

二変数ガンマ分布とその適用に関する研究 (2)

——二変数ガンマ分布の母数推定法——

長 尾 正 志・角 屋 睦

STUDY ON TWO-VARIATE GAMMA DISTRIBUTION AND ITS ENGINEERING APPLICATION (2)

—— Estimation of Population Parameters ——

By Masashi NAGAO and Mutsumi KADOYA

Synopsis

In the first report, the fundamental characteristics on two-variate gamma distribution, especially two-variate exponential distribution were discussed. In this report, as the second step, the estimation of population parameters of two-variate gamma distribution are developed for the cases that a sample from a total population is given and that a sample from a sub-population. In outline, the study is as follows:

- 1) The estimating method based on a sample from a total population, every marginal distribution of which has two parameters, is developed by using the technique of the moments method.
- 2) The estimating method based on a sample from a sub-population, in which the domain of independent variable is defined as upper tail over a given value, is developed by using the technique of the maximum likelihood method.
- 3) The applicability of the theory of two-variate gamma distribution is examined by applying it to the data of river flows.

1. は し が き

ガンマ分布は、その母数の選択に応じて、代表的な対称分布である正規分布に近い形から、指数分布のような単調減少する非対称分布に至る広範囲の形状を表現しうることから、水文統計のみならず、多くの分野でも応用面の広い分布と注目されている。しかしながら、相関を考慮した二変数ないし多変数ガンマ分布に関する研究は、実用上きわめて重要と思われるにもかかわらず、純粋数学的な興味で行なわれている若干のものを除くと、ほとんど見当たらない。われわれは第1報において二変数ガンマ分布およびその特別な場合に相当する二変数指数分布についての基礎理論について若干の考察を進めたが、本報では母集団分布の母数推定、とくに、全標本による従来の推定結果の形状母数の異なる場合への拡張、および特殊な部分標本における推定の問題を理論的に考察した結果について報告する。なお、論議の展開上、必要があれば従来の研究結果についても引用していくことにする。

2. 基 礎 式

一変数のガンマ分布のもっとも一般的な形は、次式で定義される。

$$f(x)dx = \frac{1}{\Gamma(\nu)} \left(\frac{x-v}{\sigma} \right)^{\nu-1} \exp\left(-\frac{x-v}{\sigma}\right) d\left(\frac{x-v}{\sigma}\right), \quad \left. \begin{array}{l} x \geq v \\ , \quad x < v \end{array} \right\} \dots\dots\dots (1)$$

$$= 0$$

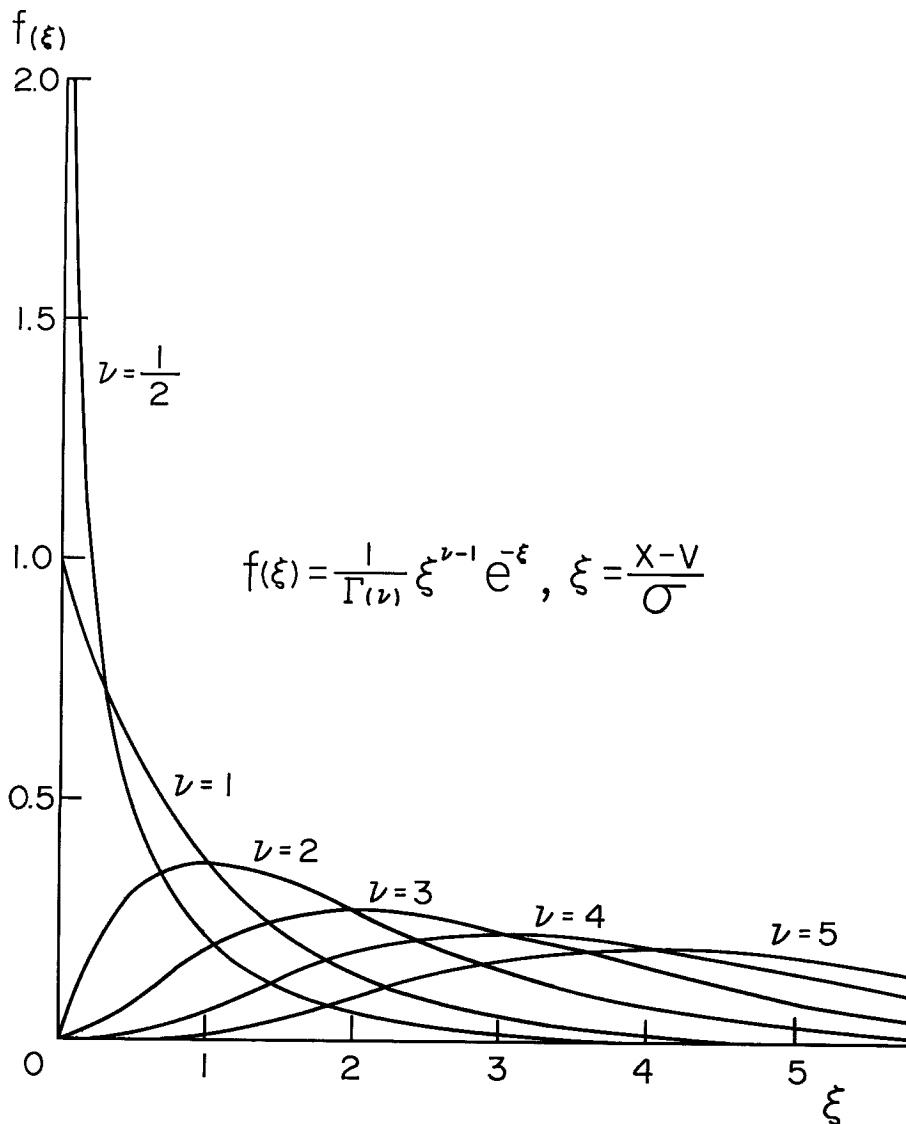


Fig. 1 Shapes of gamma distribution.

ここに、 ν は形状母数または指数、 v は原点母数、および σ は尺度母数といわれる定数である。 ν による分布の形状の変化は Fig. 1 に示したとおりである。すなわち、 ν の値が $0 < \nu < 1$ の範囲にあれば、 $f(x)$ は x の増加とともに単調に減少し、 $\nu \geq 1$ では $x = \sigma(\nu-1) + v$ で極大値をもつような非対称分布である。さらに ν が大きくなるにつれて正規分布に漸近する。とくに $\nu=1$ の場合は指数分布となる。なお、原点母数 v については普通 0 とおいても一般性を失わないから、とくに断らない限り以後 $v=0$ として議論を進める。

2.1 形状母数の異なる周辺分布をもつ二変数ガンマ分布

周辺分布が形状母数 $\nu_1 = n+m$, $\nu_2 = n$ ($m, n \geq 0$) のガンマ分布である二変数 x_1, x_2 についての密度分布関数は井沢¹⁾ によって (2) 式で与えられている。

$$f(x_1, x_2) = \frac{1}{\Gamma(n) \Gamma(m) (\sigma_1 \sigma_2)^{\frac{n+1}{2}} \sigma_1^m (1-\rho) \rho^{\frac{n-1}{2}}} \times (x_1 x_2)^{\frac{n-1}{2}} x_1^m \\ \times \exp\left\{-\frac{x_1}{\sigma_1(1-\rho)} - \frac{x_2}{\sigma_2(1-\rho)}\right\} \times \int_0^1 (1-t)^{\frac{n-1}{2}} t^{m-1} \exp\left\{\frac{\rho x_1 t}{\sigma_1(1-\rho)}\right\} \\ \times I_{n-1}\left(\frac{2\sqrt{\rho}}{1-\rho} \sqrt{\frac{x_1 x_2}{\sigma_1 \sigma_2}} (1-t)\right) dt \quad \dots\dots\dots (2)$$

