2011 年 5 月 8 日

確率分布と水文統計量の求め方

洪水流量や確率年雨量を求める場合など，年最大水文量に基づく水文統計量の求め方について，参考文献に基づき，まとめています．この文書は，理論を解説するものではなく，プログラムを作成するうえで必要となる数式を整理したものです．

目次

1. **基礎事項** 1
   - 1.1 確率年 . . . 1
   - 1.2 平均値・分散・ひずみ係数 . . . 1
   - 1.3 確率重み付き積率と L 積率 . . . 1
   - 1.4 プロッティング・ポジション公式 . . . 2

2. **確率分布モデル** 3
   - 2.1 正規分布 (N 分布) . . . 3
   - 2.2 対数正規分布 (LN3 分布) . . . 4
   - 2.3 ピアソン III 型分布 (P3 分布) . . . 6
   - 2.4 対数ピアソン III 型分布 (LP3 分布) . . . 7
   - 2.5 グンベル分布 . . . 8
   - 2.6 一般化極値分布 (GEV 分布) . . . 9
   - 2.7 平方根指数型最大値分布 (SQRT-ET 分布) . . . 10
   - 2.8 ワイブル分布 (3 母数) . . . 12

3. **ジャックナイフ法による推定** 14

4. **ブートストラップ法による推定** 14

5. **棄却検定** 15

参考文献

- ◯ 星清：水文統計解析，開発土木研究所月報，No.540，1998 年 5 月
- ◯ 星清：現場のための水文統計 (1)，開発土木研究所月報，No.540，1998 年 5 月
- ◯ 星清・新目竜一・宮原雅幸：現場のための水文統計 (2)，開発土木研究所月報，No.541，1998 年 6 月
- ◯ 合田良実・久高将信・河合弘泰：L-moments 法を用いた波浪の極値統計解析について，土木学会論文集 B2(海岸工学)，Vol.B2-65 No.1，2009，pp161-165
- ◯ Derek A. Roff(著)・野間口眞太郎 (訳)：生物学のための計算統計学 –最尤法，ブートストラップ法，無作為化法–，共立出版株式会社，2011 年 3 月 10 日

1. 基礎事項

1.1 確率年

◯ **再現期間 (確率年)T と非超過確率 p の関係**

$$T = \frac{1}{1-p} \qquad p = 1 - \frac{1}{T} \tag{1}$$

◯ **T 年確率事象 x_T が N 年間に 1 回も起こらない確率と 1 回以上起こる確率**

$$P(X < x_T)_N = \left(1 - \frac{1}{T}\right)^N \qquad \text{1 回も起こらない確率} \tag{2}$$

$$P(X \geqq x_T)_N = 1 - \left(1 - \frac{1}{T}\right)^N \qquad \text{1 回以上起こる確率} \tag{3}$$

◯ **計算事例：T 年確率事象が N 年間に 1 回も起こらない確率 P**

T	5	10	30	100	200	500	1000	5000
N	5	5	5	10	20	50	100	500
P	0.328	0.590	0.844	0.904	0.905	0.905	0.905	0.905

1.2 平均値・分散・ひずみ係数

$$\bar{x} = \frac{1}{N}\sum_{j=1}^{N} x_j \qquad \text{標本平均} \tag{4}$$

$$S^2 = \frac{1}{N}\sum_{j=1}^{N} (x_j - \bar{x})^2 \qquad \text{標本分散} \tag{5}$$

$$C_s = \frac{1}{N}\sum_{j=1}^{N} \left(\frac{x_j - \bar{x}}{S}\right)^3 \qquad \text{標本ひずみ係数} \tag{6}$$

$$\hat{\sigma}^2 = \frac{N}{N-1} S^2 \qquad \text{不偏分散} \tag{7}$$

$$\hat{\gamma} = \frac{\sqrt{N(N-1)}}{N-2} C_s \qquad \text{不偏ひずみ係数} \tag{8}$$

1.3 確率重み付き積率と L 積率

確率統計水文学では，確率重み付積率 (PWM : Probability Weighted Moments) と L 積率 (L Moments) が導入され，活用されている．

従来から用いられている平均値・分散・ひずみ係数などには，データの並び順序の概念はないが，確率重み付積率および L 積率では，昇順にデータを並び替えた順序統計量が用いられる．

母集団の確率重み付き積率 (PWM) は以下のとおり定義される．

$$\beta_r = \int_0^1 x F^r dF \qquad (r = 0, 1, 2, \dots) \tag{9}$$

また，母集団の PWM と L 積率の関係は以下のとおり関係付けられる．

$$\lambda_1 = \beta_0 \tag{10}$$

$$\lambda_2 = 2\beta_1 - \beta_0 \tag{11}$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \tag{12}$$

