大偏差原理

慶應義塾大学 理工学部　田村 要造

関西学院大学 理工学部　千代延 大造

1　はじめに

独立同分布の確率変数の和に関する基本的な極限定理に大数の法則がある。大数の法則が成立しているとき、収束先である平均値の近くでの、より詳しい収束の様子は、分散が有限であれば、中心極限定理によって与えられる。これに対し、大偏差原理は、大数の法則のように分布がデルタ分布に収束しているとき、収束先から”大きく”離れた稀な事象の確率の漸近的なふるまいをみようとするものである。

1.1 Laplace の方法

たとえば、$\mathbb{R}$ 上の密度関数 f_ε を持つ確率測度 $\mu^\varepsilon(dx) = f_\varepsilon(x)dx$ の族 $\{\mu_\varepsilon : \varepsilon > 0\}$ が、$\mu_\varepsilon \to \delta_{\{x_0\}}$ $(\varepsilon \to 0)$ を満たしているものとする。また、$f_\varepsilon(x)$ $(x \neq x_0)$ が指数的に減少するとする。すなわち、$f_\varepsilon(x) = g_\varepsilon(x)\exp(-(1/\varepsilon)I(x))$ $(I \geq 0, I(x) = 0 \Leftrightarrow x = x_0)$ とする。ただし、$g_\varepsilon > 0$, $\sup_x(\varepsilon \log g_\varepsilon(x)) \to 0$ $(\varepsilon \to 0)$ を満たすとする。この確率の漸近挙動をみるために、1 次元有界区間上での**ラプラスの方法**を思い出してみる。

$-\infty < a < b < \infty$ とする。$\mu^\varepsilon([a,b]) = \displaystyle\int_a^b g_\varepsilon(x)\mathrm{e}^{-(1/\varepsilon)I(x)}dx$ とおく。このとき、次が成立する。

(1) $\displaystyle\lim_{\varepsilon\to 0} \varepsilon \log \mu^\varepsilon([a,b]) = \operatorname*{ess.\,sup}_{[a,b]}(-I).$

(2) (精密化)
$I(x)$ は $x_0 \in (a,b)$ で唯一の最小値 I_0 をとり、x_0 の近傍で C^2-級とし、$I''(x_0) > 0$ とする。また、$g_\varepsilon(x) = g(x)$ とする。すると、

$$
\begin{aligned}
&\mathrm{e}^{(1/\varepsilon)I_0}\mu^\varepsilon([a,b]) = \int_a^b g(x)\mathrm{e}^{-(1/\varepsilon)(I(x)-I(x_0))}dx \\
=\ & \int_{\{|x-x_0|\leq\delta\}} g(x)\mathrm{e}^{-(1/\varepsilon)(I(x)-I(x_0))}dx + \int_{\{|x-x_0|>\delta\}} g(x)\mathrm{e}^{-(1/\varepsilon)(I(x)-I(x_0))}dx \\
=\ & I_1 + I_2,
\end{aligned}
$$

となり、

$$
\begin{aligned}
I_1 &= \int_{x_0-\delta}^{x_0+\delta} g(x)\mathrm{e}^{-(1/2\varepsilon)I''(x_0)(x-x_0)^2}dx + o(\sqrt{\varepsilon}) \\
&= g(x_0)\sqrt{\varepsilon}\sqrt{\frac{2\pi}{I''(x_0)}} + o(\sqrt{\varepsilon}), \\
I_2 &= O(\mathrm{e}^{-a/\varepsilon}) \quad (a > 0)
\end{aligned}
$$

となるので、

$$
Z_\varepsilon \sim \mathrm{e}^{-(1/\varepsilon)I_0}\sqrt{\varepsilon}g(x_0)\sqrt{\frac{2\pi}{I''(x_0)}}
$$

を得る。

例 1.1. (1 次元 Gauss 分布)
X を 1 次元標準正規分布に従う確率変数とし、その密度関数を $\gamma(x)$ とする。$\sqrt{\varepsilon}X$ の分布を γ_ε とすると

$$
\gamma_\varepsilon(dx) = \frac{1}{\sqrt{2\pi\varepsilon}}\mathrm{e}^{-x^2/(2\varepsilon)}dx
$$

ゆえ、$I(x) = (1/2)x^2$ として

$$
\lim_{\varepsilon\to 0}\varepsilon\log\gamma_\varepsilon([a,b]) = -\inf_{[a,b]} I
$$

が成立する。このとき、$\{\gamma_\varepsilon, \varepsilon > 0\}$ は *rate function* を I として、大偏差原理を満たすという。

1.2 1 次元独立同分布の場合

次に、1 次元独立同分布に従う確率変数列の大数の法則に対する大偏差原理を考える。μ を $\mathbb{R}$ 上の確率分布で次の Cramèr の条件を満たすものとする。

$$
\int_{\mathbb{R}} \mathrm{e}^{\lambda x}\mu(dx) < \infty \quad (\forall\lambda \in \mathbb{R}) \tag{1.1}
$$

$\int_{\mathbb{R}} x\mu(dx) = p$ とし、μ の $(1/N)\sum_{k=1}^{N} x_k$ による像測度を μ_N とおく。すると、$\mu_N \to \delta_{\{p\}}$ (法則収束)。ここで μ のキュムラント母関数を

$$
\Lambda_\mu(\lambda) = \log\int_{\mathbb{R}} \mathrm{e}^{\lambda x}\mu(dx) \tag{1.2}
$$

とおく。今 Cramèr の条件 (1.1) を仮定しているので、Λ_μ は滑らかになって、$\mu^\lambda(dx) = \left(\int_{\mathbb{R}} \mathrm{e}^{\lambda y}\mu(dy)\right)^{-1} \mathrm{e}^{\lambda x}\mu(dx)$ とおくと、

$$
\begin{aligned}
\Lambda'_\mu(\lambda) &= \int_{\mathbb{R}} x\mu^\lambda(dx), \\
\Lambda''_\mu(\lambda) &= \int_{\mathbb{R}} (x - \int_{\mathbb{R}} y\mu^\lambda(dy))^2 \mu^\lambda(dx), \\
\Lambda'''_\mu(\lambda) &= \int_{\mathbb{R}} (x - \int_{\mathbb{R}} y\mu^\lambda(dy))^3 \mu^\lambda(dx),
\end{aligned}
\tag{1.3}
$$

が成立し、特に Λ_μ は凸関数になっている。ここで、Λ_μ の **Legendre 変換**を

$$
\Lambda^*_\mu(x) = \sup\{\lambda x - \Lambda_\mu(\lambda);\ \lambda \in \mathbb{R}\}
\tag{1.4}
$$

と定める。すると、Λ^*_μ は凸関数で $\Lambda^*_\mu(p) = 0,\ \Lambda^*_\mu(x) \geq 0$ を満たす。

命題 1.2. (上からの評価)
任意の $[a, b]$ に対し

$$
\limsup_{n\to\infty} \frac{1}{N} \log \mu_N([a, b]) \leq -\inf_{[a,b]} \Lambda^*_\mu.
$$

証明　$p < a$ の場合に示す。$\Lambda_N(\lambda) = \Lambda_{\mu_N}(\lambda)$ とおく。すると Chebyshev の不等式により、

$$
\mu_N([a, \infty)) \leq \mathrm{e}^{-(\lambda a - \Lambda_N(\lambda))} = \mathrm{e}^{-N((\lambda/N)a - \Lambda_\mu(\lambda/N))}
$$

が任意の $\lambda > 0$ に対して成り立つ。$p < a$ ならば、$\sup_{\lambda>0}(\lambda a - \Lambda_\mu(a)) = \sup_{\lambda\in\mathbb{R}}(\lambda a - \Lambda_\mu(a))$ ゆえ

$$
\liminf_{N\to\infty} \frac{1}{N} \log \mu_N([a, \infty)) \leq -\Lambda^*_\mu(a)
$$

□

命題 1.3. (下からの評価)
任意の (a, b) に対し

$$
\liminf_{n\to\infty} \frac{1}{N} \log \mu_N((a, b)) \geq -\inf_{(a,b)} \Lambda^*_\mu.
$$

証明　任意の $x \in \mathbb{R}$ に対して

$$\liminf_{n\to\infty} \frac{1}{N} \log \mu_N((x-\delta, x+\delta)) \geq -\Lambda_\mu^*(x)$$

を示す。$\lambda_x \in \mathbb{R}$ を $\Lambda_\mu^*(x) = \lambda_x x - \Lambda_\mu(\lambda_x)$ なるものとする。ここで、新しい確率測度 $\mu^{\lambda_x}(dy)$ を $\mu^{\lambda_x}(dy) = \dfrac{\mathrm{e}^{\lambda_x x}}{\mathrm{e}^{\Lambda_\mu(\lambda_x)}}\mu(dy)$ で与えると、$\int_{\mathbb{R}} y d\mu^{\lambda_x}(dy) = x$ となる。

$$\begin{aligned}
\mu_N((x-\delta, x+\delta)) &\geq \mathrm{e}^{-N\lambda_x(x+\delta)} \int \cdots \int_{\{\sum_{k=1}^N y_k \in (N(x-\delta), N(x+\delta))\}} \mathrm{e}^{\lambda_x \sum_{k=1}^N y_k} \mu^{\otimes N}(dy_1 \ldots dy_N) \\
&= \mathrm{e}^{-N(\lambda_x(x+\delta)-\Lambda_\mu(\lambda_x))} (\mu^{\lambda_x})^{\otimes N} \{|\frac{1}{N}\sum_{k=1}^N y_k - x| < \delta\}
\end{aligned}$$

ゆえ

$$\begin{aligned}
&\liminf_{N\to\infty} \frac{1}{N} \log \mu_N((x-\delta, x+\delta)) \\
\geq\ & -\lambda_x(x+\delta) + \Lambda_\mu(\lambda_x) + \liminf_{n\to\infty} \frac{1}{N} \log (\mu^{\lambda_x})^{\otimes N} \{|\frac{1}{N}\sum_{k=1}^N y_k - x| < \delta\} \\
\geq\ & -\Lambda_\mu^*(x) - \delta\lambda_x
\end{aligned}$$

□

注意 1.4. ここで述べた 1 次元の独立同分布の確率変数の和の大偏差原理に関して、*Cramèr* の条件 (1.1) が部分的にしか成立しない場合には *Petrov* [28], *Nagaev* [27] 等の結果がある。さらに、ここでの設定とは異なるが、*tail* が多項式オーダーで減少する場合 (即ち、非常に大きな一つの値が和の大部分をしめるような場合) に対しても、*tail* との関連で大偏差に関する結果が知られている。([8] 参照)

2　大偏差原理の一般論

ここでは、大偏差原理の一般論を主に Deuschel-Stroock [3] と Dembo-Zeitouni [2] から抜粋する。詳しくはそれらの本をみてください。

2.1 大偏差原理の定義

ここでは後で Sanov の定理を扱うので、Deuschel-Stroock [3] に従った設定にする。

(i) X : 可測ノルム $||\cdot||$ を持つ局所凸線形位相空間

(ii) $E \subset X$ は閉凸部分集合で、X の位相と同値な距離 ρ でポーランド空間となり次を満たす。

- 任意の $\alpha \in [0,1]$ に対し $\rho(\alpha p_1 + (1-\alpha)p_2, \alpha q_1 + (1-\alpha)q_2,) \leq \rho(p_1,p_2) \wedge \rho(q_1,q_2)$
- $\rho(p,q) \leq ||p-q||$

例 2.1. Σ をポーランド空間とし、$X = \mathcal{M}(\Sigma)$ (Σ 上の符号付き有界測度に弱位相を入れた空間)、$E = \mathcal{M}_1^+(\Sigma)$ (Σ 上の確率測度全体に弱位相を入れた空間)、$\rho = Prohorov$ の距離、$||\cdot|| = ||\cdot||_{var}$.

