac{d}{dt}\left(\frac{1}{g(t)^2}\frac{d\varphi}{dt}\right) = \frac{\varphi}{\lambda}$$

境界条件は

$$\varphi(0) = 0 \ , \quad \varphi^{(1)}(T) = 0$$

変数を *Liouville* 変換

$$x = \int_0^t |g(u)|\,du \ , \quad y = \frac{\varphi(t)}{\sqrt{|g(t)|}}$$

を行なって、問題は

$$\begin{cases} -\dfrac{d^2y}{dx^2} + z(x)y = \dfrac{y}{\lambda} \\ y(0) = 0 \ , \quad y^{(1)}(a) = y^* \end{cases}$$

$$\left(a = \int_0^T |g(u)|\,du,\ z = \sqrt{|g|}\,\frac{d^2}{dt^2}\sqrt{|g|}\ , \quad y^* \text{は定数} \right)$$

となる．補題２の方法で．

$$\frac{1}{\lambda_n} = \frac{n^2\pi^2}{a^2} + O(1) = n^2\pi^2\left[\int_0^T |g(u)|\,du\right]^{-2} + O(1)$$

$$\therefore \qquad \lambda_n = \frac{1}{\pi^2}\left[\int_0^T |g(u)|\,du\right]^2 n^{-2} + o(n^{-2})$$

従って、$\xi(t)$ $(0 \leqq t \leqq T)$ のε－エントロピーは、

$$H_\varepsilon(\xi) = \frac{2}{\pi^2}\left[\int_0^T |g(u)|\,du\right]^2 \frac{1}{\varepsilon^2} + o\left(\frac{1}{\varepsilon^2}\right)$$

<u>Remark</u> *Lévy* の $M_5(t) = \int_0^T \sqrt{3}\left(\frac{2}{3} - \frac{u}{t} + \frac{1}{3}\frac{u^3}{t^3}\right) dB(u)$ （[12]参照）と与えられるような場合には、時間区間 $[0,1]$ で考えるとき、

$$r_5(t,s) = \frac{s}{2} - \frac{s^2}{5t} + \frac{s^4}{70t^3} \qquad (0 \leqq s \leqq t \leqq 1)$$

となる。この $M_5(t)$ は $t=0$ で、§7 の仮定 1、2、をみたさない。しかし、

Sem. on Probab.
Vol.29 1968年
P1-95

この時も上記のように微分方程式の境界値問題に変形していくと、

$$\text{微分方程式}\quad L(\varphi) \equiv -\frac{d^3}{dt^3}\left(\frac{t^4}{12}\frac{d^3\varphi}{dt^3}\right) + \frac{d^2}{dt^2}\left(t^2\frac{d^2\varphi}{dt^2}\right) - \frac{d}{dt}\left(2\frac{d\varphi}{dt}\right) = \frac{\varphi}{\lambda}$$

境界条件 ： $\varphi(0)=0$, $\varphi^{(1)}(0)$, $\varphi^{(2)}(0)$ は有限
$t=1$ では以前と同様にして導かれる条件、

となり、$t=0$ で最高次の係数 $\frac{t^4}{12}$ が 0 となってしまう、いわゆる *singlar* な *Sturm-Liouville* 問題となる。

このとき、$\frac{1}{\lambda_n}$ の上からの評価は先の議論と全く同じにして得られる。$\frac{1}{\lambda_n}$ の下からの評価は別の工夫を要する（ *Kaji* [11]） が、結局このときも、

$$\lambda_n \asymp n^{-6}$$

が導かれて、

$$H_\varepsilon(M_5) \asymp \varepsilon^{-\frac{2}{5}}$$

を得る。

§ 8、多次元パラメーターの *Brown* 運動の ε-エントロピー

§5、定理 5.1 のあとでのべたように、多次元パラメーターの確率過程の場合にも定理 5.1 と類似の結果がなりたつ。たとえば、有界開領域（または球面のようなもの） $K\subset R^d$ 上で定義された (ω, A)-可測 $(A\in K)$ な確率過程 $\xi(A,\omega)$ で、$E\int_K \xi(A,\omega)^2 dA<\infty$ （dA は R^d のルベーグ測度）をみたすものについてはそうした結果がなりたつ。さらに、平均連続な *gauss* 過程 *) の場合にも定理 5.2 に類似な結果、したがって定理 5.3 に類似な結果がなりたつ。この § では、これらの結果を用いて、平均連続な、多次元パラメーターの *gauss* 過程の例として、多次元パラメータの *Brown* 運動 $X(A)$、$A\in R^d$ において、A を R^d の単位球面 S^{d-1} 上に制限した *gauss* 過程の ε-エントロピーを求めてみよう。この § の表題の $X(A)$ 自身の ε-エントロピーの評価は未だなされていない。

*) ここで考える確率過程は、$E\,\xi(A,\omega)\equiv 0$ をみたすとする。

Sem. on Probab.
Vol.29 1968年
P1-95

まず、定理の形で今の場合における定理5.3に類似の結果をのべておこう。以下、$X(A)$ というのは d 次元パラメーターの Brown 運動のパラメーターを S^{d-1} に制限した Gauss 過程と考えることにする。

定理 8.1

$$H_\varepsilon(X) = \inf\left\{ I(X,Y)\,;\ E\int_{S^{d-1}} |X(A)-Y(A)|^2 d\sigma(A)^{*)} \leqq \varepsilon^2 \right\}$$
$$= \frac{1}{2}\sum \log\left(\frac{\lambda_k}{\theta^2} \vee 1\right)$$

ここで、$\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq 0$ は $E\,X(A)X(B) = \Gamma(A,B)$ を核とする $L^2[S^{d-1}]$ での積分作用素の固有値、θ^2 は (4.6) からきまる数とする.

以下、$X(A)$ の ε-エントロピーの評価、したがって λ_k の評価を、論理の筋道としては逆であるが、d 次元パラメーターの Brown 運動の、球面調和関数を用いての展開を用いて行う。

H.P.Mckean**) によれば

$$(8.1)\qquad X(A) = \sum_{n\geqq 0}\sum_{l=1}^{D(n)} x_n^l(1)\,h_n^l(A), \qquad A \in S^{d-1}$$

と展開できる。ここで、$h_n^l(A)$ は n 次の球面調和関数で、

$$(8.2)\qquad \int_{S^{d-1}} h_n^l(A)\,h_m^k(A)\,d\sigma(A) = \begin{cases} 1 & l=k,\ n=m \text{ のとき.} \\ 0 & \text{otherwise} \end{cases}$$

をみたす。$D(n)$ は n 次の球面調和関数のうち 1 次独立なものの個数で、

$$(8.3)\qquad D(n) = (2n-2+d)\,\frac{(n-3+d)!}{(d-2)!\,n!} \qquad (d \geqq 2,\ n \geqq 0)$$

ただし、$d=2$ のとき $D(0)=1$ とする。また、$x_n^l(1) \equiv x_n^l$ $(n \geqq 0,\ 1 \leqq l \leqq D(n))$ は互に独立な gaussian random variable で

$$(8.4)\qquad x_n^l = C(d)\int_0^1 C_n(u)\,dB_n^l(u)$$

---

*) $d\sigma$ は S^{d-1} 上の一様確率測度

**) *Brownian motion with a several-dimensional time*, Theory Prob. Appl. vol 8, no 4 (1963) 335-354.

Sem. on Probab.
Vol.29 1968年
P1-95

と表現される．ここで，$C(d)$ は次元 d のみに関係する数で，

$$(8.5) \qquad C(d)^2 = 2(d-1)\int_0^{\frac{\pi}{2}} \sin^{d-2}\theta\, d\theta$$

$B_n^l(u)$ $(u \geqq 0,\ 1 \leqq l \leqq D(n))$ は互に独立な標準 Brown 運動であり，

$$(8.6) \qquad C_n(u) = \frac{\int_0^{\cos^{-1}u} p_n(\cos\theta)\sin^{d-2}\theta\, d\theta}{\int_0^{\pi} \sin^{d-2}\theta\, d\theta} \qquad (n \geqq 0)$$

$$p_n(\cos\theta) = C_n^{\frac{d-2}{2}}(\cos\theta) \Big/ C_n^{\frac{d-2}{2}}(1)$$

ここで，$C_n^{\nu}(\cdot)$ は Gegenbauer の多項式である。

展開 (8.1) と X_n^l の独立性（そして $EX_n^l = 0$）を用いて，

$$(8.7) \qquad \Gamma(A,B) = \sum_{n \geqq 0}\sum_{l=1}^{D(n)} E(X_n^l)^2 R_n^l(A) R_n^l(B)$$

を得るが，Mercer の展開定理によれば，$\Gamma(A,B)$ *) を核とする積分作用素の固有値 λ_n^l $(n \geqq 0,\ 1 \leqq l \leqq D(n))$ が $E(X_n^l)^2$ に等しいことを (8.7) は意味している。したがって，以下では $E(X_n^l)^2$ の評価を行う。そして，次の定理がなりたつ。

<u>定理 8.2</u>

$$(8.8) \qquad E(X_n^l)^2 = c\, n^{-d} + o(n^{-d}) \qquad (1 \leqq l \leqq D(n))$$

ここで，c は次元 d のみに関係する正の定数である．

<u>定理 8.3</u>

$$(8.9) \qquad H_\varepsilon(X) = \frac{d}{(d-1)!(d-1)}\left(\frac{2cd}{(d-1)!}\right)^{d-1} \cdot \frac{1}{\varepsilon^{2(d-1)}} + o\left(\frac{1}{\varepsilon^{2(d-1)}}\right)$$

ここで，c は定理 8.2 のものである．

定理 8.2 から定理 8.3 を導くのは，定理 8.1 にしたがって評価することだけであるから証明を省略する．(8.9) での $\varepsilon^{-2(d-1)}$ の係数が標準 Brown

---

*) $\Gamma(A,B) = \frac{1}{2}(|A| + |B| - |A-B|)$ であるから，$\Gamma(A,B)$ は連続である．
$|\cdot|$ は R^d でのベクトルの長さを表わす。

Sem. on Probab.
Vol.29 1968年
P1-95

運動の場合の ε^{-2} の係数にくらべて複雑なのは、固有値の *multiplicity* のためである。また、(8.8) での c は相当に複雑な形をとるので単に c とのみ書いておくことにした。

$d=2,3$ の場合における定理 8.2 の証明

($d=2$) このとき、(8.6) の $p_n(\cos\theta) = \cos n\theta$, $C_n(u) = \frac{1}{n\pi} \times \sin(n\cos^{-1}u)$ となることから、

$$E(x_n^l)^2 = \frac{1}{n^2\pi}\int_0^1 \sin^2(n\cos^{-1}u)\,du$$

$$= \frac{1}{n^2\pi}\int_0^{\frac{\pi}{2}} \sin^2 n\theta \sin\theta\, d\theta = \frac{1}{2\pi}\cdot\frac{1}{n^2} + o\left(\frac{1}{n^2}\right)$$

($d=3$) このとき、$p_n(\cos\theta) = P_n(\cos\theta)$、ここで $P_n(\cdot)$ は n 次の Legendre 多項式となるので $C_n(u) = \frac{1}{2(2n+1)}(P_{n-1}(u) - P_{n+1}(u))$ 、これから、Legendre 多項式の直交性を用いて、

$$E(x_n^l)^2 = \frac{1}{(2n+1)^2}\left(\int_0^1 P_{n+1}(u)^2\,du + \int_0^1 P_{n-1}(u)^2\,du\right)$$

$$= \frac{1}{4}\cdot\frac{1}{n^3} + o\left(\frac{1}{n^3}\right)$$

$d \geqq 4$ の場合における定理 8.2 の証明

このとき (8.4) から、

$$E(x_n^l)^2 = C(d)^2\int_0^1 C_n(u)^2\,du$$

$$= \frac{d-1}{2}\left[(d-3)!\right]^2\left[\int_0^{\frac{\pi}{2}} \sin^{d-2}\theta\,d\theta\right]^{-1}\left(n^{-2d+6} + o(n^{-2d+6})\right)\cdot I_d$$

$$I_d \equiv \int_0^1\left\{\int_0^{\cos^{-1}u} C_n^{\frac{d-2}{2}}(\cos\theta)\sin^{d-2}\theta\,d\theta\right\}^2 du$$

となることがわかる。したがって、I_d を評価すればよいが、そのためには I_d の被積分関数を考察する必要がある。

(I) $d = 2p+2$ ($p \geqq 1$) のとき、

$$I_p(u) \equiv \int_0^{\cos^{-1}u} C_n^{\frac{d-2}{2}}(\cos\theta)\sin^{d-2}\theta\,d\theta$$

$$= \int_0^{\cos^{-1}u} C_n^{p}(\cos\theta)\sin^{2p}\theta\,d\theta$$

Sem. on Probab.
Vol.29 1968年
P1-95

ここで、$C_n^p(\cos\theta)\sin^{2p}\theta$ は、Gegenbauer の多項式の漸化式

$$\sin^2\theta\, C_n^{\nu+1}(\cos\theta) = \frac{1}{2\nu}\left\{(n+2\nu)C_n^{\nu}(\cos\theta) - (n+1)\cos\theta\, C_{n+1}^{\nu}(\cos\theta)\right\} \tag{8.10}$$

