

水文統計における多変数確率分布理論

—— 二変数統計を中心として ——

長 尾 正 志

1. 二変数の確率分布とその応用

1.1 多変数統計と水工計画

治水・利水を総合して一般に水工計画では、複数個の水文量を同時に勘案すべき場合が多く、必然的に多変数統計の手法が要求される。しかしながら、現実の解析では、従来の計画降雨量や計画高水流量の策定で代表されるように、単変数統計としての解析が大部分であった。また、たとえ多変数統計手法が適用されたとしても、変量間の相関を無視した独立事象と仮定して、いわば単変数手法の組み合わせとしての適用に過ぎなかったようである。

その主な理由には、根底となる多変数統計の理論および応用手法に関する研究が、特殊な場合を除いて、今なお不十分であることが挙げられる。すなわち、正規分布を除けば、一般的な多変数統計理論の展開はほとんど未解明といって過言でない。しかも、水文量は、一般的にいて、原形のままで正規分布に従うことはむしろまれといわれる。しかし、近年、正規分布以外の多変数理論も漸次研究が進展し、ある程度までは理論的取り扱いが可能になりつつある。⁽¹⁾

さて、水工計画における多変数統計の応用を大きく類別すると以下の三つに分けられる。

イ. 多変量からなる結合生起確率の評価

ロ. 指定変量による条件付変量の推定

ハ. 多変量群から性格の共通した数群への判別

すなわち、イ.では、計画基本量としての水文量が複数個存在し、これらが計画に対して好都合や不都合な状態に判別できる場合、各状態に対する諸量の結合生起の可能性を評価するような事例である。たとえば、堤防による治水計画では、河道を流下する洪水のピーク時の水位、流量、流速などが破堤・溢流に関係し、これら諸量の結合確率が問題とされる。

つぎに、ロ.では、1個ないし複数個の変量が指定された場合に、問題とされる1個ないし複数個の変量の大きさやその出現の可能性を推定する事例である。正規分布の場合の回帰直線の利用などがその好例であり、近年、注目されつつある水文資料の模擬発生はこの問題の応用面に関連する。

さらに、ハ.では、多数の観測値群を比較的性格の共通した小数の群に分類するといった事例で、たとえば、流域内の多数の雨量観測所の観測値を整理・統合するといった場合であり、多変量正規母集団の仮定に基づく多変量解析論の応用が考えられている。

なお、多変数統計と単変数統計との基本的な相異は、相関母数の有無である。しかも、二変数の場合には、相関母数は、たとえば標本相関係数といった単変量による推定ができるが、三変数以上になれば、重相関係数や偏相関係数といった複数量による表現が必要である。したがって、多変数分布を記述する母数の個数が、変量の次元が増加するに伴い飛躍的に増すとともに、各周辺分布のみならず、複雑な条件付分布を取り扱わねばならないことが解析を非常に複雑にさせる原因である。

そこで、以下では、上記イ.とロ.を中心としてその基礎と考えられる二変数理論とその水工計画への適用を主体として、代表的な連続分布である正規分布とガンマ分布について具体的に説明する。

1.2 二変数正規分布とその応用

この分布の最初の統計的取り扱い、GaltonとDicksonによるといわれるが、正規分布以外の多変数統計の理論もほとんどこの分布を基礎として派生したとみてよい。したがって、その理論的諸特性もかなり研究されており、純粋数学的には大多数の問題が解明されているといっても過言ではない。ただ実際の応用に際して留意すべき問題が若干残されているが、詳細は後述することにし、ここでは、以後の考察に必要な基礎的特性についてとりまとめておく。

a. 基礎式

(i) 分布関数

二変数 x_1, x_2 に関する二変数正規分布の密度分布は次式で与えられる。

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left\{ \left(\frac{x_1-m_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x_1-m_1}{\sigma_1} \right) \left(\frac{x_2-m_2}{\sigma_2} \right) + \left(\frac{x_2-m_2}{\sigma_2} \right)^2 \right\} \right\} \quad \dots\dots\dots(1)$$

ここで、 $m_1, m_2; \sigma_1^2, \sigma_2^2; \rho$ はそれぞれ x_1, x_2 の平均値、分散、相関係数に対応した定数であることが知られている。これより、 x_i ($i=1, 2$) の周辺分布、および x_j を与えた場合の x_i の条件分布 ($i \neq j; i, j=1, 2$) は、それぞれ以下のような正規分布となる。

$$f_i(x_i) \equiv \int_{-\infty}^{\infty} f(x_i, x_j) dx_j = \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left\{ -\frac{(x_i-m_i)^2}{2\sigma_i^2} \right\} \quad \dots\dots\dots(2)$$

$$f(x_i | x_j) \equiv f(x_1, x_2) / f_j(x_j) \\ = \frac{1}{\sqrt{2\pi}\sigma_i\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2\sigma_i^2(1-\rho^2)} \left\{ x_i - m_i - \rho \frac{\sigma_i}{\sigma_j} (x_j - m_j) \right\}^2 \right\} \quad \dots\dots\dots(3)$$

条件付分布も周辺分布と同じ形状になるという上記特性は、ほとんど正規分布のみに認められる応用上有用な性質である。すなわち、その規準化表示により、数表などを利用した水文資料の人為的作製などに活用される。

なお、以上の分布関数に含まれる未知定数(母数)を推定するには、普通、積率解や最尤解が用いられるが、これらは大抵の類書に詳しいので、ここでは省略し、その結果を利用した応用面を中心として記述する。

(ii) 線型結合和の分布

二変数 x_1, x_2 が式(1)の密度分布に従うとき、線型結合和 $z = Ax_1 + Bx_2 + D$ (A, B, D : 定数) の分布を示す。最終結果の密度分布は以下のとおりとなる。

$$f(z) = (a/\sqrt{1+b^2+2\rho b}) \cdot \varphi((az-d)/\sqrt{1+b^2+2\rho b}) \quad \dots\dots\dots(4)$$

ただし、関数 $\varphi(x)$ は規準正規分布の密度分布で次式

$$\varphi(x) = (1/\sqrt{2\pi}) \exp(-x^2/2) \quad \dots\dots\dots(5)$$

で定義される。また、定数 a, b, d は次式で与えられる。

$$a \equiv (A\sigma_1)^{-1}, \quad b \equiv (B\sigma_2)/(A\sigma_1), \quad d \equiv (Am_1 + Bm_2 + D)/(A\sigma_1) \quad \dots\dots\dots(6)$$

結局、相関の存在する場合の線型結合和 z も正規分布に従い、 z の平均値には線型変換の形が、また分散には変換の形のみならず相関係数の大きさが関係することが分る。たとえば、 $A, B > 0$ の場合、相関係数 ρ が $+1$ から -1 に減ずるにつれて分散も減少することになる。

その傾向を、規準化変量

$$\xi = (x_1 - m_1) / \sigma_1, \quad \eta = (x_2 - m_2) / \sigma_2 \quad \cdots \cdots (7)$$

の和 z の累積分布

$$\begin{aligned} F(z) &= \Phi\left(\frac{z}{\sqrt{2(1+\rho)}}\right) \\ &= \frac{1}{2\sqrt{\pi(1+\rho)}} \int_{-\infty}^z \exp\left\{-\frac{t^2}{4(1+\rho)}\right\} dt \quad \cdots \cdots (8) \end{aligned}$$

として、図-1に示す。ただし、 $\Phi(x)$ は規準正規分布 $\varphi(x)$ の累積分布である。

b. 確率積分

二変数 x_1, x_2 の結合密度分布 $f(x_1, x_2)$ が既知の際、 x_1, x_2 面内の部分領域 D の内部に変数が含まれる確率 $P(D)$ は、

$$P(D) = \iint_{(x_1, x_2) \in D} f(x_1, x_2) dx_1 dx_2$$

で計算されるが、それには普通煩雑な数値積分の演算を要する。しかし、最近、式(7)の規準化変量 ξ, η による二変数正規分布の表示

$$f(\xi, \eta; \rho) = (1/2\pi\sqrt{1-\rho^2}) \exp\left[-(\xi^2 - 2\rho\xi\eta + \eta^2)/2(1-\rho^2)\right] \quad \cdots \cdots (9)$$

で $\xi \geq h, \eta \geq k$ となる上側確率

$$U(h, k; \rho) \equiv \int_h^\infty \int_k^\infty f(\xi, \eta; \rho) d\xi d\eta \quad \cdots \cdots (10)$$

およびそれに関連した数表が発表されたので、計算はかなり容易になったといえよう。

すなわち、各種の ρ, h, k に対する上側確率 U の値、および ξ, η が独立の場合に次式で定義される V 関数

$$V(h, a) \equiv \int_0^h \varphi(\xi) \int_0^a \varphi(\eta) d\xi d\eta$$

が各種の h, a について数表となっている。ただし、 U 関数と V 関数の関係は、

$$\begin{aligned} U(h, k; \rho) &\equiv V(h, (h/k - \rho)/\sqrt{1-\rho^2}) + V(k, (h/k - \rho)/\sqrt{1-\rho^2}) \\ &\quad + \sin^{-1} \rho / 2\pi + \left[\Phi(h) + \Phi(k) \right] / 2 - 1/4 \end{aligned}$$