この例では次が成立。

命題 2.2. $X^* = C_b(\Sigma)$

証明 $\lambda \in X^*$ に対し、$\varphi_\lambda(\sigma) = \lambda(\delta_{\{\sigma\}})$ ととると $\varphi_\lambda \in C_b(\Sigma)$ となる。 □

定義 2.3. $I:\ E \to [0,\infty]$ が

(i) **rate function** とは I が下半連続なこと。

(ii) **good rate function** とは、*rate function* で任意の $L > 0$ に対し、$\{p \in E:\ I(p) \leq L\}$ がコンパクトになること。

$\{Q_\varepsilon:\ \varepsilon > 0\} \subset \mathcal{M}_1^+(E)$ とする。

定義 2.4. (i) $\{Q_\varepsilon:\ \varepsilon > 0\}$ が *rate function* I で**大偏差原理 (large deviation principle)** を満たすとは、
(1) $\forall G\ open\ \subset E:\ \liminf_{\varepsilon\to 0} \varepsilon \log Q_\varepsilon(G) \geq -\inf_G I$
(2) $\forall F\ closed\ \subset E:\ \limsup_{\varepsilon\to 0} \varepsilon \log Q_\varepsilon(F) \leq -\inf_F I$
が成り立つこと。

(ii) $\{Q_\varepsilon:\ \varepsilon > 0\}$ が *rate function* I で**弱い意味の大偏差原理 (weak large deviation principle)** を満たすとは、上の (1) と
$(2)'$ $\forall K\ compact\ \subset E:\ \limsup_{\varepsilon\to 0} \varepsilon \log Q_\varepsilon(K) \leq -\inf_K I$
が成り立つこと。

まず次が成立する。

命題 2.5. $\{Q_\varepsilon:\ \varepsilon > 0\}$ が *l.d.p.* を満たすとすると、*rate function* は一意的。

例 2.6. $B = C([0,1];\mathbb{R})$, $B^* = \mathcal{M}[0,1]$ とし、$(W(t), t \in [0,1])$ をブラウン運動、γ をその B 上の分布とする。$\varepsilon > 0$ に対し、$\sqrt{\varepsilon}W(\cdot)$ の分布を γ_ε とおく。*Cameron-Martin* 空間を $H = \{h : [0,1] \to \mathbb{R} : h(0) = 0, h$ は絶対連続, $\int_0^1 |\dot{h}|^2 dt < \infty\}$ 、その内積を $(h_1, h_2)_H = \displaystyle\int_0^1 \dot{h}_1(t)\dot{h}_2(t)dt$ とする。さらに、$\iota^* : B^* \to H$ を
$\iota^*(\nu)(t) = \displaystyle\int_0^t (f_\nu(1) - f_\nu(s))ds, \quad f_\nu(t) = \int_{(0,t]} \nu(ds)$ とおくと次が成立する。

$$\int_B \mathrm{e}^{\sqrt{-1}\langle w,\nu\rangle}\gamma(dw) = \mathrm{e}^{-(1/2)||\iota^*(\nu)||_H^2} \qquad (\forall \nu \in B^*)$$

これより、シフトを $(\theta w)(t) = w(t) + h(t)$ とおくと、$\gamma \circ \theta^{-1} \ll \gamma \Leftrightarrow h \in H$ で $d(\gamma \circ \theta^{-1})/d\gamma = \exp(\widetilde{(h,w)}_H - (1/2)||h||_H^2)$ (**Cameron-Martin**) となり、直感的に $\gamma(dw) = \exp(-(1/2)\int_0^1 |\dot{h}(t)|^2 dt)$"$dw$" と思えるので、例 1.1 と同様に考えると、次の結果が予想される。

(Schilder の定理)

$$(2.1) \qquad I(h) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^1 |\dot{h}(t)|^2 dt, & h \in H \\ \infty, & \text{その他} \end{cases}$$

とおくと、$\{\gamma_\varepsilon,\ \varepsilon > 0\}$ は I を *rate function* として *l.d.p.* を満たす。

定義 2.7. $\{Q_\varepsilon :\ \varepsilon > 0\} \subset \mathcal{M}_1^+(E)$ が **exponentially tight** とは、任意の $L > 0$ に対してコンパクト集合 $K_L \subset E$ があって

$$\limsup_{\varepsilon \to 0} \varepsilon \log Q_\varepsilon(K_L^c) \le -L$$

とできること。

命題 2.8. $\{Q_\varepsilon :\ \varepsilon > 0\}$ は *rate function* I で *weak l.d.p.* を満たしているとする。また、$\{Q_\varepsilon :\ \varepsilon > 0\}$ は *exponentially tight* とする。このとき、

(i) I は *good rate function* であり、

(ii) $\{Q_\varepsilon :\ \varepsilon > 0\}$ は *good rate function* I で *l.d.p.* を満たす。

証明 (i)

$$\inf_{K_{L+1}^c} I \ge -\limsup_{\varepsilon \to 0} \varepsilon \log Q_\varepsilon(K_{L+1}^c) \ge L + 1$$

従って、$\{q : I(q) \leq L\} \subset K_{L+1}$ かつ閉ゆえ $\{q : I(q) \leq L\}$ はコンパクト。
(ii) F を E の閉集合とする。$Q_\varepsilon(F) \leq Q_\varepsilon(F \cap K_L) + Q_\varepsilon(K_L^c)$ ゆえ

$$\limsup_{\varepsilon \to 0} \varepsilon \log Q_\varepsilon(F) \leq -(\inf_{F \cap K_L} I \wedge L) \leq -(\inf_F I \wedge L)$$

$\square$

2.2 大偏差原理の遺伝

定理 2.9. (contraction principle)
(E', ρ') を距離空間とし、$f : \ E \to E'$ を連続写像とする。

(i) I を E 上の *good rate function* とし、

$$I'(p') = \inf\{I(p) : \ f(p) = p'\}$$

とおくと、I' は E' 上の *good rate function* になる。

(ii) $\{Q_\varepsilon, \varepsilon > 0\} \subset \mathcal{M}_1^+(E)$ が *good rate function* I で *l.d.p.* を満たしているならば、$\{Q_\varepsilon \circ f^{-1}, \varepsilon > 0\} \subset E'$ は *good rate function* I' で *l.d.p.* を満たす。

証明 (i) I good ゆえ $I'(p') < \infty$ に対し、$f(p) = p'$, $I'(p') = I(p)$ なる $p \in E$ がある。$p'_n \to p'_0$ とすると $f(p_n) = p'_n$, $I'(p'_n) = I(p_n)$ なる $\{p_n\} \subset E$ がとれ、必要なら部分列をとって $p_n \to p_0$ とできる。$f(p_0) = p'_0$ ゆえ $I(p_0) \geq I'(p'_0)$。I 下半連続ゆえ $\liminf I'(p'_n) \geq I'(p'_0)$ となり、I' も下半連続となる。また、$\{p' : I'(p') \leq L\} = \{f(p) : I(p) \leq L\}$ ゆえ $\{p' : I'(p') \leq L\}$ はコンパクト。
(ii) $A' \subset E'$ に対し、$\inf_{p' \in A'} I'(p') = \inf_{f(p) \in A'} I(p)$ となり、また G' が開ならば $f^{-1}(G')$ も開、F' が閉ならば $f^{-1}(F')$ も閉ゆえ l.d.p. が成立する。 $\square$

例 2.10. 例 2.6 の応用として次の *Ventcel-Freidlin* の結果を述べる。
$b : \mathbb{R}^d \to \mathbb{R}^d$ を有界、*Lipschitz* 連続関数とし、連続写像 $X : C([0,1] : \mathbb{R}^d) \to C([0,1]; \mathbb{R}^d)$ を

$$X(w,t) = w(t) + \int_0^t b(X(w,s))ds$$

で定め、$P_\varepsilon = \gamma_\varepsilon \circ X^{-1}$ とおく。即ち、P_ε でみると、$W(t)$ をブラウン運動として、

$$X(t) = \sqrt{\varepsilon} W(t) + \int_0^t b(X(s))ds$$

となる。
ここで

$$(2.2) \quad I_b = \begin{cases} \dfrac{1}{2} \displaystyle\int_0^1 |\dot{h}(t) - b(h(t)|^2 dt & h(t) \text{ が絶対連続で、} \dot{h} \in L^2[0,1] \\ \infty & \text{それ以外} \end{cases}$$

と定めると $\{P_\varepsilon, \varepsilon > 0\}$ は I_b を *rate function* として *l.d.p.* を満たす。

定理 2.11. (inverse contraction principle)
(E',ρ') を距離空間とし、$\{Q'_\varepsilon,\varepsilon>0\}\subset\mathcal{M}_1^+(E')$ は *exponetially tight* とする。$g:E'\to E$ を連続な全単射とし、$\{Q'_\varepsilon\circ g^{-1},\varepsilon>0\}\subset\mathcal{M}_1^+(E)$ が *rate function* I で *weak l.d.p.* を満たすとする。
このとき $I'(p')=I(g(p'))$ とおくと、$\{Q'_\varepsilon,\varepsilon>0\}$ は I' を *good rate function* として *l.d.p.* を満たす。

証明 I' は下半連続ゆえ rate function. $\{Q'_\varepsilon\}$ が exponentially tight なので、rate function I' で weak l.d.p. を満たすことをいえばよい。
(上からの評価)
$K'\subset E'$ コンパクトとすると $g(K')$ もコンパクト。

$$\begin{aligned}\limsup_{\varepsilon\to0}\varepsilon\log Q'_\varepsilon(K')&=\limsup_{\varepsilon\to0}\varepsilon\log Q'_\varepsilon\circ g^{-1}(g(K'))\\&\le-\inf_{g(K')}I=-\inf_{K'}I'\end{aligned}$$

(下からの評価)
$p'\in E'$ を $I'(p')=a<\infty$ なるものとする。$\{Q'_\varepsilon\}$ は exponentially tight なので、$\exists K'_{a+1}\subset E'$ コンパクトで、

$$-\inf_{g({K'_{a+1}}^c)}I\le\liminf_{\varepsilon\to0}\varepsilon\log Q'_\varepsilon({K'_{a+1}}^c)<-(a+1)$$

成立。よって $p'\in K'_{a+1}$. G' を p' の近傍とする。$g:K'_{a+1}\to g(K'_{a+1})$ homeo なので、$G\subset g(G'\cap K'_{a+1})\cup g(K'_{a+1})^c=g(G'\cup {K'_{a+1}}^c)$ となる $g(p')$ の E での近傍 G がとれる。$\{Q'_\varepsilon\circ g^{-1}\}$ の下からの評価より、

$$\liminf_{\varepsilon\to0}\varepsilon\log Q'_\varepsilon\circ g^{-1}(G)\ge-\inf_G I=\inf_{g^{-1}(G)}I'\ge-I'(p').$$

また、

$$\begin{aligned}\liminf_{\varepsilon\to0}\varepsilon\log Q'_\varepsilon\circ g^{-1}(G)&\le\liminf_{\varepsilon\to0}\varepsilon\log(Q'_\varepsilon(G')+Q'_\varepsilon({K'_{a+1}}^c))\\&\le\left(\liminf_{\varepsilon\to0}\varepsilon\log Q'_\varepsilon(G')\right)\vee\left(\limsup_{\varepsilon\to0}\varepsilon\log Q'_\varepsilon({K'_{a+1}}^c)\right)\end{aligned}$$

よって、$I'(p')\le\liminf_{\varepsilon\to0}\varepsilon\log Q'_\varepsilon(G')$. □

系 2.12. *E 上に 2 つの距離 ρ_1,ρ_2 があって ρ_1 の方が ρ_2 より強いとする。$\{Q_\varepsilon,\varepsilon>0\}$ が $\mathcal{M}_1^+(E,\rho_1)$ の exponentially tight な族とし、rate function I で (E,ρ_2) 上 weak l.d.p. を満たすとする。このとき、$\{Q_\varepsilon,\varepsilon>0\}$ は (E,ρ_1) 上 good rate function I で、l.d.p. を満たす。*