ここに、 $\Gamma(\nu)$ はガンマ関数、 $I_\nu(z)$ は ν 次の変形ベッセル関数であり、また σ_1, σ_2, ρ は定数である。とくに ρ は、定義式より、 $0 \leq \rho \leq 1$ でなければならない。この分布関数を基礎にして種々の統計量を導くことができるが、ここでは以後直接関係する積率のみを導いておく。

まず、原点のまわりの p, q 次の積率はつぎのように書ける。

$$\nu_{pq} \equiv \int_0^\infty \int_0^\infty x_1^p x_2^q f(x_1, x_2) dx_1 dx_2 \\ = C_1 \int_0^1 (1-t)^{\frac{n-1}{2}} t^{m-1} \left\{ \int_0^\infty x_1^{\frac{n-1}{2}+m+p} \exp\left\{-\frac{(1-\rho)t}{\sigma_1(1-\rho)} x_1\right\} \right. \\ \times \left. \left(\int_0^\infty x_2^{\frac{n-1}{2}+q} \exp\left\{-\frac{x_2}{\sigma_2(1-\rho)}\right\} \times I_{n-1}\left(\frac{2\sqrt{\rho(1-t)}}{\sqrt{\sigma_1 \sigma_2}(1-\rho)} \sqrt{x_1} \sqrt{x_2}\right) dx_2 \right) dx_1 \right\} dt \\ \dots\dots\dots (3)$$

ただし

$$C_1 = \{\Gamma(n) \Gamma(m) (\sigma_1 \sigma_2)^{\frac{n+1}{2}} \sigma_1^m (1-\rho) \rho^{\frac{n-1}{2}}\}^{-1}$$

ところで、(3) 式中の x_2 に関する積分項を I_2 とおくと、これは積分可能であって次式のようになる²⁾。

$$I_2 = \frac{\Gamma(n+q)}{\Gamma(n)} \rho^{\frac{n-1}{2}} (1-\rho)^{q+1} \sigma_1^{-\frac{n-1}{2}} \sigma_2^{\frac{n+1}{2}+q} (1-t)^{\frac{n-1}{2}} x_1^{\frac{n-1}{2}} \\ \times {}_1F_1\left(n+q; n; \frac{\rho(1-t)}{\sigma_1(1-\rho)} x_1\right)$$

ただし、 ${}_1F_1$ は合流型超幾何関数である。この結果を (3) 式に代入すると x_1 に関する積分項 I_1 が残るが、これも積分できて次式のようになる³⁾。

$$I_1 = \Gamma(n+m+p) \left\{ \frac{\sigma_1(1-\rho)}{1-\rho t} \right\}^{n+m+p} F\left(n+q, n+m+p; n; \rho \frac{1-t}{1-\rho t}\right)$$

ただし F は超幾何関数である。そこで F に関する公式⁴⁾、

$$F\left(n+q, n+m+p; n; \rho \frac{1-t}{1-\rho t}\right) \\ = \left(\frac{1-\rho t}{1-\rho}\right)^{n+m+p+q} F\left(-m-p, -q; n; \frac{\rho(1-t)}{1-\rho t}\right)$$

および変数変換 $(1-t)/(1-\rho t) = x$ を用いて (3) 式を積分すると⁵⁾、求める積率は結局次式で与えられる。

$$\nu_{pq} = \frac{\Gamma(n+m+p) \Gamma(n+q) (1-\rho)^{n+q} \sigma_1^p \sigma_2^q I(p, q)}{\{\Gamma(n)\}^2 \Gamma(m)} \left\{ \begin{array}{l} I(p, q) = \int_0^1 x^{n-1} (1-x)^{m-1} (1-\rho x)^{-n-m-q} G(p, q) dx \\ G(p, q) = F(-m-p, -q; n; \rho x) \end{array} \right\} \dots\dots\dots (4)$$

ところで、 $I(p, q)$ 内の積分演算を一般に実行することは困難であるが、 p, q が低次の場合には比較的容易に行なうことができる。ここでは以後の計算に必要な 2 次以下の積率を求めておく。すなわち、 p, q に具体的に 0, 1, 2 を代入すると $G(p, q)$ は ρx に関するつぎのような 2 次以下のべき級数で表わせる。

$$G(1, 0) = 1$$

$$G(0, 1) = 1 + \frac{m}{n} \rho x, \quad G(0, 2) = 1 + \frac{2m}{n} \rho x + \frac{m(m-1)}{n(n+1)} \rho^2 x^2$$

$$G(1, 1) = 1 + \frac{m+1}{n} \rho x$$

そこで, 積分公式

$$\int_0^1 x^{\lambda-1} (1-x)^{\mu-1} (1-\beta x)^{-\nu} dx = \frac{\Gamma(\lambda)\Gamma(\mu)}{\Gamma(\lambda+\mu)} F(\nu, \lambda; \lambda+\mu; \beta)$$

を用いて整理すると, 求める 2 次以下の積率は積率は結局つぎのように表わせる。

$$\left. \begin{aligned} \nu_{10} &= (n+m)\sigma_1 = \nu_1\sigma_1, & \nu_{01} &= n\sigma_2 = \nu_2\sigma_2 \\ \nu_{20} &= (n+m)(n+m+1)\sigma_1^2 = \nu_1(\nu_1+1)\sigma_1^2 \\ \nu_{02} &= n(n+1)\sigma_2^2 = \nu_2(\nu_2+1)\sigma_2^2 \\ \nu_{11} &= n(n+m+\rho)\sigma_1\sigma_2 = (\nu_1+\rho)\nu_2\sigma_1\sigma_2 \end{aligned} \right\} \dots\dots\dots (5)$$

2.2 形状母数の等しい周辺分布をもつ二変数ガンマ分布

さて実用上重要な二変数ガンマ分布は, 二変数の周辺分布の形状の等しいもの, 換言すれば形状母数 ν_1, ν_2 の等しい場合である。その場合の結果は前報⁶⁾ で報告しているが, 以下の考察の基礎として重要であるので最終的な結果のみを記しておく, 密度分布が (3)' 式, p, q 次の積率, およびその 2 次以下のものが (4)' および (5)' 式である。

$$f(x_1, x_2) = \frac{1}{\Gamma(\nu)(\sigma_1\sigma_2)^{\frac{\nu+1}{2}}(1-\rho)\rho^{\frac{\nu-1}{2}}} \times (x_1x_2)^{\frac{\nu-1}{2}} \times \exp \left\{ -\frac{x_1}{\sigma_1(1-\rho)} - \frac{x_2}{\sigma_2(1-\rho)} \right\} \times I_{\nu-1} \left(\frac{2\sqrt{\rho}}{1-\rho} \sqrt{\frac{x_1x_2}{\sigma_1\sigma_2}} \right) \dots\dots\dots (3)'$$

$$\nu_{pq} = \frac{\Gamma(\nu+p)\Gamma(\nu+q)}{\{\Gamma(\nu)\}^2} (1-\rho)^{p+q+\nu} \sigma_1^p \sigma_2^q F(\nu+p, \nu+q; \nu; \rho) \dots\dots\dots (4)'$$

$$\left. \begin{aligned} \nu_{10} &= \nu\sigma_1, & \nu_{01} &= \nu\sigma_2 \\ \nu_{02} &= \nu(\nu+1)\sigma_1^2, & \nu_{02} &= \nu(\nu+1)\sigma_2^2 \\ \nu_{11} &= (\nu+\rho)\nu\sigma_1\sigma_2 \end{aligned} \right\} \dots\dots\dots (5)'$$

