経営と経済　第95巻 第３・４号　2016年３月

一様分布族のもとでの事後分布に対する大偏差原理

式　見　拓　仙

Abstract

In this article, we consider a family of uniform distributions as a statistical model. Assuming that the prior distribution has a smooth, positive density on the parameter space, we prove the large deviation principle of the posterior distributions. We derive the rate functions explicitly. To this end, we apply the Gärtner-Ellis theorem to the posterior distributions of the reciprocal of the parameter.

Keywords: posterior distributions, large deviations, the Gärtner-Ellis theorem, Laplace's method

1．はじめに

事後分布に対する大偏差原理は Ganesh and O'Connell (1999)，Ganesh and O'Connell (2000)，Macci (2010)，Macci (2011) らによるいくつかの研究成果がある．Ganesh and O'Connell (1999) はデータの取り得る空間が有限集合である場合に事後分布の大偏差原理を導いている．Ganesh and O'Connell (2000) は事前分布が Dirichlet であるケースについて論じている．Macci (2010) は threshold モデルを統計モデルとする場合に，最尤推定量と事後分布の大偏差原理を証明している．また，同論文には１次元の自然指数分布族を統計モデルとする場合についても，事後分布の大偏差原理を導いている．そのいくつかの具体例は Macci (2011) に挙げられている．

統計モデルとして，一様分布族 $U(0;\theta)$; $\theta>0$ を仮定すると，事後分布の大偏差原理は Macci (2010，Proposition 3) の特別なケースとして得られる．本論文では，この結果を Gärtner-Ellis の定理を応用例として示す．

必要な定義や結果を概括しておく．S を位相空間，$\mathscr{B}(S)$ を S の Borel-σ-代数とする．I を S から $[0,\infty]$ への関数とする．任意の $M\geqslant 0$ に対して，$\{I\leqslant M\}$ がコンパクトであるという条件を満たすとき，I を S 上のレート関

数と呼ぶ．(μ_n) を S 上の分布列とする．以下の条件が成り立つとき，(μ_n) は S 上で，レート関数 I を持つ大偏差原理 (large deviation principle) を満たすという：S のすべての閉集合 F に対して，

$$\limsup_n \frac{1}{n} \log \mu_n(F) \leqq -\inf_{x \in F} I(x).$$

であり，さらに S のすべての開集合 G に対して，

$$\liminf_n \frac{1}{n} \log \mu_n(G) \leqq -\inf_{x \in G} I(x).$$

大偏差原理という用語は，以後，LDP と略記する．Gärtner-Ellis の定理は LDP を導くうえで，応用範囲の広い有用な道具である．次章にて，ある事後分布の LDP を導くために利用する．その応用を念頭おいて，$\mathbb{R}$ 上の分布列 (μ_n) に対して，Gärtner-Ellis の定理を述べておく．μ_n のキュムラント母関数を $K_n(t)$ とする．すべての $t \in \mathbb{R}$ に対し，$\lim_n K_n(nt)/n$ が $\bar{\mathbb{R}} = [-\infty, \infty]$ に存在するならば，それを $K(t)$ で表すことにする．K の有効領域は $\mathscr{D}_K = \{t \in \mathbb{R} : K(t) < \infty\}$ によって定義される．K の Fenchel-Legendre 変換は

$$K^*(t^*) = \sup_{t \in \mathbb{R}} [t^* t - K(t)]$$

によって定義される．K^*も凸関数になる．K^*の有効領域も $\mathscr{D}_{K^*} = \{t^* \in \mathbb{R} : K^*(t^*) < \infty\}$ によって定義される．$s^* \in \mathscr{D}_{K^*}$が K^*の exposed point であるとは，ある $s \in \mathbb{R}$ が存在して，

$$\forall t^* \neq s^* K^*(t^*) - K^*(s^*) > s(t^* - s^*) \tag{1.1}$$

が成り立つことである．K^*の exposed point に対して，(1.1) を成り立たしめる s を s^*の exposing hyperplane とよぶ．K^*の exposed point のうち，その exposing hyperplane が $\mathscr{D}_K$ の内部に属するような点の全体を $\mathscr{E}_{K^*}$で表すことにする．

