7Ｋ－6

ラプラス分布のパラメトリック適合度検定法

原田　治行
(九州工業大学工学部)

１．はじめに

通常、ラプラス分布の適合度は、ノンパラメトリックな検定法である　χ^2検定やKolmogorov-Smirnov(KS)検定[1]で行われている。本論文では、ラプラス分布の母数を推定する式が、モーメント法と最尤法では異なり、漸近特性では、値が一致する性質を利用して、上記の２つの方法で求めた母数を変数とする統計量Ｕを定義し、データ数が多いければ、これに基づいてパラメトリックなラプラス分布の検定が出来ることを統計的仮説検定により示した。

２．ラプラス分布の母数と漸近特性

分布の母数を推定する方法には、モーメント法と最尤法[1]がある。代表的な確率分布である正規分布、二項分布、ポアッソン分布、幾何分布などは両者の推定式が一致することが知られている。しかし、ラプラス分布の場合は両者の推定式は一致しない。ラプラス分布の母数は、平均値μと標準偏差σである。ラプラス分布を(1)式に示す。

$$f(x)=\frac{1}{\sqrt{2}\,\sigma}e^{-\frac{\sqrt{2}}{\sigma}|x-\mu|} \tag{1}$$

2.1　モーメント法による母数推定

これは、標本モーメントと母集団分布の対応するモーメントを等しいと置き、その方程式を母数に付いて解く方法である。結果を(2),(3)式に示す。

$$\mu_1=\frac{\sum_{i=1}^{N}X_i}{N} \tag{2}$$

$$\sigma_1=\sqrt{\frac{\sum_{i=1}^{N}(X_i-\mu_1)^2}{N}} \tag{3}$$

ただし、X_i(i=1,・・・,N)は、標本データである。それぞれ、標本平均、標本標準偏差に等しい。

2.2　最尤法による母数の推定

これは、対数尤度関数を求め、それを推定する母数で偏微分し、ゼロと置いた連立方程式を解く方法である。結果を(4),(5)式に示す。

$$\mu_2 = \text{メジアン} \tag{4}$$

$$\sigma_2=\frac{\sqrt{2}\sum_{i=1}^{N}|X_i-\mu_2|}{N} \tag{5}$$

2.3 母数の漸近特性

N→∞の時に$\mu_1,\mu_2,\sigma_1,\sigma_2$の漸近特性を求める。ラプラス分布は左右対称なので平均値μ_1と、メジアンμ_2は等しと置けるので$\mu_1=\mu_2=\mu$と置く。また、(3),(5)式はそれぞれ、(6),(7)式と置ける。

$$\sigma_1=\sqrt{\int_{-\infty}^{\infty}(x-\mu)^2 f(x)\,dx} \tag{6}$$

$$\sigma_2=\sqrt{2}\int_{-\infty}^{\infty}|x-\mu|f(x)\,dx \tag{7}$$

(6),(7)式を各々積分すると$\sigma_1=\sigma_2=\sigma$となり互いに等しくなる。

３．統計量Ｌ

3.1 統計量Ｌの定義

統計量Ｌを(8)式で定義する。

$$L \triangleq \frac{1}{2}\left(\frac{\sigma_2}{\sigma_1}\right)^2 \tag{8}$$

N→∞の漸近特性では、Ｌは 1/2 になる。ここで、$\mu_1=\mu_2=\mu$が近似的に成り立つくらいにＮが十分に大きいとすると（条件Ａ）、Ｌは(9)式で近似出来る。

$$L \cong \frac{\left\{\sum_{i=1}^{N}|X_i-\mu|\right\}^2}{N\sum_{i=1}^{N}(X_i-\mu)^2} \tag{9}$$

3.2 統計量Ｌの分布

統計量Ｌの範囲を求める。(10)式に示すコーシの不等式に於て、$b_i=1$, $a_i=|x_i-\mu|$と置くとＬの範囲は、(11)式で表せる。

$$\sum a_i^2\sum b_i^2 \geq \left(\sum a_i b_i\right)^2 \tag{10}$$

$$1 > L > 0 \tag{11}$$

次に、Ｅ［Ｌ］とＥ［L^2］を求める。Ｅ［$|x-\mu|^n$］は(12)式で表せる。

$$E[|x-\mu|^n]=\left(\frac{\sigma}{\sqrt{2}}\right)^n n! \tag{12}$$

(9)式の分母・分子は互いに独立であることと、(12)式と多項定理より、(13),(14)式が求まる。

$$E[L]=\frac{N+1}{2N}=\frac{1+\frac{1}{N}}{2} \tag{13}$$

Parametric Test for Goodness of Fit of Laplace's Distribution
Haruyuki HARADA
Kyushu Institute of Technology

$$E[L^2]=\frac{(N+1)(N+2)(N+3)}{4N^2(N+5)}=\frac{1+\frac{6}{N}+\frac{11}{N^2}+\frac{6}{N^3}}{4+\frac{20}{N}} \quad (14)$$

