統計数理（2009）
第 57 巻 第 2 号 443–447

［研究ノート］

観測ノイズが極値分布に影響を与えないための十分条件

西山 陽一[†]・志村 隆彰[†]

（受付 2009 年 4 月 23 日；改訂 6 月 18 日；採択 6 月 24 日）

要　　旨

従来の極値理論は，確率変数列 $\{X_i\}$ が厳密に観測できるという理想的な想定のもとで研究がなされてきた．しかし，現実のデータは常に観測誤差が入る危険性にさらされている．この論文では，確率変数列 $\{X_i\}$ に対する極値分布と観測ノイズ e_i の入ったデータ $\{X_i+e_i\}$ に対する極値分布が同一になるための十分条件を与える．条件はノイズ e_i の Orlicz ノルムを用いて記述される．我々の結果により，観測誤差が入っている状況において，理想的な想定における極値理論をそのまま適用してよいための条件がはっきりする．

キーワード： 極値統計量，観測ノイズ，Orlicz ノルム．

1. 問題提起

$\{X_i\}$ は任意の実数値確率変数列であるとする．ある定数列 a_n, b_n および非退化分布 G が存在して，極値統計量 $M_n = \max_{1\le i\le n} X_i$ に対して

$$\frac{M_n - b_n}{a_n} \to^d G \tag{1.1}$$

が成立していることを仮定する．独立同一分布列の場合を念頭にこういう仮定をおいたが，従属確率変数列でもかまわないとする．この論文では $a_n \to \infty$ である状況のみを考える．独立同一分布列の場合には，この仮定は分布が long tailed であること，すなわち任意の実数 k に対して $x\to\infty$ とするとき $P(X>x+k) \sim P(X>x)$ が満たされることと同値であり，そのためには分布が subexponential であれば十分である（志村, 2009）．

この論文では，実際にはノイズの入ったデータ $\tilde{X}_i = X_i + e_i$ が観測されることを想定する．ただし $\{e_i\}$ は任意の実数値確率変数列であるとする（独立同一分布でなくてもよいし，$\{X_i\}$ に依存してもよいとする）．このデータに対する極値統計量 $\tilde{M}_n = \max_{1\le i\le n} \tilde{X}_i$ に対して上と同じ a_n, b_n, G について

$$\frac{\tilde{M}_n - b_n}{a_n} \to^d G \tag{1.2}$$

が成立するための条件を考察する．

ここで，この問題を考える動機をふたつ述べる．ひとつ目は，現実のデータ解析ではつねにノイズが入る危険性があるため，極値理論を使ってよいための条件をはっきりさせる必要を感じることである．すなわち，例えば X_i が河川の水位を表す確率変数である場合，X_i の値が厳

[†] 統計数理研究所：〒190–8562 東京都立川市緑町 10–3

密に観測できるということは現実的に考えられず，つねに丸め誤差や測定器による誤差が入っている．従来においては X_i が厳密に測定できるという仮想的な前提のもとで極値理論が適用されてきたが，我々の結果により，定数列 a_n のオーダーに応じて，許される誤差の範囲が明確に記述できて，観測ノイズをもつデータに対する極値理論(より具体的には(1.1)や(1.2)の形をした漸近理論およびそれにもとづく統計的手続き)の適用が正当化される．

ふたつ目は数学的興味である．極値理論は，歴史的には $\{X_i\}$ が独立に同一分布 F に従う場合に研究されてきたが，近年，($\tilde{X}_i = X_i + e_i$ の形とは限らない)定常時系列 $\{\tilde{X}_i\}$ であってその定常分布が F である場合について，独立同一分布の場合に帰着させるというアプローチで研究がなされている．その中で，$a_n^{-1}(M_n - b_n) \to^d G$ かつ $a_n^{-1}(\tilde{M}_n - b_n) \to^d H$ であったとしても $G = H$ とは限らず，extremal index と呼ばれるある定数 $\theta \in [0,1]$ が存在して $G = H^\theta$ となることもしばしばあることが知られている(Embrechts et al., 1997, の Theorem 4.4.1 およびそれに対する Remark を見よ)．一方 $G = H$ となる場合もある(Embrechts et al., 1997, の Theorem 4.4.6 を見よ)．こういったアプローチは，時系列 $\{\tilde{X}_i\}$ が定常であるという大きな前提がある上に，例えば Embrechts et al. (1997)に書いてある「漸近的独立条件 $D(u_n)$」(211 頁)や「anti-clustering 条件 $D'(u_n)$」(213 頁)を満たさなければならないという制約があった．これに対し，我々は定常性を仮定せず，$\tilde{X}_i = X_i + e_i$ という関係と e_i に対する簡明なモーメント条件だけから $G = H$ となるための十分条件を提示する．

