

ショットノイズモデルによる模擬発生時系列の分布特性

神戸大学工学部 正員 神田 徹
兵 庫 県 正員 ○山内良太

1. まえがき

ある時間単位の変量がガンマ分布に従うときには、その k 倍($k \geq 1$)の時間単位の変量もガンマ分布で近似でき、そのパラメータを推定できることが理論的に示されている¹⁾。しかし、厳密にはガンマ分布に従わない水文量に対してこの理論を適用する場合、推定された分布特性値がその水文量の真の分布特性値とどれだけ異なるかを示すことが本文の第一の目的である。このため、水文量時系列としてはショット・ノイズ過程に従う流量時系列を対象とした。次に、水文量時系列がショット・ノイズ過程に従うものとするとき、そのモデルパラメータの推定において、標本から求める積率の誤差がパラメータの値にどの程度影響を及ぼすかについて調べた。

2. ショット・ノイズ過程の特性値²⁾

ショット・ノイズ過程は(1)式で定義され、このモデルで模擬発生される連続的過程 $X(t)$ は(2)式のガンマ分布(2母数)に従う。

$$X(t) = \sum_{i=1}^{N(t)} y_i \exp\{-b(t-\tau_i)\} \quad (1)$$

$$f\{X(t)\} = (1/\theta)\{1/\Gamma(v/b)\}\{X(t)/\theta\}^{v/b-1} \exp\{-X(t)/\theta\} \quad (2)$$

ここに、 $N(t)$ は事象発生率 v のポアソン過程、 y は降雨量に相当するもので平均 θ の指数分布に従い、 b は低減率である。実際の流量資料は(3)式のように期間 T における平均値 X_t として与えられる。この平均化過程 X_t の積率とパラメータの関係は(4)式で与えられる。

$$X_t = \frac{1}{T} \int_0^T X(s) ds \quad (3)$$

$$\left. \begin{aligned} \text{平均値: } E\{X_t\} &= v\theta/b & \text{分散: } \text{Var}\{X_t\} &= (v\theta^2/b) \cdot 2\{bT - (1-e^{-bT})\} / (bT)^2 \\ \text{ひずみ係数: } C\{X_t\} &= (b/v)^{1/2} \cdot 3\{bT - (1-e^{-bT})\} - (1/2)(1-e^{-bT})^2 / [2\{bT - (1-e^{-bT})\}^3]^{1/2} \\ \text{lag-1自己相関係数: } \rho\{X_t, X_{t+1}\} &= (1-e^{-bT})^2 / [2\{bT - (1-e^{-bT})\}] \end{aligned} \right\} \quad (4)$$

また、(5)式で定義される $X(t)$ の期間 T での総量 $S_t(T)$ については、その積率とパラメータの関係は(6)式で与えられる。なお、ひずみ係数、自己相関係数は(4)式と同じである。

$$S_t(T) = \int_0^T X(s) ds \quad (5)$$

$$\text{平均値: } E\{S_t(T)\} = (v\theta/b)T \quad \text{分散: } \text{Var}\{S_t(T)\} = (v\theta^2/b)[2\{bT - (1-e^{-bT})\}]/b^2 \quad (6)$$

3. ショット・ノイズ過程変量に対するガンマ分布の近似度

いま、 X_t を単位期間 $T=1$ での平均化過程とすれば $S_t(1) = X_t$ であるので、以下においては一般的に $S_t(T)$ について検討する。 $S_t(T)$ の分布を次式のガンマ分布(2母数)で近似する。

$$f\{S_t(T)\} = (1/\alpha)\{1/\Gamma(\theta)\}\{S_t(T)/\alpha\}^{\theta-1} \exp\{-S_t(T)/\alpha\} \quad (7)$$

ここに、 α は尺度母数 ($\alpha > 0$)、 β は形状母数 ($\beta > 0$) である。このとき、 $S_t(T)$ の積率とパラメータの関係は(8)式で与えられる。

$$\left. \begin{aligned} \text{平均値: } \mu &= \alpha\beta & \text{分散: } \sigma^2 &= \alpha^2\beta \\ \text{ひずみ係数: } \gamma &= 2/\beta^{1/2} \end{aligned} \right\} (8)$$

ここで、平均値と分散が(6)式の値と等しいようなガンマ分布を対象として、この分布の近似度を次式のひずみ係数の相対誤差で評価する。

$$e = \frac{\gamma - C\{S_t(T)\}}{C\{S_t(T)\}} = \frac{4\{bT - (1 - e^{-bT})\}^2}{3bT\{bT - (1 - e^{-bT}) - (1/2)(1 - e^{-bT})^2\}} - 1 \quad (9)$$

