３．最尤法によるパラメータ推定

1. 推定量のもつべき特性

試験データを得る目的の一つとして、そのデータが属する母集団分布の性質を知りたい場合があります。

知りたい母集団の性質としては 1）分布の位置を表す指標（平均、中央値、最頻値等）、2）分布の広がりを表す指標（分散、標準偏差等）、3）分布の母数などがありますが、これらはデータから推定（estimation）することになります。

推定はデータの関数である推定量（estimator）を用いて行われます。データは確率変数のため、推定量はある確率分布に従うことになります。関数である推定量にデータを代入して得られる値を推定値（estimate）と言います。

当然ながらその推定量はできるだけ母集団の性質を正確に表すものを選ぶ必要がありますが、良い推定量とはどのような特性を持てば良いのでしょうか？

特性として 1）標本の大きさ n を固定して、その未知パラメータを算出し、この作業を繰り返し行った場合の特性。（小標本特性）、2）標本の大きさ n が非常に大きくなっていく場合の特性。（大標本特性）の 2 つを考えることが必要です。

真のパラメータを θ、確率分布である推定量を $\hat{\theta}$ とした場合 $E\left[\hat{\theta}\right]=\theta$ となる性質をもつ推定量を**不偏推定量**と呼びます。（$E[\ \ ]$：期待値）

ここで、x_i を平均 μ 、分散 σ^2 の母集団からの無作為標本とします。標本平均は $\bar{x}=\frac{1}{n}\cdot\sum_{i=1}^{n}x_i$ 、標本分散は $\frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2$ となり、期待値は $E\left[\bar{x}\right]=\frac{1}{n}\cdot\sum_{i=1}^{n}E\left[x_i\right]=\frac{1}{n}\cdot n\cdot\mu=\mu$ となり、$E\left[\hat{\theta}\right]=\theta$ を満たしますから標本平均（$\bar{x}$）は不偏推定量です。

分散の不偏推定量を導きましょう。そのために $x_i=\left(x_i-\mu\right)+\mu$ として両辺を二乗し、その期待値が 2）式となることを利用します。

$$x_i{}^2=\left\{\left(x_i-\mu\right)+\mu\right\}^2=\left(x_i-\mu\right)^2+\mu^2+2\cdot\left(x_i-\mu\right)\cdot\mu \qquad \cdot\cdot\cdot\ 1)$$

$$E\left[x_i^{\ 2}\right]=E\left[\left(x_i-\mu\right)^2\right]+E\left[\mu^2\right]+E\left[2\cdot\left(x_i-\mu\right)\cdot\mu\right]=\sigma^2+\mu^2 \qquad \cdots 2)$$
$$\because E\left[\left(x_i-\mu\right)\right]=0$$

さらに、標本平均を $\overline{x}=\frac{1}{n}\cdot\sum_{i=1}^{n}x_i=\frac{1}{n}\cdot\sum_{i=1}^{n}\left\{\left(x_i-\mu\right)+\mu\right\}=\mu+\frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i-\mu\right)$ として、その二乗値は 3）式となります。

$$\overline{x}^2=\mu^2+\frac{1}{n^2}\cdot\sum_{i=1}^{n}\sum_{j=1}^{n}\left(x_i-\mu\right)\cdot\left(x_j-\mu\right)+\frac{2\cdot\mu}{n}\cdot\sum_{i=1}^{n}\left(x_i-\mu\right)^2 \qquad \cdots 3)$$

その期待値 $E\left[\overline{x}^2\right]$ は 4）式を使って 5）式で表現できることになります。

$$E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\left(x_i-\mu\right)\cdot\left(x_j-\mu\right)\right]=E\left[\sum_{i=1}^{n}\left(x_i-\mu\right)^2+\sum_{i=1}^{n}\sum_{j\neq j}^{n}\left(x_i-\mu\right)\cdot\left(x_j-\mu\right)\right]=\sigma^2+0=\sigma^2 \qquad \cdots 4)$$

$$E\left[\overline{x}^2\right]=\mu^2+\frac{1}{n^2}\cdot\left(n\cdot\sigma^2\right)=\mu^2+\frac{\sigma^2}{n} \qquad \cdots 5)$$

標本平方和 $\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2$ は 6）式で表現でき、その期待値は 1）式、5）式の結果を代入して 7）式となりますから、標本平方和の期待値を $n-1$ で除してやることにより分散（σ^2）の不偏推定量が得られることになります。

$$\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2=\sum_{i=1}^{n}x_i^{\ 2}-2\cdot\overline{x}\cdot\sum_{i=1}^{n}x_i+n\cdot\overline{x}^2=\sum_{i=1}^{n}x_i^{\ 2}-n\cdot\overline{x}^2 \qquad \cdots 6)$$

$$E\left[\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\right]=E\left[\sum_{i=1}^{n}x_i^{\ 2}\right]-n\cdot E\left[\overline{x}^2\right]=n\cdot\sigma^2+\mu^2-n\cdot\left(\mu^2+\frac{\sigma^2}{n}\right)=\left(n-1\right)\cdot\sigma^2 \qquad \cdots 7)$$

すなわち、n 個のデータ $x_1,x_2,\cdots,x_n$ からの母分散の不偏推定量（V_{es}）は 7）式で与えられます。

$$V_{es}=\frac{1}{n-1}\cdot\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2 \qquad \cdots 8)$$

ところで、無作為標本の加重平均 $\overline{x}_w$（$\overline{x}_w=\sum_{i=1}^{n}w_i\cdot x_i,\sum_{i=1}^{n}w_i=1$）は不偏推定量です。なぜなら 8）式となるからです。

$$E\left[\overline{x}_w\right]=E\left[\sum_{i=1}^{n}w_i\cdot x_i\right]=\sum_{i=1}^{n}w_i\cdot E\left[x_i\right]=\sum_{i=1}^{n}w_i\cdot\mu=\mu \qquad \cdot\cdot\cdot\ 9)$$

このことから、m 個のデータ $x_1, x_2, \cdots, x_m$ が得られた場合、母分散（μ）の不偏推定量は $\overline{x}=\frac{1}{m}\cdot\sum_{i=1}^{m}x_i$ でも、$\overline{x}_w=\sum_{i=1}^{m}w_i\cdot x_i$ でも良いことになります。それではまずいということで、不偏推定量の中でバラツキ（分散）が最も少ないものを使おうと考えられました。この推定量を有効推定量（efficient estimator）と呼びます。

分散が最も小さいかどうかを調べる有力な方法としては、クラメール・ラオの不等式があります。不偏推定量を$\hat{\theta}$、分散を$V\left[\hat{\theta}\right]$、確率密度関数を $f\left(x\right)$ とした場合 9）式で表現されます。

$$V\left(\hat{\theta}\right)\geq\frac{1}{n\cdot E\left[\frac{\partial\ln\left\{f\left(x\right)\right\}}{\partial\theta}\right]^2}=\frac{1}{-n\cdot E\left[\frac{\partial^2\ln\left\{f\left(x\right)\right\}}{\partial\theta^2}\right]} \qquad \cdot\cdot\cdot\ 10)$$

すなわち、$V\left[\hat{\theta}\right]$の最小値は$\frac{1}{n\cdot E\left[\frac{\partial\ln\left\{f\left(x\right)\right\}}{\partial\theta}\right]^2}$で計算できることを示しています。

