各種分布のパラメータの最尤推定（最尤推定大演習）

得られたデータ $x_\lambda\ (\lambda=1,\cdots,N)$ が、特定の分布に従うかどうかを調べる際、分布のパラメータが既知であるとは限らない。そのため、多くの場合、与えられたデータを用いて各種分布のパラメータを推定し、その下で検定の問題を考えることになると思われる。そこで、メニュー［分析－基本統計－分布と検定］で表示される図 1 の分析実行メニューに、パラメータを自動的に推定する機能を加えた。分布を選んで「推定」ボタンをクリックすると左のテキストボックスに推定値が表示される。

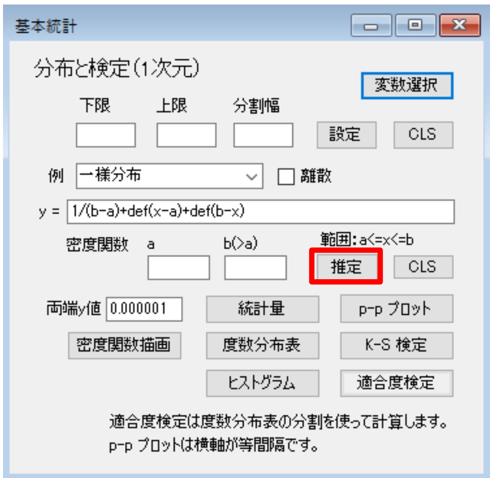

図 1　分布と検定実行メニュー

ここでは分布毎にパラメータを推定するための方法を具体的に与えておく。

正規分布　$(-\infty < x < \infty)$

密度関数：$f(x\,|\,\mu,\sigma^2)=\dfrac{1}{\sqrt{2\pi\sigma^2}}\exp[-(x_\lambda-\mu)^2/2\sigma^2]$

尤度関数：$L=\dfrac{1}{(2\pi\sigma^2)^{N/2}}\exp\left[-\dfrac{1}{2\sigma^2}\displaystyle\sum_{\lambda=1}^{N}(x_\lambda-\mu)^2\right]$

対数尤度：$\log L=-\dfrac{1}{2\sigma^2}\displaystyle\sum_{\lambda=1}^{N}(x_\lambda-\mu)^2-\dfrac{N}{2}\log(2\pi\sigma^2)$

スコアベクトル $\mathbf{U}$ と情報行列 $\mathfrak{I}$

$$\boldsymbol{\beta}=\begin{pmatrix}\mu\\ \sigma^2\end{pmatrix},\qquad \mathbf{U}=\begin{pmatrix}\partial\log L/\partial\mu\\ \partial\log L/\partial\sigma^2\end{pmatrix},\qquad \mathfrak{I}=-\begin{pmatrix}\partial^2\log L/\partial\mu^2 & \partial^2\log L/\partial\mu\partial\sigma^2\\ \partial^2\log L/\partial\mu\partial\sigma^2 & \partial^2\log L/\partial(\sigma^2)^2\end{pmatrix}$$

$$\partial\log L/\partial\mu=\frac{1}{\sigma^2}\sum_{\lambda=1}^{N}(x_\lambda-\mu)=0 \qquad\qquad \mu=\frac{1}{N}\sum_{\lambda=1}^{N}x_\lambda$$

$$\partial\log L/\partial\sigma^2=\frac{1}{2\sigma^4}\sum_{\lambda=1}^{N}(x_\lambda-\mu)^2-\frac{N}{2\sigma^2}=0 \qquad\qquad \sigma^2=\frac{1}{N}\sum_{\lambda=1}^{N}(x_\lambda-\mu)^2$$

以上で解析的に求めることが可能であるが、プログラムでは練習問題としてニュートン・ラフソン法を用いて計算を試している。

$$\partial^2 \log L / \partial \mu^2 = -\frac{N}{\sigma^2}$$

$$\partial^2 \log L / \partial \mu \partial \sigma^2 = -\frac{1}{\sigma^4} \sum_{\lambda=1}^{N} (x_\lambda - \mu) \to 0$$

$$\partial^2 \log L / \partial (\sigma^2)^2 = -\frac{1}{\sigma^6} \sum_{\lambda=1}^{N} (x_\lambda - \mu)^2 + \frac{N}{2\sigma^4} \to -\frac{N}{2\sigma^4}$$