上式を図示すれば図-1のようであり、 $T=1$ 、すなわち X_t については近似度は良好といえるが、 T が大きくなるにつれて近似度は悪くなっている。

4. ガンマ分布理論による積率推定の精度

単位期間の変量がガンマ分布(2母数)に従うとき、その T 倍の期間の変量は十分の近似でガンマ分布(2母数)に従い、両者の母数間には次式の関係がある。¹⁾

$$\left. \begin{aligned} \rho(T) &= \frac{\rho(1)[T\{1-\rho(1)^2\} - 2\rho(1)\{1-\rho(1)^T\}]}{T^2\{1-\rho(1)\}^2} = g(\rho(1), T) \\ \frac{\xi(1)}{\xi(T)} &= \frac{\eta(T)}{T\eta(1)} = \frac{\rho(T)}{\rho(1)} \end{aligned} \right\} (10)$$

ここに、 ξ は形状母数、 η は尺度母数、 ρ は lag-1 自己相関係数で、カッコ内の 1、 T はそれぞれ単位期間、その T 倍の期間を示す。上式の関係より、 $\{\xi(1), \eta(1), \rho(1)\}$ か $\{\xi(T), \eta(T), \rho(T)\}$ のいずれか一方が既知ならば、他方の母数が求まり分布形が定まる。

(1) 単位期間の分布からその T 倍の期間の分布を推定する場合

ショット・ノイズ過程変量 X_t がガンマ分布に従うとみなして $S_t(T)$ の積率を上記の式から推定した場合、その推定値の真の積率に対する誤差は(11)、(12)式で与えられる。

$$e_r(\sigma^2) = \frac{\sigma^2(T) - \text{Var}\{S_t(T)\}}{\text{Var}\{S_t(T)\}} = \frac{[T\{1-\rho(1)^2\} - 2\rho(1)\{1-\rho(1)^T\}]\{b-1+e^{-b}\}}{\{1-\rho(1)\}^2(bT-1+e^{-bT})} - 1 \quad (11)$$

$$e_r(\gamma) = \frac{\gamma(T) - C\{S_t(T)\}}{C\{S_t(T)\}} = \frac{4\sqrt{T\{1-\rho(1)^2\} - 2\rho(1)\{1-\rho(1)^T\}}\sqrt{b-1+e^{-b}}\sqrt{(bT-1+e^{-bT})^3}}{3bT\{1-\rho(1)\}\{bT - (1 - e^{-bT}) - (1/2)(1 - e^{-bT})^2\}} - 1 \quad (12)$$

ここに、 $e_r(\sigma^2)$ は分散の相対誤差、 $e_r(\gamma)$ はひずみ係数の相対誤差、 $\sigma^2(T)$ 、 $\gamma(T)$ はそれぞれ(10)式より求まる分散、ひずみ係数である。上式を図示すれば図-2、3のようであり