また、 φ, Φ はそれぞれ規準正規分布の密度分布、累積分布である。

さて、このような数表および単変数の理論を活用すれば、計画意図に適、不適の状態に対応する確率

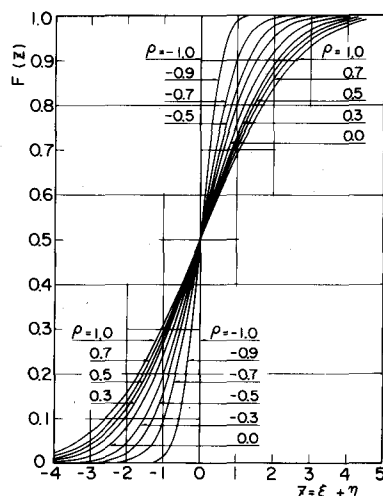


図-1 二変数正規分布における規準化変量の和の累積分布

の数値積分がかなりの精度で可能となる。たとえば、上流に貯水池の存在する河道を含む水系の治水計画では、計画基本量として、河道ではその区間を流下する洪水のピーク流量、貯水池では流入洪水の総流量を採用する場合に、各変量の正規化変換を行なった後、計画の危険側確率を評価する⁽⁸⁾といった利用の仕方は多々あると思われる。

c. 回 帰 特 性

二変数正規分布の重要な応用として、回帰特性の利用がある。その内容も、初期には欠測部の資料補填などに限定されていたが、今日では数多い水文学相互間の空間的関連性や時系列的な因果関係の究明など非常に広範囲に及んでいる。

一般に y の x への回帰曲線とは、 x が所与の際の y の推定値 $\hat{y}$ を、条件付平均値 $E(y|x)$ の軌跡として、次式のように定義するものである。

$$y = \hat{y} = E(y|x) \equiv \int_{-\infty}^{\infty} y \cdot f(y|x) dy \quad \cdots \cdots \cdots (11)$$

とくに、 x と上記 $y = \hat{y}$ との関係が次式の直線

$$y = \alpha + \beta(x - \bar{x}), \quad (\alpha, \beta: \text{定数}, \bar{x}: x \text{ の平均値}) \quad \cdots \cdots \cdots (12)$$

となる場合を線型回帰と称し、取り扱いの簡単のため多用される。

一方、 x から y を推定するのに、定数 $a, b, c, \cdots$ を含んだ関数 $y = g(x; a, b, c, \cdots)$ を用いたとする。その推定精度を二乗誤差で評価し、最小二乗誤差を与えるような定数の推定値 $\hat{a}, \hat{b}, \hat{c}, \cdots$ を採用した関数 $\hat{y} = g(x; \hat{a}, \hat{b}, \cdots)$ と回帰曲線の定義式とは完全に合致することが知られており、このことが回帰曲線の調法される根拠となっている。

さて、 x, y が二変数正規分布に従う場合、 x を与えたときの y の条件付平均値 $E(y|x)$ は、条件付分布の式(3)より、次式の直線で表わされる。

$$E(y|x) = m_y + \rho(\sigma_y/\sigma_x)(x - m_x) \quad \cdots \cdots \cdots (13)$$

したがって、この場合、回帰曲線は一般に回帰直線となる。しかも、回帰曲線からの変動性を表現する条件付変量の標準偏差 $D(y|x)$ は、式(3)より、次式

$$D(y|x) = \sigma_y \sqrt{1 - \rho^2} \quad \cdots \cdots \cdots (14)$$

のように指定変数 x に無関係な定数となる。等分散性と称されるこの正規分布の基本的性質は、資料の模擬発生などに二変数正規モデルを採用する際の簡便性への根拠を与える。

なお、回帰曲線の標本による推定式を示しておく。まず式(12)で最小二乗誤差基準により、 α, β の推定値を求めるが、それには、標本に対する x, y の共分散、 x の分散をそれぞれ

$$\mu_{xy} = \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) / N, \quad S_x^2 = \sum_{i=1}^N (x_i - \bar{x})^2 / N$$

と記すことにより、次式で与えられる。

$$\hat{\alpha} = \bar{y}, \quad \hat{\beta} = \mu_{xy} / S_x^2 \quad \cdots \cdots \cdots (15)$$

一方、標本相関係数 r の定義は

$$r \equiv \mu_{xy} / S_x \cdot S_y \quad (S_y: y \text{ の標本標準偏差})$$

であるから、二変数正規分布だけでなく、一般に回帰曲線が回帰直線で表わされる場合には、回帰直線は次式で推定できることになる。

$$y = \bar{y} + r(S_y / S_x)(x - \bar{x}) \quad \cdots \cdots \cdots (16)$$

1.3 二変数ガンマ分布とその応用

a. 二変数ガンマ分布

水文統計の立場から、正規分布に次いで重要な応用面をもつ連続分布はガンマ分布で、これは密度分布形における非対称性を基礎にしたものといえる。この分布は、母数の選択に応じて、正規分布に近い形から、指数分布などの逆J字型のような非対称分布に至る非常に広範囲の形を網羅することから、気象統計、人口統計などで極めて有効と目されている。たとえば、各種継続時間に対する降水量分布がこれに従うことを示した報告も多い。

しかし、変量相互間の相関を考慮した二変数あるいは多変数ガンマ分布に関する研究は、あまり顕著ではない。⁽⁴⁾ 著者の知る範囲では、Wicksellによる特性関数の誘導⁽⁵⁾がその始まりであるが、本格的には、ガンマ型の二変数分布の誘導およびLaguerre多項式や変形ベッセル関数による表示を行なったKibble⁽⁶⁾の研究、さらに、一般的な多変数への拡張を試みたKrishnamoorthyとParthasarathyの研究⁽⁷⁾、わが国では気象統計の立場から二変数ガンマ分布の基本的特性を考察した井沢の研究^{(8), (9)}が注目される程度であろう。しかも、以上はいずれも純粋数学的な興味に基づいた研究で、実際問題への適用という観点からは極めて不満足な段階に留まっているように思える。ここでは、実用上必要な二変数の理論および応用に際して重要な基礎事項を、正規分布と対比しながら、説明しておく。

(i) 基礎式

1. 分布関数

二変数 x_1, x_2 がそれぞれ形状母数 ν_1, ν_2 ($\nu_1 \geq \nu_2$)、尺度母数 σ_1, σ_2 および原点母数がともに0のガンマ分布に従う場合、 x_1, x_2 の結合生起の密度分布は、井沢により以下のように示されている。⁽¹⁰⁾ ただし、 $\nu_1 = n + m, \nu_2 = n$ ($n, m \geq 0$) としておく。

$$f(x_1, x_2) = \frac{1}{\Gamma(n) \Gamma(m) (\sigma_1 \sigma_2)^{\frac{n+1}{2}} \sigma_1^m (1-\rho)^{\frac{n-1}{2}} \rho^{\frac{n-1}{2}}} \times (x_1 x_2)^{\frac{n-1}{2}} x_1^m \\ \times \exp \left\{ -\frac{x_1}{\sigma_1(1-\rho)} - \frac{x_2}{\sigma_2(1-\rho)} \right\} \int_0^1 (1-t)^{\frac{n-1}{2}} t^{m-1} \exp \left\{ \frac{\rho x_1 t}{\sigma_1(1-\rho)} \right\} \\ \times I_{n-1} \left(\frac{2\sqrt{\rho}}{1-\rho} \sqrt{\frac{x_1 x_2}{\sigma_1 \sigma_2} (1-t)} \right) dt \quad \dots\dots\dots (17)$$

式中の $\Gamma(\nu)$ 、 $I_\nu(z)$ はそれぞれ引数 ν のガンマ関数、 ν 次の第1種変形ベッセル関数である。なお、 ρ は相関に関係する定数(相関母数)で、定義より明らかなように、 $0 \leq \rho \leq 1$ である。

上式より、 x_i ($i=1, 2$) の周辺分布の密度分布、累積分布、および x_2 が所与の場合の x_1 の条件付分布は、それぞれ次式のようになる。

$$\left. \begin{aligned} f(x_i) &= \frac{1}{\Gamma(\nu_i) \sigma_i^{\nu_i}} x_i^{\nu_i-1} \exp \left(-\frac{x_i}{\sigma_i} \right), \\ F(x_i) &= \frac{1}{\Gamma(\nu_i)} \gamma(\nu_i, x_i / \sigma_i) \end{aligned} \right\} \quad (i=1, 2) \quad \dots\dots\dots (18)$$

$$f(x_1 | x_2) = \{ \Gamma(m) \}^{-1} \sigma_1^{-\frac{n+1}{2}-m} \sigma_2^{-\frac{n-1}{2}} (1-\rho)^{-1} \rho^{-\frac{n-1}{2}} \exp \left\{ -\frac{x_1}{\sigma_1(1-\rho)} - \frac{\rho x_2}{\sigma_2(1-\rho)} \right\} \\ \times x_1^{\frac{n-1}{2}+m} x_2^{-\frac{n-1}{2}} \int_0^1 (1-t)^{\frac{n-1}{2}} t^{m-1} \exp \left\{ \frac{\rho x_1 t}{\sigma_1(1-\rho)} \right\} \\ \times I_{n-1} \left(\frac{2\sqrt{\rho}}{1-\rho} \sqrt{\frac{x_1 x_2}{\sigma_1 \sigma_2}} (1-t) \right) dt \quad \dots\dots\dots (19)$$