初期値は $\mu_0 = 0,\ \sigma_0^2 = 1$ を用いている。

注） $\partial^2 \log L / \partial \sigma^2 = -\dfrac{2N}{\sigma^2}$ $\qquad \partial^2 \log L / \partial \sigma^2 = 4\sigma^2\, \partial^2 \log L / \partial (\sigma^2)^2$

パラメータの推定にはどちらの値を用いるべきだろうか。

χ^2 分布 $(0 < x < \infty)$ パラメータが離散的

密度関数： $f(x) = \dfrac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} \exp(-x/2)$

尤度関数： $L = \dfrac{1}{2^{Nn/2}\Gamma(n/2)^N} \displaystyle\prod_{\lambda=1}^{N} x_\lambda^{n/2-1} \exp(-x_\lambda/2)$

対数尤度： $\log L = \dfrac{n-2}{2} \displaystyle\sum_{\lambda=1}^{N} \log(x_\lambda) - \dfrac{1}{2} \sum_{\lambda=1}^{N} x_\lambda - \dfrac{Nn}{2} \log 2 - N \log \Gamma(n/2)$

$\chi^2 \sim \chi_n^2$ のとき、$E(\chi^2) = n$ の性質を用いて、

$n = \lfloor x + 0.5 \rfloor$ 　　注） $\lfloor x \rfloor$ は x を越えない最大の整数

これを元に $(1\le)\ n-5 \le n_{\max} \le n+5$ の範囲で最大の対数尤度を与える自由度 $n_{\max}$ を求めている。

F分布 $(0 < x < \infty)$

密度関数： $f(x) = \dfrac{1}{B(n_1/2, n_2/2)} \left(\dfrac{n_1}{n_2}\right)^{n_1/2} \dfrac{x^{n_1/2-1}}{(1 + x n_1/n_2)^{(n_1+n_2)/2}}$

尤度関数： $L = \dfrac{1}{B(n_1/2, n_2/2)^N} \left(\dfrac{n_1}{n_2}\right)^{Nn_1/2} \displaystyle\prod_{\lambda=1}^{N} \dfrac{x_\lambda^{n_1/2-1}}{(1 + x_\lambda n_1/n_2)^{(n_1+n_2)/2}}$

対数尤度：

$$\log L = \left(\frac{n_1}{2} - 1\right) \sum_{\lambda=1}^{N} \log(x_\lambda) - \frac{n_1+n_2}{2} \sum_{\lambda=1}^{N} \log(1 + x_\lambda n_1/n_2) + \frac{Nn_1}{2} \log(n_1/n_2) - N \log B(n_1/2, n_2/2)$$

$E[X] = \dfrac{n_2}{n_2 - 2}$ $(n_2 > 2)$, $V[X] = \dfrac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2-2)^2(n_2-4)}$ $(n_2 > 4)$ 　を利用して、

$n_2 = \dfrac{2E[X]}{E[\mathrm{X}] - 1}$, $n_1 = \dfrac{2n_2^2(n_2-2)}{(n_2-2)^2(n_2-4)V[X] - 2n_2^2}$

これを元に、ぶれが大きいので、$(1\le) n_i - 20 \le n_{i\max} \le n_i + 20$ の範囲で対数尤度を最大化す

る $n_{i\max}$ を求めている。

t 分布 $(-\infty < x < \infty)$

密度関数： $f(x) = \dfrac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\left(1+\dfrac{x^2}{n}\right)^{-(n+1)/2}$

尤度関数： $L = \left(\dfrac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\right)^N \prod_{\lambda=1}^{N}\left(1+\dfrac{x_\lambda^2}{n}\right)^{-(n+1)/2}$

対数尤度： $\log L = -\dfrac{n+1}{2}\sum_{\lambda=1}^{N}\log\left(1+\dfrac{x_\lambda^2}{n}\right) + N\log\left(\dfrac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\right)$