と、$\sin\theta\, C_n^1(\cos\theta) = \sin(n+1)\theta$ を用いて、

$$C_n^p(\cos\theta)\sin^{2p}\theta$$
$$= 2^{-(p-1)}\frac{1}{(p-1)!}\sum_{k=1}^{p} A_k^p \cos^{k-1}\theta\sin\theta\sin(n+k)\theta$$

ここで、A_k^p は n の $p-1$ 次の多項式で、$A_k^p = \binom{p-1}{k-1}n^{p-1} + o(n^{p-1})$ と表わされる。さらに、$\cos^{k-1}\theta\sin\theta\sin(n+k)\theta$ を $\{\cos m\theta\}$ の1次結合として表わして積分すれば

$$I_p(u) = \sum_{k=0}^{p}\frac{B_k^p}{n+2k}\sin(n+2k)\alpha \qquad \alpha = \cos^{-1}u$$

となる。ここで、B_k^p は n の $p-1$ 次の多項式でその最高次の係数は

$$b_k^p \equiv 2^{-(p-1)}\frac{1}{(p-1)!}\cdot\sum_{l=k}^{p}\binom{p-1}{l-1}\binom{l}{k}\frac{l-2k}{l\cdot 2^l}$$

である。これから、積分変数を α に変えて積分すれば

$$I_d = \int_0^{\frac{\pi}{2}} I_p(\alpha)^2\sin\alpha\, d\alpha$$
$$= \sum_{k,l=0}^{p}\frac{B_k^p B_l^p}{(n+2k)(n+2l)}\int_0^{\frac{\pi}{2}}\sin(n+2k)\alpha\sin(n+2l)\alpha\sin\alpha\, d\alpha$$
$$= n^{2p-4}\sum_{k,l=0}^{p}\frac{b_k^p\, b_l^p}{2\{1-4(k-l)^2\}} + o(n^{2p-4})$$
$$= n^{d-6}\cdot\frac{1}{2}\sum_{k,l=0}^{p}\frac{b_k^p\cdot b_l^p}{1-4(k-l)^2} + o(n^{d-6})$$

ここで、$\sum_{k,l=0}^{p}\frac{b_k^p\cdot b_l^p}{1-4(k-l)^2}$ は b_k^p, $0 \leqq k \leqq p$, の正値2次形式であることが容易にわかる。そして、$b_0^p > 0$ である。これで、今の場合に定理が証明された。

(II) $d = 2p+3$ $(p \geqq 1)$ のとき、

(I) と同様に、

Sem. on Probab.
Vol.29 1968年
P1-95

$$I_p(u) = \int_0^{\cos^{-1}u} C_n^{\frac{d-2}{2}}(\cos\theta)\sin^{d-2}\theta\, d\theta$$

$$= \int_0^{\cos^{-1}u} C_n^{p+\frac{1}{2}}(\cos\theta)\sin^{2p+1}\theta\, d\theta$$

を考察する．

半整数の Gegenbauer 多項式は Legendre の陪多項式を用いて

$$C_n^{p+\frac{1}{2}}(\cos\theta) = \frac{2^p\cdot p!}{(2p)!\sin^p\theta} P_{n+p}^p(\cos\theta)$$

と表わせ、また定義により $P_{n+p}^p(x) = (1-x^2)^{p/2}\dfrac{d^p}{dx^p}P_{n+p}(x)$ であるから、積分変数を $x=\cos\theta$ に変換して $I_p(u)$ を書き直すと

$$I_p(u) = \frac{2^p\cdot p!}{(2p)!}\int_u^1 (1-x^2)^p\frac{d^p}{dx^p}P_{n+p}(x)\,dx$$

となるが、この積分の被積分関数を Legendre の多項式の漸化式

$$(8.11)\qquad (1-x^2)P_n'(x) = n(P_{n-1}(x) - xP_n(x))$$

と、Legendre の微分方程式を k 回微分して得られる方程式

$$(8.12)\qquad (1-x^2)\frac{d^{k+2}}{dx^{k+2}}P_n(x) - 2(k+1)x\frac{d^{k+1}}{dx^{k+1}}P_n(x) + (n+k+1)(n-k)\frac{d^k}{dx^k}P_n(x) = 0$$

を用いて変形すると、それを n の多項式と見れば、

$$(8.13)\qquad \begin{cases} (1-x^2)^p\dfrac{dx^p}{dx^p}P_{n+p}(x) = \pm(1-x^2)^{\frac{p}{2}}P_{n+p}(x)\,n^p + o(n^p) & (p=\text{偶数}) \\ (1-x^2)^p\dfrac{d^p}{dx^p}P_{n+p}(x) = \pm(1-x^2)^{\frac{p-1}{2}}(-xP_{n+p}(x) + P_{n+p-1}(x))\,n^p + o(n^p) & (p=\text{奇数}) \end{cases}$$

と書けることがわかる。また、$I_p(u)$ の積分の部分を

$$\int_u^1 (1-x^2)\frac{d^p}{dx^p}P_{n+p}(x)\,dx$$

$$= -(1-u^2)\frac{d^{p-1}}{dx^{p-1}}P_{n+p}(u) + 2p\int_u^1 x(1-x^2)^{p-1}\frac{d^{p-1}}{dx^{p-1}}P_{n+p}(x)\,dx$$

と変形しておく。これを用いれば、

$$I_d = \int_0^1 I_p(u)^2\,du$$

Sem. on Probab.
Vol.29 1968年
P1-95

$$= \left(\frac{2^p p!}{(2p)!}\right)^2 \times \left\{ \int_0^1 \left[(1-u^2)^p \frac{d^{p-1}}{du^{p-1}} P_{n+p}(u) \right]^2 du \right.$$

$$- 4p \int_0^1 (1-u^2) \frac{d^{p-1}}{du^{p-1}} P_{n+p}(u) \left[\int_u^1 x(1-x^2)^{p-1} \frac{d^{p-1}}{dx^{p-1}} P_{n+p}(x)\,dx \right] du$$

$$\left. + 4p^2 \int_0^1 \left[\int_u^1 x(1-x^2)^{p-1} \frac{d^{p-1}}{dx^{p-1}} P_{n+p}(x)\,dx \right]^2 du \right\}$$

は (8.13)を用いて n の次数を評価すると、p が奇数のときには

$$(8.14) \qquad \int_0^1 \left[(1-u^2)^p \frac{d^{p-1}}{du^{p-1}} P_{n+p}(u) \right]^2 du$$

$$= n^{2(p-1)} \int_0^1 (1-u^2)^{p+1} P_{n+p}(u)^2\, du + o(n^{2(p-1)})$$

2番目と3番目の積分についても n の最高次の次数は $n^{2(p-1)}$ で、それらの係数はそれぞれ、

$$(8.15) \qquad \left\{ \begin{array}{l} \int_0^1 (1-u^2)^{\frac{p+1}{2}} P_{n+p}(u) \left(\int_u^1 x (1-x^2)^{\frac{p-1}{2}} P_{n+p}(x)\,dx \right) du \\ \int_0^1 \left[\int_u^1 x(1-x^2)^{\frac{p-1}{2}} P_{n+p}(x)\,dx \right]^2 du \end{array} \right.$$

となる。ここで、(8.14) の右辺の積分は、Legendre の多項式の漸化式

$$(8.16) \qquad x P_n(x) = \frac{1}{2n+1} \left((n+1) P_{n+1}(x) + n P_{n-1}(x) \right)$$

と直交関係式

$$(8.17) \qquad \int_0^1 P_m(x) P_n(x)\,dx = \left\{ \begin{array}{ll} 0, & m \neq n,\ m-n=\text{偶数} \\ \frac{1}{2n+1}, & m = n \end{array} \right.$$

を用いて評価すれば

$$(8.18) \qquad \frac{1}{2n} \sum_{r=0}^{p+1} (-1)^r \binom{p+1}{r} 2^{-2r} \binom{2r}{r} + o\left(\frac{1}{n}\right)$$

となって、n^{-1} の大きさであることがわかる。(8.15)の 2つの積分は Schwarz の不等式を用いて評価できて、結局

$$(8.19) \qquad I_d = c(d)\, n^{2p-3} + o(n^{2p-3}) = c(d)\, n^{d-6} + o(n^{d-6})$$

と評価できることが示される。ここで $c(d)$ は次元 d によって決まる正の定数である。また、p が偶数のときにも、(8.13) の下の式を用いて評価することにより、(8.19) のなりたつことが示せる。なおこの場合には、

Sem. on Probab.
Vol.29 1968年
P1-95

(8.18) に対応する式は、

(8.20) $$\frac{1}{2n}\sum_{r=0}^{p}(-1)^r\binom{p}{r}2^{-(2r+1)}\frac{(2r)!}{r!\,(r+1)!}$$

となっている。

これで定理8.2の証明が終ったことになる。

Sem. on Probab.
Vol.29 1968年
P1-95

第4章 Stationary Gaussian process の ε-エントロピー

この章では stationary gaussian process の ε-エントロピーについて述べる。定常性よりこの process の ε-エントロピーは作用素 K_T：

$$K_T\varphi(t)=\int_0^T r(t-s)\varphi(s)\,ds$$

（ここで $r(t)$ は process の covariance function）の固有値を使って計算される。

§9では上のタイプの固有値問題に関する Widom [31] の結果を用い、ある条件をみたす spectral density をもつ process の ε-エントロピーを具体的に求め、spectral density $f(\lambda)$ の $\lambda\to\infty$ のときの収束の速さの、ε-エントロピーの $\varepsilon\to0$ のときの発散の速さへの反映の様子を見ていく。

§10では§5で定義した単位時間当りの ε-エントロピーが、discrete spectral をもつ process に対しては常に 0、continuous spectral をもつ process に対しては常に存在し、spectral density を使って表わされることを示す。

§9 Stationary Gaussian process の ε-エントロピー

この§では、$\xi\equiv\{\xi(t);-\infty<t<\infty\}$ を実数値をとる平均連続な stationary gaussian process とし、$E\xi(t)\equiv0$、$E\xi(t+s)\xi(s)=r(t)$ とする。

§5で述べたように、$\xi^T\equiv\{\xi(t);\,0\leqq t\leqq T\}$ $(T<\infty)$ の ε-エントロピー $H_\varepsilon(\xi^T)$ は $L^2[0,T]$ 上の積分作用素 K_T：

$$(9.1)\qquad K_T\varphi(t)=\int_0^T r(t-s)\varphi(s)\,ds\qquad \varphi\in L^2[0,T]$$

Sem. on Probab.
Vol.29 1968年
P1-95

の固有値 $\{\lambda_n \equiv \lambda_n(T)\}_{n\geq 1}$ により (5.10) で計算される。

stationary process の *covariance function* $\Gamma(t)$ は

$$(9.2) \qquad \Gamma(t) = \int_{-\infty}^{\infty} e^{it\lambda}\, dF(\lambda)$$

とスペクトル分解ができる。ここで dF は R^1 上の有界測度で *spectral measure* と呼ばれる。特に dF が絶対連続 (*Lebesgue measure* に関して) の場合その *density* (*spectral density* という) を $f(\lambda) = F'(\lambda)$ とおくと

$$(9.2') \qquad \Gamma(t) = \int_{-\infty}^{\infty} e^{it\lambda} f(\lambda)\, d\lambda$$

とできる。($f(\lambda) \in L^1(R^1)$, $f(\lambda) = f(-\lambda) \geq 0$ である)

最初に次のことに注意しておく。

<u>補題 9.1</u>

$\xi_j \equiv \{\xi_j(t)\}$ $(j = 1,2)$ を *spectral density* $f_j(\lambda)$ をもつ平均連続な *stationary gaussian process* とする。もし任意の λ に対して $f_1(\lambda) \geq f_2(\lambda)$ ならば、$H_\varepsilon(\xi_1^T) \geq H_\varepsilon(\xi_2^T)$ である。

<u>証明</u> ξ_j に対応する作用素 (9.1) を K_j、K_j の固有値を $\{\lambda_{j,n}\}_{n\geq 1}$ (大きい順に並べる) とすると

$$\lambda_{j,n} = \sup_{\substack{\varphi_1,\cdots,\varphi_{n-1} \\ \varphi_k \perp \varphi_l \ (k\neq l)}} \ \inf_{\varphi \perp \varphi_1,\cdots,\varphi_{n-1}} \frac{(K_j \varphi, \varphi)}{\|\varphi\|}$$

である。(*Riesz, Sz-Nagy* [24] 参照) ところで

$$\begin{aligned}(K_j\varphi, \varphi) &= \int_0^T\!\!\int_0^T \Gamma_j(t-s)\, \varphi(s)\varphi(t)\, ds\, dt \\ &= \int_0^T\!\!\int_0^T\!\!\int_{-\infty}^{\infty} e^{i(t-s)} f_j(\lambda)\, \varphi(s)\varphi(t)\, d\lambda\, ds\, dt \\ &= \int_{-\infty}^{\infty} f_j(\lambda) \left|\int_0^T e^{it\lambda}\varphi(t)\, dt\right|^2 d\lambda\end{aligned}$$

従って、$\forall \varphi \in L^2[0,T]$ に対して、$(K_1\varphi, \varphi) \geq (K_2\varphi, \varphi)$ である。故に $\forall n$ に対し、$\lambda_{1,n} \geq \lambda_{2,n}$ となり (5.10) より明らかに $H_\varepsilon(\xi_1^T) \geq H_\varepsilon(\xi_2^T)$ である。

covariance function $\Gamma(t)$ が (9.2') の型のとき 即ち連続スペクトルのみをもつとき、作用素 (9.1) に関し次の *Widom* [30] の結果があり それを用いて ε-エントロピー $H_\varepsilon(\xi^T)$ が計算できる。