ここに、 $\gamma(\nu, x)$ は第1種不完全ガンマ関数である。条件付分布はもはやガンマ分布に従わず、指定変数によって形状を変化させる。このような傾向は通常の変数水文量に関してよく経験される。

ロ. 分布特性値

ここでは、積率の誘導結果と分布特性値を示す。まず、二変数 x_1, x_2 について原点のまわりの p, q 次の積率は以下のように求められる。

$$\nu_{pq} \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1)^p (x_2)^q f(x_1, x_2) dx_1 dx_2 \\ = \frac{\Gamma(n+m+p) \Gamma(n+q) (1-\rho)^{n+q} \sigma_1^p \sigma_2^q}{\{ \Gamma(n) \}^2 \Gamma(m)} \\ \times \int_0^1 x^{n-1} (1-x)^{m-1} (1-\rho x)^{-n-m-q} \cdot F(-m-p, -q; n; \rho x) dx \quad \dots\dots\dots (20)$$

ただし、 $F(a, b; c; z)$ は超幾何関数である。したがって、具体的に4次以下の積率を求めると次式となる。

$$\left. \begin{aligned} \nu_{10} &= \nu_1 \sigma_1, & \nu_{01} &= \nu_2 \sigma_2 \\ \nu_{20} &= \nu_1(\nu_1+1) \sigma_1^2, & \nu_{02} &= \nu_2(\nu_2+1) \sigma_2^2 \\ \nu_{11} &= \nu_2(\nu_1+\rho) \sigma_1 \sigma_2, & & \\ \nu_{30} &= \nu_1(\nu_1+1)(\nu_1+2) \sigma_1^3, & \nu_{03} &= \nu_2(\nu_2+1)(\nu_2+2) \sigma_2^3 \\ \nu_{40} &= \nu_1(\nu_1+1)(\nu_1+2)(\nu_1+3) \sigma_1^4, & \nu_{04} &= \nu_2(\nu_2+1)(\nu_2+2)(\nu_2+3) \sigma_2^4 \end{aligned} \right\} \dots\dots\dots (21)$$

これより、主要な分布特性値は以下のように求められる。

$$\left. \begin{aligned} \text{平均値} \quad m_1 &\equiv \nu_{10} = \nu_1 \sigma_1, & m_2 &\equiv \nu_{01} = \nu_2 \sigma_2 \\ \text{分散} \quad D_1^2 &\equiv \nu_{20} - \nu_1^2 = \nu_1 \sigma_1^2, & D_2^2 &\equiv \nu_{02} - \nu_2^2 = \nu_2 \sigma_2^2 \\ \text{相関係数} \quad r &\equiv (\nu_{11} - \nu_{10} \nu_{01}) / D_1 D_2 = \sqrt{\nu_2 / \nu_1} \cdot \rho \\ \text{ひずみ係数} \quad \gamma_{10} &= 2 / \sqrt{\nu_1}, & \gamma_{01} &= 2 / \sqrt{\nu_2} \\ \text{過剰係数} \quad \gamma_{20} &= 6 / \nu_1, & \gamma_{02} &= 6 / \nu_2 \end{aligned} \right\} \dots\dots\dots (22)$$

ハ. 回帰特性値

回帰特性に関する条件付変量の平均値、分散を、特性関数を媒介として誘導する。まず、条件付変量 $(x_1 | x_2)$ に対する特性関数は次式で求められる。

$$Q(z) \equiv \int_0^{\infty} \exp(izx_1) f(x_1 | x_2) dx_1 \\ = (1 - i\sigma_1 z)^{-m} \left\{ 1 - i\sigma_1(1-\rho)z \right\}^{-n} \exp \left[-\frac{\rho}{\sigma_2(1-\rho)} \left\{ 1 - \frac{1}{1 - i\sigma_1(1-\rho)z} \right\} x_2 \right] \quad (23)$$

したがって、この結果より、条件付平均値は、

$$E(x_1 | x_2) = \nu_1 \sigma_1 + \rho(\sigma_1 / \sigma_2)(x_2 - \nu_2 \sigma_2) \quad \dots\dots\dots(24)$$

となるから、 x_2 に対する x_1 の回帰曲線は直線となる。同様に、 x_2 に対する x_1 の条件付分散を求めると、次式となる。

$$\begin{aligned} D^2(x_1 | x_2) &\equiv E\{(x_1 | x_2)^2\} - \{E(x_1 | x_2)\}^2 \\ &= \{(\nu_1 - \nu_2) + \nu_2(1 - \rho)^2\} \sigma_1^2 + 2(\sigma_1^2 / \sigma_2) \rho(1 - \rho)x_2 \quad \dots\dots\dots(25) \end{aligned}$$

すなわち、二変数正規分布では、条件付分散が指定変数に関係しないという等分散性が成立したのに対して、二変数ガンマ分布では、条件付分散は指定変数の増加に伴って線型的に増加することが顕著な相違である。

(ii) ガンマ型変量の和分布

二変量に関する母数が異なる場合の解は誘導が困難で、以下では、形状母数、尺度母数がそれぞれ共通な場合の分布特性を考察する。前出の式(17)で、 $\nu_1 = \nu_2 = \nu$ 、 $\sigma_1 = \sigma_2 = \sigma$ の場合、変量の和 $z = x_1 + x_2$ の密度分布 $f_\nu(z)$ は、井沢により次式で表わされる。

$$f_\nu(z) = \frac{\pi^{\frac{1}{2}}}{\Gamma(\nu) \sigma^{\nu+\frac{1}{2}} (1-\rho)^{\frac{1}{2}} \rho^{\frac{2\nu-1}{4}}} \left(\frac{z}{2}\right)^{\nu-\frac{1}{2}} \exp\left\{-\frac{z}{\sigma(1-\rho)}\right\} \cdot I_{\nu-\frac{1}{2}}\left(\frac{\sqrt{\rho}}{\sigma(1-\rho)} z\right) \quad (26)$$

さて、水工計画では変量の出現の可能性を知る必要上超過確率が、したがって累積分布が問題とされることが多い。そこで、 ν が自然数および半奇数の場合の累積分布を、もとの累積分布と比較するために、式(18)のような第1種不完全ガンマ関数 $\gamma(\nu, x)$ を使って、表示しておく。

まず ν が自然数 $n (=1, 2, \dots)$ の場合、累積分布 $F(z)$ は、次式のように有限級数で表示できる。

$$F(z) = F_n(z) \equiv \int_0^z f_n(z) dz = \sum_{k=0}^{n-1} H_k(z) \quad \dots\dots\dots(27)$$

ただし、

$$H_k(z) \equiv a_k \cdot \gamma(n-k, z/\sigma(1+\sqrt{\rho})) + a_{-k} \cdot \gamma(n-k, z/\sigma(1-\sqrt{\rho})) \quad \dots\dots\dots(28)$$

また、 a_k 、 a_{-k} は次式のような定数である。

$$\left. \begin{aligned} a_k &= (-1)^k \frac{(n+k-1)!}{(n-1)! k! (n-1-k)!} \frac{(1-\rho)^k (1+\sqrt{\rho})^{n-k}}{(2\sqrt{\rho})^{n+k}} \\ a_{-k} &= (-1)^n \frac{(n+k-1)!}{(n-1)! k! (n-1-k)!} \frac{(1-\rho)^k (1-\sqrt{\rho})^{n-k}}{(2\sqrt{\rho})^{n+k}} \end{aligned} \right\} \quad \dots\dots\dots(29)$$

ついで、 ν が半奇数の場合 ($\nu = n + 1/2$; $n = 0, 1, 2, \dots$) には、 $F(z)$ はつぎのような無限級数で表示できる。

$$F_{n+\frac{1}{2}}(z) = \frac{n!}{(2n)!} (1-\rho)^{n+\frac{1}{2}} \sum_{k=0}^{\infty} \frac{\rho^k}{k! (n+k)! 2^k} \cdot \gamma\left(2n+2k+1, \frac{z}{\sigma(1-\rho)}\right) \quad (30)$$

いずれの場合も、累積分布は $\gamma(\nu, x)$ の線型結合として表現されることになる。

すなわち、正規分布では各成分変量と和が同じ分布族に従う、いわゆる再生性が成立する、のに対して、ガンマ分布ではもはや再生性は厳密には成立しない。しかし、以下の特別な相関母数の場合には、

和の累積分布の形より再生性が成り立つことが誘導できる。

$$\left. \begin{aligned} \rho \rightarrow 0 \text{ (完全独立)} : F_\nu(z) &= \left\{ \Gamma(2\nu) \right\}^{-1} \cdot \gamma(2\nu, z/\sigma) \\ \rho \rightarrow 1 \text{ (完全従属)} : F_\nu(z) &= \left\{ \Gamma(\nu) \right\}^{-1} \cdot \gamma(\nu, z/2\sigma) \end{aligned} \right\} \quad \dots\dots\dots (31)$$

図-2は、上式で、 $\sigma=1$ とおいた場合の z を標準化変量 ξ と記した累積分布の変化の状態を示したもので、形状母数 n の増加に伴って $F_n(\xi)$ は次第に減少する。