標本値の確率重み付き積率は以下のとおりであり，標本値についても PWM と L 積率の関係は同様に用い

られる．

$$b_0 = \frac{1}{N}\sum_{j=1}^{N} x_{(j)} \tag{13}$$

$$b_1 = \frac{1}{N(N-1)}\sum_{j=1}^{N}(j-1)x_{(j)} \tag{14}$$

$$b_2 = \frac{1}{N(N-1)(N-2)}\sum_{j=1}^{N}(j-1)(j-2)x_{(j)} \tag{15}$$

ここに，$x_{(j)}$ は N 個の標本を昇順に並べ替えたときの，小さいほうから j 番目の値を示す．

1.4 プロッティング・ポジション公式

$$F[x_{(i)}] = \frac{i-\alpha}{N+1-2\alpha} \tag{16}$$

N	標本数
i	標本値を昇順に並べたときの小さいほうからの順位
$x_{(i)}$	i 番目の順位標本値
$F[x_{(i)}]$	プロッティング・ポジション (非超過確率相当)
α	1〜1 の定数

公式名	Weibull	Blom	Cunnane	Gringorten	Hazen
α	0	0.375	0.40	0.44	0.5

2. 確率分布モデル

2.1 正規分布 (N 分布)

水文統計量 x が，正規分布に従う場合の母数推定方法を示す.

(1) 確率密度関数

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \cdot \exp\left[-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right] \tag{17}$$

(2) 確率分布関数

$$F(x) = \Phi\left(\frac{x-\mu_x}{\sigma_x}\right) \qquad \Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} \exp\left(-\frac{1}{2}t^2\right) dt \tag{18}$$

(3) 非超過確率 p に対する水文統計量 x_p

$$z = \frac{x-\mu_x}{\sigma_x} \qquad \rightarrow \qquad x = \mu_x + \sigma_x z \tag{19}$$

$$x_p = \mu_x + \sigma_x z_p \qquad z_p \text{ は } p = \Phi(z) \text{ となる } z \text{ の値} \tag{20}$$

(4) 母数推定 (L 積率法)

$$b_0 = \frac{1}{N}\sum_{j=1}^{N} x_{(j)} \qquad b_1 = \frac{1}{N(N-1)}\sum_{j=1}^{N}(j-1)x_{(j)} \qquad b_2 = \frac{1}{N(N-1)(N-2)}\sum_{j=1}^{N}(j-1)(j-2)x_{(j)} \tag{21}$$

ここに，$x_{(j)}$ は N 個の標本を昇順に並べ替えたときの，小さいほうから j 番目の値を示す.

上記 $b_i = \beta_i$ として，以下の関係により λ_i を算定する.

$$\lambda_1 = \beta_0 \qquad \lambda_2 = 2\beta_1 - \beta_0 \qquad \lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \tag{22}$$

母数は，以下に示す L 積率と母数の関係により推定する.

$$\begin{cases} \mu_x = \lambda_1 \\ \sigma_x = \sqrt{\pi}\lambda_2 \end{cases} \tag{23}$$

2.2 対数正規分布 (LN3 分布)

水文統計量 x が，分布下限 a を持つ 3 母数対数正規分布に従う場合の母数推定方法を示す.

(1) 確率密度関数

$$f(x) = \frac{1}{(x-a)\sqrt{2\pi}\sigma_y} \cdot \exp\left\{-\frac{1}{2}\left[\frac{\ln(x-a)-\mu_y}{\sigma_y}\right]^2\right\} \qquad y = \ln(x-a) \tag{24}$$

(2) 確率分布関数

$$F(x) = \Phi\left(\frac{\ln(x-a)-\mu_y}{\sigma_y}\right) \qquad \Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}\exp\left(-\frac{1}{2}t^2\right)dt \tag{25}$$

(3) 非超過確率 p に対する水文統計量 x_p

$$z = \frac{\ln(x-a)-\mu_y}{\sigma_y} \qquad \rightarrow \qquad x = a + \exp(\mu_y + \sigma_y z) \tag{26}$$

$$x_p = a + \exp(\mu_y + \sigma_y z_p) \qquad z_p \text{ は } p = \Phi(z) \text{ となる } z \text{ の値} \tag{27}$$

(4) 母数推定 (岩井法)

$$\begin{cases} a = \dfrac{x_{(1)}\cdot x_{(N)} - {x_m}^2}{x_{(1)} + x_{(N)} - 2x_m} & x_{(1)} + x_{(N)} - 2x_m > 0 \\ \mu_y = \dfrac{1}{N}\sum_{j=1}^{N}\ln(x_j - a) & \\ {\sigma_y}^2 = \dfrac{1}{N}\sum_{j=1}^{N}\left[\ln(x_j-a) - \mu_y\right]^2 & \end{cases} \tag{28}$$

ここに，$x_{(1)}$，$x_{(N)}$，x_m は，それぞれ標本最小値，標本最大値，メディアン (中央値) である．x_j など添字に括弧がついていない標本値は，順位は関係ない統計量であるが，実務計算では，最小・最大・メディアンを求めるのに小さい順に並び替えたほうが便利であるため $x_{(j)}$ の形となろう.

