

極値統計学

高橋 倫也 (神戸大学・海事科学部)
r-taka@maritime.kobe-u.ac.jp

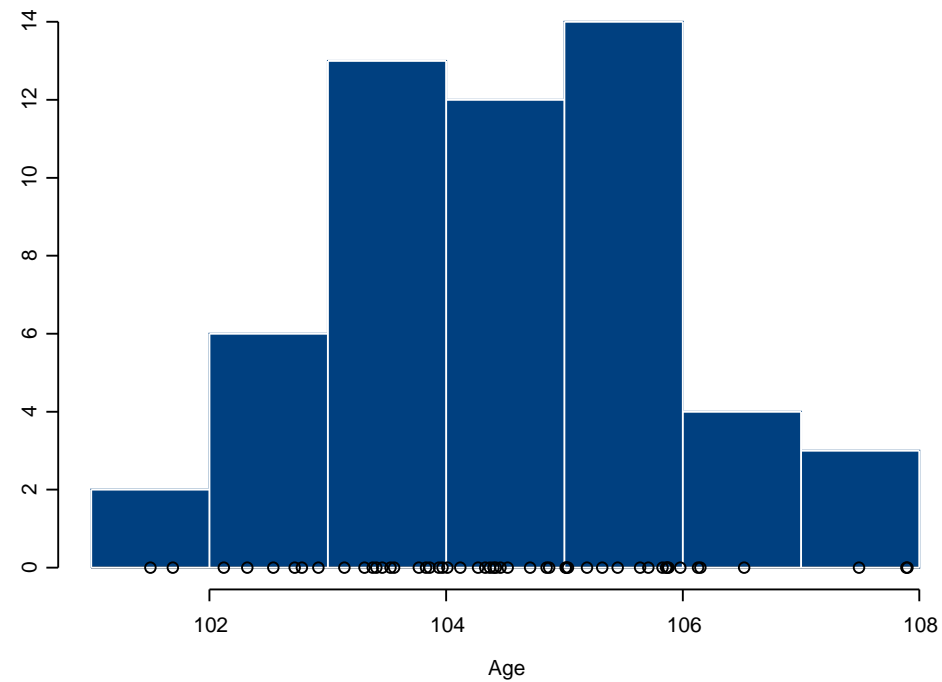
平成 2 0 年 8 月

Einmahl and Magnus (2007)

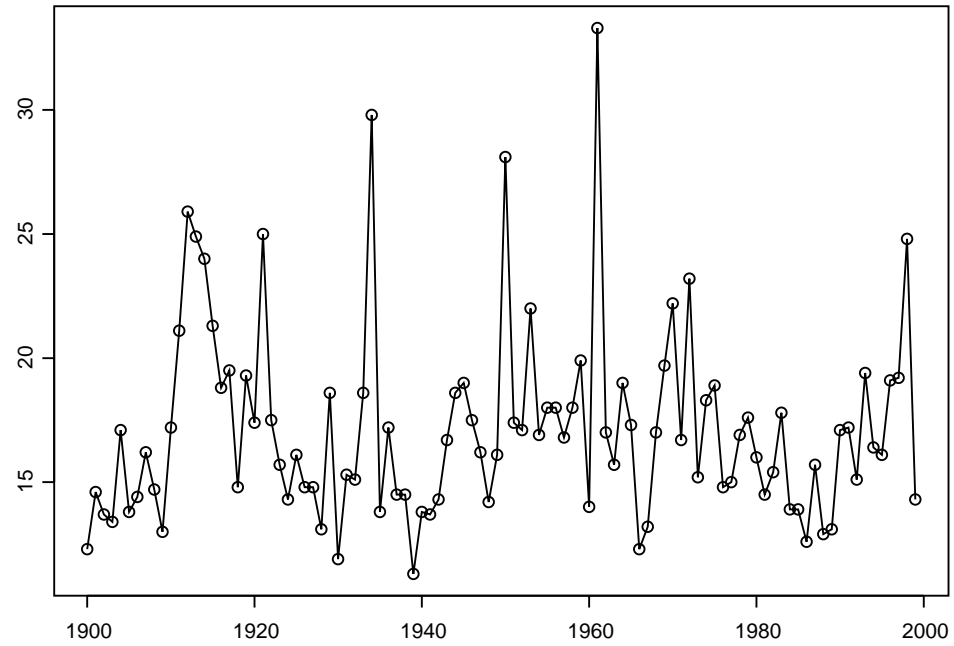
Records in athletics through extreme-value theory.

百メートル走は9秒29まで短縮可能？ 数学者が予測 04:35 産経新聞

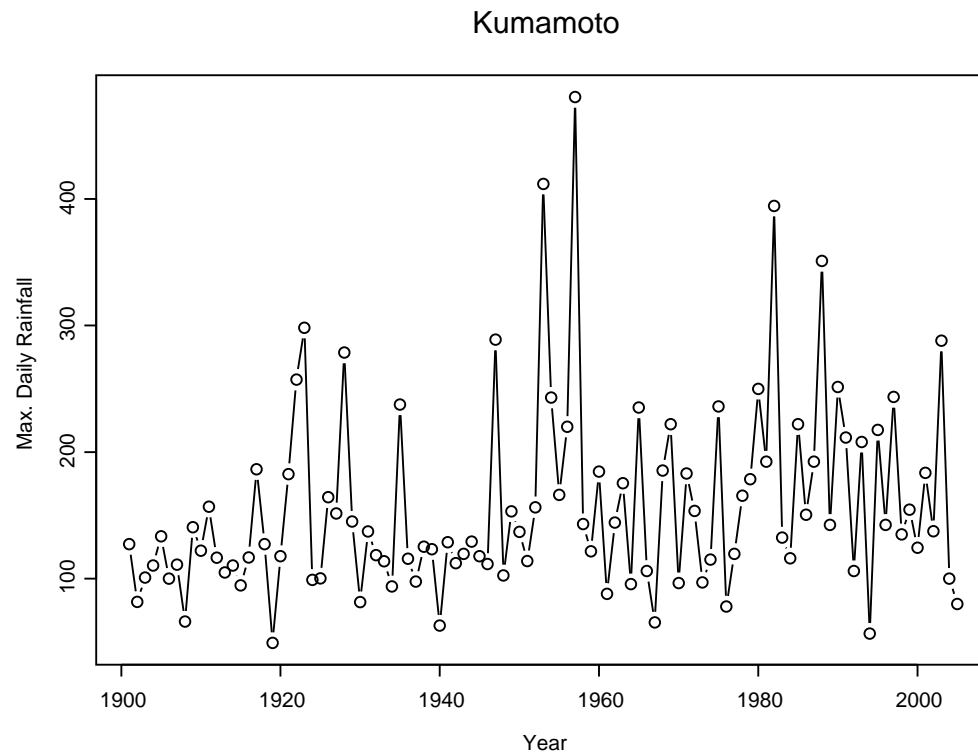
百メートルは9秒29で走れるが、男子マラソンはもう限界。ティルベルク大(ドイツ)のオランダ人数学者、ジョン・アインマール教授がこのほど、陸上の各種世界記録の限界値の研究結果を発表し、男子百メートルでアサファ・パウエル(ジャマイカ)の持つ9秒77の記録は9秒29まで短縮できるとの見通しを示した。



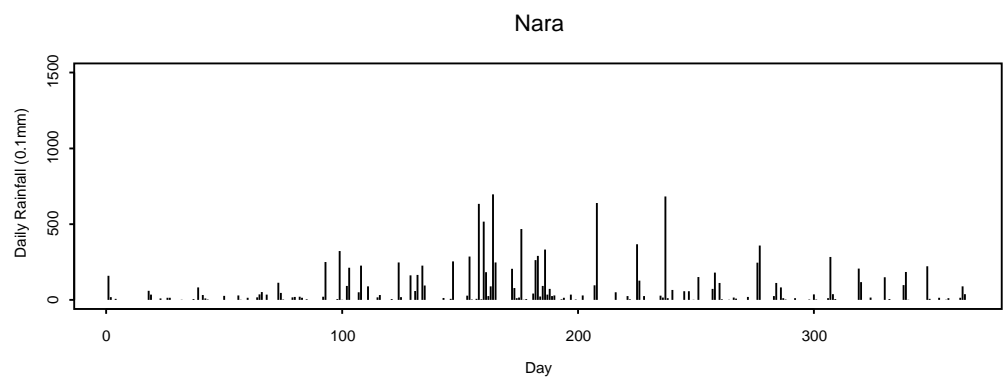
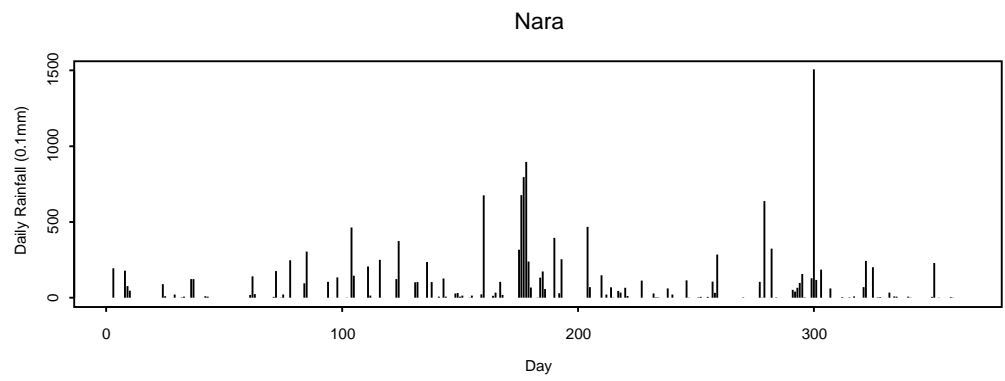
Castillo, Hadi, Balakrishnan and Sarabia (2004) p.11–12.
The yearly oldest ages at death in Sweden during the period from
1905 to 1958 for women. (54名, 101.50歳 ~ 107.90歳)



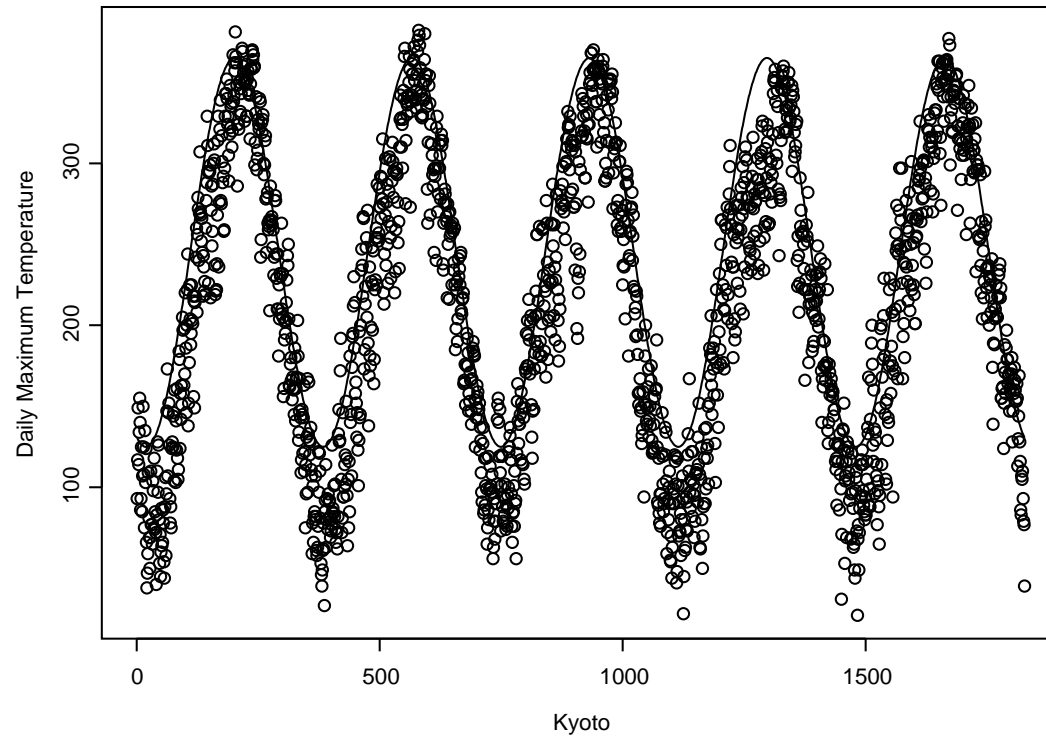
大阪の年最大風速 (m/s) , 1900年 ~ 1999年 .



熊本の年最大日雨量 (0.1mm) , 1901年から2005年 .



奈良の日降水量 (0.1mm) 1961年 (上) 1962年 (下) .



京都の日最高気温 (単位 0.1度C) , 2000年1月1日 ~ 2004年12月31日 .

目標

データから，どのような大きな値がどれくらいの確率でいつ頃出現するか知りたい．

極値理論

データすべてに適合する分布は無いが，ゆるい条件の下で，ある期間または領域での最大データに適合する分布や，十分大きな閾値を超過するデータに適合する分布は存在する．

これらの分布は一般極値分布，一般パレート分布と呼ばれ，これらをデータに適合させある長い期間（または広い領域）で，ある大きな値が平均 1 回出現するであろうと予測する事は可能．

1 変量の場合の極値統計学

Colesの本の極値理論に基づく1変量データ解析法について紹介する

A. 独立で同一分布に従う1変量確率変数列に関する極値理論の基礎

B. 尤度による推測法と実データ解析

1 . はじめに

リスクは，一般に異常に大きな（または小さな）値が観測されたときに発生．
観測データの平均ではなく大きな値の変動が重要．
例えば1年間の観測データから一番大きな値のものを取り出す．
これを何年か繰り返して収集したものを区分最大値（block maximum）
データという．

1 変量の独立で同一分布の場合 （1920年代～1940年代）
Fréchet , Fisher and Tippett , von Mises , Gnedenko 等
極値極限分布（極値分布）には3つの型がある．

分布がそれらの極値分布の吸引領域に属するための必要十分条件．

3つの型の極値分布を一つの式で表した分布を一般極値分布といい，区分最大
値データの解析に適用されている．

2 変量への拡張 (1960年前後)

Geffroy, Tiago de Oliveira, Sibuya 等により行われた。

Gumbel (1958) の本の出版により極値統計学は気象学, 水文学, 諸工学 (材料, 腐食, 土木, 建築, 信頼性) に広まった。

水文学の分野：“Peaks Over Threshold (POT)”法。
十分大きな閾値を超過するデータ（水準超過値 (threshold exceedances) データという）が指数分布で近似出来るという事実に基づくデータ解析手法。

一般パレート分布が Balkema and de Haan (1974) と Pickands (1975) により導出された。この一般パレート分布を水準超過値データ解析に適用する。

上位 r 個のデータを用いて利用可能データを増やす
Weissman (1978), Smith (1986), Tawn (1988) 等

極値理論の数学的な基礎になる正則変動関数の理論
多変量極値分布の応用についての研究
ファイナンスへの応用の研究。

2 . モデルと準備

独立で同一分布に従う 1 変量確率変数列を考える .

$$X_1, X_2, \dots, X_n, \dots$$

同一分布を次の様にする :

$$F(x) = P(X_i \leq x), \quad i = 1, 2, \dots$$

ここで , X_1, X_2, \dots, X_n を大きさの順に列べた順序統計量を考える :

$$X_{(1:n)} \leq X_{(2:n)} \leq \dots \leq X_{(n:n)}$$

極値統計量

$$Z_n = X_{(n:n)} = \max \{ X_1, X_2, \dots, X_n \} = \max_{1 \leq i \leq n} X_i$$

n を大きくしていくと Z_n は分布 F の上端点 x_+ へ収束 :

$$Z_n \rightarrow x_+ = \sup\{x : F(x) < 1\} \leq \infty, \quad n \rightarrow \infty.$$

Z_n を正規化 : 数列 $a_n > 0$, $b_n \in \mathbb{R}$ ($n = 1, 2, \dots$) と退化していない分布 $G(x)$ を持つ確率変数 Z が存在して ,

$$\frac{Z_n - b_n}{a_n} \xrightarrow{d} Z, \quad n \rightarrow \infty.$$

すなわち

$$P\left(\frac{Z_n - b_n}{a_n} \leq x\right) \rightarrow P(Z \leq x) = G(x).$$

G : 極値分布 (extreme value distribution) , 分布 F は極値分布 G の最大値吸引領域 (maximum domain of attraction) に属する ($F \in MDA(G)$) .
(a_n, b_n) : 吸引係数 .

基本的な問題

- (1) 極値分布 $G(x)$ はどのような分布か？
- (2) $F \in MDA(G)$ となるための F の必要十分条件は何か？
- (3) 吸引係数 a_n, b_n はどう決めるか？

注意 1 . 以下では最大の場合のみを議論する .

最小 $W_n = \min_{1 \leq i \leq n} X_i$ に関する結果は , 関係式

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

より , 最大の場合からすぐ得られる . 例えば , 最小の場合の極値分布 W は , G を (対応する) 最大の極値分布とすると $W(x) = 1 - G(-x)$ で求まる .

正規化した極値統計量の分布関数：

$$\begin{aligned} P\left(\frac{Z_n - b_n}{a_n} \leq x\right) &= P(Z_n \leq a_n x + b_n) = P\left(\max_{1 \leq i \leq n} X_i \leq a_n x + b_n\right) \\ &= P(X_i \leq a_n x + b_n, \quad i = 1, 2, \dots, n) \\ &= \prod_{i=1}^n P(X_i \leq a_n x + b_n) \\ &= F^n(a_n x + b_n). \end{aligned}$$

弱収束は分布関数の収束 :

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x), \quad \forall x \in \mathbb{R}.$$

これから , $a_n x + b_n \uparrow x_+ = \sup\{x : F(x) < 1\} \leq \infty$.

$y = a_n x + b_n$ とおくと ,

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

$X_i^* = \mu + \sigma X_i$, $i = 1, 2, \dots, n$. $\mu \in \mathbb{R}$, $\sigma > 0$.

$$\implies Z_n^* = \max_{1 \leq i \leq n} X_i^* = \mu + \sigma \max_{1 \leq i \leq n} X_i = \mu + \sigma Z_n.$$

$0 < G(x) < 1$ のとき ,

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x) \iff \lim_{n \rightarrow \infty} n[1 - F(a_n x + b_n)] = -\log G(x).$$

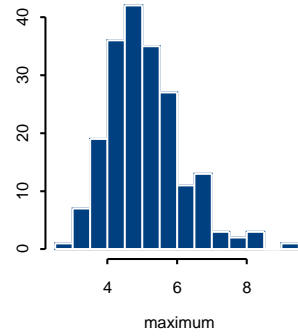
このことは ,

$$F^n(a_n x + b_n) = \left\{ 1 + \frac{-n[1 - F(a_n x + b_n)]}{n} \right\}^n \rightarrow \exp \left\{ -\lim_{n \rightarrow \infty} n[1 - F(a_n x + b_n)] \right\}$$

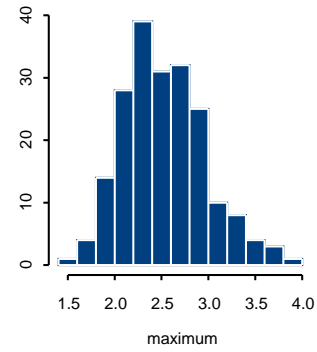
より明らか . ここで , 次の事実を使った :

$$x_n \rightarrow x, \quad n \rightarrow \infty \implies \lim_{n \rightarrow \infty} \left(1 + \frac{x_n}{n} \right)^n = e^x.$$

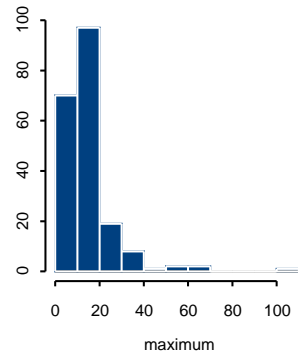
Block Maximum from Exp(1)



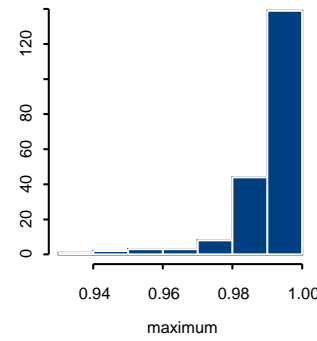
Block Maximum from N(0, 1)



Block Maximum from Par(2)



Block Maximum from U(0, 1)



100個の乱数の最大値を200個生成したときのヒストグラム。

極値分布

具体的な分布の極値統計量の漸近分布（極値分布）

F : 分布関数 , $f(x)$: 密度関数

例 1 . 標準指数分布 $\text{Exp}(1)$

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

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

ここで $a_n = 1$, $b_n = \log n$.

例 2 . パレート分布 ($\alpha > 0$)

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

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

ここで $a_n = n^{1/\alpha}$, $b_n = 0$.

例3 . ベータ分布 ($\alpha > 0$)

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

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

ここで $a_n = n^{-1/\alpha}$, $b_n = 1$.

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 :

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

注意 2 . 次の式が成り立つ : $\forall x \in \mathbb{R}$ と $\forall n \in \mathbb{N}$ に対して ,

$$\Lambda^n(x + \log n) = \Lambda(x), \quad \Phi_\alpha^n(n^{1/\alpha}x) = \Phi_\alpha(x), \quad \Psi_\alpha^n(n^{-1/\alpha}x) = \Psi_\alpha(x).$$

注意 3 . 非退化な分布 G が極値分布 $\iff \exists a_n > 0, b_n \in \mathbb{R} (n = 1, 2, \dots)$
s.t.

$$G^n(a_n x + b_n) = G(x), \quad \forall x \in \mathbb{R} \quad \text{and} \quad \forall n = 1, 2, \dots$$

この関数方程式は最大値安定性 (max-stable) と呼ばれ, この解が定理 1 の 3 つの極値分布になる .