Sem. on Probab.
Vol.29 1968年
P1-95

補題 9.2 (Widom [30])*)

$r(t)$ が (9.2′) の型で与えられたときの作用素 (9.1) の n 番目の大きさの固有値を $\lambda_n \equiv \lambda_n(T)$ $(n=1,2,\cdots)$ とすると $f(\lambda)$ の型に応じて次のことが成立する。

(I)
$$f(\lambda)=\begin{cases}1 & |\lambda|\leqq M\\ 0 & |\lambda|>M\end{cases} \qquad (M>0 \text{ は定数})$$

のとき

$$(9.3) \qquad \log\lambda_n \overset{**)}{\sim} -2n\log n \qquad (\text{as } n\to\infty)$$

(II) 連続かつ $[0,\infty)$ で単調減少な偶関数 $f_0(\lambda)$ があって、$f(\lambda)\sim f_0(\lambda)$ $(\text{as } \lambda\to\infty)$ かつ

$$\frac{\log f(\lambda)}{\lambda} \longrightarrow 0 \qquad (\text{as } \lambda\to\infty)$$

のとき

$$(9.4) \qquad \lambda_n \sim 2\pi f\left(\frac{\pi n}{T}+o(n)\right) \qquad (\text{as } \lambda\to\infty)$$

(III)
$$\frac{\log f(\lambda)}{\lambda} \longrightarrow -\alpha \qquad (\text{as } \lambda\to\infty) \quad (0<\alpha<\infty \text{ は定数})$$

のとき、

$$(9.5) \qquad \log\lambda_n \sim -n\pi\frac{K\left(\operatorname{sech}\frac{\pi T}{2\alpha}\right)}{K\left(\tanh\frac{\pi T}{2\alpha}\right)} \qquad (\text{as } n\to\infty)$$

ここで $K(r)$ は第一種完全連続積分 i.e

$$(9.6) \qquad K(r)=\int_0^{\frac{\pi}{2}}(1-r^2\sin^2\theta)^{-\frac{1}{2}}\,d\theta$$

(IV) λ が十分大きなところで $-\log f(\lambda)$ は convex で、

$$\frac{\log f(\lambda)}{\lambda} \longrightarrow -\infty \quad (\text{as } \lambda\to\infty)$$

のとき、

---

*) Widom [30] では積分区間が $[-1,1]$ の場合を扱っているが、証明を見れば区間が $[0,T]$ の場合に直せることが容易にわかる。

**) ～は比が 1 に近づくことを表わす。

Sem. on Probab.
Vol.29 1968年
P1-95

(9.7) $\log \lambda_n \sim \log f(\sigma_n) \qquad (\text{as } n \to \infty)$

ここでσ_nは次式の *unique* な解。

$$-\log f(\sigma) = 2n \log \frac{n}{\sigma}$$

証明 省略。

上の補題を用いればε-エントロピー$H_\varepsilon(\xi^T)$について次のことが導ける。

定理 9.1

$\xi = \{\xi(t)\,;\ -\infty < t < \infty\}$は *spectral density* $f(\lambda)$ をもつ、実数値をとる平均連続な *stationary gaussian process* とする。このとき、

(I) $|\lambda| > M$ ($M > 0$ は定数) で $f(\lambda) = 0$ のとき

(9.8) $H_\varepsilon(\xi^T) \sim \dfrac{1}{2} \dfrac{(\log \frac{1}{\varepsilon})^2}{\log \log \frac{1}{\varepsilon}} \qquad (\text{as } \varepsilon \to 0)$

(II) $f(\lambda)$は次をみたす。

(i) 連続かつ $[0, \infty)$で単調減少な偶関数 $f_0(\lambda)$ があり、

$$f(\lambda) \sim f_0(\lambda) \qquad (\text{as } \lambda \to \infty)$$

(ii)
$$\lim_{\lambda \to \infty} \frac{\log f(\lambda)}{\lambda} = 0$$

(iii)
$$f(\lambda + o(\lambda)) \sim f(\lambda) \qquad (\text{as } \lambda \to \infty)$$

(iv) ξの単位時間当りのε-エントロピー $\overline{H}_\varepsilon(\xi)$ *) が次をみたす。

$$\lim_{\delta, \varepsilon \to 0} \frac{\overline{H}_{(1+\delta)\varepsilon}(\xi)}{\overline{H}_\varepsilon(\xi)} = 1$$

このとき、$^\forall T$ (*fixed*) に対して、

(9.9) $\dfrac{1}{T} H_{\sqrt{T}\varepsilon}(\xi^T) \sim \overline{H}_\varepsilon(\xi) \qquad (\text{as } \varepsilon \to 0)$

---

*) §10で述べるように、$\overline{H}_\varepsilon(\xi)$ は次式で求まる。

$$\overline{H}_\varepsilon(\xi) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \log\left(\frac{f(\lambda)}{\theta^2} \vee 1\right)$$

ただし、θ^2は次式で定まる。

$$\int_{-\infty}^{\infty} (f(\lambda) \wedge \theta^2)\, d\lambda = \varepsilon^2$$

Sem. on Probab.
Vol.29 1968年
P1-95

特に、

(9.10) $f(\lambda) \sim c\,\lambda^{-(\alpha+1)} (\log\lambda)^{-\beta}$ $(\text{as } \lambda\to\infty)$ ($c>0$, $\alpha>0$, $-\infty<\beta<\infty$ は定数)

のとき、

(9.11) $H_\varepsilon(\xi^T) \sim c_0\, \varepsilon^{-\frac{2}{\alpha}} (\log\frac{1}{\varepsilon})^{-\frac{\beta}{\alpha}}$ $(\text{as } \varepsilon\to\infty)$

ただし、

(9.12) $c_0 = \frac{\alpha+1}{2\pi} \left(\frac{2^{1-\beta} c(\alpha+1)\, T^{\alpha+1}}{\alpha^{1-\beta}} \right)^{\frac{1}{\alpha}}$

(III) $\lim_{\lambda\to\infty} \frac{\log f(\lambda)}{\lambda} = -\alpha$ ($0<\alpha<\infty$ は定数) のとき.

(9.13) $H_\varepsilon(\xi^T) \sim c_0 (\log\frac{1}{\varepsilon})^2$ $(\text{as } \varepsilon\to 0)$

ただし、

(9.14) $c_0 = \frac{1}{\pi} \frac{K(\tanh\frac{\pi T}{2\alpha})}{K(\operatorname{sech}\frac{\pi T}{2\alpha})}$ ($K(\cdot)$ は (9.6) で定義されたもの)

<u>証明</u> 作用素 (9.1) の固有値を $\{\lambda_n \equiv \lambda_n(T)\}_{n\geq 1}$ (大きい順に並べる) とするとき、$N=N(\varepsilon)$, $\theta=\theta(\varepsilon)$ を

(9.15) $\varepsilon^2 = \sum_{n=1}^{\infty} (\lambda_n \wedge \theta^2) \equiv N\theta^2 + \sum_{n=N+1}^{\infty} \lambda_n$

となるように定めると、定理 5.3 より、

(9.16) $H_\varepsilon(\xi^T) = \frac{1}{2} \sum_{n=1}^{N} \log\frac{\lambda_n}{\theta^2}$

(I) $f(\lambda)$ が (I) の仮定をみたすとき補題 9.2.(I) より

$$\log\lambda_n = -2n\log n + o(n\log n)$$

従って、$\lambda_N \geqq \theta^2 \geqq \lambda_{N+1}$ に注意すれば

(9.17)
$$\begin{aligned} H_\varepsilon(\xi^T) &= \frac{1}{2}\sum_{n=1}^{N} \log\lambda_n - \frac{N}{2}\log\theta^2 \\ &= -\sum_{n=1}^{N} n\log n + o(n\log n) + N^2\log N + o(N^2\log N) \\ &= \frac{1}{2} N^2 \log N + o(N^2\log N) \end{aligned}$$

一方 $\varepsilon^2 = N\theta^2 + \sum_{n=N+1}^{\infty} \lambda_n = N\theta^2 + o(N\theta^2)$ だから、

$$\log\varepsilon^2 = \log N + \log\theta^2 + o(1) = -2N\log N + o(N\log N)$$

Sem. on Probab.
Vol.29 1968年
P1-95

従って

$$(9.18) \qquad N = \frac{\log\frac{1}{\varepsilon}}{\log\log\frac{1}{\varepsilon}} + o\left(\frac{\log\frac{1}{\varepsilon}}{\log\log\frac{1}{\varepsilon}}\right)$$

(9.17) (9.18) より (9.8) を得る。

(II) $g(\lambda) = g(-\lambda) \geqq 0$ なる $g \in L^1(R^1)$ に対し

$$(9.19) \qquad \tilde{H}_\varepsilon(g) = \int_{-\infty}^{\infty} \log\left(\frac{g(\lambda)}{\theta^2} \vee 1\right)$$

ただし、θ^2 は次式で決める。

$$(9.20) \qquad \int_{-\infty}^{\infty} (g(\lambda) \wedge \theta^2)\, d\lambda = \varepsilon^2$$

と定義すると次のことが容易にわかる。

もし、$g(\lambda) = a^2 k(\lambda)$ $(-\infty < \lambda < \infty)$ ならば

$$(9.21) \qquad \tilde{H}_{a\varepsilon}(g) = \tilde{H}_\varepsilon(k)$$

もし、$g(\lambda) \leqq k(\lambda)$ $(-\infty < \lambda < \infty)$ ならば

$$(9.22) \qquad \tilde{H}_\varepsilon(g) \leqq \tilde{H}_\varepsilon(k)$$

もし、$g(\lambda) = k(\lambda)$ for $|\lambda| > M$ (M は定数) ならば、$^\exists \varepsilon_0 > 0$, $^\exists C$ (定数)、

$$(9.23) \qquad \tilde{H}_\varepsilon(g) = \tilde{H}_\varepsilon(k) + C \quad \text{for } \varepsilon < \varepsilon_0$$

さらに次のことがわかる。

もし、$g(\lambda) \sim k(\lambda)$ (as $\lambda \to \infty$) ならば、$^\exists \delta = \delta(\varepsilon)$ ($\downarrow 0$ as $\varepsilon \downarrow 0$)

$$(9.24) \qquad (1-\delta)\tilde{H}_{(1+\delta)\varepsilon}(k) \leqq \tilde{H}_\varepsilon(g) \leqq (1+\delta)\tilde{H}_{(1-\delta)\varepsilon}(k)$$

∵) 仮定から $^\forall \delta_0 > 0$ に対し、$^\exists M$.

$$\frac{1}{(1+\delta_0)^2} k(\lambda) \leqq g(\lambda) \leqq \frac{1}{(1-\delta)^2} k(\lambda) \qquad \text{for } |\lambda| > M$$

従って (9.21) (9.22) (9.23) から、$^\exists \varepsilon_0 > 0$, $^\exists C = C(\delta_0)$ (ε には無関係)

$$(9.25) \qquad \tilde{H}_{(1+\delta_0)\varepsilon}(k) - C \leqq \tilde{H}_\varepsilon(g) \leqq \tilde{H}_{(1-\delta_0)\varepsilon}(k) + C \quad \text{for } \varepsilon < \varepsilon_0$$

(9.25) から (9.24) が導ける。

さて、(II) の証明に入ろう。仮定と補題 9.2, (II) より

$$(9.26) \qquad \lambda_n \sim 2\pi f\left(\frac{\pi n}{T} + o(n)\right) \sim 2\pi f\left(\frac{\pi n}{T}\right) \quad (\text{as } n \to \infty)$$

そこで

Sem. on Probab.
Vol.29 1968年
P1-95

$$g(\lambda) = \pi f\left(\frac{\pi\lambda}{T}\right)$$

とおき、(9.16) (9.15) と (9.19) (9.20) を比較し、(9.24) を導いたのと同様に (9.26) より、$^{\exists}\delta = \delta(\varepsilon)$ ($\downarrow 0$ as $\varepsilon \downarrow 0$)

$$(9.27) \qquad \frac{1}{4}(1-\delta)\tilde{H}_{(1+\delta)\varepsilon}(g) \leqq H_{\varepsilon}(\xi^T) \leqq \frac{1}{4}(1+\delta)\tilde{H}_{(1-\delta)\varepsilon}(g)$$

ところで

$$\tilde{H}_{\varepsilon}(g) = \int_{-\infty}^{\infty} \log\left(\frac{\pi f(\frac{\pi\lambda}{T})}{\theta^2} \vee 1\right) d\lambda$$

$$= \frac{T}{\pi}\int_{-\infty}^{\infty} \log\left(\frac{f(\lambda)}{\frac{\theta^2}{\pi}} \vee 1\right) d\lambda$$

$$\varepsilon^2 = \int_{-\infty}^{\infty}(g(\lambda) \wedge \theta^2)\, d\lambda = \int_{-\infty}^{\infty}\left(\pi f\left(\frac{\pi\lambda}{T}\right) \wedge \theta^2\right) d\lambda = T\int_{-\infty}^{\infty}\left(f(\lambda) \wedge \frac{\theta^2}{\pi}\right) d\lambda$$