3. 全標本による母数推定

3.1 形状母数の異なる場合

まず積率解であるが, これは標本と母集団の積率を等置することによって次式のように求められる。

$$\left. \begin{aligned} \hat{\nu}_1 &= \frac{(\bar{x}_1)^2}{\bar{x}_1^2 - (\bar{x}_1)^2}, & \hat{\nu}_2 &= \frac{(\bar{x}_2)^2}{\bar{x}_2^2 - (\bar{x}_2)^2} \\ \hat{\sigma}_1^2 &= \frac{\bar{x}_1^2 - (\bar{x}_1)^2}{\bar{x}_1}, & \hat{\sigma}_2^2 &= \frac{\bar{x}_2^2 - (\bar{x}_2)^2}{\bar{x}_2} \\ \hat{\rho} &= \frac{\bar{x}_1(\bar{x}_1\bar{x}_2 - \bar{x}_1\bar{x}_2)}{\bar{x}_2\{\bar{x}_1^2 - (\bar{x}_1)^2\}} = \sqrt{\frac{\hat{\nu}_1}{\hat{\nu}_2}} r \end{aligned} \right\} \dots\dots\dots (6)$$

ただし r は標本相関係数である。通常, $0 < \rho < 1$ かつ $n, m > 0$ と考えてよいから, (6) 式の最後の等式は, m/n が大, すなわち分布形状の違いが大になればなるほど, 標本相関係数 r の存在範囲が 0 と, 1 より小さいある値 $\hat{\rho}\sqrt{1/(1+m/n)}$ との狭い範囲内に限定されてくることを表わしている。これは, 二変数正規分布の場合に標本相関係数が必ず -1 と +1 の範囲であることに比べて, 顕著な差異である。

また, ν_1, ν_2 が既知の場合には (6) 式は (6)' 式のように表わされる。

$$\left. \begin{aligned} \hat{\sigma}_1 &= \frac{\bar{x}_1}{\nu_1}, & \hat{\sigma}_2 &= \frac{\bar{x}_2}{\nu_2} \\ \hat{\sigma}_1^2 &= \frac{\bar{x}_1^2}{\nu_1(\nu_1+1)}, & \hat{\sigma}_2^2 &= \frac{\bar{x}_2^2}{\nu_2(\nu_2+1)} \\ \hat{\rho} &= \nu_1 \frac{\bar{x}_1\bar{x}_2}{\bar{x}_1\bar{x}_2} - \nu_1 = \sqrt{\frac{\nu_1}{\nu_2}} r \end{aligned} \right\} \dots\dots\dots (6)'$$

なお、この場合の最尤解は、現在のところ、誘導は困難であるが、相関母数 ρ の推定値が積率解と一致しないことは、つぎの形状母数の等しい場合の結果から類推される。

3.2 形状母数の等しい場合

この場合の解は、前記の一般の場合において、 $\nu_1 = \nu_2 = \nu$ とおくことによって求められ、その結果はすでに前報に発表した⁷⁾が、後述の部分標本に対する論議を進める上で基礎となるので最終結果のみを記しておく。

a. 積率解

$$\left. \begin{aligned} \hat{\nu} &= \frac{(\bar{x}_1)^2}{x_1^2 - (\bar{x}_1)^2} = \frac{(\bar{x}_2)^2}{x_2^2 - (\bar{x}_2)^2} \\ \hat{\sigma}_1^2 &= \frac{x_1^2 - (\bar{x}_1)^2}{x_1}, \quad \hat{\sigma}_2^2 = \frac{x_2^2 - (\bar{x}_2)^2}{x_2} \\ \hat{\rho} &= \frac{\bar{x}_1 \bar{x}_2 - x_1 x_2}{\sqrt{x_1^2 - (\bar{x}_1)^2} \sqrt{x_2^2 - (\bar{x}_2)^2}} = r \end{aligned} \right\} \dots\dots\dots (7)$$

なお形状母数 ν を既知とする場合には、上式は次式のように表わされる。

$$\left. \begin{aligned} \hat{\sigma}_1^2 &= \frac{\bar{x}_1^2}{\nu}, \quad \hat{\sigma}_2^2 = \frac{\bar{x}_2^2}{\nu} \\ \hat{\sigma}_1^2 &= \frac{\bar{x}_1^2}{\nu(\nu+1)}, \quad \hat{\sigma}_2^2 = \frac{\bar{x}_2^2}{\nu(\nu+1)} \\ \hat{\rho} &= \nu \frac{\bar{x}_1 \bar{x}_2}{x_1 x_2} - \nu = r \end{aligned} \right\} \dots\dots\dots (7)'$$

b. 最尤解

最尤解としては形状母数 ν が既知の場合の解が次式のように求められている。

$$\left. \begin{aligned} \hat{\sigma}_1^2 &= \frac{\bar{x}_1^2}{\nu}, \quad \hat{\sigma}_2^2 = \frac{\bar{x}_2^2}{\nu} \\ \nu \sqrt{\hat{\rho}} &= \frac{1}{n} \sum_{i=1}^n \frac{I_\nu \left(\frac{2\sqrt{\hat{\rho}} \xi_i \eta_i}{1-\hat{\rho}} \right)}{I_{\nu-1} \left(\frac{2\sqrt{\hat{\rho}} \xi_i \eta_i}{1-\hat{\rho}} \right)} \sqrt{\xi_i \eta_i} \end{aligned} \right\} \dots\dots\dots (8)$$

ただし ξ_i, η_i は次式で与えられる規準化変数である。

$$\xi_i = x_{1i}/\hat{\sigma}_1, \quad \eta_i = x_{2i}/\hat{\sigma}_2 \quad \dots\dots\dots (9)$$

すなわち、(7)' と (8) 式を対比すると尺度母数 σ についての推定値は積率解、最尤解で一致するが、相関母数 ρ については一致しないことがわかる。

4. 部分標本による母数推定

以上は全標本が手に入った場合の推定法であったが、実際にはすべての標本が分かっている使用できる場合はむしろ少ない。たとえば、治水計画において、高水のピーク流量の分布が問題となるが、普通、高水観測は警戒水位などのある限界以上の水位に対して実施されている場合が多い。また、実用解析上からも、なるべく少数の標本で推定できれば好都合である。

ここでは、部分的な標本として、変数 x_1 に関してある限界値 x_{1c} 以上の標本がある場合、換言すれば、全標本数 N の x_1 のうち上位 n 個しか分かっていない場合に母数をどのように推定すべきかを考える。

さて、一般にこのような順序統計量の大きい方または小さい方が分らない場合の母数推定の問題は、一次的には、いわゆる中途打ち切り標本の理論 (censored sampling theory) が、気象統計、人口統計などの分野で、とくに下位標本に対して応用されている。^{7), 8)} ここでは、それを上位標本に置き直して考えてみる。説明の便宜上、 x_1 を制約付変数、 x_2 を制約外変数と呼んでおく。

4.1 制約付変数に関する母数推定

a. 原点母数が未知の場合

まず一般的な考察をするために、 x_1 に対する周辺分布 $g(x_1)$ を原点母数 v_1 を含めた次式で定義しておく。

$$g(x_1) = \frac{1}{\Gamma(\nu_1)\sigma_1^{\nu_1}} (x_1 - v_1)^{\nu_1-1} \exp\left(-\frac{x_1 - v_1}{\sigma_1}\right), \quad x_1 \geq v_1 \quad \dots\dots\dots (10)$$