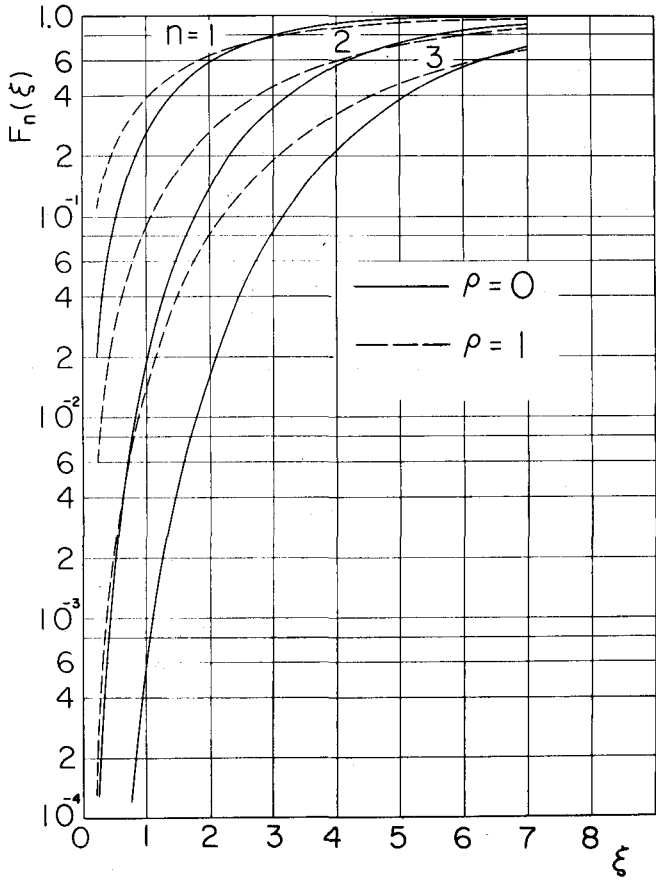


図-2 二変数ガンマ分布の和の分布 ($\rho=0$ および1の場合)

(iii) 分布のあてはめ⁽¹¹⁾

イ. 形状母数の異なる場合

観測標本に理論分布を適合させるための母数の推定について述べる。この場合、現在のところ最尤解は求められていないので、積率解のみの結果を示す。式(17)の密度分布に従う母集団から N 組の標本 (x_{1i}, x_{2i}) ($i=1, 2, \dots, N$)を取り出す。式(2)を使って積率を求めることにより、積率解は以下のように表わされる。

$$\left. \begin{aligned} \hat{\nu}_i &= \frac{(\bar{x}_i)^2}{x_i^2 - (\bar{x}_i)^2}, & \hat{\sigma}_i &= \frac{\overline{x_i^2} - (\bar{x}_i)^2}{\bar{x}_i} \quad (i=1, 2) \\ \hat{\rho} &= \sqrt{\hat{\nu}_2 / \hat{\nu}_1} \cdot r, & \overline{x_i^j} &\equiv \sum_{k=1}^N (x_{ik})^j / N \quad (i, j=1, 2) \end{aligned} \right\} \dots\dots\dots (32)$$

ただし、 r は標本相関係数で、 $\nu_1 \geq \nu_2$ を仮定している。

なお、形状母数 ν_1, ν_2 が既和の場合には、上式は以下のように簡単になる。

$$\left. \begin{aligned} \hat{\sigma}_i &= \bar{x}_i / \nu_i, & \hat{\sigma}_i^2 &= \overline{x_i^2} / \nu_i (\nu_i + 1), \quad (i=1, 2) \\ \hat{\rho} &= \sqrt{\nu_2 / \nu_1} \cdot r \end{aligned} \right\} \dots\dots\dots (33)$$

ロ. 形状母数の等しい場合

前記の特別な場合として、形状母数が共通（対称ガンマ分布）の解を積率解と最尤解について示す。次式のとおりで、この場合には、相関母数の推定値と標本相関係数が完全に一致する。

$$\left. \begin{aligned} \hat{\nu} &= \frac{(\bar{x}_i)^2}{x_i^2 - (\bar{x}_i)^2}, & \hat{\sigma}_i &= \frac{\overline{x_i^2} - (\bar{x}_i)^2}{\bar{x}_i} \\ \hat{\rho} &= \frac{\overline{x_1 x_2} - \bar{x}_1 \bar{x}_2}{\sqrt{\overline{x_1^2} - (\bar{x}_1)^2} \sqrt{\overline{x_2^2} - (\bar{x}_2)^2}} \equiv r \end{aligned} \right\} \dots\dots\dots (34)$$

なお、形状母数が既和の際には、次式が解となる。

$$\left. \begin{aligned} \hat{\sigma}_i &= \bar{x}_i / \nu, & \hat{\sigma}_i^2 &= \overline{x_i^2} / \nu (\nu + 1) \quad (i=1, 2) \\ \hat{\rho} &= \nu \overline{x_1 x_2} / \bar{x}_1 \bar{x}_2 - \nu \equiv r \end{aligned} \right\} \dots\dots\dots (35)$$

ハ. 最 尤 解

一般には求め難いので、形状母数を既和とした解を示す。標本組 (x_{1i}, x_{2i}) ($i=1, 2, \dots, N$) より結合確率として尤度 $L = \prod_{i=1}^N f(x_{1i}, x_{2i})$ を求め、尤度方程式を解くと、尺度母数、相関母数の最尤解はそれぞれ次式のように求められる。

$$\left. \begin{aligned} \hat{\sigma}_j &= \bar{x}_j / \nu \quad (j=1, 2) \\ \nu \sqrt{\hat{\rho}} &= \frac{1}{N} \sum_{i=1}^N \frac{I_\nu(2\sqrt{\hat{\rho}} \xi_i \eta_i / (1-\hat{\rho}))}{I_{\nu-1}(2\sqrt{\hat{\rho}} \xi_i \eta_i / (1-\hat{\rho}))} \sqrt{\xi_i \eta_i} \end{aligned} \right\} \dots\dots\dots (36)$$

ただし、 $\xi_i = x_{1i} / \hat{\sigma}_1, \eta_i = x_{2i} / \hat{\sigma}_2$

すなわち、尺度母数は単独変量の平均値として推定し得るが、相関母数は等式中に陰の形で含まれるために、試算的に求めざるを得ず、扱いやすい形とはいえない。したがって、実用的には、労力をかけて厳密な最尤解を求めるより、概要を把握するだけであれば、むしろ積率解を使う方が簡便であろう。

また、形状母数が既知の場合の最尤解と積率解とを比較すると、尺度母数については両者が完全に一致するが、相関母数については、積率解が標本相関係数の定義式と一致するのに反して、最尤解では一般に一致しないことが分る。

b. 二変数指数分布

指数分布はガンマ分布の特別な場合であり、解析的にもっとも単純な形である。水工計画上、数日程度以下数時間位までの短時間降水量、降水や無降水の日間隔、積雪深などの計画基本量として重要な水文量がこれによく適合するといわれる。また、応用上重要な極値分布の理論誘導の出発点としての意義も大きい。こうした意味から、最近Gumbel, Freund らによって周辺分布が指数分布に従うある種の二変数指数分布の研究が見受けられ、^{(12), (13)} その解析結果から従来の正規分布理論に偏重した相関性の吟味や回帰分析のあり方に批判が生じつつある。しかし、彼らが考察の対象とした二変数指数分布の理論モデルは非常に特殊なものであり、その結果得られた結論も極めて限定されたものであり普遍性に乏しい。

ここでは、前記の二変数ガンマ分布の成果を利用して、二変数指数分布の基礎的特性の概要を示し、実用化への方針を略述しておく。

(i) 基礎式

1. 分布関数

二変数指数分布の密度分布、周辺分布および条件付分布はそれぞれつぎのようになる。ただし、 $i \neq j$ で、 $i, j = 1, 2$ である。

$$\left. \begin{aligned} f(x_1, x_2) &= \frac{1}{\sigma_1 \sigma_2 (1-\rho)} \exp \left\{ -\frac{x_1}{\sigma_1 (1-\rho)} - \frac{x_2}{\sigma_2 (1-\rho)} \right\} \cdot I_0 \left(\frac{2\sqrt{\rho}}{1-\rho} \sqrt{\frac{x_1 x_2}{\sigma_1 \sigma_2}} \right) \\ f(x_i) &= \sigma_i^{-1} \exp(-x_i / \sigma_i), \quad F(x_i) = \exp(-x_i / \sigma_i) \\ f(x_i | x_j) &= \frac{1}{\sigma_i (1-\rho)} \exp \left\{ -\frac{x_i}{\sigma_i (1-\rho)} - \frac{\rho x_j}{\sigma_j (1-\rho)} \right\} \cdot I_0 \left(\frac{2\sqrt{\rho}}{1-\rho} \sqrt{\frac{x_1 x_2}{\sigma_1 \sigma_2}} \right) \end{aligned} \right\} \quad (37)$$