条件Aを満たすために、E［L］に於てO（N^{-2}）以下を無視すると、E［L^2］に於いては、O（N^{-4}）以下を無視することになる。ところが、V［L］を求めると(15)式となり矛盾する。

$$V[L]=E[L^2]-\{E[L]\}^2=\frac{-N^2+1}{4N^2(N+5)}<0 \quad (15)$$

そこで、E［L］に於て、O（N^{-1}）以下を無視すると、E［L］,E［L^2］,V［L］は(16)式となる。

$$E[L]=\frac{1}{2},\ E[L^2]=\frac{N+6}{4(N+5)},\ V[L]=\frac{1}{4(N+5)} \quad (16)$$

$N\to\infty$の漸近特性では、V［L］は、ゼロとなる。以上より、統計量LはNが大きければ、0と1の間の数であり平均値は1/2、$N\to\infty$で分散はゼロに近づくので、分布が一山型で滑らかに変化する密度を持つと言える。よって、(17)式のベータ分布で近似することが出来る。

$$f_\beta(L)=\frac{1}{B(p,q)}L^{p-1}(1-L)^{q-1},\ (0<L<1) \quad (17)$$

(p,q)を母数とするベータ分布の最初の２つのモーメントは、p/(p+q),および、p(p+1)/(p+q)/(p+q+1)なので、それぞれ(16)式のE［L］,E［L^2］と等しいと置くと(18)式となる。

$$p=q=\frac{N}{2}+2\doteqdot\frac{N}{2} \quad (18)$$

4．統計量U

4.1 ベータ分布からF分布への変換

統計量Lを(19)式で表す変数変換を行えば、統計量Tは自由度(2p,2q)のF分布になることが知られている。[2]

$$T=\frac{qL}{p(1-L)}=\frac{L}{1-L} \quad (19)$$

4.2 F分布から単位正規分布への変換

(20)式で表されるEdward Paulsonによる正規変換[2]によれば、自由度(2p,2q)を持つF分布は自由度が３以上の時に近似的に単位正規分布になる。

$$U=\frac{(1-W_2)S-(1-W_1)}{\sqrt{W_2S^2+W_1}} \quad (20)$$

ただし、$S=T^{1/3}$, $W_1=1/9p$, $W_2=1/9q$

(20)式に(8)式を代入して整理すれば、(21)式となる。

$$U=\sqrt{\frac{9}{2}N}\ \frac{\left(\frac{\sigma_2^2}{2\sigma_1^2-\sigma_2^2}\right)^{1/3}-1}{\sqrt{1+\left(\frac{\sigma_2^2}{2\sigma_1^2-\sigma_2^2}\right)^{2/3}}} \quad (21)$$

5．シミュレーション

Nが大きければ、統計量Uが単位正規分布に従うことを統計的仮説検定に基づきテストする。「変量Xがラプラス分布に従う時、データ数Nが大きくなれば変量Xの統計量Uの分布は単位正規分布に従う。」と言う仮説の基にχ^2検定とKS検定を行う。データ数Nのラプラス乱数列を500回発生させて500個の統計量Uよりχ^2検定とKS検定を行った結果を表１に示す。χ^2検定は区間数を11に取っているので自由度は８となり、1%水準値は20.1である。また、KS検定の1%水準値は0.073である。共に有意水準1%で仮説が採択されるのはNが約500以上の時であることが分かる。

表１　統計量Uのχ^2値とKS値

N	χ^2値	KS値
100	5.44	0.127
200	19.19	0.113
300	6.84	0.107
400	7.70	0.117
500	9.79	0.057
600	6.99	0.069
700	10.53	0.070
800	7.04	0.064
900	5.47	0.068

6．むすび

本論文では、まずラプラス分布の母数を推定する式が、モーメント法と最尤法では異なるが、漸近特性では値が一致することを示した。そして、２つの方法で求めた母数を変数とする統計量Uを定義しデータ数Nが約500以上であれば、Uの分布は単位正規分布に従うことを、統計的仮説検定により示した。よって、統計量Uによりラプラス分布の適合度が単位正規分布のパーセンタイルにより検定できる。

謝辞

本研究に対して機会を与えて下さった太田教授に感謝します。

参考文献

(1)A.Papoulis:"PROBABILITY & STATISTICS" ,PRINTICE HALL,Inc.(1990)

(2)G.P.Wadsworth,J.G.Bryan:"Applications of Probability and Random Values", McGraw-Hill Book Co.(1974)