2. 主定理

主定理を述べるために，Orlicz ノルム $||Z||_\psi$ を定義しよう．ψ は恒等的にゼロではない下に凸な非減少関数であって $\psi(0) = 0$ かつある定数 c に対して $\limsup_{x,y\to\infty} \psi(x)\psi(y)/\psi(cxy) < \infty$ が成り立つものとする．Z は実数値確率変数であるとする．このとき，

$$||Z||_\psi = \inf\{C > 0 :\ E\psi(|Z|/C) \le 1\}$$

と定義する．特に $\psi(z) = z^p$ のときには L_p ノルムになる．その他の Orlicz ノルムの重要な例としては $\psi_p(z) = e^{z^p} - 1$ (ただし $p \ge 1$)の場合がある．このとき $\psi_p^{-1}(z) = (\log(1+z))^{1/p}$ であり，また

$$P(|Z| > x) \le K e^{-Cx^p}, \quad \forall x \ge 0$$

ならば $||Z||_{\psi_p} \le ((1+K)/C)^{1/p}$ である(例えば van der Vaart and Wellner, 1996, の Lemma 2.2.1 を見よ)．一般に $E|Z| \le \psi^{-1}(1)||Z||_\psi$ が成り立つことが定義と Jensen の不等式から容易にわかる．

ここで主定理を述べる．

定理 1. $\psi^{-1}(n)/a_n \to 0$ を仮定する．もし $\sup_i ||e_i||_\psi < \infty$ ならば, (1.2)は成立する．

証明. まず最大値 $\max_{1\le i\le n} X_i$ を達成する i を $i_0 = i_0(\omega)$ (実現 ω に依存)と記すと

$$\max_{1\le i\le n} X_i - \max_{1\le i\le n} |e_i| \le X_{i_0} + e_{i_0} \le \max_{1\le i\le n}\{X_i + e_i\} \le \max_{1\le i\le n} X_i + \max_{1\le i\le n} |e_i|$$

であるから，$|M_n - \tilde{M}_n| \le \max_{1\le i\le n} |e_i|$ である．よって $\max_{1\le i\le n} |e_i|/a_n \to^p 0$ を示せば十分である．任意の $\varepsilon > 0$ に対し

$$\begin{aligned} P\left(\frac{\max_{1\le i\le n}|e_i|}{a_n} > \varepsilon\right) &\le \frac{E\max_{1\le i\le n}|e_i|}{a_n\varepsilon} \\ &\le \frac{\psi^{-1}(1)||\max_{1\le i\le n}|e_i|||_\psi}{a_n\varepsilon} \end{aligned}$$

である．ここで van der Vaart and Wellner (1996) の Lemma 2.2.2 より，ψ にのみ依存する定数 K が存在して，右辺は

$$\frac{\psi^{-1}(1)K\psi^{-1}(n)\max_{1\le i\le n}||e_i||_\psi}{a_n\varepsilon}$$

で押さえられ，これは仮定より $n\to\infty$ とするときゼロに収束する．証明は完了した．□

上記の結果は，その証明が容易であるがゆえに，既に知られているのではと思う読者もいるかもしれない．しかし，既存の文献にはこの主張は見あたらない．極値理論に Orlicz ノルムを結びつける発想は新しく，また鍵となった van der Vaart and Wellner (1996) の Lemma 2.2.2 もその動機であった経験過程の理論への応用以外にはまだ十分に認知されていないように見える．我々の結果は今後発刊される極値理論の教科書に書かれ得るような簡明さと重要性をあわせもっていると確信する．