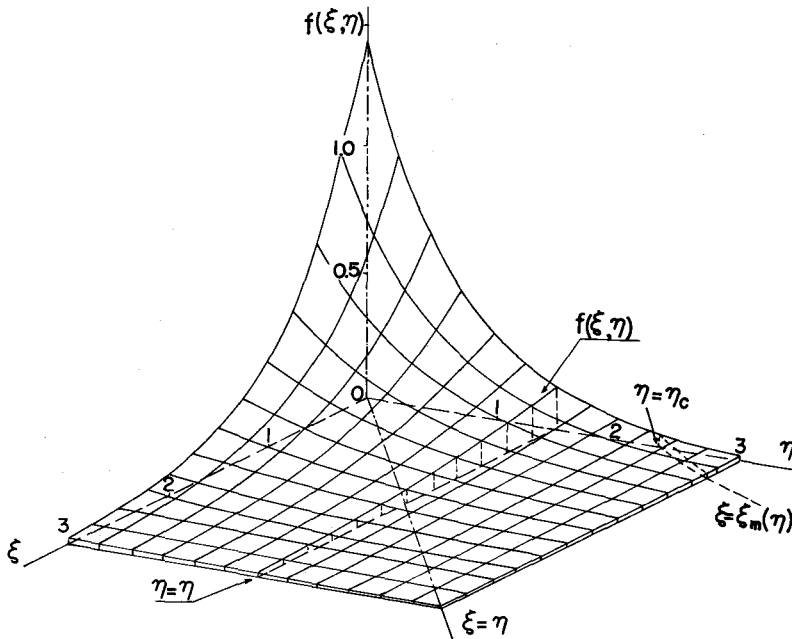


図-3 二変数指数分布の密度分布曲面 ($\rho = 0.3$ の場合)

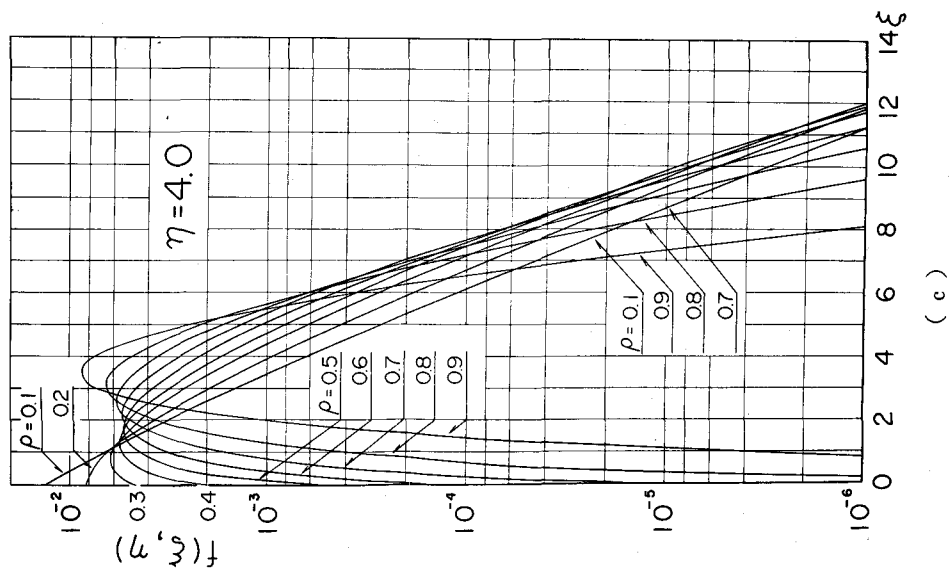
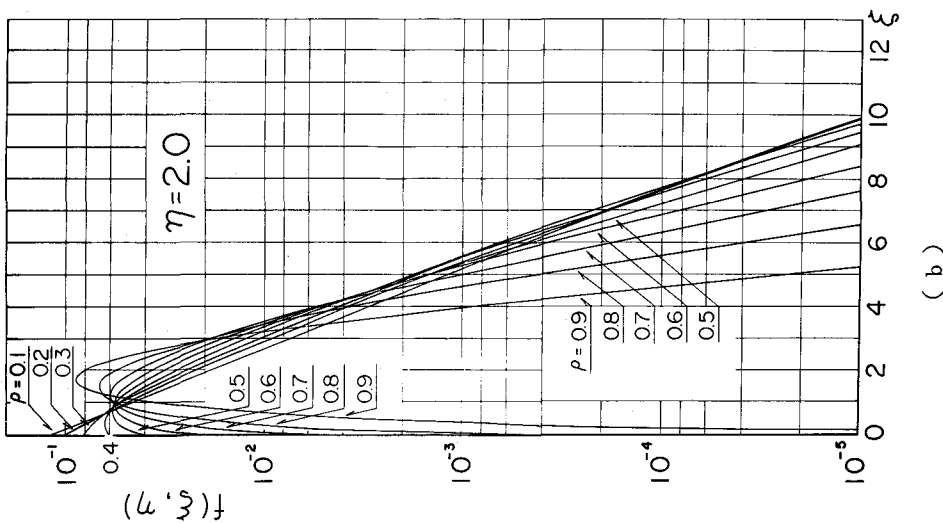
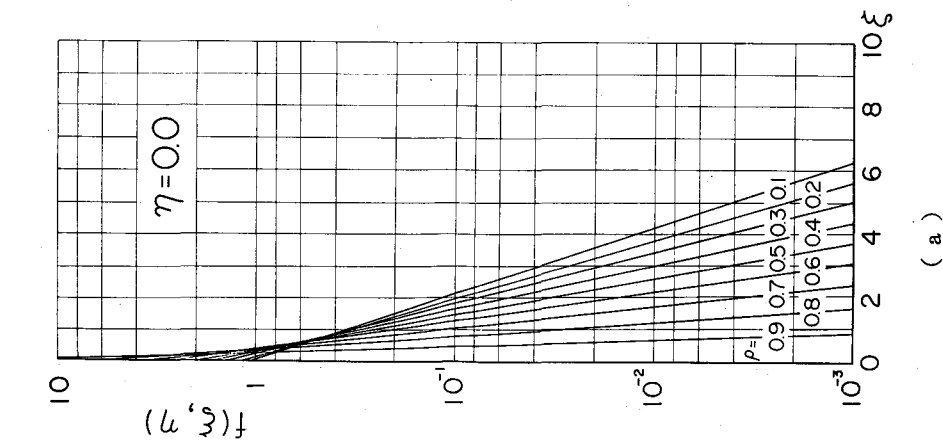


図-4 二変数指数分布の密度分布 ($\eta = \text{const}$)

その密度分布曲面の概要を $\rho = 0.3$ として図-3に示す。ただし、座標には規準変換量 $\xi = x_1/\sigma_1$, $\eta = x_2/\sigma_2$ を用いている。

さらに詳細に密度曲面を把握するために、曲面を η の指定値における切断面の図形で表現する。まず、 $\eta = 0$ では、 ξ が増すにつれて $f(\xi, \eta)$ は指数関数的に減ずる。ついで、 η が増すに伴い、その減少割合は指数減少の割合より鈍くなり、ある一定の限界値 $\eta = \eta_c$ から最頻値が $\xi = \xi_m(\eta)$ において出現することとなる。その模様を図-4(a), (b), (c)に示す。

ロ. 分布特性値と回帰特性

x_1, x_2 に関する p, q 次の積率 ν_{pq} は次式

$$\nu_{pq} = p! q! (1-\rho)^{p+q+1} \sigma_1^p \sigma_2^q \cdot F(p+1, q+1; 1; \rho)$$

となることを利用して、主要な分布特性値は以下のように求められる。

$$\left. \begin{array}{l} \text{平均値} \quad m_i = \sigma_i^2, \quad \text{分散} \quad D_i^2 = \sigma_i^2 \quad (i=1, 2) \\ \text{相関係数} \quad r = \rho, \quad \text{ひずみ係数} 2, \quad \text{過剰係数} 6 \end{array} \right\} \dots\dots\dots (38)$$

回帰特性としては、たとえば x_j を指定した際の x_i の条件付平均値、条件付分散は以下のようになる。($i \neq j; i, j = 1, 2$)

$$\left. \begin{array}{l} E(x_i | x_j) = \sigma_i + \rho(\sigma_i/\sigma_j)(x_j - \sigma_j) \\ D^2(x_i | x_j) = \sigma_i^2 + (1-\rho)^2 + 2(\sigma_i^2/\sigma_j) \rho(1-\rho)x_j \end{array} \right\} \dots\dots\dots (39)$$

ハ. 条件付非超過確率

計算の実用化のために、条件付非超過確率を定義し、その応用について説明する。規準化変量 $\xi = x_1/\sigma_1$, $\eta = x_2/\sigma_2$ で表示すれば、条件付非超過確率は次式で定義される。

$$\begin{aligned} F(\xi | \eta) &\equiv \int_0^\xi f(\xi | \eta) d\xi \\ &= \frac{1}{1-\rho} \exp\left(-\frac{\rho\eta}{1-\rho}\right) \cdot \int_0^\xi \exp\left(-\frac{\xi}{1-\rho}\right) \cdot I_0\left(\frac{2\sqrt{\rho\eta}}{1-\rho} \sqrt{\xi}\right) d\xi \end{aligned} \dots\dots\dots (40)$$

この $F(\xi | \eta)$ は、 ξ, η のみならず ρ にも依存するので、実用化には詳細な数表化が要求される。そこで、要求精度に応じた区間 $(0, \xi)$ の細分点での $f(\xi | \eta)$ を求めて数値積分を実施した。計算結果は ρ, ξ および η の値に対する $F(\xi | \eta)$ の値、あるいは ρ, η および $F(\xi | \eta)$ を与えた場合の ξ の値として数表化されている。