3つの極値分布は, 次の1つの式, von Mises-Jenkinson 表現 ((1936), (1955)) で表せる.

$$G_\gamma(x) = \exp\{-(1 + \gamma x)^{-1/\gamma}\}, \quad 1 + \gamma x > 0, \quad -\infty < \gamma < \infty.$$

分布 $G_\gamma(x)$ を持つ確率変数を Z とすると:

$$Z = (E^{-\gamma} - 1)/\gamma, \quad E \sim \text{Exp}(1); \quad \gamma^{-1} \log(1 + \gamma Z) = Z_0 \sim \Lambda.$$

また, 3つの極値分布の型との関係は:

$$\Lambda(x) = G_0(x), \quad \Phi_\alpha(x) = G_{1/\alpha}(\alpha(x-1)), \quad \Psi_\alpha(x) = G_{-1/\alpha}(\alpha(x-1)).$$

吸引領域

極値分布の最大値吸引領域に関する必要十分条件は次のようになる：

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

$$F \in MDA(\Phi_\alpha) \iff F(x) < 1, \forall x \quad \text{and} \quad \lim_{x \rightarrow \infty} \frac{1 - F(tx)}{1 - F(x)} = t^{-\alpha}, \quad t > 0.$$

$$F \in MDA(\Psi_\alpha) \iff x_+ < \infty \quad \text{and} \quad \lim_{x \uparrow x_+} \frac{1 - F(x_+ - (x_+ - x)t)}{1 - F(x)} = t^\alpha, \quad t > 0.$$

$$F \in MDA(\Lambda) \iff \exists s(\cdot) > 0 \quad \text{s.t.} \quad (*) \quad \lim_{x \uparrow x_+} \frac{1 - F(x + ts(x))}{1 - F(x)} = e^{-t}.$$

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

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

注意 4 . 吸引係数は次で求まる :

$$F \in MDA(\Phi_\alpha), \quad F(a_n) = 1 - 1/n, \quad b_n = 0.$$

$$F \in MDA(\Psi_\alpha), \quad F(x_+ - a_n) = 1 - 1/n, \quad b_n = x_+.$$

$$F \in MDA(\Lambda), \quad F(b_n) = 1 - 1/n, \quad a_n = s(b_n).$$

標本数が確率変数の場合

独立で同一分布に従う確率変数列 $\{X_i\}_{i=1}^{\infty}$ の母集団分布 F は $MDA(G)$ に属し、標本数 N は $\{X_i\}_{i=1}^{\infty}$ とは独立でポアソン分布 $Po(\lambda)$ に従う、すなわち、

$$P(N = i) = e^{-\lambda} \frac{\lambda^i}{i!}, \quad i = 0, 1, 2, \dots$$

とする。

標本数が $N \geq 1$ の条件の下で、 $Z_N = \max\{X_1, \dots, X_N\}$ の分布 $F^N(y)$:

$$\begin{aligned} F^N(y) &= P(Z_N \leq y \mid N \geq 1) = \frac{P(\max\{X_1, \dots, X_N\} \leq y, N \geq 1)}{P(N \geq 1)} \\ &= \frac{1}{1 - e^{-\lambda}} \left(e^{-\lambda\{1-F(y)\}} - e^{-\lambda} \right). \end{aligned}$$

ここで, a_n, b_n を F の吸引係数として, $y \equiv a_{[\lambda]}x + b_{[\lambda]}$ ($[\lambda]$ は λ を超えない最大の整数) とすると,

$$\lim_{\lambda \rightarrow \infty} \lambda \{1 - F(a_{[\lambda]}x + b_{[\lambda]})\} = \lim_{\lambda \rightarrow \infty} \frac{\lambda}{[\lambda]} [\lambda] \{1 - F(a_{[\lambda]}x + b_{[\lambda]})\} = -\log G(x).$$

$\lambda \rightarrow \infty$ のとき

$$F^N(a_{[\lambda]}x + b_{[\lambda]}) \rightarrow \exp(\log G(x)) = G(x).$$

よって,

$$P(Z_N \leq y \mid N \geq 1) \approx G\left(\frac{y - b_{[\lambda]}}{a_{[\lambda]}}\right).$$

一般に標本数が確率変数の場合でもその期待値が十分大であれば，最大値データは極値分布に従うとして良い．

腐食の分野：一定面積中の孔食の個数は確率変数とみなせるが，各一定面積中の最大孔食深さデータに極値分布を適合させて解析出来る．

一般極値分布

定義 1 . 次の分布を一般極値 (generalized extreme value) 分布 といい $GEV(\mu, \sigma, \xi)$ ($-\infty < \mu < \infty, \sigma > 0, -\infty < \xi < \infty$) で表す .

$$G(z) = \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} = G_\xi \left(\frac{z - \mu}{\sigma} \right), \quad 1 + \xi(z - \mu)/\sigma > 0.$$

ただし , G_ξ は標準一般極値分布 $GEV(0, 1, \xi)$ の分布関数

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

とする . μ は位置パラメータ , σ は尺度パラメータ , ξ は形状パラメータである .

$G(z)$ は,

$\xi < 0$ のときは Weibull 分布で $z < \mu - \sigma/\xi$,

$\xi = 0$ のときは次から Gumbel 分布で $-\infty < z < \infty$,

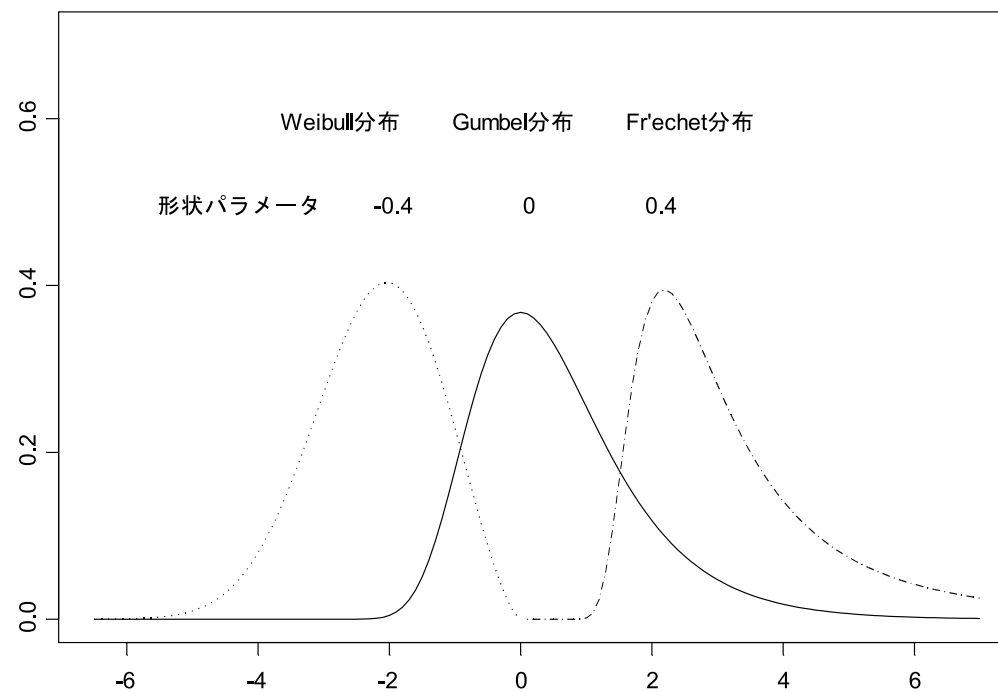
$$G_0((z - \mu)/\sigma) = \lim_{\xi \rightarrow 0} G_\xi((z - \mu)/\sigma) = \exp\{-\exp[-(z - \mu)/\sigma]\}$$

$\xi > 0$ の場合は Fréchet 分布で $z > \mu - \sigma/\xi$.

一般極値分布 $G_\xi(z)$ の密度関数は,

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

となる.



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 .

一般パレート分布

定義 2 . 次の分布を一般パレート (generalized Pareto) 分布といい $GP(\sigma, \xi)$ ($\sigma > 0, -\infty < \xi < \infty$) 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

$$H_\xi(y) = 1 - (1 + \xi y)^{-1/\xi}, \quad 1 + \xi y > 0,$$

とする . σ は尺度パラメータ, ξ 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 $0 < y < -\sigma/\xi$,

$\xi = 0$ のときは次より指数分布で $0 < y < \infty$,

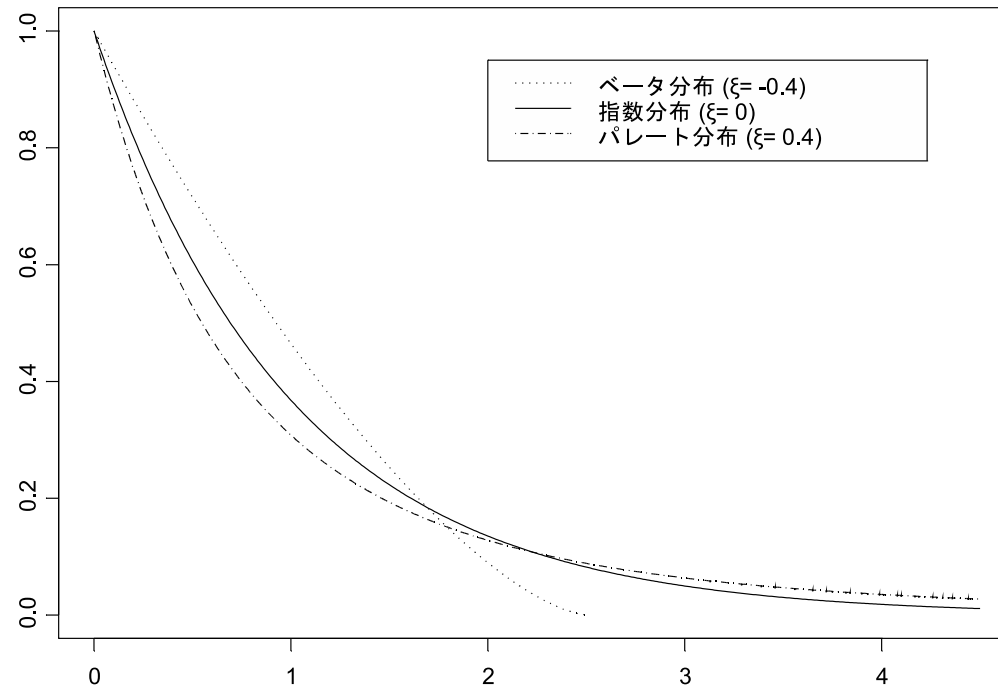
$$H_0(y/\sigma) = \lim_{\xi \rightarrow 0} H_\xi(y/\sigma) = 1 - e^{-y/\sigma}$$

$\xi > 0$ の場合はパレート分布で $0 < y < \infty$.

一般パレート分布 $H_\xi(y)$ の密度関数は ,

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

となる .



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

定理 3 . [von Mises (1936) , Smith (1990)]

$$\phi(x) := \frac{1 - F(x)}{f(x)}, \quad \phi'(x) = \frac{d}{dx}\phi(x)$$

$$\lim_{x \rightarrow x_+} \phi'(x) = \xi \implies F \in MDA(G_\xi).$$

証明 . 次の関係がある :

$$\frac{d}{dx}\{\log(1-F(x))\} = \frac{-f(x)}{1-F(x)} \implies \log(1-F(x)) = -\int_{x_-}^x \frac{f(t)}{1-F(t)} dt.$$

ただし $x_- = \inf\{x : F(x) > 0\}$. よって ,

$$1 - F(x) = \exp \left[-\int_{x_-}^x \frac{dt}{[1 - F(t)]/f(t)} \right] = \exp \left[-\int_{x_-}^x \frac{dt}{\phi(t)} \right].$$

ゆえに

$$\begin{aligned} \frac{1 - F(u + x\phi(u))}{1 - F(u)} &= \exp \left[- \int_{x_-}^{u+x\phi(u)} \frac{dt}{\phi(t)} \right] / \exp \left[- \int_{x_-}^u \frac{dt}{\phi(t)} \right] \\ &= \exp \left[- \int_u^{u+x\phi(u)} \frac{dt}{\phi(t)} \right] = \exp \left[- \int_0^x \frac{ds}{\phi(u + s\phi(u))/\phi(u)} \right]. \end{aligned}$$

ここで $u < t < u + x\phi(u)$, $0 < s = (t - u)/\phi(u) < x$. また, ϕ が微分可能だから,

$$\frac{\phi(u + s\phi(u))}{\phi(u)} = 1 + \frac{\phi(u + s\phi(u)) - \phi(u)}{\phi(u)} = 1 + \int_0^s \phi'(u + w\phi(u)) dw = 1 + s\phi'(y).$$

ただし $u < y < u + x\phi(u)$.

$$\frac{1 - F(u + x\phi(u))}{1 - F(u)} = \exp \left[- \int_0^x \frac{ds}{1 + s\phi'(y)} \right] = \exp \left[- \int_1^{1+x\phi'(y)} \frac{dt}{\phi'(y)t} \right]$$

($t = 1 + s\phi'(y)$ とする)

$$= \exp \left[-\frac{1}{\phi'(y)} \log(1 + x\phi'(y)) \right] = (1 + x\phi'(y))^{-1/\phi'(y)}.$$

ここで $n[1 - F(b_n)] = 1$, $a_n = \phi(b_n)$ とおくと, $b_n \uparrow x_+$ で $\phi'(b_n) \rightarrow \xi$ であるから,

$$\frac{n[1 - F(b_n + x\phi(b_n))]}{n[1 - F(b_n)]} \approx (1 + x\phi'(b_n))^{-1/\phi'(b_n)},$$

$$\lim_{n \rightarrow \infty} n[1 - F(a_n x + b_n)] = (1 + \xi x)^{-1/\xi} = -\log G_\xi(x).$$

したがって

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G_\xi(x).$$

ここで, $F(b_n) = 1 - 1/n$, $a_n = [1 - F(b_n)]/f(b_n) = n[1 - F(b_n)]/nf(b_n) = 1/nf(b_n)$. \square

注意 5 .

例 3 のベータ分布（上限のある連続分布）の場合は $\xi = -1/\alpha < 0$.

例えば，一様分布の場合は $\xi = -1$.

正規分布，指数分布，ワイブル分布，ガンマ分布，対数正規分布の様に分布の右裾が指数的に減少する連続分布の場合は $\xi = 0$.

例 2 のパレート分布のように分布の右裾がべき乗で減少する連続分布の場合は $\xi > 0$.

教科書に出てくる連続分布はほとんどどれかの極値分布の吸引領域に属している . 離散分布（例えばポアソン分布等）は一般に極値分布の吸引領域には属さない .

例 4 . 次の密度関数を持つワイブル分布を考える :

$$f(x) = cx^{c-1} \exp(-x^c), \quad x \geq 0, \quad c > 0.$$

このとき , $F(x) = 1 - \exp(-x^c)$ で

$$\phi(x) = \frac{1}{cx^{c-1}}$$

より

$$\phi'(x) = \frac{1-c}{cx^c} \rightarrow 0, \quad x \rightarrow \infty.$$

したがって , ワイブル分布は Gumbel 分布 ($\xi = 0$) の吸引領域に属する .

注意 6 . Penultimate 近似 : 証明の中で ,

$$n[1 - F(a_n x + b_n)] \approx (1 + \xi_n x)^{-1/\xi_n}, \quad \xi_n = \phi'(b_n).$$

よって ,

$$F^n(a_n x + b_n) \approx \exp[-(1 + \xi_n x)^{-1/\xi_n}] = G_{\xi_n}(x).$$

G_{ξ_n} による近似の方が G_{ξ} による近似よりも良い場合がある . これを penultimate 近似という . 特に $\xi = 0$ の場合 , $\xi_n > 0$ または $\xi_n < 0$ のとき極値分布の型が違
う . ところで , 正規分布 ($\xi = 0$) の場合に Gumbel 分布への収束が遅いことを
Fisher and Tippett (1928) は注意している . Penultimate 近似に関する詳し
い研究は Cohen (1982) にある . データに特定の極値分布 (例えば Gumbel 分
布) を適合するのではなく , 一般極値分布を適合してデータから ξ_n を推定した方
が推測精度が良くなる場合がある .

注意7 . 一般パレート分布による近似 : 証明の中で次を示した .

$$\frac{1 - F(u + x\phi(u))}{1 - F(u)} = [1 + x\phi'(y)]^{-1/\phi'(y)}, \quad x > 0.$$

ここで , $y = x\phi(u)$, $\sigma = \phi(u)$ とおくと ,

$$P(X > u+y | X > u) = \frac{1 - F(u + y)}{1 - F(u)} = \left(1 + \frac{\phi'(y)y}{\phi(u)}\right)^{-1/\phi'(y)} \approx \left(1 + \xi \frac{y}{\sigma}\right)^{-1/\xi}.$$

すなわち ,

$$F_u(y) = P(X - u \leq y | X > u) \approx 1 - (1 + \xi y/\sigma)^{-1/\xi} = H_\xi(y/\sigma).$$

$F \in MDA(G_\xi)$ のとき , 分布 F の右裾の条件付き分布は一般極値分布 G_ξ と同じ形状パラメータ ξ の一般パレート分布 H_ξ で近似できる .

裾指数 (tail index)

分布 F が一般極値分布 G_ξ の吸引領域に属する ($F \in MDA(G_\xi)$) とき, F の右裾の条件付き分布は一般パレート分布 H_ξ で近似できた. この共通のパラメータ ξ を F の裾指数という.

リスクの分布が F で, その裾指数を ξ とする. Teugels は国際統計協会 (ISI, 1999) の講演で「飛行場における風速 (F)」のリスクを次の様に分類した:

$\xi < 0$ ならば安全 (リスクには上限がある).

$\xi = 0$ ならば要注意 (リスクには上限が無く, 大きなリスクが発生する可能性がある).

$\xi > 0$ ならば危険 (リスクには上限が無く, 非常に大きなリスクが発生しやすい).

点過程による極値理論

\mathbb{R}^d 上の確率ベクトル列 $\{Y_i\}$ にたいして,

$$N(A) := \sum_i I_A(Y_i), \quad A \subset \mathbb{R}^d,$$

とおく. ここで, I_A は集合 A の定義関数:

$$I_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

($N(A)$ は集合 A に含まれる $\{Y_i\}$ の個数.)

$N = N(\cdot)$ を点過程 (point process) と呼ぶ.

N : **Poisson**過程 \iff (i) $N(A) \sim \text{Po}(\Lambda(A))$, $\forall A$: Borel集合, (ii) 互いに交わらない Borel 集合 B_1, \dots, B_n に対して $N(B_1), \dots, N(B_n)$ が互いに独立.

非負の測度 $\Lambda(\cdot)$: 平均値測度

任意の Borel 集合 A にたいして

$$\Lambda(A) = \int_A \lambda(y) dy$$

のとき, $\lambda(y)$ を Poisson 過程の強度と言う.

$\lambda(y) = \text{定数}$ のとき Poisson 過程は定常.

定理 4 . X_1, X_2, \dots を i.i.d. 確率変数列とし, $Z_n = \max_{1 \leq i \leq n} X_i$ に対して $a_n > 0, b_n \in \mathbb{R}$ が存在して

$$P\{(Z_n - b_n)/a_n \leq z\} \rightarrow G(z),$$

とする . すなわち, $F \in MDA(G)$. ただし

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

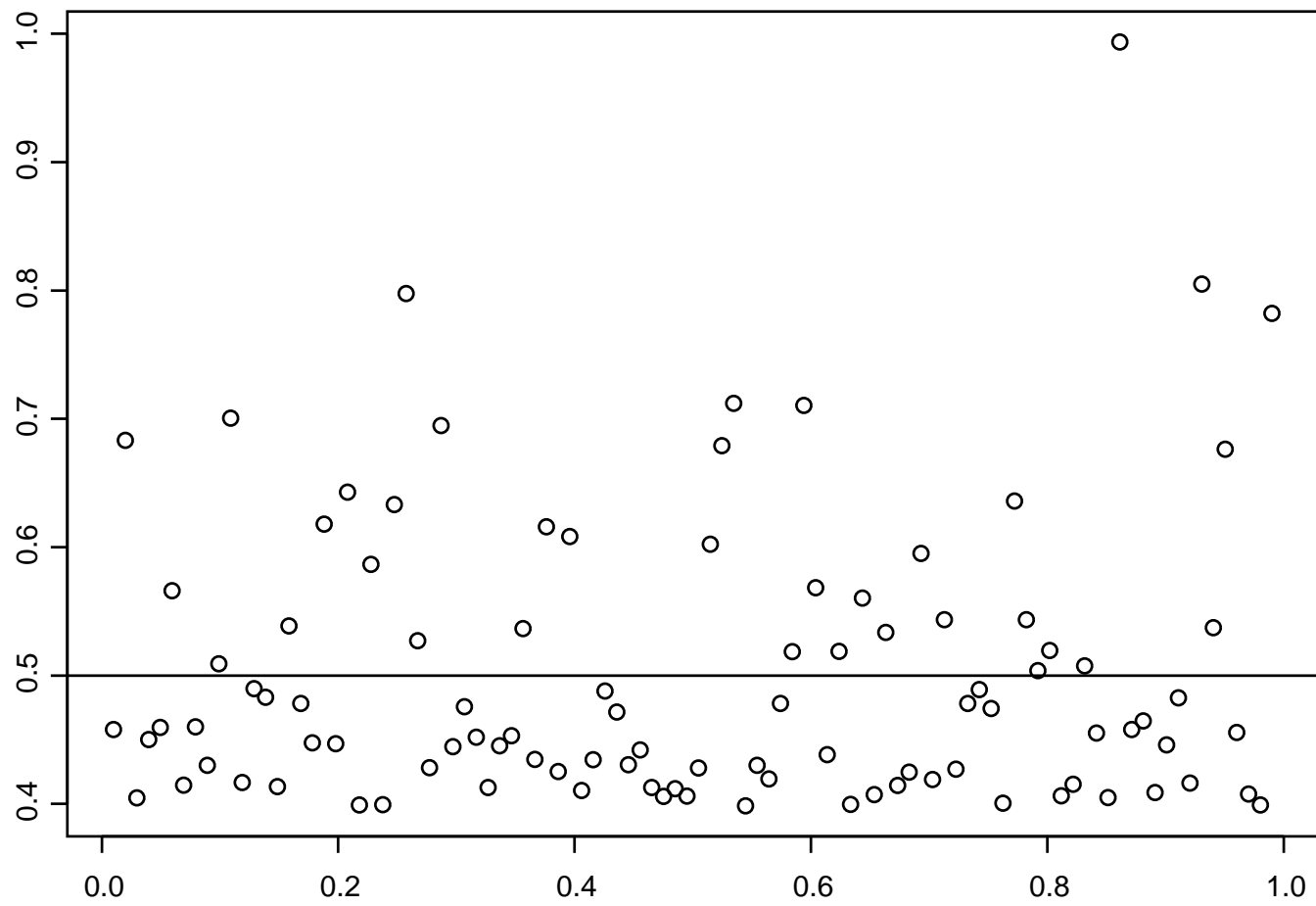
また, z_-, z_+ をそれぞれ分布 G の下, 上端点とする . このとき, 点過程列

$$N_n = \left\{ \left(\frac{i}{n+1}, \frac{X_i - b_n}{a_n} \right) : i = 1, \dots, n \right\}$$

は, 任意の $u > z_-$ の領域 $(0, 1) \times (u, z_+)$ で, $A = [t_1, t_2] \times (z, z_+)$ の強度が次で与えられる Poisson 過程に収束する .

$$\Lambda(A) = (t_2 - t_1) \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi}.$$

Point Process



証明 . $A = [0, 1] \times (u, z_+)$ という領域を考える . ただし , u は十分大 . N_n の n 個の各点が A に落ちる確率 p は

$$p = P\{(X_i - b_n)/a_n > u\} = 1 - F(a_n u + b_n) = \frac{n[1 - F(a_n u + b_n)]}{n}$$
$$\approx \frac{1}{n}(-\log G(u)) = \frac{1}{n} \left[1 + \xi \left(\frac{u - \mu}{\sigma} \right) \right]^{-1/\xi} .$$

X_i は互いに独立であったから , $N_n(A)$ は二項分布 $B(n, p)$ に従う . 一方 , $n \rightarrow \infty$ のとき , $N_n(A)$ は Poisson 分布に収束する .

$$N_n(A) \sim B(n, p) \xrightarrow{d} \text{Po}(np), \quad n \rightarrow \infty .$$

ここで

$$np = \Lambda(A) = \left[1 + \xi \left(\frac{u - \mu}{\sigma} \right) \right]^{-1/\xi} .$$

時間軸に関して点過程は定常であるから, $A = [t_1, t_2] \times (u, z_+)$, $[t_1, t_2] \subset [0, 1]$ の形の領域での $N_n(A)$ の極限分布は $Po(\Lambda(A))$ となる. ただし

$$\Lambda(A) = (t_2 - t_1) \left[1 + \xi \left(\frac{u - \mu}{\sigma} \right) \right]^{-1/\xi}.$$

領域 A の形と互いに交わらない集合での個数は互いに独立になる事に注意すると定理が得られる. \square

GEVモデル

$A_z = (0, 1) \times (z, \infty)$ を考える . 事象 $\{(Z_n - b_n)/a_n \leq z\}$ と $\{N_n(A_z) = 0\}$ は等しい . よって

$$\begin{aligned} P\{(Z_n - b_n)/a_n \leq z\} &= P\{N_n(A_z) = 0\} \rightarrow P\{N(A_z) = 0\} = \exp\{-\Lambda(A_z)\} \\ &= \exp\left\{-\left[1 + \xi \left(\frac{z - \mu}{\sigma}\right)\right]^{-1/\xi}\right\}. \end{aligned}$$

正規化した極値統計量の極限分布は一般極値分布になる .