(5) 母数推定 (積率法)

$$\bar{x} = \frac{1}{N}\sum_{j=1}^{N}x_j \qquad S_x^2 = \frac{1}{N}\sum_{j=1}^{N}(x_j-\bar{x})^2 \qquad C_{sx} = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{x_j-\bar{x}}{S_x}\right)^3 \tag{29}$$

$$\mu_x = \bar{x} \qquad \sigma_x = [N/(N-1)]^{1/2}S_x \qquad \gamma_x = \frac{\sqrt{N(N-1)}}{N-2}C_{sx} \tag{30}$$

不偏ひずみ係数 γ_x については，以下に示す，Bobee と Robitaille による偏り補正した式 (対数正規分布の場合) も用いられる．( ) 内で B に乗ずるのは C_{sx} の 3 乗であることに注意.

$$\gamma_x = C_{sx}(A + B\cdot {C_{sx}}^3) \tag{31}$$

$$\text{ここに} \qquad A = 1.01 + 7.01/N + 14.66/N^2 \qquad B = 1.69/N + 74.66/N^2 \tag{32}$$

ここで，積率と母数の関係は以下の通り.

$$\mu_x = a + \epsilon\sqrt{\phi} \qquad \sigma_x = \epsilon\sqrt{\phi(\phi-1)} \qquad \gamma_x = (\phi+2)\sqrt{\phi-1} \tag{33}$$

$$\epsilon = \exp(\mu_y) \qquad \phi = \exp\left({\sigma_y}^2\right) \tag{34}$$

よって，母数は以下に示す関係より推定する．

$$\phi = \left[\beta + \sqrt{\beta^2 - 1}\right]^{1/3} + \left[\beta - \sqrt{\beta^2 - 1}\right]^{1/3} - 1 \qquad \left(\beta = 1 + \frac{{\gamma_x}^2}{2}\right) \tag{35}$$

$$\epsilon = \frac{\sigma_x}{\sqrt{\phi(\phi - 1)}} \tag{36}$$

$$\begin{cases} \sigma_y = \sqrt{\ln \phi} \\ \mu_y = \ln \epsilon \\ a = \mu_x - \epsilon\sqrt{\phi} \end{cases} \tag{37}$$

2.3 ピアソン III 型分布 (P3 分布)

水文統計量 x が，ピアソン III 型分布：Pearson type 3 distribution (ガンマ分布) に従う場合の母数推定方法を示す.

(1) 確率密度関数

$$f(x) = \frac{1}{|a| \cdot \Gamma(b)} \left(\frac{x-c}{a} \right)^{b-1} \cdot \exp\left(-\frac{x-c}{a} \right) \qquad a > 0 : c \leqq x < \infty \tag{38}$$

(2) 確率分布関数

$$F(x) = G\left(\frac{x-c}{a} \right) \qquad G(w) = \frac{1}{\Gamma(b)} \int_0^w t^{b-1} \exp(-t) dt \qquad (a > 0) \tag{39}$$

(3) 非超過確率 p に対する水文統計量 x_p

$$w = \frac{x-c}{a} \qquad \rightarrow \qquad x = c + aw \tag{40}$$

$$x_p = c + aw_p \qquad w_p \text{ は } p = G(w) \text{ となる } w \text{ の値} \tag{41}$$

(4) 母数推定 (積率法)

$$\bar{x} = \frac{1}{N} \sum_{j=1}^{N} x_j \qquad S_x^2 = \frac{1}{N} \sum_{j=1}^{N} (x_j - \bar{x})^2 \qquad C_{sx} = \frac{1}{N} \sum_{j=1}^{N} \left(\frac{x_j - \bar{x}}{S_x} \right)^3 \tag{42}$$

$$\mu_x = \bar{x} \qquad \sigma_x = [N/(N-1)]^{1/2} S_x \qquad \gamma_x = \frac{\sqrt{N(N-1)}}{N-2} C_{sx} \tag{43}$$

不偏ひずみ係数 γ_x については，以下に示す，Bobee と Robitaille による偏り補正した式 (ピアソン III 型分布の場合) も用いられる．( ) 内で B に乗ずるのは C_{sx} の 2 乗であることに注意.

$$\gamma_x = C_{sx}(A + B \cdot {C_{sx}}^2) \tag{44}$$

$$\text{ここに} \qquad A = 1 + 6.51/N + 20.2/N^2 \qquad B = 1.48/N + 6.77/N^2 \tag{45}$$

ここで，積率と母数の関係は以下の通り.

$$\mu_x = c + a \cdot b \qquad {\sigma_x}^2 = a^2 \cdot b \qquad \gamma_x = \frac{2a}{|a|\sqrt{b}} \tag{46}$$

以上より，母数は以下に示す関係より推定する.

$$\begin{cases} b = 4/{\gamma_x}^2 & (b > 0) \\ a = \sigma_x/\sqrt{b} & (\gamma_x < 0 \rightarrow a = -\sigma_x/\sqrt{b} < 0) \\ c = \mu_x - ab \end{cases} \tag{47}$$

ここで，$\gamma_x < 0$ の場合は $a < 0$ であり，w_p は $1 - p$ に対する値とすることに注意する.

なお，γ_x の絶対値が小さい場合，対数ピアソン III 型と同じ現象が発生する可能性がある．ただし数値的には未確認.