補題1.1 (Gärtner-Ellis). (μ_n) を $\mathbb{R}$ 上の分布列とする．もし，すべての $t \in \mathbb{R}$ に対し，$K(t)$ が存在し，0 が $\mathscr{D}_K$ の内点ならば，K^*は $\mathbb{R}$ の凸レート関数であり，$\mathbb{R}$ のすべての閉集合 F に対し，

$$\limsup_n \frac{1}{n} \log \mu_n(F) \leqslant -\inf_{t^* \in F} K^*(t^*)$$

であり，$\mathbb{R}$ のすべての開集合 G に対して，

$$\liminf_n \frac{1}{n} \log \mu_n(G) \geqslant -\inf_{t^* \in \mathscr{E}_{K^*} \cap G} K^*(t^*)$$

である．

この証明や一般の位相線型空間上の分布列に対する Gärtner-Ellis の定理については Dembo and Zeitouni (1998) を見よ．

$K^*(t^*)$ の $\Theta \in \mathscr{B}(\mathbb{R})$ への制限を $K^*(\theta)$ で表すことにする．また，$(\mathbb{R}, \mathscr{B}(\mathbb{R}))$上の分布列 (μ_n) に対し，μ_n の Θ への制限をμ_n^Θ で表すことにする．

系1.2. $\Theta \in \mathscr{B}(\mathbb{R})$ であるとし，(μ_n) は $\mu_n(\Theta)=1, n=1,2,...$ を満たす$(\mathbb{R}, \mathscr{B}(\mathbb{R}))$上の分布列とする．さらに，すべての $t \in \mathbb{R}$ に対し，$K(t)$ が存在し，0 は $\mathscr{D}_K$ の内点であると仮定する．もし，$\mathscr{D}_{K^*} \subset \Theta$ であるならば，$K^*(\theta)$は Θ上のレート関数であり，Θのすべての閉集合 F に対して，

$$\limsup_n \frac{1}{n} \log \mu_n^\Theta(F) \leqslant -\inf_{\theta \in F} K^*(\theta)$$

が成り立つ．もし，Θが $\mathbb{R}$ の開集合で，$\mathscr{E}_{K^*} = \mathscr{D}_{K^*}$ならば，$\Theta$のすべての開集合 G に対して，

$$\liminf_n \frac{1}{n} \log \mu_n^\Theta(G) \geqslant -\inf_{\theta \in G} K^*(\theta)$$

が成り立つ．

証明. $\mathscr{D}_{K^*}\subset\Theta$より，すべての$M\geqslant 0$に対して，$\{t^*\in\mathbb{R}:K^*(t^*)\leqslant M\}\subset\Theta$であるから，$\{t^*\in\mathbb{R}:K^*(t^*)\leqslant M\}$は$\Theta$のコンパクト集合である．また，$\{\theta\in\Theta:K^*(\theta)\leqslant M\}=\{t^*\in\mathbb{R}:K^*(t^*)\leqslant M\}$であるから，$K^*(\theta)$が$\Theta$上のレート関数であることがわかる．

もし，FがΘの閉集合ならば，$F=\Theta\cap F'$となる$\mathbb{R}$の閉集合F'が存在する．よって，Gärtner-Ellis の定理より，

$$\limsup_n \frac{1}{n}\log \mu_n^{\Theta}(F)=\limsup_n \frac{1}{n}\log \mu_n^{\Theta}(\Theta\cap F')=\limsup_n \frac{1}{n}\log \mu_n(\Theta\cap F')$$

$$=\limsup_n \frac{1}{n}\log \mu_n(F')\leqslant -\inf_{t^*\in F'} K^*(t^*)$$

$$=-\left\{\inf_{t^*\in F'\cap\Theta} K^*(t^*)\wedge\inf_{t^*\in F'\cap\Theta^c} K^*(t^*)\right\}$$

$$=-\inf_{t^*\in F'\cap\Theta} K^*(t^*)=-\inf_{\theta\in F} K^*(\theta).$$

次にΘは$\mathbb{R}$の開集合であるとする．このとき，Θの開集合Gは$\mathbb{R}$の開集合でもあるので，Gärtner-Ellis の定理および$\mathscr{E}_{K^*}=\mathscr{D}_{K^*}$という仮定より，

$$\liminf_n \frac{1}{n}\log \mu_n^{\Theta}(G)=\liminf_n \frac{1}{n}\log \mu_n(G)\geqslant \inf_{t^*\in\mathscr{D}_{K^*}\cap G} K^*(t^*)$$