正規分布の母平均 μ の有効推定量を算出してみましょう。確率密度関数は 10）式で与えられます。$\frac{1}{-n\cdot E\left[\frac{\partial^2\ln\left\{f\left(x\right)\right\}}{\partial\theta^2}\right]}$は確率密度関数を μ で偏微分をして 11）式となります。

$$f_{norm}\left(x,\mu,\sigma^2\right)=\frac{1}{\sqrt{2\cdot\pi\cdot\sigma^2}}\cdot exp\left[\frac{-1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}\right] \qquad \cdot\cdot\cdot\ 11)$$

$$\frac{\partial\ln\left\{f\left(x\right)\right\}}{\partial\mu}=\frac{x-\mu}{\sigma^2}\quad\rightarrow\quad\frac{\partial^2\ln\left\{f\left(x\right)\right\}}{\partial\mu^2}=-\frac{1}{\sigma^2} \qquad \cdot\cdot\cdot\ 12)$$

すなわち、μ の不偏推定量の最小分散は$\frac{1}{-n\cdot E\left[\frac{\partial^2\ln\left\{f\left(x\right)\right\}}{\partial\theta^2}\right]}=\frac{1}{-n\cdot\left(-\frac{1}{\sigma^2}\right)}=\frac{\sigma^2}{n}$であること

になります。

そして、標本平均の分散は 4）式から 12）式となり 11）式と同値になりますので、母平均μの不偏推定量の最小分散に一致しますから標本平均（$\bar{x}$）は有効推定量となります。

$$V\left(\bar{x}\right)=E\left[\left(\bar{x}-\mu\right)^2\right]=E\left[\bar{x}^2-2\cdot\bar{x}\cdot\mu+\mu^2\right]=\left(\mu^2+\frac{\sigma^2}{n}-\mu^2\right)=\frac{\sigma^2}{n} \qquad \cdot\cdot\cdot \quad 13)$$

このように、確率密度関数$f\left(x\right)$を仮定できて有効推定量を算出できる場合には推定量の良し悪しの議論はこれ以上進める必要はないことにしましょうというのが暗黙の了解事項です。

しかしながら、確率密度関数$f\left(x\right)$を仮定できても有効推定量を算出できない場合や確率密度関数$f\left(x\right)$が仮定できない（しない）場合にはもう一歩議論を進める必要があります。

例えば、正規分布と仮定できない場合の標本平均と加重平均の分散の最小値はどちらかという問題です。

上述したように加重平均$\bar{x}_w=\sum_{i=1}^{n}w_i\cdot x_i$は$\sum_{i=1}^{m}w_i=1$なる条件付の不偏推定量です。加重平均$\bar{x}_w$の分散を算出してみましょう。定義から加重平均の分散は次式となります。

$$\begin{aligned}V\left(\bar{x}_w\right)&=E\left[\bar{x}_w-E\left(\bar{x}_w\right)\right]^2=E\left[\sum_{i=1}^{n}w_i\cdot x_i-\mu\right]^2=E\left[\sum_{i=1}^{n}w_i\cdot x_i-\mu\cdot\sum_{i=1}^{n}w_i\right]^2\\&=E\left[\sum_{i=1}^{n}w_i\cdot\left(x_i-\mu\right)\right]^2=E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}w_i\cdot w_j\cdot\left(x_i-\mu\right)\cdot\left(x_j-\mu\right)\right]\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}w_i\cdot w_j\cdot E\left[\left(x_i-\mu\right)\cdot\left(x_j-\mu\right)\right]=\sigma^2\cdot\sum_{i=1}^{n}{w_i}^2\end{aligned}$$

$\because x_i,x_j$は独立　$E\left[\left(x_i-\mu\right)\cdot\left(x_j-\mu\right)\right]=0 \quad i\neq j$

加重平均の分散は$V\left(\bar{x}_w\right)=\sigma^2\cdot\sum_{i=1}^{n}{w_i}^2$となることが分かりました。

では分散の最小値はどうなるのでしょうか？　$\sum_{i=1}^{m}w_i=1$の条件付の最小値問題を解くことになります。

条件付の最小値はラグランジュの未定乗数法を使い算出しますが、結論だけを述べると

$w_i = \dfrac{1}{n} \quad \left(i = 1, 2, \cdots n \right)$ となります。ですから、加重平均の分散が最小値になる推定量は

$\overline{x}_w = \sum_{i=1}^{n} w_i \cdot x_i = \sum_{i=1}^{n} \dfrac{1}{n} \cdot x_i = \dfrac{1}{n} \cdot \sum_{i=1}^{n} x_i$ となります。

このことから加重平均の中で標本平均が最も小さい分散を与えることが分かります。加重平均は一般的な用語でいうと線形です。そこで、線形で最も小さな分散を与える不偏推定量は最良線形不偏推定量（Best Linear Unbiased Estimator <BLUE>）と呼ばれます。すなわち、標本平均は最良線形不偏推定量です。（正規分布とは仮定していません！）

当然、有効推定量も最良線形不偏推定量（BLUE）も不偏推定量です。不偏推定量は $E\left[\hat{\theta}\right] = \theta$ なる性質を持っていますのでこれで推定量の良し悪しの基準にしようとするものです。

一方、推定量（$\hat{\theta}$）と真の値（θ）との差の二乗和（$E\left[\left(\hat{\theta} - \theta\right)^2\right]$）は平均平方誤差（mean square error<MSE>）と呼ばれますが、この平均平方誤差を最小とする推定量（MSE 最小推定量）が最も良い推定量とする考え方もあります。

分散（σ^2）の MSE 最小推定量は、$\dfrac{1}{\sigma^2} \cdot \sum_{i=1}^{n} \left(x_i - \overline{x} \right)$ が自由度 $n-1$ の χ^2 分布となることを利用して $\dfrac{1}{n+1} \cdot \sum_{i=1}^{n} \left(x_i - \overline{x} \right)^2$ となります。

また後述しますが、最尤推定量は $\dfrac{1}{n} \cdot \sum_{i=1}^{n} \left(x_i - \overline{x} \right)^2$ で、上述したように不偏推定量は $\dfrac{1}{n-1} \cdot \sum_{i=1}^{n} \left(x_i - \overline{x} \right)^2$ です。

n は自然数のため、$\dfrac{1}{n-1} > \dfrac{1}{n} > \dfrac{1}{n+1}$ となります。このため、不偏推定量＞最尤推定量＞MSE 最小推定量の順に小さくなります。

このように、色々な推定量を選ぶ基準があり、これらの基準を全て満たす推定量があるわけではありません。

すなわち、『全員が満足する推定量はない。』、『ベストな推定量はない。』ことになります。

重要なことは、採用した推定量はどのような方法で算出されたかを押さえ、その推定量の性質を理解しておくことかと思います。

2.　最尤法によるパラメータ推定

最尤推定量と呼ばれる推定量があります。最尤推定量はデータ数nが大きい場合、1）元の分布が何であれ正規分布に従う性質（漸近正規性）2）漸近的に不偏推定量になる性質があることが証明されている推定量で、確率密度関数が想定される場合には最も有力な推定量となります。

2.1　算出方法の基礎

確率密度関数（$f(x,\theta)$）を持つ確率分布からの観測値を$(x_1, x_2, \cdots, x_n)$とした場合に13）式で定義される関数を尤度関数と呼び、この関数値が最大になるようにθを推定する方法を最大尤度法、略して最尤法（maximum likelihood method）と言います。

$$L_x = f(x_1,\theta) \cdot f(x_2,\theta) \cdots f(x_n,\theta) = \prod_{i=1}^{n} f(x_i,\theta) \qquad \cdot\cdot\cdot \quad 14)$$

そして、最尤法で求めたられた推定量は最尤推定量と呼ばれます。文字通り最も尤もらしい推定量です。

尤度関数はデータx_iを固定してθの関数と考えますが、逆にθを固定してx_iの関数と見た場合には同時確率密度関数と呼ばれています。なお、確率変数はお互いに独立と仮定しています。