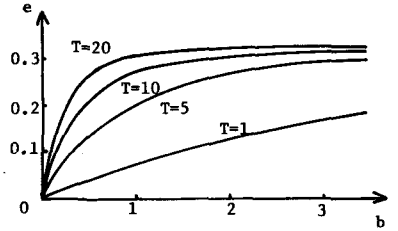


図-1 $X_t, S_t(T)$ のひずみ係数の相対誤差

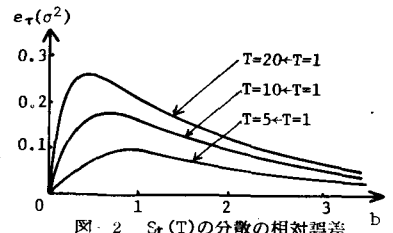


図-2 $S_t(T)$ の分散の相対誤差

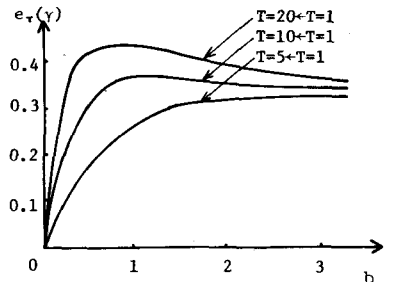


図-3 $S_t(T)$ のひずみ係数の相対誤差

ひずみ係数の相対誤差は X_t のガンマ分布近似の相対誤差 (図-1) とほとんど変わらないのに対して、分散の相対誤差は b が小さい範囲で大きくなっている。

(2) T 倍の期間の分布から単位期間の分布を推定する場合

(1) とは逆に、 T 倍の期間の分布をガンマ分布とみなして単位期間の変量 X_t の積率を推定した場合には、その誤差は次式、及び図-4, 5 で与えられる。

$$e_1(\sigma^2) = \frac{\sigma^2(1) - \text{Var}\{X_t\}}{\text{Var}\{X_t\}} = \frac{2\{bT - (1 - e^{-bT})\}^2 g^{-1}(P(T), T)}{T^2(1 - e^{-bT})^2 \{b - (1 - e^{-b})\}} - 1 \quad (13)$$

$$e_1(\gamma) = \frac{\gamma(1) - C\{X_t\}}{C\{X_t\}} = \frac{4\{bT - (1 - e^{-bT})\} \sqrt{2\{b - (1 - e^{-b})\}^3} g^{-1}(P(T), T)}{3bT(1 - e^{-bT}) \{b - (1 - e^{-b}) - (1/2)(1 - e^{-b})^2\}} - 1 \quad (14)$$

図のように、ひずみ係数の相対誤差は $S_t(T)$ のガンマ分布近似における相対誤差 (図-1) とほとんど変わらないのに対して、分散は T が大きいほど、また b が小さい場合にかなり大きな誤差を生じる。

以上、積率の相対誤差によって分布の近似度を検討したが、分布形そのものの対応については講演時に示す。

5. 積率推定の誤差がパラメータの推定に及ぼす影響

X_t あるいは $S_t(T)$ の母集団積率の値が正しく推定できれば、(4), (6) 式から b , v , θ の値を得ることができる。

しかし、母集団積率は観測資料 (標本) から推定するので、その推定誤差のために b , v , θ の推定に誤差が生じる。以下では、両者の誤差の関係について検討する。なお、標本から母集団積率を推定する場合の誤差の特性 (特に、偏り) については別の機会に報告する。

パラメータ b , v , θ は時間単位 T の変量の積率 (平均: $m = E\{S_t(T)\}$, 分散: $s^2 = \text{Var}\{S_t(T)\}$, lag-1 自己相関係数: $\rho_1 = \rho\{S_t(T), S_{t+1}(T)\}$) から次式より求まる ((4), (6) 式)。

$$\frac{(1 - e^{-bT})^2}{2\{bT - (1 - e^{-bT})\}} = \rho_1 \quad (15)$$

$$\theta = \frac{s^2 b^2 T}{2m\{bT - (1 - e^{-bT})\}} \quad (16)$$

$$v = \frac{2m^2\{bT - (1 - e^{-bT})\}}{s^2 b T^2} \quad (17)$$

このとき、積率の誤差とパラメータ b , v , θ の誤差の関係は次式で与えられる。

$$\frac{\Delta b}{b} = \frac{(1 - e^{-bT})(bT - 1 + e^{-bT})}{bT(e^{-2bT} + 2bTe^{-bT} - 1)} \frac{\Delta \rho_1}{\rho_1} = f_b(b, T) \frac{\Delta \rho_1}{\rho_1} \quad (18)$$

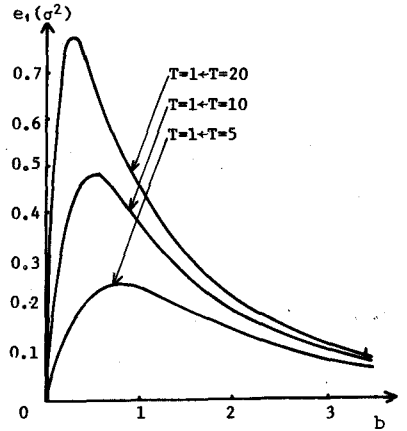


図-4 X_t の分散の相対誤差

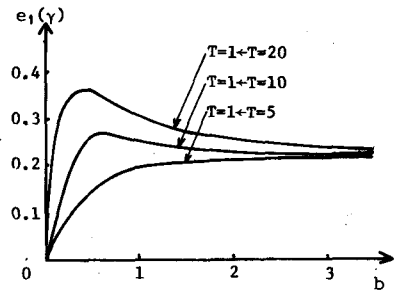


図-5 X_t のひずみ係数の相対誤差

$$\begin{aligned}\frac{\Delta\theta}{\theta} &= -\frac{\Delta m}{m} + \frac{\Delta s^2}{s^2} + \frac{(1-e^{-bT})\{bT-2+e^{-bT}(2+bT)\}}{bT(e^{-2bT}+2bTe^{-bT}-1)} \frac{\Delta\rho_i}{\rho_i} \\ &= -\frac{\Delta m}{m} + \frac{\Delta s^2}{s^2} + f_\theta(b,T) \frac{\Delta\rho_i}{\rho_i}\end{aligned}\quad (19)$$

