数理解析研究所講究録
第167巻 1972年 71-80

極値統計量の一様漸近分布について

創価大・経　池田貞雄
香川大・経　松縄規

§1. 序

$X_{N,1} < X_{N,2} < \cdots < X_{N,N}$ を steadily increasing な一次元連続分布（pdf. $f(x)$, cdf. $F(x)$）からの random sample に基づく、大きさ N の順序統計量とする。本稿では lower n extremes $X_{N(n)} = (X_{N,1}, \cdots, X_{N,n})$ および upper m extremes $Y_{N(m)} = (X_{N,N-m+1}, \cdots, X_{N,N})$（$n$, m は N に depend してよい）の $(B)_d$ の意味での一様漸近分布を与える。特に n と m が fix される場合には，n 番最小値 $X_{N,n}$ または m 番最大値 $X_{N,N-m+1}$（n, $m \geq 1$）を適当に標準化したものが $N \to \infty$ で proper な極限分布を持つならば，それぞれは極値極限分布として可能な三つのタイプの分布関数のいずれか一つへの法則収束であるという事実は Gnedenko [1] および Smirnov [6] による周知の結果であるが，これはかなり一般的な条件下で $(B)_d$-型収束に拡張され得ることを示す。

§2.　極値の一様漸近分布

確率積分変換 $Z_{N,i} = F(X_{N,i})$, $i = 1, \cdots, N$ によって $(0,1)$ 上の一様分布からの大きさ N の random sample に基づく順序統計量 $Z_{N,1} < Z_{N,2} < \cdots < Z_{N,N}$ を考える．まず始めに，

$$(2\cdot 1)\qquad U_{(n)} = (Z_{N,1}, \cdots, Z_{N,n}),\quad V_{(m)} = (Z_{N,N-m+1}, \cdots, Z_{N,N})$$

の一様漸近分布を考察する。なおこゝで $n+m \leqq N$ で，n および m はそれぞれ N の増加と共に変化してよい。

ところで，確率変数系 $\{U_{(n)}, V_{(m)}\}$ は μ, λ を $0 \leqq \mu, \lambda < 1$ なる定数とし，$N \to \infty$ で $n/N \to \mu$ (resp. $n/N \to 0$) および $m/N \to 0$ (resp. $m/N \to \lambda$) ならば $(B)_d$-型漸近独立である (Ikeda & Matsunawa [4]) から，$U_{(n)}$, $V_{(m)}$ の漸近分布を別々に考えればよい。

$U_{(n)}$, $V_{(m)}$ の pdf. はそれぞれ

$$(2\cdot 2)\qquad g_N(u_{(n)}) = \frac{N!}{(N-n)!}(1-u_n)^{N-n},\quad (0 < u_1 < \cdots < u_n < 1),$$

および

$$(2\cdot 3)\qquad h_N(v_{(m)}) = \frac{N!}{(N-m)!}v_1^{N-m},\qquad (0 < v_1 < \cdots < v_m < 1)$$

で与えられる．こゝに $u_{(n)} = (u_1, \cdots, u_n)$, $v_{(m)} = (v_1, \cdots, v_m)$.

さて，$\tilde{U}_{(n)}$, $\tilde{V}_{(m)}$ を pdf. がそれぞれ

$$(2\cdot 4)\qquad \tilde{g}_N(u_{(n)}) = \frac{N^n}{\zeta_{N,n}} e^{-Nu_n},\quad (0 < u_1 < \cdots < u_n < 1),$$

および

$$(2\cdot5)\qquad \tilde{h}_N(v_{(m)}) = \frac{N^m}{\zeta_{N,m}}\cdot e^{-N(1-v_1)},\quad (0<v_1<\cdots<v_m<1)$$

で与えられる確率ベクトルとする．こゝに，

$$(2\cdot6)\qquad \zeta_{N,k} = \frac{1}{\Gamma(k)}\int_0^N x^{k-1}e^{-x}dx,\quad (k=n,m).$$

