情報論的学習理論テクニカルレポート 2009
Technical Report on Information-Based Induction Sciences 2009 (IBIS2009)

極値統計学*
Statistics of Extremes

高橋倫也[†]
Rinya Takahashi

Abstract: Statistics of extremes deals with disastrous rare events like heavy rains, strong winds, earthquakes, etc. For planning embankments and other constructions, or to sell the insurance to cover the loss, the assessment of the possible severeness is essential. Frequently, the possibility of the events, which people have never experienced, must be estimated, that is we have to extrapolate statistically available data. A standard procedure is to make use of the two types of basic distributions, the *generalized extreme value distributions* (GEV), and the *generalized Pareto distributions* (GP). According to the available data one of two distributions is adopted, the GEV for block maxima and the GP for threshold exceedances.

Keywords: 極値理論，一般極値分布，一般パレート分布，再現レベル，バリューアットリスク

1　まえがき

数理統計学の入門では，データに適合させる分布としてまず正規分布が出てくる．正規分布がデータに適合可能であることは中心極限定理によって説明される．正規分布の平均と分散は標本平均と不偏分散で推定され，平均や分散に回帰構造が仮定される場合はパラメータは最尤法で推定される．一方，統計的信頼性理論では寿命データにワイブル分布が適合される．ワイブル分布がデータに適合できることは最弱リンクモデルにより説明される．

極値統計学の目的は，長期間または広領域等における最大値（または最小値）に関する推測である．最大値に関する情報を有しているデータとして，標本を同じ大きさのブロックに分けたときの各ブロックごとの最大値データ（これを**区分最大値データ**（block maxima) とよぶ)，または標本の中で十分大きな閾値を超えるデータ（これを**水準超過データ** (threshold exceedances or Peaks Over Threshold) とよぶ）の 2 つのタイプが考えられる．極値理論により，区分最大値データには一般極値分布を，また水準超過データには一般パレート分布を適合させる．通常パラメータは最尤法で推定する．一般に，極値統計学で興味があるのは，一般極値分布や一般パレート分布のパラメータではなく，それらの分布の上側微小確率点（再現レベルとバリューアットリスクとよばれるもの）である．

以下 2 節では，2 つのタイプの極値データに適合させる 2 つのタイプの分布について紹介する．3 節ではデータに一般極値分布と一般パレート分布を適合させる場合の推測法について述べる．なお，実データ解析や詳しい文献，ソフトの紹介はスライドで行う．

2　極値理論

分布関数 F を持つ確率変数 X を考える．分布 F からの確率標本を $X_1, X_2, \ldots, X_n$ とし，極値統計量を

$$Z_n = \max\{X_1, X_2, \ldots, X_n\}$$

とする．このとき分布 F の上限 (upper boundary) を $x_F = \sup\{x : F(x) < 1\} \leq \infty$ とすると，$Z_n \to x_F$ $(n \to \infty)$ となる．そこで，Z_n を規準化した場合の極限分布を考える．分布 F に依存する定数列 $a_n > 0$，$b_n \in \mathbb{R}$ $(n = 1, 2, \ldots)$ が存在して，非退化な分布 G に対して

$$\begin{aligned} P\left(\frac{Z_n - b_n}{a_n} \leq x\right) &= P(Z_n \leq a_n x + b_n) \\ &= F^n(a_n x + b_n) \xrightarrow{\mathcal{L}} G(x) \end{aligned}$$

となるとき，G を**極値分布** (exreme value distribution, EVD) とよぶ．これを記号で $F \in \mathcal{D}(G)$ と書き，F は

---
*企画セッション「金融リスクと統計的学習」

†神戸大学　海事科学部，〒 658-0022 神戸市東灘区深江南町 5-1-1, tel. 078-431-6267, e-mail r-taka@maritime.kobe-u.ac.jp,
Faculty of Maritime Sciences, Kobe University, 5-1-1, Fukae-Minami-Machi, Higashi-Nada-ku, Kobe, 658-0022.