2.3 Varadhan の定理

Laplace の方法の無限次元版として次の Varadhan の定理がある。

定理 2.13. (Varadhan)
$\{Q_\varepsilon, \varepsilon > 0\} \subset \mathcal{M}_1^+(E)$ が *good rate function* I で *l.d.p.* を満たしているとし、$\Phi : E \to \mathbb{R}$ は有界連続関数とする。このとき、

$$\lim_{\varepsilon \to 0} \varepsilon \log \int_E \exp\left(\frac{1}{\varepsilon}\Phi(q)\right) Q_\varepsilon(dq) = \sup_E(\Phi - I)$$

証明 ($\leq$) 任意の $p \in E$, $\delta > 0$ に対し、p の近傍 G で $\inf_G \Phi \geq \Phi(p) - \delta$ なるものがとれる。

$$\begin{aligned}\liminf_{\varepsilon \to 0} \varepsilon \log \int_E \exp\left(\frac{1}{\varepsilon}\Phi\right) dQ_\varepsilon &\geq \liminf_{\varepsilon \to 0} \varepsilon \log \int_G \exp\left(\frac{1}{\varepsilon}\Phi\right) dQ_\varepsilon \\ &\geq \Phi(p) - \delta + \liminf_{\varepsilon \to 0} \varepsilon \log Q_\varepsilon(G) \geq \Phi(p) - \delta - I(p)\end{aligned}$$

($\geq$) $L > 0$ に対し、$K_L = \{q \in E : I(q) \leq L\}$ はコンパクト。$p \in K_L$, $\delta > 0$ に対し、p の近傍 G_p で $\inf_{\overline{G_p}} I \geq I(p) - \delta$, $\sup_{\overline{G_p}} \Phi \leq \Phi(p) + \delta$ を満たすものがとれる。K_L はコンパクトなので、有限個の $p_1, p_2, \ldots . p_n \in K_L$ で $K_L \subset \cup_{i=1}^n G_{p_i}$ とできる。よって

$$\begin{aligned}\int_E \mathrm{e}^{(1/\varepsilon)\Phi} dQ_\varepsilon &\leq \sum_{i=1}^n \int_{G_{p_i}} \mathrm{e}^{(1/\varepsilon)\Phi} dQ_\varepsilon + \int_{(\cup_{i=1}^n G_{p_i})^c} \mathrm{e}^{(1/\varepsilon)\Phi} dQ_\varepsilon \\ &\leq \sum_{i=1}^n \mathrm{e}^{(1/\varepsilon)(\Phi(p_i)+\delta)} Q_\varepsilon(\overline{G_{p_i}}) + \mathrm{e}^{(1/\varepsilon)||\Phi||_\infty} Q_\varepsilon\left((\cup_{i=1}^n G_{p_i})^c\right)\end{aligned}$$

ゆえ

$$\begin{aligned}&\limsup_{\varepsilon \to 0} \varepsilon \log \int_E \mathrm{e}^{(1/\varepsilon)\Phi} dQ_\varepsilon \\ &\quad \leq \max\left\{\Phi(p_i) - I(p_i) + 2\delta,\ ||\Phi||_\infty - L : 1 \leq i \leq n\right\} \\ &\quad \leq \sup_E(\Phi - I) \quad (\text{as } L \to \infty).\end{aligned}$$

□

注意 2.14. *Varadhan* の定理は、より詳しくは以下の形で成立。

(i) (下からの評価)
$\{Q_\varepsilon, \varepsilon > 0\}$ が *rate function* I で *l.d.p.* の下からの評価を満たすとし、Φ は下半連続とする。すると次が成立。

$$\liminf_{\varepsilon \to 0} \varepsilon \log \int_E \mathrm{e}^{(1/\varepsilon)\Phi} dQ_\varepsilon \geq \sup_E(\Phi - I) \tag{2.3}$$

(ii) (上からの評価)
$\{Q_\varepsilon, \varepsilon > 0\}$ が *good rate function* I で *l.d.p.* の上からの評価を満たすとし、Φ は上半連続で

$$\lim_{L\to\infty} \limsup_{\varepsilon\to 0} \varepsilon \log \int_{\{\Phi\ge L\}} \mathrm{e}^{(1/\varepsilon)\Phi} dQ_\varepsilon = -\infty \tag{2.4}$$

を満たすとする。このとき

$$\limsup_{\varepsilon\to 0} \varepsilon \log \int_E \mathrm{e}^{(1/\varepsilon)\Phi} dQ_\varepsilon \le \sup_E (\Phi - I) \tag{2.5}$$

が成り立つ。
また、(2.4) が成り立つためには次が成立すれば十分。ある $\alpha > 1$ があって

$$\sup_{0<\varepsilon\le 1} \left(\int_E \mathrm{e}^{(\alpha/\varepsilon)\Phi} dQ_\varepsilon \right)^\varepsilon < \infty. \tag{2.6}$$

2.4 大偏差原理の成立

ここでは 大偏差原理 の下からの評価と、弱い意味の大偏差原理成立のための条件を与える。

定理 2.15. $B_r(p)$ を $p \in E$ を中心、半径 $r > 0$ の E の球とし、その全体を $\mathcal{B} = \{B_r(p) : p \in E, r > 0\}$ とする。さらに、$\mathcal{L}(B) = -\liminf_{\varepsilon\to 0} \varepsilon \log Q_\varepsilon(B)$, $I(p) = \sup_{B\in\mathcal{B}, B\ni p} \mathcal{L}(B)$ とおく。このとき次が成立する。

(i) I は *rate function* で、$\{Q_\varepsilon :\ \varepsilon > 0\}$ は *rate funcion* を I として *l.d.p.* の下からの評価を満たす。

(ii) さらに

$$I(p) = \sup_{B\in\mathcal{B}, B\ni p} \left(-\limsup_{\varepsilon\to 0} \varepsilon \log Q_\varepsilon(B) \right) \tag{2.7}$$

が成り立つとすると $\{Q_\varepsilon :\ \varepsilon > 0\}$ は *rate function* I で *weak l.d.p.* を満たす。

証明 (i) $I(p) > a$ とする。すると $\mathcal{L}(B) > a, p \in B$ なる $B \in \mathcal{B}$ があり、$q \in B$ ならば $I(q) \ge \mathcal{L}(B) > a$ より、$B \subset \{p; I(p) > a\}$. よって $\{p; I(p) > a\}$ は開集合となり、I は下半連続。ここで $G \subset E$ を開集合とし、$p \in G$ とする。すると $B \subset G, B \ni p$ なる $B \in \mathcal{B}$ があるので、

$$\liminf_{\varepsilon\to 0} \varepsilon \log Q_\varepsilon(G) \ge \liminf_{\varepsilon\to 0} \varepsilon \log Q_\varepsilon(B) \ge -\mathcal{L}(B) \ge -I(p)$$

(ii) $K \subset E$ をコンパクト、$\delta > 0$ とし、$I^\delta = (I(p) - \delta) \wedge \left(\frac{1}{\delta}\right)$ とおく。(2.7) より、任意の $q \in K$ に対し $q \in B_q, -\limsup_{\varepsilon \to 0} \varepsilon \log Q_\varepsilon(B_q) \geq I^\delta(q)$ なる $B_q \in \mathcal{B}$ が存在する。$K \subset \cup_{i=1}^n B_{q_i}$ となる $q_1, q_2, \ldots, q_n \in K$ がとれるので、

$$\begin{aligned}
\limsup_{\varepsilon \to 0} \varepsilon \log Q_\varepsilon(K) &\leq \max_{1 \leq i \leq n} \left\{ \limsup_{\varepsilon \to 0} \varepsilon \log Q_\varepsilon(B_{q_i}) \right\} \\
&\leq - \min_{1 \leq i \leq n} I^\delta(q_i) \leq - \inf_K I^\delta
\end{aligned}$$

$\delta \to 0$ として上からの評価が示された。 □

系 2.16. 任意の $B \in \mathcal{B}$ に対して $\lim_{\varepsilon \to 0} \varepsilon \log Q_\varepsilon(B)$ が存在するならば、定理 2.15 の (ii) が成立する。

2.5 Legendre 変換

ここでは Legendre 変換を用いて、rate function が凸関数のときに、その形と大偏差原理の上からの評価を与える。
一般に Legendre 変換に対し次の双対性が成立する。

定理 2.17. $f : X \to (-\infty, +\infty]$ は下半連続で凸とする。$g : X^* \to (-\infty, +\infty]$ を次で定義する。

$$g(\lambda) = \sup\{\langle \lambda, x \rangle - f(x) : x \in X\}$$

すると次が成立する。

$$f(x) = \sup\{\langle \lambda, x \rangle - g(\lambda) : \lambda \in X^*\}$$

大偏差原理が成り立っていると次が成立する。

定理 2.18. $\{Q_\varepsilon :\ \varepsilon > 0\} \subset \mathcal{M}_1^+(E)$ が凸の *good rate function* I で *l.d.p.* を満たしており、

$$\sup_{\varepsilon \in (0,1]} \left(\int_E \mathrm{e}^{(1/\varepsilon)\langle \lambda, x \rangle} Q_\varepsilon(dx) \right)^\varepsilon < \infty \tag{2.8}$$

を満たすとする。このとき次が成立する。

(i) $\lambda \in X^*$ に対し、次の極限が存在する。

$$\Lambda(\lambda) = \lim_{\varepsilon \to 0} \varepsilon \log \int_E \mathrm{e}^{(1/\varepsilon)\langle \lambda, x \rangle} Q_\varepsilon(dx)$$

(ii) また $\lambda \in X^*$ に対し、

$$\Lambda(\lambda) = \sup\{\langle \lambda, q \rangle - I(q) : q \in E\}$$

(iii) さらに

$$I(q) = \Lambda^*(q) = \sup\{\langle\lambda, q\rangle - \Lambda(\lambda) : \lambda \in X^*\}$$

証明　(2.8) より Varadhan の定理 (注意 2.14 (ii)) が使えて

$$\lim_{\varepsilon\to 0} \varepsilon \log \int_E \mathrm{e}^{(1/\varepsilon)\langle\lambda, x\rangle} Q_\varepsilon(dx) = \sup_E(\langle\lambda, q\rangle - I(q))$$

これと定理 2.17 より結果を得る。 □

逆に、Legendre 変換より、上からの評価が次のように得られる。

定理 2.19. $\{Q_\varepsilon :\ \varepsilon > 0\} \subset \mathcal{M}_1^+(E)$ が、任意の $\lambda \in X^*$ に対し極限

$$\Lambda(\lambda) = \lim_{\varepsilon\to 0} \varepsilon \log \int_E \mathrm{e}^{(1/\varepsilon)\langle\lambda, x\rangle} Q_\varepsilon(dx)$$

を持つとする。すると、
(i) Λ は凸となり、

$$\Lambda^*(q) = \sup\{\langle\lambda, q\rangle - \Lambda(\lambda) : \lambda \in X^*\} \tag{2.9}$$

とおくと
(ii) Λ^* は非負値、下半連続、凸で
(iii) 任意のコンパクト集合 K に対し

$$\limsup_{\varepsilon\to 0} \varepsilon \log Q_\varepsilon(K) \le -\inf_K \Lambda^*$$