$\zeta_{N,k}$ に関して

<u>補題 2·1</u>　$k/N\to0$ ならば $\zeta_{N,k}\to1$, $(N\to\infty)$,

が成立する．また，(2·2)〜(2·5) からK-L情報量が容易に計算されて，次の結果を得る．

<u>定理 2·1</u>

(i)　$n/N\to0$ ならば $U_{(n)}\sim\tilde{U}_{(n)}\ (\mathbb{B})_d$, $(N\to\infty)$,

(ii)　$m/N\to0$ ならば $V_{(m)}\sim\tilde{V}_{(m)}\ (\mathbb{B})_d$, $(N\to\infty)$,

(iii)　$(n+m)/N\to0$ ならば

$$(U_{(n)}, V_{(m)}) \sim (\tilde{U}_{(n)})(\tilde{V}_{(m)})\ (\mathbb{B})_d,\ (N\to\infty),$$

こゝに，$(\tilde{U}_{(n)})(\tilde{V}_{(m)})$ は $(n+m)$-次元の確率ベクトルでそのpdf. は次式で与えられる：

$$(2\cdot7)\qquad \tilde{p}_N(u_{(n)}, v_{(m)}) = \tilde{g}_N(u_{(n)})\,\tilde{h}_N(v_{(m)}).$$

Ikeda [2] から直ちに次が従う．

<u>系 2·1</u>　各Nに対して $\varphi_N(x_{(n+m)})$ を $(n+m)$-次元実数空間上の可測関数とする．$(n+m)/N\to0$ ならば

$$(2\cdot8)\qquad \varphi_N((U_{(n)}, V_{(m)})) \sim \varphi_N((\tilde{U}_{(n)})(\tilde{V}_{(m)}))\ (\mathbb{B})_d,\ (N\to\infty).$$

次に上述の諸結果を元の分布の場合へ拡張する事を考えよう。lower n extremes および upper m extremes をそれぞれ

(2・9) $\quad X_{(n)}=(X_{N,1},\cdots,X_{N,n}),\quad Y_{(m)}=(X_{N,N-m+1},\cdots,X_{N,N})$

と置く。$F(x)$ が steadily increasing であるから逆関数 $F^{-1}(z)$ が存在して，

$$X_{(n)}=F^{-1}(U_{(n)})=(F^{-1}(Z_{N,1}),\cdots,F^{-1}(Z_{N,n}))$$

であり，系2・1より，もし $n/N\to 0$ ならば

(2・10) $\quad X_{(n)}\sim F^{-1}(\tilde{U}_{(n)})\ (=\tilde{X}_{(n)},\ \text{say}),\ (\mathbb{B})_d\quad (N\to\infty).$

同様に $m/N\to 0$ ならば

(2・11) $\quad Y_{(m)}\sim F^{-1}(\tilde{V}_{(m)})\ (=\tilde{Y}_{(m)},\ \text{say}),\ (\mathbb{B})_d\quad (N\to\infty)$

となる。$\tilde{X}_{(n)}$，$\tilde{Y}_{(m)}$ の pdf. はそれぞれ

(2・12) $\quad \tilde{g}_N^*(x_{(n)})=\dfrac{N^n}{\zeta_{N,n}}e^{-NF(x_n)}\cdot\prod_{i=1}^{n}f(x_i),\ (-\infty<x_1<\cdots<x_n<\infty),$

および

(2・13) $\quad \tilde{h}_N^*(y_{(m)})=\dfrac{N^m}{\zeta_{N,m}}e^{-N(1-F(x_1))}\cdot\prod_{j=1}^{m}f(y_j),\ (-\infty<y_1<\cdots<y_m<\infty)$

となる。こゝに $x_{(n)}=(x_1,\cdots,x_n)$，$y_{(m)}=(y_1,\cdots,y_m)$.