これら数表を用いれば、1.2のbで記述した二変数正規分布の確率積分に対応した確率計算が、この場合も可能となる。すなわち、任意領域 D を規準化座標 (ξ, η) で表示すると、二変数指数分布の場合、領域 D に対する確率 $P(D)$ は次式で求められる。

$$\begin{aligned} P(D) &\equiv \int \int_D f(\xi, \eta) d\xi d\eta = \int_{\eta_l}^{\eta_u} f(\eta) \left\{ \int_{\xi_{il}}^{\xi_{iu}} f(\xi | \eta) d\xi \right\} d\eta \\ &= \sum_i \exp(-\eta_i) \{ F(\xi_{iu} | \eta_i) - F(\xi_{il} | \eta_i) \} \cdot \Delta \eta_i \end{aligned} \dots\dots\dots (41)$$

ここに、 η_u, η_l は η の変域の上下限を、 ξ_{iu}, ξ_{il} は D を ξ 軸に平行な直線で細分した際の i 番目要素の上下限、 $\Delta \eta_i$ は i 番目要素の η 方向間隔を示す。 $F(\xi | \eta)$ は数表から適当な補間法によって求めることができる。

さらに、たとえば日降水量などを計画基本量とした場合、二変数指数分布としての模擬発生が有用となるが、⁽⁹⁾ その際の正規分布との相違を述べておく。

まず、指定変数 x より従属変数 y を推定する場合、 (x, y) が二変数正規分布に従えば、 y の推定値は、その回帰特性により、次式

$$y = m_y + \rho(\sigma_y/\sigma_x)(x - m_x) + \sigma_y \sqrt{1 - \rho^2} \cdot t \quad \dots\dots\dots(42)$$

によればよい。ここに、 t は規準正規乱数である。すなわち、右辺第3項がランダム特性、それ以外が回帰特性を示す。二変数指数分布の場合、上式に対して、式(39)より、次式が適用できるはずである。

$$y = m_y + \rho(\sigma_y/\sigma_x)(x - m_x) + \sigma_y \sqrt{(1 - \rho)^2 + 2\rho(1 - \rho)x/\sigma_x} \cdot \varepsilon \quad \dots\dots\dots(43)$$

上式中の ε は、形式的には式(42)の t に対応するが、もはや ρ 、 x に無関係ではなく、条件付非超過確率 $F(y|x)$ の関数である。いま、 y 、 x の規準化変量を $\xi = y/\sigma_y$ 、 $\eta = x/\sigma_x$ と記すと、 y は次式のように $F(y|x)$ すなわち $F(\xi|\eta)$ の関数としてはじめて推定できることになる。

$$y = y[F(y|x)] = y[F(\xi|\eta)] \quad \dots\dots\dots(44)$$

したがって、この場合にも条件付非超過確率 $F(\xi|\eta)$ の数表は重要な役割を果すわけである。

なお、二変数指数分布についても、和分布や母数推定など実用的に重要な事項も多々あるが、前述の二変数ガンマ分布の結果より容易に類推できるので記述は省略しておく。

さて、以上の分布理論の応用として降雨のDAD解析への適用を例として説明する。

2. DAD解析への二変数ガンマ分布理論の応用

2.1 計画降雨とDAD解析

a. DAD解析の概念

治水計画において計画基本量となる豪雨の統計的性格を解明することは極めて重要である。ところで、降雨量は一般にある期間内の総量として観測され、その期間によって統計的性格をかなり異にする。普通の河川計画の対象としては日程度以下の降雨が、とくに中小河川では精々数時間以下が問題とされる。

さて、降雨現象そのものは場所的・時間的に著しく変動する。そこで、所要地点の洪水時の最大流量を推定する目的で、かなり巨視的な観点から、流域面積(Area)と降雨継続時間(Duration)を基本要因として、降雨量(Depth)を推定するというDAD解析の概念が発生したのは極めて自然といえる。しかも、当初は、所要規模の流域面積と継続時間に対する最大降雨量を既往資料から推定することを意図していたが、これら相互関係の複雑さや最大値の推定の困難さのために、経験公式の使用を前提とした資料の統計的処理による解析に留まっているのが現状といえる。最近とくに問題となる中小河川災害の原因と目される中間スケール、メソスケールなどの豪雨現象の降雨機構の解明が現在のところ必ずしも十分でないことを勘案すると、このような問題は多変数統計の問題と考えざるを得ないであろう。なお、以後では、流域面積と降雨量の関係は別の機会に譲り、降雨量と継続時間との関係を考察する。

b. 降雨強度式と統計的意義

継続時間と降雨量の関係の研究は、降雨強度曲線の経験公式にその典型をみることが出来る。さて、降雨強度曲線とは、任意の降雨継続時間に対応した降雨強度を一つの曲線式で表現しようとしたもので、慣用されているものは多種多様である。代表的なものとして、Talbot型、Sherman型、その変形と目される久野・石黒の型式、あるいはそれらを一般化したCleveland型などが挙げられよう。⁽¹⁷⁾

ところで、これら降雨強度曲線の形式およびそれに含まれる定数は全く経験的に決定されるといってよいが、当然対象とする降雨頻度により相違するはずである。これに対して石黒らは、あらゆる降雨強度曲線はその勾配を表わす特性係数に時間降雨強度を乗じて表現できるという考えから、特性係数法という確率的概念を提案している。⁽¹⁸⁾また、曲線式中の定数に直接的に確率年の導入を試みたものに、

FairやBernardによる式など種々の提案がなされてきた。すなわち、これら諸公式は基本的には経験式に準拠しながら、形式的にその定数項に統計的意義を導入しようとしたものと理解できる。しかし、

前提となる経験式の成立そのものに対する考察はほとんど不問に付されたままであったといつてよい。このように降雨時系列の特性に対する理論的配慮が薄弱な限り、たとえ確率の概念を形式的に付与したとしても普遍性のある計画降雨の策定は望み難いように思える。

最近、短時間におけるハイレググラフの形状に関する統計的性質を、独立降雨量の和分布の応用として、理論的に究明しようとする研究が石原・友杉によって試みられている。これは、降雨量を球に分割期間を箱に擬して、球を箱へランダムに配分するといういわゆる壺モデルとしての定式化であるが、観測資料による理論検証において、基礎仮定とした独立性は時間スケールが小さくなるとともに満足され難くなるという結果を報告している。⁽¹⁹⁾

一方、一連降雨における短時間雨量の実測によって、降雨時系列に自己相関性が無視し得ないといふ種々の報告⁽²⁰⁾を勘案すると、降雨配分過程においても、自己相関性の導入が不可欠であろう。このような考えの下に、ごく最近、ガンマ分布の和分布もガンマ分布族で近似できることを独立変量の和分布を基礎として誘導し、主として日単位以上の水文量時系列に対する時間スケールと分布形の関係を論じた室田・江藤らの注目すべき研究がある。ここでは、さらに理論的に厳密に前述した相関を考慮した二変数ガンマ分布理論の応用として、日単位程度以下で時間程度位までの短時間降雨の配分問題を説明してみよう。⁽²¹⁾

a. 降雨量時系列の確率分布モデル⁽²²⁾

(i) 降雨量時系列の和分布

いま、時刻 t における降雨強度を $r(t)$ とすると、通常入手し得るのは t までのある時間間隔 T にわたる積算値 $R(T, t)$ で、それは次式で表わされる。

$$R(T, t) = \int_{t-T}^t r(s) ds \quad \dots\dots\dots (45)$$

ここでは、 t は正とし、離散的に $t = T, 2T, 3T, \dots$ と採ることとする。さらに、 $i \equiv t/T$ 、 $R(T, t) \equiv R_i$ ($i = 1, 2, \dots$) と略記し、 R_1 と R_2 の和分布を考える。なお、 R_1 と R_2 は周辺分布が共通なガンマ分布に従うとしておく。 R_i ($i = 1, 2$) の密度分布は、式(18)で、 $\nu_i \equiv \nu$ 、 $\sigma_i \equiv \sigma$ 、 $x_i \equiv R_i$ (≥ 0) とおけばよく、結合分布は同様に式(17)で与えられる。以後、 R_i がこのようなガンマ分布族に従うことを $R_i \in G(\nu, \sigma; \rho)$ と表記する。

さて、和 $z = R_1 + R_2$ の密度分布は、すでに式(20)で与えられており、 ν が自然数や半奇数の際には累積分布も明らかで、このように相関をもつガンマ型変量の和はもはや厳密にはガンマ分布に従わない。しかし、日ないし時間程度の期間降雨量がほぼガンマ分布で近似し得ることはよく報告されているので、この経験に相応するような近似解を考えてみる。

まず、和 z の特性関数は

$$\begin{aligned} Q_z(\theta) &\equiv \int_{-\infty}^{\infty} \exp(i\theta z) \cdot f(\nu, \sigma, \rho; z) dz \\ &= \left[1 - \sigma^2(1-\rho)\theta^2 - 2\sigma\theta i \right]^{-\nu} \quad \dots\dots\dots (46) \end{aligned}$$

で、これより、和の平均値 m 、分散 d^2 、ひずみ係数 C_s に添字(2)を付すと、それぞれ以下ようになる。

$$\begin{aligned} m(2) &= 2\nu\sigma, \quad d^2(2) = 2\nu\sigma^2(1+\rho) \\ C_s(2) &= 2 \left\{ 4 - 3(1-\rho) \right\} (2\nu)^{-\frac{1}{2}} (1+\rho)^{-\frac{3}{2}} \quad \dots\dots\dots (47) \end{aligned}$$

さて、平均値と分散を近似式と厳密式で合致させるという条件から、和分布もガンマ分布族で近似し得るとした近似式の母数および累積分布はそれぞれ次式のように求められる。ただし、ダッシュは近似値