なお、この 4.1 では簡単のために、以後添字 1 を省略し、たとえば x_1 を x のように、表現する。また、標本に小さい方から $x_{(1)}, x_{(2)}, \dots\dots$ のように添字を付して記すことにする。したがって、総数 N の標本 $x_{(j)}$ ($j=1, 2, \dots, N$) の大きい方から n 個が分っている場合、大きい方から i 番目 ($i=1, 2, \dots, n$) すなわち小さい方から $N-i+1$ 番目の順序統計量 $x_{(N-i+1)}$ の分布は、(10) 式で x に $x_{(N-i+1)}$ を代入すればよい。したがって、 x について上位 n 個の標本のもつ尤度は、標本 $x_{(N-n+1)}, x_{(N-n+2)}, \dots, x_{(N)}$ の同時分布として次式で与えられる。

$$\begin{aligned} & L(\nu, \sigma; x_{(N-n+1)}, x_{(N-n+2)}, \dots, x_{(N-1)}, x_{(N)}) \\ &= \left(\frac{N}{N-n} \right) \left[\int_v^{x_{(N-n+1)}} g(u) du \right]^{N-n} \cdot n! \prod_{j=N-n+1}^N g(x_{(j)}) \\ &= \frac{N!}{(N-n)!} \{ \Gamma(\nu) \}^{-N} \sigma^{-\nu n} \left[\gamma\left(\nu, \frac{x_{(N-n+1)} - v}{\sigma}\right) \right]^{N-n} \\ & \quad \times \left\{ \prod_{j=N-n+1}^N (x_{(j)} - v) \right\}^{\nu-1} \exp\left\{ -\frac{1}{\sigma} \sum_{j=N-n+1}^N (x_{(j)} - v) \right\} \end{aligned}$$

ただし、 $\gamma(\nu, x)$ は不完全ガンマ関数で、

$$\gamma(\nu, x) = \int_0^x t^{\nu-1} e^{-t} dt \quad \dots\dots\dots (11)$$

である。簡単のために標本最小値 $x_{(N-n+1)}$ を x_c と略記し、さらに無次元化した調和平均 R 、相乗平均 S および相加平均 T を次式のように定義しておく。

$$\left. \begin{aligned} R &= (x_c - v) \cdot \frac{1}{n} \sum_{j=N-n+1}^N \frac{1}{x_{(j)} - v} \\ S &= \frac{1}{x_c - v} \left\{ \prod_{j=N-n+1}^N (x_{(j)} - v) \right\}^{1/n} \\ T &= \frac{1}{x_c - v} \cdot \frac{1}{n} \sum_{j=N-n+1}^N (x_{(j)} - v) \end{aligned} \right\} \quad \dots\dots\dots (12)$$

これらの記号を使うと、尤度 L は結局次式のように書ける。

$$\begin{aligned} L &= \frac{N!}{(N-n)!} \{ \Gamma(\nu) \}^{-N} \sigma^{-\nu n} \left[\gamma\left(\nu, \frac{x_c - v}{\sigma}\right) \right]^{N-n} \\ & \quad \times \{ n(x_c - v) S \}^{\nu-1} \exp\left\{ -\left(\frac{x_c - v}{\sigma}\right) T \right\} \quad \dots\dots\dots (13) \end{aligned}$$

一方、標本比 n/N と標本最小値 x_c に対する規準化変量 $\xi_c = (x_c - v)/\sigma$ の間には

$$\begin{aligned} -\frac{n}{N} &= \int_{x_c}^{\infty} g(x) dx \\ &= \int_{\xi_c}^{\infty} \frac{1}{\Gamma(\nu)} t^{\nu-1} e^{-t} dt = \frac{\Gamma(\nu, \xi_c)}{\Gamma(\nu)} \quad \dots\dots\dots (14) \end{aligned}$$

の関係がある。ただし $\Gamma(\nu, \xi)$ は第 2 種不完全ガンマ関数で、第 1 種ガンマ関数との間には次式の関係がある。

$$\Gamma(\nu, \xi) = \Gamma(\nu) - \gamma(\nu, \xi)$$

したがって、未知母数 v, ν および σ を求めるための関係式

$$\partial \log L / \partial v = \partial \log L / \partial \nu = \partial \log L / \partial \sigma = 0$$

によって、結局次式が誘導できる。

$$\left. \begin{aligned} \frac{\xi_c}{\nu-1} - \frac{\xi_c^\nu e^{-\xi_c}}{(\nu-1)\Gamma(\nu, \xi_c)} &= R \\ \exp \left[\frac{d}{d\nu} \left\{ \log_e \left(\frac{\Gamma(\nu, \xi_c)}{\Gamma(\nu)} \right) \right\} + \phi(\nu) \right] / \xi_c &= S \\ \frac{\xi_c^{\nu-1} e^{-\xi_c}}{\Gamma(\nu, \xi_c)} + \frac{\nu}{\xi_c} &= T \end{aligned} \right\} \dots\dots\dots (15)$$

ただし、 $\phi(\nu)$ は *psi* 関数で

$$\phi(\nu) = \Gamma'(\nu) / \Gamma(\nu) \dots\dots\dots (16)$$

で定義される。

したがって、原理的には (15) 式を連立させて、 ν, ν および ξ_c を解けばよいが、未知母数 ν が右辺に陰に含まれるために試行錯誤的な繰返し計算を必要とし、実用的には、つぎのような図式解法が便利である。

まず、前もって ξ_c, ν 面上で R, S および T の曲線群を求めておく。つぎに ν を ν' と仮定し、 $x_{(j)} - \nu'$ を求めると (15) 式の右辺 R, S, T が計算できる。そこで R, S 曲線群の交点の座標として ξ_c, ν の第 1 次近似値 ξ_c', ν' を、他方 S, T より同様に ξ_c'', ν'' を求める。こうしてえられた ξ_c' と ξ_c'' 、あるいは ν' と ν'' は一般に異なっているから、これが一致するまで、すなわち $\xi_c' = \xi_c''$ 、 $\nu' = \nu''$ となるまで ν' を変えて試算を繰返し最終的な収束値、 ν, ξ_c および ν を求める。さらに、 σ の推定値は次式で求められる。

$$\hat{\sigma} = (x_{(N-n+1)} - \nu) / \xi_c \dots\dots\dots (17)$$

この図式解法に必要な ξ_c, ν 面上での R, S および T の曲線群を計算すると、**Figs. 2, 3** および **4** のようになる。なお、この計算に必要な不完全ガンマ関数は K. Pearson の数表⁵⁾ などを利用して求められたものである。

したがって、理論的には、このような試算によって未知母数 ν, ν, σ が求められるが、実際の計算にはかなり労力を要する。そこで、もう少し簡単に解のえられる場合として、原点母数が既知あるいは原点母数を

