

二変数ガンマ分布における密度分布と相関母数の関連性

名古屋工業大学 正員 長尾正志

1. 研究の経緯と目的

水工計画において、ガンマ分布は、形状母数の選択に応じて、一方的に減少する逆J字型分布、指数分布、正の歪をもつ単一ピーク分布、さらに、形状が正規分布に近いものまで、広い範囲にわたる分布型を近似しうるので、極めて有用性が高い。二変数ガンマ分布は、その2次元への拡張である。もちろん、周辺分布がガンマ分布で、所与の母数をもつ二変数ガンマ分布は確定できず、種々の理論分布が存在しうる。著者は、Wicksell, Kibbleらの提案になる理論分布式を利用しているが、この理由は、理論的に健全で、回帰特性が線型的であるなど実用性に優れると考えるからである。

しかし、この式は、第1種変形ベッセル関数を含むために、取扱いが難しいという批判もある。もちろん、現代の電算の能力からみれば、さほど本質的な難点とも思えないが、その特性を明確にするとともに、比較的容易に計算しうる工夫も必要と考え、密度分布と相関母数の関連性を調べてみた。

2. 二変数ガンマ分布の密度分布曲面

以下、形状母数の等しい場合についてのみ扱う。密度分布と相関母数の関連性を理解するために、相関母数の相違による密度分布曲面への影響を、3次元のコンピュータ・グラフィックス(陰線処理済み)で示す。水平軸には、標準化変量 $\xi = x_1/\lambda_1$, $\eta = x_2/\lambda_2$ (x_i :ガンマ変量, λ_i :尺度母数, $i=1, 2$) をとり、二変数指数分布と二変数ガンマ分布に分けて説明する。

2.1 二変数指数分布 a. 理論式 標準化変量 ξ, η による結合密度分布は次式で示される。

$$f(\xi, \eta) = (1-p)^{-1} \exp\{-(\xi+\eta)/(1-p)\} \cdot I_0(2\sqrt{p\xi\eta}/(1-p)) \quad (I_0: \text{変形ベッセル関数}) \quad (1)$$

b. 密度分布曲面 変域 $0 \leq \xi, \eta \leq 3$, 相関母数 $p = 0.1, 0.5, 0.9$ の曲面を図-1に示す。縦軸 f は水平軸の2倍に拡大し、 f 軸を中心とした回転角は 30° , f 軸の $\xi \sim \eta$ 面に直交した方向からの回転角は 40° , ξ 軸の分画数は60としている。

図より、相関母数 p が1(完全従属)に近いと、条件付分布は単峰分布形となり、密度曲面が対称軸 $\xi = \eta$ の近傍に集中する。また、 $p = 0$ に近づくと、結合分布は、それぞれの周辺分布 $f_1(\xi) = e^{-\xi}$, $f_2(\eta) = e^{-\eta}$ である指数分布の積として積に近づいていく様子が読みとれよう。

2.2 二変数ガンマ分布 a. 理論式 $I_{\nu-1}(x)$ を第1種変形ベッセル関数として、次式となる。

$$f(\xi, \eta) = \Gamma(\nu)^{-1} \cdot (1-p)^{-(\nu-1)/2} \cdot \exp\{-(\xi+\eta)/(1-p)\} \cdot (\xi\eta)^{(\nu-1)/2} \cdot I_{\nu-1}(2\sqrt{p\xi\eta}/(1-p)) \quad (2)$$

ここでは、 $\nu = 2$ (単一ピークの分布) の場合を示す。上式より、その密度分布は次式で表わされる。

$$f(\xi, \eta) = (1-p)^{-1/2} \cdot \exp\{-(\xi+\eta)/(1-p)\} \cdot I_1(2\sqrt{p\xi\eta}/(1-p)) \quad (3)$$

b. 密度分布曲面 密度分布曲面を図-2に示す。縦軸は水平軸の2.0倍とし、 f 軸を中心とした回転角は 30° , f 軸の $\xi \sim \eta$ 面に直交した方向からの回転角は 50° , ξ 軸の分画数は50である。この場合、正規分布でみられるように、相関母数 p が1に近いと、対称軸 $\xi = \eta$ に集中し、 p が0に近づくと、結合分布は、周辺分布 $f_1(\xi) = \xi e^{-\xi}$, $f_2(\eta) = \eta e^{-\eta}$ の積に近づいていく様子がみえる。