G の**吸引領域に属する**という．また，a_n，b_n を基準化係数という．

よく知られている分布の場合の極値分布について考える．分布関数を F，その密度関数を f とする．

例１． 標準指数分布 Exp(1)

$$F(x) = 1 - e^{-x}, \quad f(x) = e^{-x}, \ x \geq 0.$$

$$\begin{aligned} F^n(x + \log n) &= \left\{1 - e^{-(x+\log n)}\right\}^n \\ &= \left\{1 + \frac{-e^{-x}}{n}\right\}^n \\ &\to e^{-e^{-x}}, \ n \to \infty. \\ &\qquad (-\infty < x < \infty). \end{aligned}$$

$a_n = 1$，$b_n = \log n$ である．

例２． Pareto 分布　$(\forall\, \alpha > 0)$

$$F(x) = 1 - 1/x^\alpha, \quad f(x) = \alpha x^{-\alpha-1}, \ x \geq 1.$$

$$\begin{aligned} F^n(n^{1/\alpha}x) &= \left\{1 - \frac{1}{(n^{1/\alpha}x)^\alpha}\right\}^n = \left\{1 + \frac{-x^{-\alpha}}{n}\right\}^n \\ &\to e^{-x^{-\alpha}}, \ n \to \infty. \quad (x > 0). \end{aligned}$$

$a_n = n^{1/\alpha}$，$b_n = 0$ である．

例３． ベータ分布　$(\forall\, \alpha > 0)$

$$F(x) = 1-(1-x)^\alpha, \quad f(x) = \alpha(1-x)^{\alpha-1}, \ 0 \leq x \leq 1.$$

$$\begin{aligned} F^n(n^{-1/\alpha}x + 1) &= \left\{1 - (-n^{-1/\alpha}x)^\alpha\right\}^n \\ &= \left\{1 + \frac{-(-x)^\alpha}{n}\right\}^n \\ &\to e^{-(-x)^\alpha}, \ n \to \infty. \quad (x \leq 0). \end{aligned}$$

$a_n = n^{-1/\alpha}$，$b_n = 1 = x_F$ である．

極値分布は（位置と尺度を別として）上の例の 3 つの型になる事が知られている．

定理１． [Fréchet (1927), Fisher and Tippett (1928), Gnedenko (1943);“ Trinity Theorem”]　極値分布 $G(x)$ は次の 3 つの型に限る：

$$\begin{aligned} \Lambda(x) &= \exp(-\exp(-x)), & x \in \mathbb{R}, \\ \Phi_\alpha(x) &= \exp(-x^{-\alpha}), & x \geq 0, \ \alpha > 0, \\ \Psi_\alpha(x) &= \exp(-(-x)^\alpha), & x \leq 0, \ \alpha > 0. \end{aligned}$$

分布 Λ，Φ_α，Ψ_α はそれぞれ Gumbel，Fréchet,（負の）Weibull 分布とよばれる．

極値分布の吸引領域に属するための必要十分条件は次のようになる：

定理２． [Gnedenko (1943), de Haan (1970)]

$F \in \mathcal{D}(\Phi_\alpha) \iff x_F = \infty$ and

$$\lim_{x \to \infty} \frac{1 - F(tx)}{1 - F(x)} = t^{-\alpha}, \ \forall\, t > 0.$$

$F \in \mathcal{D}(\Psi_\alpha) \iff x_F < \infty$ and

$$\lim_{x \uparrow x_F} \frac{1 - F(x_F - (x_F - x)t)}{1 - F(x)} = t^\alpha, \ \forall\, t > 0.$$

$F \in \mathcal{D}(\Lambda) \iff \exists\, s(\cdot) > 0$ s.t.

$$(*) \quad \lim_{x \uparrow x_F} \frac{1 - F(x + ts(x))}{1 - F(x)} = e^{-t}.$$

$$(*) \implies \int_x^{x_F} (1 - F(y))dy < \infty, \quad x < x_F.$$

$$\implies s(x) = \int_x^{x_F} \frac{1 - F(y)}{1 - F(x)} dy \quad \text{satisfies } (*).$$

定理 1 の 3 つの極値分布は次の 1 つのパラメータ ($\xi \in \mathbb{R}$) を持つ分布で表される．これを一般極値 (generalized extreme value, GEV) 分布とよぶ．

$$G_\xi(x) = \exp\left\{-(1 + \xi x)^{-1/\xi}\right\}, \quad 1 + \xi x > 0,$$

ここで，$\xi = 0$ の場合は

$$G_0(x) = \lim_{\xi \to 0} G_\xi(x) = \exp\{-\exp(-x)\} = \Lambda(x)$$

とする．また，

$$\Phi_\alpha(x) = G_{1/\alpha}(\alpha(x-1)), \quad \Psi_\alpha(x) = G_{-1/\alpha}(\alpha(x+1))$$