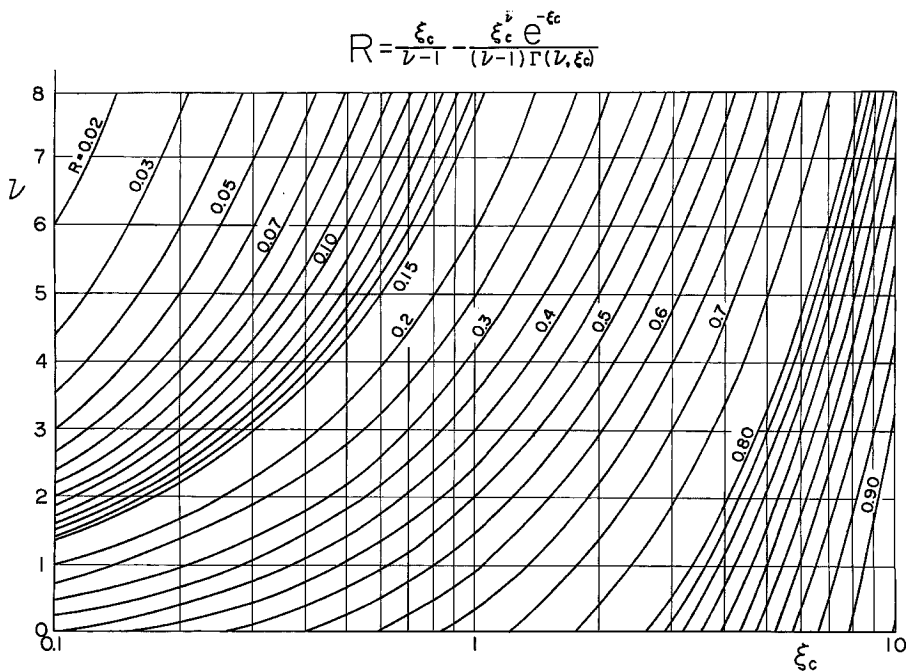
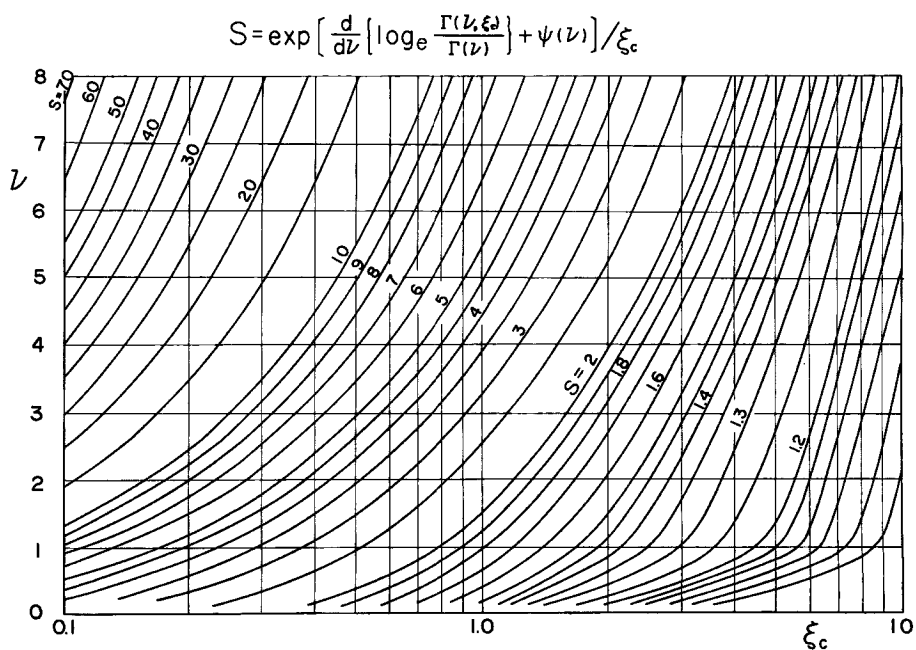
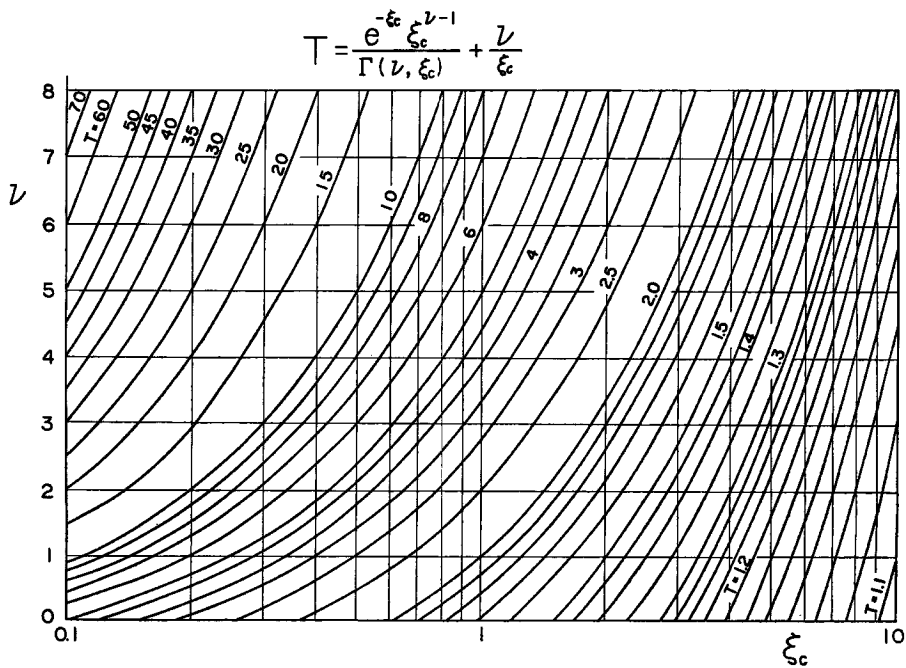


Fig. 2 Function of R .

Fig. 3 Function of S .Fig. 4 Function of T .

0 と仮定した場合についてみて考える。

b. 原点母数が既知の場合

原点母数が既知の場合は、前記の未知の場合において $v=0$ とおいたものを考えればよい。また、最尤解として v に関する微分項は不要であるから、(15) 式における R に関する等式は必要でない。

すなわち、次式

$$S = \frac{1}{x_{(N-n+1)}} \left\{ \prod_{j=N-n+1}^N x_{(j)} \right\}^{1/n}, \quad T = \frac{1}{x_{(N-n+1)}} \frac{1}{n} \sum_{j=N-n+1}^N x_{(j)} \quad \dots\dots\dots (12)'$$

で S, T を求め、(15) 式中の S, T に関する等式

$$\left. \begin{aligned} \exp \left[\frac{d}{d\nu} \left\{ \log_e \left(\frac{\Gamma(\nu, \xi_c)}{\Gamma(\nu)} \right) \right\} + \psi(\nu) \right] / \xi_c &= S \\ \frac{\xi_c^{\nu-1} e^{-\xi_c}}{\Gamma(\nu, \xi_c)} + \frac{\nu}{\xi_c} &= T \end{aligned} \right\} \quad \dots\dots\dots (15)'$$

を連立させて求めた解 $(\hat{\xi}_c, \hat{\nu})$ の $\hat{\nu}$ がまず所要の形状母数の最尤解となる。つぎに、尺度母数 σ の最尤解は、えられた $\hat{\xi}_c$ を (17) 式に代入し、かつ $v=0$ とおくことによって求められる。

ところで、すでに ξ_c, ν 面上における S, T の曲線群は Figs. 3, 4 のようにえられている。そこで、(12)' から S, T を求め、2 曲線 $S=\text{const.}, T=\text{const.}$ の交点の座標として $\hat{\xi}_c, \hat{\nu}$ を図式的に求めることができる。

