単峰性をもつ歪んだ分布族

藤澤　洋徳　数理・推論研究系　准教授

はじめに

対称な分布は正規分布をはじめとして様々な分布が提案されている．対称でない歪んだ分布もデータの存在範囲が有限区間や半開区間の場合は様々に提案されている．しかしながら，データの存在範囲が実数全体の場合には，歪んだ分布の提案は，いまだに発展途上段階にある．そのような分布の族の代表例としては歪対称分布族がある．この分布族は非常にハンディである．しかしながら，単峰性が自然に保証されていないなどの欠点を持っている．本研究では，この問題点を克服しつつ，以下のような特徴をもった分布族を提案する．

提案する分布族の特徴

単峰性が保証されている．正規化定数が簡単である．位置･尺度･歪度･尖度などに関するパラメータの役割をはっきりすることができる．歪度に関して単調性をもつ．欲しい性質をもつ分布の作成が簡単である．乱数の発生が簡単である．回帰モデルのノイズ分布として使いやすい．

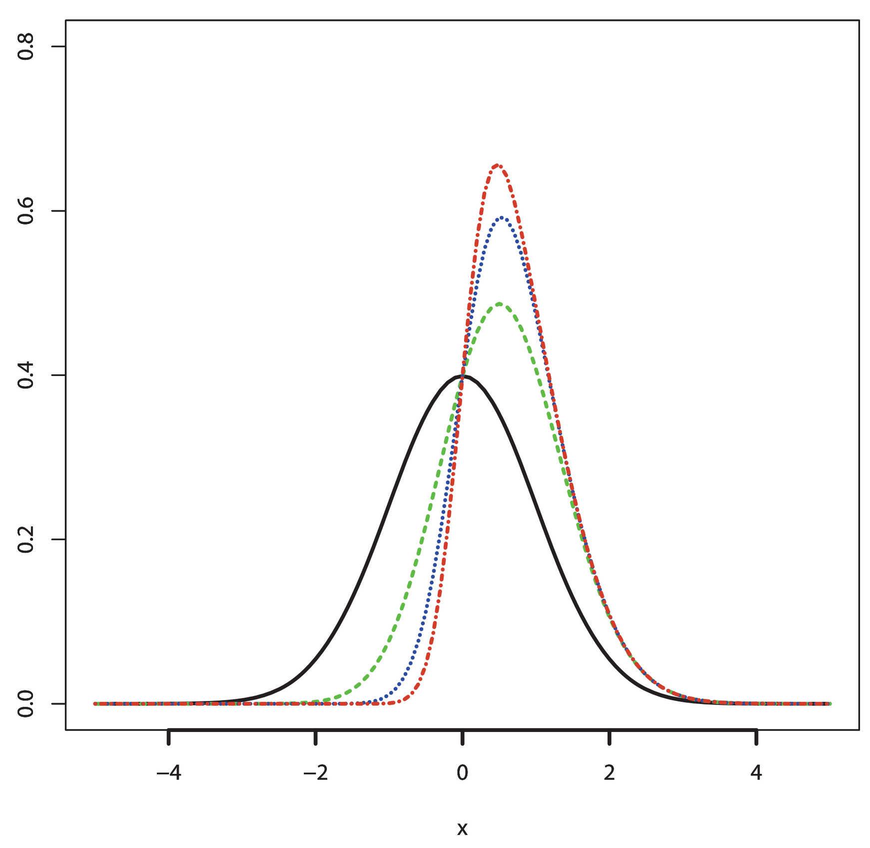
歪対称分布

歪対称分布の密度関数はたとえば次で表現される：

$$f_G(x;\lambda) = 2G(\lambda x)f(x).$$

ここで，$f(x)$は原点対称の密度関数であり，$G(x)$は原点対称な密度関数をもつ分布関数である．

特に$f(x)$と$G(x)$が標準正規分布の密度関数と分布関数である場合の密度関数$f_G(x;\lambda)$の動向を図で表現しておく．



なお，この場合は，正規分布の特殊性から単峰性は保証されているが，一般的には単峰性は保証されない．モードが同じ場所に留まらずに右に左に動いているのも気になる点である．