従って、

$$\tilde{H}_{\varepsilon}(g) = 4T\, \overline{H}_{\varepsilon/\sqrt{T}}(\xi)$$

(9.27) より、

$$T(1-\delta)\overline{H}_{\frac{(1+\delta)\varepsilon}{\sqrt{T}}}(\xi) \leqq H_{\varepsilon}(\xi^T) \leqq T(1+\delta)\overline{H}_{\frac{(1-\delta)\varepsilon}{\sqrt{T}}}(\xi)$$

$$\therefore \qquad (1-\delta)\overline{H}_{(1+\delta)\varepsilon}(\xi) \leqq \frac{1}{T}H_{\sqrt{T}\varepsilon}(\xi^T) \leqq (1+\delta)\overline{H}_{(1-\delta)\varepsilon}(\xi)$$

従って仮定(iv) より (9.9)を得る。

特に、$f(\lambda)$ が (9.10) をみたすときには、(i) ～ (iv) は明らかに成り立ち、(9.9)より具体的に計算し、(9.11)を得る。

(Ⅲ) $f(\lambda)$ が(Ⅲ) の仮定をみたすとき、補題 9.2 (Ⅲ)より

$$\log \lambda_n = -cn + o(n)$$

ただし、(9.14)の c_0 に対し、$c = \frac{1}{c_0}$. $N = N(\varepsilon)$ を (9.15) で定めると、

$$\sum_{n=N+1}^{\infty} \lambda_n = \sum_{N+1}^{\infty} e^{-cn+o(n)} = O(e^{-cN})$$

$$N\theta^2 = O(Ne^{-cN})$$

Sem. on Probab.
Vol.29 1968年
P1-95

$$\therefore \qquad \varepsilon^2 = N\theta^2 + \sum_{N+1}^{\infty} \lambda_n = O(N e^{-cN})$$

$$\therefore \qquad N = \frac{2}{c} \log \frac{1}{\varepsilon} + o\left(\log \frac{1}{\varepsilon}\right)$$

従って、

$$H_\varepsilon(\xi^T) = \frac{1}{2} \sum_1^N \log \lambda_n - \frac{N}{2} \log \theta^2$$

$$= \frac{1}{2}\left(-c \sum_1^N n + o\left(\sum_1^N n\right)\right) + \frac{c}{2} N^2 + o(N^2)$$

$$= \frac{c}{4} N^2 + o(N^2) = c_0 \left(\log \frac{1}{\varepsilon}\right)^2 + o\left(\left(\log \frac{1}{\varepsilon}\right)^2\right)$$

§ 10 Stationary Gaussian process の単位時間当りの ε－エントロピー

この § では、$\xi \equiv \{\xi(t); -\infty < t < \infty\}$ を実数値をとる平均連続な、*stationary gaussian process* とし、$E\xi(t) \equiv 0$、$E\xi(t+s)\xi(s) = r(t)$ とする。

単位時間当りの ε－エントロピー $\overline{H}_\varepsilon(\xi)$ は (5,12) で定義されるが、今の場合は定常性より、次式の極限が存在するときその極限で定義してよい。

$$(10,1) \qquad \overline{H}_\varepsilon(\xi) = \lim_{T\to\infty} \frac{1}{T} H_{\sqrt{T}\varepsilon}(\xi^T)$$

まず、*discrete spectrum* をもつ *process* に対して $H_\varepsilon(\xi) = 0$ となることを示そう。

spectral measure dF が *point mass* のみをもつとき，即ち $r(t)$ が

$$(10,2) \qquad r(t) = \sum_{n=-\infty}^{\infty} c_n \cos \lambda_n t$$

$$\left(c_n \geq 0,\ \sum_{n=-\infty}^{\infty} c_n < \infty,\quad c_n = c_{-n},\quad \lambda_n = -\lambda_{-n}\right)$$

と表わされるとき、ξ は *discrete spectrum* をもつという。

<u>定理 10,1</u> (*Baba*)

ξ を *discrete spectrum* をもつ *process* とすると

Sem. on Probab.
Vol.29 1968年
P1-95

(10.3) $\overline{H}_\varepsilon(\xi) = 0$

証明 仮定より

(10.4) $I(t) = c_0 + 2\sum_{n=1}^{\infty} c_n \cos \lambda_n t$

(10.5) $\xi(t) = \xi_0 + \sum_{n=1}^{\infty} (\xi_n \cos \lambda_n t + \eta_n \sin \lambda_n t)$

と展開できる。ここで、ξ_0は$N(0, c_0)$，$\xi_n, \eta_n\ (n \geq 1)$ は $N(0, 2c_n)$ に従う *random variable* で、$\xi_n, \eta_m\ (n \geq 0,\ m \geq 1)$は互に独立。*gaussian system* $\tilde{\xi} = \{\xi_0, \xi_1, \eta_1, \xi_2, \eta_2, \cdots\}$ に対し、*gaussian system* $\tilde{\dot{\xi}} = \{\dot{\xi}_0, \dot{\xi}_1, \dot{\eta}_1 \cdots\cdots \dot{\xi}_N, \dot{\eta}_N\}$ を $\dot{\xi}_n, \dot{\eta}_m$ は平均 0，$(\xi_n, \dot{\xi}_n), (\eta_m, \dot{\eta}_m)$ $(n = 0, 1, \cdots N,\ m = 1, \cdots, N)$ は互に独立なもので、

$$E|\xi_n - \dot{\xi}_n|^2 = E|\eta_m - \dot{\eta}_m|^2 = 2\theta^2 \quad (n = 0, 1, \cdots, N,\ m = 1, \cdots, N)$$

N, θ^2 は ε に *depend* して決まる数で

$$2(N+1)\theta^2 + 2\sum_{n \geq N+1} c_n = \varepsilon^2$$

なるものとする。*)

次に$\tilde{\xi}$から *discrete spectrum* をもつ *stationary gaussian process* $\dot{\xi}(t)\{\dot{\xi}(t)\}$ を次式で定義する。

(10.6) $\dot{\xi}(t) = \dot{\xi}_0 + \sum_{n=1}^{N} (\dot{\xi} \cos \lambda_n t + \dot{\eta}_n \sin \lambda_n t)$

この作り方から、$^\forall T > 0$ に対し、

$$I(\xi^T, \dot{\xi}^T) \leq I(\xi, \dot{\xi}) = I(\tilde{\xi}, \tilde{\dot{\xi}}) < \infty$$

であり、さらに

$$\int_0^T E|\xi(t) - \dot{\xi}(t)|^2 dt = \left(2(N+1)\theta^2 + 2\sum_{n \geq N+1} c_n\right) T = \varepsilon^2 T$$

従って、

$$H_{\sqrt{T}\varepsilon}(\xi^T) \leq I(\xi^T, \dot{\xi}^T) \leq I(\xi, \dot{\xi})$$

$$\therefore \quad \lim_{T\to\infty} \frac{1}{T} H_{\sqrt{T}\varepsilon}(\xi^T) \leq \lim_{T\to\infty} \frac{1}{T} I(\xi, \dot{\xi}) = 0$$

---

*) $\tilde{\xi}$に対し、$\dot{\xi}$ は $H_\varepsilon(\tilde{\xi}) = I(\tilde{\xi}, \dot{\xi})$ をみたしているものである。(ただし $c_0 \geq c_1 \geq c_2 \geq \cdots\cdots$ とする。)

Sem. on Probab.
Vol.29 1968年
P1-95

次に spectral density をもつ process の $\overline{H}_\varepsilon(\xi)$ を計算しよう。

§9と同様、作用素 K_T :

(10.7) $$K_T\varphi(t) = \int_0^T r(t-s)\varphi(s)\,ds$$

の固有値を大きさの順に並べ、それを $\lambda_n \equiv \lambda_n(T)$ ($n=1,2,\cdots$) とおく。$N_T(x_1,x_2)$ で、$x_2 \geqq \lambda_n(T) > x_1$ なる $\lambda_n(T)$ の個数を表わし、$N_T(x) = N_T(x,\infty)$ とする。$E(x_1,x_2) = \{x;\ x_2 \geqq 2\pi f(x) > x_1\}$ とし、$E(x) = E(x,\infty)$ とする。このとき次のことが成り立つ。

補題 10.1 (Pinsker [19])

$m(E(x))$ ($m(\cdot)$ は Lebesgue measure) の連続点 x に対し

(10.8) $$\lim_{T\to\infty}\frac{1}{T}N_T(x) = \frac{1}{2\pi}m(E(x))$$

証明 省略、(Pinsker [19]、定理 11.2.2 及びその Remark 参照)

この補題を使って次の定理が証明できる。

定理 10.2 (Pinsker [20])

$\xi \equiv \{\xi(t)\}$ を spectral density $f(\lambda)$ をもつ stationary Gaussian process とすると、

(10.9) $$\overline{H}_\varepsilon(\xi) = \frac{1}{4\pi}\int_{-\infty}^{\infty}\log\left(\frac{f(\lambda)}{\theta^2}\vee 1\right)d\lambda$$

ただし、θ^2 は次式で定める。

(10.10) $$\int_{-\infty}^{\infty}(f(\lambda)\wedge\theta^2)\,d\lambda = \varepsilon^2$$

証明 まず、$m(E(x))$ の連続点 x_1, x_2 に対し、

(10.11) $$\lim_{T\to\infty}\frac{1}{T}\sum_{x_2\geqq\lambda_n(T)>x_1}\lambda_n(T) = \int_{E(x_1,x_2)}f(\lambda)\,d\lambda$$

が成り立つことを示そう。

$^\forall\varepsilon>0$ (fix) に対して $[x_1,x_2]$ の分点 $\{a_j\}_{j=0}^{n}$ を $x_1 = a_0 < a_1 < \cdots < a_n = x_2$, a_j は $m(E(x))$ の連続点, $a_j - a_{j-1} < \dfrac{2\pi\varepsilon}{m(E(x_1,x_2))}$ をみたすようにとる。このとき補題 10.1 より、ある $T_0 > 0$ が存在し、$T > T_0$ に対し次式が成り立つ。

Sem. on Probab.
Vol.29 1968年
P1-95

$$\left|\frac{1}{T} N_T(a_{j-1}, a_j) - \frac{1}{2\pi} m(E(a_{j-1}, a_j))\right| < \frac{\varepsilon}{n x_2}$$

$$\therefore \quad \frac{1}{T} \sum_{x_2 \geqq \lambda_n(T) > x_1} \lambda_n(T) \geqq \sum_{j=1}^{n} a_{j-1}\left(\frac{1}{2\pi} m(E(a_{j-1}, a_j)) - \frac{\varepsilon}{n x_2}\right)$$

$$\geqq \sum_{j=1}^{n} \frac{a_{j-1}}{2\pi} m(E(a_{j-1}, a_j)) - \varepsilon \geqq \int_{E(x_1, x_2)} f(\lambda)\, d\lambda - 2\varepsilon$$

同様に、

$$\frac{1}{T} \sum_{x_2 \geqq \lambda_n(T) > x_1} \lambda_n(T) \leqq \int_{E(x_1, x_2)} f(\lambda)\, d\lambda + 2\varepsilon$$

故に (10、11) は成立。

・ ところで、

(10,12) $$H_{\sqrt{T}\varepsilon}(\xi^T) = \frac{1}{2} \sum_{n=1}^{\infty} \log\left(\frac{\lambda_n(T)}{\theta_T^2} \vee 1\right)$$

ここで、

(10、13) $$\sum_{n=1}^{\infty} (\lambda_n(T) \wedge \theta_T^2) = T\varepsilon^2$$

即ち

(10、13′) $$\sum_{n=1}^{\infty} \left(\frac{\lambda_n(T)}{T} \wedge \frac{\theta_T^2}{T}\right) = \varepsilon^2$$

$\forall \varepsilon > 0$ (fix) に対し、θ, θ_T を各々 (10、10), (10, 13) で定めるとき、

(10、14) $$\lim_{T\to\infty} \theta_T^2 = 2\pi\theta^2$$

を示そう。

$m(E(x))$ の不連続点に対し、任意の近さで $m(E(x))$ の連続点が存在するから、初めから $2\pi\theta^2$, θ_T^2 は $m(E(x))$ の連続点として以下の議論をしてよい。

(10、15) $$\varepsilon^2 = \int_{-\infty}^{\infty} (f(\lambda) \wedge \theta^2)\, d\lambda = \int_{E(0, 2\pi\theta^2)} f(\lambda)\, d\lambda + m(E(2\pi\theta^2)) \cdot \theta^2$$

(10、11) と補題 10,1 より、$\forall \delta > 0$ に対し T が十分大きければ、

$$\left|\int_{E(0, 2\pi\theta^2)} f(\lambda)\, d\lambda + m(E(2\pi\theta^2)) \cdot \theta^2 - \frac{1}{T} \sum_{\lambda_n(T) \leqq 2\pi\theta^2} \lambda_n(T)\right.$$

$$\left. - \frac{2\pi}{T} N_T(2\pi\theta^2)\, \theta^2 \right| < \delta$$

Sem. on Probab.
Vol.29 1968年
P1-95

(10.13′) と (10.15) より、

$$\left| \sum_{\lambda_n(T) \leqq \theta_T^2} \frac{\lambda_n(T)}{T} + \frac{\theta_T^2}{T} N_T(\theta_T^2) - \frac{1}{T} \sum_{\lambda_n(T) \leqq 2\pi\theta^2} \lambda_n(T) - \frac{2\pi}{T} N_T(2\pi\theta^2)\cdot\theta^2 \right|$$
$$< \delta$$