3. 密度分布の相関母数による多項式展開

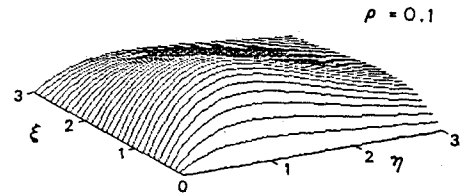
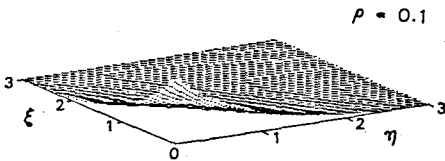
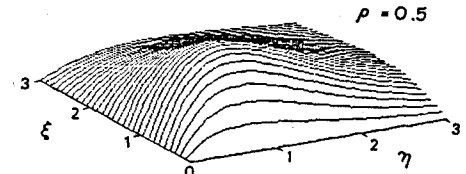
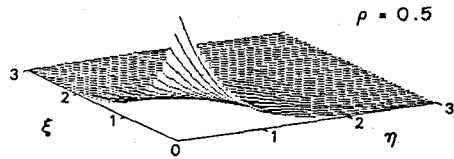
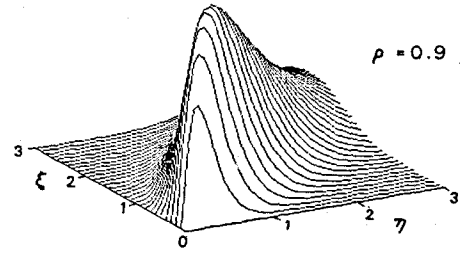
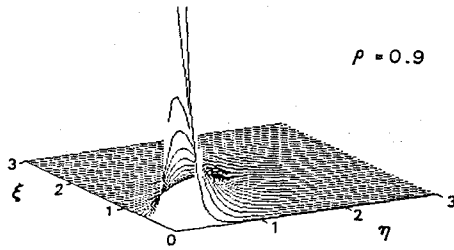


図-1 二変数指数分布の密度分布曲面

図-2 二変数ガンマ分布の密度分布曲面

相関母数による密度分布への影響をみるために、相関母数による直交多項式としての展開を示す。

3.1 基礎式 $f(\xi, \eta) = \{ \Gamma(\nu) \}^{-2} e^{-\xi-\eta} (\xi \eta)^{\nu-1} \sum_{n=0}^{\infty} n B(n, \nu) \cdot L_n^{\nu-1}(\xi) \cdot L_n^{\nu-1}(\eta) \rho^n$ (4)

ここに、 $B(x, y) = \Gamma(x) \cdot \Gamma(y) / \Gamma(x+y)$: ベータ関数、 $L_n^{\alpha}(x) = \sum_{m=0}^n (-1)^m \binom{n+\alpha}{n-m} x^m / m!$: ラゲール多項式 (5)

3.2 相関母数を用いた有限多項式近似 紙数の関係上、二変数指数分布のみを説明する。このとき、(4)式は、 $f(\xi, \eta) = e^{-\xi-\eta} \sum_{n=0}^{\infty} L_n(\xi) \cdot L_n(\eta) \rho^n$ (6) となる。ここに、 $L_n(x) = L_n^1(x) = \sum_{m=0}^n (-1)^m \binom{n}{n-m} x^m / m!$: ラゲール多項式である。無限級数の(5)式を、有限項 n までで近似した場合の相対誤差 E_R は次式で求められる。 $E_R = e^{-\xi-\eta} \sum_{n=0}^{\infty} L_n(\xi) \cdot L_n(\eta) \rho^n / f(\xi, \eta)$ (7) なお、上式右辺の f は式(1)で与えられる。この E_R を種々の n, ρ および ξ, η について計算した。計算結果の概略は以下のとおりである。

- ξ, η の変域としては、 $\xi = 0$ で $f(0, \eta) = (1-\rho)^{-1} \exp \{-\eta / (1-\rho)\}$ と分かるから、常識的な範囲 0.5 (0.5) 5.0 とする。また、対称性より、 $\xi \leq \eta$ とし、誤差許容基準を±15%とする。
- $\rho = 0.3$ のとき、 $n=1$ で誤差基準外のものは55ケースのうち4なので、ほぼ線形近似でよい。もちろん、 $n=2$ (2次近似) とすれば、基準外はなくなり、最大相対誤差も4%程度に収まる。
- $\rho = 0.5$ のとき、 $n=2$ で基準外 8/55、 $n=3$ で 3/55で、3項までで近似でほぼよい。
- $\rho = 0.7$ のとき、 $n=10$ 位にしてほぼ誤差基準に収まるが、こんな莫大な計算をするより、直接変形ベッセル関数を、 $I_{\nu}(z) = \sum_{n=0}^{\infty} z^{\nu+2n} / \{ 2^{\nu+2n} n! \Gamma(\nu+n+1) \}$ で計算する方が得策であろう。現実には、相関母数はあまり大きくないことが多いが、以上の結果より、 $\rho \leq 0.3$ 位なら線形近似、0.5位までで3次近似、それ以上の ρ の場合には、直接変形ベッセル関数を求める方がよいであろう。