尤度関数は積の関数ですが 13）式の両辺を自然対数で取ると和の関数として表現でき、L_xを最大にするθ_{es}も、L_xの対数を最大にするθ_{es}も同値ですから、13）式の対数を取った 14)式を最大にするθ_{es}を求めても良いことになります。（LL_xは対数尤度関数と呼ばれます。）

$$LL_x = \ln\left(\prod_{i=1}^{n} f(x_i,\theta_{es})\right) = \sum_{i=1}^{n} \ln\left(f(x_i,\theta_{es})\right) \qquad \cdot\cdot\cdot \quad 15)$$

すなわち、LL_xをθで偏微分した方程式（最尤方程式）を満たすθを求め、最大値になるかを調べることにより最尤推定量を求めることになります。通常、対数尤度関数は上に凸なので最尤方程式を満たすθは最大値となります。

$$\frac{\partial}{\partial\theta} LL_x = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \ln\left(f(x_i,\theta)\right) = 0 \qquad \cdot\cdot\cdot \quad 16)$$

なお、最尤推定量（maximum likelihood estimator）はその頭文字を使って MLE と記述される場合があります。

2.2　完全データの最尤推定量

1）正規分布、対数正規分布

$\left(x_1, x_2, \cdots, x_n\right)$ を平均が μ_x、分散が V_x の正規分布からの標本（データ）とします。

その確率密度関数は $f\left(x_i, \mu_x, V_x\right)=\frac{1}{\sqrt{2\cdot\pi\cdot V_x}}\cdot\exp\left(-\frac{\left(x_i-\mu_x\right)^2}{2\cdot V_x}\right)$ で表現されますから、対数尤度関数は 16）式となります。（分散を σ^2 と記載すると、標準偏差 σ で偏微分するという恐れがあるため、ここでは分散（V）で表現しました。）

$$LL_{norm}=\sum_{i=1}^{n}\ln\left(\frac{1}{\sqrt{2\cdot\pi\cdot V_{es}}}\cdot\exp\left(-\frac{\left(x_i-\mu_{es}\right)^2}{2\cdot V_{es}}\right)\right)=n\cdot\ln\left(\frac{1}{\sqrt{2\cdot\pi\cdot V_{es}}}\right)-\frac{1}{2\cdot V_{es}}\cdot\sum_{i=1}^{n}\left(x_i-\mu_{es}\right)^2$$

・・・ 17）

未知パラメータは μ_{es}, V_{es} ですから、17）式の最尤方程式が得られます。

$$\begin{aligned}\frac{\partial}{\partial\mu_{es}}LL_{norm}&=\frac{1}{V_x}\cdot\sum_{i=1}^{n}\left(x_i-\mu_{es}\right)=0\\ \frac{\partial}{\partial V_{es}}LL_{norm}&=-\frac{n}{2\cdot V_{es}}+\frac{1}{2\cdot V_{es}^2}\cdot\sum_{i=1}^{n}\left(x_i-\mu_{es}\right)=0\end{aligned}$$

・・・ 18）

17）式の最尤方程式を解けば 18）式となります。すなわち、正規分布の $\mu, V=\sigma^2$ の最尤推定量は標本平均、標本分散（不偏分散ではありません）となります。

$$\mu_{es}=\frac{1}{n}\cdot\sum_{i=1}^{n}x_i \quad , \quad V_{es}=\frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i-\mu_{es}\right)^2$$

・・・ 19）

同様に対数正規分布の λ, ξ の最尤推定量は 19）式となり、対数正規分布の標本平均、標本分散と同値になります。

$$\lambda_{es}=\frac{1}{n}\cdot\sum_{i=1}^{n}\ln\left(x_i\right) \quad , \quad \xi_{es}{}^2=\frac{1}{n}\cdot\sum_{i=1}^{n}\left(\ln\left(x_i\right)-\lambda_{es}\right)^2$$

・・・ 20）

2）ワイブル分布

a）2 母数ワイブル分布

2 母数ワイブル分布の確率密度関数は 20）式で表現されますので、その対数尤度関数は 21）式となります。

$$f_{2pwei}(x,m,\eta) := \left(\frac{x}{\eta}\right)^{m-1}\cdot\frac{m}{\eta}\cdot exp\left[-\left(\frac{x}{\eta}\right)^{m}\right] \qquad 0 < x < \infty \quad 0 < m < \infty \quad 0 < \eta < \infty \qquad \cdots\ 21)$$

$$LL_{\ 2pwei} = \sum_{i=1}^{n} ln\left[\left(\frac{x_i}{\eta_{\ es}}\right)^{m_{\ es}-1}\cdot\frac{m_{\ es}}{\eta_{\ es}}\cdot exp\left[-\left(\frac{x_i}{\eta_{\ es}}\right)^{m_{\ es}}\right]\right] \qquad \cdots\ 22)$$

18）式を η_{es} で偏微分して 0 とおくと η_{es} は 22）式となります。

$$\frac{d}{d\eta_{es}}LL_{2pwei} = -n\cdot\frac{m_{es}}{\eta_{es}} + \sum_{i=1}^{n}\left[(x_i)^{m_{es}}\cdot\left(\frac{1}{\eta_{es}}\right)^{m_{es}}\cdot\frac{m_{es}}{\eta_{es}}\right] = 0 \ \rightarrow \ \eta_{es} = \left[\frac{\sum_{i=1}^{n}(x_i)^{m_{es}}}{n}\right]^{\frac{1}{m_{es}}}$$

$$\cdots\ 23)$$

同様にして m_{es} で偏微分すると 23）式が得られます。

$$\frac{1}{m_{\ es}} + \frac{1}{n}\cdot\sum_{i=1}^{n} ln(x_i) - \frac{\sum_{i=1}^{n}\left[(x_i)^{m_{\ es}}\cdot ln(x_i)\right]}{\sum_{i=1}^{n}(x_i)^{m_{\ es}}} = 0 \qquad \cdots\ 24)$$

すなわち、2 母数ワイブル分布の m,η の最尤推定値は始めに 23）式から m_{es} を算出し、η_{es} に関する 22）式に代入することにより η_{es} を算出することができます。

ｂ） ３母数ワイブル分布

３母数ワイブル分布の確率密度関数は 24）式で与えられ対数尤度関数は 25）式となります。

$$f_{3pwei}(x,m,\eta,\gamma) := \left(\frac{x-\gamma}{\eta}\right)^{m-1}\cdot\frac{m}{\eta}\cdot exp\left[-\left(\frac{x-\gamma}{\eta}\right)^{m}\right] \qquad \begin{matrix}\gamma < x < \infty \\ 0 < m < \infty \\ 0 < \eta < \infty\end{matrix} \qquad \cdots 25)$$

$$LL_{3pwei} = \sum_{i=1}^{n} ln\left[\left(\frac{x_i-\gamma_{es}}{\eta_{es}}\right)^{m_{es}-1}\cdot\frac{m_{es}}{\eta_{es}}\cdot exp\left[-\left(\frac{x_i-\gamma_{es}}{\eta_{es}}\right)^{m_{es}}\right]\right] \qquad \cdots 26)$$

LL_{3pwei} を $\gamma_{es}, m_{es}, \eta_{es}$ で偏微分すれば（3.26)、(3.27)、(3.28）式となります。

$$\frac{d}{d\gamma_{es}}LL_{3pwei} = \sum_{i=1}^{n}\left[\frac{-m_{es}}{x_i-\gamma_{es}}+\frac{1}{x_i-\gamma_{es}}+\left(\frac{x_i-\gamma_{es}}{\eta_{es}}\right)^{m_{es}}\cdot\frac{m_{es}}{x_i-\gamma_{es}}\right] = 0 \qquad \cdots 27)$$

$$\frac{d}{dm_{es}}LL_{3pwei} = -n\cdot ln(\eta_{es}) + \frac{n}{m_{es}} + \sum_{i=1}^{n}\left[ln(x_i-\gamma_{es}) - \left(\frac{x_i-\gamma_{es}}{\eta_{es}}\right)^{m_{es}}\cdot ln\left(\frac{x_i-\gamma_{es}}{\eta_{es}}\right)\right] = 0$$

$$\cdots 28)$$

$$\frac{d}{d\eta_{es}}LL_{3pwei} = -n\cdot\frac{m_{es}}{\eta_{es}} + \frac{m_{es}}{\eta^{m_{es}+1}}\cdot\sum_{i=1}^{n}(x_i-\gamma_{es})^{m_{es}} = 0 \qquad \cdots 29)$$