従って補題 10.1 より、$\theta_T^2 \leqq 2\pi\theta^2$ のとき、十分大きな T に対して

$$\delta > \frac{1}{T}(2\pi\theta^2 - \theta_T^2) N_T(2\pi\theta^2) > (2\pi\theta^2 - \theta_T^2)\left(\frac{1}{2\pi} m(2\pi\theta^2) - \delta\right)$$
$$\equiv M(2\pi\theta^2 - \theta_T^2) \qquad (M \text{ は } T \text{ に無関係})$$

同様に、$\theta_T^2 \geqq 2\pi\theta^2$ なる十分大きな T に対しては、

$$\delta > M'(\theta_T^2 - 2\pi\theta^2) \qquad (M' \text{ は } T \text{ に無関係})$$

従って、(10.14) が成立。

最後に (10.9) を証明する。

$^\forall\delta_1 > 0$ (fix) に対し、$0 = a_0 < a_1 < \cdots$ を、a_j ($j = 1, 2, \cdots$) は $m(E(x))$ の連続点で、$a_j - a_{j-1} < \delta_1$ととる。このとき $a_k \leqq 2\pi\theta^2 < a_{k+1}$ なる k が存在し、(10.14) より T が十分大きければ $a_{k-1} < \theta_T^2 < a_{k+1}$ かつ $|\theta_T^2 - 2\pi\theta^2| < \delta_1$ である。従ってこのとき、

$$\frac{1}{T} H_{\sqrt{T}\varepsilon}(\xi^T) = \frac{1}{2T} \sum_{n=1}^{N_T(\theta_T^2)} \log \frac{\lambda_n(T)}{\theta_T^2}$$

$$\leqq \frac{1}{2T} \sum_{j \geqq k-1} \left\{ \log \frac{a_{j+1}}{2\pi\theta^2 - \delta_1} \right\} \frac{T}{2\pi}\left(m(E(a_j, a_{j+1})) + \delta_2 \right)$$

$$\leqq \frac{1}{4\pi} \int_{-\infty}^{\infty} \log\left(\frac{f(\lambda)}{\theta^2} \vee 1\right) d\lambda + \delta_3$$

($T \to \infty$ のとき、$\delta_2, \delta_3 \to 0$)、同様に、

$$\frac{1}{T} H_{\sqrt{T}\varepsilon}(\xi^T) \geqq \frac{1}{4\pi} \int_{-\infty}^{\infty} \log\left(\frac{f(\lambda)}{\theta^2} \vee 1\right) d\lambda - \delta_4$$

($T \to \infty$ のとき、$\delta_4 \to 0$)、従って $T \to \infty$ として (10.9) を得る。

上の定理により *spectral density* が与えられれば、常に単位時間当りの ε-エントロピーは計算されるが、ここでいくつかの場合にそれを具体的に求める。(定理 9.1 との比較を考えながら) いずれも初等的な計算

Sem. on Probab.
Vol.29 1968年
P1-95

で求まるので証明は省略する。

(I) $f(\lambda) = 0$ for $|\lambda| > M$ (Mは定数) のとき、

(10.17) $$\overline{H}_\varepsilon(\xi) \sim \frac{A}{2\pi} \log \frac{1}{\varepsilon} \quad (\text{as } \varepsilon \to 0)$$

ここで、 $A = m(\{\lambda;\ f(\lambda) \neq 0\})$

(II) $f(\lambda) \sim c\,\lambda^{-(\alpha+1)}(\log \lambda)^{-\beta}$ (as $\lambda \to \infty$) ($c>0$, $\alpha>0$, $-\infty<\beta<\infty$ は定数) のとき、

(10.18) $$\overline{H}_\varepsilon(\xi) \sim c_0\, \varepsilon^{-\frac{2}{\alpha}} \left(\log \frac{1}{\varepsilon}\right)^{-\frac{\beta}{\alpha}} \quad (\text{as } \varepsilon \to 0)$$

ただし、 $$c_0 = \frac{1}{2\pi}\left(\frac{2^{1-\beta}\, c\, (\alpha+1)}{\alpha^{1-\beta}}\right)^{\frac{1}{\alpha}}$$

(III) $\dfrac{\log f(\lambda)}{\lambda} \to -\alpha$ (as $\lambda \to \infty$) ($0<\alpha<\infty$ は定数) のとき

(10.19) $$\overline{H}_\varepsilon(\xi) \sim \frac{1}{\alpha\pi}\left(\log \frac{1}{\varepsilon}\right) \quad (\text{as } \varepsilon \to 0)$$

Sem. on Probab.
Vol.29 1968年
P1-95

第5章 Gaussian process の絶対連続性, path の連続性との関係

§11 Gaussian process の絶対連続性, path の連続性との関係

2つの Gaussian process の間の絶対連続性, Gaussian process の path の連続性については、これまでに多くの研究がなされている。この§ではこれらの性質と ε-エントロピーとの関係について述べる。

最初に互に絶対連続な2つの Gaussian process の ε-エントロピーの, ε↓0 のときの無限大に発散する速さは等しい(若干の条件の下で)ことを示す。(定理11.4及びその系)

path の連続性と ε-エントロピーの関係については、これまでのところ何の結果も得られていないので、ここでは stationary Gaussian process の path の連続性, ε-エントロピーの各々に関して得られている結果を並記しその対応関係をながめるにとどめる。

[I] 絶対連続性との関係

$\xi \equiv \{\xi(t)\,;\,0 \leqq t \leqq T\}$, $\tilde{\xi} \equiv \{\tilde{\xi}(t)\,;\,0 \leqq t \leqq T\}$ $(T < \infty)$ を2つの実数値をとる平均連続な Gaussian process とする。$\Omega = R^{[0,T]}$, $\mathcal{B}$ を $R^{[0,T]}$ の cylinder set から生成される最小の σ-algebra とする。このとき ξ, $\tilde{\xi}$ は $(\Omega, \mathcal{B})$ 上の gauss measure P, $\tilde{P}$ を定める。

(11.1) $\quad m(t) = E\xi(t) \equiv 0 \qquad \Gamma(s,t) = E\xi(s)\xi(t)$

(11.2) $\quad \tilde{m}(t) = \tilde{E}\tilde{\xi}(t) \qquad \tilde{\Gamma}(s,t) = \tilde{E}(\tilde{\xi}(s)-\tilde{m}(s))(\tilde{\xi}(t)-\tilde{m}(t))$

とおく(E, $\tilde{E}$ は各々 P, $\tilde{P}$ による平均) $L^2\ (\equiv L^2(P))$ は $\{\xi(t)\,;\,0 \leqq t \leqq T\}$ の張る Hilbert 空間で, 内積は $(\xi(s),\xi(t))_P = E\xi(s)\xi(t)$ で定義したものとする。同様に, 内積を $(\xi(s),\xi(t))_{\tilde{P}} = \tilde{E}(\xi(s)-\tilde{E}\xi(s))(\xi(t)-\tilde{E}\xi(t))$ で定義した Hilbert 空間を $L^2(\tilde{P})$ と記す。

Sem. on Probab.
Vol.29 1968年
P1-95

Gaussian process の絶対連続性に関しては次の定理が知られている。

定理 11.1 （ Rozanov 〔25〕）

$\xi \equiv \{\xi(t)\,;\, 0 \leqq t \leqq T\}$ と $\tilde{\xi} \equiv \{\tilde{\xi}(t)\,;\, 0 \leqq t \leqq T\}$ が互に絶対連続なためには、次の２つの条件が成り立つことが必要かつ十分である．

(A) $(F\xi(s), F\xi(t))_p = \tilde{r}(s,t) \qquad 0 \leqq s,t \leqq T$

によって、L^2上に可逆有界対称正値作用素が定義され、I − F は *Hilbert-Schmidt* 型である．（I は恒等作用素）

(B) $f(\xi(s)) = \tilde{m}(s) \qquad 0 \leqq s \leqq T$

によって、L^2上の線型汎関数が定義され，F の固有元 $\{\eta_n\}$ に対し $\sum_{n=1}^{\infty} f(\eta_n)^2 < \infty$ である．

証明　省略（ Rozanov〔25〕, Sato〔27〕参照）

K（$\tilde{K}$）を $r(s,t)$（$\tilde{r}(s,t)$）を *Kernel* とする $L^2[0,T]$ 上の積分作用素、即ち

(11.3) $\quad K\varphi(t) = \int_0^T r(s,t)\varphi(s)\,ds \qquad \left(\tilde{K}\varphi(t) = \int_0^T \tilde{r}(s,t)\varphi(s)\,ds\right)$

とし、K（$\tilde{K}$）の固有値（大きい順に並べる）を $\{\lambda_n\}_{n\geqq 1}$（$\{\tilde{\lambda}_n\}_{n\geqq 1}$）、対応する固有関数系（c.o.n.s. にとる）を $\{\varphi_n(t)\}$（$\{\tilde{\varphi}_n(t)\}$）とする。

§5で述べたように，$\xi, \tilde{\xi}$ の ε－エントロピー $H_\varepsilon(\xi)$, $H_\varepsilon(\tilde{\xi})$ は、固有値 $\{\lambda_n\}$, $\{\tilde{\lambda}_n\}$ によって決まる。そこで、$\xi \sim \tilde{\xi}$（互に絶対連続）のとき、$\{\lambda_n\}$ と $\{\tilde{\lambda}_n\}$ の間にどのような関係が成り立つかをみていこう。

無限次元行列

$$(11.4) \qquad S = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \sqrt{\lambda_2} & \\ 0 & & \ddots \end{pmatrix} \qquad \tilde{S} = \begin{pmatrix} \sqrt{\tilde{\lambda}_1} & & 0 \\ & \sqrt{\tilde{\lambda}_2} & \\ 0 & & \ddots \end{pmatrix}$$

を ℓ^2上の作用素と考えると次のことが成り立つ．

定理 11.2

ξ と $\tilde{\xi}$ が互に絶対連続ならば、I − A が *Hilbert-Schmidt* 型なる ℓ^2 上の対称作用素 A 及び ℓ^2上の直交作用素 B が存在し、

(11.5) $\quad \tilde{S}^2 = B^* S A S B$

とできる．（B^* は B の共役作用素）

Sem. on Probab.
Vol.29 1968年
P1-95

証明

(11.6) $\quad \xi_n = \frac{1}{\sqrt{\lambda_n}} \int_0^T \xi(t) \varphi_n(t)\, dt \qquad \left(\tilde{\xi}_n = \frac{1}{\sqrt{\tilde{\lambda}_n}} \int_0^T \tilde{\xi}(t) \varphi_n(t)\, dt \right)$

とおくと、§5で述べたように $\{\xi_n\}$ $(\{\tilde{\xi}_n\})$ は $L^2(P)$, $(L^2(\tilde{P}))$ の c.o.n.s. である。又 $\mathcal{L}^2 = L^2([0,T]\times\Omega, \mathcal{B}_T\times\mathcal{B}, dm\times dP)$ ($\mathcal{B}_T$ は $[0,T]$ の Borel 集合族、dm は Lebesgue measure) とおくと、

(11.7) $\quad \xi(t) = \sum_{n=1}^{\infty} \sqrt{\lambda_n}\, \varphi_n(t)\, \xi_n \qquad in\ \mathcal{L}^2$

である。

定理11.1の F に対し、F の固有値、固有元 ($L^2(P)$ の c.o.n.s. にとる) を $\{\mu_n\}$, $\{\eta_n\}$ とすると

(11.8)
$$F\eta_n = \mu_n \eta_n , \quad \sum_{n=1}^{\infty} (1-\mu_n)^2 < \infty$$
$$(\eta_n, \eta_m)_P = \delta_{nm} , \quad (\eta_n, \eta_m)_{\tilde{P}} = (F\eta_n, F\eta_m)_P = \delta_{nm}\mu_n^2$$

である。$(\xi_n, \eta_m)_P = c_{nm}$, $(c_{nm})_{n,m\geq 1} = C$ とおくと C は直交行列、i.e.

$$\sum_{n=1}^{\infty} c_{\ell n} c_{mn} = \sum_{n=1}^{\infty} c_{n\ell}\, c_{nm} = \delta_{\ell m}$$

である。$\{ \varphi_n(t)\eta_m(\omega) ; n, m = 1, 2, \cdots \}$ は $\mathcal{L}^2$ の c.o.n.s. だから (11.7) より

$$\xi(t) = \sum_{n,m=1}^{\infty} \sqrt{\lambda_n}\, c_{nm}\, \varphi_n(t)\, \eta_m \qquad in\ \mathcal{L}^2$$

従って、

(11.9) $\quad F\xi(t) = \sum_{n,m=1}^{\infty} \sqrt{\lambda_n}\, c_{nm}\, \mu_m\, \varphi_n(t)\, \eta_m \qquad in\ \mathcal{L}^2$

故に定理11.1 (A)より

$$\tilde{\Gamma}(s,t) = (F\xi(s), F\xi(t))_P$$
$$= \sum_n \sum_{\ell,m} \sqrt{\lambda_\ell}\sqrt{\lambda_m}\, c_{\ell n} c_{mn}\, \mu_n^2\, \varphi_\ell(s)\varphi_m(t)$$

$\tilde{\Gamma}(s,t) \in L^2([0,T]\times[0,T])$ かつ $\{ \varphi_\ell(s)\varphi_m(t) ; \ell, m = 1, 2, \cdots \}$ は $L^2([0,T]\times[0,T])$ の c.o.n.s. であることから和の順序が交換でき、