GPモデル

平均値測度を次の様に分解する： $\Lambda(A_z) = \Lambda_1([t_1, t_2]) \times \Lambda_2((z, \infty))$. ただし

$$\Lambda_1([t_1, t_2]) = t_2 - t_1 \quad \text{and} \quad \Lambda_2((z, \infty)) = \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} .$$

このとき , $\tilde{\sigma} = \sigma + \xi(u - \mu)$ とすると

$$P\left\{ (X_i - b_n)/a_n > z \mid (X_i - b_n)/a_n > u \right\}$$

$$\rightarrow \frac{\Lambda_2(z, \infty)}{\Lambda_2(u, \infty)} = \frac{[1 + \xi(z - \mu)/\sigma]^{-1/\xi}}{[1 + \xi(u - \mu)/\sigma]^{-1/\xi}} = \left[1 + \xi \left(\frac{z - u}{\tilde{\sigma}} \right) \right]^{-1/\xi} .$$

上位 r 個の同時漸近分布 $r(> 1)$: 固定

順序統計量の上位 r 個 $(X_{(n:n)}, X_{(n-1:n)}, \dots, X_{(n-r+1:n)})$ の同時漸近分布

極値統計量 $Z_n = X_{(n:n)}$ の分布が一般極値分布 $GEV(\mu, \sigma, \xi)$ で近似されるとき, 順序統計量の上位 r 個 $(X_{(n:n)}, X_{(n-1:n)}, \dots, X_{(n-r+1:n)})$ の同時分布は, 次の同時密度関数を持つ分布で近似される:

$z^{(1)} \geq z^{(2)} \geq \dots \geq z^{(r)}$ として, $\xi \neq 0$ のとき,

$$f_{\xi}(z^{(1)}, \dots, z^{(r)}) = \exp \left\{ - \left[1 + \xi \left(\frac{z^{(r)} - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \times \prod_{k=1}^r \frac{1}{\sigma} \left[1 + \xi \left(\frac{z^{(k)} - \mu}{\sigma} \right) \right]^{-1/\xi - 1},$$

ただし, $1 + \xi(z^{(k)} - \mu)/\sigma > 0$, $k = 1, 2, \dots, r$.

$\xi = 0$ のとき ,

$$f_0(z^{(1)}, \dots, z^{(r)}) = \exp \left\{ - \exp \left[- \left(\frac{z^{(r)} - \mu}{\sigma} \right) \right] \right\} \times \prod_{k=1}^r \frac{1}{\sigma} \exp \left[- \left(\frac{z^{(k)} - \mu}{\sigma} \right) \right].$$

この同時密度関数を持つ分布を $r\text{GEV}(\mu, \sigma, \xi)$ で表す .

同時密度関数には一般極値分布と同じパラメータ (μ, σ, ξ) しか含まれていない .

強定常時系列

$\{X_i\}_{i=1}^{\infty}$ を周辺分布が F の強定常時系列とする：
任意の i_1, \dots, i_n と $k, n \in \mathbb{N} = \{1, 2, \dots\}$ に対して

$$(X_{i_1}, \dots, X_{i_n}) \stackrel{d}{=} (X_{i_1+k}, \dots, X_{i_n+k}).$$

$\{\tilde{X}_i\}_{i=1}^{\infty}$ を独立で同一分布 F に従う確率変数列とし、

$$Z_n = \max\{X_1, \dots, X_n\}, \quad \tilde{Z}_n = \max\{\tilde{X}_1, \dots, \tilde{X}_n\},$$

とおく。

定理 . 任意の $i < j < i' < j'$ に対して

$$\max \{X_i, \dots, X_j\} \quad \text{と} \quad \max \{X_{i'}, \dots, X_{j'}\}$$

が $i' - j \rightarrow \infty$ のときに漸近的に独立になるならば , ある数列 $a_n > 0$, $b_n \in \mathbb{R}$ に対して

$$\lim_{n \rightarrow \infty} P \{ (\tilde{Z}_n - b_n) / a_n \leq x \} = G(x)$$

\iff

$$\lim_{n \rightarrow \infty} P \{ (Z_n - b_n) / a_n \leq x \} = G^\theta(x), \quad \exists \theta \in (0, 1].$$

定理のパラメータ θ を極値指数 (extremal index) という . 強定常時系列では , 大きな値の X_i に続く値が大きくなりがちであるために , 極大値の塊が生じる (clustering) . この極大値の塊の平均個数は $1/\theta$ になる .

データ解析

推定

$\mathcal{F} = \{f(x; \theta) : \theta \in \Theta\}$: 母集団分布の統計モデル

$f(x; \theta)$: パラメータ θ を持つ関数形が既知の密度関数

統計モデルからの確率変数 X の n 個の確率標本を $\{X_1, X_2, \dots, X_n\}$ とし, その実現値 (データ) を $\{x_1, x_2, \dots, x_n\}$ とする .

k 次元パラメータ $\theta = (\theta_1, \theta_2, \dots, \theta_k) \in \Theta \subset \mathbb{R}^k$ の場合を考える .

最尤推定

尤度関数

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n f(x_i; \boldsymbol{\theta})$$

対数尤度関数は

$$l(\boldsymbol{\theta}) = \sum_{i=1}^n \log f(x_i; \boldsymbol{\theta})$$

パラメータ $\boldsymbol{\theta}$ を最尤法で推定する。次の連立方程式の解が最尤推定量：

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \theta_i} = 0, \quad i = 1, 2, \dots, k.$$

一般に，最尤推定量は漸近的に最適な推定量であり，尤度が書ければ最尤推定値を数値計算で求めることが出来る．

最尤推定量（値）を $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k)$ で表わす．

定理 5 . 分布族 \mathcal{F} についての適当な正則条件の下で最尤推定量は漸近正規推定量になる . すなわち , n が十分大のとき

$$\hat{\theta} \sim N_k(\theta, I_E(\theta)^{-1}).$$

ここで ,

$$I_E(\theta) = \begin{bmatrix} e_{11}(\theta) & e_{12}(\theta) & \cdots & e_{1k}(\theta) \\ e_{21}(\theta) & e_{22}(\theta) & \cdots & e_{2k}(\theta) \\ \vdots & \vdots & & \vdots \\ e_{k1}(\theta) & e_{k2}(\theta) & \cdots & e_{kk}(\theta) \end{bmatrix},$$

$$e_{ij}(\theta) = E \left\{ -\frac{\partial^2}{\partial \theta_i \partial \theta_j} l(\theta) \right\} = nE \left\{ -\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f(X; \theta) \right\}.$$

記号 \sim は “近似的に \cdots に従って分布する ” を表す .

行列 $I_E(\theta)$ を 期待情報行列 という .

一般に，行列 $I_E(\boldsymbol{\theta})$ は次の 観測情報行列 で $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ として推定する：

$$I_O(\boldsymbol{\theta}) = \begin{bmatrix} -\frac{\partial^2}{\partial\theta_1^2}l(\boldsymbol{\theta}) & -\frac{\partial^2}{\partial\theta_1\partial\theta_2}l(\boldsymbol{\theta}) & \cdots & -\frac{\partial^2}{\partial\theta_1\partial\theta_k}l(\boldsymbol{\theta}) \\ -\frac{\partial^2}{\partial\theta_2\partial\theta_1}l(\boldsymbol{\theta}) & -\frac{\partial^2}{\partial\theta_2^2}l(\boldsymbol{\theta}) & \cdots & -\frac{\partial^2}{\partial\theta_2\partial\theta_k}l(\boldsymbol{\theta}) \\ \vdots & \vdots & & \vdots \\ -\frac{\partial^2}{\partial\theta_k\partial\theta_1}l(\boldsymbol{\theta}) & -\frac{\partial^2}{\partial\theta_k\partial\theta_2}l(\boldsymbol{\theta}) & \cdots & -\frac{\partial^2}{\partial\theta_k^2}l(\boldsymbol{\theta}) \end{bmatrix} .$$

ここで，

$$-\frac{\partial^2}{\partial\theta_i\partial\theta_j}l(\boldsymbol{\theta}) = -\sum_{l=1}^n \frac{\partial^2}{\partial\theta_i\partial\theta_j} \log f(x_l; \boldsymbol{\theta}).$$

応用ではパラメータ θ の値ではなくその関数 $\phi = g(\theta)$ を推定したい場合がある .

特に極値統計学では分布 $f(x; \theta)$ の再現レベルと呼ばれる上側微少確率点の推定が目的の場合が多い .

ここでは , $\phi \in \mathbb{R}^1$ の場合を考える .

定理 6 . パラメータ θ の最尤推定量を $\hat{\theta}$ とし , $\phi = g(\theta)$ をスカラーとする . このとき , ϕ の最尤推定量は $\hat{\phi} = g(\hat{\theta})$ となる .

定理7 . (デルタ法) n が十分大のとき , パラメータ θ の最尤推定量を $\hat{\theta}$ とし , その漸近分散を V_{θ} とする . このとき , スカラー $\phi = g(\theta)$ の最尤推定量は

$$\hat{\phi} = g(\hat{\theta}) \sim N(\phi, V_{\phi})$$

となる . ただし $V_{\phi} = \nabla\phi^T V_{\theta} \nabla\phi$ で

$$\nabla\phi = \left[\frac{\partial\phi}{\partial\theta_1}, \dots, \frac{\partial\phi}{\partial\theta_k} \right]^T$$

として , $\hat{\theta}$ で求める .

プロフィール尤度関数による推測

$\theta = (\theta_1, \dots, \theta_k)$ の対数尤度を形式的に $l(\theta_i, \theta_{-i})$ で表す。
 θ_{-i} はパラメータ θ から θ_i を除いたすべての成分からなるもの。

パラメータ θ_i の プロファイル対数尤度 (profile log-likelihood)

$$l_p(\theta_i) = \max_{\theta_{-i}} l(\theta_i, \theta_{-i})$$

一般化：パラメータ θ を2つの成分 $(\theta^{(1)}, \theta^{(2)})$ に分割． $\theta^{(1)}$ は k_1 次元で
 $\theta^{(2)}$ は $k - k_1$ 次元． $\theta^{(1)}$ のプロフィール対数尤度

$$l_p(\theta^{(1)}) = \max_{\theta^{(2)}} l(\theta^{(1)}, \theta^{(2)}).$$

定理 8 . $\hat{\theta}$ を k 次元パラメータ $\theta = (\theta^{(1)}, \theta^{(2)})$ の最尤推定量とする . ただし , $\theta^{(1)}$ は θ の k_1 次元の成分とする . このとき , 適当な正則条件の下で , n が十分大のとき

$$D_p(\theta^{(1)}) = 2 \left\{ l(\hat{\theta}) - l_p(\theta^{(1)}) \right\} \sim \chi_{k_1}^2$$

が成立する .

応用 1 : 信頼区間

1次元の成分 θ_i に対する, 近似的に信頼係数 $(1 - \alpha)$ の信頼区間

$$\left\{ \theta_i : D_p(\theta_i) \leq \chi_1^2(\alpha) \right\} = \left\{ \theta_i : \max_{\boldsymbol{\theta}_{-i}} l(\theta_i, \boldsymbol{\theta}_{-i}) \geq l(\hat{\boldsymbol{\theta}}) - \chi_1^2(\alpha)/2 \right\}$$

$\chi_1^2(\alpha)$ は χ_1^2 分布の上側 α 確率点である .

応用 2 : モデル選択

パラメータ θ のモデルを \mathcal{M}_1 とし, この部分モデルの \mathcal{M}_0 は θ の k_1 個の成分を制約した (例えば 0 とおいた) ものとする . パラメータを $\theta = (\theta^{(1)}, \theta^{(2)})$ と分割 . モデル \mathcal{M}_0 では最初の k_1 次元の成分は 0 . ここで $D_p(\theta^{(1)}) > \chi_{k_1}^2(\alpha)$ のとき有意水準 α でモデル \mathcal{M}_1 を採用する .

モーメント法とPWM法

これらの方法により得られた推定量は漸近的な精度には問題がある。また、分布のパラメータが時間と共に変化するというような構造を入れた場合の推定が困難。しかし、最尤推定値を求める際のニュートン法の初期値として用いられることがある。

モーメント法 分布 $F(x; \theta)$ を持つ確率変数 X のモーメント ($E(X^p)$ または $V(X)$ で θ の関数になる) とデータのモーメントが等しいとおいて、パラメータ θ の推定値を作る。2次元のパラメータの場合は、次の平均 \bar{x} と分散 s^2 が用いられることが多い:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

PWM法 分布 $F(x) = F(x; \theta)$ を持つ確率変数 X の確率加重積率 (probability-weighted moment, PWM) は, 実数 p, r, s に対して

$$M_{p, r, s} = E [X^p \{F(X)\}^r \{1 - F(X)\}^s] = \int_0^1 (F^{-1}(u))^p u^r (1 - u)^s du$$

で定義される .

この推定法は分布関数の逆関数が陽に求まるとき使える . PWM $M_{1, r, s}$ を用いると通常モーメント法の場合よりもパラメータとモーメントの関係が簡単になる場合がある . 式 $u^r(1-u)^s$ は u または $1-u$ のべき乗の線形結合で表せるから $M_{1, r, 0}$ または $M_{1, 0, s}$ を用いればよい .

$$\alpha_r := M_{1, r, 0} = E [X \{F(X)\}^r], \quad r = 0, 1, 2, \dots$$

$$\beta_s := M_{1, 0, s} = E [X \{1 - F(X)\}^s], \quad s = 0, 1, 2, \dots$$

を考える . $r = s = 0$ のときは $\alpha_0 = \beta_0 = E(X)$.

α_r, β_s の推定 : $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ をデータの順序統計量として

$$a_r = \frac{1}{n} \sum_{i=1}^n \frac{(i-1)(i-2)\cdots(i-r)}{(n-1)(n-2)\cdots(n-r)} x_{(i)}, \quad r = 1, 2, \dots$$

$$b_s = \frac{1}{n} \sum_{i=1}^n \frac{(n-i)(n-i-1)\cdots(n-i-s+1)}{(n-1)(n-2)\cdots(n-s)} x_{(i)}, \quad s = 1, 2, \dots$$

または (ただし γ と δ は適当な定数)

$$\hat{\alpha}_r = \frac{1}{n} \sum_{i=1}^n p_{i,n}^r x_{(i)}, \quad r = 0, 1, \dots$$

$$\hat{\beta}_s = \frac{1}{n} \sum_{i=1}^n (1 - p_{i,n})^s x_{(i)}, \quad s = 0, 1, \dots$$

$$p_{i,n} = \frac{i + \gamma}{n + \delta}, \quad i = 1, 2, \dots, n.$$

モデル診断

推測結果は適合したモデルに依存する．モデルの適合を調べる必要がある．推測に用いたデータの適合をみてモデルの妥当性を判断する．

$\{x_1, x_2, \dots, x_n\}$: 未知の母集団分布 F からの n 個のデータ

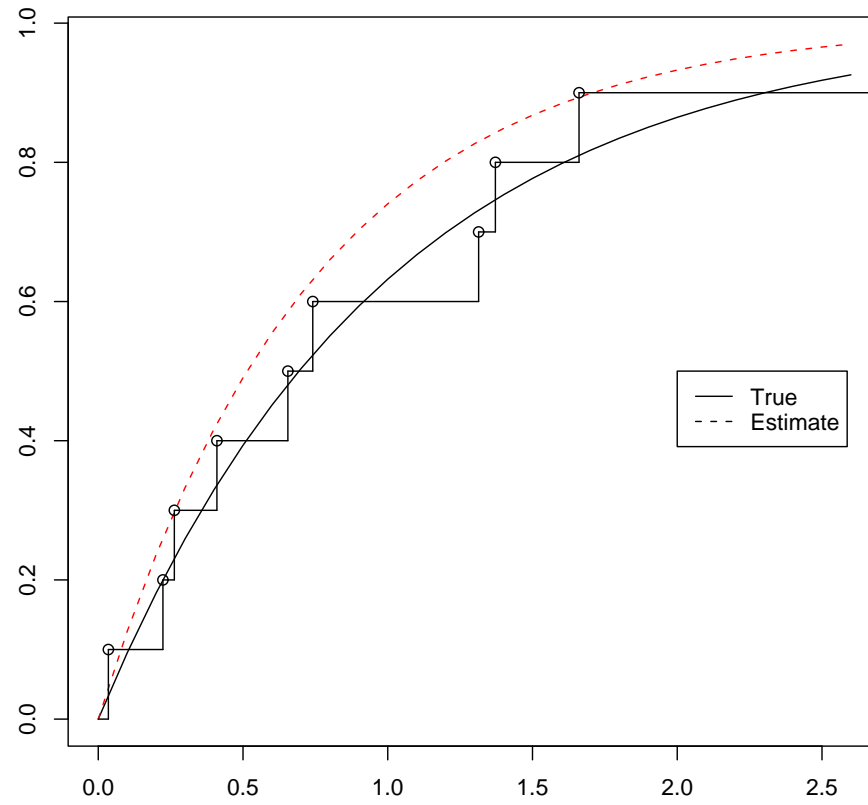
$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$: 順序統計量

\hat{F} : 統計モデルの下で F を (例えば最尤法で) 推定したもの

定義 3 . 分布 F の経験分布関数は ,

$$\tilde{F}(x) = \frac{1}{n+1} \sum_{i=1}^n I_{(-\infty, x]}(x_i)$$

である .



標準指数分布 H_0 , $GP(1, 0)$, の経験分布関数 .

定義4 . 確率プロット (probability plot , PP plot) は , 次の点をプロットしたものである :

$$\left\{ \left(\hat{F}(x_{(i)}), \frac{i}{n+1} \right) : i = 1, \dots, n \right\}.$$

定義5 . 確率点プロット (quantile plot , QQ plot) は次の点をプロットしたものである :

$$\left\{ \left(\hat{F}^{-1} \left(\frac{i}{n+1} \right), x_{(i)} \right) : i = 1, \dots, n \right\}.$$

ここで , $\hat{F}^{-1}(i/(n+1))$, $x_{(i)}$ とともに F の $i/(n+1)$ 確率点の推定値である .

極値理論の下での推測

一般極値 (GEV) モデル

区分最大値データ $\{z_1, z_2, \dots, z_n\}$ に一般極値分布 $\text{GEV}(\mu, \sigma, \xi)$,

$$G(z) = \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} = G_\xi \left(\frac{z - \mu}{\sigma} \right), \quad 1 + \xi(z - \mu)/\sigma > 0,$$

を適合させる場合の推測 .

極値統計学では , パラメータ (μ, σ, ξ) の直接の推定よりも , 次の極値分布の上側 (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 $1 - p$ 確率点 z_p , すなわち

$$G_{\xi} \left(\frac{z_p - \mu}{\sigma} \right) = 1 - p$$

は次のように表される :

$$z_p = \begin{cases} \mu + \sigma \left[\left\{ -\log(1 - p) \right\}^{-\xi} - 1 \right] / \xi, & \xi \neq 0, \\ \mu + \sigma \left[-\log \left\{ -\log(1 - p) \right\} \right], & \xi = 0. \end{cases}$$

この確率点 z_p は再現期間 (return period) $1/p$ の再現レベル (return level) という . 例えば年最大値データを扱うとき , 再現期間 $1/p = 200$ 年の再現レベル $z_{1/200}$ は 200年に平均1度現れる様な (大きな) 値と解釈できる .

一般に $p \ll 1/n$ 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 .

$y_p = -\log(1 - p)$ とおくと ,

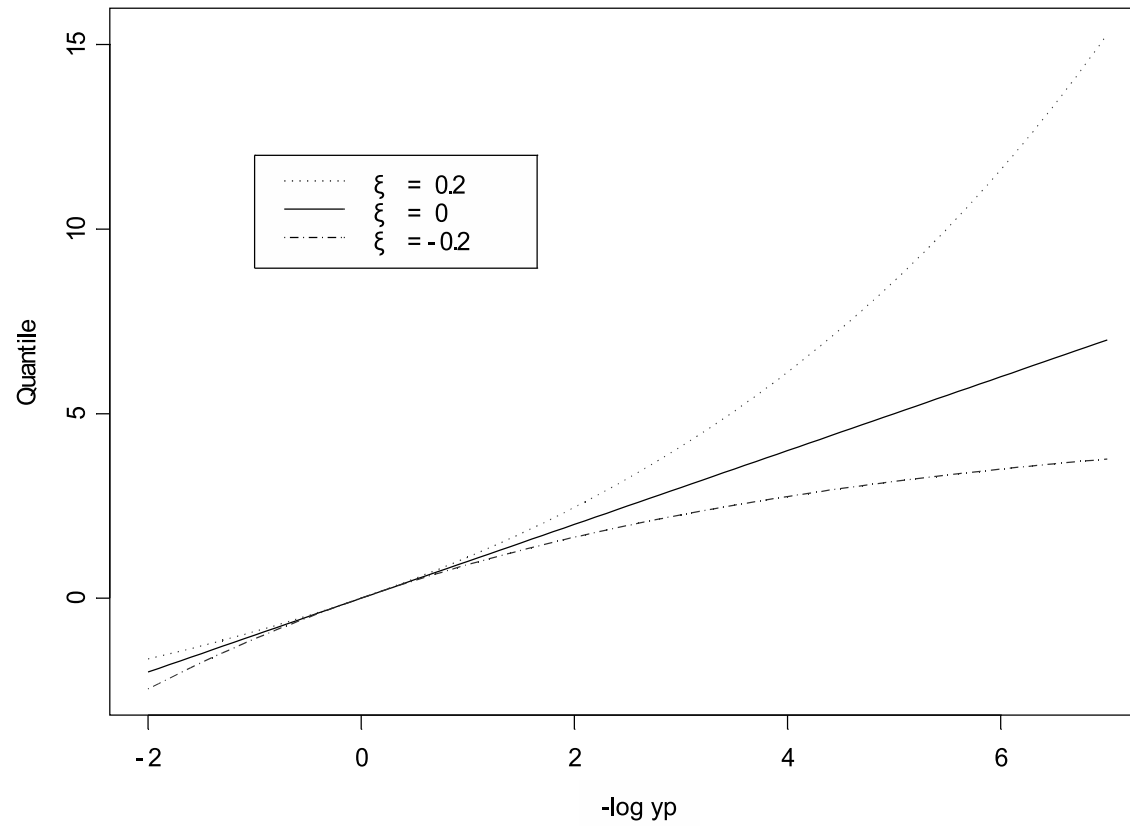
$$z_p = \begin{cases} \mu + \sigma(y_p^{-\xi} - 1)/\xi, & \xi \neq 0, \\ \mu + \sigma(-\log y_p), & \xi = 0, \end{cases}$$

となる . このとき

$$\{(-\log y_p, z_p); 0 < p < 1\}$$

をプロットしたものを再現レベルプロット (return level plot) という .