である．

一般極値分布は**最大値安定** (max–stable) **性**という性質を持つ：任意の n に対して $A_n > 0$ と $B_n \in \mathbb{R}$ が存在して

$$G_\xi^n(A_n x + B_n) = G_\xi(x)$$

となる．この関数方程式の解が定理 1 の 3 つの型の極値分布になる．

一般極値分布を用いると，

$$F^n(a_n x + b_n) \to G_\xi(x).$$

よって n が十分大のとき，$a_n x + b_n = z$ とおくと

$$P(Z_n \leq z) = F^n(z) \approx G_\xi\left(\frac{z - b_n}{a_n}\right).$$

ここで「G_ξ による近似が十分であるためには n がどれくらい大きければよいか？」が重要な問題になる．

一方，

$$\begin{aligned} F^n(a_nx+b_n) &= \left[1-\{1-F(a_nx+b_n)\}\right]^n \\ &= \left[1-n\{1-F(a_nx+b_n)\}/n\right]^n \\ &\to \exp\left[-\lim_{n\to\infty} n\{1-F(a_nx+b_n)\}\right] \end{aligned}$$

から，

$$\lim_{n\to\infty} n\{1-F(a_nx+b_n)\} = -\log G_\xi(x) = (1+\xi x)^{-1/\xi}.$$

ここで，$x=0$ とおくと

$$\lim_{n\to\infty} n\{1-F(b_n)\} = 1.$$

上の 2 式の比をとると，

$$\lim_{n\to\infty} \frac{n\{1-F(a_nx+b_n)\}}{n\{1-F(b_n)\}} = (1+\xi x)^{-1/\xi}.$$

よって n が十分大のとき，

$$\begin{aligned} P(X>a_nx+b_n \,|\, X>b_n) &= \frac{1-F(a_nx+b_n)}{1-F(b_n)} \\ &\approx (1+\xi x)^{-1/\xi}. \end{aligned}$$

すなわち，

$$P(X-b_n \le a_nx \,|\, X>b_n) \approx 1-(1+\xi x)^{-1/\xi}.$$

この右辺の分布

$$H_\xi(x) = 1+\log G_\xi(x) = 1-(1+\xi x)^{-1/\xi}$$

を一般パレート (Generalized Pareto, GP) 分布とよぶ．

よって b_n が十分大のとき

$$P(X-b_n \le y \,|\, X>b_n) \approx H_\xi\left(\frac{y}{a_n}\right).$$

ここでも「H_ξ による近似が十分であるためには b_n がどれくらい大きければよいか？」が重要な問題になる．

分布 F の**条件付き超過分布関数** F_u を

$$F_u(y) = P(X-u \le y \,|\, X>u), \quad 0 \le y \le x_F - u$$

とする．このとき上で述べたことは次のようになる：

定理３． [Pickands (1975)]

$$F \in \mathcal{D}(G_\xi) \iff \lim_{u\to x_F} F_u(a(u)y) = H_\xi(y), \quad \forall y \ge 0,\ F_u(a(u)y) < 1.$$

ただし，$a(\cdot)$ は a_n を連続化した適当な正の関数．

定理 2 より，裾の重たい分布の t 分布や Pareto 分布等は Fréchet 分布 $(\xi>0)$ の吸引領域に属する．裾の確率が指数的に減少する正規分布，指数分布，ガンマ分布等は Gumbel 分布 $(\xi=0)$ の吸引領域に属する．そして，上限のあるベータ分布は Weibull 分布 $(\xi<0)$ の吸引領域に属する．また定理 3 より，上で述べた分布の十分大きな値での条件付き超過分布は一般パレート分布 H_ξ $(\xi>0,\ \xi=0,\ \xi<0)$ で近似出来る．すなわち，一般極値分布と一般パレート分布は，$\xi>0$ の場合は裾が重たい分布に，$\xi=0$ の場合は裾が軽い分布に，そして $\xi<0$ の場合は上限のある分布に対応する．

最小に関しては，

$$\min\{X_1, X_2, \ldots, X_n\} = -\max\{-X_1, -X_2, \ldots, -X_n\}$$

の関係式を利用すれば同様の結果が得られる．

3 推測

極値統計学の目的は，未知の母集団分布の右裾（または左裾）の推測である．極値データに対して統計的推測を行う準備として，まず一般極値分布族と一般パレート分布族を考える．

定義１． 次の分布を**一般極値分布**といい GEV(μ, σ, ξ) で表す．

$$G(z) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} = G_\xi\left(\frac{z-\mu}{\sigma}\right),$$

$$1+\xi(z-\mu)/\sigma > 0.$$

ただし，G_ξ は標準一般極値分布 GEV(0, 1, ξ) の分布関数．$\mu \in \mathbb{R}$ は位置，$\sigma>0$ は尺度，$\xi \in \mathbb{R}$ は形状パラメータである．

$G(z)$ は，$\xi<0$ のときは Weibull 分布で $z<\mu-\sigma/\xi$, $\xi=0$ のときは Gumbel 分布で $-\infty<z<\infty$，$\xi>0$ のときは Fréchet 分布で $z>\mu-\sigma/\xi$.