(11.10) $\quad \tilde{\Gamma}(s,t) = \sum_{\ell,m} \sum_n \sqrt{\lambda_\ell}\sqrt{\lambda_m}\, c_{\ell n} c_{mn}\, \mu_n^2\, \varphi_\ell(s)\varphi_m(t)$

一方

$$\tilde{\Gamma}(s,t) = \sum_n \tilde{\lambda}_n\, \tilde{\varphi}_n(s)\tilde{\varphi}_n(t)$$

従って、

(11.11) $\quad \int_0^T\!\!\int_0^T \tilde{\Gamma}(s,t)\tilde{\varphi}_p(s)\tilde{\varphi}_q(t)\, ds\, dt = \delta_{pq}\tilde{\lambda}_p$
$$= \sum_{\ell,m,n} \sqrt{\lambda_\ell \lambda_m}\, c_{\ell n} c_{mn}\, \mu_n^2 \int_0^T\!\!\int_0^T \varphi_\ell(s)\varphi_m(t)\tilde{\varphi}_p(s)\tilde{\varphi}_q(t)\, ds\, dt$$

$\int_0^T \varphi_n(t)\tilde{\varphi}_m(t)\, dt = b_{nm}$, $(b_{nm})_{n,m\geq 1} = B$ とおくと、B は ℓ^2 上の直交作

Sem. on Probab.
Vol.29 1968年
P1-95

素で、(11.11) より

(11.12) $\quad \delta_{pq}\tilde{\lambda}_p = \sum_{l,m,n} \sqrt{\lambda_l \lambda_m}\, c_{ln} c_{mn} \mu_n^2\, b_{lp} b_{mq}$

従って、$D = \begin{pmatrix} \mu_1 & & 0 \\ & \mu_2 & \\ 0 & & \ddots \end{pmatrix}$ とおけば (11.12) は

$$\tilde{S}^2 = B^* S C D^2 C^* S B$$

を意味する．従って $A = CD^2C^*$ とおけば

$$\tilde{S}^2 = B^* S A S B$$

である。しかるに、$I - F$ が *Hilbert-Schmidt* 型だから、$I - D$，従って $I - D^2$ も *Hilbert-Schmidt* 型であり、$I - A = C(I - D^2)C^*$ も *Hilbert-Schmidt* 型となる．

<u>Remark</u> B は直交作用素だから、B^*SASB と SAS の固有値は、同じである。従って，($\tilde{S}^2$の) 固有値 $\{\tilde{\lambda}_n\}$ は SAS の固有値として求まる。

以上のことから、結局、$\{\lambda_n\}$ と $\{\tilde{\lambda}_n\}$ の関係を調べるには、$\tilde{S}^2$ と SAS の固有値の関係を調べればよい。そのために、補題を準備する．

<u>補題 11.1</u>

S_1, S_2 は *Hilbert* 空間上の完全連続な自己共役作用素，$S = S_1 + S_2$ とし、S, S_1, S_2 の n 番目の大きさの正固有値を λ_n^+, $\lambda_{1,n}^+$, $\lambda_{2,n}^+$ ，n 番目の負固有値を λ_n^-, $\lambda_{1,n}^-$, $\lambda_{2,n}^-$ とすると

(11.13) $\quad \lambda_{n+m-1}^+ \leqq \lambda_{1,n}^+ + \lambda_{2,m}^+$

(11.14) $\quad \lambda_{n+m-1}^- \geqq \lambda_{1,n}^- + \lambda_{2,m}^-$

<u>証明</u> (詳しくは *Riesz, Sz-Nagy* [24] 参照)

一般に、λ_n^+ (λ_n^-) は

$$\lambda_n^+ = \inf_{\substack{\varphi_1,\cdots,\varphi_{n-1} \\ \varphi_i \perp \varphi_j}} \ \sup_{\substack{\varphi \\ \varphi \perp \varphi_1,\cdots,\varphi_{n-1}}} \frac{(S\varphi,\varphi)}{\|\varphi\|^2}$$

$$\left(\lambda_n^- = \sup_{\prime\prime} \ \inf_{\prime\prime} \ \frac{(S\varphi,\varphi)}{\|\varphi\|^2} \right)$$

で求まり、実際に inf sup (sup inf) は，$\varphi_1, \cdots, \varphi_{n-1}, \varphi$ として $\lambda_1^+, \cdots, \lambda_n^+$ ($\lambda_1^-, \cdots, \lambda_n^-$) に対応する固有関数 $\varphi_1^+, \cdots, \varphi_n^+$ ($\varphi_1^-, \cdots, \varphi_n^-$) をとったときに

Sem. on Probab.
Vol.29 1968年
P1-95

attain されることから、(11.13),(11.14) は容易に導かれる。

補題 11.2

S は *Hilbert* 空間上の完全連続な自己共役な作用素、A は $I-A$ が完全連続である自己共役作用素とし、SAS は正値とする. S^2, SAS の n 番目に大きい固有値を各々 λ_n, μ_n とすると、$^\forall\delta>0$ に対し、$^\exists N=N(\delta)$: 自然数

$$(11.15)\qquad (1-\delta)\lambda_{n+N} \leq \mu_n \leq (1+\delta)\lambda_{n-N} \qquad (n=N+1, N+2, \cdots)$$

証明 $I-A$ の固有値を $\{\nu_n\}$ ($|\nu_1|\geq|\nu_2|\geq\cdots$ とする) 対応する固有元を $\{\varphi_n\}$ (c.o.n.s. にとる) とし、$H_n=\mathfrak{S}_{1\leq j\leq n}\{\varphi_j\}$ とする. $I-A$ は完全連続だから、$^\forall\delta>0$ に対し $^\exists N=N(\delta)$

$$(11.16)\qquad |\nu_n|\leq\delta \qquad \text{if}\quad n>N$$

作用素 A_1, A_2 を次の様に定義する。

$$(11.17)\qquad A_1=\begin{cases} I & \text{on } H_N \\ A & \text{on } H_N^\perp \end{cases} \qquad A_2=\begin{cases} A-I & \text{on } H_N \\ 0 & \text{on } H_N^\perp \end{cases}$$

$A=A_1+A_2$ だから

$$(11.18)\qquad SAS = SA_1S + SA_2S$$

A_2 の値域は N 次元、従って SA_2S の値域も高々 N 次元だから、SA_2S の n 番目の大きさの正(負)固有値を $\mu_{2,n}^+$ ($\mu_{2,n}^-$) とすると

$$(11.19)\qquad \mu_{2,n}^+=\mu_{2,n}^-=0 \qquad \text{if}\quad n>N$$

一方、A_1 については定義より

$$(1-\delta)I\leq A_1\leq(1+\delta)I$$

従って、一般に $A\geq0$ なら $SAS^*\geq0$ であることに注意すれば、

$$(11.20)\qquad (1-\delta)S^2\leq SA_1S\leq(1+\delta)S^2$$

故に SA_1S の n 番目に大きい固有値を $\mu_{1,n}$ とすると

$$(11.21)\qquad (1-\delta)\lambda_n\leq\mu_{1,n}\leq(1+\delta)\lambda_n \qquad n=1,2,\cdots$$

(11.18) に補題 11.1 を用い、(11.19) に注意すれば、

$$\mu_n\leq\mu_{1,n-N}+\mu_{2,N+1}^+=\mu_{1,n-N} \qquad n>N$$
$$\mu_n\geq\mu_{1,n+N}+\mu_{2,N+1}^-=\mu_{1,n+N} \qquad n=1,2,\cdots$$

Sem. on Probab.
Vol.29 1968年
P1-95

よって

(11.22) $\mu_{1,n+N} \leqq \mu_n \leqq \mu_{1,n-N} \qquad n > N$

(11.21) (11.22) より (11.15) を得る。

定理 11.2 と補題 11.2 より、次の定理を得る。

<u>定理 11.3</u>

ξ と $\tilde{\xi}$ が互に絶対連続のとき，ξ，$\tilde{\xi}$ に対応する固有値を $\{\lambda_n\}$, $\{\tilde{\lambda}_n\}$ (大きい順に並べる) とすると，$\forall \delta > 0$ に対し、$\exists N = N(\delta)$:

(11.23) $(1-\delta)\lambda_{n+N} \leqq \tilde{\lambda}_n \leqq (1+\delta)\lambda_{n-N} \qquad n > N$

<u>証明</u> 定理 11.2，その *Remark*，補題 11.2 より明らか。

この定理は，2つの *gaussian process* が互に絶対連続ならば、2つの *process* に対応する固有値の *asymptotic* な挙動が等しいことを示している。このことを ε-エントロピーの言葉で表わそう。

gaussian process の ε-エントロピーはその固有値を使って、

(11.24) $H_\varepsilon(\xi) = \frac{1}{2}\sum_{n=1}^{\infty} \log\left(\frac{\lambda_n}{\theta^2} \vee 1\right)$

ただし、

(11.25) $\sum_{n=1}^{\infty} (\lambda_n \wedge \theta^2) = \varepsilon^2$

で計算されるから、結局 $\{\lambda_n\}$ と $\{\tilde{\lambda}_n\}$ が (11.23) をみたしているとき、$H_\varepsilon(\xi)$ と $H_\varepsilon(\tilde{\xi})$ の間にどんな関係が成り立つかを調べればよい。

<u>定理 11.4</u>

ξ と $\tilde{\xi}$ が互に絶対連続ならば、$\forall \varepsilon > 0$, $\forall \delta > 0$ に対し、$\exists \delta_0 \equiv \delta_0(\delta, \varepsilon)$: δ を固定したとき，$\delta_0(\delta,\varepsilon) \downarrow 0$ as $\varepsilon \downarrow 0$ で、

(11.26) $(1-\delta_0) H_{(1+\delta)(1+\delta_0)\varepsilon}(\xi) \leqq H_\varepsilon(\tilde{\xi}) \leqq (1+\delta_0) H_{(1-\delta)(1-\delta_0)\varepsilon}(\xi)$

<u>系</u>

ξ と $\tilde{\xi}$ が互に絶対連続で、ξ の ε-エントロピー $H_\varepsilon(\xi)$ が、

(11.27) $\lim_{\varepsilon,\delta\to 0} \dfrac{H_{(1+\delta)\varepsilon}(\xi)}{H_\varepsilon(\xi)} = 1$

をみたすとき、次式が成立する。

Sem. on Probab.
Vol.29 1968年
P1-95

$$(11.28) \qquad \lim_{\varepsilon\to 0} \frac{H_\varepsilon(\xi)}{H_\varepsilon(\tilde{\xi})} = 1$$

定理の証明のために、二つの gaussian process を ξ, $\tilde{\xi}$ とし応する固有値を $\{\lambda_n\}$, $\{\tilde{\lambda}_n\}$ （大きい順に並べる）とし、固有値の間の関係が、ε- エントロピーにどう反映するかに関する次の補題を準備する。

<u>補題 11.3</u>

(I) すべての n に対し、$\lambda_n \leqq \tilde{\lambda}_n$ ならば、$H_\varepsilon(\xi) \leqq H_\varepsilon(\tilde{\xi})$

(II) すべての n に対し、$\lambda_n = a^2\tilde{\lambda}_n$ ならば、$H_{a\varepsilon}(\xi) = H_\varepsilon(\tilde{\xi})$

(III) ある番号Nから先の n に対し、$\lambda_n = \tilde{\lambda}_n$ ならば、$^{\exists}\varepsilon_0$, $^{\forall}\varepsilon < \varepsilon_0$ に対し、$H_\varepsilon(\xi) = H_\varepsilon(\tilde{\xi}) + C$，ここで $C = \frac{1}{2}\sum_{n=1}^{N} \log\frac{\lambda_n}{\tilde{\lambda}_n}$ で、ε に無関係な定数.

(IV) すべての n に対し、$\tilde{\lambda}_n = \lambda_{n+N}$ （N は定数）ならば、$^{\exists}\delta_1, \delta_2$; $\delta_i \equiv \delta_i(\varepsilon) \downarrow 0$ as $\varepsilon \downarrow 0$ $(i=1,2)$

$$H_{(1+\delta_1)\varepsilon}(\xi) \leqq (1+\delta_2) H_\varepsilon(\tilde{\xi})$$

<u>証明</u> (I),(II),(III) は (11.24),(11.25)から直ちにわかる。(IV) は多少の計算を要するが、(11.24),(11.25) から容易に導かれる.