$$\begin{aligned}\frac{\Delta v}{v} &= 2\frac{\Delta m}{m} - \frac{\Delta s^2}{s^2} + \frac{(1-e^{-bT})\{1-e^{-bT}(1+bT)\}}{bT(e^{-2bT}+2bTe^{-bT}-1)} \frac{\Delta\rho_i}{\rho_i} \\ &= 2\frac{\Delta m}{m} - \frac{\Delta s^2}{s^2} + f_v(b,T) \frac{\Delta\rho_i}{\rho_i}\end{aligned}\quad (20)$$

上式中の $f_b(b,T)$, $f_\theta(b,T)$, $f_v(b,T)$ の値をそれぞれ図-6, 7, 8 に示す。上式よりわかるように, b には $\Delta\rho/\rho_i$ のみが影響し, θ と v には $\Delta\mu/\mu$, $\Delta\sigma^2/\sigma^2$, $\Delta\rho/\rho_i$ がそれぞれ (1 : 1 : $f_\theta(b,T)$), (2 : 1 : $f_v(b,T)$) の割合で影響を及ぼす。このように, $\Delta\rho/\rho_i$ の影響は b , T にも関係する。 T が大きくなるにつれて $f_b(b,T)$, $f_\theta(b,T)$, $f_v(b,T)$ はそれぞれ -1, -1, 0 に漸近する。また, T が小さくなるにつれて $f_b(b,T)$, $f_v(b,T)$ は絶対値が大きくなり, $f_\theta(b,T)$ は -0.5 に漸近する。つまり, ρ_i の影響に関しては, T が大きい場合には時系列の尺度に関するパラメータ θ に及ぼす影響が大きく, T が小さい場合には時系列の形状に関するパラメータ b , v に及ぼす影響が大きくなる。従って, 短い時間単位の標本からパラメータを推定する場合には, ρ_i の推定に高い精度が要求される。一定の観測期間の資料からパラメータを推定する場合, 短い時間単位の変量を用いる方が長い時間単位の変量よりも大きな標本が得られるので, 推定値のばらつきは小さい。従って, パラメータ推定の問題は, 積率推定における偏り補正の問題に帰着する。

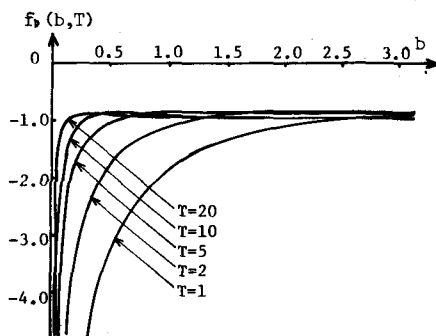


図-6 $f_b(b,T)$ の値

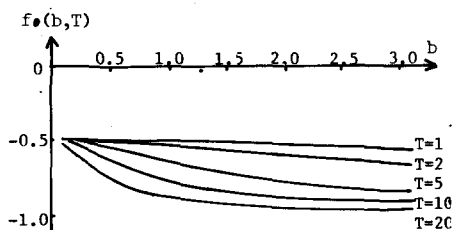


図-7 $f_\theta(b,T)$ の値

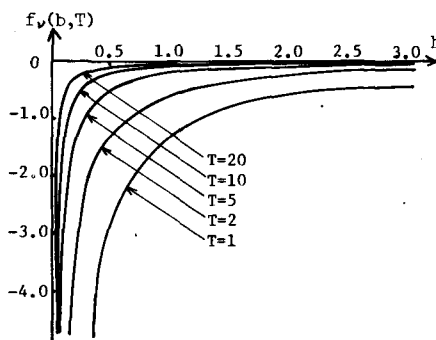


図-8 $f_v(b,T)$ の値

参考文献

- 1) 長尾正志：確率雨量配分率曲線の理論的推定，土木学会論文報告集，No. 243, 1975.
- 2) Weiss, G. : Shot Noise Model for the Generation of Synthetic Streamflow Data, W. R. R., Vol. 13, No. 1, 1977.