標準一般極値分布 $G_\xi(z)$ の密度関数は，$\xi \ne 0$ のとき

$$\begin{aligned} g_\xi(z) &= (1+\xi\, z)^{-1/\xi-1} \exp\left\{-(1+\xi\, z)^{-1/\xi}\right\}, \\ &\qquad 1+\xi z>0, \end{aligned}$$

で，$\xi=0$ のとき

$$\begin{aligned} g_0(z) &= \exp\left\{-z-\exp(-z)\right\}, \\ &\qquad -\infty<z<\infty, \end{aligned}$$

となる．**図１**は一般極値分布 GEV(−2.5, 1, −0.4)（上限 0），GEV(0, 1, 0)，そして GEV(2.5, 1, 0.4)（下限 0）の密度関数である．

定義２． 次の分布を**一般パレート分布**といい GP(σ, ξ) で表す．

$$H(y) = 1 - \left(1 + \xi\,\frac{y}{\sigma}\right)^{-1/\xi} = H_\xi\left(\frac{y}{\sigma}\right), \quad 1 + \xi y/\sigma > 0.$$

ただし，H_ξ は標準一般パレート分布 GP$(1, \xi)$ の分布関数．$\sigma > 0$ は尺度，$\xi \in \mathbb{R}$ は形状パラメータである．

$H(y)$ は，$\xi < 0$ のときはベータ分布で $0 < y < -\sigma/\xi$，$\xi = 0$ のときは $H_0(y/\sigma) = \lim_{\xi \to 0} H_\xi(y/\sigma) = 1 - e^{-y/\sigma}$ より指数分布で $0 < y < \infty$，$\xi > 0$ のときは Pareto 分布で $0 < y < \infty$．

標準一般パレート分布 $H_\xi(y)$ の密度関数は，

$$h_\xi(y) = \begin{cases} (1 + \xi\, y)^{-1/\xi - 1}, & 1 + \xi y > 0, \quad \xi \neq 0, \\ \exp(-y), & 0 < y < \infty, \quad \xi = 0, \end{cases}$$

となる．**図２**は，標準一般パレート分布 GP$(1, \xi)$，$\xi = -0.4,\ 0,\ 0.4$ の密度関数である．

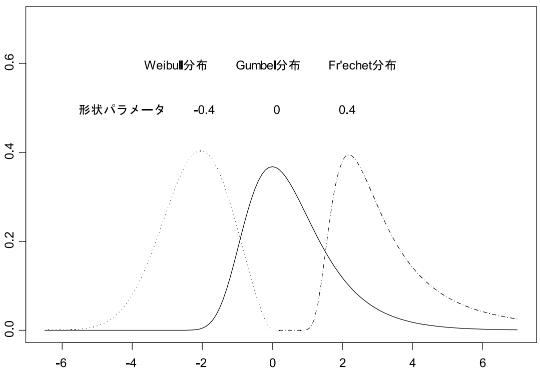

図 1: 一般極値分布 GEV$(-2.5, 1, -0.4)$（上限 0），GEV$(0, 1, 0)$，GEV$(2.5, 1, 0.4)$（下限 0）の密度関数.

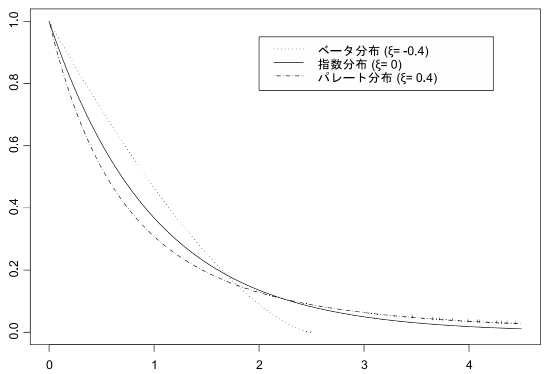

図 2: 一般パレート分布 GP$(1, \xi)$，$\xi = -0.4,\ 0,\ 0.4$ の密度関数.