平均： $E[X] = 0$

分散：$V[X] = \dfrac{n}{n-2}$ を利用して、 $n = \dfrac{2V[X]}{V[X]-1}$

これを元に $(1\le)\ n-5 \le n_{\max} \le n+5$ の範囲で最大の対数尤度を与える自由度 $n_{\max}$ を求めている。

ガンマ分布 $(0 < x < \infty)$

密度関数： $f(x) = \dfrac{1}{b^a\Gamma(a)} x^{a-1}\exp(-x/b)$

尤度関数： $L = \dfrac{1}{[b^a\Gamma(a)]^N}\prod_{\lambda=1}^{N} x_\lambda^{a-1}\exp(-x_\lambda/b)$

対数尤度： $\log L = (a-1)\sum_{\lambda=1}^{N}\log x_\lambda - \dfrac{1}{b}\sum_{\lambda=1}^{N} x_\lambda - Na\log b - N\log\Gamma(a)$

$$\partial\log L/\partial a = \sum_{\lambda=1}^{N}\log x_\lambda - N\log b - N\,\Gamma'(a)/\Gamma(a)$$

$$\partial\log L/\partial b = 1/b^2\sum_{\lambda=1}^{N} x_\lambda - N\,a/b$$

$$\partial^2\log L/\partial a^2 = -N\left[\Gamma''(a)/\Gamma(a) - \Gamma'(a)^2/\Gamma(a)^2\right]$$

$$\partial^2\log L/\partial a\partial b = -N/b$$

$$\partial^2\log L/\partial b^2 = -2/b^3\sum_{\lambda=1}^{N} x_\lambda + N\,a/b^2$$

初期値は $a_0 = 0.5,\ b_0 = 0.5$ を用いている。

逆ガンマ分布 $(0 \le x \le 1)$

密度関数： $f(x) = \dfrac{b^a}{\Gamma(a)} x^{-a-1}\exp(-b/x)$

尤度関数： $L=\left[b^a/\Gamma(a)\right]^N \prod_{\lambda=1}^{N} x_\lambda^{-a-1}\exp(-b/x_\lambda)$

対数尤度： $\log L=-(a+1)\sum_{\lambda=1}^{N}\log x_\lambda - b\sum_{\lambda=1}^{N}(1/x_\lambda)+Na\log b-N\log\Gamma(a)$

$\partial\log L/\partial a=-\sum_{\lambda=1}^{N}\log x_\lambda+N\log b-N\,\Gamma'(a)/\Gamma(a)$

$\partial\log L/\partial b=-\sum_{\lambda=1}^{N}(1/x_\lambda)+N\,a/b$

$\partial^2\log L/\partial a^2=-N\left[\Gamma''(a)/\Gamma(a)-\Gamma'(a)^2/\Gamma(a)^2\right]$

$\partial^2\log L/\partial a\partial b=N/b$

$\partial^2\log L/\partial b^2=-N\,a/b^2$

初期値は $a_0=0.5,\ b_0=0.5$ を用いている。

ベータ分布 $(0\le x\le 1)$

密度関数： $f(x)=\dfrac{x^{a-1}(1-x)^{b-1}}{B(a,b)}=\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}$

尤度関数： $L=\dfrac{1}{B(a,b)}\prod_{\lambda=1}^{N}x_\lambda^{a-1}(1-x_\lambda)^{b-1}$

対数尤度： $\log L=(a-1)\sum_{\lambda=1}^{N}\log x_\lambda+(b-1)\sum_{\lambda=1}^{N}\log(1-x_\lambda)-N\log B(a,b)$

$\partial\log L/\partial a=\sum_{\lambda=1}^{N}\log x_\lambda-N\dfrac{B(a,b)_a}{B(a,b)}$

$\partial\log L/\partial b=\sum_{\lambda=1}^{N}\log(1-x_\lambda)-N\dfrac{B(a,b)_b}{B(a,b)}$

$\partial^2\log L/\partial a^2=-N\left(\dfrac{B(a,b)_{aa}}{B(a,b)}-\dfrac{B(a,b)_a^2}{B}\right)$

$\partial^2\log L/\partial a\partial b=-N\left(\dfrac{B(a,b)_{ab}}{B(a,b)}-\dfrac{B(a,b)_a B(a,b)_b}{B(a,b)^2}\right)$