3. 例

例をいくつか述べる．

命題 2. $a_n\sim n^{1/\alpha}$ (ただし $\alpha>0$) とする．$p>(\alpha\vee 1)$ とし，$\sup_i E|e_i|^p<\infty$ を仮定すれば，(1.2) は成立する．

証明. 定理 1 において $\psi(z)=z^p$ と取ればよい．□

$\{X_i\}$ が独立同一分布に従う場合の命題 2 の例としては Cauchy 分布 ($\alpha=1$) および Pareto 分布 ($\alpha>0$ が一般) があり，これらは極値分布 G が Fréchet 分布となる代表例である．後者は α が大きいほど裾が軽いが，その場合 p を大きく取らなければならない (つまり $\{e_i\}$ の裾を軽くしなければならない)．一般に，極値分布 G が Fréchet 分布となる場合には a_n は $n^{1/\alpha}$ と緩慢変動関数の積になることが知られている (Embrechts et al., 1997, の 131 頁)．この α に対応して $p>(\alpha\vee 1)$ を選ぶと上述の命題の主張がそのまま成り立つ．
次の例はより裾が軽い場合を扱うものである．

命題 3. ある $p\ge 1$ に対して $(\log n)^{1/p}/a_n\to 0$ とする．ある $K,C>0$ が存在して全ての i に対し

$$P(|e_i|>x)\le Ke^{-Cx^p},\quad \forall x\ge 0$$

を仮定すれば，(1.2) は成立する．

証明. 定理 1 において $\psi(z)=\psi_p(z)=e^{z^p}-1$ と取り，$||e_i||_{\psi_p}\le((1+K)/C)^{1/p}$ に注意すればよい．□

$\{X_i\}$ が独立同一分布に従う場合の命題 3 の例としては Weibull-like 分布 (の一部), Benktander-type-I 分布，Benktander-type-II 分布があり，これらの場合，極値分布 G は Gumbel 分布となる (Embrechts et al., 1997, の 153–157 頁を見よ)．極値分布が Gumbel 分布となるための特徴付けおよびそのときの a_n,b_n の決まり方については Embrechts et al. (1997) の Theorem 3.3.26 を参照のこと．

4. 結論

我々の結果は，得られた a_n に対応して関数 ψ を $\psi^{-1}(n)/a_n\to 0$ となるようにとると，それに関するノイズの Orlicz ノルムが有界であることが (1.2) を得るためには十分であることを

主張している．我々の結果はもとの $\{X_i\}$ が独立同一分布列でなくても適用可能である．例えば，定常列の極値分布に関する一般論が前述の Embrechts et al. (1997) の 4.4 節にあるし，また今後もさまざまな従属列に対する極値理論が確立されていくことが予想されるが，我々の結果は $\{X_i\}$ が極値理論の枠組みに入り，かつ $a_n \to \infty$ である限りにおいて，ノイズが付加したデータを扱うことを常に可能にするものである．

参 考 文 献

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events*, Springer-Verlag, Berlin Heidelberg.

志村隆彰 (2009). 最大値吸引領域の吸引係数について，極値理論の工学への応用 (6)，統計数理研究所共同研究リポート，No. 224, 79–84.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*: *With Applications to Statistics*, Springer-Verlag, New York.

A Sufficient Condition for Observation Noise Not to Affect the Extreme Value Distribution

Yoichi Nishiyama and Takaaki Shimura

The Institute of Statistical Mathematics

In the usual context of extreme value theory, it is assumed that random variables $\{X_i\}$ can be observed exactly. On the other hand, real data are always exposed at risk of being perturbed by observation noise. We present a sufficient condition under which the extreme value distribution of a sequence $\{X_i\}$ is the same as that of data $\{X_i + e_i\}$ with observation noise e_i. The condition is described in terms of the Orlicz norm. Our result gives a condition under which one may safely believe the extreme value theory which has been studied under the ideal situation that the data are exactly observed.

Key words: Extreme value statistics, observation noise, Orlicz norm.