28）式から 29）式が得られます。

$$\eta_{es} = \left[\frac{\sum_{i=1}^{n}(x_i-\gamma)^{m_{es}}}{n}\right]^{\frac{1}{m_{es}}} \qquad \cdots 30)$$

29）式を 26)、27）式に代入してそれぞれ 30）式、31）式が得られます。

$$\frac{-\sum_{i=1}^{n}\left[(x_i-\gamma_{es})^{m_{es}}\cdot ln(x_i-\gamma_{es})\right]}{\sum_{i=1}^{n}(x_i-\gamma_{es})^{m_{es}}}\cdot n + \sum_{i=1}^{n} ln(x_i-\gamma_{es}) + \frac{n}{m_{es}} = 0 \qquad \cdots 31)$$

$$\frac{(-m_{es}+1)}{n \cdot m_{es}} \cdot \sum_{i=1}^{n} \frac{1}{x_i - \gamma_{es}} + \frac{\sum_{i=1}^{n} (x_i - \gamma_{es})^{m_{es}-1}}{\sum_{i=1}^{n} (x_i - \gamma_{es})^{m_{es}}} = 0 \quad \cdots 32)$$

すなわち、30)、31）式の連立方程式から m_{es}, λ_{es} を求め、その結果を 19）式に代入して η_{es} を算出することになります。

2.3　打ち切りデータの最尤推定量

打ち切りデータ（censored data）とは観測値がある値以上あるいはある値以下である個数は分かっているが実際の値は得られていないときのデータのことを言います。

1) 保証荷重試験を行った場合、2) 強度試験で試験機の能力から破壊荷重が得られない場合、3) 長期荷重調整係数試験で破壊していない試験体がある場合、4) せん断強度を算出するための試験で他の要因（例えば曲げ）で破壊し、せん断強度が得ることができなかった場合などが打ち切りデータの構造を持っています。

打ち切りデータは切れた分布（truncated distribution）からのデータと混同される場合がありますので注意が必要です。

値が得られているデータ n 個（$x_L < x_1, x_2, \cdots, x_n < x_U$）、$x_L$ 以下の個数を m_L 、x_U 以上の個数を m_U とします。

x_L 以下の m_L 個の確率は確率分布関数（$F(x_L)$）を使えば $F(x_L)^{m_L}$ 、x_U 以上の確率は $\int_{x_U}^{\infty} f(z)dz = 1 - F(x_U)$ で与えられますから、個数 m_U では $(1-F(x_U))^{m_U}$ となります。このため、尤度関数、対数尤度関数は 32)、33）式となります。

$$L_X = F_X(x_L)^{m_L} \cdot \left\{ \prod_{i=1}^{n} f_X(x_i) \right\} \cdot \left\{ 1 - F_X(x_U) \right\}^{m_U} \quad \cdots 33)$$

ここに、$f_X(x), F_X(x)$ は確率密度関数、確率分布関数

$$LL_X = \ln\left[L_X\right] = m_L \cdot \ln\left(F_X\left(x_L\right)\right) + m_U \cdot \ln\left(1 - F_X\left(x_U\right)\right) + \sum_{i=1}^{n} \ln\left(f_X\left(x_i\right)\right) \quad \cdot\cdot\cdot \quad 34)$$

最尤推定量は対数尤度関数の最大値として定義されますから 33）式を未知パラメータで偏微分して最尤方程式から完全データと同じ手順で算出されます。

12）式と異なり 33）式には確率分布関数 $F_X\left(x\right)$ が入っていますので、$F_X\left(x\right)$ が簡単に偏微分できない正規分布、対数正規分布などの場合には少々工夫が必要になります。そこで先にワイブル分布を取り上げます。

なお、以下では、x_U 以上の個数を m_U 、値が得られているデータ n 個（$x_1, x_2, \cdots, x_n < x_U$ ）とした片側の打ち切りデータとします。

2.3.1　ワイブル分布

２母数ワイブル分布の確率密度関数は 34）式、確率分布関数は 35）式で与えられます。

$$f_{2pwei}(x, m, \eta) := \left(\frac{x}{\eta}\right)^{m-1} \cdot \frac{m}{\eta} \cdot exp\left[-\left(\frac{x}{\eta}\right)^{m}\right] \qquad \begin{array}{l} 0 < x < \infty \\ 0 < m < \infty \\ 0 < \eta < \infty \end{array} \qquad \cdot\cdot\cdot \quad 35)$$

$$F_{2pwei}(x, m, \eta) = 1 - exp\left[-\left(\frac{x}{\eta}\right)^{m}\right] \qquad \cdot\cdot\cdot \quad 36)$$

そして、対数尤度関数は 36）式になります。完全データの場合と異なるのは第１項目です。

$$LL_{2p} = m_U \cdot \left[-x_U^{m_{es}} \cdot \left(\frac{1}{\eta_{es}}\right)^{m_{es}}\right] + \sum_{i=1}^{n} ln\left[\left(\frac{x_i}{\eta_{es}}\right)^{m_{es}-1} \cdot \frac{m_{es}}{\eta_{es}} \cdot exp\left[-\left(\frac{x_i}{\eta_{es}}\right)^{m_{es}}\right]\right]$$

$$\cdot\cdot\cdot \quad 37)$$

対数尤度関数を η_{es} で偏微分して 0 と置き整理すると 37）式、m_{es} で偏微分して 0 と置き整理すると 38）式となります。

$$\eta_{es} = \left[\frac{m_U \cdot x_U^{m_{es}} + \sum_{i=1}^{n} (x_i)^{m_{es}}}{n} \right]^{\frac{1}{m_{es}}} = \left(\frac{\alpha}{n} \right)^{\frac{1}{m}} = \left(\frac{m_U \cdot x_U^{m_{es}} + Sxm}{n} \right)^{\frac{1}{m_{es}}} \quad \cdots \; 38)$$

$$\frac{m_U \cdot x_U^{m_{es}} \cdot ln(x_U) + \sum_{i=1}^{n} \left[(x_i)^{m_{es}} \cdot ln(x_i) \right]}{m_U \cdot x_U^{m_{es}} + \sum_{i=1}^{n} (x_i)^{m}} - \frac{1}{m_{es}} - \frac{1}{n} \cdot \sum_{i=1}^{n} ln(x_i) = 0 \quad \cdots \; 39)$$

最尤推定量m_{es}は（3.38）式を解くことにより算出できます。そしてm_{es}が算出できれば 39）式に代入してηの最尤推定量η_{es}を算出することが可能になります。

同様に、3 母数ワイブル分布の最尤推定量は39)、40)式の連立方程式からm_{es},γ_{es}を算出し、.41）式からη_{es}を求めることができます。

$$\frac{ln(x_U - \gamma_{es}) \cdot m_U \cdot (x_U - \gamma_{es})^{m_{es}} + \sum_{i=1}^{n} \left[(x_i - \gamma_{es})^{m_{es}} \cdot ln(x_i - \gamma_{es}) \right]}{m_U \cdot (x_U - \gamma_{es})^{m_{es}} + \sum_{i=1}^{n} (x_i - \gamma_{es})^{m_{es}}} - \frac{1}{m_{es}} - \frac{1}{n} \cdot \sum_{i=1}^{n} ln(x_i - \gamma_{es}) = 0$$