単峰な歪対称分布族

対応する密度関数は次で表現される：

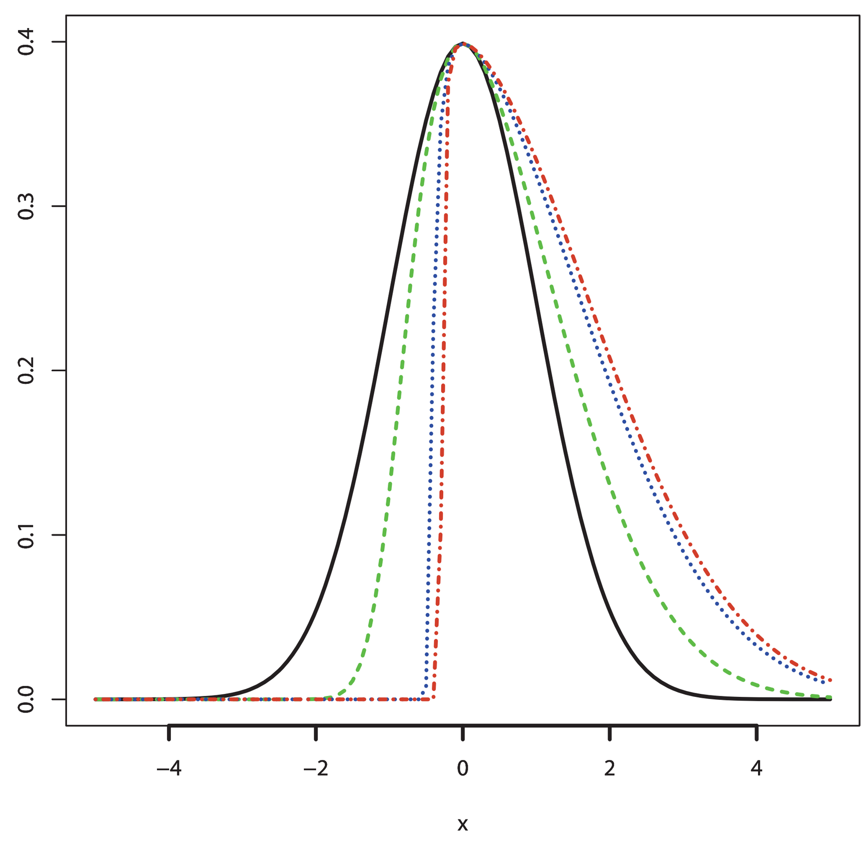
$$f_r(x;\lambda) = f(r(x;\lambda)).$$

ここで，$f(x)$はやはり原点対称の密度関数であり，$r(x)$はその逆関数$r^{-1}(y) = s(y)$が次の性質を満たす関数とする：$s(y) = y + H(y)$と表現した時に$h(y) = H'(y)$は奇関数であり$h(y) = h(-y)$をみたす．

ここで$f(x)$が標準正規分布の密度関数であり関数$H(y)$が以下であるときの密度関数$f_r(x;\lambda)$の密度関数の動向を図で表現しておく：

$$H(y;\lambda) = a_\lambda \frac{\sqrt{1+\lambda^2 y^2}-1}{\lambda}, \quad a_\lambda = 1 - e^{-\lambda^2}.$$

なお，関数Hは，結果として生じる密度関数$f_r(x;\lambda)$に対する性質を適当に要求すると，自然に考えだされる関数の一つである．実際には，要求に応じて，適当に関数Hを構成することで，様々な歪んだ分布を構成することができる．（詳細は論文を参照されたい．）