が成立する。

証明　(i) は Hölder の不等式より、また (ii) は $x \mapsto \langle\lambda, x\rangle - \Lambda(\lambda)$ の上限ゆえ下半連続、凸となる。
(iii) $p \in E, \delta > 0$ に対し、$\langle\lambda, p\rangle - \Lambda(\lambda) \ge (1/\delta) \wedge (\Lambda^*(p) - \delta/2)$ となる Λ をとる。この λ に対し $\langle\lambda, p - q\rangle \le \delta/2$ $(\forall q \in \overline{B_r(p)})$ となる $r > 0$ をとる。すると Chebyshev の不等式より

$$\begin{aligned}\varepsilon \log Q_\varepsilon(\bar{B}_r(p)) &\le -\langle\lambda, p\rangle + \frac{\delta}{2} + \varepsilon \log \int_E \mathrm{e}^{(1/\varepsilon)\langle\lambda, q\rangle} Q_\varepsilon(dq) \\ &\le -\langle\lambda, p\rangle + \Lambda(\lambda) + \delta \le (\Lambda^*(p) + \frac{3}{2}\delta) \wedge (\frac{1}{\delta} + \delta)\end{aligned}$$

K をコンパクト集合とすると、 $K \subset \cup_{i=1}^n B_{r_i}(p_i)$ となる $p_1, \ldots, p_n \in K$, $r_1, \ldots, r_n > 0$ がある。よって

$$\begin{aligned}\limsup_{\varepsilon\to 0} \varepsilon \log Q_\varepsilon(K) &\le \limsup_{\varepsilon\to 0} \varepsilon \log \sum_{i=1}^n Q_\varepsilon(\overline{B_{r_i}(p_i)}) \\ &\le \max_{1\le i\le n}\{-(\Lambda^*(p_i) + \frac{3}{2}\delta), -\frac{1}{\delta}\} \le -\inf_K \Lambda^*\end{aligned}$$

□

系 2.20. $\{Q_\varepsilon:\ \varepsilon>0\}\subset\mathcal{M}_1^+(E)$ が *exponentially tight* で、任意の $\lambda\in X^*$ に対し

$$\Lambda(\lambda)=\lim_{\varepsilon\to0}\varepsilon\log\int_E \mathrm{e}^{(1/\varepsilon)\langle\lambda,x\rangle}Q_\varepsilon(dx)$$

が存在するならば、(2.9) で与えた Λ^* は $\{Q_\varepsilon:\ \varepsilon>0\}$ の *l.d.p.* の上からの評価を与える。

3 Sanov の定理

ここでは大偏差原理の例として Sanov の定理について Deuschel-Stroock [3]、Dembo-Zeitouni [2] に従って述べる。
例 2.1 の設定で扱う。すなわち、Σ をポーランド空間、$E=\mathcal{M}_1^+(\Sigma), X=\mathcal{M}(\Sigma), ||\alpha||=||\alpha||_{var},\ (\alpha\in X)$ とする。$\{X_1,X_2,\dots\}$ を Σ-値の独立同分布の確率変数列とし、その共通の分布を $\mu\in\mathcal{M}_1^+(\Sigma)$ とする。また $L_N:\Sigma^{\otimes N}\to\mathcal{M}_1^+(\Sigma)$ を

$$L_N(\underline{\mathrm{x}}_N)=\frac{1}{N}\sum_{k=1}^N\delta_{\{x_i\}} \tag{3.1}$$

で定め、$Q_N=\mu^{\otimes N}\circ L_N^{-1}\in\mathcal{M}_1^+(\mathcal{M}_1^+(\Sigma))$ とおく。すると、大数の法則より

$$Q_N\to\delta_{\{\mu\}}\quad(\text{法則収束})$$

が成り立つ。
weak upper bound
ここでも、$V\in C_b(\Sigma)$ に対してキュムラント母関数を考える。
$\int_{\Sigma^{\otimes N}}\exp(N\langle L_N,V\rangle)d\mu^{\otimes N}=(\int_\Sigma \mathrm{e}^V d\mu)^N=\mathrm{e}^{N\Lambda_\mu(V)}$ ゆえ

$$\frac{1}{N}\log\int_{\Sigma^{\otimes N}}\exp(N\langle L_N,V\rangle)d\mu^{\otimes N}=\Lambda_\mu(V),$$
$$\Lambda_\mu^*(\nu)=\sup\{\int_\Sigma Vd\nu-\Lambda_\mu(V):V\in C_b(\Sigma)\}.$$

とおけば、定理 2.19 より rate function Λ_μ^* は $\{Q_N,N\geq1\}$ の l.d.p. の weak upper bound を与える。

例 3.1. 有限点上の分布 $\Sigma=\{\sigma_1,\dots,\sigma_n\},\ \mu(\{\sigma_k\})=\mu_k$ の場合は
$\Lambda(\lambda)=\log\sum_{k=1}^n\mathrm{e}^{\lambda_k}\mu_k,\ \Lambda^*(q)=\sup_{(\lambda_k)}\left(\sum_{k=1}^n\lambda_kq_k-\Lambda(\lambda)\right)$ ゆえ、λ_k で微分して $\Lambda^*(q)=\sum_{k=1}^n f_k\log f_k\mu_k$ *(ただし $q_k=f_k\mu_k$)* を得る。

ここで、相対エントロピー H を次で与える。

$$H(\nu|\mu) = \begin{cases} \displaystyle\int_\Sigma f(\sigma)\log f(\sigma)d\mu(\sigma) & \nu \ll \mu,\ f = \dfrac{d\nu}{d\mu}, \\ & \qquad \log f \in L^1(\Sigma, d\nu) \\ \infty & \text{それ以外} \end{cases} \tag{3.2}$$

命題 3.2. $\Lambda^*(\nu) = H(\nu|\mu)$

証明　初めに $x\log x$ が凸関数であること、及び $\log x$ が単調増加の凹関数であることより、$\nu_\theta = \theta\mu + (1-\theta)\nu$ とおくと、$\lim_{\theta\to 0} H(\nu_\theta|\mu) = H(\nu|\mu)$ が成立することに注意する。
まず ν が μ に関して絶対連続なら $\Lambda^*_\mu(\nu) \leq H(\nu|\mu)$ を示す。上にのべたことにより、$f \geq \theta\ (>0)$ としてよい。$V \in C_b(\Sigma)$ に対して Jensen の不等式より

$$\mathrm{e}^{\int_\Sigma V d\nu - H(\nu|\mu)} \leq \int_\Sigma \mathrm{e}^{V - \log f} d\nu = \int_\Sigma \mathrm{e}^V d\mu$$

より

$$\int_\Sigma V d\nu - \log\int_\Sigma \mathrm{e}^V d\mu \leq H(\nu|\mu) \quad (\forall V \in C_b(\Sigma)). \tag{3.3}$$

次に、$\Lambda^*_\mu(\nu) < \infty$ ならば ν は μ に絶対連続になることを示す。$\Lambda^*_\mu(\nu) < \infty$ とする。すると、有界連続関数で近似することで、E 上の任意の有界可測関数 φ に対し

$$\int_E \varphi d\nu - \log\int_E \mathrm{e}^\varphi d\mu \leq \Lambda^*_\mu(\nu) \tag{3.4}$$

が成り立つ。$\mu(A) = 0$ とする。$\varphi(\sigma) = a\mathbf{1}_A(\sigma)\ (a>0)$ とおくと (3.4) より $a\nu(A) \leq \Lambda^*_\mu(\nu)\ (\forall a > 0)$ となり、$\nu \ll \mu$ を得る。
最後に $H(\nu|\mu) \leq \Lambda^*_\mu(\nu)$ を示す。$\Lambda^*_\mu(\nu) < \infty$ としてよい。上に示したように $\nu \ll \mu$ ゆえ、$d\nu/d\mu = f$ とおく。$f \geq \theta\ (\theta > 0)$ としてよい。f が有界ならば、(3.4) に $\varphi = \log f$ を代入して、$H(\nu|\mu) \leq \Lambda^*_\mu(\nu)$ を得る。f が有界でないときには、$f_n = f \wedge n$ とおいて Fatou を用いると結果を得る。 □

exponential tightness
一般につぎの命題が成立する。

命題 3.3. $\{R_\varepsilon, \varepsilon \in (0,1)\} \subset \mathcal{M}_1^+(E)$ とする。ある $M \geq 1, \beta \geq 1$ があって、任意の有界可測関数 $V : \Sigma \to \mathbb{R}$ に対し次が成立しているとする。

$$\sup_{0<\varepsilon<1}\left\{\int_E \exp\left(\frac{1}{\varepsilon}\int_\Sigma V(\sigma)\nu(d\sigma)\right) R_\varepsilon(d\nu)\right\}^\varepsilon \leq M\int_\Sigma \mathrm{e}^{\beta V} d\mu \tag{3.5}$$

このとき、次が成立する。

(i) Σ 上の有界可測関数 $V_m,\ m=1,2,\dots$ が一様有界で、$V_m(\sigma)\to 0, \mu$-$a.e.$ とする。このとき、部分列 $(V_{m(\ell)})$ がとれて、任意の $\varepsilon\in(0,1), L>1$ に対して、

$$R_\varepsilon\left\{\nu\in\mathcal{M}_1^+(\Sigma):\sup_{\ell\geq L}\ell\int_\Sigma V_{m(\ell)}d\nu\geq 1\right\}\leq \mathrm{e}^{(1/\varepsilon)L}$$

(ii) 特に、任意の L に対し、コンパクト集合 $C_L\subset E$ がとれて $R_\varepsilon(C_L)\leq \mathrm{e}^{-L/\varepsilon}$.

証明　(i) 任意の可測な $V:\Sigma\to\mathbb{R}$ と $\delta>0$ に対して、Chebyshev の不等式と (3.5) より

$$\left(R_\varepsilon\left\{\nu\in\mathcal{M}_1^+(\Sigma):\int_\Sigma Vd\nu\geq\delta\right\}\right)^\varepsilon\leq \mathrm{e}^{-T\delta}\left(\int_E \mathrm{e}^{(T/\varepsilon)\int_\Sigma Vd\nu}R_\varepsilon(d\nu)\right)^\varepsilon$$
$$\leq \mathrm{e}^{-T\delta}M\int_\Sigma \mathrm{e}^{\beta TV}d\mu$$

なので、$\ell\in\mathbb{Z}^+$ に対して $\delta=1/\ell, T=\ell(\ell+1+\log(2M))$ ととり、$m(\ell)$ を $\int_\Sigma \mathrm{e}^{\beta TV_{m(\ell)}}d\mu\leq 2$ ととると

$$R_\varepsilon\left\{\nu\in E:\ell\int_\Sigma V_{m(\ell)}d\nu\geq 1\right\}$$
$$\leq\left(M\mathrm{e}^{-(\ell^2+\ell+\ell\log(2M))/\ell}\int_\Sigma \mathrm{e}^{\beta TV_{m(\ell)}}d\mu\right)^{(1/\varepsilon)}\leq \mathrm{e}^{-(\ell+1)/\varepsilon}$$

(ii) $K_m, m=1,2,\dots$ を Σ のコンパクト集合で、$\mu(K_m^c)\to 0$ なるものとする。$V_m=\mathbf{1}_{K_m^c}$ ととると、(i) より部分列 $K_{m(\ell)},\ell=1,2,\dots$ がとれて

$$R_\varepsilon\left\{\nu\in E:\nu(K_{m(\ell)}^c)\geq 1\right\}\leq \mathrm{e}^{-(\ell+1)/\varepsilon}$$

となるので、$C_L=\cap_{\ell\geq L}\{\nu\in E:\nu(K_{m(\ell)})\geq 1-1/\ell\})$ とおくと

$$R_\varepsilon(C_L^c)=R_\varepsilon\left(\cup_{\ell\geq L}\{\nu\in E:\nu(K_{m(\ell)}^c)>1/\ell\}\right)$$
$$\leq\sum_{\ell\geq L}\mathrm{e}^{-(\ell+1)/\varepsilon}\leq \mathrm{e}^{-L/\varepsilon}.$$