・・・ 40)

$$\frac{m_U \cdot (x_U - \gamma_{es})^{m_{es}-1} + \sum_{i=1}^{n} (x_i - \gamma_{es})^{m_{es}-1}}{m_U \cdot (x_U - \gamma_{es})^{m_{es}} + \sum_{i=1}^{n} (x_i - \gamma_{es})^{m_{es}}} + \frac{1 - m_{es}}{n \cdot m_{es}} \cdot \sum_{i=1}^{n} \frac{1}{x_i - \gamma_{es}} = 0 \quad \cdots \; 41)$$

$$\eta_{es} = \left[\frac{m_U \cdot (x_U - \gamma_{es})^{m_{es}} + \sum_{i=1}^{n} (x_i - \gamma_{es})^{m_{es}}}{n} \right]^{\frac{1}{m_{es}}} \quad \cdots \; 42)$$

2.3.2　正規分布、対数正規分布

対数尤度関数は 42）式となります。（$\Phi(\ )$：標準正規分布の確率分布関数）

$$LL_{norm} = (n-r)\cdot\left\{1-\Phi\left(\frac{x_0-\mu_{es}}{\sqrt{V_{es}}}\right)\right\}+\sum_{i=1}^{r}\ln\left(\frac{1}{\sqrt{2\cdot\pi\cdot V_{es}}}\exp\left(-\frac{(x_i-\mu_{es})^2}{2\cdot V_{es}}\right)\right) \quad ・・・ 43）$$

μ_{es} で偏微分して 0 と置き整理すると 43）式となります。

$$\mu_{es} = Y\cdot\sigma_{es}+\mu_x \quad ・・・ 44）$$

ただし、 $Y = \frac{m_U}{n}\cdot\frac{\phi\left(\frac{x_U-\mu_{es}}{\sigma_{es}}\right)}{1-\Phi\left(\frac{x_U-\mu_{es}}{\sigma_{es}}\right)}$ 　　$\mu_x = \frac{1}{n}\cdot\sum_{i=1}^{n}(x_i)$

また、対数尤度関数を V_{es} で偏微分して 0 と置き整理すると 44）式が得られます。

$$V_{es} = Y\cdot(x_U-\mu_{es})\cdot\sigma_{es}+(\mu_{es}-\mu_x)^2+V_x \quad ・・・ 45）$$

ただし、$V_x = \frac{1}{n}\cdot\sum_{i=1}^{n}(x_i-\mu_x)^2$

44）式に 43）式を代入して整理すると 45）式となります。

$$V_x = \sigma_{es}\cdot(\sigma_{es}-Y\cdot x_U+Y\cdot\mu_x) \quad ・・・ 46）$$

$u_0 = \frac{x_0-\mu_{es}}{\sigma_{es}}$ より $\sigma_{es} = \frac{x_0-\mu_{es}}{u_0}$ となりますから 45）式より 46）式となります。

$$\frac{V_x}{(x_U-\mu_x)^2} = \frac{1-Y\cdot(Y+u_U)}{(Y+u_U)^2} \quad ・・・ 47）$$

すなわち、46）式の左辺は既知の値ですから収束計算して u_U を算出することができます。そして u_U が求まれば Y が算出できます。

さらに 43）式と $\sigma_{es} = \frac{x_U-\mu_{es}}{u_U}$ から 47）式となることから σ_{es} が算出可能になり、μ_{es} は 43）式に代入して求めることができるようになります。

$$\sigma_{es} = \frac{x_U-\mu_x}{Y+u_U} \quad ・・・ 48）$$

以上のことをまとめると算出手順は以下のようになります。

A) x_U 以下のデータが n 個（$x_1, x_2, \cdots, x_n$）から平均、分散を計算する。

$$\mu_x = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \quad , \quad V_x = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(x_i - \mu_x \right)^2$$

B) u_U を次式から収束計算して求める。

$$\frac{V_x}{\left(x_U - \mu_x \right)^2} = \frac{1 - Y \cdot \left(Y + u_U \right)}{\left(Y + u_U \right)^2}$$

ただし、$Y = \dfrac{m_U}{n} \cdot \dfrac{\phi\left(u_U\right)}{\Phi\left(u_U\right)}$ （$\phi(\ ), \Phi(\ )$：標準正規分布の確率密度関数、確率分布関数）

c） σ_{es} を次式から求める。

$$\sigma_{es} = \frac{x_U - \mu_x}{Y + u_U}$$

d） μ_{es} を次式から求める。

$$\mu_{es} = Y \cdot \sigma_{es} + \mu_x$$

対数正規分布の場合には $\ln\left(x\right)$ が正規分布に従うことから、算出手順は以下のようになります。

A) x_U 以下のデータが n 個（$x_1, x_2, \cdots, x_n$）から平均、分散を計算する。

$$\lambda_x = \frac{1}{n} \cdot \sum_{i=1}^{n} \ln\left(x_i\right) \quad , \quad {\xi_x}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\ln\left(x_i\right) - \lambda_x \right)^2$$

B) u_U を次式から収束計算して求める。 $\dfrac{{\xi_x}^2}{\left(\lambda_x - ln\left(x_U\right) \right)^2} = \dfrac{1 - Y \cdot \left(Y + u_U \right)}{\left(Y + u_U \right)^2}$

ただし、$u_U = \dfrac{\ln(x_U) - \lambda_{es}}{\xi_{es}}$ 、$Y = \dfrac{m_U}{n} \cdot \dfrac{\phi\left(u_U\right)}{\Phi\left(u_U\right)}$ （$\phi(\ ), \Phi(\ )$：標準正規分布の確率密度関数、確率分布関数）

c） ξ_{es} を次式から求める。 $\xi_{es} = \dfrac{ln\left(x_U\right) - \lambda_x}{Y + u_U}$

d） λ_{es} を次式から求める。 $\lambda_{es} = Y \cdot \xi_{es} + \lambda_x$

2.4　切れた分布データの最尤推定量

ある区間のデータのみが得られている場合の最尤推定量の算出方法です。打ち切りデータと異なるのは、区間$\left[x_L, x_U\right]$　$\left(x_L < x_U\right)$以外の情報（個数等）が得られてないということです。

確率密度関数$f_0\left(x,\theta\right)$をもつ分布の区間$\left[x_L, x_U\right]$以外のデータが取り除かれている分布確率密度関数は 48）式で示されます。（θ：パラメータ）

$$f_{ab}\left(x,\theta\right)=\begin{cases}\dfrac{1}{F_0(x_U,\theta)-F_0(x_L,\theta)}\cdot f_0\left(x,\theta\right) & \left(x_L \le x \le x_U\right) \\ 0 & \left(\text{その他}\right)\end{cases} \quad \cdots \quad 49)$$

ここに、$F_0\left(x,\theta\right)$は確率分布関数。

対数尤度関数は打ち切りデータの 33）式の第 3 項目と同じ意味を持つ 49）式となります。

$$LL_x=\ln\left(\prod_{i=1}^{n} f_{ab}\left(x_i,\theta_{es}\right)\right)=\sum_{i=1}^{n}\ln\left(f_{ab}\left(x_i,\theta_{es}\right)\right)=\sum_{i=1}^{n}\ln\left(\frac{f_0\left(x_i,\theta_{es}\right)}{F_0(x_U,\theta_{es})-F_0(x_L,\theta_{es})}\right) \quad \cdots \quad 50)$$

最尤推定量θ_{es}は対数尤度関数を偏微分した方程式の解となることは上述した通りですが、$F_0\left(x,\theta\right)$の偏微分が容易でも方程式が複雑になったり、また、正規分布、対数正規分布のように$F_0\left(x,\theta\right)$の偏微分が開いた形で表現できない場合や方程式を開いた形で表現できないことなどがあります。このため、通常は対数尤度関数の（制約付）最大値問題として数値解析することになります。