を意味する。

$$\nu(2)' = 2\nu / (1+\rho), \quad \sigma(2)' = \sigma(1+\rho) \quad \dots\dots\dots(48)$$

$$F(\nu', \sigma', \rho; z) = \{\Gamma(\nu(2)')\}^{-1} \cdot \gamma(\nu(2)', z/\sigma(2)') \quad \dots\dots\dots(49)$$

以上の近似の誤差をひずみ係数の近似度で評価すると、

$$\alpha[C_s(2)] = C_s(2)' / C_s(2) = (1+\rho)^2 / (1+3\rho) \quad \dots\dots\dots(50)$$

により、近似度は最小値でも 0.889 と、実用的には十分に近似度は高い。

さらに表現を単純化して、雨量時系列

$$\{R(T=1, t=i)\} \equiv \{R(1, i)\} \quad (i=1, 2, \dots)$$

を $\{R(i)\}$ と略記し、同様な考えで、和の個数を倍加する。一般に、単位期間の 2^n 倍 (n は自然数) の期間長に対する降雨量 $R(1)+R(2)+\dots+R(2^n)$ が

$$R(1)+R(2)+\dots+R(2^n) \in G(\nu(2^n)', \sigma(2^n)'; \rho(2^n))$$

ならば、

$$\nu(2^n)' = 2^n \cdot \nu(1) / \prod_{i=1}^n \{1+\rho(2^{i-1})'\}, \quad \sigma(2^n)' = \sigma(1) \cdot \prod_{i=1}^n \{1+\rho(2^{i-1})'\} \quad \dots\dots\dots(51)$$

が成り立つ。なお、その場合のひずみ係数は、

$$\left. \begin{aligned} C_s(2^n)' &\equiv 2 / \sqrt{\nu(2^n)'} = C_s(1) / \sqrt{\prod_{i=1}^n \{1+\rho(2^{i-1})'\} / 2} \\ \text{ただし } C_s(1) &= 2 / \sqrt{\nu(1)} \end{aligned} \right\} \quad \dots\dots\dots(52)$$

となる。

以上により、ある基準期間に対する降雨量を単位期間に分割した際の雨量分布の形状を推定するには、前記の経過を逆に辿ればよい。ただし、そのためには、任意の期間雨量時系列に対する自己相関特性の理論的解明を必要とする。

b. 降雨時系列の自己相関モデル⁽²⁾

(i) 弱定常性と線型 1 次自己回帰モデル

雨量時系列 $\{R(i)\} \equiv \{R(T=1, t=i)\}$ に対する 2 次以下の積率に関する統計量に着目する。

まず平均値と分散は

$$\left. \begin{aligned} m &\equiv E[R(i)] = \nu(1) \cdot \sigma(1) \\ d^2 &\equiv E[\{R(i) - m\}^2] = \nu(1) \cdot \{\sigma(1)\}^2 \end{aligned} \right\} \quad \dots\dots\dots(53)$$

ついで、相関母数については、積率解では通常の標本相関係数と合致することから、次式のような単位期間ずれの自己相関係数で表わされる。

$$\rho \equiv E[\{R(i) - m\}\{R(i+1) - m\}] / d^2 \quad \dots\dots\dots(54)$$

以後、これら諸量が時点 i に無関係 (弱定常性) であると仮定する。

さらに、あい継続する時系列 $R(i)$, $R(i+1)$ の関係として、いま採用している分布では回帰曲線が線型回帰をなすことを勧奨して、次式の線型 1 次の自己回帰モデルを前提とする。

ただし、 $\omega(i)$ は線型回帰よりの偏差項、すなわちランダム項で $E[\omega(i)] \equiv 0$ である。

$$\frac{R(i+1) - m}{d} = \rho \frac{R(i) - m}{d} + \omega(i) \quad \dots\dots\dots(55)$$

(ii) 継続期間と自己相関係数の関係

これには、基準期間長 S が単位期間長の 2^n 倍の場合 (n は自然数で、以後倍数系列分割の場合とい

う) とその他の任意分割の場合についての理論結果を示す。

イ. 倍数系列分割の場合

一般に、 $S = 2^n$ 単位の期間長の降雨量に対して、次式の漸化式の成り立つことが示される。

$$\rho(2^i) = \rho(2^{i-1}) \cdot \{1 + \rho(2^{i-1})\} / 2 \quad (i = 1, 2, \dots, n) \quad \dots\dots\dots 56$$

ロ. 任意分割の場合

分割をしない場合 ($n = 0$) を含めて、それと $n = 1, 2$ で式56に適合する多項式として、 $\rho(S)$ の近似値を求めると、

$$\rho(S) = \rho(1) [S \{1 - \rho(1)^2\} - 2\rho(1) \{1 - \rho(1)^S\}] / S^2 \{1 - \rho(1)\}^2 \quad \dots\dots\dots 57$$

のようになる。すなわち、上式は非分割 ($S = 1$) を含めて、 $S = 2, 4$ については式56と完全に一致し、さらに、1日程度の基準期間より1時間程度までの雨量配分に利用し得ることが、両式の対比を通じて、示される。

結局、任意の基準期間長 S に対する自己相関係数 $\rho(S)$ から単位期間 (S 等分割期間) に対する自己相関係数 $\rho(1)$ の推定が可能となった。したがって、基準期間長に対する降雨量分布が $G(\nu(S), \sigma(S); \rho(S))$ として既知ならば、式56あるいは57より、 $\rho(1)$ が求まる。さらに、式51より、

$$\nu(1) / \nu(S) = \sigma(S) / S \cdot \sigma(1) = \rho(S) / \rho(1) \quad \dots\dots\dots 58$$

であるから、 $\nu(1), \sigma(1)$ が分り、結局、

任意の単位期間に分割した際の分布

$G(\nu(1), \sigma(1); \rho(1))$ が求められることになる。

エ. 雨量配分率曲線の理論的推定

以後、新たに基準期間雨量を $R^{(1)}$ 、その S 等分割期間雨量を $R^{(s)}$ と記す。前述の結果より、 $R^{(1)}$ の二変数ガンマ分布としての母数が推定されると、所与の超過確率 F に対する $R^{(s)}$ が推定できるが、表現の便宜上、これを配分率 $R^{(s)} / R^{(1)}$ として表示しよう。なお、基準期間長 S に通常のように24時間を採用すれば、任意の継続時間 T と分割数 S との関係はつぎのとおりである。

$$T = 24 / S,$$

$$\text{あるいは } S = 24 / T \quad \dots\dots\dots 59$$

しかも、このように基準期間に日程度を問題とすれば、雨量分布の形状は指数分布ないしそれより若干歪んだ逆J字型となるのが通例である。著者は、 $\nu(S) = 1, 3/4, 1/2$ に対して、各種の相関母数 ρ 、非超過確率 F についての配分率曲線を算出・図示してその実用化を計っているが、図-5に $\nu(S) = 1$ の例を示すに止める。

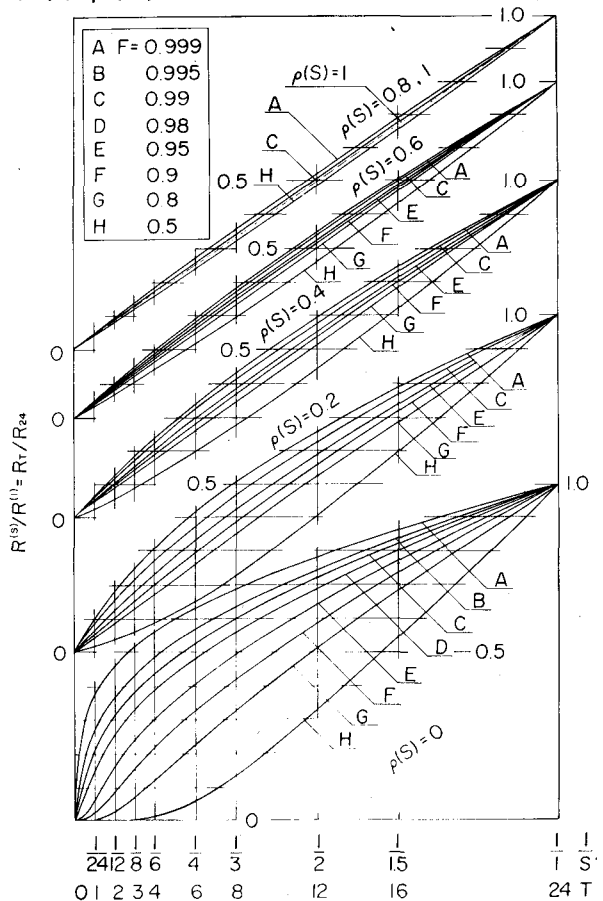


図-5 確率雨量配分率曲線 ($\nu(S) = 1$ の場合)

以下では、さらにこれを理論的な推定式として誘導した結果を述べておく。⁽²³⁾

(i) 規準化確率雨量の近似的表現

基準期間の S 分割雨量 $R^{(S)}$ を、 $\xi = R^{(S)} / \sigma(1)$ と規準化表示すると、 ξ の非超過確率 F は次式で計算できる。

$$1 - F(\xi) = \Gamma(\nu, \xi) / \Gamma(\nu), \quad \nu \equiv \nu(1) \quad \dots\dots\dots(60)$$