再現レベルプロットは , $\xi < 0$ のときは凹 (concave) で $p \rightarrow 0$ ($-\log y_p \rightarrow \infty$) のとき上限 $\mu - \sigma/\xi$ に近づく , $\xi = 0$ のときは直線で , $\xi > 0$ のときは凸 (convex) で $p \rightarrow 0$ のとき ∞ に発散する .



GEV(0, 1, ξ) , $\xi = -0.2, 0, 0.2$ の再現レベルプロット .

最尤法 GEV(μ, σ, ξ)

対数尤度： $\xi \neq 0$ のとき

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

$$1 + \xi(z_i - \mu)/\sigma > 0, \quad i = 1, \dots, n,$$

$\xi = 0$ のとき

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

対数尤度を最大にする最尤推定値 $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$ または $(\hat{\mu}, \hat{\sigma})$ を求める .

GEV(μ, σ, ξ) 期待情報行列 (Prescott and Walden, 1980) :

$$\frac{n}{\sigma^2 \xi^2} \begin{pmatrix} \xi^2 p & -\xi \{p - 2\Gamma(2 + \xi)\} & \sigma \xi \left(q - \frac{p}{\xi} \right) \\ " & 1 - 2\Gamma(2 - \xi) + p & \sigma \left[1 - \gamma + \frac{1 - \Gamma(2 + \xi)}{\xi} - q + \frac{p}{\xi} \right] \\ " & " & \sigma^2 \left[\frac{\pi^2}{6} + \left(1 - \gamma + \frac{1}{\xi} \right)^2 - \frac{2q}{\xi} + \frac{p}{\xi^2} \right] \end{pmatrix}$$

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 $p = (1 + \xi)^2 \Gamma(1 + 2\xi)$, $q = \Gamma(2 + \xi) \{ \Psi(1 + \xi) + (1 + \xi) / \xi \}$, $\gamma = 0.5772157\dots$ Euler の定数 .

分布族 $\{\text{GEV}(\mu, \sigma, \xi), -\infty < \mu < \infty, \sigma > 0, -\infty < \xi < \infty\}$: 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 (Smith, 1985) .

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 .

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ことが出来る . これから , (μ, σ, ξ) の各パラメータの信頼区間を求める .

再現レベル z_p の最尤推定値は $y_p = -\log(1 - p)$ として ,

$$\hat{z}_p = \begin{cases} \hat{\mu} + \hat{\sigma} \{ y_p^{-\hat{\xi}} - 1 \} / \hat{\xi}, & \hat{\xi} \neq 0, \\ \hat{\mu} + \hat{\sigma} \{ -\log y_p \}, & \hat{\xi} = 0. \end{cases}$$

デルタ法より , 分散は $V(\hat{z}_p) \approx \nabla z_p^T V \nabla z_p$ となる .
ここで V は $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$ の分散共分散行列で

$$\begin{aligned} \nabla z_p^T &= \left[\frac{\partial z_p}{\partial \mu}, \frac{\partial z_p}{\partial \sigma}, \frac{\partial z_p}{\partial \xi} \right] \\ &= \left[1, (y_p^{-\xi} - 1) / \xi, \sigma y_p^{-\xi} (-\log y_p) / \xi - \sigma (y_p^{-\xi} - 1) / \xi^2 \right] \end{aligned}$$

を $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$ で求める .

$\hat{\xi} < 0$ の場合は 分布の上端点 $z_0 = \mu - \sigma/\xi$ の推定

$$\hat{z}_0 = \hat{\mu} - \hat{\sigma}/\hat{\xi}$$

が重要 . この場合は

$$\nabla z_p^T = \left[1, -1/\xi, \sigma/\xi^2 \right]$$

となる .

一般に, ξ や z_p 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.

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ればよい.

ξ の近似的に信頼係数 95% の信頼区間は

$$\begin{aligned} & \left\{ \xi : 2 \left\{ l(\hat{\mu}, \hat{\sigma}, \hat{\xi}) - \max_{\mu, \sigma} l(\mu, \sigma, \xi) \right\} \leq \chi_1^2(0.05) = 3.841 \right\} \\ & = \left\{ \xi : \max_{\mu, \sigma} l(\mu, \sigma, \xi) \geq l(\hat{\mu}, \hat{\sigma}, \hat{\xi}) - 1.921 \right\} \end{aligned}$$

となる.

再現レベル z_p のプロファイル尤度を求めるために、

$$\mu = z_p - \sigma \left[y_p^{-\xi} - 1 \right] / \xi$$

から、パラメータ (μ, σ, ξ) を (z_p, σ, ξ) へ変える。このとき、対数尤度は

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

と表される。これから、 z_p の近似的に 95% の信頼区間は

$$\begin{aligned} & \left\{ z_p : 2 \left\{ l(\hat{z}_p, \hat{\sigma}, \hat{\xi}) - \max_{\sigma, \xi} l(z_p, \sigma, \xi) \right\} \leq \chi_1^2(0.05) \right\} \\ & = \left\{ z_p : \max_{\sigma, \xi} l(z_p, \sigma, \xi) \geq l(\hat{z}_p, \hat{\sigma}, \hat{\xi}) - \chi_1^2(0.05)/2 \right\} \end{aligned}$$

となる。

PWM法

次の PWM を考える (Hosking et al., 1985) :

$$\alpha_r := M_{1,r,0} = E[X\{G(X)\}^r], \quad r = 0, 1, 2, \dots$$

一般極値分布の逆関数は ,

$$G^{-1}(u) = \begin{cases} \mu + \sigma \{(-\log u)^{-\xi} - 1\} / \xi, & \xi \neq 0, \\ \mu + \sigma \log(-\log u), & \xi = 0, \end{cases}$$

で , $\xi \neq 0$ のとき PWM は ,

$$\alpha_r = \int_0^1 G^{-1}(u) u^r du = \frac{\mu + \sigma \{(r+1)^\xi \Gamma(1-\xi) - 1\} / \xi}{r+1}, \quad \xi < 1.$$

これから

$$\begin{cases} \alpha_0 & = \mu + \sigma \{ \Gamma(1 - \xi) - 1 \} / \xi, \\ 2\alpha_1 - \alpha_0 & = \sigma(2^\xi - 1) \Gamma(1 - \xi) / \xi, \\ (3\alpha_2 - \alpha_0) / (2\alpha_1 - \alpha_0) & = (3^\xi - 1) / (2^\xi - 1). \end{cases}$$

1, 2 番目の式から,

$$\sigma = \frac{(2\alpha_1 - \alpha_0) \xi}{\Gamma(1 - \xi)(2^\xi - 1)}, \quad \mu = \alpha_0 + \frac{\sigma \{ 1 - \Gamma(1 - \xi) \}}{\xi}.$$

3 番目の式で, $-1/2 < \xi < 1/2$ のとき

$$\xi = 7.8590 c - 2.9554 c^2, \quad c = \frac{\log 2}{\log 3} - \frac{2\alpha_1 - \alpha_0}{3\alpha_2 - \alpha_0}$$

とすれば, その誤差は 0.0009 以下.

データの順序統計量を $z_{(1)} \leq z_{(2)} \leq \cdots \leq z_{(n)}$ とするとき,

$$\hat{\alpha}_r = \frac{1}{n} \sum_{i=1}^n \left(\frac{i - 0.35}{n} \right)^r z_{(i)}, \quad r = 0, 1, 2$$

で PWM's の α_r , $r = 0, 1, 2$ を推定. これらを用いて PWM 推定量 $\hat{\xi}_*$, $\hat{\sigma}_*$, $\hat{\mu}_*$ が得られる. 例えば,

$$\hat{\xi}_* = 7.8590 c - 2.9554 c^2, \quad c = \frac{\log 2}{\log 3} - \frac{2\hat{\alpha}_1 - \hat{\alpha}_0}{3\hat{\alpha}_2 - \hat{\alpha}_0},$$

である.

検定問題

次の検定を考える：

$$H_0 : \xi = 0 \quad \text{vs.} \quad H_1 : \xi \neq 0$$

極値データ $\{z_1, z_2, \dots, z_n\}$ が Gumbel 分布からの標本と見なせるかどうかの検定。

PWM 推定量 $\hat{\xi}_*$ は $H_0 : \xi = 0$ の下で漸近的に正規分布 $N(0, 0.5633/n)$ に従うから、

$$\frac{\hat{\xi}_*}{\sqrt{0.5633/n}} \sim N(0, 1).$$

漸近的に有意水準 10% の検定は

$$\left| \hat{\xi}_* / \sqrt{0.5633/n} \right| > 1.645 \iff H_0 \text{ reject}$$

となる . Hosking et al. (1985) のシミュレーション実験によると , この検定方式の結果は良好である .

Hosking (1984) は , この検定問題に関して種々の検定法の検出力をシミュレーション実験で調べた結果 , Bartlett 修正をした尤度比検定がベストであったと述べている .

モデル診断

GEVモデルでは $F \in MAD(G_\xi)$ を仮定するが，これをデータから検証することは出来ない．

モデルとデータが適合しているかどうかに関する以下の様な診断を行う必要はある．

確率プロット

$$\left\{ \left(\hat{G}(z_{(i)}), \frac{i}{n+1} \right) : i = 1, \dots, n \right\}$$

ここで，

$$\hat{G}(z_{(i)}) = \exp \left\{ - \left[1 + \hat{\xi} \left(\frac{z_{(i)} - \hat{\mu}}{\hat{\sigma}} \right) \right]^{-1/\hat{\xi}} \right\}.$$

確率点プロット

$$\left\{ \left(\hat{G}^{-1} \left(\frac{i}{n+1} \right), z_{(i)} \right) : i = 1, \dots, n \right\}$$

ここで,

$$\hat{G}^{-1} \left(\frac{i}{n+1} \right) = \hat{\mu} + \hat{\sigma} \left[\left\{ -\log \left(\frac{i}{n+1} \right) \right\}^{-\hat{\xi}} - 1 \right] / \hat{\xi}.$$

再現レベルプロット $0 < p < 1$ に対して, \hat{z}_p を再現レベル z_p の最尤推定値として

$$\left\{ (-\log y_p, \hat{z}_p) : 0 < p < 1 \right\}$$

データのヒストグラムと推定した密度関数を描いた図もモデル診断に利用できる。

問題点

区分最大法 (block maximum method)

「ブロックの大きさをどう決めるか？」

ブロックを小さくすると利用データ数は多くなり、分散は小さくなるが一般極値分布への適合は悪くなり偏りは大きくなる。

ブロックを大きくすると利用データ数は少なくなり、一般極値分布への適合は良くなり偏りは小さくなるが分散は大きくなる。

ブロックの大きさはこれらの事を考慮し現実に応じて決める必要がある。

この方法には最大値データのみを用いることから「それ以外のデータの情報を用いない」という欠点と、一般に「データ数が少ない」という問題がある。

非定常の場合の極値 パラメータが時間に依存するモデル

GEVモデル

$$Z_t \sim \text{GEV}(\mu(t), \sigma(t), \xi(t)), \quad t = 1, 2, \dots, m.$$

ここで, たとえば $\mu(t)$ として

$$\mu(t) = \beta_0 + \beta_1 t \quad \text{または} \quad \mu(t) = \beta_0 + \beta_1 t + \beta_2 t^2$$

等を, また $\sigma(t)$ については

$$\sigma(t) = \exp(\beta_0^* + \beta_1^* t)$$

を考える.

パラメータが他の変量に依存する (共変量) モデルも同様に考えられる.

推測

GEVモデル

$$Z_t \sim \text{GEV}(\mu(t), \sigma(t), \xi(t)), \quad t = 1, 2, \dots, m.$$

$\mu(t)$, $\sigma(t)$, $\xi(t)$ は上の様に表されているとし, すべてのパラメータを β とする.

尤度は

$$L(\beta) = \prod_{t=1}^m \frac{1}{\sigma(t)} g_{\xi(t)} \left(\frac{z_t - \mu(t)}{\sigma(t)} \right).$$

$g_{\xi}(z)$ は $\text{GEV}(0, 1, \xi)$ の密度関数.

尤度関数を β に関して最大化し最尤推定値 $\hat{\beta}$ を求める.

2つのモデルに $\mathcal{M}_0 \subset \mathcal{M}_1$ の関係があるとき，漸近的に

$$2\left\{l_1(\mathcal{M}_1) - l_0(\mathcal{M}_0)\right\} \sim \chi_{k_1}^2$$

となる．ただし $l_i(\mathcal{M}_i)$ はモデル \mathcal{M}_i ($i = 0, 1$) の下での最大対数尤度で， k_1 はモデル \mathcal{M}_1 と \mathcal{M}_0 のパラメータの次元の差である．

直感的にパラメータが1個増えたとき対数尤度が2以上増えればそのパラメータは有意である．

モデル診断 **GEVモデル**

$$Z_t \sim \text{GEV}(\mu(t), \sigma(t), \xi(t)), \quad t = 1, 2, \dots, m,$$

このとき

$$Z_0 = \frac{1}{\xi(t)} \log \left\{ 1 + \xi(t) \left(\frac{Z_t - \mu(t)}{\sigma(t)} \right) \right\} \sim \text{GEV}(0, 1, 0) : \text{標準 Gumbel 分布}$$

$$P(Z_0 \leq z) = G_0(z) = \exp(-\exp(-z)), \quad -\infty < z < \infty.$$

ここで, z_1, z_2, \dots, z_m を観測値として, $\hat{\mu}(t), \hat{\sigma}(t), \hat{\xi}(t)$ を最尤推定値 $\hat{\beta}$ から求めたものとする

$$\tilde{z}_t = \frac{1}{\hat{\xi}(t)} \log \left\{ 1 + \hat{\xi}(t) \left(\frac{z_t - \hat{\mu}(t)}{\hat{\sigma}(t)} \right) \right\}, \quad t = 1, 2, \dots, m$$

は標準 Gumbel 分布からのデータと見なせる . これらを大きさの順に並べたものを

$$\tilde{z}_{(1)} \leq \tilde{z}_{(2)} \leq \dots \leq \tilde{z}_{(m)}$$

とする .

確率プロットは ,

$$\left\{ \left(G_0(\tilde{z}_{(i)}), \frac{i}{m+1} \right) : i = 1, \dots, m \right\}$$

となる . ただし

$$G_0(z_{(i)}) = \exp(-\exp(-\tilde{z}_{(i)})).$$

確率点プロットは

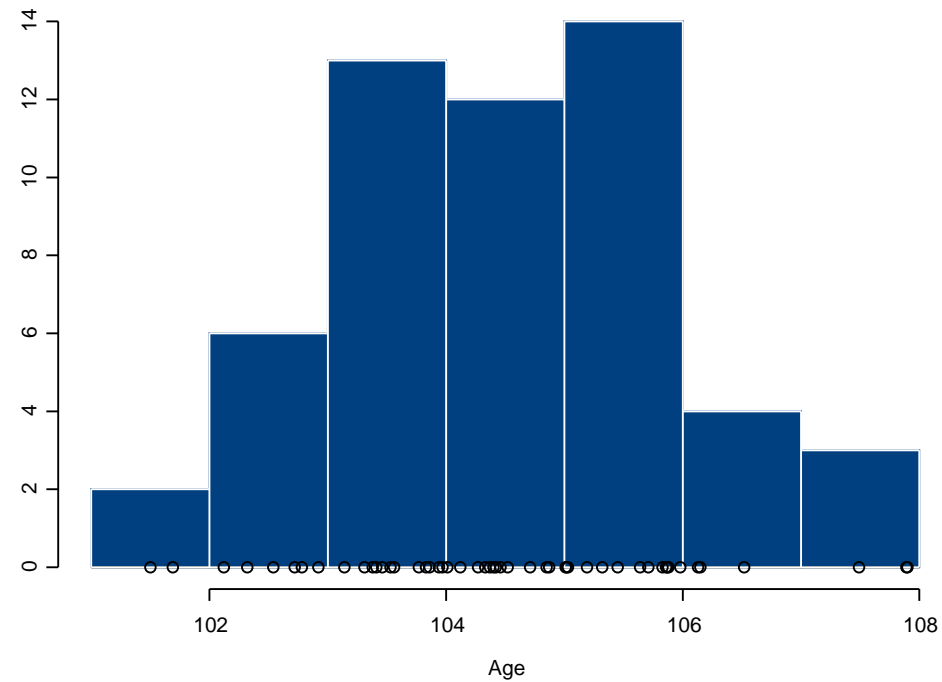
$$\left\{ \left(G_0^{-1} \left(\frac{i}{m+1} \right), \tilde{z}_{(i)} \right) : i = 1, \dots, m \right\}$$

となる . ただし

$$G_0^{-1} \left(\frac{i}{m+1} \right) = -\log(-\log(i/(m+1))).$$

確率プロットは Gumbel 分布の選択に関して不変であるが、確率点プロットはそうではない。

データ解析例 1



The yearly oldest ages at death in Sweden during the period from 1905 to 1958 for women. (54名, 101.50歳 ~ 107.90歳).

データに一般極値分布を適合して解析する .

最尤法でパラメータを推定 :

最大対数尤度 -95.904

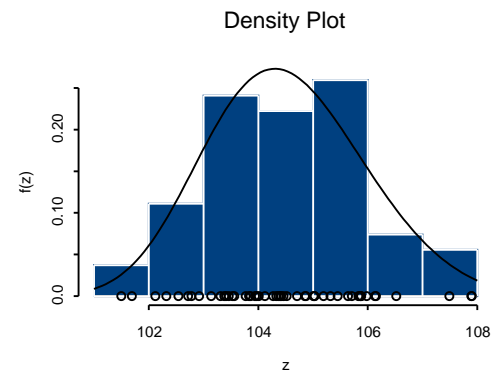
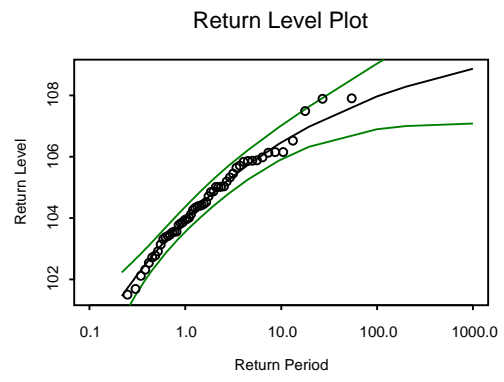
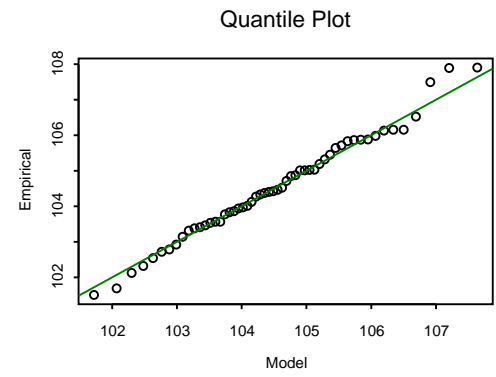
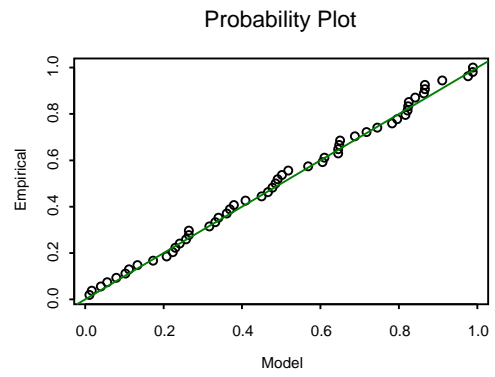
最尤推定値とその標準誤差

$$\hat{\mu} = 103.966 (0.209), \quad \hat{\sigma} = 1.3834 (0.147), \quad \hat{\xi} = -0.221 (0.0886).$$

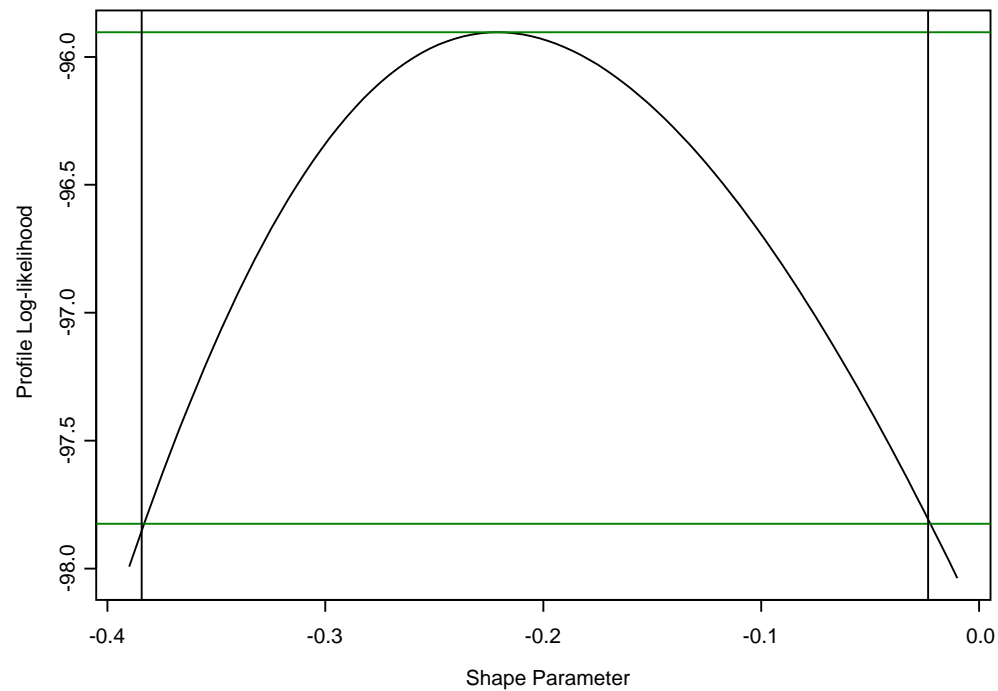
ξ の推定値が負である . このとき分布の上限は

$$\hat{\mu} - \hat{\sigma}/\hat{\xi} = 103.966 + 1.3834/0.221 = 110.226$$

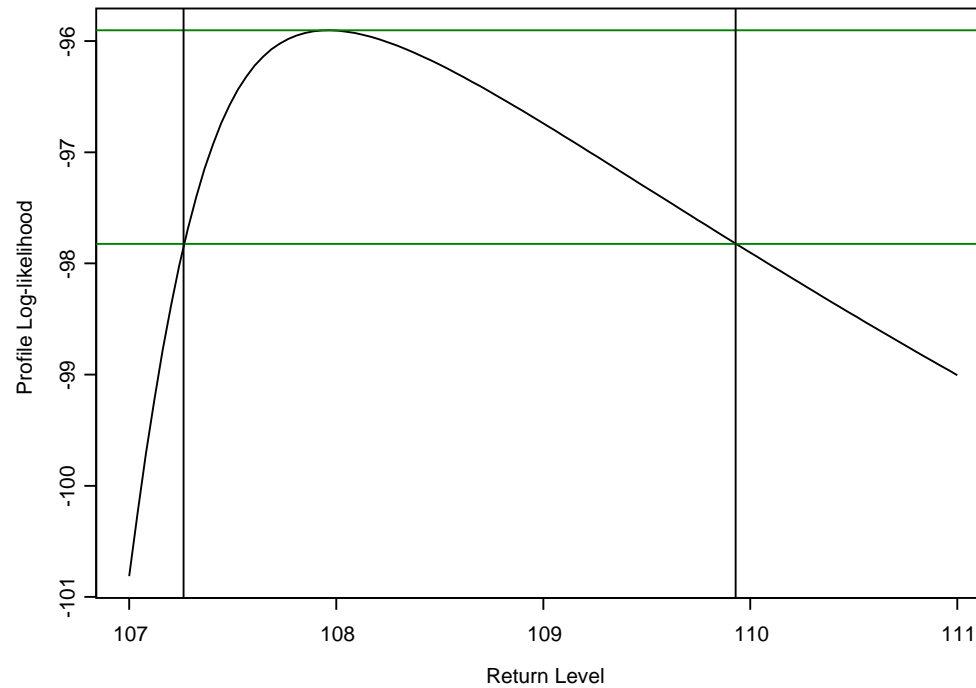
と推定される . ちなみにデータの最大値は 107.90 .