3．赤池情報量規準（AIC)による分布形選択

完全データ、打ち切りデータ、切れてデータに対して正規分布、対数正規分布、ワイブル分布などを仮定して最尤法からパラメータを推定する方法について述べました。では、どの分布形がモデルとしてふさわしいかを知りたいことがあります。

その方法としては、分布形を仮定して検定するカイ 2 乗法（χ^2法）や、分布形を仮定せず、観

測されたデータの順位のみの情報から検定する、コロモゴルフ－スミルノフ（Koimogorov－Smirnov）検定（以下　ＫＳ検定）が用いられてきました。

χ^2 検定、ＫＳ検定のどちらも、理論モデルと観測データの差に対応する値が、ある有意水準を仮定した場合の確率より小さい場合に、その帰無仮説を有意として、仮定した分布が妥当であるかどうかを判定するものです。

このため、ある有意水準で受けいれられた仮定分布が、別の有意水準では受け入れられないことが生じることがあるため、χ^2 検定、ＫＳ検定どちらも、仮定分布を絶対評価するものではありませんが、検定に要する計算量が少ないためかよく用いられていました。

最尤法は対数尤度関数を最大化するパラメータを求めることでしたが、その対数尤度関数に推定パラメータを代入した値を最大対数尤度（値）と言い、分布モデルを M として $MLL(M)$ なる記号で表現します。

そして、(推定) パラメータの数を k として、50) 式の赤池情報量規準 (AIC Akaie Information Criterion)を定義します。

$$AIC(M) = -2 \times MLL(M) + 2 \times k \qquad \cdots \quad 51)$$

分布モデルを n 個想定すると、AIC は n 個算出されますが、この中で最も小さな AIC となる分布モデルが真のモデルに最も近いとする判定量です。正確には、モデル M_i とモデル M_j の $AIC(M_i)$、$AIC(M_j)$ の差の絶対値が 1 以上ならば、AIC の小さいモデルが真のモデルに近いということになります。このため、$AIC(M_i)=1000$、$AIC(M_j)=1003$ ならば差の絶対値が 3 のため、モデル M_i が真のモデルに近いと判定されますが、$AIC(M_i)=0.1$、$AIC(M_j)=0.3$ の場合には差の絶対値が 0.2 のため、どちらのモデルが真のモデルに近いかは判定できません。

AIC 法は赤池弘次氏により提案されたモデル選択の方法で、その応用範囲は非常に広く、統計学に対する日本の最大の貢献の一つと言われています。

最尤法による推定量（最尤推定量）を θ_{es} とした場合、最大対数尤度（$MLL(x,\theta_{es})$）は 51)

式で表現されますので、50)式は 52)式となります。

$$MLL\left(x,\theta_{es}\right)=\sum_{i=1}^{n}\ln\left[f\left(x_i,\theta_{es}\right)\right] \qquad \cdot\cdot\cdot\ 52)$$

$$AIC\left(x,\theta_{es},k\right)=-2\times\sum_{i=1}^{n}\ln\left[f\left(x_i,\theta_{es}\right)\right]+2\times k \qquad \cdot\cdot\cdot\ 53)$$

ここに　θ_{es} ：最尤推定量、 $f\left(x,\theta_{es}\right)$ ：確率密度関数、 k ：推定パラメータ数

４．計算例

4.1　初期値

最尤推定量を算出するために方程式を解いたり、最大値問題として解く必要がありますが、初期値を与えないと数値解析ができません。そして、その初期値の与え方としては試行錯誤するほかありませんが、経験上、推定パラメータが 2 個の正規分布、対数正規分布、2P ワイブル分布は、標本平均と標本分散から計算できる値を初期値として与えると比較的収束がうまくいくようです。

・正規分布　$\mu_{es0}=\frac{1}{m}\cdot\sum_{i=1}^{m}x_i=\mu_x \quad \sigma_{es0}{}^2=\frac{1}{m}\cdot\sum_{i=1}^{m}\left(x_i-\mu_x\right)^2$

・対数正規分布　$\lambda_{es0}=\frac{1}{m}\cdot\sum_{i=1}^{m}\ln\left(x_i\right)=\lambda_x \quad \xi_{es0}{}^2=\frac{1}{m}\cdot\sum_{i=1}^{m}\left(\ln\left(x_i\right)-\lambda_x\right)^2$

・2P ワイブル分布

ステップ 1

$$\frac{\sqrt{\Gamma\left(\frac{m_{es0}+2}{m_{es0}}\right)-\Gamma\left(\frac{m_{es0}+1}{m_{es0}}\right)^2}}{\Gamma\left(\frac{m_{es0}+1}{m_{es0}}\right)}=\frac{\sigma_{es0}}{\mu_{es0}}$$

から m_{es0} を求める。

ステップ 2

$$\eta_{es0}=\frac{\mu_{es0}}{\Gamma\left(\frac{m_{es0}+1}{m_{es0}}\right)}$$

から η_{es0} を求める。

・3P ワイブル

3P ワイブル分布の場合には推定したいパラメータが$\eta_{es}, m_{es}, \gamma_{es}$の 3 つですので、標本平均と標本分散の 2 つの値からだけでは 3 つの初期値を与えることができません。

そこで、ワイブル確率紙にプロットしてγ_{es0}を求める手法を応用した方法を以下に示します。

1） m 個のデータ $x_1, x_2, \cdots, x_m$ を小さい順に並べかえ、順序統計量$t_1, t_2, \cdots, t_m$を作る。

2）$Y_i = \ln\left(\ln\left(\dfrac{1}{1-F(t_i)}\right)\right), X_i = \ln\left(t_i - \gamma_{es0}\right)$ と置いた場合の相関係数が最大となるγ_{es0}を求める。

ここで$F\left(t_i\right)$は順序統計量の累積分布関数で平均値ランク、メジアンランクが考えられます。お奨めはメジアンランクでその近似値は$\dfrac{i-0.3}{n+0.4}$です。

3） $\mu_m = \gamma_{es0} + \eta_{es0} \cdot \Gamma\left(\dfrac{m_{es0}+1}{m_{es0}}\right)$ $,\sigma_{es0} = \eta_{es0} \cdot \sqrt{\Gamma\left(\dfrac{m_{es0}+2}{m_{es0}}\right) - \Gamma\left(\dfrac{m_{es0}+1}{m_{es0}}\right)^2}$ の関係から、η_{es0}, m_{es0}を算出する。

この手法は、確率紙を用いて推定する方法を厳密（？）にした方法です。

4.2　算出手順、算出方法

算出手順として、1）パラメータ算出、2）AIC 算出、3）図化を想定します。

パラメータ算出方法としてデータ構造と分布形に対してまとめると、表－１のようになります。

表－ 1　算出方法

項目	正規分布	対数正規分布	2P ワイブル分布
完全データ	厳密解	厳密解	①方程式を解く ②最大化関数で解く
打ち切りデータ	①方程式を解く ②最大化関数で解く	①方程式を解く ②最大化関数で解く	①方程式を解く ②最大化関数で解く
切れたデータ	②最大化関数で解く	②最大化関数で解く	②最大化関数で解く

最も汎用的な方法は、②最大化関数で解く方法となります。

完全データ扱いの場合では、上述したように最尤推定値は

1）正規分布

$$\mu_{es} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \quad , \quad V_{es} = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(x_i - \mu_{es} \right)^2$$

2）対数正規分布

$$\lambda_{es} = \frac{1}{n} \cdot \sum_{i=1}^{n} \ln\left(x_i \right) \quad , \quad \xi_{es}{}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\ln\left(x_i \right) - \lambda_{es} \right)^2$$