4.2 その他の母数推定

つぎに残された未知母数、すなわち制約外変量の周辺分布に関する母数および相関母数、の推定を同時標本がすべて既知の場合と部分的にしか知られていない場合に分けて考える。なお、簡単のために、形状母数を等しく ν 、また原点母数はともに 0 と仮定しておく。前記の結果によって σ_1, ν は既知となるから、推定すべき母数は σ_2 と ρ である。

a. 同時標本がすべて既知の場合

まず、 $x_1 (\geq x_{1c})$ に対応する x_2 の同時標本 (x_1, x_2) がすべて既知の場合を考える。密度分布が (3)' 式で与えられる二変数ガンマ分布の場合、指定変数 x_1 に対する従属変数 x_2 の条件付平均値 $E(x_2|x_1)$ は次式の直線となることが知られている⁵⁾。

$$E(x_2|x_1) = \nu\sigma_2 + \rho \frac{\sigma_2}{\sigma_1} (x_1 - \nu\sigma_1) \quad \dots\dots\dots (18)$$

ところで、これを x_1 の関数と考えた x_2 x_2 面内の軌跡、 $x_2 = E(x_2|x_1)$ 、が回帰曲線の定義であるが、それは同時に、 x_1 より x_2 を推定する際の自乗誤差を最小にするものであることが示されている¹⁰⁾。簡単のために、(18) 式による x_2 の推定式を次式のように表現しておく。

$$\left. \begin{aligned} y &= au + b \\ a &= \rho\sigma_2, \quad b = \nu\sigma_2, \quad u = x_1/\sigma_1 - \nu, \quad y = x_2 \end{aligned} \right\} \quad \dots\dots\dots (19)$$

さて、前述したように、上式が回帰直線であることから定数 a, b は存在する同時標本 $(u_i, y_i) (i=1, 2, \dots, n)$ から最小自乗的に求められる。

すなわち、回帰直線による推定の自乗誤差

$$I = \sum_{i=1}^n (y_i - au_i - b)^2$$

が定数 a, b に対して最小になる条件

$$\partial I / \partial a = \partial I / \partial b = 0$$

は次式で与えられる。

$$\hat{a} = \frac{\bar{u}y - \bar{u}\bar{y}}{\bar{u}^2 - (\bar{u})^2}, \quad \hat{b} = \frac{\bar{u}^2\bar{y} - \bar{u}\bar{u}y}{\bar{u}^2 - (\bar{u})^2} \quad \dots\dots\dots (20)$$

したがって、求める σ_2, ρ は次式で与えられる。

$$\left. \begin{aligned} \hat{\sigma}_2 &= \frac{\bar{u}^2 \bar{y} - \bar{u} \bar{u} \bar{y}}{\nu \{u^2 - (\bar{u})^2\}} = \frac{\xi^2 \bar{x}_2 - \xi \bar{\xi} \bar{x}_2 + \nu (\xi \bar{x}_2 - \bar{\xi} \bar{x}_2)}{\nu \{\xi^2 - (\bar{\xi})^2\}} \\ \hat{\rho} &= \frac{\nu (\bar{u} \bar{y} - \bar{u} \bar{u} \bar{y})}{\bar{u}^2 \bar{y} - \bar{u} \bar{u} \bar{y}} = \frac{\nu (\xi \bar{x}_2 - \bar{\xi} \bar{x}_2)}{\xi^2 \bar{x}_2 - \xi \bar{\xi} \bar{x}_2 + \nu (\xi \bar{x}_2 - \bar{\xi} \bar{x}_2)} \end{aligned} \right\} \dots\dots\dots (21)$$

ただし、上式において、 $\xi = x_1/\sigma_1$ であり、また、たとえば $\bar{\xi}$ は n 個の ξ_i ($i=1, 2, \dots, n$) についての平均値を意味している。

b. 同時標本が部分的にしか既知でない場合

さて、 x_1 に対する x_2 の同時標本が部分的にしか分らない場合、たとえば x_2 が x_{2c} 以上というような制約が加わった場合の推定を考えてみる。この場合には厳密な推定は困難であるが、入手しうるすべての標本のもつ確率情報を最大限有効に活用するような推定方法によらざるをえない。そこで、ここでは入手しうるすべての同時標本 (x_{1i}, x_{2i}) ($i=1, 2, \dots, m$) を全標本と考えた最尤解として以下の近似的な推定法が考えられる。

まず、制約付変量を 4.1 で求めた σ_1 を使って、 $\xi = x_1/\sigma_1$ で規準化しておく、 ξ と x_2 の同時分布は、(3)' 式より、次式で表わされる。

$$\begin{aligned} f(\xi, x_2) &= \frac{1}{\Gamma(\nu) \sigma_2^{\frac{\nu+1}{2}} (1-\rho)^{\frac{\nu-1}{2}} \rho^{\frac{\nu-1}{2}}} \exp \left\{ -\frac{\xi}{1-\rho} - \frac{x_2}{\sigma_2(1-\rho)} \right\} \\ &\times (\xi x_2)^{\frac{\nu-1}{2}} \times I_{\nu-1} \left(\frac{2\sqrt{\rho}}{1-\rho} \sqrt{\xi x_2} \right) \dots\dots\dots (22) \end{aligned}$$

したがって、 m 個の標本のもつ尤度は

$$P = \prod_{i=1}^m f(\xi_i, x_{2i})$$

で求められるから、 $\partial \log P / \partial \sigma_2 = \partial \log P / \partial \rho = 0$ とおき、 σ_2 を消去して整理すると、 ρ に関する式として次式がえられる。

$$-\frac{2\rho}{1-\rho} \bar{\xi} = \bar{G}, \quad G(z_i) = \frac{z_i I_\nu(z_i)}{I_{\nu-1}(z_i)} \dots\dots\dots (23)$$

ただし、

$$z_i = \frac{2\sqrt{\rho}}{1-\rho} \sqrt{\xi_i x_{2i}} \left\{ \frac{\rho \bar{\xi} + \nu(1-\rho)}{x_2} \right\}^{\frac{1}{2}} \dots\dots\dots (24)$$

であり、たとえば $\bar{G}$ は $G(z_i)$ ($i=1, 2, \dots, m$) の m 個の平均値を意味する。さらに、上式から ρ の最尤解 $\hat{\rho}$ が求められれば σ_2 の最尤解 $\hat{\sigma}_2$ は次式で求められる。

$$\hat{\sigma}_2 = \frac{x_2}{\hat{\rho} \bar{\xi} + \nu(1-\hat{\rho})} \dots\dots\dots (25)$$

なお、 x_2 についても、たとえば上位から s 個といった標本抽出が可能で、周辺分布に関する母数 σ_2 も既知となる場合には、 x_2 の規準化 $\eta = x_2/\sigma_2$ が可能である。その場合の ρ の最尤解は、(22) で $\sigma_2=1, x_2=\eta$ とおき、 ρ のみについて解くと次式がえられる。

$$2\nu\rho(1-\rho) - 2\rho\bar{\xi} - 2\rho\eta + (1-\rho^2) \bar{G} = 0$$

上式は、未知母数 ρ に関する 2 次方程式であるが、いまの場合正の ρ のみを採用しうから

$$\rho = \frac{-\{\bar{\xi} + \eta - 1\} + \sqrt{(\bar{\xi} + \eta - 1)^2 + \bar{G}(G + 2\nu)}}{\bar{G} + 2\nu} \dots\dots\dots (26)$$

を満足する $\rho = \hat{\rho}$ が最尤解を与える。ただし、この場合

$$z_i = \frac{2\sqrt{\rho}}{1-\rho} \sqrt{\xi_i \eta_i} \quad (i=1, 2, \dots, m) \dots\dots\dots (27)$$