データ解析の診断 .



プロファイル尤度に基づく形状パラメータ ξ の 95%信頼区間 .
[$-0.384, -0.023$]. ($\hat{\xi} = -0.221 (0.0886)$)



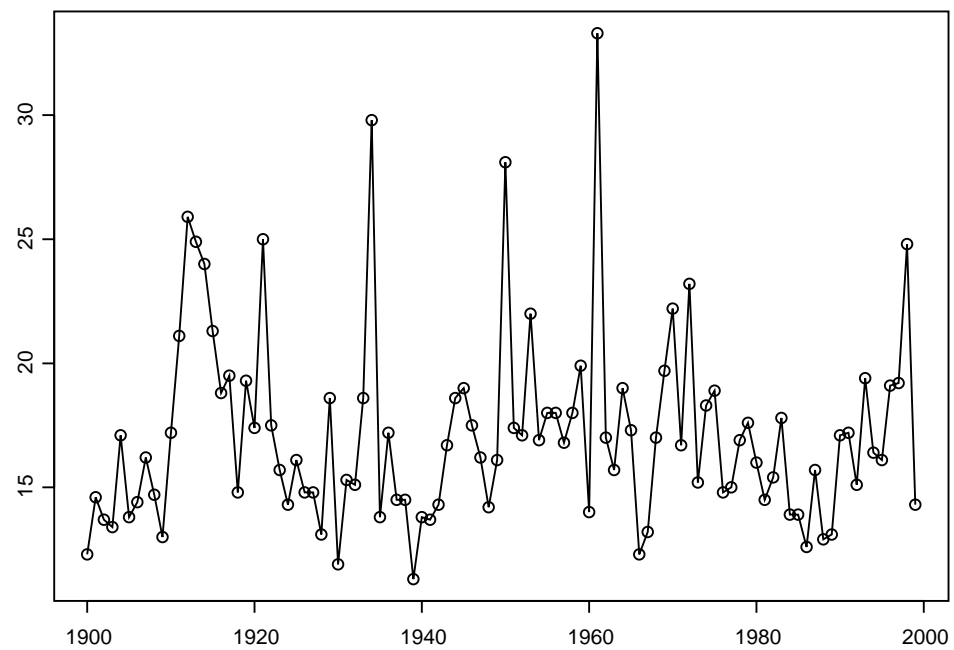
プロファイル尤度に基づく 100 年再現レベルの 95%信頼区間 .
[107.263, 109.929] ($\hat{z}_{1/100} = 107.96$)

データ解析例 2

大阪の1900年から1999年までの百年間の年最大風速(m/s)データ(大阪管区気象台編(1982), 石原他(2002))をGEVモデルで解析する。

気象データに関しては観測基準や観測機器が変わっていることがあるが, ここではこれらの影響を無視する。

気象データ等の品質に関しては木下武雄(2004)を参照。



大阪の年最大風速 (m/s) , 1900年 ~ 1999年 .
データの最大値は 33.3(m/s) .

一般極値分布 $GEV(\mu, \sigma, \xi)$ を適合させる .

最尤推定値は $(\hat{\mu}, \hat{\sigma}, \hat{\xi}) = (15.349, 2.550, 0.111)$, 最大対数尤度は -257.78 で , 分散共分散行列は

$$V = \begin{bmatrix} 0.08235 & 0.03024 & -0.00686 \\ 0.03024 & 0.04703 & -0.00231 \\ -0.00686 & -0.00231 & 0.00567 \end{bmatrix}$$

となった . この行列の対角成分が $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$ の分散である .

位置パラメータの年への回帰モデル $\mu = \beta_0 + \beta_1 \text{年}$ を調べて見た . この場合の最大対数尤度は -257.352 . また , μ, σ 共に年に依存するモデルも調べてみたが最大対数尤度はほとんど増えなかった . したがって , このデータでは μ, σ とも年に依存しないとみなせる .

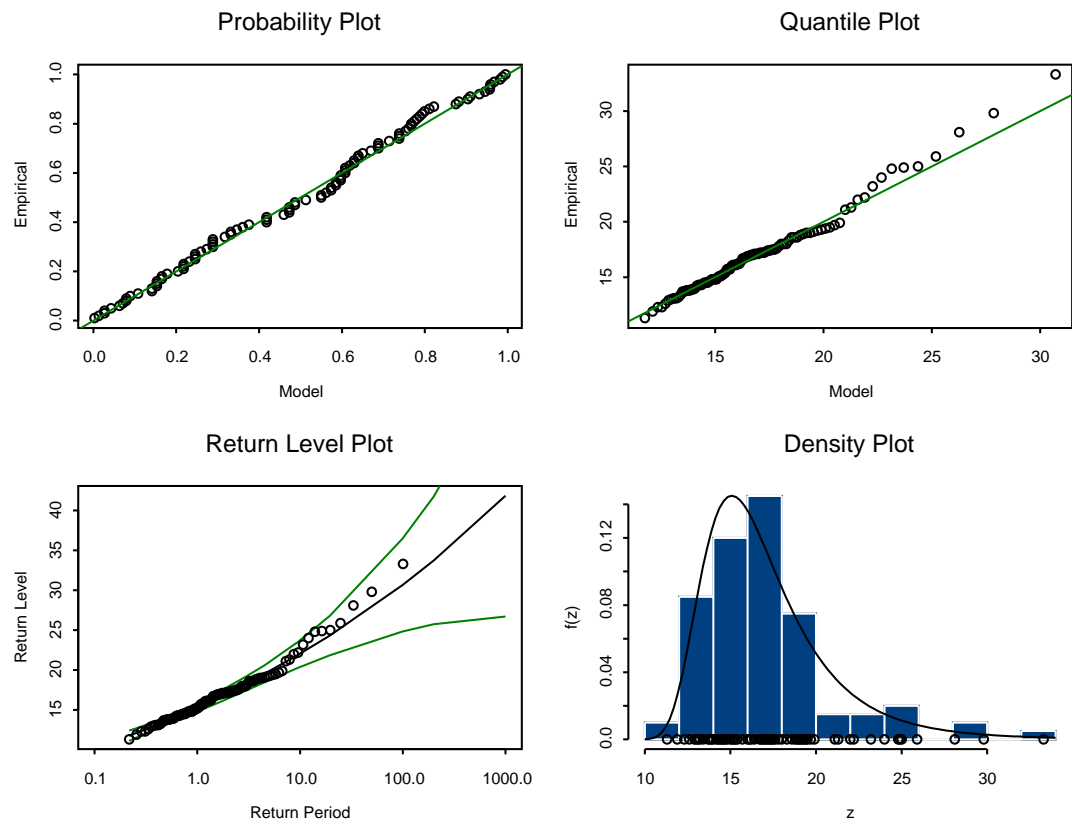


図5 大阪の年最大風速データ解析の診断。

大阪の年最大風速では $\hat{\xi} = 0.111 > 0$ で Fréchet 分布が適合する .

ξ の 95%信頼区間は , 最尤推定量の漸近正規性より

$$\xi \in 0.111 \pm 1.96\sqrt{0.00567} = [-0.037, 0.259].$$

プロファイル尤度に基づく 95%信頼区間は 図 6 より $[-0.021, 0.274]$.

プロファイル尤度に基づく信頼区間は右にずれている .

一般にプロファイル尤度に基づく信頼区間の方が精度が良い .

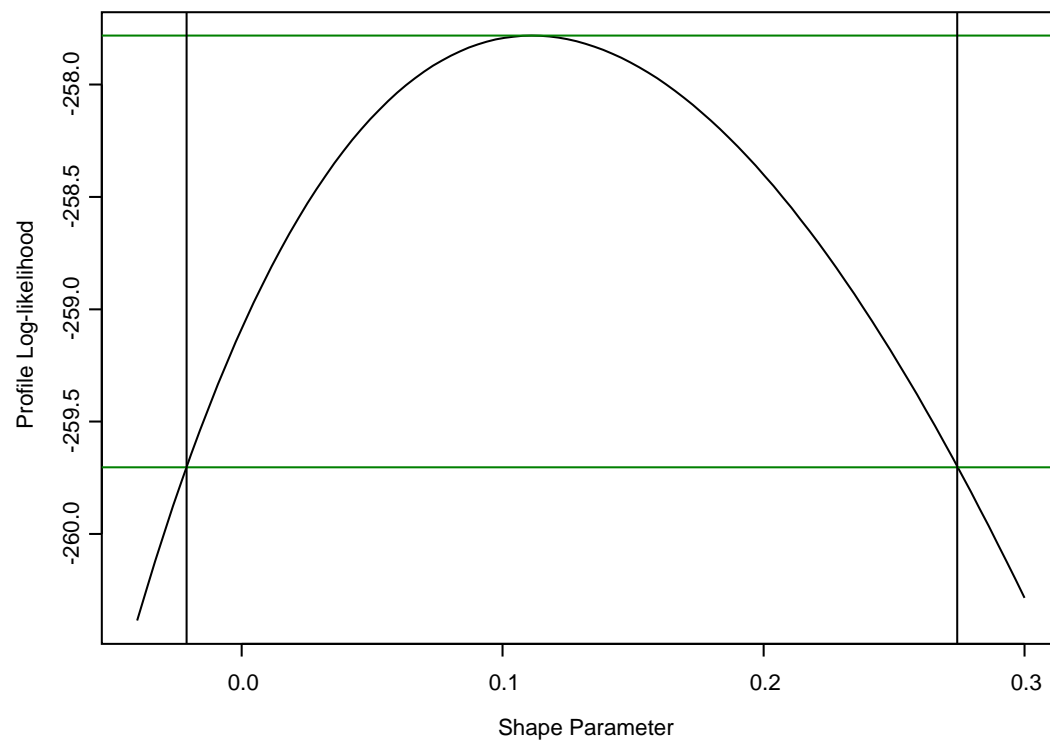


図6 大阪の年最大風速の形状パラメータ ξ のプロファイル尤度。
 $\hat{\xi} = 0.111 > 0$, 95%信頼区間は $[-0.021, 0.274]$. ($[-0.037, 0.259]$)

再現期間200年の再現レベル $z_{1/200}$ の最尤推定値は33.74で、漸近正規性に基づく95%信頼区間は

$$z_{1/200} \in \hat{z}_{1/200} \pm 1.96 \sqrt{\nabla z_{1/200}^T V \nabla z_{1/200}} = [25.75, 41.73].$$

プロファイル尤度に基づく95%信頼区間は図7より[28.29, 46.90] .

プロファイル尤度に基づく信頼区間は右にずれている .

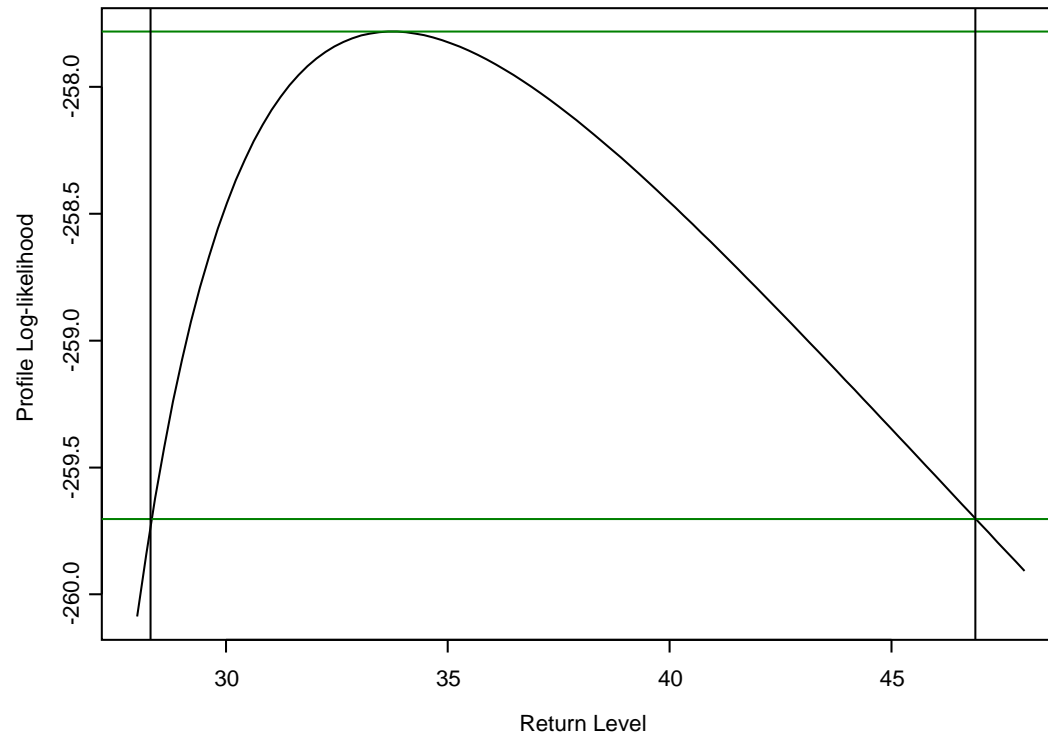
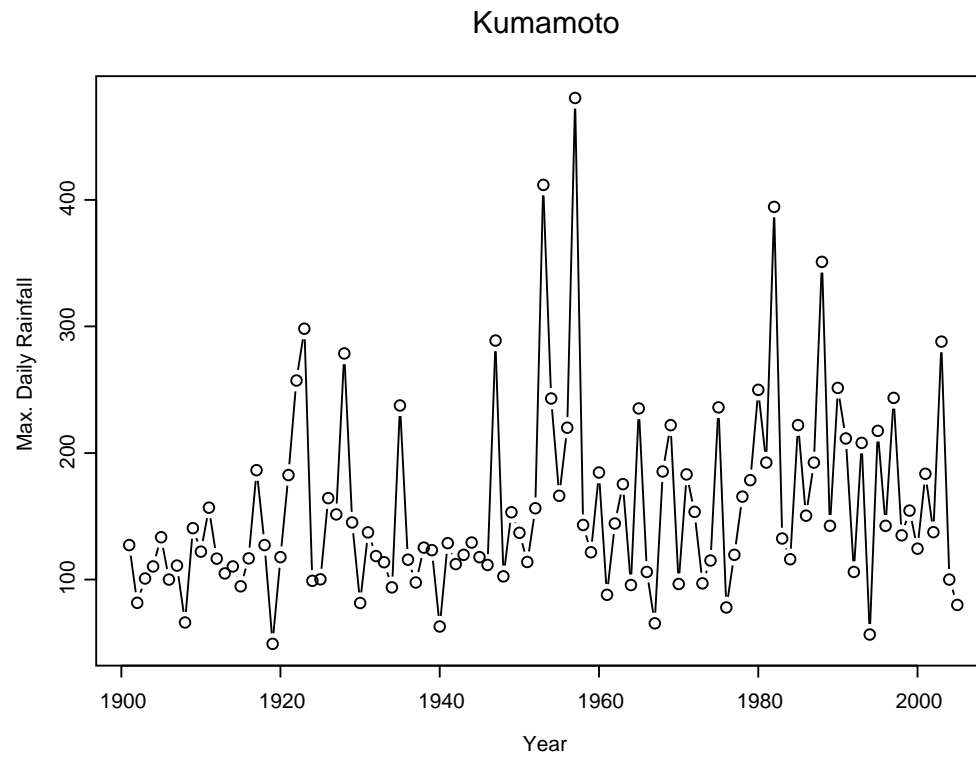


図7 大阪の年最大風速の200年再現レベルのプロファイル尤度 .
 $\hat{z}_{1/200} = 33.74$, 95%信頼区間は $[28.29, 46.90]$. ($[25.75, 41.73]$)

データ解析例 3



熊本の年最大日雨量 (0.1mm) , 1901年から2005年 .

非定常の GEV モデル

$$Z_t \sim \text{GEV}(\mu(t), \sigma(t), \xi), \quad t = 1, 2, \dots, m.$$

次のモデルを考えパラメータを最尤法で推定 .

モデル1 : $\mu(t) = \mu, \sigma(t) = \sigma.$

モデル2 : $\mu(t) = \beta_0 + \beta_1 t, \sigma(t) = \sigma.$

モデル3 : $\mu(t) = \beta_0 + \beta_1 t + \beta_2 t^2, \sigma(t) = \sigma.$

モデル4 : $\mu(t) = \beta_0 + \beta_1 t, \sigma(t) = \exp(\beta_2 + \beta_3 t).$

モデル5 : $\mu(t) = \beta_0 + \beta_1 t + \beta_2 t^2, \sigma(t) = \exp(\beta_3 + \beta_4 t).$

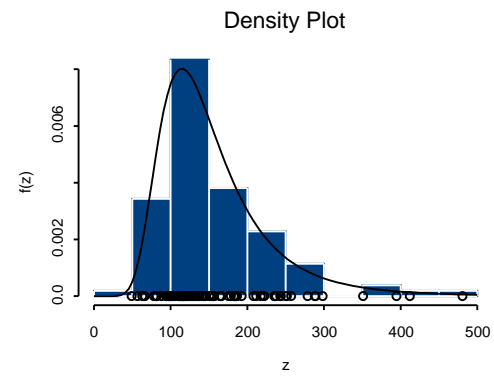
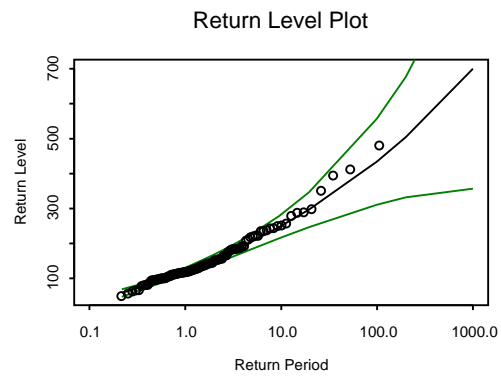
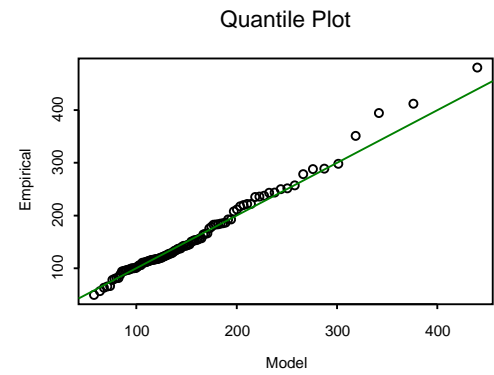
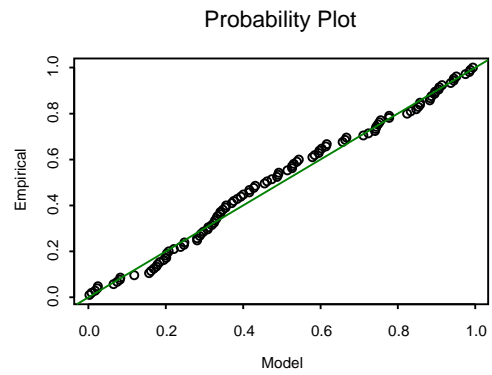
これらのモデルの下での最大対数尤度は次の様になった .