3）2P ワイブル分布

①方程式を解く。

$$\frac{1}{m_{es}} + \frac{1}{n} \cdot \sum_{i=1}^{n} ln\left(x_i\right) - \frac{\sum_{i=1}^{n} \left[\left(x_i\right)^{m_{es}} \cdot ln\left(x_i\right) \right]}{\sum_{i=1}^{n} \left(x_i\right)^{m_{es}}} = 0$$ から m_{es} 算出し $$\eta_{es} = \left[\frac{\sum_{i=1}^{n} \left(x_i\right)^{m_{es}}}{n} \right]^{\frac{1}{m_{es}}}$$ に

m_{es} を代入して、η_{es} を求める。

②最大化関数で解く。

完全データ扱いをする場合の関数例（paracalfll(x))を以下に示します。

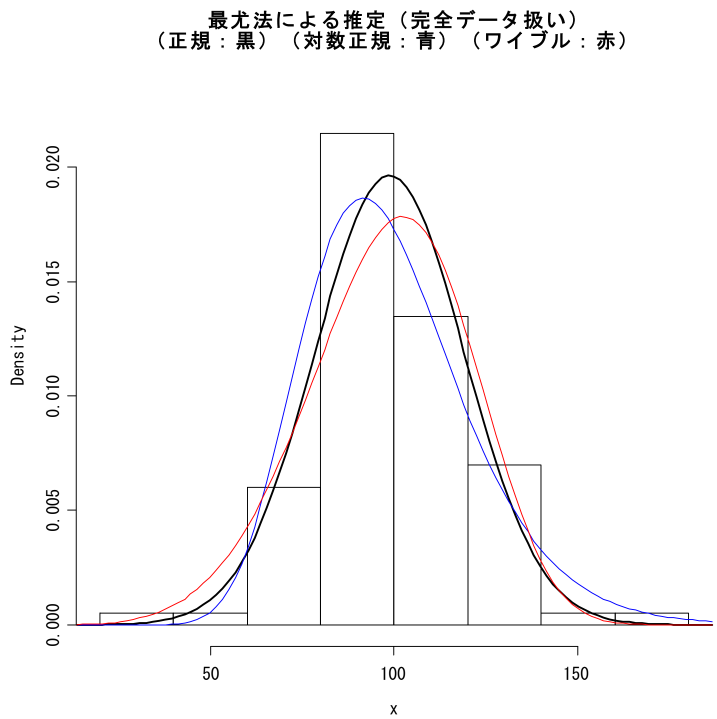
```
#完全データ扱い
paracalfull<-function(x){#最尤法によるパラメータ推定（完全データ扱い）
n<-length(x)#データ大きさ
#対数尤度関数の定義（完全）
LLnorm<-function(para){#正規分布
atai<-sum(log(dnorm(x,para[1],para[2])))
return(atai)
}
LLlog<-function(para){#対数正規分布
atai<-sum(log(dlnorm(x,para[1],para[2])))
return(atai)
}
LL2pweibull<-function(para){#2P ワイブル分布
atai<-sum(log(dweibull(x,para[1],para[2])))
return(atai)
}

fm<-function(mes){#fm 関数(2P ワイブル用)
s1<-gamma(1+1/mes)
```

```
s2<-gamma(1+2/mes)
atai<-((n-1)/n)^0.5*sd(x)/mean(x)-(s2-s1^2)^0.5/s1
return(atai)
}

#正規分布
μes<-mean(x)
σes<-((n-1)/n)^0.5*sd(x)
LLnorm.max<-LLnorm(c(μes,σes))#最大対数尤度
AICnorm<-2*2-2*LLnorm.max

#対数正規分布
lx<-log(x)#データ log 化
λes<-mean(lx)
ξes<-((n-1)/n)^0.5*sd(lx)
LLlog.max<-LLlog(c(λes,ξes))#最大対数尤度
AIClog<-2*2-2*LLlog.max

#2P ワイブル分布
mes0<-1
mes0<-uniroot(fm,c(1,15))$root#m 初期値
ηes0<-mean(x)/gamma(1+1/mes0)#η初期値
para<-c(mes0,ηes0)#パラメータ初期値
para<-optim(para, LL2pweibull, method="SANN",control=list(fnscale=-1))$par#2P ワイブル分布のパラメー
タ推定"シミュレーテッドアニーリング法"
para2p<-optim(para, LL2pweibull, method="BFGS",control=list(fnscale=-1))$par#2P ワイブル分布のパラメ
ータ推定"準ニュートン法"
η2es<-para2p[2]
m2es<-para2p[1]
LL2pweibull.max<-LL2pweibull(para2p)#最大対数尤度
AIC2pweibull<-2*2-2*LL2pweibull.max

#最小 AIC
minAIC<-min(AICnorm,AIClog,AIC2pweibull)
if(minAIC==AICnorm){fname<-"正規分布"}
if(minAIC==AIClog){fname<-"対数正規分布"}
if(minAIC==AIC2pweibull){fname<-"2P ワイブル分布"}


#図化
hist(x,main="最尤法による推定（完全データ扱い)¥n（正規：黒）（対数正規：青）（ワイブル：赤）",
freq=F,ylim=c(0,1.2*dnorm(μes,μes,σes)))
curve(dnorm(x,μes,σes),add=T,col="black",lwd=ifelse(minAIC==AICnorm,2,1))
curve(dlnorm(x,λes,ξes),add=T,col="blue",lwd=ifelse(minAIC==AIClog,2,1))
curve(dweibull(x,m2es,η2es),add=T,col="red",lwd=ifelse(minAIC==AIC2pweibull,2,1))

kekka<-c(
c("μes"=μes,"σes"=σes,"AICnorm"=AICnorm),
c("λes"=λes,"ξes"=ξes,"AIClog"=AIClog),
c("η2es"=η2es,"m2es"=m2es,"AIC2pweibull"=AIC2pweibull),
c("AIC 最小分布:"=fname)
)
return(kekka)
}
```

2P ワイブル分布のパラメータ算出では、頑強性が高い 1）シミュレーテッドアニーリング法と 2）準ニュートン法を用いています。

完全データ扱いをする場合の関数（paracalfll(x))を使って 1）パラメータ算出、2）AIC 算出、3）図化をした場合の例を以下に示します。

```
x<-rnorm(100,100,20)#正規分布乱数発生
> kekka<-paracalfull(x)
> kekka
                μes                 σes              AICnorm                  λes
 "98.7675620852965" "20.2945009100568" "889.857698284915" "4.56920050264825"
                 ξes                AIClog                 η2es                 m2es
"0.227628607360735" "905.619828523218" "106.875469751047" "5.08040209618786"
         AIC2pweibull         AIC 最小分布:
 "896.931757590939"            "正規分布"
```