2.4 対数ピアソン III 型分布 (LP3 分布)

水文統計量 x の対数変換値 $y = \ln x$ が，ピアソン III 型分布に従う場合の母数推定方法を示す．

(1) 確率密度関数

$$f(x) = \frac{1}{|a| \cdot \Gamma(b) \cdot x} \left(\frac{\ln x - c}{a} \right)^{b-1} \cdot \exp\left(-\frac{\ln x - c}{a} \right) \qquad a > 0 : \exp(c) < x < \infty \tag{48}$$

(2) 確率分布関数

$$F(x) = G\left(\frac{\ln x - c}{a} \right) \qquad G(w) = \frac{1}{\Gamma(b)} \int_0^w t^{b-1} \exp(-t) dt \qquad (a > 0) \tag{49}$$

(3) 非超過確率 p に対する水文統計量 x_p

$$w = \frac{\ln x - c}{a} \qquad \rightarrow \qquad x = \exp(c + aw) \tag{50}$$

$$x_p = \exp(c + aw_p) \qquad w_p \text{ は } p = G(w) \text{ となる } w \text{ の値} \tag{51}$$

(4) 母数推定 (積率法)

$$y_j = \ln x_j \qquad \bar{y} = \frac{1}{N} \sum_{j=1}^{N} y_j \qquad S_y^2 = \frac{1}{N} \sum_{j=1}^{N} (y_j - \bar{y})^2 \qquad C_{sy} = \frac{1}{N} \sum_{j=1}^{N} \left(\frac{y_j - \bar{y}}{S_y} \right)^3 \tag{52}$$

$$\mu_y = \bar{y} \qquad \sigma_y = [N/(N-1)]^{1/2} S_y \qquad \gamma_y = \frac{\sqrt{N(N-1)}}{N-2} C_{sy} \tag{53}$$

不偏ひずみ係数 γ_x については，以下に示す，Bobee と Robitaille による偏り補正した式 (ピアソン III 型分布の場合) も用いられる．( ) 内で B に乗ずるのは C_{sy} の 2 乗であることに注意．

$$\gamma_y = C_{sy}(A + B \cdot {C_{sy}}^2) \tag{54}$$

$$\text{ここに} \qquad A = 1 + 6.51/N + 20.2/N^2 \qquad B = 1.48/N + 6.77/N^2 \tag{55}$$

ここで，積率と母数の関係は以下の通り．

$$\mu_y = c + a \cdot b \qquad {\sigma_y}^2 = a^2 \cdot b \qquad \gamma_y = \frac{2a}{|a|\sqrt{b}} \tag{56}$$

以上より，母数は以下に示す関係より推定する．

$$\begin{cases} b = 4/{\gamma_y}^2 & (b > 0) \\ a = \sigma_y/\sqrt{b} & (\gamma_y < 0 \rightarrow a = -\sigma_y/\sqrt{b} < 0) \\ c = \mu_y - ab \end{cases} \tag{57}$$

ここで，$\gamma_y < 0$ の場合は $a < 0$ であり，w_p は $1 - p$ に対する値とすることに注意する．

なお，γ_y の絶対値が小さい場合，b が非常に大きな数値となり，ガンマ分布の%点計算が収束しないケースもあり得る．このため，b が大きな数値の場合は，以下の示す Wilson-Hilferty 変換により非超過確率 p に対する水文統計量 x_p を算定する．

$$x_p = \exp(\mu_y + \sigma_y \cdot K_p) \qquad K_p = \frac{2}{\gamma_y} \left(1 + \frac{\gamma_y z_p}{6} - \frac{{\gamma_y}^2}{36} \right) - \frac{2}{\gamma_y} \tag{58}$$

ここに，z_p は $N(0,1)$ に従う標準正規変量である．Wilson-Hilferty 変換を用いるか否かの境界は，$b < 10,000$ ば目安になる．ひずみ係数 $\gamma_y < 0$ の場合でも，標準正規変量 z_p は p に対して求めればよいし，また K_p の計算でも，γ_y を $|\gamma_y|$ とする必要な無い．すなわち，p も γ_y も計算値そのままを入力すればよい．

2.5 グンベル分布

水文統計量 x が，グンベル分布 (Gumbel distribution) に従う場合の母数推定方法を示す.

(1) 確率密度関数

$$f(x) = \frac{1}{a}\exp\left[-\frac{x-c}{a} - \exp\left(-\frac{x-c}{a}\right)\right] \qquad -\infty < x < \infty \tag{59}$$

(2) 確率分布関数

$$F(x) = \exp\left[-\exp\left(-\frac{x-c}{a}\right)\right] \tag{60}$$

(3) 非超過確率 p に対する水文統計量 x_p

$$p = \exp\left[-\exp\left(-\frac{x-c}{a}\right)\right] \qquad \rightarrow \qquad x = c - a\ln[-\ln(p)] \tag{61}$$

$$x_p = c - a\ln[-\ln(p)] \tag{62}$$

(4) 母数推定 (L 積率法)

$$b_0 = \frac{1}{N}\sum_{j=1}^{N} x_{(j)} \qquad b_1 = \frac{1}{N(N-1)}\sum_{j=1}^{N}(j-1)x_{(j)} \qquad b_2 = \frac{1}{N(N-1)(N-2)}\sum_{j=1}^{N}(j-1)(j-2)x_{(j)} \tag{63}$$

ここに，$x_{(j)}$ は N 個の標本を昇順に並べ替えたときの，小さいほうから j 番目の値を示す.