モデル1 :	-578.204	パラメータ 3 個	$\chi_1^2(0.05) = 3.841$
モデル2 :	-576.973	パラメータ 4 個	$\chi_2^2(0.05) = 5.991$
モデル3 :	-576.527	パラメータ 5 個	
モデル4 :	-573.376	パラメータ 5 個	
モデル5 :	-573.290	パラメータ 6 個	

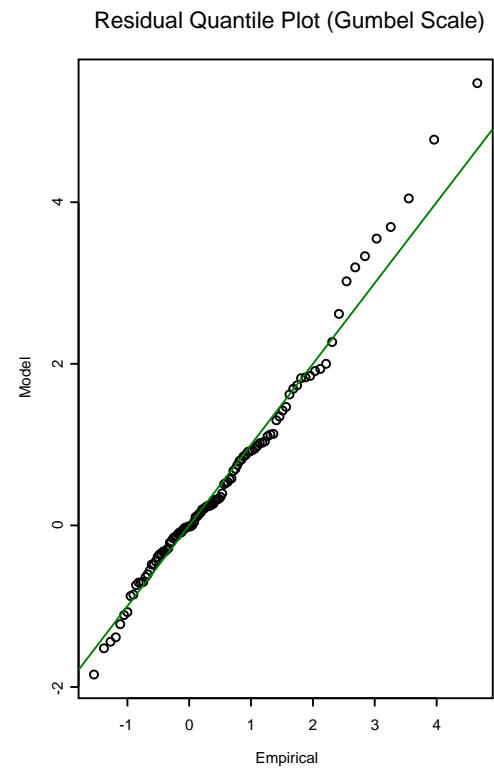
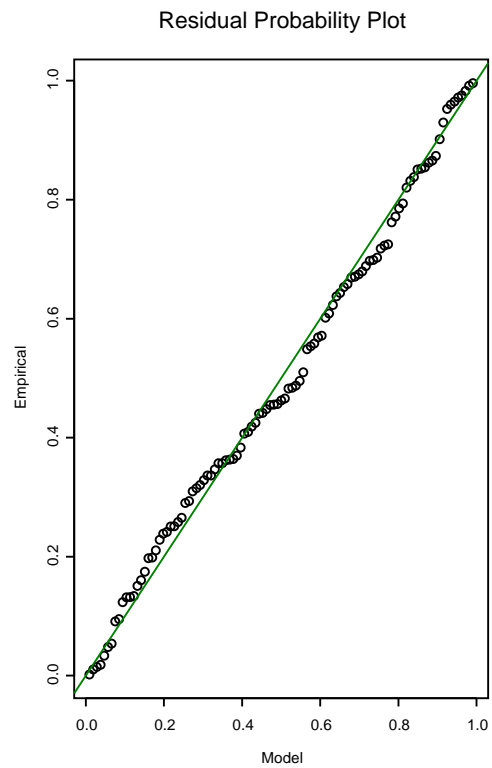
これからモデル4が選ばれる : すなわち熊本のデータでは

$$\hat{\mu}(t) = 123.85 + 23.33t, \quad \hat{\sigma}(t) = \exp(3.820 + 0.377t), \quad \hat{\xi} = 0.112.$$

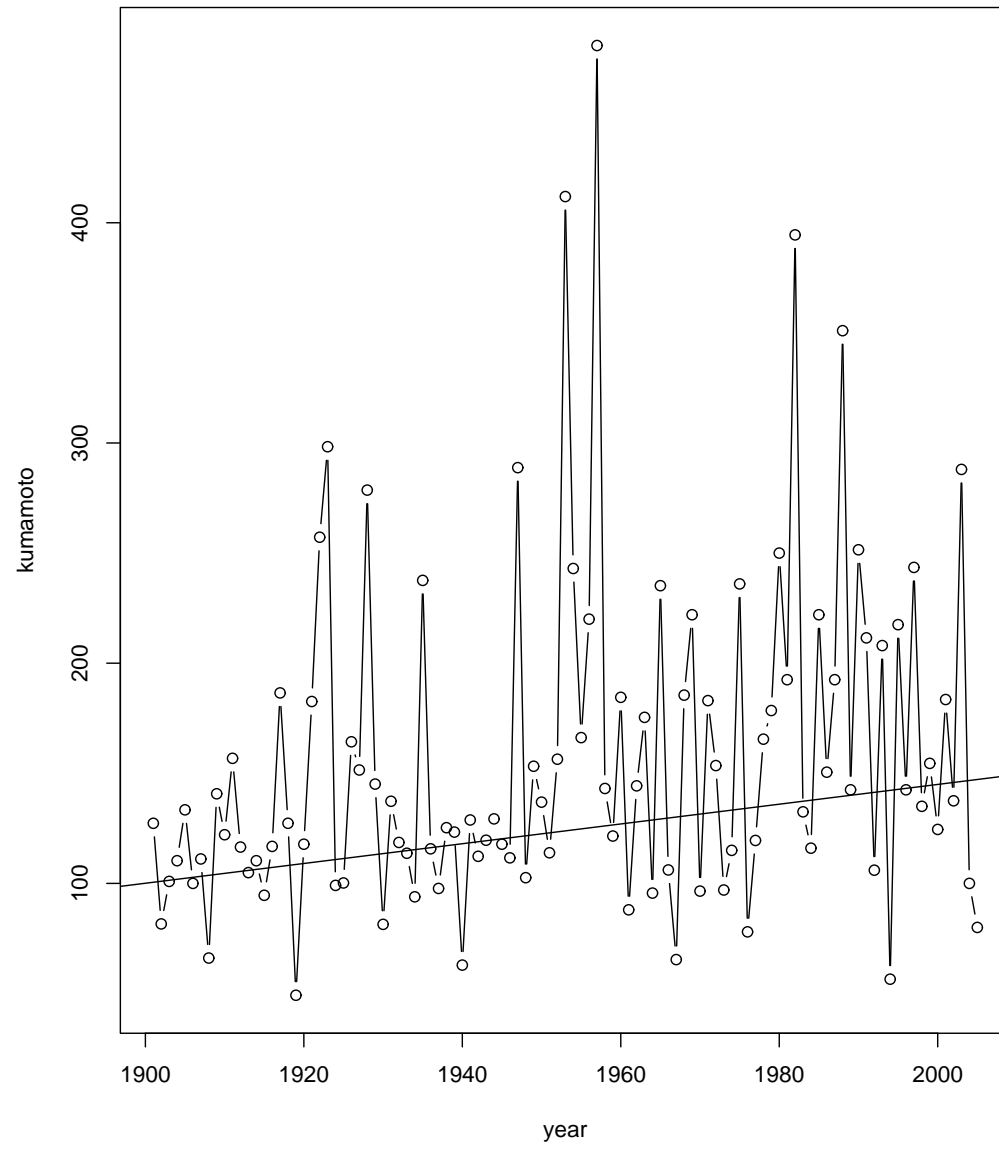
ただし $t = (\text{year} - 1953)/52$, と考えられる .



モデル1の診断：熊本の年最大日雨量，1901年から2005年．



モデル4の診断：熊本の年最大日雨量，1901年から2005年．



一般パレート (GP) モデル

水準超過値データ $\{y_1, y_2, \dots, y_n\}$ に一般パレート分布 $GP(\sigma, \xi)$

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

を適合させる場合の推測 .

一般パレート分布の性質

$$Y \sim \text{GP}(\sigma, \xi)$$

$\xi < 1$ で平均は存在

$$E(Y) = \int_0^{y_+} (1 - H(y)) dy = \int_0^{y_+} \left(1 + \xi \frac{y}{\sigma}\right)^{-1/\xi} dy = \frac{\sigma}{1 - \xi}.$$

ただし, $y_+ = \sup\{y : H(y) < 1\}$.

$\xi < 1/2$ のとき分散は存在

$$V(Y) = \frac{\sigma^2}{(1 - \xi)^2(1 - 2\xi)}.$$

$U \sim U(0, 1)$ を一様分布とすると

$$\left(1 + \xi \frac{Y}{\sigma}\right)^{-1/\xi} = U \implies \frac{1}{\xi} \log \left(1 + \xi \frac{Y}{\sigma}\right) = -\log U, \quad Y = \frac{\sigma(U^{-\xi} - 1)}{\xi}.$$

この式は一般パレート分布の平均，分散，情報量等を計算するとき有用．

$$Y - v | Y > v \sim \text{GP}(\sigma + \xi v, \xi) \quad (v > 0)$$

これは次から明らか，

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

$e(v) = E(Y - v | Y > v)$: Y の 平均超過 (mean excess) 関数

$\tilde{e}(v)$: Y の メジアン超過 (median excess) 関数

$$P(Y - v \leq \tilde{e}(v) | Y > v) = 1/2$$

$Y - v | Y > v \sim \text{GP}(\sigma + \xi v, \xi)$ より

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

$$\tilde{e}(v) = \begin{cases} \sigma(2^\xi - 1)/\xi + (2^\xi - 1)v, & \xi \neq 0, \\ \sigma \log 2, & \xi = 0. \end{cases}$$

最尤法 $GP(\sigma, \xi)$

対数尤度： $\xi \neq 0$ のとき

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

$$1 + \xi y_i/\sigma > 0, \quad i = 1, 2, \dots, n,$$

$\xi = 0$ のとき

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

対数尤度を最大にする最尤推定値 $(\hat{\sigma}, \hat{\xi})$ または $\hat{\sigma}$ を求める .

$\xi > -0.5$ の場合は最尤推定量が漸近有効推定量になる (Smith, 1985) .

すなわち, $\xi > -0.5$ のとき, 最尤推定量 $(\hat{\sigma}, \hat{\xi})$ は漸近有効で, 漸近的に平均 (σ, ξ) で分散共分散行列

$$V = \frac{1}{n} \begin{pmatrix} 2\sigma^2(1 + \xi) & -\sigma(1 + \xi) \\ -\sigma(1 + \xi) & (1 + \xi)^2 \end{pmatrix}$$

の 2 変量正規分布に従う .

応用上 $\xi \leq -0.5$ となることは稀で, パラメータの推定は最尤法で行えばよい .

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 . これを用いて信頼区間を構成することができる .

$\{x_1, x_2, \dots, x_{n_o}\}$ を未知の母集団分布 F からの n_o 個のデータとする。このとき分布 F が適当な条件を満たせば、十分な大きな閾値 u 以上の水準超過値データ $y_i = x_{[i]} - u, i = 1, 2, \dots, n_u$ は一般パレート分布 $GP(\sigma, \xi)$ に従うとみなせる。そこで、閾値 u が求めれば上で述べた最尤法でパラメータを推定すればよい。

Smith (1987) は仮定 $F \in MDA(G_\xi)$ の下で、水準値超過データに基づく分布 F の裾指数（と上端点、 $\xi < 0$ の場合）の推定精度について $\xi > 0, \xi = 0$ と $\xi < 0$ の場合に分けて詳しい議論をしている。

m 観測再現レベル

一般パレート分布のパラメータ (σ, ξ) の直接の推定よりも，次の $GP(\sigma, \xi)$ に基づく上側（微少）確率点の推定が目的の場合が多い．

十分大きな u に対して，条件つき確率変数 $X - u | X > u$ の分布が一般パレート分布 $GP(\sigma, \xi)$ に従うとする．すなわち，

$$P(X > x | X > u) = \left[1 + \xi \left(\frac{x - u}{\sigma} \right) \right]^{-1/\xi}, \quad x > u.$$

このとき，

$$P(X > x) = \zeta_u \left[1 + \xi \left(\frac{x - u}{\sigma} \right) \right]^{-1/\xi},$$

ただし， $\zeta_u = P(X > u)$ ．

x_m : m 個のデータで平均 1 回超えられるレベル

$$\zeta_u \left[1 + \xi \left(\frac{x_m - u}{\sigma} \right) \right]^{-1/\xi} = \frac{1}{m}$$

より

$$x_m = u + \frac{\sigma}{\xi} \left[\left(m \zeta_u \right)^\xi - 1 \right].$$

$\xi = 0$ のときは,

$$x_m = u + \sigma \log(m \zeta_u).$$

m は $x_m > u$ となる十分大きな数 .

x_m : m 観測再現レベル (m -observation return level)

m 観測再現レベルプロット

$$\left\{ (\log m, x_m); x_m > u \right\}$$

m 観測再現レベルプロットは, $\xi < 0$ のときは凹で, $\xi = 0$ のときは直線で, $\xi > 0$ のときは凸になる.

自然データを扱う場合は N 年再現レベルを考えた方がよい. 1 年に n_y 個観測する場合は $m = N \times n_y$. この場合の N 年再現レベルは

$$z_N = \begin{cases} u + \sigma \left[(N n_y \zeta_u)^\xi - 1 \right] / \xi, & \xi \neq 0, \\ u + \sigma \log (N n_y \zeta_u), & \xi = 0. \end{cases}$$

σ, ξ, ζ_u の推定値が必要. σ と ξ は最尤推定値を用い, $\hat{\zeta}_u = n/n_o$ とすれば良い.

閾値 u を超える数は二項分布 $B(n_o, \zeta_u)$ に従うから $\hat{\zeta}_u$ は最尤推定値でその分散は

$$V(\hat{\zeta}_u) \approx \hat{\zeta}_u(1 - \hat{\zeta}_u)/n_o$$

となる。したがって、 $(\hat{\zeta}_u, \hat{\sigma}, \hat{\xi})$ の分散共分散行列は近似的に

$$V^* = \begin{bmatrix} \hat{\zeta}_u(1 - \hat{\zeta}_u)/n_o & \mathbf{0}^T \\ \mathbf{0} & V \end{bmatrix}.$$

ここで、 $\mathbf{0}^T = (0, 0)$ で、 V は $(\hat{\sigma}, \hat{\xi})$ の分散共分散行列。

これを用いるとデルタ法から \hat{x}_m の分散を次の様に求めることが出来る :

$$V(\hat{x}_m) \approx \nabla x_m^T V^* \nabla x_m.$$

ただし

$$\nabla x_m = \begin{bmatrix} \frac{\partial x_m}{\partial \zeta_u} \\ \frac{\partial x_m}{\partial \sigma} \\ \frac{\partial x_m}{\partial \xi} \end{bmatrix} = \begin{bmatrix} \sigma m^\xi \zeta_u^{\xi-1} \\ \{(m\zeta_u)^\xi - 1\} / \xi \\ \sigma (m\zeta_u)^\xi \log(m\zeta_u) / \xi - \sigma \{(m\zeta_u)^\xi - 1\} / \xi^2 \end{bmatrix}$$

で , $(\hat{\zeta}_u, \hat{\sigma}, \hat{\xi})$ で求める .

m 観測再現レベルの精度の良い信頼区間を求めるにはプロファイル尤度を考える必要がある．そのために x_m の定義式から

$$\sigma = \begin{cases} \frac{(x_m - u)\xi}{(m\zeta_u)^\xi - 1}, & \xi \neq 0, \\ \frac{x_m - u}{\log(m\zeta_u)}, & \xi = 0, \end{cases}$$

として，対数尤度 $l(x_m, \xi)$ を考える． x_m を固定して $\max_{\xi} l(x_m, \xi)$ 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．

モーメント法 $GP(\sigma, \xi)$

$\xi < 1/2$ のとき, モーメント推定値は

$$\frac{\sigma}{1 - \xi} = \bar{y}, \quad \frac{\sigma^2}{(1 - \xi)^2(1 - 2\xi)} = s^2$$

から

$$\hat{\sigma}_M = \frac{1}{2}\bar{y}(\bar{y}^2/s^2 + 1), \quad \hat{\xi}_M = \frac{1}{2}(1 - \bar{y}^2/s^2).$$

ただし \bar{y} と s^2 はそれぞれ平均と分散.

PWM法

次の PWM を考える (Hosking et al., 1987) :

$$\beta_s = M_{1,0,s} = E[Y\{1-H(Y)\}^s] = \int_0^1 H^{-1}(u)(1-u)^s du = \frac{\sigma}{(s+1)(s+1-\xi)}.$$

ただし $\xi < 1$. このとき , $s = 0, 1$ を用いて

$$\beta_0 = \frac{\sigma}{1-\xi}, \quad \beta_1 = \frac{\sigma}{2(2-\xi)}$$

から

$$\sigma = \frac{2\beta_0\beta_1}{\beta_0 - 2\beta_1}, \quad \xi = 2 - \frac{\beta_0}{\beta_0 - 2\beta_1}.$$

ここで、データの順序統計量を $y_{(1)} \leq y_{(2)} \leq \cdots \leq y_{(n)}$ とするとき、

$$\hat{\beta}_0 = \bar{y}, \quad \hat{\beta}_1 = \frac{1}{n} \sum_{i=1}^n \left(1 - \frac{i - 0.35}{n}\right) y_{(i)}$$

で PWM's の β_0 と β_1 を推定すればよい。これを用いて PWM 推定値 $\hat{\sigma}_*$, $\hat{\xi}_*$ が得られる。

閾値の決定

一般パレート分布をデータに適合する方法は、水準超過法 (threshold method) と呼ばれている。最適な閾値 u の決定の問題がある。

閾値 u を小さくとるとデータ数は多くなり、分散は小さくなるが一般パレート分布への適合は悪くなり偏りは大きくなる。

閾値 u を大きくとるとデータ数は少なくなり、一般パレート分布への適合は良くなり偏りは小さくなるが分散は大きくなる。

閾値はこれらの事を考慮し現実に応じて決めなければならない。

1 . 平均超過プロット (mean excess plot)

データで u を超えるものを $x_{[1]}, x_{[2]}, \dots, x_{[n_u]}$ とし, x_{max} でデータの最大値を表す . 値 $u < x_{max}$ に対して標本平均超過, すなわち

$$\left\{ \left(u, \frac{1}{n_u} \sum_{i=1}^{n_u} (x_{[i]} - u) \right) : u < x_{max} \right\}$$

をプロットする . この図で, ある u 以上ではプロットが直線と見なせるとき, この u を閾値と決める .

$$Y \sim \text{GP}(\sigma, \xi) \quad E(Y - v | Y > v) = \frac{\sigma}{1 - \xi} + \frac{\xi}{1 - \xi} v$$

2 . メジアン超過プロット (median excess plot)

値 $u < x_{max}$ に対して標本メジアン超過 , すなわち

$$\left\{ \left(u, \text{median}_{1 \leq i \leq n_u} \{ x_{[i]} - u \} \right) : u < x_{max} \right\}$$

をプロットする . この図で , ある u 以上ではプロットが直線と見なせるとき , この u を閾値と決める .

$$\tilde{e}(v) = \begin{cases} \sigma(2^\xi - 1)/\xi + (2^\xi - 1)v, & \xi \neq 0, \\ \sigma \log 2, & \xi = 0. \end{cases}$$

3 . パラメータ ξ と σ^* の安定性

値 $u < x_{max}$ に対して, データ $\{x_{[i]} - u\}_{i=1}^{n_u}$ に $GP(\sigma_u, \xi)$ を適合させ最尤法でパラメータ σ_u と ξ を推定する. 値 u を変化させて $(u, \hat{\sigma}^*)$ と $(u, \hat{\xi})$ をプロットする. ただし $\sigma^* = \sigma_u - \xi u$. このとき, ある u 以上では推定値 $\hat{\sigma}^*$ と $\hat{\xi}$ が共に一定であると見なせるとき, この u を閾値と決める.

$$Y - v | Y > v \sim GP(\sigma + \xi v, \xi) \quad (v > 0), \quad \sigma_v = \sigma + \xi v.$$

4 . Pareto 確率紙 と Hillプロット ($\xi > 0$ の場合)

大きい方から m 個のデータ , $x^{(1)} \geq x^{(2)} \geq \dots \geq x^{(m)}$, を選び ,

$$\left\{ \left(-\log \left(\frac{i}{m+1} \right) , \log x^{(i)} \right) , \quad i = 1, 2, \dots, m \right\}$$

をプロットする . この図で , ある $x^{(k+1)}$ より右側ではプロットが直線と見なせるとき , この $x^{(k+1)}$ を閾値と決める .

このプロットの傾きは , Hill推定値 (Hill, 1975)

$$\hat{\xi}_H(k) = \frac{1}{k} \sum_{i=1}^k \left(\log x^{(i)} - \log x^{(k+1)} \right)$$

で推定できる . そこで , 各 k について $\hat{\xi}_H(k)$ をプロットしそれが一定であると見なせる最大の k を決め , $x^{(k+1)}$ を閾値とする .

モデル診断 GP(σ, ξ)

GPモデルの下で, $y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(n)}$ をデータの順序統計量とする.

確率プロット

$$\left\{ \left(\widehat{H}(y_{(i)}), \frac{i}{n+1} \right) : i = 1, \dots, n \right\}$$

ただし

$$\widehat{H}(y_{(i)}) = 1 - \left(1 + \widehat{\xi} \frac{y_{(i)}}{\widehat{\sigma}} \right)^{-1/\widehat{\xi}}.$$

確率点プロット

$$\left\{ \left(\widehat{H}^{-1} \left(\frac{i}{n+1} \right), y_{(i)} \right) : i = 1, \dots, n \right\}$$

ただし

$$\widehat{H}^{-1} \left(\frac{i}{n+1} \right) = \widehat{\sigma} \left[\left(\frac{n+1-i}{n+1} \right)^{-\widehat{\xi}} - 1 \right] / \widehat{\xi}.$$

再現レベルプロットは, m が十分大のときの $\{(\log m, \widehat{x}_m)\}$. ただし

$$\widehat{x}_m = u + \widehat{\sigma} \left[\left(m \widehat{\zeta}_u \right)^{\widehat{\xi}} - 1 \right] / \widehat{\xi}.$$

非定常の場合の極値 パラメータが時間に依存する GP モデル

GP モデル

日ごとの観測値 X_1, X_2, \dots を考える . t 日の季節を $s(t)$ で表す .

季節モデル

$$X_t - u_{s(t)} \mid X_t > u_{s(t)} \sim \text{GP}(\sigma_{s(t)}, \xi_{s(t)}).$$

ただし $u_{s(t)}$ は季節 $s(t)$ での閾値 , $\sigma_{s(t)}$ と $\xi_{s(t)}$ は季節 $s(t)$ での一般パレート分布のパラメータである .

モデル診断 GPモデル

$$Y_t \sim \text{GP}(\sigma(t), \xi(t))$$

を考える．このとき

$$Y_0 = \frac{1}{\xi(t)} \log \left\{ 1 + \xi(t) \frac{Y_t}{\sigma(t)} \right\} \sim \text{GP}(1, 0) : \text{標準指数分布}$$

$$P(Y_0 \leq y) = H_0(z) = 1 - \exp(-y), \quad 0 < y < \infty.$$

ここで $y_{t_1}, y_{t_2}, \dots, y_{t_k}$ を観測値として, $\hat{\sigma}(t), \hat{\xi}(t)$ を最尤推定値 $\hat{\beta}$ から求めたものとする

$$\tilde{y}_{t_k} = \frac{1}{\hat{\xi}(t)} \log \left\{ 1 + \hat{\xi}(t) \frac{y_{t_i}}{\hat{\sigma}(t)} \right\}, \quad i = 1, 2, \dots, k$$

は標準指数分布からのデータと見なせる．そこで, これらを大きさの順に並べたも

のを

$$\tilde{y}_{(1)} \leq \tilde{y}_{(2)} \leq \cdots \leq \tilde{y}_{(k)}$$

とする .