また、C_L のコンパクト性は、$\{\nu\in E:\nu(K_{m(\ell)})\geq 1-1/\ell\}$ は $\mathcal{M}_1^+(\Sigma)$ の tight family で閉ゆえコンパクト。よって C_L も $\mathcal{M}_1^+(\Sigma)$ のコンパクト集合となる。 □

系 3.4. $\{Q_N, N\geq 1\}$ は *exponentially tight family.*

証明　$R_{1/N} = Q_N$ とおくと、

$$\int_{\mathcal{M}_1^+(\Sigma)} \mathrm{e}^{N\int_\Sigma V d\nu} \mathbb{Q}_N(d\nu) = \left(\int_\Sigma \mathrm{e}^{V(\sigma)} \mu(d\sigma) \right)^N$$

ゆえ、(3.5) を満たす。 □

lower bound
ここでは、2 章の一般論を用いて、大偏差原理の成立をみる。
$\mathcal{C}_0 = \{A \subset \mathcal{M}_1^+(\Sigma) : A$ は凸で開, $A \neq \emptyset\}$ とおく。

補題 3.5. (i) 任意の凸集合 $A \subset E$ に対し、$N \mapsto Q_N(A)$ は優乗法的 *i.e.* $Q_{M+N}(A) \geq Q_M(A) Q_N(A)$。
(ii) $A \in \mathcal{C}_0$ ならば、全ての N に対して $Q_N(A) = 0$ であるか、またはある N_0 があって、$Q_N(A) > 0$ $(\forall N \geq N_0)$ が成り立つ。

証明　(i)

$$L_{M+N}^{M+1} = \sum_{\ell=M+1}^{M+N} \delta_{\{x_\ell\}}$$

とおく。A は凸ゆえ $L_M \in A,\ L_{M+N}^{M+1} \in A$ ならば

$$L_{M+N} = \frac{M}{M+N} L_M + \frac{N}{M+N} L_{M+N}^{M+1} \in A$$

なので

$$\begin{aligned} &\mu^{\otimes M}\{L_M \in A\} \mu^{\otimes N}\{L_N \in A\} = \mu^{\otimes (M+N)}\{L_M \in A, L_{M+N}^{M+1} \in A\} \\ &\quad \leq \mu^{\otimes (M+N)}\{L_{M+N} \in A\}. \end{aligned}$$

(ii) $Q_M(A) > 0$ とする。すると $Q_M(K) > 0$ となるコンパクト凸集合 $K \subset A$ が存在する。$0 < 2\delta < \rho(K, A^c)$ と δ をとる。任意の N に対し $N = Ms + r,\ 0 \leq r \leq M - 1$ とする。

$$\begin{aligned} Q_N(A) &= \mu^{\otimes N}\{L_N \in A\} \geq \mu^{\otimes N}\{L_{Ms} \in K, ||L_{Ms} - L_N|| < \delta\} \\ &\geq \mu^{\otimes Ms}\{L_{Ms} \in K, \frac{r}{N}||L_{Ms}|| < \frac{\delta}{2}\} \mu^{\otimes r}\{\frac{r}{N}||L_r|| < \frac{\delta}{2}\} \\ &= \mu^{\otimes Ms}\{L_{Ms} \in K\} \qquad (N > 2M/\delta \text{ ととる}) \\ &\geq (\mu^{\otimes M}\{L_M \in K\})^s > 0 \end{aligned}$$

□

ここでつぎの補題に注意する。

補題 3.6. $f : \mathbb{Z}^+ \to \mathbb{R}$ が劣加法的である (i.e. $f(m+n) \leq f(m) + f(n)$) ならば、
$\lim_{n\to\infty} \dfrac{f(n)}{n}$ が存在して、$\inf_{n\to\infty} \dfrac{f(n)}{n}$ に等しい。

命題 3.7. $I(\nu) = \lim_{r\to 0} \mathcal{L}(B_r(\nu))$ $(= \sup\{\mathcal{L}(A) : \nu \in A \in \mathcal{C}_0\})$ とおくと、I は凸の *rate function* で $\{Q_N, N \geq 1\}$ は I を *rate function* として *weak l.d.p.* を満たす。

証明 補題 3.5、補題 3.6 と系 2.16 より、I は $\{Q_N, N \geq 1\}$ の weak l.d.p. を与える rate function になる。よって、I の凸性を示せばよい。$\nu_1, \nu_2 \in E$ をとり、$\nu = \frac{1}{2}\nu_1 + \frac{1}{2}\nu_2$ とおく。$\nu \in A \in \mathcal{C}_0$ ととし、$\nu_1 \in A_1, \nu_2 \in A_2$, $A \supset \frac{1}{2}A_1 + \frac{1}{2}A_2$ となる $A_1, A_2 \in \mathcal{C}_0$ をとる。すると

$$
\begin{aligned}
\mathcal{L}(A) &= -\lim_{N\to\infty} \frac{1}{2N} \log Q_{2N}(A) \\
&\leq -\lim_{n\to\infty} \frac{1}{2N} \log \mu^{\otimes 2N}\{L_N \in A_1, L_{2N}^{N+1} \in A_2\} \\
&= -\lim_{N\to\infty} \frac{1}{2}\left(\frac{1}{N}\log Q_N(A_1) + \frac{1}{N}\log Q_N(A_2)\right) \\
&\leq \frac{1}{2} I(\nu_1) + \frac{1}{2} I(\nu_2)
\end{aligned}
$$

となり、凸性が示された。 □

定理 3.8. (Sanov)
Σ をポーランド空間、μ をその上の確率測度とし、$H(\cdot|\mu)$ を (3.2) で定まる μ に対する相対エントロピーとする。$Q_N = \mu^{\otimes N} \circ (L_N)^{-1}$ とおくと、$\{Q_N, N \geq 1\}$ は $H(\cdot|\mu)$ を凸の *good rate function* として大偏差原理を満たす。

証明 系 3.4 と命題 3.7 より、命題 3.7 で与えた I は凸の good rate function となる。よって、定理 2.17 と定理 2.18 より $I(\nu) = \Lambda^*(\nu)$ で、命題 3.2 より定理が示される。 □

注意 3.9. *Sanov* の結果を *Markov* 過程に拡張した大偏差原理は *Donsker-Varadhan* によって与えられている。また、これをさらに拡張した *hypermixing* 過程に対する大偏差原理が *Chiyonobu-Kusuoka* [16] で与えられている。

4 Sanov の定理による Laplace の方法の精密化

ここでは Sanov の定理に関する Laplace の方法 (Varadhan の定理) の精密化について述べる。

4.1 設定と結果

この章でも 3 章と同じ設定で考える。ただし、Σ はコンパクト距離空間とする。Sanov の定理 3.8 が成立しているとき、Varadhan の定理によって次の Laplace の方法の第一近似が与えられる。

定理 4.1. (Varadhan)
$F : \mathcal{M}_1^+(\Sigma) \to \mathbb{R}$ を有界連続関数とする。このとき次が成立する。

$$\frac{1}{N}\log\int_{\Sigma^{\otimes N}} \mathrm{e}^{NF(L_N)}d\mu^{\otimes N} = -\inf\{H(\nu|\mu) - F(\nu) : \nu \in \mathcal{M}_1^+(\Sigma)\} \tag{4.1}$$

そこでここではその精密化として次のオーダーである定数項の係数を求めることを目的とする。(4.1) の右辺の値を $-b_F$ とおき、

$$K_F = \{\nu \in \mathcal{M}_1^+(\Sigma) : H(\nu|\mu) - F(\nu) = b_F\} \quad (K_F \neq \emptyset \text{ とする})$$

とおく。$H(\cdot|\mu)$ が good rate function ゆえ、 K_F はコンパクト集合となる。以下 $H(\cdot|\mu)$ を $I(\cdot)$ で表す。1 章でも見たように、精密化を行うには、滑らかさが必要になるので、$\mathcal{M}_1^+(\Sigma)$ を適当な Banach 空間に埋め込む必要がある。ここでは、以下のような Hilbert 空間を考える、
$\{\psi_n\}$ を $L^2(d\mu)$ の完備直交基底、$a = (a_n)$ を実数列で次を満たすものとする。

$$\text{(i) } a_n > 0, \quad \text{(ii) } \lim_{n\to\infty} a_n = 0, \quad \text{(iii) } \sum_{n=1}^{\infty} a_n ||\psi_n||_\infty^2 = 1. \tag{4.2}$$

$\nu, \lambda \in \mathcal{M}(\Sigma)$ に対して

$$\langle \nu, \lambda\rangle_a = \sum_{n=1}^{\infty} a_n \left(\int_\Sigma \psi_n d\nu\right)\left(\int_\Sigma \psi_n d\lambda\right) \tag{4.3}$$

と定め、$H_a = \overline{\mathcal{M}(\Sigma)}^{||\cdot||_a}$ とおく。すると、$\mathcal{M}_1^+(\Sigma) \subset H_a$ はコンパクト部分集合となり、連続な $F : \mathcal{M}_1^+(\Sigma) \to \mathbb{R}$ は $\widetilde{F} : H_a \to \mathbb{R}$ に拡張できる。
次の条件を考える。

Ass. 1 $F : H_a \to \mathbb{R}$ は 2 階 Fréchet 可微分とする。

Ass. 2 任意の $\kappa \in K_F$ で ”$D^2(I - F)(\kappa) > 0$”.

この仮定のもとで、次の定理が成り立つ。

定理 4.2. Ass.1, Ass.2 を仮定する。すると次が成立する。

(i) $\sharp K_F > \infty$ となる。これを、$K_F = \{\kappa_1, \ldots, \kappa_n\}$ とおく。

(ii) $E^{\mu^{\otimes N}}\left[\mathrm{e}^{NF(L_N)}\right] = \mathrm{e}^{-Nb_F}\left(\sum_{i=1}^{n} a_i + o(1)\right)$
ここで、$a_i = \left(\det(I - D^2F(\kappa_i) \circ S_{\kappa_i})\right)^{-1/2}$,
ただし、S_{κ_i} は後の (4.6) で定義される作用素。

注意 4.3. あとで見るように、作用素 S_{κ_i} は核型なので、det は定義できる。

$(\Sigma^{\otimes\infty}, \mathcal{B}(\Sigma)^{\otimes\infty})$ 上の確率測度 Q_N を

$$Q_N(A) = \frac{E^{\mu^{\otimes\infty}}[\mathrm{e}^{NF(L_N)}\mathbf{1}_A]}{E^{\mu^{\otimes N}}[\mathrm{e}^{NF(L_N)}]} \tag{4.4}$$

で定義すると、その収束先は、上の定理 4.2 より次のように定まる。

系 4.4.

$$\lim_{N\to\infty} Q_N = \sum_{i=1}^{n} p_i \kappa_i^{\otimes\infty}, \quad p_i = \frac{a_i}{\sum_{j=1}^{n} a_j}$$

ここで、a_i は定理 4.2 で与えたもの。

Ass. 2 の正確な意味と、定理 4.2 に出てきた作用素 S_κ について、以下で説明する。
$X, X_1, X_2, \ldots$ を分布 μ を持つ独立同分布の確率変数とする。$\varphi \in C_b(\Sigma)$ を $\int_\Sigma \varphi d\mu = 0$ を満たすものとする。キュムラント母関数は

$$\Lambda_N(\varepsilon) = \frac{1}{N}\log E[\mathrm{e}^{\varepsilon\sum_{i=1}^{N}\varphi(X_i)}] = \log E[\mathrm{e}^{\varepsilon\varphi(X)}]$$

となるので、ε で微分して、

$$\Lambda_N'(\varepsilon) = \frac{E[\varphi(X)\mathrm{e}^{\varepsilon\varphi(X)}]}{E[\mathrm{e}^{\varepsilon\varphi(X)}]} = E^\varepsilon[\varphi(X)]$$

$$\text{ただし、}\ \mu^\varepsilon(dx) = \frac{\mathrm{e}^{\varepsilon\varphi(x)}}{\int_\Sigma \mathrm{e}^{\varepsilon\varphi(y)\mu(dy)}}\mu(dx)$$