3.1 一般極値（GEV) モデル

区分最大値データ $\{z_1, z_2, \ldots, z_n\}$ に一般極値分布 GEV(μ, σ, ξ) を適合させる場合の推測について述べる．ここでは，母集団分布は一般極値分布の吸引領域に属していると仮定し，最大値データは一般極値分布に従うと考えている．しかし,「吸引領域に属する」はデータが存在しない領域に関する仮定で，それをデータから検証することは出来ない．同様の事が一般パレート分布の場合にも言える．

最尤法

一般極値分布 GEV(μ, σ, ξ) の対数尤度は，$\xi \neq 0$ のとき

$$\begin{aligned} l(\mu, \sigma, \xi) &= -n \log \sigma - (1 + 1/\xi) \sum_{i=1}^{n} \log\left[1 + \xi\left(\frac{z_i - \mu}{\sigma}\right)\right] \\ &\quad - \sum_{i=1}^{n}\left[1 + \xi\left(\frac{z_i - \mu}{\sigma}\right)\right]^{-1/\xi}, \\ &\qquad 1 + \xi(z_i - \mu)/\sigma > 0, \quad i = 1, \ldots, n, \end{aligned}$$

で，$\xi = 0$ のとき

$$l(\mu, \sigma) = -n \log \sigma - \sum_{i=1}^{n}\left(\frac{z_i - \mu}{\sigma}\right) - \sum_{i=1}^{n} \exp\left\{-\left(\frac{z_i - \mu}{\sigma}\right)\right\},$$

となる．これらの対数尤度を最大にする最尤推定値 $(\widehat{\mu}, \widehat{\sigma}, \widehat{\xi})$ または $(\widehat{\mu}, \widehat{\sigma})$ を求める．

情報行列は次のようになる (Prescott and Walden, 1980)：

$$\frac{n}{\sigma^2 \xi^2}\begin{bmatrix} e^*_{11} & e^*_{12} & e^*_{13} \\ \prime\prime & e^*_{22} & e^*_{23} \\ \prime\prime & \prime\prime & e^*_{33} \end{bmatrix},$$

$$\begin{aligned} &e^*_{11} = \xi^2 p, \quad e^*_{12} = \xi\{\Gamma(2 + \xi) - p\}, \\ &e^*_{13} = \sigma\xi(p/\xi - q), \quad e^*_{22} = 1 - 2\,\Gamma(2 + \xi) + p, \\ &e^*_{23} = \sigma\left[\frac{\Gamma(2 + \xi) - 1}{\xi} + q - \frac{p}{\xi} - 1 + \gamma\right], \\ &e^*_{33} = \sigma^2\left[\frac{\pi^2}{6} + \left(1 - \gamma + \frac{1}{\xi}\right)^2 - \frac{2q}{\xi} + \frac{p}{\xi^2}\right]. \end{aligned}$$

ただし (μ, σ, ξ) の順で，$\Gamma(\cdot)$ はガンマ関数，$\Psi(r) = d \log \Gamma(r)/dr$ で $p = (1 + \xi)^2 \Gamma(1 + 2\xi)$，$q = \Gamma(2 + \xi)\{\Psi(1 + \xi) + (1 + \xi)/\xi\}$，$\gamma = 0.5772157...$ Euler の定数.

分布族 GEV(μ, σ, ξ)，$\mu \in \mathbb{R}$，$\sigma > 0$，$\xi \in \mathbb{R}$ は最尤推定量の漸近理論に関する通常の正則条件を満たしていない．しかし，$\xi > -0.5$ の場合は最尤推定量は一致推定量で漸近正規性を持ち漸近有効推定量になることが Smith (1985) により示されている．応用上 $\xi \leq -0.5$ となることは稀で，特に自然現象では $-0.5 < \xi < 0.5$ となることが多いと言われている．したがって，パラメータの推定は最尤法で行えばよい．

各最尤推定値の標準誤差の推定値は観測情報行列から求める方が精度が良い．これから (μ,σ,ξ) の各パラメータの信頼区間が求まる．

再現レベル

極値統計学では，次の上側（微少）確率点の推定が目的の場合が多い．

一般極値分布 GEV$(\mu,\,\sigma,\,\xi)$ の $1-1/T$ 確率点 R_T，すなわち

$$G(R_T)=G_\xi\left(\frac{R_T-\mu}{\sigma}\right)=1-1/T$$

は次のように表される：

$$R_T=\begin{cases}\mu+\sigma\big[\{-\log(1-1/T)\}^{-\xi}-1\big]/\xi, & \xi\neq 0,\\ \mu+\sigma\big[-\log\{-\log(1-1/T)\}\big], & \xi=0.\end{cases}$$

この確率点 R_T は**再現期間** (return period) T の **再現レベル** (return level) とよばれている．例えば年最大データを扱うとき，再現期間 $T=200$ 年の再現レベル R_{200} は 200 年に平均 1 度超えられる様な（大きな）値と解釈できる．一般に $n\ll T$ の場合を考える，これはデータの存在しない領域の推測（外挿）になる．