確率プロットは ,

$$\left\{ \left(1 - \exp(-\tilde{y}_{(i)}) , \frac{i}{k+1} \right) : i = 1, \dots, k \right\}$$

で , 確率点プロットは

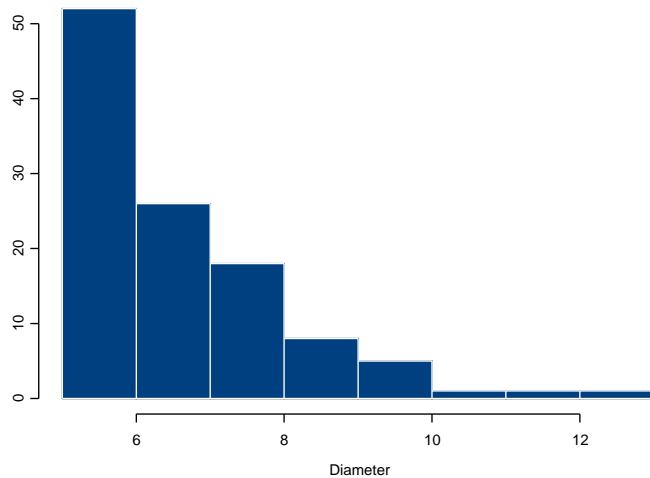
$$\left\{ \left(-\log \left(1 - \frac{i}{k+1} \right) , \tilde{y}_{(i)} \right) : i = 1, \dots, k \right\}$$

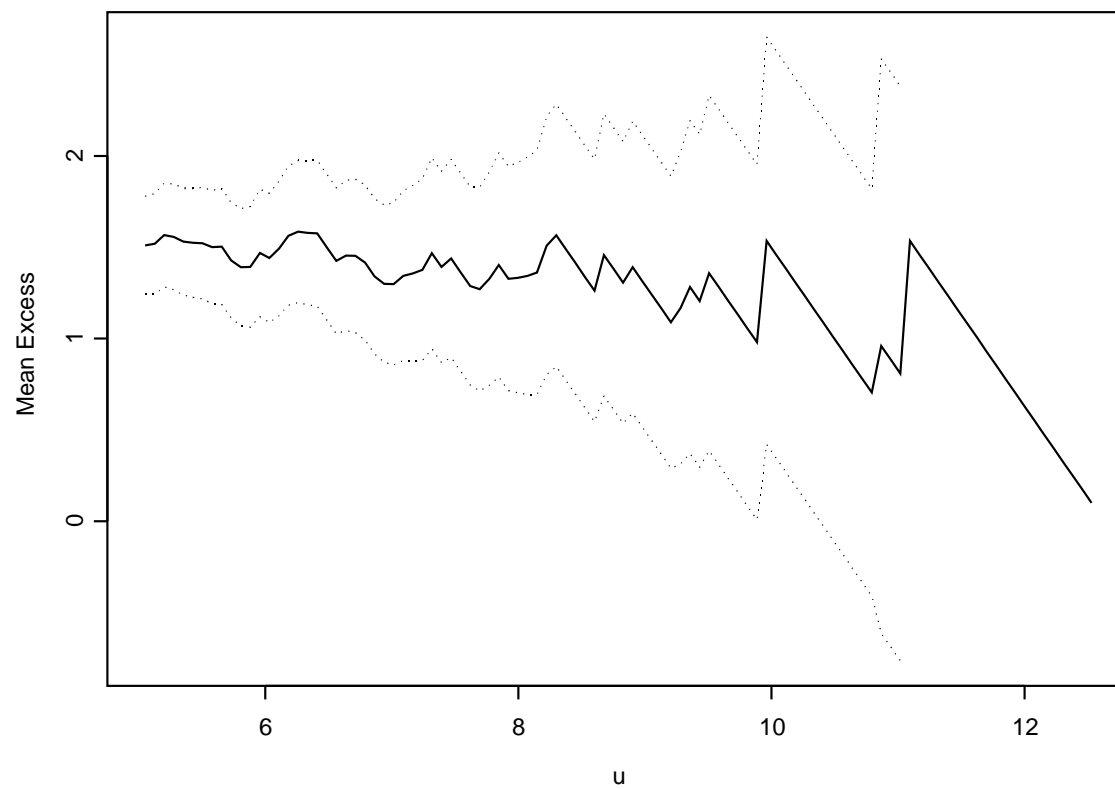
となる .

データ解析例 1

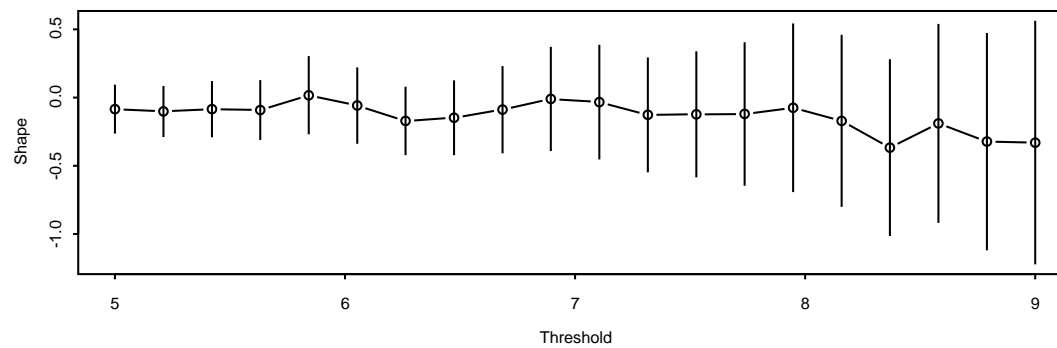
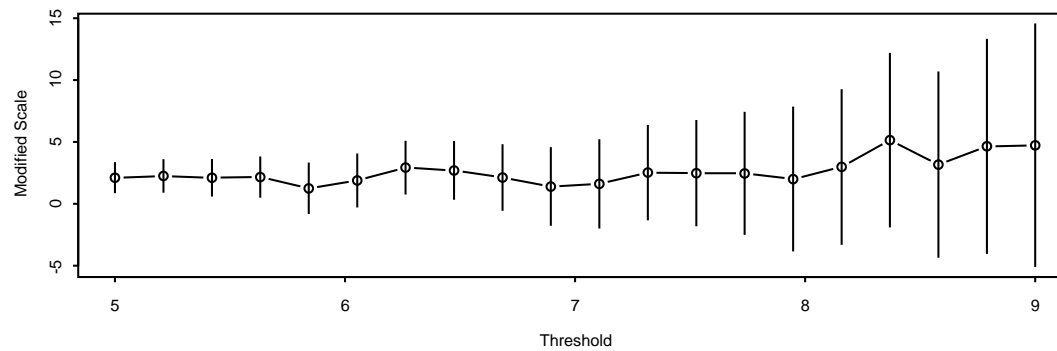
Anderson & Coles (2002) のデータ : Beirlant et al. (2004)

clean bearing steel の磨いた表面上で観測された切断介在物で直径が $5\mu\text{m}$ 以上のすべての観測値からなる．データのサイズは 112 ．





直径データ（平面）の標本平均余命関数．



修正尺度 σ^* と形状パラメータ ξ の最尤推定値 .

GPモデル

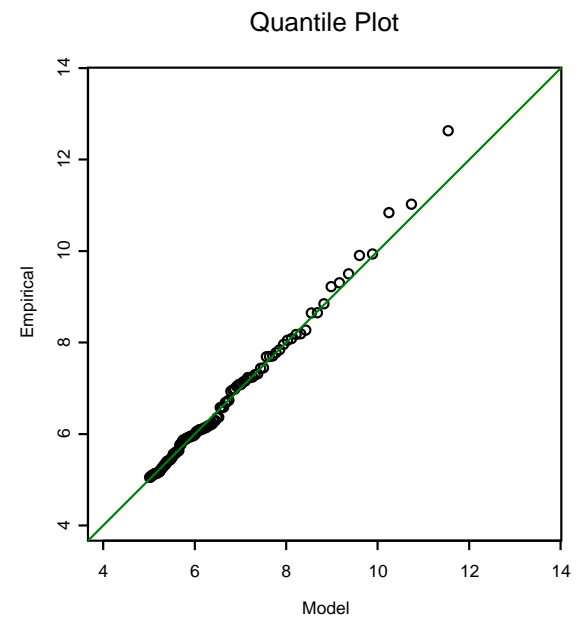
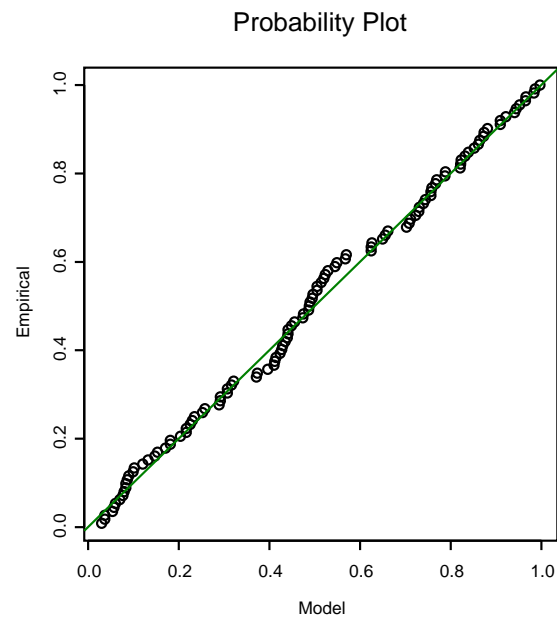
閾値を 5 として一般パレート分布を適合させて解析すると、最尤推定値（標準誤差）は

$$\hat{\xi} = -0.085 (0.0914), \quad \hat{\sigma} = 1.68 (0.220)$$

で、最大対数尤度は -160.6 となった。

診断のために PP と QQ プロットを描かせた。

一般パレート分布の適合は悪くない。



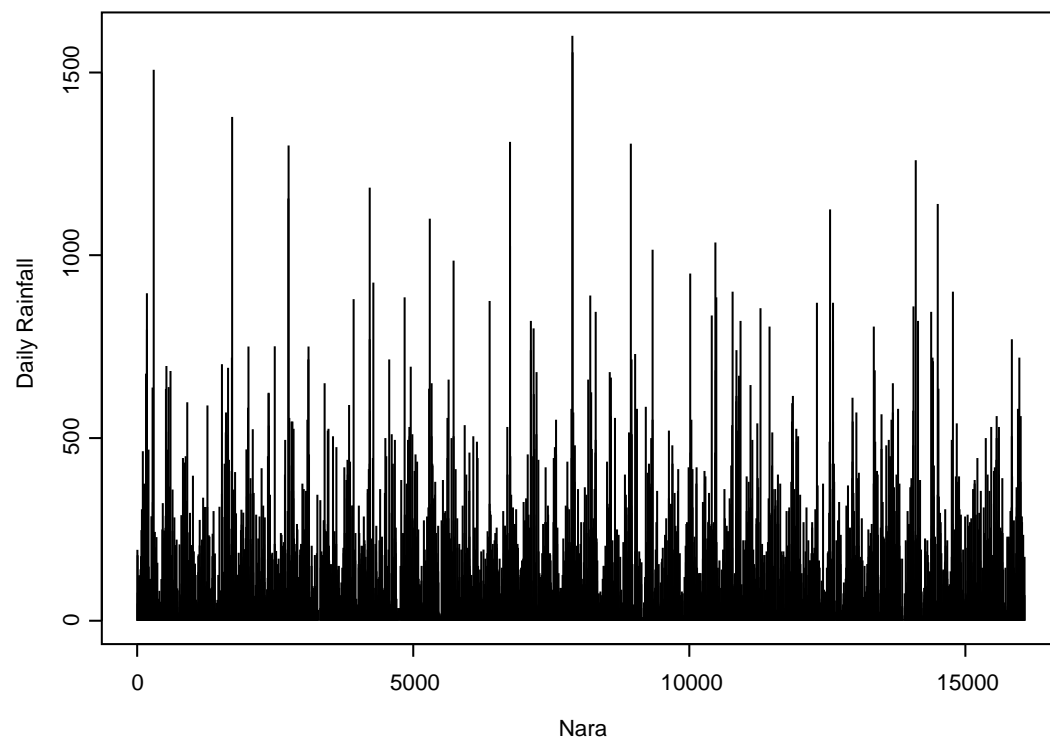
一般パレート分布による解析の診断 .

データ解析例 2

ここでは日降水量（単位 0.1mm）データに GP モデルを適用して解析する．

図は奈良の 1961 年 1 月 1 日から 2004 年 12 月 31 日までの日降水量データ（気象データベース，地上観測（財）気象業務支援センター）のプロットである．

このデータを一般パレート分布を用いて解析する．



奈良の日降水量データ (0.1mm) , 1961年1月1日 ~ 2004年12月31日 .
1000以上のデータはまばら . 500以上のデータも少ない .

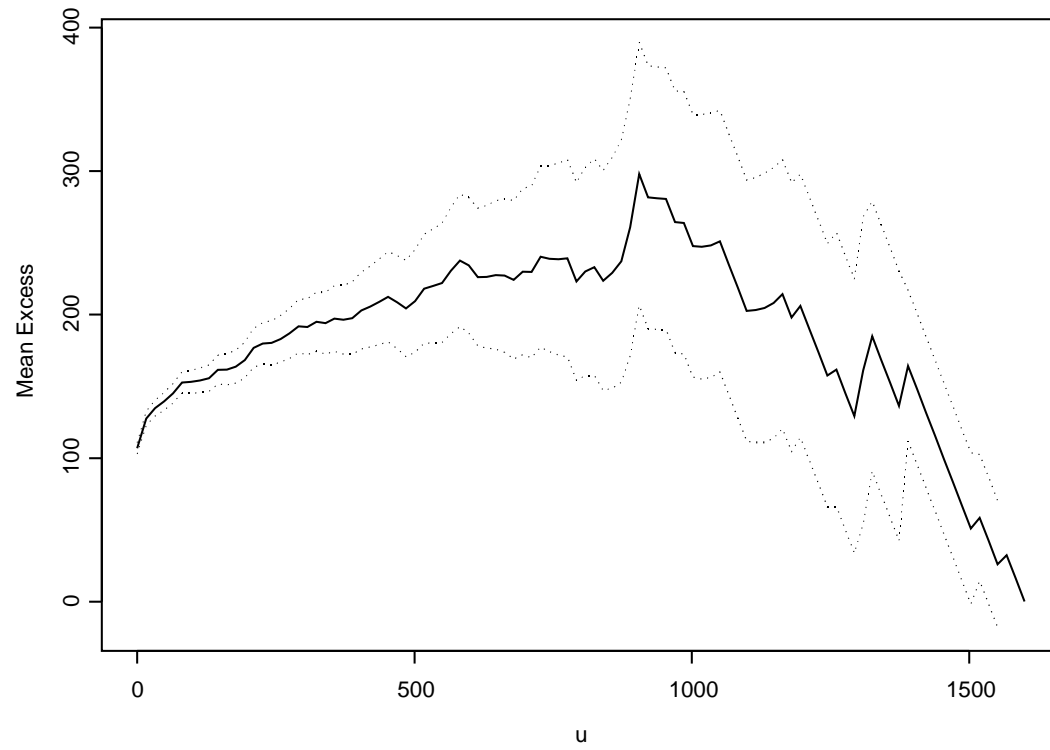


図9 奈良の日降水量データの平均超過プロット。

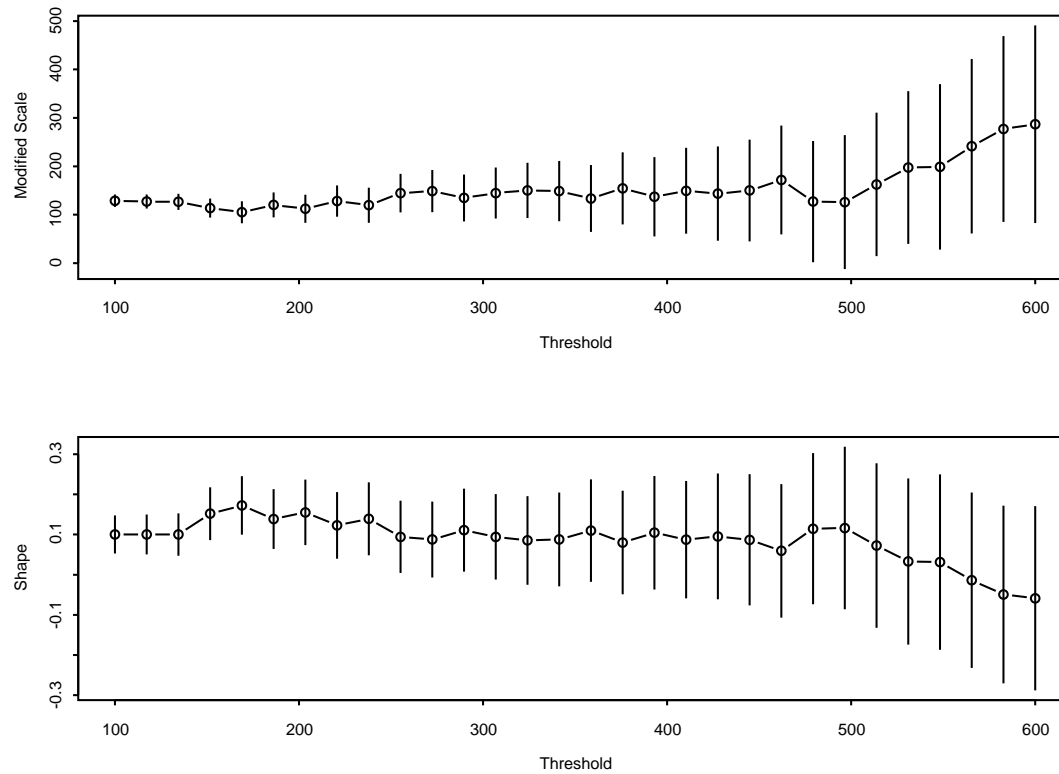


図 1 0 奈良の日降水量データの修正尺度 σ^* と形状パラメータ ξ の最尤推定値 .

図9は、データの平均超過プロットである。閾値 u として 300 あたりが考えられる。Coles (2001) による、閾値 u を変化させて修正パラメータ σ^* と形状パラメータ ξ を描かせたものが次の図10である。これらの図から閾値 $u = 300$ を採用。

453個の超過データからパラメータ (σ, ξ) の最尤推定値を求めた。最大対数尤度は -2839.646 で推定値 (標準誤差) は次の様になった：

$$\hat{\sigma} = 180.65 (12.541), \quad \hat{\xi} = 0.072 (0.0512).$$

これらの推定値を用いてモデル診断を行ったのが図11である。奈良の日降水量の水準超過値データにはGPモデルが適合している事がわかる。

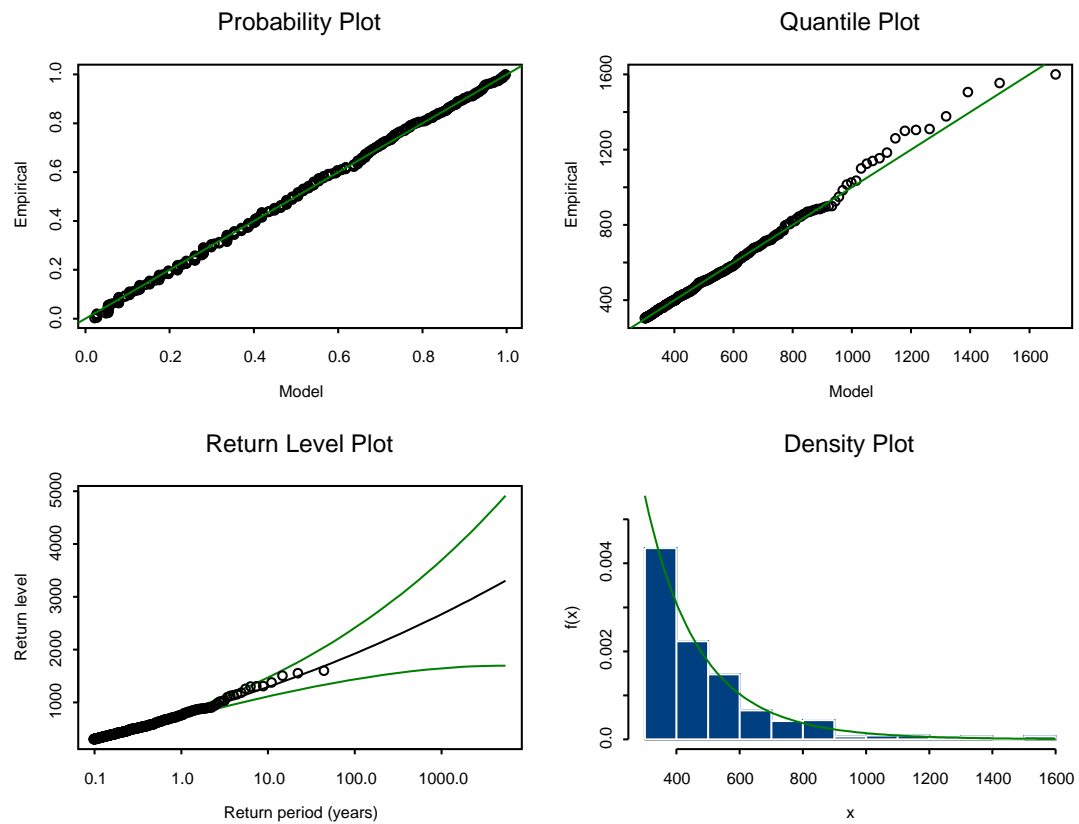


図 1 1 奈良の日降水量データ解析の診断 .

形状パラメータ ξ の推定値は正．かなり大きな日降水量が発生する可能性がある．

ξ の信頼区間を求めるためにプロファイル尤度関数を描かせたものが 図 1 2 である．これから ξ の 95% 信頼区間は $[-0.020, 0.183]$ となった．

100年再現レベルの最尤推定値は $\hat{z}_{1/100} = 1925$ で，その信頼区間を求めるために描かせたのが図 1 3 で，これから 95% 信頼区間は $[1567, 2633]$ となった．

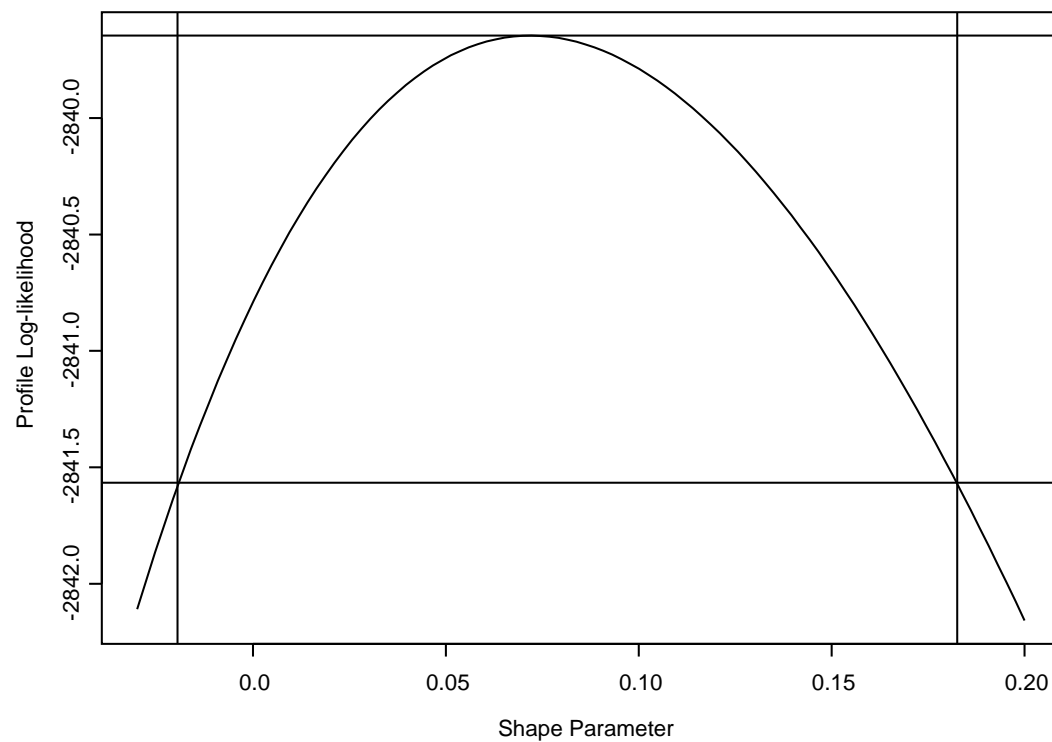


図 1 2 奈良の日降水量データの形状パラメータ ξ のプロファイル尤度 .
 $\hat{\xi} = 0.072 (0.0512)$, 95% 信頼区間は $[-0.020, 0.183]$.