上記 $b_i = \beta_i$ として，以下の関係により λ_i を算定する.

$$\lambda_1 = \beta_0 \qquad \lambda_2 = 2\beta_1 - \beta_0 \qquad \lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \tag{64}$$

母数は，以下に示す L 積率と母数の関係により推定する.

$$\begin{cases} a = \lambda_2/\ln 2 \\ c = \lambda_1 - 0.5772a \end{cases} \tag{65}$$

2.6 一般化極値分布 (GEV 分布)

水文統計量 x が，一般化極値分布 (Generalized Extreme Value distribution) に従う場合の母数推定方法を示す.

(1) 確率密度関数

$$f(x) = \frac{1}{a}\left(1 - k\frac{x-c}{a}\right)^{1/k-1} \cdot \exp\left[-\left(1 - k\frac{x-c}{a}\right)^{1/k}\right] \qquad (k \neq 0) \tag{66}$$

(2) 確率分布関数

$$F(x) = \exp\left[-\left(1 - k\frac{x-c}{a}\right)^{1/k}\right] \qquad (k \neq 0) \tag{67}$$

(3) 非超過確率 p に対する水文統計量 x_p

$$p = \exp\left[-\left(1 - k\frac{x-c}{a}\right)^{1/k}\right] \qquad \rightarrow \qquad x = c + \frac{a}{k} \cdot \left\{1 - [-\ln(p)]^k\right\} \tag{68}$$

$$x_p = c + \frac{a}{k} \cdot \left\{1 - [-\ln(p)]^k\right\} \tag{69}$$

(4) 母数推定 (L 積率法)

$$b_0 = \frac{1}{N}\sum_{j=1}^{N} x_{(j)} \qquad b_1 = \frac{1}{N(N-1)}\sum_{j=1}^{N}(j-1)x_{(j)} \qquad b_2 = \frac{1}{N(N-1)(N-2)}\sum_{j=1}^{N}(j-1)(j-2)x_{(j)} \tag{70}$$

ここに，$x_{(j)}$ は N 個の標本を昇順に並べ替えたときの，小さいほうから j 番目の値を示す.

上記 $b_i = \beta_i$ として，以下の関係により λ_i を算定する.

$$\lambda_1 = \beta_0 \qquad \lambda_2 = 2\beta_1 - \beta_0 \qquad \lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \tag{71}$$

母数は，以下に示す L 積率と母数の関係により推定する.

$$\begin{cases} k = 7.8590d + 2.9554d^2 \qquad \text{ここに} \qquad d = \dfrac{2\lambda_2}{\lambda_3 + 3\lambda_2} - \dfrac{\ln(2)}{\ln(3)} \\ a = \dfrac{k\lambda_2}{(1 - 2^{-k}) \cdot \Gamma(1+k)} \\ c = \lambda_1 - \dfrac{a}{k} \cdot [1 - \Gamma(1+k)] \end{cases} \tag{72}$$

2.7 平方根指数型最大値分布 (SQRT-ET 分布)

水文統計量 x が，平方根指数型最大値分布 (SQRT exponential-type distribution of maximum) に従う場合の母数推定方法を示す.

(1) 確率密度関数

$$f(x) = \frac{ab}{2} \exp\left[-\sqrt{bx} - a\left(1+\sqrt{bx}\right)\exp\left(-\sqrt{bx}\right)\right] \qquad (x \geqq 0) \tag{73}$$

(2) 確率分布関数

$$F(x) = \exp\left[-a\left(1+\sqrt{bx}\right)\exp\left(-\sqrt{bx}\right)\right] \qquad (x \geqq 0) \tag{74}$$

(3) 非超過確率 p に対する水文統計量 x_p

$$p = \exp\left[-a\left(1+\sqrt{bx}\right)\exp\left(-\sqrt{bx}\right)\right] = \exp\left[-a(1+t_p)\exp(-t_p)\right] \qquad (t_p = \sqrt{bx}) \tag{75}$$

$$\rightarrow \qquad x = \frac{{t_p}^2}{b} \qquad \ln(1+t_p) - t_p = \ln\left[-\frac{1}{a}\ln(p)\right] \tag{76}$$

$$x_p = \frac{{t_p}^2}{b} \qquad \ln(1+t_p) - t_p = \ln\left[-\frac{1}{a}\ln(p)\right] \tag{77}$$

■(参考) t_p の求め方

$$g(t_p) = \ln(1+t_p) - t_p - \ln\left[-\frac{1}{a}\ln(p)\right] \tag{78}$$

とおくと

$$g'(t_p) = \frac{1}{1+t_p} - 1 \tag{79}$$

であり，関数 $g(t_p)$ は単調減少関数であることがわかる．また，通常，$g(0) > 0$ であるため，Newton-Raphson 法により，以下の繰り返しにより $g(t_p) = 0$ となる t_p を求められる.

$$t_{p(n+1)} = t_{p(n)} - \frac{g(t_{p(n)})}{g'(t_{p(n)})} \qquad (n) \text{ は繰り返し回数のカウンタ} \tag{80}$$

t_p の初期値としては，非超過確率 p が比較的大きい領域での x_p を知りたいため，標本値 x の最大値を用いて $t_p = \sqrt{b \cdot x_{max}}$ とすることにより，効率的に t_p を求められる.