$\partial^2\log L/\partial b^2=-N\left(\dfrac{B(a,b)_{bb}}{B(a,b)}-\dfrac{B(a,b)_b^2}{B}\right)$

初期値の設定で、平均値が 0 に近い場合は 1,5、1 に近い場合は 5,1、0.5 に近い場合は 0.5, 0.5 などを使う。小さい方から大きい方へ近づけて行くことは問題ないが、大きい方から小さい方へ近づけて行く際にはエラーが出る。

ワイブル分布 $(0<x<\infty)$ **（失敗例）**

通常の a,b を使って最尤法を試みた。

密度関数： $f(x)=(a/b)\left(x/b\right)^{a-1}\exp\left[-\left(x/b\right)^a\right]$

尤度関数： $L=(a/b)^N\prod_{\lambda=1}^{N}\left(x_\lambda/b\right)^{a-1}\exp\left[-\left(x_\lambda/b\right)^a\right]=a^Nb^{-Na}\prod_{\lambda=1}^{N}x_\lambda^{a-1}\exp[-x_\lambda^a b^{-a}]$

対数尤度： $\log L=(a-1)\sum_{\lambda=1}^{N}\log x_\lambda-b^{-a}\sum_{\lambda=1}^{N}x_\lambda^a+N\log a-Na\log b$

$$\partial\log L/\partial a=\sum_{\lambda=1}^{N}\log x_\lambda+\log b\cdot b^{-a}\sum_{\lambda=1}^{N}x_\lambda^a-b^{-a}\sum_{\lambda=1}^{N}\log x_\lambda\cdot x_\lambda^a+N/a-N\log b$$

$$\partial\log L/\partial b=ab^{-a-1}\sum_{\lambda=1}^{N}x_\lambda^a-Na/b$$

$$\partial^2\log L\big/\partial a^2=-(\log b)^2b^{-a}\sum_{\lambda=1}^{N}x_\lambda^a+2\log b\cdot b^{-a}\sum_{\lambda=1}^{N}\log x_\lambda\cdot x_\lambda^a$$
$$-b^{-a}\sum_{\lambda=1}^{N}(\log x_\lambda)^2\cdot x_\lambda^a-N\big/a^2$$

$$\partial^2\log L\big/\partial a\partial b=(1-a\log b)b^{-a-1}\sum_{\lambda=1}^{N}x_\lambda^a+ab^{-a-1}\sum_{\lambda=1}^{N}\log x_\lambda\cdot x_\lambda^a-N/b$$

$$\partial^2\log L\big/\partial b^2=-a(a+1)b^{-a-2}\sum_{\lambda=1}^{N}x_\lambda^a+Na\big/b^2$$

この方法は、収束が思うように行かず、エラーとなった。

ワイブル分布 $(0<x<\infty)$　**再度**

上記の失敗を踏まえ、生存時間分析で用いたパラメータの推定法を利用する。

密度関数： $f(x)=(a/b)\left(x/b\right)^{a-1}\exp\left[-\left(x/b\right)^a\right]$

尤度関数：
$$L=(a/b)^N\prod_{\lambda=1}^{N}\left(x_\lambda/b\right)^{a-1}\exp\left[-\left(x_\lambda/b\right)^a\right]=a^Nb^{-Na}\prod_{\lambda=1}^{N}x_\lambda^{a-1}\exp[-x_\lambda^a b^{-a}]$$
$$=a^Ne^{N\beta}\prod_{\lambda=1}^{N}x_\lambda^{a-1}\exp[-x_\lambda^a e^\beta]$$

$$\log L(a,b)=\sum_{\lambda=1}^{N}\left[\left(\log a+(a-1)\log x_\lambda+\beta\right)-x_\lambda^a e^\beta\right]$$
$$=(a-1)\sum_{\lambda=1}^{N}\log x_\lambda-e^\beta\sum_{\lambda=1}^{N}x_\lambda^a+N\log a+N\beta$$

$$\frac{\partial}{\partial a}\log L=\sum_{\lambda=1}^{N}\log x_\lambda-e^\beta\sum_{\lambda=1}^{N}\log x_\lambda\cdot x_\lambda^a+N/a$$