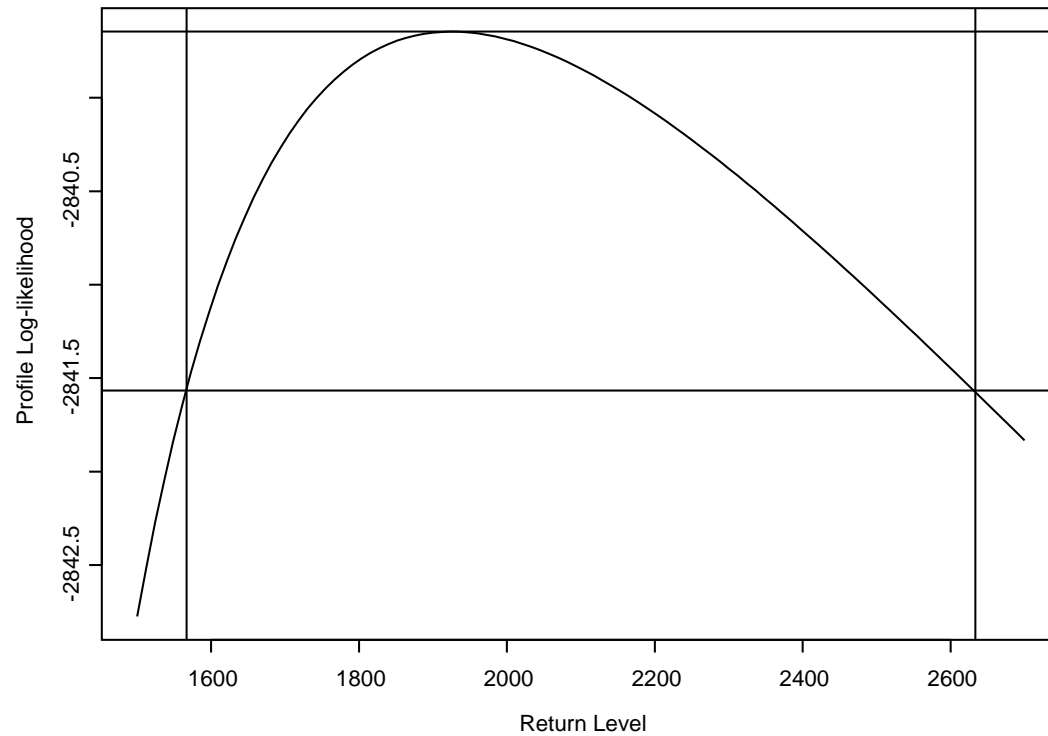
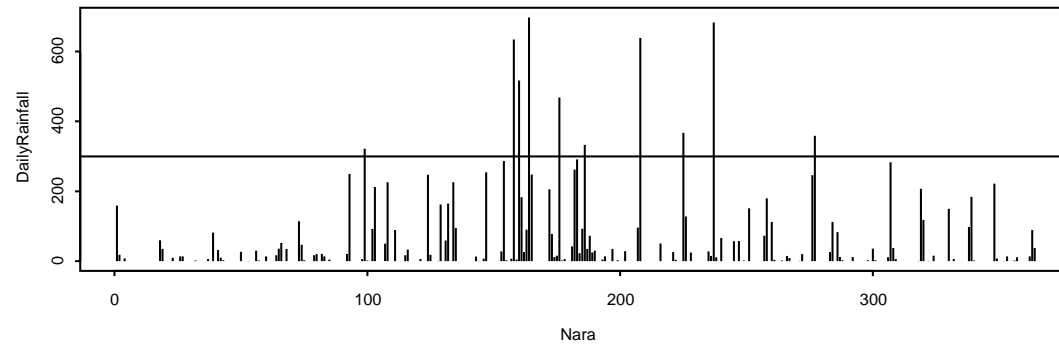
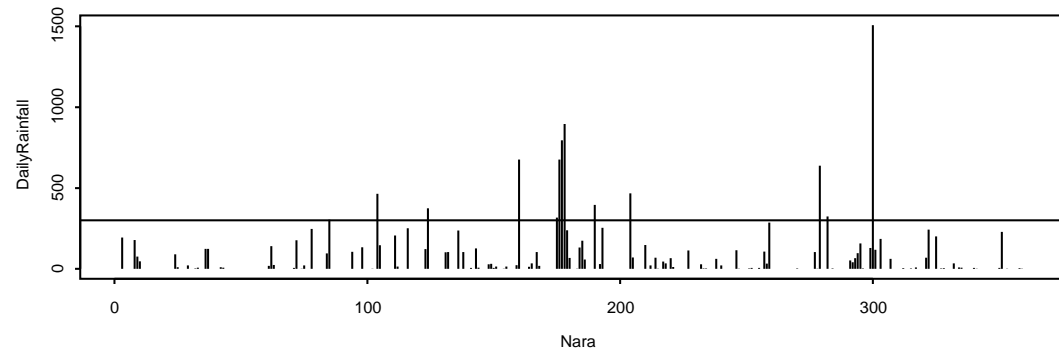


図 1 3 奈良の日降水量データの再現期間 $z_{1/100}$ のプロファイル尤度 .
 $\hat{z}_{1/100} = 1925$, 95% 信頼区間は $[1567, 2633]$.

GPモデル：データは互いに独立で同一分布 $GP(\sigma, \xi)$ からの標本と仮定．

独立性を持たせるためには連続して閾値を超える値があるのは好ましくない．

図 1 4 は 1961 年と 1962 年のデータに閾値 300 を描いたものである．一部同じ雨（定義は難しい!?!）からと思われる水準超過値がある．そこで，1 日空いていれば他の雨からのものとみなし連続した日で超過値がある場合はそれらの中の最大値のみ用いる事にして，新たに 404 個（453 個から 49 個減）のデータを GP モデルで解析してみる．



奈良の日降水量データ (0.1mm) , 1961年(上)と1962年(下) .

同じ閾値 300 を用いてパラメータを推定：最大対数尤度は -2548.728 で最尤推定値（標準誤差）は

$$\hat{\sigma} = 189.54 (13.901), \quad \hat{\xi} = 0.064 (0.0540).$$

パラメータの推定値は最初に求めたものにほぼ等しい。この推定値によるモデル診断は図 1 5。また ξ と $z_{1/100}$ の信頼区間を求めるためにそれぞれのプロファイル尤度を描かせたのが図 1 6 と図 1 7 である。これから、 ξ と $z_{1/100}$ の 95% 信頼区間はそれぞれ $[-0.032, 0.182]$, $[1559, 2649]$ となった。

奈良の日降水量データではデータの修正による影響はほとんどなかった。

従属性のあるデータから独立と見なせるデータを取り出すことを **declustering** というが一般論が無く難しい問題である（Coles (2001) §5.3.2 参照。）

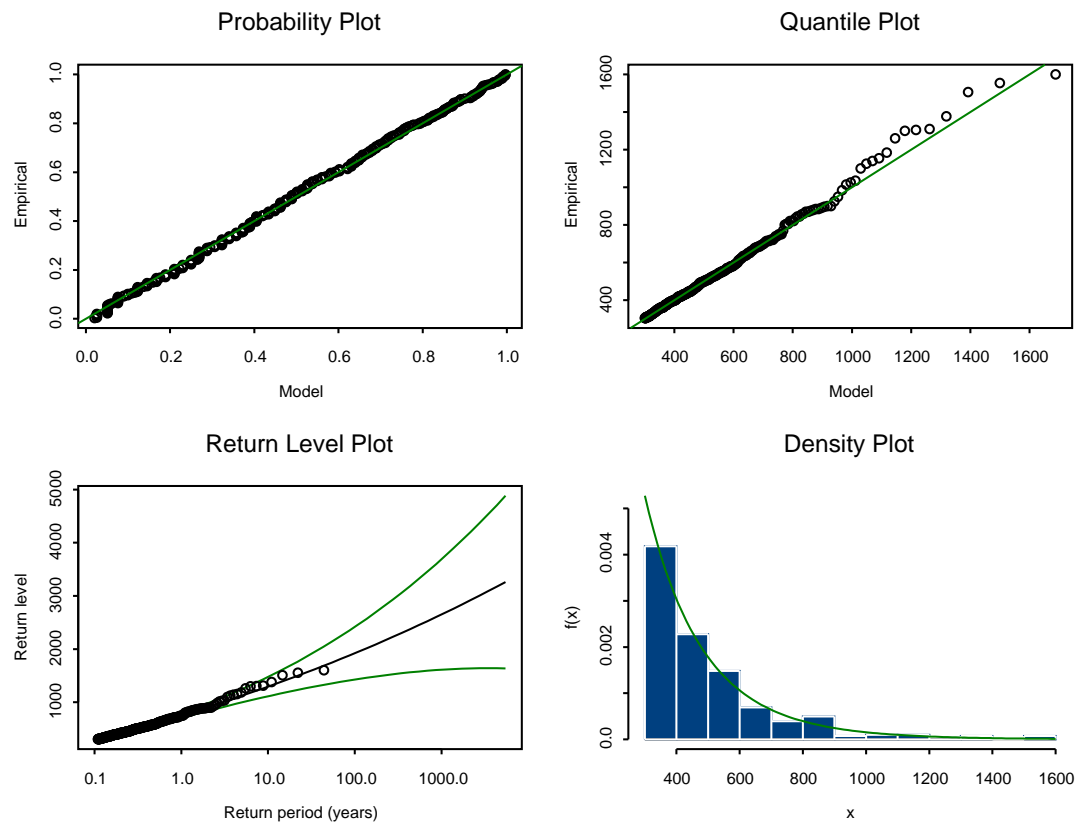


図 1 5 奈良の修正日降水量データ解析の診断 .

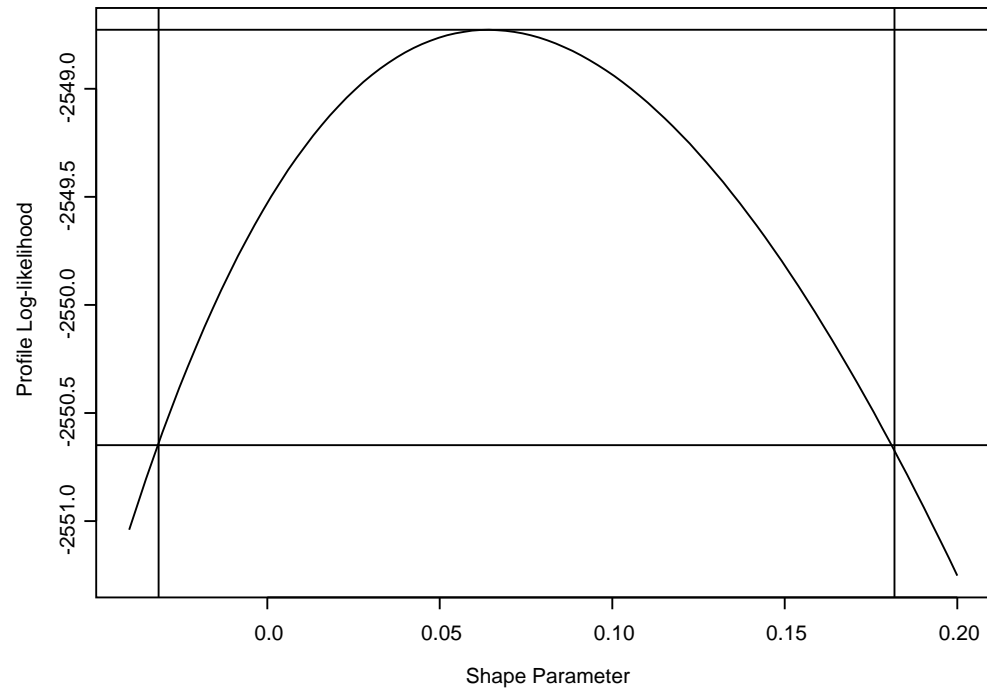


図 1 6 奈良の修正日降水量データの形状パラメータ ξ のプロファイル尤度 .
 $\hat{\xi} = 0.064$ (0.0540) , 95% 信頼区間は $[-0.032, 0.182]$.
 ($[-0.020, 0.183]$)

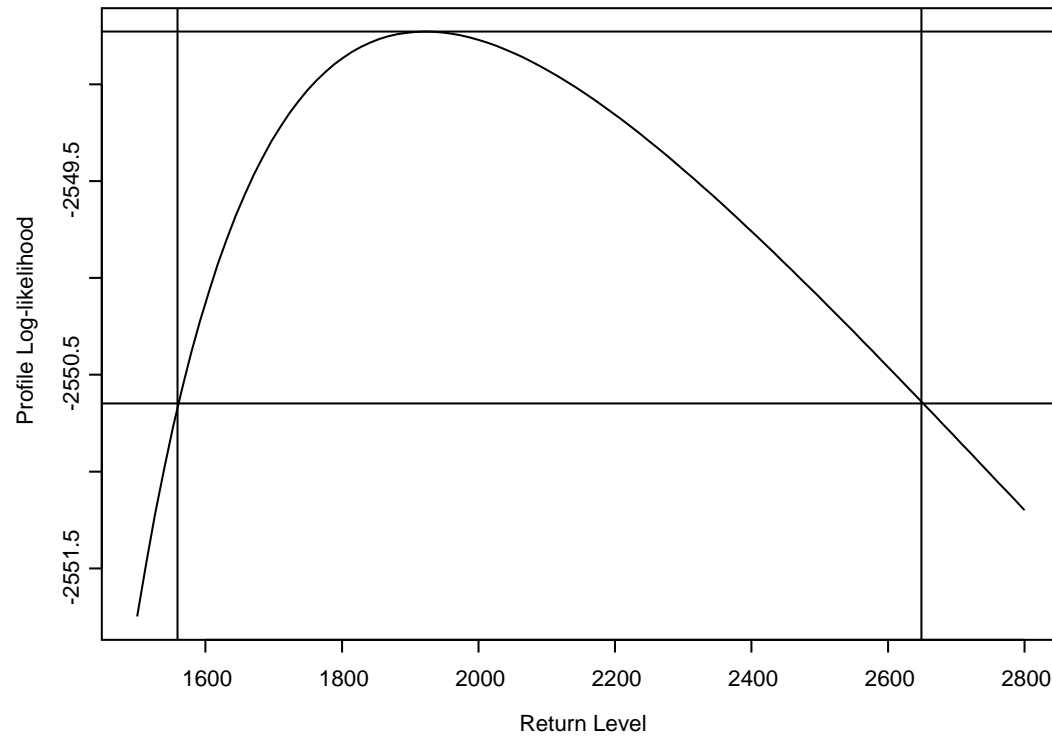


図 1 7 奈良の修正日降水量データの再現期間 $z_{1/100}$ のプロファイル尤度 .
 $z_{1/100}$ の 95% 信頼区間は [1559, 2649] . ([1567, 2633])

点過程 (PP) モデル

十分大きな閾値 u を決めると, 領域 $(0, 1) \times (u, \infty)$ では N_n は Poisson 過程 N で近似できた. 領域 $A = (0, 1) \times (u, \infty)$ に入っている点を

$$\left\{ (t_1, x_1), \dots, (t_{N(A)}, x_{N(A)}) \right\}$$

とする. 領域 A 内では $N_n \approx N$ であるから, 近似的な尤度が計算できる.

$A = [0, 1] \times (u, \infty)$ としたとき

$$\Lambda(A) = \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi}$$

であった.

自然現象で年に n_y 個観測する場合の尤度は

$$L_A(\mu, \sigma, \xi; x_1, \dots, x_n) = \exp \left\{ -\Lambda(A) \right\} \prod_{i=1}^{N(A)} \lambda(t_i, x_i)$$

$$\propto \exp \left\{ -n_y \left[1 + \xi \left(\frac{u - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \prod_{i=1}^{N(A)} \frac{1}{\sigma} \left[1 + \xi \left(\frac{x_i - \mu}{\sigma} \right) \right]^{-1/\xi - 1}$$

となる .

この尤度関数を最大化して最尤推定値 $(\hat{\mu}, \hat{\sigma}, \hat{\xi})$ を求める . 標準誤差や信頼区間も通常の方法で求まる .

PPモデル

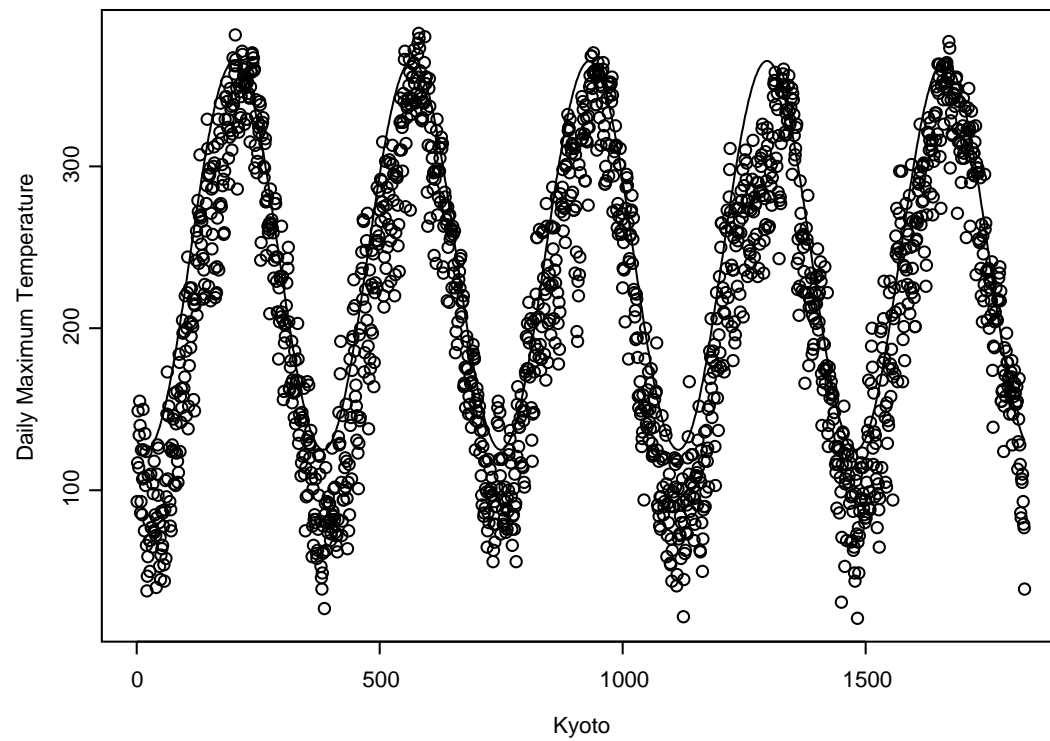
京都の日最高気温（単位 0.1 度C）データをPPモデルを適用して解析する．

図 1 8 は京都の2000年1月1日から2004年12月31日までの日最高気温データと時間に依存する閾値を描いたものである．

気温は明らかに周期性を持っている．そこで，Coles (2001) の本にあるように時間に依存する閾値関数を試行錯誤で探し，次の関数を採用した．

$$u(t) = 245 + 120 \sin(2\pi(t - 110)/365.25)$$

各期間で，ほぼ同じ割合でデータがこの閾値関数を超える様にしなければいけないが難しい．



京都の日最高気温 (0.1度C) , 2000年1月1日 ~ 2004年12月31日 .

この閾値を用いて Coles (2001) にならいパラメータに構造をいれた :

$$\begin{aligned}\mu(t) &= \beta_0 + \beta_1 \sin(2\pi t/365.25) + \beta_2 \cos(2\pi t/365.25), \\ \log \sigma(t) &= \beta_3 + \beta_4 \sin(2\pi t/365.25) + \beta_5 \cos(2\pi t/365.25).\end{aligned}$$

ここで , 形状パラメータの ξ は不変としておく .

Coles の S-Plus のプログラムを用いて解析したところ最大対数尤度は -149.03 で , 各パラメータの最尤推定値 (標準誤差) は次の様になった :

$$\begin{aligned}\hat{\beta}_0 &= 287.29 (1.100), \quad \hat{\beta}_1 = -28.38 (1.377), \quad \hat{\beta}_2 = -103.07 (2.033), \\ \hat{\beta}_3 &= 1.233 (0.1335), \quad \hat{\beta}_4 = 0.233 (0.0465), \quad \hat{\beta}_5 = 0.277 (0.0684), \\ \text{そして, } \hat{\xi} &= -0.451 (0.0473).\end{aligned}$$

このデータ解析結果の診断は図 19 である . 解析結果は良好である .

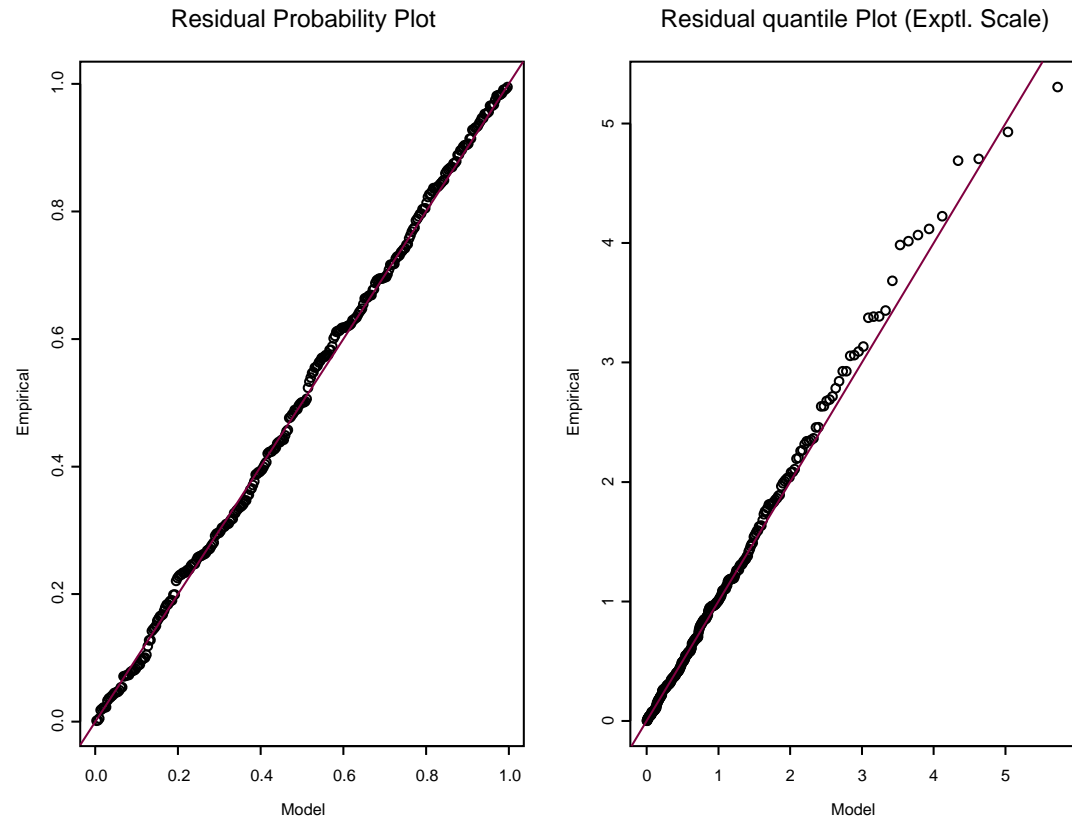


図 1 9 京都の日最高気温データ解析の診断図 .

他のモデルを調べてみたが，上記の位置と尺度パラメータに周期性を入れたモデルが最良であった．

大阪の日最高気温に関する解析結果を得た．