を得る。同様に、

$$\Lambda_N''(\varepsilon) = E^\varepsilon[(\varphi(X) - E^\varepsilon[\varphi(X)])^2] = \mathrm{Var}^\varepsilon[\varphi(X)]$$
$$\Lambda_N'''(\varepsilon) = E^\varepsilon[(\varphi(X) - E^\varepsilon[\varphi(X)])^3]$$

を得る。これより、次が得られる。

補題 4.5.

$$\left|\frac{1}{N}\log E\left[\mathrm{e}^{\varepsilon\sum_{n=1}^{N}\varphi(X_n)}\right]-\frac{\varepsilon^2}{2}\int_\Sigma \varphi(x)^2\mu(dx)\right| \quad \leq ||\varphi||_\infty^3\varepsilon^3$$

補題 4.6. $\int_\Sigma \varphi d\mu = 0, \int_\Sigma r_3 d\mu = 0, r_3(\varepsilon) = o(\varepsilon)||\varphi||_\infty^3$ なる φ, r_3 に対して $d\mu^\varepsilon = (1+\varepsilon\varphi + r_3(\varepsilon))d\mu$ とおく。このとき

$$I(\mu^\varepsilon) = H(\mu^\varepsilon|\mu) = \frac{\varepsilon^2}{2}\int_\Sigma \varphi^2 d\mu + o(\varepsilon^2)$$

同様にして

補題 4.7. $\nu \ll \mu$ に対し $d\nu/d\mu = f\ (>0)$ とし、$d\nu^\varepsilon = f(1+\varepsilon\varphi + r_3)d\mu$ とする。ただし、$\int_\Sigma \varphi d\nu = 0, \int_\Sigma r_3 d\nu = 0, r_3 = o(\varepsilon)||\varphi||_\infty^3$ を満たすものとする。すると、

$$I(\nu^\varepsilon) = I(\nu) + \int_\Sigma (\varepsilon\varphi + r_3)\log f d\mu + \frac{\varepsilon^2}{2}\int_\Sigma \varphi^2 f d\mu + o(\varepsilon^2)$$

次に、作用素 S_κ について述べる。一般に $\varphi \in C_b(\Sigma)$ に対して $\nu \mapsto \int_\Sigma \varphi d\nu$ を H_a でみると連続になるとは限らぬが、(4.2) に出てきた ψ_n に対しては

$$\widehat{\psi_n} = \frac{1}{a_n}\psi_n d\mu$$

とおくと、$\widehat{\psi_n} \in H_a$ で、$\langle\widehat{\psi_n}, \nu\rangle_a = \int_\Sigma \psi_n d\nu$ が成立する。$C_0(\Sigma) = \{\psi_n$ の有限線形和$\}$ とおくと、任意の $\varphi \in C_0(\Sigma)$ に対して $\widehat{\varphi} \in H_a$ が存在して、全ての $\nu \in \mathcal{M}_1^+(\Sigma)$ に対して $\langle\widehat{\varphi}, \nu\rangle_a = \int \varphi d\nu$ が成立する。またさらに、

$$\ell_n = \frac{1}{\sqrt{a_n}}\psi_n d\mu \tag{4.5}$$

とおくと、$\{\ell_n\}$ は H_a の正規直交基底となる。
$\nu \in \mathcal{M}_1^+(\Sigma)$ に対して $S_\nu : H_a \to H_a$ を

$$S_\nu \ell_n = \sum_{m=1}^{\infty}\left(\sqrt{a_m a_n}\frac{1}{2}\int_\Sigma \psi_m\psi_n dnu\right)\ell_m \tag{4.6}$$

で定義する。すると

補題 4.8. $S_\nu : H_a \to H_a$ は核型作用素になる。

$I - F$ が K_F で最小をとることより、次を得る。

命題 4.9. (i) $\kappa \in K_F$ ならば、κ は μ に関して絶対連続で、$d\nu/d\mu = f$ とおくと、適当な定数 C をとって

$$f(x) = C\mathrm{e}^{F_1(\kappa)(x)}, \quad \text{ただし、} F_1(\kappa)(x) = DF(\kappa)[\delta_{\{x\}}]$$

(ii) $\kappa \in K_F$ ならば、$\int_\Sigma \varphi d\kappa = 0$ となる任意の $\varphi \in C_b(\Sigma)$ に対して

$$\int_\Sigma \varphi^2 d\kappa \geq D^2F(\kappa)[f\varphi, f\varphi]$$

ここで Ass. 2 を正確に述べなおすと次のようになる。
Ass. 2′ $\int_\Sigma \varphi d\kappa = 0$ を満たす恒等的に 0 でない任意の $\varphi \in C_b(\Sigma)$ に対して

$$\int_\Sigma \varphi^2 d\kappa > D^2F(\kappa)[f\varphi, f\varphi]$$

が成り立つ。

4.2 局所的な中心極限定理と定理の証明の概略

次に、$\kappa \in K_F$ の近くでの、局所的な中心極限定理を述べる。S_κ が核型作用素だったので、H_a 上の中心化された Gauss 測度 γ で、分散が S_κ であるもの i.e.

$$\int_\Sigma \langle h, \xi\rangle_a \langle k, \xi\rangle_a \gamma(d\xi) = \langle h, S_\kappa k\rangle_a$$

を満たすものが、存在する。また、補題 4.5 で $\varepsilon = N^{-1/2}$ ととることで、次の命題を得る。

命題 4.10. H_a 上で

$$\sqrt{N} L_N \to \gamma \quad \text{(in law)}$$

最後に、定理 4.2 の証明の概要を述べる。
(i) $H(\cdot|\mu)$ が good rate function であるから K_F はコンパクトとなり、Ass. 2 より $\sharp K < \infty$ となる。
(ii) $\sharp K_F = 1,\ K_F = \{\kappa\}$ とする。命題 4.9 (i) より、

$$H(\kappa|\mu) = \int_\Sigma F_1 d\kappa - \log \int_\Sigma \mathrm{e}^{F_1} d\mu$$

となり、一方任意の有界可測な Φ に対し、測度を μ から κ に変換すると、

$$E^{\mu^{\otimes N}}[\Phi(L_N)] = E^{\kappa^{\otimes N}}\left[\Phi(L_N)\mathrm{e}^{-N(\int F_1 dL_N - \log\int \mathrm{e}^{F_1} d\mu)}\right]$$

となるので

$$
\begin{aligned}
&\mathrm{e}^{Nb_F} E^{\mu^{\otimes N}}[\mathrm{e}^{NF(L_N)}] \\
&= E^{\mu^{\otimes N}}[\mathrm{e}^{N(F(L_N)-F(\kappa)+H(\kappa|\mu))}] \\
&= E^{\kappa^{\otimes N}}[\mathrm{e}^{N(F(L_N)-F(\kappa)+\int F_1 d(L_N-\kappa))}] \\
&= E^{\kappa^{\otimes N}}[\mathrm{e}^{N(F(L_N)-F(\kappa)-DF(\kappa)[L_N-\kappa]}))] \\
&= E^{\kappa^{\otimes N}}\left[\mathrm{e}^{(1/2)(D^2F(\kappa)[\sqrt{N}(L_N-\kappa),\sqrt{N}(L_N-\kappa)]+O(N^{-1/2}(\sqrt{N}||L_N-\kappa||_\infty)^3)}\right] \\
&= I_1 + I_2
\end{aligned}
$$

ここで、$I_1 = E^{\kappa^{\otimes N}}[\mathrm{e}^{-},\ ||L_N-\kappa||_a < c]$, $I_2 = E^{\kappa^{\otimes N}}[\mathrm{e}^{-},\ ||L_N-\kappa||_a \geq c]$. すると、大偏差原理が成立することより、$I_2 \to 0$ $(\forall c > 0)$ また、十分小の $c > 0$ に対し、Ass. 2 より $\mathrm{e}^{(1/2)D^2F(\kappa)[\sqrt{N}(L_N-\kappa),\sqrt{N}(L_N-\kappa)]}$ が一様可積分となって

$$
I_1 \to \int_{H_a} \mathrm{e}^{(1/2)F_2(\xi,\eta)}\gamma(d\xi)\gamma(d\eta) = (\det(I - D^2F(\kappa)\circ S_\kappa))^{-1/2}
$$

となり、定理 4.2 は証明される。

注意 4.11. (i) ここでは $I - F$ のヘシアンが非退化との条件のもとで精密化を述べたが、退化した場合も許す一般的枠組みも与えられている。([14], [23], [25] 等参照)

(ii) 一般の *Banach* 空間に値をとる独立確率変数の和に対する大偏差原理については、ここで述べた結果に対応するものとして、中心極限定理が成り立つという仮定の下での *Bolthausen* の結果 [13] 等があるが、この条件を弱め、中心極限定理が成立しない場合の大偏差原理に関しては、*Kusuoka-Liang* [19] を初め *Liang* による結果 [25], [26] 等がある。

(iii) また、この章で扱ったのは、ポテンシャルのオーダーと、*rate function* のオーダーが一致している場合であったが、これらのオーダーが異なっている場合については、千代延による結果 [17] がある。

5 Donsker-Varadhan 理論

ここでいう「Donsker-Varadhan 理論」とは「Donsker と Varadhan によるマルコフ過程の経験分布に対する大偏差原理についての 1970 年代の中頃から 80 年代にかけての一連の共同研究の成果として確立した理論」という意味である.

独立同分布な確率変数列 (i.i.d.)$X_1, X_2, \cdots$ のそれぞれの分布が $E[e^{\xi X_1}] < \infty$, $\forall \xi \in \mathbf{R}$ であるとき、その経験分布

$$L_n = \frac{1}{n}\sum_{n=1}^{\infty}\delta_{X_i}$$

に対して、大偏差原理が成立する事を見た (Sanov の定理). では、確率変数列の独立性の仮定を落とした場合になお大偏差原理が成立するだろうか. これは確率論の問題として自然で重要である.

M. D. Donsker と S. R. S. Varadhan はマルコフ過程の経験分布 (離散マルコフ過程 $X_1, X_2, \cdots$ の L_n、連続マルコフ過程 $X_t, t \geq 0$ の

$$L_t = \frac{1}{t}\int_0^t \delta_{X_s} ds$$

に対する大偏差原理が成立するための十分条件を提示し、そのクラスのマルコフ過程の経験分布に対して大偏差原理を証明した. しかし、彼らの理論の重要性は単に i.i.d. ではない確率変数列に対する経験分布の大偏差原理を証明したという事だけにあるのではない. 彼らの仕事の最も重要な点は、スペクトル理論やポテンシャル論などの解析学の諸問題、あるいはランダム系やポラロンなどの数理物理学のいくつかの問題においてマルコフ過程の大偏差原理を示す事が本質的である事を見抜き、現実に大偏差原理を用いてこれらの問題を解決したことにある. 言い換えると、これらの諸問題をある (非常に大きい) 空間上の積分のラプラス型の漸近挙動の問題であると捉える事により、

1. 大偏差原理を示す. とくに、レート関数を表示する式を得る.
2. しかる後に Varadhan の積分公式を適用する.

といったモデルによらない普遍的な方法論を確立した点に意義があるのである.