$$=-\inf_{t^*\in G} K^*(t^*)=-\inf_{\theta\in G} K^*(\theta).$$

以上で系は証明された． □

$(\Theta,\mathscr{U}),(\mathscr{X},\mathscr{A})$を可測空間，統計モデル$(P_\theta)_{\theta\in\Theta}$は$\Theta$から可測空間$\mathscr{X}$上の stochastic kernel であると仮定する．ϑはΘに値をとるパラメータで，事前分布πを持つとする．$X_1,X_2,\ldots$はϑのもとで条件付 i.i.d. 確率変数列であり，$P_\vartheta(A),A\in\mathscr{A}$は$X_1$の$\vartheta$のもとでの正則条件付確率であるとする．$(\Theta,\mathscr{U}),\pi,(\mathscr{X},\mathscr{A}),(P_\theta)_{\theta\in\Theta}$が与えられれば，確率空間$(\Omega,\mathscr{F},P)$とその上に定義された確率変数$\vartheta,X_1,X_2,\ldots$で上の条件を満たすものを容易に構成する

ことができる．$X_1, ...X_n$ のもとで ϑ の正則条件付確率が存在すれば，それを $X_1, ...X_n$ のもとでの ϑ の事後分布とよぶ．一般に，Θが Borel 空間ならば，ϑ の事後分布は存在する．

本稿の目的は，P_θ が一様分布族 $U(0,\theta), \theta>0$ であるとき，事後分布の LDP を証明することである．

2．定理とその証明

$(\Theta,\mathscr{U})=((0,\infty),\mathscr{B}(0,\infty)), (\mathscr{X},\mathscr{A})=(\mathbb{R},\mathscr{B}(\mathbb{R}))$とし，統計モデル$(P_\theta)_{\theta\in\Theta}$は一様分布族$(U(0,\theta))_{\theta\in\Theta}$であると仮定する．パラメータ ϑ の事前分布は Lebesgue 測度に関して密度関数$\pi(\theta)$ をもち，さらに，πは微分可能で，$\pi>0$ であるとする．

定理2.1． ϑ の $X_1, ...X_n$ のもとで事後分布は，確率１でΘ上で LDP を満たし，そのレート関数は

$$I(\theta)=\begin{cases}\infty & 0<\theta<\vartheta,\\ \log\theta-\log\vartheta, & \vartheta\leqslant\theta\end{cases}\tag{2.1}$$

である．

注意2.2． Θ上のレート関数(2.1)は凸関数ではない．

証明．以後，$X_{(1,n)}=\min_{1\leqslant i\leqslant n}X_i$，$X_{(n,n)}=\max_{1\leqslant i\leqslant n}X_i$と定義する．$X_1, ...X_n$ のもとでυの事後分布は

$$\pi_n(d\theta)\overset{\triangle}{=}\begin{cases}\dfrac{\theta^{-n}1_{(X_{(n,n)},\infty)}(\theta)}{\int_{X_{(n,n)}}^{\infty}\theta^{-n}\pi(\theta)d\theta}\pi(\theta)d\theta, & X_{(1,n)}>0,\\ \pi(\theta)d\theta, & X_{(1,n)}\leqslant 0\end{cases}$$

によって与えられる．$X_1, X_2, \ldots$ は ϑ のもとで条件付 i.i.d. であるから，すべての n に対し，

$$P(X_{(1,n)} \leqslant 0 \mid \vartheta) = P(X_1 \leqslant 0, \ldots X_n \leqslant 0 \mid \vartheta) = P(X_1 \leqslant 0 \mid \vartheta) \cdots P(X_n \leqslant 0 \mid \vartheta) = 0$$

であり，特に，すべての n に対し，$P(X_{(1,n)} \leqslant 0) = E[P(X_{(1,n)} \leqslant 0 \mid \vartheta)] = 0$ であることに留意しておく．以後，$\omega \notin \cup_{n \geqslant 1}\{X_{(1,n)} \leqslant 0\}$ であるとする．したがって，すべての $n \geqslant 1$ に対し，

$$\pi_n(d\theta) \propto \theta^{-n} 1_{(X_{(n,n)}, \infty)}(\theta)\pi(\theta)d(\theta) \tag{2.2}$$

である．(π_n) の LDP を証明する方針は以下の通りである．まず，$1/\vartheta$ の事後分布 $P(1/\vartheta \mid X_1, \ldots X_n)$ に対し，Gärtner-Ellis の定理を適用することにより，その LDP を示し，再度，パラメータ変換をして，(π_n) の LDP を得る．