$$\frac{\partial}{\partial\beta}\log L=-e^\beta\sum_{\lambda=1}^{N}x_\lambda^a+N$$

$$\frac{\partial^2}{\partial a^2}\log L=-e^\beta\sum_{\lambda=1}^{N}(\log x_\lambda)^2x_\lambda^a-N\big/a^2$$

$$\frac{\partial}{\partial a \partial \beta} \log L = -e^{\beta} \sum_{\lambda=1}^{N} \log x_{\lambda}\, x_{\lambda}^{a}$$

$$\frac{\partial^2}{\partial \beta^2} \log L = -e^{\beta} \sum_{\lambda=1}^{N} x_{\lambda}^{a}$$

初期値は $a_0 = 2,\ \beta = 2$ を用いている。

指数分布 $(0 < x < \infty)$

密度関数： $f(t) = a \exp(-ax)$　　（$x \geq 0$）

尤度関数： $L = a^N \prod_{i=1}^{N} \exp(-ax_{\lambda}) = a^N \exp\left(-a \sum_{\lambda=1}^{N} x_{\lambda} \right)$

対数尤度： $\log L = N \log a - a \sum_{i=1}^{N} x_{\lambda}$

$$\frac{\partial}{\partial a} \log L = \frac{N}{a} - \sum_{\lambda=1}^{N} x_{\lambda} = 0 \qquad\qquad a = N \Big/ \sum_{\lambda=1}^{N} x_{\lambda}$$

$$\frac{\partial^2}{\partial a^2} \log L = -\frac{N}{a^2}$$

この逆数は、推定値の分散を与える。（今回は使わない）

推定値は解析的に求まるが、練習問題として最尤法を用いてみる。

初期値は $a = 0.1$ を用いている。

ポアソン分布 $(0 < x < \infty)$，**整数**

確率関数： $P(x) = e^{-a}\, a^x \big/ x!$

尤度関数： $L = \prod_{\lambda=1}^{N} e^{-a}\, a^{x_{\lambda}} \big/ x_{\lambda}! = e^{-Na} \prod_{\lambda=1}^{N} a^{x_{\lambda}} \big/ x_{\lambda}!$

対数尤度： $\log L = -Na + \log a \sum_{\lambda=1}^{N} x_{\lambda} - \sum_{\lambda=1}^{N} \log x_{\lambda}!$

$$\partial \log L \big/ \partial a = -N + \frac{1}{a} \sum_{\lambda=1}^{N} x_{\lambda}$$

$$\partial^2 \log L \big/ \partial a^2 = -\frac{1}{a^2} \sum_{\lambda=1}^{N} x_{\lambda}$$

初期値は $a = 0.1$ を用いている。

2 項分布 $(0 < x < \infty)$，**整数**

まず以下の関係を使って、度数 n を求める。

$E[X] = np$，$V[X] = npq$

$$n = \frac{E[X]^2}{E[X] - V[X]}$$

次に最尤法を使って、確率 p を求める。

確率関数： $P(x) = {}_nC_x p^x (1-p)^{n-x}$

尤度関数： $L = \prod_{\lambda=1}^{N} {}_nC_{x_\lambda} p^{x_\lambda} (1-p)^{n-x_\lambda}$

対数尤度： $\log L = \sum_{\lambda=1}^{N} \log {}_nC_{x_\lambda} + \log p \sum_{\lambda=1}^{N} x_\lambda + \log(1-p) \sum_{\lambda=1}^{N} (n - x_\lambda)$

$$\partial \log L / \partial p = \frac{1}{p} \sum_{\lambda=1}^{N} x_\lambda - \frac{1}{1-p} \sum_{\lambda=1}^{N} (n - x_\lambda)$$

$$\partial^2 \log L / \partial p^2 = -\frac{1}{p^2} \sum_{\lambda=1}^{N} x_\lambda - \frac{1}{(1-p)^2} \sum_{\lambda=1}^{N} (n - x_\lambda)$$

これも解析的に解を求めることができるが、最尤法の演習とする。

初期値は $p = 0.5$ を用いている。