である。

5. 部分標本による二変数指数分布の母数推定

さて、ガンマ分布における重要な分布として指数分布があることは前述したが、以上の部分標本による二

変数ガンマ分布の母数推定において、二変数指数分布に対する結果は $\nu_1 = \nu_2 = 1$ とおくことによって容易に誘導できる。その結果を列挙するとつぎのようになる。なお、記号はすべて前と同じものを使い、また $\nu_2 = 0$ を仮定している。

5.1 制約付変量に関する母数推定

制約付変量 x_1 の周辺分布 $g(x_1)$ が次式

$$g(x_1) = \frac{1}{\sigma_1} \exp\left(-\frac{x_1 - v_1}{\sigma_1}\right) \dots\dots\dots (28)$$

で定義される指数分布の場合、全標本数 N のうちからとった上位 n 個の部分標本による v_1, σ_1 の最尤推定値は次式で求められる。

$$\left. \begin{aligned} v_1 &= x_{1(N-n+1)} - \left(\log \frac{N}{n}\right) \left(\frac{1}{n} \sum_{j=N-n+1}^N x_{1(j)} - x_{1(N-n+1)}\right) \\ \hat{\sigma}_1 &= \frac{1}{n} \sum_{j=N-n+1}^N x_{1(j)} - x_{1(N-n+1)} \end{aligned} \right\} \dots\dots\dots (29)$$

とくに、 $v_1 = 0$ と仮定した場合の σ_1 の最尤推定値は次式のように表わされる。

$$\left. \begin{aligned} \hat{\sigma}_1 &= \frac{1}{n} \sum_{j=N-n+1}^N x_{1(j)} - x_{1(N-n+1)} \\ \text{または} \quad \hat{\sigma}_1 &= \left(\frac{1}{n} \sum_{j=N-n+1}^N x_{1(j)}\right) / \left(1 + \log \frac{N}{n}\right) \end{aligned} \right\} \dots\dots\dots (30)$$

5.2 その他の母数の推定

a. 同時標本がすべて既知の場合

(29) 式によって求められた $v_1, \hat{\sigma}_1$ を使って規準化された変量 $\xi = (x_1 - \hat{v}_1) / \hat{\sigma}_1$ を用いると、 σ_2, ρ は (21) 式より次式で推定される。

$$\left. \begin{aligned} \hat{\sigma}_2 &= \frac{\bar{\xi}^2 x_2 - \bar{\xi} \bar{\xi} x_2 + \bar{\xi} x_2 - \bar{\xi} x_2}{\bar{\xi}^2 - (\bar{\xi})^2} \\ \hat{\rho} &= \frac{\bar{\xi} x_2 - \bar{\xi} x_2}{\bar{\xi}^2 x_2 - \bar{\xi} \bar{\xi} x_2 + \bar{\xi} x_2 - \bar{\xi} x_2} \end{aligned} \right\} \dots\dots\dots (31)$$

b. 同時標本が部分的にしか既知でない場合

部分的な同時標本 (ξ_i, x_{2i}) ($i=1, 2, \dots, m$) を全標本とみなした ρ の最尤解 $\hat{\rho}$ は、(23) 式を満足するものとして与えられる。ただし、(24) 式は次式で表わされる。

$$z_i = \frac{2\sqrt{\rho}}{1-\rho} \sqrt{\xi_i x_{2i}} \left\{ \frac{\rho \bar{\xi} + 1 - \rho}{x_2} \right\}^{\frac{1}{2}} \dots\dots\dots (32)$$

また、この $\hat{\rho}$ を用いて、 σ_2 の最尤解は次式で求められる。

$$\hat{\sigma}_2 = \frac{\bar{x}_2}{\hat{\rho} \bar{\xi} + 1 - \hat{\rho}} \dots\dots\dots (33)$$

なお、 x_2 についても規準化 $\eta = x_2 / \sigma_2$ が可能な場合、 ρ 最尤解 $\hat{\rho}$ は次式を満足する ρ として与えられる。

$$\rho = \frac{-(\bar{\xi} + \bar{\eta} - 1) + \sqrt{(\bar{\xi} + \bar{\eta} - 1)^2 + \bar{G}(\bar{G} + 2)}}{\bar{G} + 2} \dots\dots\dots (34)$$

ただし、上式中の $G = G(z)$ および z はそれぞれ (23), (27) 式で定義されるものである。

6. 河川流出量への適用

最後に、二変数ガンマ分布の水文量分布への有用性を調べた一つの実例を示しておく。全標本を対象とし、簡単のために、原点母数を 0 としておく。二変数として紀の川上流吉野川の寺尾地点（流域面積 243 km²）およびそれより約 6 km 離れた支川高見川の鷺家口地点（122 km²）の T 日間流出量 x_1 および x_2 を考える。この分布を、形状母数の異なる二変数ガンマ分布と仮定し、積率解としての母数推定を行なった。標本とし

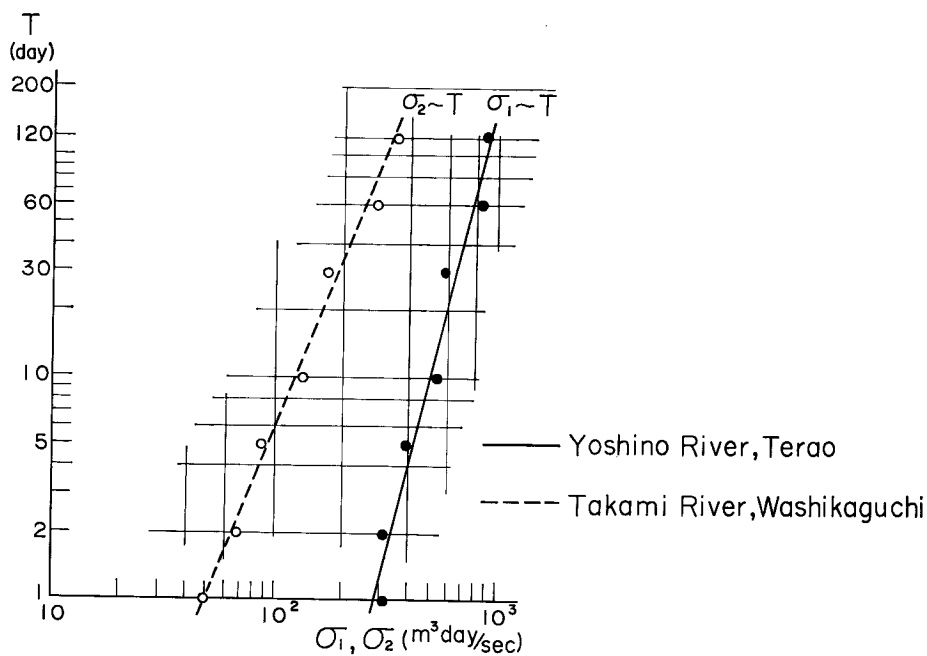


Fig. 5 Relations between scale parameter, σ , and sample duration, T , in average river flows.