(4) 母数推定 (最尤法)

母数 a，b は，以下に示す対数尤度関数 L が最大になるように定める.

$$L(a,b) = \sum_{j=1}^{N} \ln f(x_j)$$

$$= N\ln a + N\ln b - N\ln 2 - \sum_{j=1}^{N}\sqrt{bx_j} - a\left[\sum_{j=1}^{N}\exp\left(-\sqrt{bx_j}\right) + \sum_{j=1}^{N}\sqrt{bx_j}\exp\left(-\sqrt{bx_j}\right)\right] \tag{81}$$

上式 L を b に関して偏微分したものが 0 となる条件より，以下のとおり b の関数として a が定まる．これを a_1 とする．

$$\frac{\partial L}{\partial b} = 0 \qquad \rightarrow \qquad a = \frac{\sum_{j=1}^{N} \sqrt{bx_j} - 2N}{\sum_{j=1}^{N} (bx_j) \exp\left(-\sqrt{bx_j}\right)} = a_1 \tag{82}$$

また，L を a に関して偏微分したものが 0 となる条件より，これも b の関数として a が定まる．これを a_2 とする．

$$\frac{\partial L}{\partial a} = 0 \qquad \rightarrow \qquad a = \frac{N}{\sum_{j=1}^{N} \exp\left(-\sqrt{bx_j}\right) + \sum_{j=1}^{N} \sqrt{bx_j} \exp\left(-\sqrt{bx_j}\right)} = a_2 \tag{83}$$

L が最大となるのは $a_1 = a_2$ のときであるため，$h(b) = a_1(b) - a_2(b) = 0$ となる b の値を，二分法で求めることができる．なお，$a_2 > 0$ は保障されるが，a_1 は，以下の条件を満たすときに $a_1 > 0$ となるため，二分法における b の小さい側の初期値を設定する際注意する．

$$a_1 > 0 \qquad \rightarrow \qquad b > \left(\frac{2N}{\sum_{j=1}^{N} \sqrt{x_j}} \right)^2 \tag{84}$$

二分法における b の大きい側の初期値は，経験的に b が 1 程度以下であることから，以下のようにプログラムすれば良い．(C 言語)

```
/* Bisection method */
b1=bb;                          /* a1>0 となる b1（小さい側初期値）を設定 */
b2=b1+0.5;                      /* 小さい側初期値 b1+0.5 を大きい側初期値 b2 とする */
bb=0.5*(b1+b2);                 /* 中間値 bb の設定 */
f1=FSQR(nd,datax,b1,&a1,&a2); /* h(b1) の計算 */
f2=FSQR(nd,datax,b2,&a1,&a2); /* h(b2) の計算 */
ff=FSQR(nd,datax,bb,&a1,&a2); /* h(bb) の計算 */
do{                             /* 収束計算ループ  */
    if(f1*ff<0.0)b2=bb;
    if(ff*f2<0.0)b1=bb;
    if(ff==0.0)break;
    if(0.0<f1*ff&&0.0<ff*f2){b1=b2;b2=b1+0.5;} /* 見込んだ範囲に解がない場合上側範囲を */
    bb=0.5*(b1+b2);                           /* 0.5 ずつ増加させて解を検索する       */
    f1=FSQR(nd,datax,b1,&a1,&a2);
    f2=FSQR(nd,datax,b2,&a1,&a2);
    ff=FSQR(nd,datax,bb,&a1,&a2);
}while(0.001<fabs(a1-a2));                    /* 収束判定基準：|h(b)|<0.001 */
```

2.8 ワイブル分布 (3 母数)

水文統計量 x が，3 母数ワイブル分布 (Weibull distribution) に従う場合の母数推定方法を示す．

(1) 確率密度関数

$$f(x) = \frac{k}{a}\left(\frac{x-c}{a}\right)^{k-1} \exp\left[-\left(\frac{x-c}{a}\right)^{k}\right] \qquad (k \neq 0) \tag{85}$$

(2) 確率分布関数

$$F(x) = 1 - \exp\left[-\left(\frac{x-c}{a}\right)^{k}\right] \qquad (k \neq 0) \tag{86}$$

(3) 非超過確率 p に対する水文統計量 x_p

$$1 - p = \exp\left[-\left(\frac{x-c}{a}\right)^{k}\right] \qquad \rightarrow \qquad x = c + a[-\ln(1-p)]^{1/k} \tag{87}$$

$$x_p = x = c + a[-\ln(1-p)]^{1/k} \tag{88}$$

(4) 母数推定 (L 積率法：合田らの方法 *) による)

$$b_0 = \frac{1}{N}\sum_{j=1}^{N} x_{(j)} \qquad b_1 = \frac{1}{N(N-1)}\sum_{j=1}^{N}(j-1)x_{(j)} \qquad b_2 = \frac{1}{N(N-1)(N-2)}\sum_{j=1}^{N}(j-1)(j-2)x_{(j)} \tag{89}$$