定理2・1および(2・10)，(2・11)から次の結果を得る：

<u>定理2・2</u>

(i) $n/N\to 0$ ならば $X_{(n)}\sim\tilde{X}_{(n)},\ (\mathbb{B})_d\ (N\to\infty)$，

(ii) $m/N\to 0$ ならば $Y_{(m)}\sim\tilde{Y}_{(m)},\ (\mathbb{B})_d\ (N\to\infty)$，

(iii) $(n+m)/N\to 0$ ならば $(X_{(n)},Y_{(m)})\sim(\tilde{X}_{(n)})(\tilde{Y}_{(m)}),\ (\mathbb{B})_d\ (N\to\infty).$

§3. 極値の他の漸近分布族

式 (2·12), (2·13) に見るように $(\tilde{X}_{(n)})(\tilde{Y}_{(m)})$ の分布は，当然の事ながら，$F(x)$ の tail-shapes に大いに関係する．この事は $F(x)$ および $f(x)$ をそれぞれの tail-shapes が類似する或る $G(x)$ および $g(x)$ で置き替えれば，何らかの条件下で，$(X_{(n)}, Y_{(m)})$ の漸近分布として $(\tilde{X}_{(n)})(\tilde{Y}_{(m)})$ とは異なる漸近分布族が存在するであろうということを示唆している．さて，

$$(a, b) = \{x : f(x) > 0\},$$

$(a < b;\ a, b$: extended real) と置き，$\tilde{X}_{(n)}$ の pdf. を改めて次の様に記す．

$$(3 \cdot 1) \qquad p_N(x_{(n)}) = \frac{N^n}{\zeta_{N,n}} e^{-NF(x_n)} \cdot \prod_{i=1}^{n} f(x_i), \quad (a < x_1 < \cdots < x_n < b).$$

更に，領域

$$E_{(n)} = \{x_{(n)} : a < x_1 < \cdots < x_n < c \leq b\}, \ c : \text{const.},$$

の上で pdf. が，(3·1) に対応する，

$$(3 \cdot 2) \qquad \tilde{p}_N(x_{(n)}) = \frac{N^n}{\zeta_{N,n}} e^{-NG(x_n)} \cdot \prod_{i=1}^{n} g(x_i), \quad (a < x_1 < \cdots < x_n < c \leq b)$$

と一致する n-次元確率ベクトル $W_{(n)N}$ を考える．但しこゝで $G(x)$ および $g(x)$ $(\geq 0, \neq 0)$ は次の条件を満す任意に与えられた可測関数とする：

$$(3\cdot3)\qquad \int_{E_{(n)}} \tilde{p}_N(x_{(n)})\,dx_{(n)} \leqq 1.$$

なお，(3・2)に於て $g(x)$ は必ずしも $G(x)$ の導関数でなくてよい．

いま，$n/N \to 0$ $(N\to\infty)$ を仮定すれば $\lambda_N \to \infty$, $n\lambda_N/N \to 0$ $(N\to\infty)$ となる $\{\lambda_N\}$ $(N=1,2,\cdots)$ が存在するから，この様な $\{\lambda_N\}$ $(N=1,2,\cdots)$ に対して

$$(3\cdot4)\qquad D(\lambda_N)=\{x_{(n)}:\ 0<F(x_1)<\cdots<F(x_n)<\frac{n}{N}\lambda_N\}$$

と定義すると，$D(\lambda_N)$ は $\tilde{X}_{(n)}$ の漸近主領域；即ち

$$(3\cdot5)\qquad \int_{D(\lambda_N)} p_N(x_{(n)})\,dx_{(n)} \to 1,\ (N\to\infty)$$

となる．そこで，もし

$$(3\cdot6)\qquad \rho(\tilde{X}_{(n)}, W_{(n)N}\,|\,D(\lambda_N)) = \int_{D(\lambda_N)} \sqrt{p_N(x_{(n)})\,\tilde{p}_N(x_{(n)})}\,dx_{(n)} \to 1,\quad (N\to\infty),$$