<u>定理 11.4 の証明</u> ξ と $\tilde{\xi}$ が互に絶対連続だから、定理 11.3 より、$1 > {}^{\forall}\delta > 0$ に対し、$^{\exists}N$ （以下この δ と N を固定しておく）

$$(11.23') \qquad \frac{1}{(1+\delta)^2}\lambda_{n+N} \leqq \tilde{\lambda}_n \leqq \frac{1}{(1-\delta)^2}\lambda_{n-N} \qquad (n > N)$$

が成り立つ。今、ξ_1, ξ_2 を対応する固有値 $\{\lambda_{1,n}\}$, $\{\lambda_{2,n}\}$ が

$$\lambda_{1,n} = \begin{cases} (1-\delta)^2\tilde{\lambda}_n & n \leqq N \\ \lambda_{n-N} & n > N \end{cases}$$

$$\lambda_{2,n} = \begin{cases} (1+\delta)^2\tilde{\lambda}_n & n \leqq N \\ \lambda_{n+N} & n > N \end{cases}$$

であるような gaussian process とすると.

$$\frac{1}{(1+\delta)^2}\lambda_{2,n} \leqq \tilde{\lambda}_n \leqq \frac{1}{(1-\delta)^2}\lambda_{1,n} \qquad (n=1,2,\cdots)$$

Sem. on Probab.
Vol.29 1968年
P1-95

従って、補題 11.3 の (I), (II) より、

$$(11.29) \qquad H_{(1+\delta)\varepsilon}(\xi_2) \leqq H_\varepsilon(\tilde{\xi}) \leqq H_{(1-\delta)\varepsilon}(\xi_1)$$

さらに、$n > N$ に対し、$\lambda_{1,n} = \lambda_{n-N}$, $\lambda_{2,n} = \lambda_{n+N}$ だから補題 11.3 の (IV) より $\exists\ \delta_1 \equiv \delta_1(\varepsilon)$, $\delta_1' \equiv \delta_1'(\varepsilon)$, $\delta_2 \equiv \delta_2(\varepsilon)$, $\delta_2' \equiv \delta_2'(\varepsilon)$ ($\downarrow 0$ as $\varepsilon \downarrow 0$)

$$(11.30) \qquad H_\varepsilon(\xi_1) \leqq (1+\delta_1')\, H_{(1-\delta_1)\varepsilon}(\xi)$$

$$(11.31) \qquad H_\varepsilon(\xi_2) \geqq (1-\delta_2')\, H_{(1+\delta_2)\varepsilon}(\xi)$$

(11.29), (11.30), (11.31) より

$$(11.32) \qquad (1-\delta_2')\, H_{\varepsilon_2}(\xi) \leqq H_\varepsilon(\tilde{\xi}) \leqq (1+\delta_1')\, H_{\varepsilon_1}(\xi)$$

ただし、$\varepsilon_1 = (1-\delta)(1-\delta_1)\varepsilon$, $\varepsilon_2 = (1-\delta)(1+\delta_2)\varepsilon$。従って、$\delta_0 \equiv \delta_0(\delta, \varepsilon)$ を、$\delta_0 \geqq \delta_1, \delta_1', \delta_2, \delta_2'$ かつ各 δ (fix) に対し $\delta_0(\delta,\varepsilon) \downarrow 0$ as $\varepsilon \downarrow 0$ ととれば、(11.26) が成立。

<u>系の証明</u>　条件 (11.27) が成り立つときには、(11.26) と、各 δ に対し、$\lim_{\varepsilon\to 0} \delta_0(\delta,\varepsilon) = 0$ であることから、(11.28) が導かれる。

<u>*Remark 1.*</u>　系の条件 (11.27) は例えば、ε-エントロピー $H_\varepsilon(\xi)$ の *main term* が $\varepsilon^{-\alpha}(\log\frac{1}{\varepsilon})^{\beta}$ ($\alpha > 0$, $-\infty < \beta < \infty$, *or* $\alpha = 0$, $\beta \geqq 0$) の場合には、みたされている。従って、このセミナー・ノートで具体的に ε-エントロピーを計算した *gaussian process* の多くの場合に、(11.27) はみたされ、系が適用できる。

<u>*Remark 2.*</u>　2つの *process* が互に絶対連続のとき、定理 11.1 の作用素 F は、I − F が *Hilbert-Schmidt* 型であるが、定理 11.3, 定理 11.4 の証明には、I − F の完全連続性しか使っていない。従って、(11.26) 又は (11.28) は、$\xi \sim \tilde{\xi}$ よりもかなり弱い主張である。(11.26) 又は (11.28) をみたしていても、互に絶対連続でないような例は容易に構成できる。

<u>例 1</u>　ブラウン運動

ξ がブラウン運動のとき、$H_\varepsilon(\xi) = \frac{2T^2}{\pi^2}\cdot\frac{1}{\varepsilon^2} + o\left(\frac{1}{\varepsilon^2}\right)$ (§5, (5.11)) で、(11.27) がみたされる。従って、系より、$\tilde{\xi}$ をブラウン運動と互に絶対連続な *Gaussian process* とすると、

$$H_\varepsilon(\tilde{\xi}) = \frac{2T^2}{\pi^2}\cdot\frac{1}{\varepsilon^2} + o\left(\frac{1}{\varepsilon^2}\right)$$

Sem. on Probab.
Vol.29 1968年
P1-95

例 2 N重マルコフ過程

ξ を定理 7.1 の N重マルコフ過程とするとやはり系が適用でき、$\tilde{\xi}$ を ξ と互に絶対連続な *gaussian process* とすると、

$$H_\varepsilon(\tilde{\xi}) \asymp \varepsilon^{-\frac{2}{2N-1}}$$

例 3

ξ が *spectral density* $f(\lambda) \sim c\,\lambda^{-(\alpha+1)}(\log\lambda)^{-\beta}$ をもつ *stationary Gaussian process* のときも系が適用でき、$\tilde{\xi}$ を ξ と互に絶対連続な *gaussian process* とすると、§9, (9.11) より、

$$H_\varepsilon(\tilde{\xi}) \sim c_0\,\varepsilon^{-\frac{2}{\alpha}}\left(\log\frac{1}{\varepsilon}\right)^{-\frac{\beta}{\alpha}}$$

ここで c_0 は (9.12) で与えられる。

〔II〕 *path* の連続性との関係

ここでは、*stationary gaussian process* の場合に、例をあげて、*path* の連続性と ε − エントロピーの対応をながめるにとどめる。

よく知られているように、平均連続な *stationary Gaussian process* は、確率 1 で連続であるか、さもなくば確率 1 でどの長さ有限の *time interval* でも *unbounded* である。(Balyaev 〔2〕)

連続か否かの判定条件としては次のものがある。(Nisio 〔18〕)

$\xi(t)$ の *spectral measure* を $F(d\lambda)$ とし、$s_j = \sqrt{F(2^{j+1}) - F(2^j)}$ とおくと、$\xi(t)$ の *path* が連続 (確率 1 で) かつ $E\{\sup_{0\leqq t\leqq T}|\xi(t)|\} \leqq \infty$ ならば

$$\sum_{j=1}^{\infty} s_j < \infty \tag{11.32}$$

逆に、$\{s_j\}$ が単調減少のときには、(11.32) が成り立てば、$\xi(t)$ の *path* は連続 (確率 1 で) ある。

このことから、

例 1 $\xi(t)$ は次の *spectral density* $f(\lambda)$ をもつ *stationary gaussian process* とする。

Sem. on Probab.
Vol.29 1968年
P1-95

$$f(\lambda) \sim \frac{1}{\lambda \prod_{j=1}^{n-1} (\log_{(j)} \lambda)^2 (\log_{(n)} \lambda)^{2+\delta}} \qquad (\text{as } \lambda \to \infty)$$

このとき、$\delta > 0$ ならば連続であり、$\delta = 0$ ならば、*path* は各区間上で *unbounded* である。

一方この *process* の ε-エントロピーは、補題 9.1 の(II)，定理 9.1 より、

$$\log H_\varepsilon(\xi^T) \sim T^2 \varepsilon^{-2} \prod_{j=1}^{n-2} (\log_{(j)} \tfrac{1}{\varepsilon})^{-2} (\log_{(n-1)} \tfrac{1}{\varepsilon})^{-2-\delta}$$

次に *path* の *Hölder* 連続性に関する *Sirao, Watanabe* [29] の結果を使えば、次のことがわかる。

<u>例 2</u>　$\xi(t)$ は次の *spectral density* $f(\lambda)$ をもつ *stationary gaussian process* とする.

$$f(\lambda) \sim c\, \lambda^{-(1+\alpha)} (\log \lambda)^{-\beta} \quad \text{as } \lambda \to \infty \quad (0 < \alpha < 1, -\infty < \beta < \infty \text{ or, } \alpha = 1, \beta \leqq 0)$$

このとき，確率 1 で

$$\lim_{h \to 0} \sup_{0 \leqq t, t+h \leqq 1} \frac{\xi(t+h) - \xi(t)}{\sqrt{2 |\log h| \sigma(h)}} = 1$$

ただし、$\sigma(h) = E(\xi(t+h) - \xi(t))^2 \sim \dfrac{c' h^\alpha}{|\log h|^\beta}$　(c' は h に無関係)

一方この *process* の ε-エントロピーは、定理 9.1 (II) より.

$$H_\varepsilon(\xi^T) \sim c_0\, \varepsilon^{-\frac{2}{\alpha}} (\log \tfrac{1}{\varepsilon})^{-\frac{\beta}{\alpha}}$$

ただし. c_0 は (9.12) で与えられる。

Sem. on Probab.
Vol.29 1968年
P1-95

附録 ε－エントロピーの他の定義

Kolmogorov 〔13〕以外の確率分布のε－エントロピーについては、*A. Rényi* による有限次元分布に対する *dimension* の定義，そしてそれを確率過程の場合に拡張した *Rudemo*〔26〕の仕事と、最近の E.C. *Ponser*, E.R. *Rodemich*, H. *Rumsey*〔21〕〔22〕らによる *probabilistic metric space* の、ε－δエントロピー，*gaussian process* の *product* ε－エントロピーなどの定義があるが、これらは *Shannon* の *dimension rate*. *Kolmogorov-Tichomirov*〔14〕の距離空間のε－エントロピーの変形と考えられる。ここでは〔21〕,〔22〕における定義と、それらと *Kolmogorov*〔13〕による情報量を用いてのε－エントロピーとの関係について述べる。また、*dimension rate*, *Rényi*〔23〕の定義，〔14〕の距離空間のε－エントロピーについても簡単に触れる。

1. Probabilistic metric space のε－δエントロピー（〔21〕参照）

(X, μ, d) が *probabilistic metric space* であるとは、1) X は距離 d による完備、可分な距離空間である. 2) X の開集合全体から生成される *Borel* 族（の μ による完備化）上に完備な確率測度 μ が定義されている。ときにいう.

高々可算個の集合族（互に共通部分のない）$\mathfrak{U} = \{X_\alpha\}_{\alpha\in A}$ が、X のε－δ分割（$\delta \geq 0$, $\varepsilon > 0$）であるとは、1) 可測集合 X_α の直径は高々ε である。 2) $\mu(\bigcup_\alpha X_\alpha) \geq 1-\delta$ 、となっている時にいう。

$\mathcal{U}_{\varepsilon,\delta}$ ＝ X のε－δ分割全体、とし、$\mathfrak{U} \in \mathcal{U}_{\varepsilon,\delta}$ のエントロピー $H(\mathfrak{U})$ を次のように定義する： $\mathfrak{U} = \{X_n\}$, $\mu(X_k) = p_k$ $(k \geq 1)$、$q_k = p_k / \sum_{n\geq 1} p_n$ としたとき、

$$H(\mathfrak{U}) = -\sum_k q_k \log q_k$$

Sem. on Probab.
Vol.29 1968年
P1-95

probabilistic metric space $X=(X,\mu,d)$ の ε-δ エントロピーを

$$H_{\varepsilon,\delta}(X)=\inf_{U\in\mathcal{U}_{\varepsilon,\delta}} H(U)$$

で定義し、$\delta=0$ のときには特に ε-エントロピーという。

$$H_{\varepsilon}(X)=\inf_{U\in\mathcal{U}_{\varepsilon,0}} H(U)$$

このとき、$\mathcal{U}_{\varepsilon,\delta}$ が空でないこと、また $\delta>0$ ならば $H_{\varepsilon,\delta}(X)<\infty$、$^\forall\varepsilon>0$ $^\forall\delta\geq 0$ に対して、$^\exists U\in\mathcal{U}_{\varepsilon,\delta}$, $H(U)=H_{\varepsilon,\delta}(X)$ などが示される。そして〔21〕では平均連続な確率過程 $X(t)$, $t\in[0,1]$ によって導かれる, $L^2[0,1]$ 上の確率測度を持つ *probabilistic metric space* を考え、その ε-エントロピーが有限 ($^\forall\varepsilon>0$) になるための十分条件が $\sum_{n\geq 1} n\lambda_n<\infty$ であり、そのとき $H_\varepsilon<C(\Sigma n\lambda_n)\,\varepsilon^{-2}$ であることを示した。そして $\Sigma n\lambda_n=\infty$ ならば、$^\forall\varepsilon>0$ に対して $H_\varepsilon=\infty$ になるような平均連続な確率過程の存在を示した。ただし、ここで, $\{\lambda_n\}_{n\geq 1}$ は $\Gamma(s,t)=EX(s)X(t)$ を核とする $L^2[0,1]$ の積分作用素の固有値：$\lambda_1\geq\lambda_2\geq\cdots\geq 0$ である。

<u>Remark 1.</u> Shannon〔28〕は dimension rate を次のように「定義」した。

確率過程 $X(t)$, $0\leq t<\infty$ を $0\leq t\leq T<\infty$ で考えた path 空間から、ある測度 $\delta>0$ を取り除いた残りの空間の ε-網*) の元の最小個数を $N(\varepsilon,\delta,T)$ としたとき、

$$\lim_{\delta\to 0}\lim_{\varepsilon\to 0}\lim_{T\to 0}\frac{\log N(\varepsilon,\delta,T)}{T\log\varepsilon}$$