モード不変

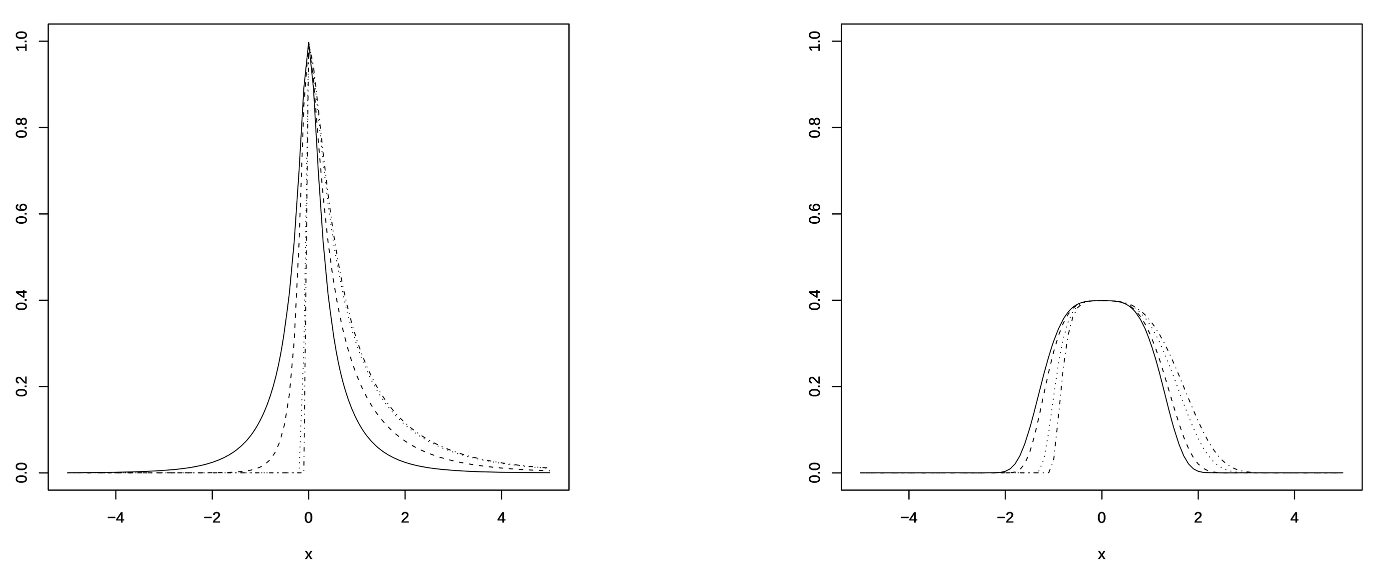
単峰性を保存しながらモードの位置が不変であることが特徴的である．この性質は，位置パラメータや尺度パラメータなどを導入したとしても，それぞれのパラメータの解釈を容易にする．ある程度のパラメータ直交性も証明できる．さらに，回帰モデルのノイズが歪んでいるときに，ノイズ分布として使いやすい長所もある．

歪度の単調性

また，歪みが自然に増えて行く様子が見て取れる．実際に，歪度パラメータλが増えるに従って，代表的な歪度の基準である三次モーメント$\mathrm{E}[X^3]$は単調に増えるし，密度関数だけから定義される別の有名な指標 (Arnold and Groeneveld, 1995) も単調に増える．

分布表現の柔軟性

密度関数fとして使える対称分布に制限はないので，適当に選ぶことで，尖っていたり，逆に尖っていなかったり，そういう分布も簡単に構成できる．



確率変数表現と乱数の発生

単峰な歪対称分布の密度関数f_rに対応する確率変数をXとし，対称な密度関数fに対応する確率変数をYとする．一様分布に従う確率変数をUとする．このとき次のような関係が得られる：

$$X = s(Y)I(U < s'(Y)/2) + s(-Y)I(U \geq s'(Y)/2).$$

この関係を利用して乱数を簡単に生成することができる．

この乱数を使って，導出した方法が，ノイズが非対称な時にどういうパフォーマンスを示すかを調べることができる．

参考文献

Fujisawa, H. and Abe, T. (2012). A family of unimodal skew-symmetric distributions with mode-invariance. ISM Research Memorandum No.1151.