再現レベル R_T の最尤推定値は，$(\mu,\,\sigma,\,\xi)$ の最尤推定値を用いて

$$\widehat{z}_T=\begin{cases}\widehat{\mu}+\widehat{\sigma}\big[y_T^{-\widehat{\xi}}-1\big]/\widehat{\xi}, & \widehat{\xi}\neq 0,\\ \widehat{\mu}+\widehat{\sigma}\big[-\log y_T\big], & \widehat{\xi}=0,\end{cases}$$

となる．ただし，$y_T=-\log(1-1/T)$．また，デルタ法よりその分散は

$$V(\widehat{R}_T)\approx\nabla R_T^{\mathrm{T}}\,V\,\nabla R_T$$

となる．ここで，V は $(\widehat{\mu},\,\widehat{\sigma},\,\widehat{\xi})$ の分散共分散行列で

$$\nabla R_T^{\mathrm{T}}=\left[\frac{\partial R_T}{\partial\mu},\,\frac{\partial R_T}{\partial\sigma},\,\frac{\partial R_T}{\partial\xi}\right]$$
$$=\Big[1,\ \ (y_T^{-\xi}-1)/\xi,\ \ \sigma y_T^{-\xi}(-\log y_T)/\xi-\sigma(y_T^{-\xi}-1)/\xi^2\Big].$$

これらを $(\widehat{\mu},\,\widehat{\sigma},\,\widehat{\xi})$ で求める．$\widehat{\xi}<0$ の場合は分布の上限の推定

$$\widehat{R}_*=\widehat{\mu}-\widehat{\sigma}/\widehat{\xi}$$

が重要で，この場合は

$$\nabla R_*^{\mathrm{T}}=\big[1,\ -1/\xi,\ \sigma/\xi^2\big]$$

となる．

一般に，ξ や R_T の信頼区間はプロファイル尤度から求めた方が精度が良いことが知られている．

形状パラメータ ξ のプロファイル尤度は，$\xi=\xi_0$ として $l(\mu,\,\sigma,\,\xi_0)$ を μ と σ に関して最大化して求める．したがって，ξ の 95%の信頼区間は近似的に

$$\begin{aligned}&\Big\{\xi:\ 2\big\{l(\widehat{\mu},\,\widehat{\sigma},\,\widehat{\xi})-\max_{\mu,\,\sigma}\,l(\mu,\,\sigma,\,\xi)\big\}\le\chi_1^2(0.05)\Big\}\\ =\ \ &\Big\{\xi:\ \max_{\mu,\,\sigma}\,l(\mu,\,\sigma,\,\xi)\ge l(\widehat{\mu},\,\widehat{\sigma},\,\widehat{\xi})-1.921\Big\}\end{aligned}$$

となる．

再現レベル R_T のプロファイル尤度を求めるために，

$$\mu=R_T-\sigma\big[y_T^{-\xi}-1\big]/\xi$$

とおき，パラメータ $(\mu,\,\sigma,\,\xi)$ を $(R_T,\,\sigma,\,\xi)$ へ換える．このとき対数尤度 $l(R_T,\,\sigma,\,\xi)$ は

$$-n\log\sigma-(1+1/\xi)\sum_{i=1}^{n}\log\left[y_T^{-\xi}+\xi\left(\frac{z_i-R_T}{\sigma}\right)\right]$$
$$-\sum_{i=1}^{n}\left[y_T^{-\xi}+\xi\left(\frac{z_i-R_T}{\sigma}\right)\right]^{-1/\xi}$$

と表される．これから，R_T の 95%の信頼区間は近似的に

$$\begin{aligned}&\Big\{R_T:\ 2\Big\{l(\widehat{R}_T,\,\widehat{\sigma},\,\widehat{\xi})-\max_{\sigma,\,\xi}\,l(R_T,\,\sigma,\,\xi)\Big\}\le\chi_1^2(0.05)\Big\}\\ =\ \ &\Big\{R_T:\ \max_{\sigma,\,\xi}\,l(R_T,\,\sigma,\,\xi)\ge l(\widehat{R}_T,\,\widehat{\sigma},\,\widehat{\xi})-1.921\Big\}\end{aligned}$$

となる．

3.2 一般パレート (GP) モデル

ここでは水準超過データ $\{y_1,\,y_2,\,\ldots\,,\,y_n\}$ に一般パレート分布 GP$(\sigma,\,\xi)$ を適合させる場合の推測について述べる．

最尤法

一般パレート分布 GP$(\sigma,\,\xi)$ の対数尤度は，$\xi\neq 0$ のとき

$$l(\sigma,\,\xi)=-n\log\sigma-(1+1/\xi)\sum_{i=1}^{n}\log(1+\xi\,y_i/\sigma),$$
$$1+\xi\,y_i/\sigma>0,\quad i=1,2,\ldots,n,$$