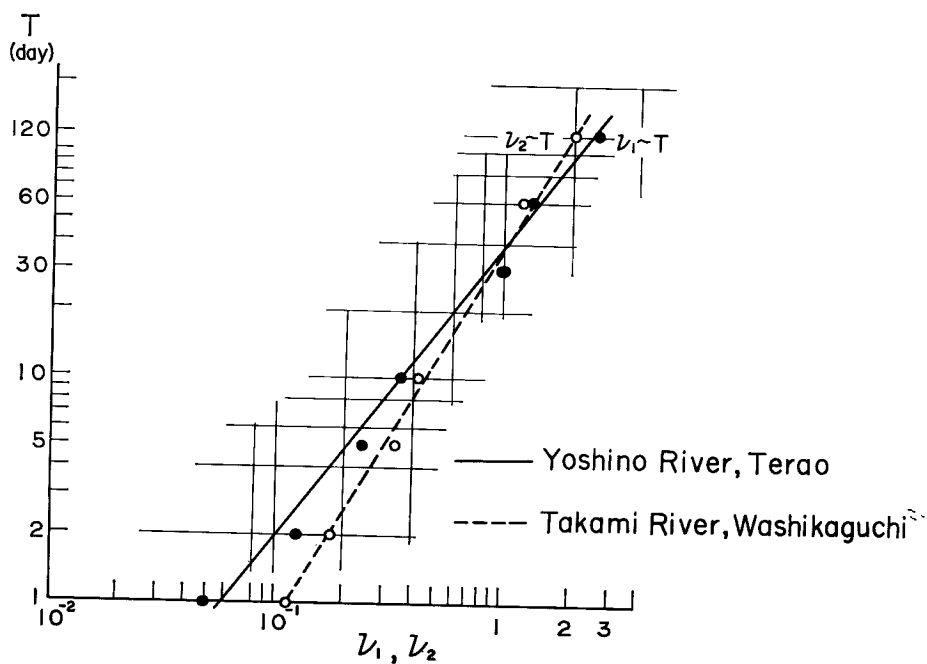


Fig. 6 Relations between shape parameter, ν , and sample duration, T , in average river flows.

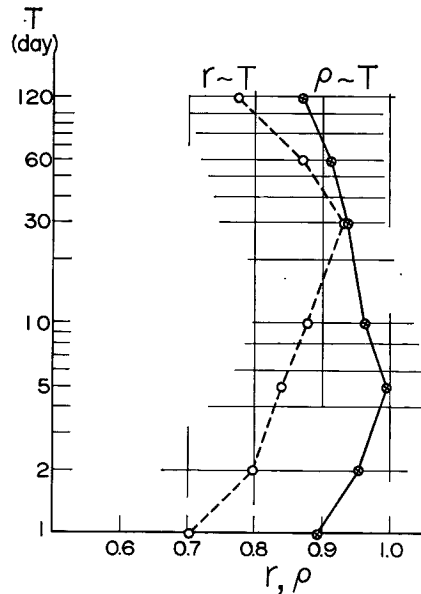


Fig. 7 Relation among correlation parameter, ρ , correlation coefficient, r , and sample duration, T , in average river flows at two station points.

て、昭和 36~38 年の 3 年間で、期間を重複しないように $T=1, 2, 5, 10, 30, 60, 120$ 日にとった流出量を $\text{m}^3 \cdot \text{day} / \text{sec}$ 単位で用いている。

まず、Fig. 5 は尺度母数 σ_1, σ_2 と標本期間 T の関係を示したものである。図中の実線および点線は、それぞれのプロット点に適合するように引いた直線で、近似的に $\sigma \sim T$ の関係を表わしているといえる。また、Fig. 6 は形状母数 ν_1, ν_2 と T の関係を示したものである。図で、 T がほぼ 30 日以下では ν が 1 以下となるから、 x_1, x_2 の周辺分布は、指数分布あるいはそれより歪んだ逆 J 字型の単調減少の分布型を示すことが分かる。したがって、このような場合には、単純な対数変換などにより変数の正規化は困難であり、直接ガンマ分布や指数分布として取扱った方が望ましいといえよう。

さらに Fig. 7 は相関母数 ρ 、標本相関係数 r と T の関係を示したものである。まず、いまの場合、二流域が近接しているために、流出量間の相関が大きいのは当然である。また、 T が 2 日、5 日と増すにつれて ρ, r は一度増加するが、さらに T が 60 日、120 日と増すと ρ, r は減少している。その理由として、いまの場合わずか 3 年間の資料で計算しているために、 T がかなり大きくなると標本数が極端に減少し、その影響が ρ, r の減少となって表われているためであろうと考えられる。したがって、相関母数の的確な推定には、それに相応した標本数が要求されるといえよう。

7. む す び

以上、二変数ガンマ分布の母数推定について理論的に考察したが、結果を要約するとつぎのとおりである。

1) 全標本による母数推定についての前報の成果を、形状母数の異なる二変数ガンマ分布に拡張して積率解を誘導し、とくに相関母数の積率解と通常の相関係数との関係を明らかにした。

2) 部分標本として片方の変量に関する上位標本を対象とし、母数推定の手順を、まず制約付変量の周辺分布に関する推定、ついで残りの母数、すなわち制約外変量の周辺分布に関する母数および相関母数、の推定に分けてつぎのように行なう方法を提示した。

2.1) 制約付変量の周辺分布に関する母数推定では、順序統計の概念を用いた censored sampling theory によって最尤解を求め、さらに、実用計算上の便宜を考え、必要な図を作製した。

2.2) 残された母数推定については、個々の制約付変量に対する他の変量の関係がすべて既知である場合 (a) と、部分的にしか既知でない場合 (b) に区別して考察した。

(a) この場合には、一般に二変数ガンマ分布の回帰曲線が直線となることを用いて、母数を最小自乗的に推定する方法を示した。

(b) この場合には、厳密な推定は困難であるが、部分標本を全標本とみなした最尤解による近似的な推定方法を示した。

3) 形状母数の異なる二変数ガンマ分布の適用例として、二地点での種々の期間における河川流出量の同時分布について、全標本による積率解としての母数推定を行ない、理論の有用性を示した。

もちろん、このような全標本あるいは部分標本によって推定された母数には、実際上何らかの推定誤差が含まれているはずで、この方面からの検討が必要である。著者は、乱教を使った電子計算機による数値実験的な手法で母数が既知の母集団を対象として、母集団から抽出された標本から種々の方法で母数を推定し、その結果を比較・検討するという形でこの問題を研究中であり、一部興味ある結果をえているが、これについては別の機会に発表したいと考えている。

参 考 文 献

- 1) 井沢竜夫：二変数の Γ -分布について (続)，気象と統計，第 4 巻，第 2 号，昭 28, pp. 15-19.
- 2) たとえば森口繁一他：数学公式Ⅲ，——特殊関数——，岩波全書，昭 33, p. 200.
- 3) Gradshteyn, I. S. and I. M. Ryzhik: Table of Integrals, Series, and Products, Academic Press, New York and London, 1965, p. 860.
- 4) 前出 3) の p. 1243.
- 5) 前出 3) の p. 286.
- 6) 長尾正志・角屋 睦：二変数ガンマ分布とその適用に関する研究(1)——二変数指数分布の基礎理論——，京都大学防災研究所年報，第 13 号 B，昭 45. 3, pp. 105-133.
- 7) Gupta, A. K.: Estimation of the Mean and Standard Deviation of Normal Population from a Censored Sample, Biometrika, Vol. 39, 1952, pp. 260-273.
- 8) Wilk, M. B., R. Gnanadesikan and M. J. Huyett: Estimation of the Gamma Distribution Using Order Statistics, Biometrika, Vol. 49, Parts 3 and 4, 1962, pp. 525-535.
- 9) 井沢竜夫：二変数の Γ -分布について (降水量の分布第 2 報)，気象と統計，第 4 巻，第 1 号，p. 12.
- 10) たとえば，Cramér, H.: The Elements of Probability Theory, John Wiley & Sons, New York, 1961, p. 138.