$1/\vartheta$ の事前分布は変数変換より，

$$\bar{\pi}(\theta) \overset{\triangle}{=} \frac{\pi(1/\theta)}{\theta^2}, \quad \theta \in \Theta$$

である $\bar{\pi}(\theta)$ は Θ 上で正であり，微分可能である．$U \in \mathscr{U}$ に対し，$U^{-1} = \{\theta^{-1} : \theta \in U\}$ と定義する．また，$0 < X_{(1,n)} \leqslant X_{(n,n)}$ であることにも留意する．任意の $U \in \mathscr{U}$ に対し，(2.2) と変数変換より，

$$\begin{aligned} P(1/\vartheta \in U \mid X_1, \ldots X_n) = P(\vartheta \in U^{-1} \mid X_1, \ldots X_n) \\ &\propto \int_{U^{-1}} \theta^{-n} 1_{(X_{(n,n)}, \infty)}(\theta)\pi(\theta)d\theta \\ &= \int_U (1/\theta)^{-n} 1_{(X_{(n,n)}, \infty)}(1/\theta)\frac{\pi(1/\theta)}{\theta^2}d\theta \\ &= \int_U \theta^n 1_{(0, 1/X_{(n,n)}, \infty)}(\theta)\bar{\pi}(\theta)d\theta \end{aligned}$$

であるから，$1/\vartheta$ の $X_1, \ldots X_n$ のもとので事後分布は

$$P(1/\vartheta \in d\theta \,|\, X_1, \ldots X_n = \frac{\theta^n 1_{(0,1/X_{(n,n)})}(\theta)\bar{\pi}(\theta)d\theta}{\int_0^{X_{(n,n)}} \theta^n \bar{\pi}(\theta)d\theta} \tag{2.3}$$

で与えられる．この事後分布を一旦$(\mathbb{R},\mathscr{B}(\mathbb{R}))$上の分布と見なして，Gärtner-Ellis の定理を適用する．事後分布(2.3)の積率母関数を$M_n(t)$で表すと，

$$M_n(nt) = \frac{\int_0^{1/X_{(n,n)}} e^{nt\theta}\theta^n \bar{\pi}(\theta)d\theta}{\int_0^{1/X_{(n,n)}} \theta^n \bar{\pi}(\theta)d\theta}$$

$$= \frac{\int_0^{1/X_{(n,n)}} \exp[n(\log\theta + t\theta]\bar{\pi}(\theta)d\theta}{\int_0^{1/X_{(n,n)}} \exp[n \log\theta]\bar{\pi}(\theta)d\theta} \tag{2.4}$$

である．(2.4)の分母の積分に対して，Laplace の方法（例えば，de Bruijn 1981; Small 2010 を参照せよ）を用いると，

$$\lim_n \frac{1}{n} \log \int_0^{1/X_{(n,n)}} \exp[n \log\theta]\bar{\pi}(\theta)d\theta = -\log\vartheta \tag{2.5}$$

を得る．同様に，(2.4)の分子に Laplace の方法を適用すると，

$$\lim_n \frac{1}{n} \log \int_0^{1/X_{(n,n)}} \exp[n(\log\theta + t\theta]\bar{\pi}(\theta)d\theta = \begin{cases} -\log(-t)-1, t < -\vartheta, \\ -\log\vartheta + \dfrac{t}{\vartheta}, \quad -\vartheta \leqslant t. \end{cases} \tag{2.6}$$

(2.5)，(2.6)より，

$$K(t) \stackrel{\triangle}{=} \lim_n \frac{1}{n} \log M_n(nt) = \begin{cases} \log\vartheta - \log(-t) - 1, & t < -\vartheta, \\ \dfrac{t}{\vartheta}, & -\vartheta \leqslant t. \end{cases}$$

$K(t)$の有効領域は，定義より$\mathscr{D}_K = \mathbb{R}$である．よって，0 は$\mathscr{D}_K$の内点である．かくして，Gärtner-Ellis の定理の十分条件が確認された．$K^*(t^*), t^* \in \mathbb{R}$を$K(t)$の Fenchel-Legendre 変換とすると，$\mathbb{R}$の任意の閉集合Fに対して，

$$\limsup_n \frac{1}{n} \log P(1/\vartheta \in F | X_1, \ldots X_n) \leqslant -\inf_{t^* \in F} K^*(t^*)$$