が成立すれば，Ikeda & Matsunawa [5] により $\tilde{X}_{(n)} \sim W_{(n)N}, (\mathbb{B})_d$ $(N\to\infty)$ が言える．以下の定理が成り立つ．

<u>定理3・1</u>　n を N に無関係に fix とする．(3・3) 及び

$$(3\cdot7)\qquad \begin{cases} \text{(i)} & |G(x)/F(x)-1| \to 0,\ (x\downarrow a), \\ \text{(ii)} & |g(x)/f(x)-1| \to 0,\ (x\downarrow a) \end{cases}$$

が満されるならば

$$(3\cdot8)\qquad X_{(n)} \sim W_{(n)N},\ (\mathbb{B})_d,\ (N\to\infty).$$

定理3·2　$n^2/N \to 0\ (N\to\infty)$ の時，(3·3) 及び

$$(3\cdot 9)\quad \begin{cases} (i) & |G(x)/F(x)-1|/F(x)\to 0,\ (x\downarrow a) \\ (ii) & |g(x)/f(x)-1|/F(x)\to 0,\ (x\downarrow a)\end{cases}$$

が満されるならば (3·8) が成立する。

upper m *extremes* に対しても同様な結果を得る。

§4.　極値統計量の極限分布族

次の様に標準化した n 番最小値

$$(4\cdot 1)\qquad (X_{N,n}-b_N)/a_N,\ (a_N>0)$$

の *proper* な極限分布族を $\mathcal{L}_n$ とする。 *Smirnov* [6] は *Gnedenko* [1] の $\mathcal{L}_1$ に関する結果を拡張して，*fix* された一般の $n\geqq 1$ に対して $\mathcal{L}_n$ が次の三つの分布からなることを示した：

$$(4\cdot 2)\quad \begin{cases} \Lambda_{n1}(x)=\begin{cases} 0 & x\leqq 0, \\ \dfrac{1}{\Gamma(n)}\displaystyle\int_0^{x^\alpha} e^{-t}t^{n-1}dt, & x>0,\ \alpha>0, \end{cases} \\ \Lambda_{n2}(x)=\begin{cases} \dfrac{1}{\Gamma(n)}\displaystyle\int_0^{|x|^{-\alpha}} e^{-t}t^{n-1}dt, & x<0,\ \alpha>0, \\ 1 & x\geqq 0, \end{cases} \\ \Lambda_{n3}(x)=\dfrac{1}{\Gamma(n)}\displaystyle\int_0^{e^x} e^{-t}t^{n-1}dt, & -\infty<x<\infty. \end{cases}$$

この結果を導くに当り，次の定理は基本的である．

定理（Smirnov）

n を fix，$X_{N,n}$ の cdf. を $\Phi_{N,n}(x)$ とし，適当に選んだ定数 a_N，b_N に対して

(4·3) $\quad \Phi_{N,n}(a_N x + b_N) \longrightarrow \Phi(x)$, $(N \to \infty)$,

（$\Phi(x)$ は proper な cdf.），が成立するための必十条件は

(4·4) $\quad v_N(x) \equiv NF(a_N x + b_N) \to v(x)$, $(N \to \infty)$,

となることである．たゞし，$v(x)$ は

(4·5) $$\frac{1}{\Gamma(n)}\int_0^{v(x)} e^{-t}\, t^{n-1}\, dt = \Phi(x)$$

で定義される，非減少かつ非負な関数である∎

定理2·2から $X_{N,n} \sim \tilde{X}_{N,n}$，$(B)_d$ $(N \to \infty)$，たゞし $\tilde{X}_{N,n}$ は pdf. が

(4·6) $$\tilde{p}_{N,n}(x) = \frac{N^n}{\zeta_{N,n}\Gamma(n)} F(x)^{n-1} \cdot e^{-NF(x)} \cdot f(x)$$