Donsker-Varadhan の一連の論文においてしめされた考え方や方法論は、さまざまな一般化 (より広いクラスの確率過程や確率場に対して大偏差原理を証明する事)、および数多くの応用 (数学や数理物理学のさまざまな問題に大偏差原理を適用する事) を生みながら発展を続けている事は言うまでもない. しかし、それらの成果を現時点での知見から見直してその全体像を提示するといったことは著者の見識・能力をはるかに超えた作業である. むしろ、この講義では、彼らの一連の論文のなかでも最も古い論文 [33] および [34] に焦点をしぼり、その内容を紹介したい. 最もシンプルなモデルが扱われているため彼らのアイデアを学びやすいと思うからである. 特にこれからこの理論を学びたい人は是非 [33] を見ていただきたい. なお、本年度 Varadhan 氏が Abel Prize を受賞されたことから、舟木先生が彼の仕事を紹介した一文を

「数学セミナー」(2007, 8 月号) に草されているいるのでそれも紹介しておきたい．

まず最初に背景を述べよう．V を $\lim_{y\to\pm\infty} V(y)=\infty$ であるような $\mathbf{R}$ 上の連続関数とする．このような V に対して固有値問題

$$\frac{1}{2}\psi''(y)-V(y)\psi(y)=-\lambda\psi(y)$$

のスペクトルは離散スペクトル (固有値)$\{\lambda_1<\lambda_2\le\cdots\}$ のみからなる．(任意の関数解析の教科書を見よ．) 対応する正規化された固有関数列を $\{\psi_1,\psi_2,\cdots\}$ としよう．このとき、放物型方程式

$$\frac{\partial u}{\partial t}=\frac{1}{2}\frac{\partial^2 u}{\partial x^2}-V(x)u$$

の基本解 $u(t,x,y)$ に対し

$$u(t,x,y)=\sum_{k=1}^{\infty} e^{-\lambda_k t}\psi_k(x)\psi_k(y)$$

が成立するが、この基本解はまた Feynmann-Kac の公式により (形式的に)

$$u(t,x,y)=E_x\big[\exp(-\int_0^t V(w_s)ds)\delta(w_t-y)\big]$$

とも表現することができる．これらの両辺の y について積分のとると、

$$E_x\big[\exp(-\int_0^t V(w_s)ds)\big]=\sum_{k=1}^{\infty} e^{-\lambda_k t}\psi_k(x)$$

である．これより、M. Kac が示した等式

$$\lim_{t\to\infty}\frac{1}{t}\log E_x\big[\exp(-\int_0^t V(w_s)ds)\big]=-\lambda_1$$

が得られる．つまり、微分作用素の最小固有値が Wiener 汎関数の積分によって表現されたことになる．一方、λ_1 は良く知られているように

$$\lambda_1=\inf\big\{\frac{1}{2}\int_{\mathbf{R}}\psi'(y)^2dy+\int_{\mathbf{R}}V(y)\psi^2(y)dy;\ \int_{\mathbf{R}}\psi^2(y)dy=1\big\}$$

とも表現される．よってこの 2 つの式をあわせると次の定理が得られる；

定理 5.1.

$$\lim_{t\to\infty}\frac{1}{t}\log E_x\big[\exp(-\int_0^t V(w_s)ds)\big]$$
$$= -\inf\big\{\frac{1}{2}\int_{\mathbf{R}}\psi'(y)^2dy+\int_{\mathbf{R}}V(y)\psi^2(y)dy;\ \psi\in L^2(\mathbf{R}),\ \int_{\mathbf{R}}\psi^2(y)dy=1\big\}.$$

この公式の両辺はそれぞれが λ_1 に等しいという関係から示されたが、直接右辺の Wiener 汎関数の積分の漸近挙動を計算することにより「右辺＝左辺」を証明できないだろうか．もしそれができるような一般論が構築できれば、最小固有値という方程式論的な量が対応しないような、より複雑な Wiener 汎関数の積分の漸近評価も可能になるはずである．M. Kac の研究以来長らく意識されていたこの問に答えたのが Donsker と Varadhan である．彼らがこの問にどのように答えをあたえたか、以下見ていこう．

$\mathcal{M}_1(\mathbf{R})$ を $\mathbb{R}$ 上の確率測度の空間とする．これは $\mathbb{R}$ 上の分布関数の空間 $\mathcal{F}$ と同一視できる．絶対連続な測度の空間を $\mathcal{F}_{\dashv|}$ と書く．$\mu(dx)=dF(x)=f(x)dx$ であるとき、μ, F および f を同一視する．以後 $\Phi:\mathcal{M}_1(\mathbf{R})\to\mathbf{R}$ は以下をみたすものとする．

Ass. 1 $0\le\Phi(\mu)\le\infty$.

Ass. 2 Φ は $M_1(\mathbf{R})$ 上下半連続.

Ass. 3 任意の $K>0$ に対して $\{\mu;\ \Phi(\mu)\le K\}$ は $M_1(\mathbf{R})$ のコンパクト集合である．

Ass. 4 $\mu_n\to\mu$ in law かつすべての n に対して $supp(\mu_n)\in[a,b]$ ならば $\Phi(\mu_n)\to\Phi(\mu)$.

定理 5.2. Ass.1-4 を満たす Φ に対して

$$\lim_{t\to\infty}\frac{1}{t}\log E_x\big[e^{-t\Phi(L_t)}\big]=-\inf_{f\in\mathcal{F}_{\dashv|}}\big\{\Phi(f)+\frac{1}{8}\int_{\mathbf{R}}\frac{f'(y)^2}{f(y)}dy\big\}.$$

この定理において、$\psi(y)=\sqrt{f(y)}$、$\Phi(\mu)=\int_{\mathbf{R}}V(y)\mu(dy)$ とすると、定理 5.1 の公式が得られる．

この公式が **1.1** で述べた「ラプラスの方法」のひとつの形であることを注意しておこう．実際、$\mathcal{F}=\mathcal{M}_1(\mathbb{R})$ の上の確率測度 $Q_x^t(A)=P_x(L_t\in A)$ が形式的に「密度関数」$Q_x^t(F)$ をもつと想像してみると、

$$
\begin{aligned}
E_x\left[e^{-t\Phi(L_t)}\right] &= \int e^{-t\Phi(F)} dQ_x^t(F) \\
&= \int e^{-t\Phi(F)} Q_x^t(F)[\mathcal{D}F]
\end{aligned}
$$

と書ける．ここで、$[\mathcal{D}F]$ は分布関数の空間の上の「ルベーグ測度」である．もしある $I : \mathcal{F} \to [0, \infty)$ が存在して $t \to \infty$ のとき

$$
Q_x^t(F)[\mathcal{D}F] \asymp \exp\left\{-tI(F)\right\}[\mathcal{D}F]
$$

であれば

$$
E_x\left[e^{-t\Phi(L_t)}\right] \asymp \int \exp\left\{-t(\Phi(F) + I(F))\right\}[\mathcal{D}F]
$$

と表現され、**1.1** で述べた積分と (関数が定義される空間がこちらのほうがはるかに大きいが) 形としては同じである．したがって **1.1.** の場合と同様

$$
\lim_{t\to\infty} \frac{1}{t} \log E_x\left[e^{-t\Phi(L_t)}\right] = -\inf_{F\in\mathcal{F}} \left\{\Phi(F) + I(F)\right\}.
$$

が成り立つと期待される．すなわち、定理 5.2 はラプラスの方法を無限次元の空間で適用した得られたものと見る事ができる．問題は、ここでのべた形式的な議論をいかに数学として正当化し I を具体的に計算するかということにある．

Upper Estimate.

分布関数の空間上の汎関数 I を

$$
I(F) = -\inf_{u\in\mathcal{U}} \int_{\mathbf{R}} \frac{u''(y)}{2u(y)} dF(y),
$$

ただし

$$
\mathcal{U} = \{u \in C_b(\mathbf{R}); u', u'' \in C_b(\mathbf{R}),\ 0 < \exists\alpha \le u \le \exists\beta\}
$$

とおく．

命題 5.3. 任意のコンパクト集合 $C \subset \mathcal{F}$ に対して

$$
\liminf_{t\to\infty} \log Q_x^t(C) \le -\inf_{F\in C} I(F).
$$

証明の概略: 任意の $u \in \mathcal{U}$ に対して Feynman-Kac の公式より

$$E_x[u(w_t)\exp(-\int_0^t (\frac{u''}{2u})(w_s)ds)] = u(x)$$

であるから、任意の $C \subset \mathcal{F}$ に対して

$$Q_x^t(C) \leq \frac{u(x)}{\alpha}\exp(t\sup_{F\in C}\int_{\mathbf{R}}(\frac{u''}{2u})(y)dF(y)$$

である．したがって、

$$\limsup_{t\to\infty}\frac{1}{t}\log Q_x^t(C) \leq \inf_{u\in\mathcal{U}}\sup_{F\in C}\int_{\mathbf{R}}(\frac{u''}{2u})(y)dF(y)$$

である．C のコンパクト性より、inf と sup を交換することができる． □

Φ に対する仮定 Ass.1-4 に注意すると、命題 5.3 より、定理 2.13. の証明と同様に

$$\lim_{t\to\infty}\frac{1}{t}\log E_x\big[e^{-t\Phi(L_t)}\big] \leq -\inf_{F\in\mathcal{F}}\big\{I(F)+\Phi(F)\big\}. \tag{5.1}$$

を得る事ができる．

Lower Estimate.

各 $x \in \mathbf{R}$ に対して、f_x を $\mathbb{R}$ 上の密度関数であって、ある $-\infty < a < x << b < \infty$ が存在して (a,b) 上 $f > 0$ かつ $[a,b]$ の外で $f = 0$ であり、

$$\int_a^b \frac{f'(y)^2}{f(y)}dy < \infty$$

である、という条件をみたすものの集まりとする．次の命題が成立する．

命題 5.4. $f \in f_x$ の $\mathcal{F}$ における任意の近傍 N_f に対し

$$\liminf_{t\to\infty}\frac{1}{t}\log P_x\{L_t \in N_f, a \leq w_s \leq b \textit{ for all } 0 \leq s \leq t\} \geq -\frac{1}{8}\int_a^b \frac{f'(y)^2}{f(y)}dy$$

が成立する．

証明の概略：$(\tilde{P}_x, z_t)$ を $b(y)=\frac{1}{2}f'(y)/f(y)$ をドリフトにもつブラウン運動、すなわち

$$\frac{1}{2}\frac{d^2}{dx^2}+\frac{1}{2}\frac{f'}{f}(x)\frac{d}{dx}$$

を生成作用素にもつ拡散過程とする．$z.$ は (a,b) 内に留まり続ける拡散過程である．ギルザノフ-丸山の公式により

$$P_x(A)=E^{\tilde{P}_x}[\exp\{-\frac{1}{2}\int_0^t\frac{f'}{f}(z_s)dz_s+\frac{1}{8}\int_0^t\Big(\frac{f'}{f}\Big)^2(z_s)ds\},A]$$

である．ここで伊藤の公式により

$$d\log f(z_s)=(\log f)'(z_s)dz_s+\frac{1}{2}(\log f)''(z_s)ds$$

であるから、

$$P_x(L_t\in N_f)=E^{\tilde{P}_x}[\Big\{\frac{f(x)}{f(z_t)}\Big\}^{1/2}\exp\{-\frac{1}{8}\int_0^t\Big(\frac{f'}{f}\Big)^2(z_s)ds-\frac{1}{4}\int_0^t\Big(\frac{f''}{f}\Big)(z_s)ds\},A]$$

である．よって

$$J(y)=\frac{1}{8}\Big(\frac{f'}{f}\Big)^2(y)-\frac{1}{4}\Big(\frac{f''}{f}\Big)(y),$$

とし、任意の $\epsilon>0$ に対して

$$S(t,\epsilon)=\Big\{\omega;\Big|\int_a^b JdL_t-\int_a^b Jfdy\Big|<\epsilon\Big\},\qquad S'(t,\epsilon)=\{\omega;L_t\in N_f\}\bigcap S(t,\epsilon)$$

とすると

$$P_x(L_t\in N_f)\geq\exp\left\{-t\left(\epsilon+\int_a^b Jfdy\right)\right\}E^{\tilde{P}_x}[\Big\{\frac{f(x)}{f(z_t)}\Big\}^{1/2};S'(t,\epsilon)].$$