が成り立ち，$\mathbb{R}$の任意の開集合Gに対して，

$$\liminf_{n} \frac{1}{n} \log P(1/\vartheta \in G | X_1, \ldots X_n) \geqslant -\inf_{t^* \in \mathscr{E}_{K^*} \cap G} K^*(t^*)$$

が成り立つ．ただし，$\mathscr{E}_{K^*}$は$K^*(t^*)$の exposed point 全体である（$\mathscr{D}_K$の内部が$\mathbb{R}$であるので，exposing hyperplane が$\mathscr{D}_K$の内部に属するという条件は不要である）

Fenchel-Legendre 変換$K^*(t^*)$は，その定義

$$K^*(t^*) = \sup_{t \in \mathbb{R}} [t^* t - K(t)]$$

にしたがって，明示的に次のように求められる．

$$K^*(t^*) = \begin{cases} \infty, & t^* \leqslant 0, \\ -(\log t^* + \log \vartheta), & 0 < t^* \leqslant 1/\vartheta, \\ \infty, & 1/\vartheta < t^*. \end{cases}$$

$\mathscr{E}_{K^*} = \mathscr{D}_{K^*} = (0, 1/\vartheta] \subset \Theta$であること，さらに$\Theta$が$\mathbb{R}$の開集合であることに注意して，系1.2より，$1/\vartheta$の事後分布は$\Theta$上で，凸関数

$$K^*(\theta) = \begin{cases} -(\log \theta + \log \vartheta), & 0 < \theta \leqslant 1/\vartheta, \\ \infty, & 1/\vartheta < \theta. \end{cases}$$

をレート関数に持つ LDP を満たす．Θの閉集合Fに対し，F^{-1}もΘの閉集合であり，Θの開集合Gに対し，G^{-1}もΘの開集合である．これらに注意し，$1/\vartheta$の事後分布に対するΘ上の LDP より，以下の結果を得る：Θのすべての閉集合Fに対して，

$$\limsup_{n} \frac{1}{n} \log \pi_n(F) = \limsup_{n} \frac{1}{n} \log P(1/\vartheta \in F^{-1} | X_1, \ldots X_n) \qquad (2.7)$$
$$\leqslant -\inf_{\theta \in F^{-1}} K^*(\theta)$$

であり，Θのすべての開集合Gに対して，

$$\liminf_n \frac{1}{n} \log \pi_n(G) = \liminf_n \frac{1}{n} \log P(1/\vartheta \in G^{-1} | X_1, \ldots X_n)$$
$$\geqslant -\inf_{\theta \in G^{-1}} K^*(\theta) \tag{2.8}$$

である．Θ 上の関数 $I(\theta)$ を $I(\theta) = K^*(1/\theta)$，$\theta \in \Theta$ によって定義すると，$I(\theta)$ は(2.1)のように表される．また，$\{\theta \in \Theta : I(\theta) \leqslant M\} = \{\theta \in \Theta : K^*(\theta) \leqslant M\}^{-1}$ であるから，$\{\theta \in \Theta : I(\theta) \leqslant M\}$ はΘのコンパクト集合である．すなわち，$I(\theta)$ はΘ上のレート関数である．さらに，任意の $U \in \mathscr{B}(\Theta)$ に対して，$\inf_{\theta \in U^{-1}} K^*(\theta) = \inf_{\theta \in U} K^*(1/\theta) = \inf_{\theta \in U} I(\theta)$ であるから，(2.7)，(2.8)より，(π_n) はΘ上で，レート関数 $I(\theta)$ をもつ LDP を満たす．$\square$

参 考 文 献

de Bruijn, N. G. (1981). *Asmptotic Methods in Analysis*. Dover, New York.

Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. Springer, Berlin.

Ganesh, A. and O'Connell, N. (1999). An inverse of Sanov's theorem. *Statist. Probab. Lett.* **42**, 201-206.

Ganesh, A. J. and O'Connell, N. (2000). A large deviation principle for Dirichlet posteriors. *Bernoulli*. **6**, 1021-1034.

Macci, C. (2010). Large deviations for estimators of some threshold parameters. *Stat. Methods. Appl*. **19**, 63-77.

Macci, C. (2011). Large deviations for estimators of unknown probabilities, with applications in risk theory. *Statist. Probab. Lett*. **81**, 16-24.

Small, C. G. (2010). *Expansions and Asymptotics for Statistics*. CRC Press, Boca Raton.