すなわち、雨量配分率の計算の基礎は、上式による非超過確率 F に対する基準化確率雨量 $\xi \equiv \xi_F$ の算出にある。これには第2種不完全ガンマ関数 $\Gamma(\nu, \xi)$ の計算を要し、 ξ_F は F について直接的に表現することはできない。そこで、概略の ξ_F の値を知るために、 ξ がかなり大きく、かつ ν が1の近傍という仮定で、 ξ_F の直接的表現を試みる。それには漸近近似式に

$$\Gamma(\nu, \xi) = \xi^{\nu-1} \cdot e^{-\xi} \left[\sum_{m=0}^{M-1} (-1)^m \Gamma(1-\nu+m) / \{\xi^m \cdot \Gamma(1-\nu)\} \right. \\ \left. + O(|\xi|^{-M}) \right] \quad [\xi \rightarrow \infty, M=1, 2, \dots] \quad \dots\dots\dots(61)$$

を用いると、最終的に、 ξ_F は次式のように F に関して陽の形で表現できる。ただし、 $\varepsilon = 1 - F$ である。

$$\xi_F = \left[-\{\log \varepsilon + \log \Gamma(\nu) + 2(1-\nu) + 1\} \right. \\ \left. + \sqrt{\{\log \varepsilon + \log \Gamma(\nu) + 2(1-\nu) - 1\}^2 + 12(1-\nu)} \right] / 2 \quad \dots\dots\dots(62)$$

なお、60式、62式の比較により、実用上問題となる $S \leq 24$ では両式は良好な一致を示すことが分っている。

(ii) 確率雨量配分率曲線の推定式

まず $S=1$ に対する ξ 、 $\nu(S)$ を ξ_0 、 ν_0 と記す。 ν_0 の近傍における ν (すなわち分割数 S があまり大きくない場合) に関して、理論誘導の出発点を式62におき若干の演算を行なえば、 ξ / ξ_0 の近似式は次式で与えられる。

$$\xi / \xi_0 = (\nu / \nu_0)^\alpha \quad \left. \begin{aligned} \alpha &\equiv (\nu_0 / 2\xi_0 \sqrt{b^2 + c}) \left[\{2 - \phi(\nu_0)\} (\sqrt{b^2 + c} + b) - 6 \right] \end{aligned} \right\} \quad \dots\dots\dots(63)$$

ただし、 $\phi(\nu) \equiv d \log \Gamma(\nu) / d\nu$ で、式中の b 、 c は以下のような超過確率 ε および基準期間雨量の形状母数 $\nu_0 \equiv \nu(S)$ の関数である。

$$b \equiv -\log \varepsilon - \log \Gamma(\nu_0) - 2(1-\nu_0) + 1, \quad c \equiv 12(1-\nu_0) \quad \dots\dots\dots(64)$$

上式に基づいて、一般的な確率雨量配分率曲線の推定式が誘導できる。

1. 完全独立の場合

説明の便宜上、基準期間に24時間を採用しておく。 $\rho(S)=0$ の場合、若干の計算の後、24時間雨量 R_{24} の任意の T 時間雨量 R_T への配分率は、以下の理論的な推定式として誘導される。

$$R_T / R_{24} = (T / 24)^\alpha \quad \dots\dots\dots(65)$$

上式中の指数 α は、式63、64で明らかのように、超過確率 ε および形状母数 $\nu(S)$ のみの関数である。結局、この場合、雨量配分率曲線は Sherman 型式に帰着して表現でき、それに含まれる経験的指数の水文統計的意義が明確にされたわけである。水工計画上、対象とされる程度の非超過確率 F および形状母数を具体的に与えたときの α の推定式およびその数値を表-1に示す。

表-1 完全独立の場合の雨量配分率曲線の指数と非超過確率および形状母数の関係

$\nu(S)$ F		1	3 / 4	1 / 2
α	Eq. (63)	$\frac{2.5772 b - 3}{\xi_0 b}$	$\frac{1.1572 (\sqrt{b^2 + 3} + b) - 2.25}{\xi_0 \sqrt{b^2 + 3}}$	$\frac{0.9909 (\sqrt{b^2 + 6} + b) - 1.5}{\xi_0 \sqrt{b^2 + 6}}$
		$b = -1n\epsilon + 1$	$b = -1n\epsilon + 0.2967$	$b = -1n\epsilon - 0.5724$
		$\xi_0 = -1n\epsilon$	$\xi_0 = (b - 2 + \sqrt{b^2 + 3}) / 2$	$\xi_0 = (b - 2 + \sqrt{b^2 + 6}) / 2$
	0.999	0.318	0.314	0.305
	0.995	0.397	0.397	0.395
	0.99	0.443	0.448	0.450
	0.95	0.610	0.629	0.647
	0.9	0.725	0.752	0.779

ロ. 自己相関を考慮した場合

$\rho(S) \approx 0$ に対する推定式は、同様に式(63)を基礎として、Sherman 型式と類似な次式で表現される。

$$R_T / R_{24} = (T / 24) (\rho_T / \rho_{24})^{1-\alpha} \quad \dots\dots\dots (66)$$

ただし、 ρ_{24} 、 ρ_T はそれぞれ24時間雨量、 T 時間雨量の相関母数である。ところで、 ρ_T / ρ_{24} は ρ_{24} 、 ν_{24} が既知ならば、 $T / 24$ のみの関数である。したがって、雨量配分率 R_T / R_{24} は、 ν_{24} 、 ρ_{24} が既知ならば、分割比 $T / 24$ のみの関数として示されたこととなる。

なお、式(66)を、基準期間雨量 $R^{(1)}$ の S 分割雨量 $R^{(s)}$ への配分率として、一般的に表現し直すと、次式となる。

$$R^{(s)} / R^{(1)} = (1 / S) \{ \rho(1) / \rho(S) \}^{1-\alpha} \quad \dots\dots\dots (67)$$

上式は雨量時系列の分布形状、自己相関性および超過確率を考慮した普遍性のある確率雨量配分率曲線の推定式といえることができる。

3. む す び

以上、二変数正規分布、二変数ガンマ分布を中心として、多変数分布理論の概要、その適用に関する基礎的方针ならびに最近の研究動向などを概述したが、紙数の関係もあって不十分な面も少なくない。とくに、標本抽出法、検定論などの方面には全く触れていないが、これらは実際問題の適用に当っては極めて重要であり、将来研究を進展させる必要がある。また、さらに一般的な多変数への拡張を計っていかねばならないと考えている。

参 考 文 献

- (1) たとえば、Johnson, N.L. and S.Kotz: 'Distributions in Statistics, vol.4 (Continuous Multivariate Distributions)', John Wiley and Sons, Inc., 1972
は最近の研究動向を伺うのに有用である。

- (2) 日本規格協会：統計数値表，E 多次元正規分布，JSA—1972，pp. 91～148，1972
- (3) 長尾正志：洪水処理計画に関する確率論的研究，京大防災研年報，第10号B，pp. 41～51，1967
- (4) たとえば，Mardia, K.V. : Families of Bivariate Distributions, Griffin's Statistical Monographs & Courses, 1970 は二変数分布の研究概要を知るのに便利である。
- (5) Wicksell, S.D. : On correlation functions of Type III, Biometrika, vol. 25, pp. 121～133, 1933
- (6) Kibble, W.F. : A two-variate Gamma-type distribution, SANKHYA, vol. 5, pp. 137～150, 1941
- (7) Krishnamoorthy, A.S. and M. Parthasarathy: A multivariate Gamma-type distribution, Annals of Mathematical Statistics, vol. 22, pp. 549～557, 1951, vol. 31, p. 229, 1960
- (8) 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，pp. 9～15, 1953
- (9) 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，pp. 15～19, 1953
- (10) 前出(8)の pp. 12～15
- (11) 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(2) — 二変数ガンマ分布の母数推定法 —，京大防災研年報，第14号B，pp. 43～56, 1971
- (12) Gumbel, E.J. : Bivariate exponential distributions, Journal of the American Statistical Association, vol. 55, pp. 698～707, 1960
- (13) Freund, J.E. : A bivariate extension of the exponential distribution, Journal of the American Statistical Association, vol. 56, pp. 971～977, 1961
- (14) 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 —，京大防災研年報，第13号B，pp. 105～133, 1970
- (15) 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(3) — 二変数指数分布の数値表 —，京大防災研年報，第14号B，pp. 57～75, 1971
- (16) 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(4) — 短時間降雨の模擬発生 —，京大防災研年報，第14号B，pp. 77～85, 1963
- (17) たとえば，土木学会：水理公式集，昭和38年増補改訂版，〔3・5・1〕降雨強度公式，pp. 384～386, 1963
- (18) 石黒政儀：特性係数法による確率降雨強度式算定法について，土木学会論文集，第74号，1961
- (19) 石原安雄・友杉邦雄：降雨の時間配分に関する確率論的考察，京大防災研年報，第14号B，pp. 87～102, 1971
- (20) 今井一郎他：降雨の統計的性質と電波の減衰，研究時報，第16巻3号，pp. 137～166, 1964
- (21) 室田明・江藤剛治・田中剛：水文量の和に関する統計的研究，土木学会論文報告集，第223号，pp. 23～31, 1974
- (22) 長尾正志：短時間豪雨分布の推定に関する二変数ガンマ分布の応用，名古屋工業大学学報，第25号，pp. 325～333, 1973
- (23) 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（投稿中）

以上