ここに，$x_{(j)}$ は N 個の標本を昇順に並べ替えたときの，小さいほうから j 番目の値を示す．

上記 $b_i = \beta_i$ として，以下の関係により λ_i を算定する．

$$\lambda_1 = \beta_0 \qquad \lambda_2 = 2\beta_1 - \beta_0 \qquad \lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \tag{90}$$

母数は，以下に示す L 積率と母数の関係により推定する．

$$\begin{cases} k = 285.3\tau^6 - 658.6\tau^5 + 622.8\tau^4 - 317.2\tau^3 + 98.52\tau^2 - 21.256\tau + 3.5160 \qquad \text{ここに} \quad \tau = \lambda_3/\lambda_2 \\ a = \dfrac{\lambda_2}{(1-2^{-1/k})\cdot\Gamma(1+1/k)} \\ c = \lambda_1 - a\cdot\Gamma(1+1/k) \end{cases} \tag{91}$$

*) 合田良実・久高将信・河合弘泰：L-moments 法を用いた波浪の極値統計解析について，土木学会論文集 B2(海岸工学)，Vol.B2-65 No.1，2009，pp161-165

なお，本文 表-2 極値分布関数の母数推定式 のワイブル分布の項で，形状母数 k の値が λ_3 の関数となっているが，$\tau_3 = \lambda 3/\lambda_2$ と思われる．また尺度母数 A の分母に k が入っているがこれも誤植と思われる．

(5) 母数推定 (最尤法)

この問題では推定すべき母数が 3 個あるため，まず最初に位置母数 c を推定し既知とした上で，形状母数 k と尺度母数 a を推定する方法をとる．

(I) 母数 c の推定

標本値 x の非超過確率 $F(x)$ は，選定したプロッティング・ポジション公式により算定する．プロッティング・ポジション公式を用いるため，標本値 x は，昇順に並べ替えておくことに注意する．

確率分布関数を変形して，

$$1 - F(x) = \exp\left[-\left(\frac{x-c}{a}\right)^k\right] \tag{92}$$

上式の両辺の対数を 2 回とることにより

$$\ln\{-\ln[1-F(x)]\} = k\ln(x-c) - k\ln a \qquad \rightarrow \qquad Y = A\cdot X + B \tag{93}$$

標本値を $Y_i = \ln\{-\ln[1-F(x_i)]\}, X_i = \ln(x_i - c)$ として直線回帰することにより，$k = A, a = \exp(-B/A)$ として求めることができる．この際，c の値を，標本値の最小値から小さいほうに変化させ，繰り返し回帰を行いながら，直線回帰による相関係数が最大となる時の c を求める c とする．

この段階で，一応 3 母数 k，a，c は求まったことになるが，更に推定精度を上げるため，c のみを固定し，次の段階で再度 k，a を算定する．ここで求めた k は，以降行う Newton-Raphson 法の初期値として利用するが，a は特に必要としない値である．

(II) 母数 k および a の推定

上記により c は定まったため，$t = x - c$ と置くことにより，

$$f(t) = \frac{k}{a}\left(\frac{t}{a}\right)^{k-1}\exp\left[-\left(\frac{t}{a}\right)^k\right] \tag{94}$$

上式の対数尤度関数 $L = \sum_{i=1}^{N}\ln f(t_i)$ において，以下の条件を満たすよう k および a を Newton-Raphson 法で推定する．k の初期値は，c を求めるときに算定した値を用いればよい．

$$\frac{\partial L}{\partial k} = 0 \qquad \rightarrow \qquad \frac{1}{k} + \frac{\sum_{i=1}^{N}\ln t_i}{N} - \frac{\sum_{i=1}^{N}[(\ln t_i)\cdot {t_i}^k]}{\sum_{i=1}^{N}{t_i}^k} = 0 \tag{95}$$

$$\frac{\partial L}{\partial a} = 0 \qquad \rightarrow \qquad a = \left(\frac{\sum_{i=1}^{N}{t_i}^k}{N}\right)^{1/k} \tag{96}$$

$$g(k) = \frac{1}{k} + \frac{T_0}{N} - \frac{T_2(k)}{T_1(k)} \qquad\qquad g'(k) = -\frac{1}{k^2} - \frac{T_3(k)\cdot T_1(k) - [T_2(k)]^2}{[T_1(k)]^2} \tag{97}$$

$$T_0 = \sum_{i=1}^{N}\ln t_i \qquad T_1(k) = \sum_{i=1}^{N}{t_i}^k \qquad T_2(k) = \sum_{i=1}^{N}[\ln t_i\cdot {t_i}^k] \qquad T_3(k) = \sum_{i=1}^{N}[\ln t_i\cdot\ln t_i\cdot {t_i}^k] \tag{98}$$