で与えられる確率変数である．よって，$(X_{N,n} - b_N)/a_N$ は $(\tilde{X}_{N,n} - b_N)/a_N$ と漸近的に $(B)_d$-同値となる．後者の cdf. は次式で与えられる．

(4·7) $$\tilde{\Phi}_{N,n}(a_N x + b_N) = \frac{1}{\zeta_{N,n}\Gamma(n)}\int_0^{NF(a_N x + b_N)} u^{n-1} e^{-u}\, du.$$

従って，$v(x)$ が (4·4) を満すならば

(4·8) $$\tilde{\Phi}_{N,n}(a_N x + b_N) \to \Phi(x), \ (N \to \infty)$$

となる。 この収束のタイプを考察しよう。

(4·2), (4·5) に見るように $v(x)$ が微分可能であるから $\Phi(x)$ の pdf. は次のようになる。

(4·9) $$\phi(x) = \frac{1}{\Gamma(n)} v(x)^{n-1} e^{-v(x)} v'(x).$$

そこで，次の事を仮定する： 或る正の定数 $c \leqq 1$ に対して

(4·10) $$\sup_{0 \leqq F(a_N x + b_N) \leqq c} \left| \frac{N a_N f(a_N x + b_N)}{v'(x)} - 1 \right| \to 0, \ (N \to \infty).$$

このとき

(4·11) $$\sup_{0 \leqq F(a_N x + b_N) \leqq c} \left| \frac{N F(a_N x + b_N)}{v(x)} - 1 \right| \to 0, \ (N \to \infty),$$

となるから，十分大きな N に対して

(4·12) $$D_N(d) \equiv \{ x : v(x) \leqq N d \} \subseteq \{ x : F(a_N x + b_N) \leqq c \}$$

となる定数 $d > 0$ が存在する。 n が N に無関係に fix されているなら，条件 (4·10) の下で $D_N(d)$ 上の affinity が

(4·13) $$\rho_N(\phi, \tilde{p}^*_{N,n} \mid D_N(d)) \to 1, \ (N \to \infty)$$

となることが示される。 こゝに $\tilde{p}^*_{N,n}$ は $(\tilde{X}_{N,n} - b_N)/a_N$ の pdf. である。 従って，(4·8) の収束は $(B)_d$ - 型であることがわかる。 以上をまとめて次の定理を得る。

定理4・1　n が N に無関係に fix とする．この時

$$(4\cdot 10) \quad \sup_{0 \leqq F(a_N x + b_N) \leqq c} \left| \frac{N a_N f(a_N x + b_N)}{v'(x)} - 1 \right| \to 0, \ (N \to \infty)$$

ならば（4・8），従って（4・3）の収束は $(B)_\alpha$-型である∎

上の結果は標準化された lower n extremes に対しても容易に拡張される．また，本節の結果と対応するものが m-番最大値，upper m extremes に対しても成立する．

REFERENCES

[1] Gnedenko, B. V. (1943). Sur la distribution du terme maximum d'une série aléatoire. Ann. Math. 44 423-453.

[2] Ikeda, S, (1963). Asymptotic equivalence of probability distributions with applications to some problems of asymptotic independence. Ann. Inst. Statist. Math. 15 87-116.

[3] Ikeda, S. (1968). Asymptotic equivalence of real probability distributions. Ann. Inst. Statist. Math. 20 339-362.

[4] Ikeda, S. and Matsunawa, T. (1970). On asymptotic independence of order statistics. Ann. Inst. Statist. Math. 22 435-449.

[5] Ikeda, S. and Matsunawa, T. (1972). On the uniform asymptotic joint normality of sample quantiles. Ann. Inst. Statist. Math. 24 33-52.

[6] Smirnov, N. V. (1952). Limit distributions for the terms of a variational series. Amer. Math. Soc. Trans. No. 67.