2.7.2　打ち切りデータ扱い

打ち切りデータ扱いをする場合の関数例（paracalcen())を以下に示します。

```
#打ち切りデータ
paracalcen<-function(x,x0,mu){#最尤法によるパラメータ推定（打ち切りデータ扱い）
n<-length(x)#データ大きさ
lx<-log(x)#データ log 化
#正規分布
LLnorm<-function(para){#対数尤度関数の定義（正規分布、打ち切り）
atai<-mu*log(1-pnorm(x0,para[1],para[2]))+sum(log(dnorm(x,para[1],para[2])))
return(atai)
}
para<-c(mean(x),((n-1)/n)^0.5*sd(x))#初期値
paranorm<-optim(para, LLnorm, control=list(fnscale=-1))$par
μes<-paranorm[1]
σes<-paranorm[2]
LLnormti<-LLnorm(c(μes,σes))
#対数正規分布
```

```
lx<-log(x)
para<-c(mean(lx),((n-1)/n)^0.5*sd(lx))#初期値
LLlog<-function(para){#対数尤度関数の定義（正規分布、打ち切り）
atai<-mu*log(1-plnorm(x0,para[1],para[2]))+sum(log(dlnorm(x,para[1],para[2])))
return(atai)
}
paralog<-optim(para, LLlog, control=list(fnscale=-1))$par
λes<-paralog[1]
ξes<-paralog[2]
LLlogti<-LLlog(c(λes,ξes))
#ワイブル分布
#2P ワイブル
LL2p<-function(para){#対数尤度関数の定義(2P ワイブル、打ち切り）
atai<-mu*log(1-pweibull(x0,para[1],para[2]))+sum(log(dweibull(x,para[1],para[2])))
return(atai)
}
#3P ワイブル
LL3p<-function(para){#対数尤度関数の定義（2P ワイブル、打ち切り）
atai<-mu*log(1-pweibull(x0-para[3],para[1],para[2]))+sum(log(dweibull(x-para[3],para[1],para[2])))
return(atai)
}
fm<-function(mes){#m 関数定義
n<-length(x)
s1<-gamma(1+1/mes)
s2<-gamma(1+2/mes)
atai<-((n-1)/n)^0.5*sd(x)/mean(x)-(s2-s1^2)^0.5/s1
return(atai)
}
cl<-function(γ){#順序統計量の相関係数
t<-sort(x)
n<-length(x)
for (i in 1:n) {F[i]<-(i-0.3)/(n+0.4)}
Y<-log(log(1/(1-F)))
X<-log(t-γ)
atai<-cor(X, Y)
return(atai)
}
mes0<-uniroot(fm,c(1,15))$root#m 初期値
ηes0<-mean(x)/gamma(1+1/mes0)#η初期値
γ<-min(x)/10#γ初期値
γes0<- nlm(cl,  γ,TRUE, fscale=-1)$estimate#γ初期値
#2P ワイブル
para<-c(mes0,ηes0)#パラメータ初期値
para2p<-optim(para, LL2p, control=list(fnscale=-1))$par#2P ワイブル分布のパラメータ推定
η2es<-para2p[2]
m2es<-para2p[1]
LL2pti<-LL2p(para2p)
#3P ワイブル
para<-c(mes0,ηes0,γes0)#パラメータ初期値
para3p<-optim(para, LL3p, control=list(fnscale=-1))$par#3P ワイブル分布のパラメータ推定
η3es<-para3p[2]
m3es<-para3p[1]
γes<-para3p[3]
LL3pti<-LL3p(para3p)
c("μes"=μes,"σes"=σes,"LLnorm"=LLnormti,
"λes"=λes,"ξes"=ξes,"LLnog"=LLlogti,
"η2es"=η2es,"m2es"=m2es,"LL2p"=LL2pti,
"η3es"=η3es,"m3es"=m3es,"γes"=γes,"LL3p"=LL3pti)
}
```

この関数を使って算出すると以下のようになります。

```
> x<-c(200,185,210,170,220,240,215,230)
> x0<-250
> mu<-2
> paracalcen(x,x0,mu)
        μes         σes      LLnorm         λes         ξes      LLnog        η2es        m2es
LL2p        η3es        m3es         γes        LL3p
 220.2632780  30.6958122 -40.8780867   5.3885886   0.1456407 -40.7157398 232.3464526   8.1863595
-41.2412687  65.6206567   1.6662752 163.7120910 -40.4867599
```

2.7.3 切断データ扱い

切断データ扱いをする場合の関数例（paracaltrun ()）を以下に示します。

```
#切断データ
paracaltrun<-function(x,xu,xl){#切断データ
n<-length(x)#データ大きさ
lx<-log(x)#データ log 化
#正規分布
LLnorm<-function(para){#対数尤度関数の定義（正規分布、切断）
atai<-sum(log(dnorm(x,para[1],para[2])/(pnorm(xu,para[1],para[2])-(pnorm(xl,para[1],para[2])))))
return(atai)
}
para<-c(mean(x),((n-1)/n)^0.5*sd(x))#初期値
paranorm<-optim(para, LLnorm, control=list(fnscale=-1))$par
μes<-paranorm[1]
σes<-paranorm[2]
LLnormti<-LLnorm(c(μes,σes))

#対数正規分布
LLlog<-function(para){#対数尤度関数の定義（対数正規分布、切断）
atai<-sum(log(dlnorm(x,para[1],para[2])/(plnorm(xu,para[1],para[2])-(plnorm(xl,para[1],para[2])))))
return(atai)
}
lx<-log(x)#データ log 化
para<-c(mean(lx),((n-1)/n)^0.5*sd(lx))#初期値
paralog<-optim(para, LLlog, control=list(fnscale=-1))$par
λes<-paralog[1]
ξes<-paralog[2]
LLlogti<-LLlog(c(λes,ξes))


#ワイブル分布
#2P ワイブル
LL2p<-function(para){#対数尤度関数の定義(2P ワイブル、打ち切り）
atai<-sum(log(dweibull(x,para[1],para[2])/(pweibull(xu,para[1],para[2])-(pweibull(xl,para[1],para

[2])))))
return(atai)
}

#3P ワイブル
LL3p<-function(para){#対数尤度関数の定義（3P ワイブル、打ち切り）
atai<-sum(log(dweibull(x-para[3],para[1],para[2])/(pweibull(xu-para[3],para[1],para[2])-(pweibull(xl-

para[3],para[1],para[2])))))
return(atai)
```

```
}

fm<-function(mes){#m 関数定義
n<-length(x)
s1<-gamma(1+1/mes)
s2<-gamma(1+2/mes)
atai<-((n-1)/n)^0.5*sd(x)/mean(x)-(s2-s1^2)^0.5/s1
return(atai)
}

cl<-function(γ){#順序統計量の相関係数
t<-sort(x)
n<-length(x)
for (i in 1:n) {F[i]<-(i-0.3)/(n+0.4)}
Y<-log(log(1/(1-F)))
X<-log(t-γ)
atai<-cor(X, Y)
return(atai)
}

mes0<-uniroot(fm,c(1,15))$root#m 初期値
ηes0<-mean(x)/gamma(1+1/mes0)#η初期値
γ<-min(x)/10#γ初期値
γes0<- nlm(cl, γ,TRUE, fscale=-1)$estimate#γ初期値

#2P ワイブル
para<-c(mes0,ηes0)#パラメータ初期値
para2p<-optim(para, LL2p, control=list(fnscale=-1))$par#2P ワイブル分布のパラメータ推定
η2es<-para2p[2]
m2es<-para2p[1]
LL2pti<-LL2p(para2p)

#3P ワイブル
para<-c(mes0,ηes0,γes0)#パラメータ初期値
para3p<-optim(para, LL3p, control=list(fnscale=-1))$par#3P ワイブル分布のパラメータ推定
η3es<-para3p[2]
m3es<-para3p[1]
γes<-para3p[3]
LL3pti<-LL3p(para3p)

c("μes"=μes,"σes"=σes,"LLnorm"=LLnormti,
"λes"=λes,"ξes"=ξes,"LLnog"=LLlogti,
"η2es"=η2es,"m2es"=m2es,"LL2p"=LL2pti,
"η3es"=η3es,"m3es"=m3es,"γes"=γes,"LL3p"=LL3pti)
}
```

この関数を使って算出すると以下のようになります。

```
> x<-c(200,185,210,170,220,240,215,230)
> xu<-250
> xl<-100
> paracaltrun(x,xu,xl)
          μes         σes      LLnorm         λes         ξes       LLnog        η2es        m2es
LL2p        η3es        m3es         γes        LL3p
  211.7082013  24.2746126 -35.6252544   5.3607076   0.1266133 -35.6329825 219.0570716  11.1628648
-35.6337261  81.6578367   0.8839113 170.0000000 -31.7870188
```