で，$\xi=0$ のとき

$$l(\sigma)=-n\log\sigma-\frac{1}{\sigma}\sum_{i=1}^{n}y_i$$

となる．これらの対数尤度を最大にする最尤推定値 $(\widehat{\sigma},\widehat{\xi})$ または $\widehat{\sigma}$ を求める．

一般パレート分布 GP$(\sigma,\,\xi)$ の Fisher 情報量行列は

$$\frac{n}{(1+\xi)(1+2\xi)}\begin{bmatrix}(1+\xi)/\sigma^2 & 1/\sigma\\ 1/\sigma & 2\end{bmatrix}$$

となる．したがって $\xi > -1/2$ ならば情報行列は有限で，最尤推定量は漸近的に分散共分散行列が

$$V_n = \frac{1}{n}\left[\begin{array}{cc} 2\sigma^2(1+\xi) & -\sigma(1+\xi) \\ -\sigma(1+\xi) & (1+\xi)^2 \end{array}\right]$$

の 2 変量正規分布に従う（Smith, 1985）．

最尤推定値の分散共分散行列は観測情報行列から求めた方が精度が良い．

閾値の決定

データに一般パレート分布を適合させるためには**閾値** (threshold) の決定が必要である．決定には，次の一般パレート分布の性質を用いる．

確率変数 Y は一般パレート分布 $\mathrm{GP}(\sigma,\,\xi)$ に従うとする．このとき $\xi < 1$ で平均は存在し

$$E(Y) = \int_0^\infty (1-H(y))dy = \frac{\sigma}{1-\xi}$$

となる．

$v > 0$ のとき，条件付き確率変数 $Y - v \,|\, Y > v$ の分布は

$$\begin{aligned} P(Y-v>y\,|\,Y>v) &= \frac{1-H(y+v)}{1-H(v)} \\ &= \frac{\bigl(1+\xi(y+v)/\sigma\bigr)^{-1/\xi}}{\bigl(1+\xi v/\sigma\bigr)^{-1/\xi}} \\ &= \left(1+\xi\frac{y}{\sigma+\xi v}\right)^{-1/\xi} \end{aligned}$$

より，一般パレート分布 $\mathrm{GP}(\sigma+\xi v,\,\xi)$ に従う．同じ形状パラメータ ξ になることに注意．また

$$\sigma = (\sigma+\xi v) - \xi v \;:\; \text{一定}$$

である．

Y の **平均超過** (mean excess) **関数** $e(v)$ は

$$e(v) = E(Y - v\,|\,Y > v)$$

で定義される．$Y - v\,|\,Y > v \;\sim \mathrm{GP}(\sigma+\xi v,\,\xi)$ より

$$e(v) = \frac{\sigma+\xi v}{1-\xi} = \frac{\sigma}{1-\xi} + \frac{\xi}{1-\xi}\,v,$$

となり，これは v の一次関数である．特に，指数分布（$\xi = 0$）では $e(v)$ は定数となる．

これらの性質を用いて，次の様な方法で閾値を決定する．

値 u を動かして，u を超過したデータに一般パレート分布を適合し形状と尺度パラメータの推定値 $(\widehat{\sigma}_u,\,\widehat{\xi}_u)$ を求める．修正尺度の推定値 $\widehat{\sigma} = \widehat{\sigma}_u - \widehat{\xi}_u u$ と $\widehat{\xi}_u$ を u に対してプロットした図で，それより右端では推定値が一定になっていると見なせる最小の値を閾値とする．

値 u を動かして，u に対して標本平均超過関数を描いた図で，それより右端が線形と見なせる最小の値を閾値とする．

バリューアットリスク

極値統計学では次の確率点の推定が目的の場合が多い．

母集団分布 F で

$$F(\mathsf{VaR}_p) = 1 - p$$

となる確率点は**バリューアットリスク** (Value–at–Risk) とよばれる．これを推定するために母集団分布 $F(x)$ を次の様に分解する：

$$F(x) = (1-F(u))F_u(y) + F(u), \quad y = x - u.$$

適当な条件の下で u を十分大きくとると，F_u は GP 分布 H_ξ で近似できる：

$$F(x) \approx (1-F(u))H_\xi\left(\frac{x-u}{\sigma}\right) + F(u).$$

上式で $=$ として，$\zeta_u = 1 - F(u)$ とおくと

$$\mathsf{VaR}_p = u + \frac{\sigma}{\xi}\left\{\left(\frac{\zeta_u}{p}\right)^{\xi} - 1\right\},$$