k が収束するまで以下の繰り返しを行う．ここに n は繰り返し回数のカウンタである．

$$k_{n+1} = k_n - \frac{g(k_n)}{g'(k_n)} \tag{99}$$

k が定まれば，次式により a が定まる．

$$a = \left(\frac{\sum_{i=1}^{N}{t_i}^k}{N}\right)^{1/k} \tag{100}$$

3. ジャックナイフ法による推定

Jackknife 法による偏りを補正した推定値 (JackKnife 推定値) と標準誤差の推定値の計算方法を示す．手順は以下の通り．

① N 個のデータ $x_{(i)}$ を用いて統計量の推定値 $\hat{\theta}$ を求める．

② i 番目のデータを除いた $N-1$ 個のデータによる推定値 $\hat{\theta}_{(i)}$ を求める．

③ $n-1$ 個のデータによる推定値 $\hat{\theta}_{(i)}$ は N 個求められるため，これらの平均 $\hat{\theta}_{(\cdot)}$ を求める．

$$\hat{\theta}_{(\cdot)} = \frac{1}{N}\sum_{i=1}^{N}\hat{\theta}_{(i)} \tag{101}$$

④ N 個のデータ $x_{(i)}$ を用いて統計量の推定値 $\hat{\theta}$ と，$N-1$ 個のデータによる推定値の平均 $\hat{\theta}_{(\cdot)}$ を用いて，以下の式で jackknife 推定値 $\bar{\theta}$ を求める．

$$\bar{\theta} = N\cdot\hat{\theta} - (N-1)\cdot\hat{\theta}_{(\cdot)} \tag{102}$$

⑤ θ の標準誤差の推定値 (SE) は，以下の通り与えられる．

$$(SE) = \sqrt{\frac{N-1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_{(i)} - \hat{\theta}_{(\cdot)}\right)^2} \tag{103}$$

4. ブートストラップ法による推定

bootstrap 法による点推定と区間推定の方法を述べる．

① N 個のデータ $x_{(i)}$ を用いて統計量の推定値 $\hat{\theta}$ を求める．

② 元の N 個のデータから繰り返しを許して N 個のデータを無作為に抽出しこれの推定値 $\theta*_{(i)}$ を求める．

③ $\theta*_{(i)}$(bootstrap 反復値) を B 個求め，これらの平均を bootstrap 法による点推定とする．

$$\hat{\theta*} = \frac{1}{B}\sum_{i=1}^{B}\theta*_{(i)} \tag{104}$$

④ bootstrap 反復値の分布は，母数の分布を表していると考えられるため，bootstrap 反復値を最小値から最大値まで昇順に並べ，反復値が所要の下側確率・上側確率となる限界を見つけ，これを信頼限界とする．この方法は，母数の分布を正規分布と仮定する必要はない．この方法はパーセンタイル法 (percentile method) と呼ぶ．

5. 棄却検定

ある特定の分布に従う標本の中に，希にしか起こらない特に大きな値あるいは小さな値が含まれている場合，これを同一母集団のデータとして採用すべきか棄却してよいかの検定を行う手順を述べる．

課題 ある母集団から採取した N 個の標本値に対し，その最大値を検定対象値 x_ϵ として棄却検定を行う．

① 検定対象値を除いた $N-1$ 個のデータより確率分布の母数推定を行い，確率分布関数 $F(x)$ を推定する．

② 推定した確率分布関数より，検定対象値の上側確率 (超過確率)q を推定する．

③ 検定対象値の上側確率が等しくなる標準正規分布の%点を算定し，これを u_ϵ とする．すなわち以下の関係を満足する u_ϵ を求める．ここに Φ は標準正規分布の確率分布関数である．

$$\text{上側確率} \quad q = 1 - F(x_\epsilon) = 1 - \Phi(u_\epsilon) \tag{105}$$

④ 検定対象値が他の $N-1$ 個の標本値と同一の母集団であるなら，次式で表される F 値は自由度 $(1, M-1)$ の F 分布に従うはずである．ここに，$M(=N-1)$ は検定対象値を除いたデータ数であることに注意.

$$F = \left(\frac{M-1}{M+1}\right) \cdot {u_\epsilon}^2 \tag{106}$$

よって以下を満足する F 分布の上側確率 2ϵ を求める．

$$F^1_{M-1}(2\epsilon) = \left(\frac{M-1}{M+1}\right) \cdot {u_\epsilon}^2 \tag{107}$$

⑤ 危険率 β に対する棄却限界 ϵ の値を求める．

$$\epsilon_0 = 1 - (1-\beta_0)^{1/N} \qquad (\beta \text{ は通常 } 5\%) \tag{108}$$

⑤ 検定対象値より定めた ϵ と棄却限界 ϵ_0 を比較し，以下の判定を行う．

$$\begin{cases} \epsilon \leqq \epsilon_0 & \text{棄却限界より検定対象値の上側確率は小さく} \\ & \text{検定対象値は棄却できる (解析に用いない)} \\ \epsilon > \epsilon_0 & \text{棄却限界より検定対象値の上側確率は大きく} \\ & \text{検定対象値は棄却できない (解析に用いる)} \end{cases} \tag{109}$$