を $X(t)$ の dimension rate という。

これは明らかに数学者を満足させる定義ではないが〔21〕の ε-エントロピーは、これの精密化への一つの方向であると考えられよう。なお、

---

*) 距離空間 X の部分集合 A が X の部分集合 Y の ε-網であるとは、$^\forall y\in Y$ に対して、$^\exists a\in A$, $d(y,a)\leq\varepsilon$ をみたしているときにいう。また、X の部分集合族 γ が Y の ε-被ふくであるとは、$Y\subset\bigcup_{\Gamma\in\gamma}\Gamma$ かつ $^\forall\Gamma\in\gamma$ に対して Γ の直径は 2ε を超えないときにいう。

Sem. on Probab.
Vol.29 1968年
P1-95

Rényi 〔23〕の場合には測度 $\delta > 0$ を取り除かなくとも有限個の ε-被ふく(前頁*)ができるような場合を考察している。また ε-被ふくの仕方にもいろいろな制限をつけてある。

<u>*Remark 2.*</u> *Kolmogorov-Tichomirov* 〔14〕による距離空間の ε-エントロピーは距離空間の全有界な部分集合 X について、$N(\varepsilon, X)$ を X の ε-被ふくの元の最小個数としたとき、

$$\log N(\varepsilon, X)$$

で定義される。

2、平均連続ガウス過程の *product* ε-エントロピー

$X(t)$, $t \in [0,1]$ を $EX(t) \equiv 0$, なる平均連続なガウス過程とする。そのとき、

$$(A.1') \qquad X(t) = \sum_{n \geq 1} \sqrt{\lambda_n^2}\, \xi_n(\omega)\, \varphi_n(t) \qquad (\text{in } L^2(\Omega))$$

と展開される。ここで、λ_n, $\varphi_n(t)$, $n \geq 1$ は $EX(s)X(t) = \Gamma(t,s)$ を核とする $L^2[0,1]$ での積分作用素の固有値と対応する固有関数である。また、$\forall \varepsilon > 0$ に対して、$\varepsilon_n \geq 0$, $n \geq 1$ を $\sum \varepsilon_n^2 \leqq \varepsilon^2$ をみたす数列とする。そのとき、$L^2[0,1]$ に $X(t)$から導かれる確率測度の入った *probabilistic metric space* を考えその空間の直積 ε-分割を、$X(t)$の φ_n 座標の ε_n 分割(n 番目の直線の巾 ε_nの区間による分割)の直積で定義する。そして Π'_ε を $\{\varepsilon_n\}$ のすべての取り方，巾 ε_nの区間のすべての取り方による直積 ε-分割の全体とする。次に、Π_ε を Π'_ε の中で可算個の *hypercube* の確率が1になっているようなもの全体とする。そのとき、$\Pi_\varepsilon \ni U$ のエントロピー $H(U)$ が定義できる。そのとき、$X = X(t)$の直積 ε-エントロピー $J_\varepsilon(X)$ を

$$J_\varepsilon(X) = \begin{cases} \inf\limits_{U \in \Pi_\varepsilon} H(U) & (\Pi_\varepsilon \neq \phi \text{ のとき}) \\ \infty & (\Pi_\varepsilon = \phi \text{ のとき}) \end{cases}$$

で定義する。

このとき、$H_\varepsilon(X)$ を上の *probabilistic metric space* の 1、の意味での ε-エントロピーとすれば、$H_\varepsilon(X)$の定義から、

$$(A.2) \qquad H_\varepsilon(X) \leqq J_\varepsilon(X)$$

Sem. on Probab.
Vol.29 1968年
P1-95

となるが、$J_\varepsilon(X)$については、精密な評価ができて、〔22〕でKolmogorov〔13〕Pinsker〔20〕にあるのと類似な結果が得られている。また、$J_\varepsilon(X)<\infty$の必要十分条件が

$$-\sum_{n\geq 1}\lambda_n\log\lambda_n<\infty$$

であることが示されている。一方Kolmogorovのε-エントロピーの場合、平均連続なガウス過程のε-エントロピーは常に有限であった。また、$H_\varepsilon(X)$と$J_\varepsilon(X)$がεのorderとしては同じであることも示されている。

3、Kolmogorovのε-エントロピーとの関係

以下、ある距離空間の値をとる確率変数Xが1、の意味でのprobabilistic metric spaceを導くような場合を考える、そして、$H_\varepsilon(X)$をXのKolmogorovのε-エントロピー，$\tilde{H}_\varepsilon(X)$をこのprobabilistic metric spaceのε-エントロピーとする，そのとき、

(A.3) $$H_\varepsilon(X)\leqq\tilde{H}_\varepsilon(X)\leqq J_\varepsilon(X)$$

の関係が成り立つことを示そう。

そのために次の補題を用いる。

補題

ξは可測空間$(X,\mathcal{B}_X)$ *) の値をとる確率変数，fを$(X,\mathcal{B}_X)$から可測空間$(Y,\mathcal{B}_Y)$ *) への可測関数とする．そのとき、

(A.4) $$I(\xi,f(\xi))=I(f(\xi),f(\xi))$$

が成り立つ．

Remark　情報量については一般に、$I(\xi,\eta)\geqq I(f(\xi),\eta)$の関係があるが（定理1.3.(Ⅲ)）、補題の場合には、fが1：1でなくともつねに等号が成り立つのである，ただし、左辺＝∞のときには，右辺＝∞の意味で等号が成り立つと考える．

証明　以下$f(\xi)\equiv\eta$と書くことにする．$I(\xi,\eta)$の定義より、

---

*) $\mathcal{B}_X$, $\mathcal{B}_Y$はそれぞれ、X，Yの部分集合族からなるσ-algebraとし、さらに、$\forall x\in X$, $\forall y\in Y$に対し、$\{x\}\in\mathcal{B}_X$, $\{y\}\in\mathcal{B}_Y$を仮定する．

Sem. on Probab.
Vol.29 1968年
P1-95

$$I(\xi,\eta) = \sup \sum_{i,j} P_{\xi\eta}(A_i \times B_j) \log \frac{P_{\xi\eta}(A_i \times B_j)}{P_\xi(A_i) P_\eta(B_j)}$$

$$= \sup \sum_{\substack{A_i \cap f^{-1}(B_j) \neq \phi \\ P_\eta(B_j) > 0}} P_{\xi\eta}(A_i \times B_j) \log \frac{P_{\xi\eta}(A_i \times B_j)}{P_\xi(A_i) P_\eta(B_j)}$$

ここで，補題 1.1 を用いれば

$$= \sup \sum_{\substack{A_i \subset f^{-1}(B_j) \\ P_\eta(B_j) > 0}} P_{\xi\eta}(A_i \times B_j) \log \frac{P_{\xi\eta}(A_i \times B_j)}{P_\xi(A_i) P_\eta(B_j)}$$

ところがこのとき、$P_{\xi\eta}(A_i \times B_j) = P(\xi \in A_i, f(\xi) \in B_j) = P(\xi \in A_i) = P_\xi(A_i)$ だから、上式は、

$$= \sup \sum_{\substack{A_i \subset f^{-1}(B_j) \\ P_\eta(B_j) > 0}} P_{\xi\eta}(A_i \times B_j) \log \frac{1}{P_\eta(B_j)}$$

ここで，$\{A_i\}$, $\{f^{-1}(B_j)\}$ が X の分割であり，$\log P_\eta(B_j)$ は i に無関係であることから、上式の和は、

$$-\sum_{P_\eta(B_j) > 0} P_\eta(B_j) \log P_\eta(B_j)$$

となることがわかる．最後に、容易に示される不等式：$p_i \geqq 0$ $(1 \leqq i \leqq n)$ $\sum_{i=1}^n p_i \leqq 1$ に対して，$-(p_1 + \cdots + p_n) \log (p_1 + \cdots + p_n) \leqq -\sum_{i=1}^n p_i \log p_i$ を用いれば，P_η が atom y_i $(i \geqq 1)$ のみからなるときには，

$$I(\xi,\eta) = -\sum_i P_\eta(\{y_i\}) \log P_\eta(\{y_i\}) = I(f(\xi), f(\xi))$$

また、P_η が atom のみから成っていないときには、

$$I(f(\xi), f(\xi)) = \infty = I(\xi, f(\xi))$$

が容易に示される。

<u>(A.3) の証明</u>　(A.2) によって第二の不等式は成り立っているから，第一の不等式のみを示せばよい。それには、$^\forall U \in \mathcal{U}_{\varepsilon,0}$ に対して、$H(U) \geqq H_\varepsilon(\xi)$ を示せば十分である．$U = \{X_n\}$ とおき、$^\forall n \geqq 1$ に対して、$x_n \in X_n$ を固定する．そして確率変数 X に対して、X'を、$E d(X,X')^2 \leqq \varepsilon^2$ をみたすように次のようにして作る。$X \in X_n$ のとき、$X' = x_n$、$^\forall n \geqq 1$ ，これから、

Sem. on Probab.
Vol.29 1968年
P1-95

$$H_\varepsilon(X) = \inf_{Ed(X,Y)^2 \leqq \varepsilon^2} I(X,Y) \leqq I(X,X')$$

X'はXの関数であるから、補題から、$I(X,X') = I(X',X')$となるが、

$$I(X',X') = H(X') = H(U)$$

これで(A.3)が示された。

Sem. on Probab.
Vol.29 1968年
P1-95

引用文献

[1] Baba,Y. ε-entropy of the Brownian motion with the multi-dimensional spherical parameter. Nagoya Math. J. 32 (1968), 31-40.

[2] Belyaev,Yu.K. Continuity and Hölder's conditions for sample functions of stationary Gaussian processes. Proc. 4-th Berkeley Symp. (1961), 23-33.

[3] Courant,R. and Hilbert,D. Mehoden der mathematischen Physik. Springer, (1931).

[4] Dobrušin,R.L. General formulation of Shannon's basic theorems of the theory of information. Amer. Math. Soc. Tanslation Ser. 2, 33 (1963).

[5] Doob,J.L. The elementary Gaussian processes, Ann. Math. Stat. 15 (1944), 229-282.

[6] Geerish,A.M. and Schultheiss,P.M. Information rates of non-Gaussian processes. IEEE Transac. Inform. Theory, IT-10 (1964), 265-271.

[7] Gelfand,I.M. Kolmogorov,A.N. and Yaglom,A.M. Количество информации и энтропия для непрерывных распределений. Доклд. Трет. Всес. Мате. Съез. (1956)

[8] Gelfand,I.M. and Yaglom,A.M. On the computation of amount of information about random functions contained in another such functions. Amer. Math. Soc. Translation,Ser.2, 12 (1959), 199-245.

[9] Hida,T. Canonical representations of Gaussian processes and their applications. Memoirs Coll. Sci. Kyoto Univ. Ser. A, 13 Math. (1960), 109-155.

[10] Kahane,J.P. Séries de Fourier aleatoires. Lecture note, (1963).

[11] Kazi,K. Kolmogorov's ε-entropy of some Gaussian processes. Ann. Inst. Stat. Math. 19 (1967), 479-503.

[12] Kazi,K. On ε-entropy of diffusion processes. to appear.

Sem. on Probab.
Vol.29 1968年
P1-95

[13] Kolmogorov,A.N. Theory of transmission of information. Amer. Math. Soc. Translation, Ser. 2, 33 (1963), 291-321.

[14] Kolmogorov,A.N. and Tihomirov,V.M. ε-entropy and ε-capacity of sets in function spaces. Amer. Math. Soc. Translation, Ser. 2, 17 (1961), 277-364.

[15] Lévy,P. A special problem of Brownian motion and a general theory of Gaussian random functions. Proc. Third Berkeley Symp. (1956), 133-175.

[16] Linikov,Yu.N. Evaluation of ε-entropy of random variables for small ε. Problems Inform. Transmis. 1 (1965), 12-18.

[17] McKean,H.P. Brownian motion with a several-dimensional time. Theory Prob. Appl. 8 (1963), 335-354.

[18] Nisio,M. 確率論セミナー 4月シンポジウム 講演 (1968)

[19] Pinsker,M.S. Information and information stability of random variables and processes. Holden-Dey Inc. (1964).

[20] Pinsker,M.S. Гауссовские источники. Проблемы Передачи Информ. 14 (1963), 59-100.

[21] Ponser,E.C. Rodemich,E.R. and Rumsey,H. Epsilon entropy of stochastic processes. Ann.Math.Stat. 38 (1967), 1000-1020.

[22] Ponser,E.C. Rodemich,E.R. and Rumsey,H. Product entropy of Gaussian distributions. to appear.

[23] Rényi,A. On the dimension and entropy of probability distributions. Acta. Math. Acad. Sci. Hung. 10 (1959), 193-215.

[24] Riesz,F. and Sz.Nagy,B. Functional analysis. (1952).

[25] Rozanov,Yu.A. On the density of one Gaussian measure with respect to another. Theory Prob. Appl. 7 (1962), 82-87.

Sem. on Probab.
Vol.29 1968年
P1-95

[26] Rudemo,M. Dimension and entropy for a class of stochastic processes. Magyer Tud. Akad. Math. Kutato Int. Kozl. 9 (1964),73-88.

[27] Sato,H. Gauss 測度の絶対連続性. Seminar on Prob. 24 (1966).

[28] Shannon,C.E. and Weaver,W. The mathematical theory of commnication. Univ. Ill. Press, (1949).

[29] Sirao,T. and Watanabe,H. On the Hölder continuity of stationary Gaussian processes. Proc. Japan Acad. 44 (1968), 482-484.

[30] Widom,H. Asymptotic behavior of the eigenvalues of certain integral equations II. Arch. Rational Mech. Anal. 17 (1964),215-229.