である．生成作用素 $\frac{1}{2}\frac{d^2}{dx^2}+b(x)\frac{d}{dx}$ をもつ拡散過程に対し

$$\frac{1}{2}u''-[b(x)u(x)]'=0$$

の解 u がこの拡散過程の不変分布 $u(x)dx$ を与えるから、今の場合 $f(x)dx$ が $(\tilde{P}_x,z.)$ の不変測度である．したがって、エルゴード定理より

$$\lim_{t\to\infty}\tilde{P}_x(S'(t,\epsilon))=1$$

であり、命題を得る． □

命題 5.3 より、再び定理 2.13. の証明と同様に

$$(5.2) \qquad \lim_{t\to\infty}\frac{1}{t}\log E_x\big[e^{-t\Phi(L_t)}\big] \geq -\inf_{f\in F}\Big\{\Phi(f)+\frac{1}{8}\int_a^b\frac{f'(y)^2}{f(y)}dy\Big\}$$

を得る事ができる．上からの評価 (5.1) および下からの評価 (5.2) より、次の補題をしめせば定理 (5.2) の証明が終わる．

補題 5.5. $-\infty \leq a < b \leq \infty$ とする．(a,b) 上 $f>0$ かつ $[a,b]$ 外で $f=0$ であるような密度関数 $f\in C^1(\mathbb{R})$ に対して

$$I(F)=\frac{1}{8}\int_a^b\frac{f'(y)^2}{f(y)}dy.$$

証明の概略：$f=e^h$ と書くと

$$I(f)=-\inf_h\int_a^b\frac{(e^h)''}{e^h}fdx=-\frac{1}{2}\inf_h\int_a^b(h''+h'^2)fdx$$

である．よって、$h=\frac{1}{2}\log f-W$ とおくと

$$I(f)=-\inf_W\int_a^b\left(\frac{f''}{4}-\frac{f'^2}{8f}-(W'f)'+\frac{W'^2f}{2}\right)dx$$

と変形される．積分の第 1 項、第 3 項は消えるので、$W'=0$ ととって補題を得る． □

最後に、円周上の滑らかなドリフトをもつ 1 次元ブラウン運動に対する大偏差原理の結果を述べる．この結果は周期的ポテンシャルを持つ Schrodinger 作用素のグリーン関数に対するあるポテンシャル論的な問題に応用されている ([35])．

$b\in C^\infty(\mathbf{S}^1)$ とし、$\{P_x,\ x\in\mathbf{S}_1,(x_t)_{t\geq 0}\}$ を $\mathbf{S}^1$ 上の

$$\frac{1}{2}\frac{d^2}{dx^2}+b(x)\frac{d}{dx}$$

を生成作用素とする拡散過程とする．(これは、b を $\mathbb{R}$ 上に周期的に拡張して $\mathbb{R}$ 上の拡散過程と同一視できる．) これに対して次が成立する．

定理 5.6. 上の拡散過程の経験分布 L_t に対して定理5.2と同じ形の大偏差原理が成立する．ただしレート関数 I は

$$I(F) = \frac{1}{8}\int_0^1 \frac{f'^2}{f}(x)dx + \frac{1}{2}\int_0^1 [b^2 + b'](x)f(x)dx - \frac{1}{2}\frac{\left(\int_0^1 b(x)dx\right)^2}{\int_0^1 \frac{1}{f(x)}dx}.$$

によりあたえられる．

証明はブラウン運動の場合とほとんど同じ考え方で実行することができる．多次元の場合にもレート関数の具体的な表示式は知られている ([34])．結果だけ述べておく．

定理 5.7. $b_j \in C^\infty(\mathbb{R}^d)$, $j = 1, \cdots n$ とする．$L = \frac{1}{2}\Delta + b\cdot\nabla$ に対して

$$I(\mu) = -\inf_u \int_{\mathbb{R}^d} \frac{Lu}{u}(x)\mu(dx)$$

とおく．$\mu(dx) = f(x)dx$, $f \in C^\infty(\mathbb{R}^d)$ である μ に対して

$$I(f) = \frac{1}{8}\int_{\mathbb{R}^d} \frac{|\nabla f|^2}{f}dx + \frac{1}{2}\int_{\mathbb{R}^d}[b^2 + \nabla\cdot b]fdx - \frac{1}{2}\inf_{W\in C^\infty(\mathbb{R}^d)}\int_{\mathbb{R}^d}|b - \nabla W|^2 fdx$$

がなりたつ．

最後に、極めて中途半端ではあるが、今思いつく関連する話題や研究についていくつか触れる．この理論はM. Kacの数学をその源流の一つにもっている．M. Kacの講義録 [9] は、彼の数学ののひとつの要約になっており、極めて興味深い書物である．そこで述べられている内容の多くがDonsker-Varadhan理論と密接な関係を持っている．例えば、scattering lengthの半古典極限の話題は [37] で論じられている．また、スペクトル論やポテンシャル論と大偏差原理の関連では、竹田先生達の一連の研究があることは言うまでもない ([30], [31], [32] など)．多様体上の拡散過程に対するDonsker-Varadhan理論ということでも文献を挙げだすときりがないが、筆者にとって身近なものとして、日本におけるマルコフ過程の大偏差原理の最初期の論文の一つである [36], 市原先生のピン止めされた拡散過程の大偏差原理の研究 [38] や、最近の楠岡-桑田-田村の確率線積分の大偏差原理の研究を挙げておく．

ただし、この節の最初に述べたように、この講義 (ノート) は Donsker-

Varadhan 理論の発展の全体を俯瞰する総合報告ではない．充実した文献リストをもつ本として [3] を挙げておきたい．

:

References

[1] J. M. Bismut, Large Deviations and Malliavin Calculus, Birkhäuser (1984).

[2] A. Dembo and O. Zeitouni, Large Deviations Techniques and Applications, 2nd ed., Springer Verlag (1998).

[3] J. D. Deuschel and D. W. Stroock, Large Deviations, Academic Press (1984).

[4] R. S. Ellis, Entropy, Large Deviations, and Statistical Mechanics, Springer Verlag (1985).

[5] F. den Hollander, Large Deviations, American Mathematical Society (2000).

[6] D. W. Stroock, An Introduction to the Theory of Large Deviations, Springer (1984).

[7] S. R. S. Varadhan, Large Deviations and Applications, SIAM (1984).

[8] V. Vinogradov, Refined Large Deviation Limit Theorems, Longman Scientific & Technical (1994).

[9] M. Kac, Integration in function spaces and some of its applications, Lezioni Fermiane vol.4, Pisa (1980),

[10] G. Ben Arous and M. Brunaud, *Méthode de Laplace: étude variationnelle des fluctuations de diffusions de type "champ moyen".* Stoch. Stoch. Rep. **31**(1990), pp. 79–144.

[11] G. Ben Arous, J. D. Deuschel and D. W. Stroock, *Precise asymptotics in large deviation.* Bull. Sci. Math. **117**(1990), pp. 107–124.

[12] G. Ben Arous and A. Gionet, *A large deviation for Wigner's law and Voiculescu's non-commutative entropy.* Prob. Th. Rel. Fields **108**(1997), pp. 517–542.

[13] E. Bolthausen, *Laplace approximations for sums of independent random vectors.* Prob. Th. Rel Fields **72**(1986), pp. 305–318.

[14] E. Bolthausen, *Laplace approximations for sums of independent random vectors II.* Prob. Th. Rel Fields **76**(1986), pp. 167–206.

[15] E. Bolthausen, J. D. Deuschel and Y. Tamura, *Laplace approximations for large deviations of nonreversible Markov processes. Part I.* Ann. Probab. **23**(1995), pp. 236–267.

[16] T. Chiyonobu and S. Kusuoka, *The large deviation principle for hypermixing processes.* Prob. th. Rel. Fields **78**(1988), pp. 627–649.

[17] T. Chiyonobu, *A limit formula for a class of Gibbs measures with long range pair interactions.* J. Math. Sci. ,Univ. Tokyo **7**(2000), pp. 463–486.

[18] T. Kumagai, *Short time asymptotic behavior and large deviation for Brownian motion on some affine nested fractals.* Publ. Res. Inst. Math. Sci. **33**(1997), pp. 223–240.

[19] S. Kusuoka and S. Liang, *Laplace approximations for sums of independent random vectors.* Prob. th. Rel. Fields **116**(2000), pp. 221–238.

[20] S. Kusuoka and S. Liang, *Laplace approximations for diffusion processes on torus: nondegenerate case.* J. Math. Sci., Univ. Tokyo **8**(2001), pp. 43–70.

[21] S. Kusuoka and Y. Tamura, *Gibbs measures for mean field potentials .* J. Fac. Sci., Univ. Tokyo SectIA **31**(1984), pp. 223–245.

[22] S. Kusuoka and Y. Tamura, *Symmetric Markov processes with mean field potentials.* J. Fac. Sci., Univ. Tokyo SectIA **34**(1987), pp. 371–389.

[23] S. Kusuoka and Y. Tamura, *Precise estimate for large deviation of Donsker-Varadhan type.* J. Fac. Sci., Univ. Tokyo SectIA **38**(1991), pp. 533–565.

[24] S. Kusuoka and D. W. Stroock, *Precise asymptotics of certain Wiener functionals.* J. Funct. Anal. **99**(1991), pp. 1–74.

[25] S. Liang, *Laplace approximations for sums of independent random vectors–the degenerate case.* J. Math. Sci., Univ. Tokyo **7**(2000), pp. 195–220.

[26] S. Liang, *Large deviation principles for a type of diffusion processes on Euclidean spaces.* J. Math. Sci., Univ. Tokyo **10**(2003), pp. 555–577.

[27] A. V. Nagaev, *Cramèrlarge deviations when the extreme conjugate distribution is heavy-tailed.* Theory Probab. Appl. **43**(1999), pp. 405–421.

[28] V. V. Petrov, *On the probability of large deviations for sums of independent random variable.* Theory Probab. Appl. **10**(1965), pp. 405–421.

[29] M. Takeda, *On a large deviation for symmetric Markov processes with finite life time.* Stoch. Stoch. Rep. **59**(1996), pp. 143–167.

[30] M. Takeda, *Large deviation principle for additive functionals of Brownian motion corresponding to Kato measures.* Potential Analysis,**19**(1), (2003), 51-67

[31] M.Takeda, Y.Shiozawa, *Variational formula for Dirichlet forms and estimates of principal eigenvalues for symmetric stable processes.*Potential Analysis, **23** ,(2005),135-151

[32] M.Takeda, K.Tsuchida *Differentiability of spectral functions for symmetric* α*-stable processes.* Trans. Amer. Math. Soc., **359**(8),(2007),4031-4054

[33] M. D. Donsker and S. R. S. Varadhan, *Asymptotic evaluation of certain Wiener integrals for large time.* in "Functional Integration and Its Applications," Proceedings of the International Conference Held at Cumberland Lodge. (1974), Edited by A.M. Arthurs, Clarenton, Oxford, pp. 15–33.

[34] M. D. Donsker and S. R. S. Varadhan, *On the principal eigenvalue of second-order ellipic differential operators.* Comm. Pure Appl. Math. **29**(1976), pp. 595–621.

[35] C. Schroeder, *Green's function for the Schrodinger operator with periodic potential.* Jour. Funct. Anal. **77**(1988), pp. 60–87.

[36] H. Okura,*Theorems on large deviations for a family of stochastic processes converging to a Markov process*, Osaka J. Math., **22,** 533-573 (1985)

[37] Y. Takahashi, *An integral representation on the path space for scattering length.*, Osaka J. Math. **27**, 373-379 (1990)

[38] K. Ichihara, *Large deiviation for pinned covering diffusion.* Bull. Sci. math. **125**, 6-7 (2001) 529-551