となる．

閾値 u を決定し，閾値を超過するデータでパラメータの最尤推定値 $(\widehat{\sigma},\,\widehat{\xi})$ を求める．また，ζ_u は N_u/n で推定する．ここで n は標本サイズで N_u は閾値 u を超過したデータ数である．これらを代入して最尤推定値

$$\widehat{\mathsf{VaR}}_p = u + \frac{\widehat{\sigma}}{\widehat{\xi}}\left\{\left(\frac{\widehat{\zeta}_u}{p}\right)^{\widehat{\xi}} - 1\right\},$$

を得る．

$\widehat{\mathsf{VaR}}_p$ の標準誤差はデルタ法により求める．ここで，$(\widehat{\zeta}_u,\,\widehat{\sigma},\,\widehat{\xi})$ の分散共分散行列は近似的に

$$V^* = \left[\begin{array}{cc} \widehat{\zeta}_u(1-\widehat{\zeta}_u)/n & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & V_{N_u} \end{array}\right]$$

となる．ただし $\mathbf{0}^{\mathrm{T}} = (0,\,0)$．これを用いると VaR_p の最尤推定量の分散はデルタ法から次の様に求めることができる：

$$V(\widehat{\mathsf{VaR}}_p) \approx \nabla \mathsf{VaR}_p^{\mathrm{T}}\, V^* \nabla \mathsf{VaR}_p,$$

ただし

$$\nabla \mathsf{VaR}_p^{\mathrm{T}} = \left[\frac{\partial \mathsf{VaR}_p}{\partial \zeta_u}, \frac{\partial \mathsf{VaR}_p}{\partial \sigma}, \frac{\partial \mathsf{VaR}_p}{\partial \xi}\right]$$

で，$(\widehat{\zeta}_u, \widehat{\sigma}, \widehat{\xi})$ で評価する．これから信頼区間を求めることも出来るが，一般には次のプロファイル尤度を用いた方が精度が良い．

プロファイル尤度を用いた信頼区間は GEV モデルの場合と同様に求まる．例えば，ξ の 95%信頼区間は近似的に

$$\{\xi : \max_{\sigma} l(\sigma, \xi) \geq l(\widehat{\sigma}, \widehat{\xi}) - 1.921\}$$

となる．

一方 VaR_p の信頼区間は

$$\sigma = \frac{(\mathsf{VaR}_p - u)\xi}{(\zeta_u/p)^{\xi} - 1}$$

として，対数尤度 $l(\mathsf{VaR}_p, \xi)$ を考える．VaR_p を固定して $\max_{\xi} l(\mathsf{VaR}_p, \xi)$ を求めればよい．ここで応用上 ζ_u のバラツキは小さいので簡単のために無視している．

4　おわりに

1 変量の場合の極値理論とそれに基づく極値に関する統計的推測について述べた．ここでは紹介しなかったが「上位 r の極値データ」を用いる手法（高橋・渋谷 (2004) 参照）も応用上よく使われる．極値理論に関しては de Haan (1970) や Embrechts et al. (2001) が詳しい．また，統計的推測に関しては Coles (2001) や Beirlant et al. (2004) が詳しい．

参考文献

[1] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of Extremes, Theory and Applications.* Wiley.

[2] Coles, S. G. (2001). *An Introduction to Statistical Modeling of Extreme Values.* Springer.

[3] Embrechts, P., Klüppelberg, C. and Mikosch, T. (2001). *Modelling Extremal Events for Insurance and Finance*, 3rd ed. Springer.

[4] de Haan, L. (1970). *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes.* Mathematical Centre Tracts 32, Mathematisch Centrum. Amsterdam.

[5] Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Cambridge Philos. Soc.* **24**, 180–190.

[6] Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Math. Polon.* **6**, 93–116.

[7] Gnedenko, B. (1943). Sur la distibution limite du terme maximum d'une serie aleatoire. *Ann, Math.* **44**, 423–453. Translated and reprinted in: *Breakthroughs in Statistics*, Vol.I, 1992, eds. S. Kotz and N. L. Johnson, Springer-verlag, pp. 195–225.

[8] Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.* **3**, 119–131.

[9] Prescott, P. and Walden, A. T. (1980). Maximum likelihood estimation of the parameters of the generalized extreme value distribution. *Biometrika* **67**, 723–724.

[10] Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika* **72**, 67–90.

[11] 高橋倫也，渋谷政昭 (2004). 上位 r 個の観測値に基づく確率点の推定．統計数理，第 52 巻，第 1 号，93–116.