2018年6月15日　統計数理研究所　オープンハウス

２変量角度データのためのコピュラ

加藤　昇吾　数理・推論研究系　准教授

はじめに

２変量角度データ

２変量角度データとは，２つの角度のペア (θ_1, θ_2) $(-\pi < \theta_1, \theta_2 \le \pi)$ として表される観測の集合のことをいう．

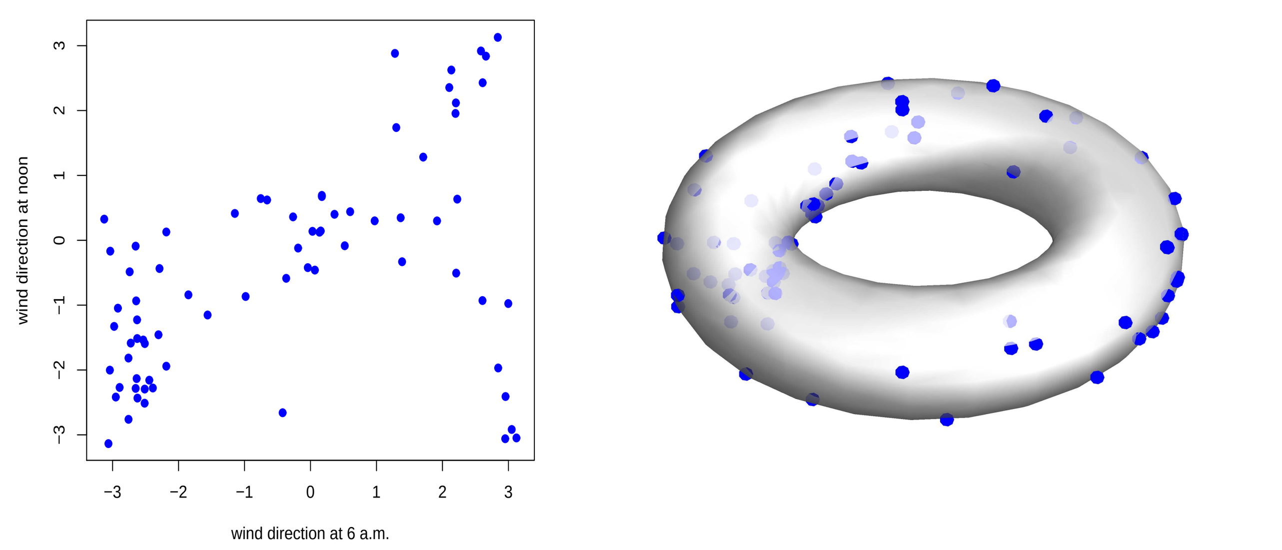

図１．ヒューストン（アメリカ）の気象観測所にて観測された午前６時と正午の風向データ．個々の観測が，左図では $(-\pi, \pi]^2$ 上の点として，右図ではトーラス上の点として表わされている．

コピュラ

２変量コピュラとは，それぞれの周辺分布が $[0, 1]$ 上の一様分布となる２変量分布関数のことをいう．

２変量コピュラは，変量間の依存関係を表すための分布関数である．

任意の２変量分布関数は，コピュラとそれぞれの周辺分布関数を用いて表現できることが知られている（スクラーの定理）．

研究の目的

角度には周期性があるため，２変量角度データに（周期性を持たない）２変量コピュラをそのまま当てはめても，満足な当てはめは期待できない．

そこで本研究では，多様な変量間の依存関係を表すことができる２変量角度データのためのコピュラを提案する．

なお本研究は，M.C. Jones教授(The Open University)，Arthur Pewsey准教授(University of Extremadura)との共同研究である．

２変量角度データのためのコピュラ

以下，２変量角度データのためのコピュラを，それぞれの周辺分布が $(-\pi, \pi]$ 上の一様分布となる２変量分布（関数）として定義する．

確率密度関数

以下の確率密度関数を考える．

$$f(\theta_1, \theta_2) = \frac{1}{4\pi^2} \sum_{m,n=-\infty}^{\infty} \phi(m, n)\, e^{-i(m\theta_1 + n\theta_2)}, \quad -\pi < \theta_1, \theta_2 \le \pi. \qquad (1)$$

ここに，$\phi(m, n)\,(\in \mathbb{C})$ は，任意の (θ_1, θ_2) に対して $f(\theta_1, \theta_2) \ge 0$，および，$\int_{(-\pi,\pi]^2} f(\theta_1, \theta_2) d\theta_1 d\theta_2 = 1$ を満たすように定義されているとする．

分布 (1) の性質

【定理１】確率ベクトル (Θ_1, Θ_2) が密度関数 (1) を持つとする．このとき，(Θ_1, Θ_2) の特性関数は，

$$E\left(e^{i(m\Theta_1 + n\Theta_2)}\right) = \phi(m, n),$$

で与えられる．ここに，m, n は任意の整数とする．

【定理２】密度関数 (1) を持つ分布がコピュラとなるのは，

$$\phi(m, 0) = \begin{cases} 1, & m = 0, \\ 0, & m \ne 0, \end{cases} \qquad \phi(0, n) = \begin{cases} 1, & n = 0, \\ 0, & n \ne 0, \end{cases}$$

が成り立つときで，またそのときに限る．

定理１より，$\overline{\phi(m, n)} = \phi(-m, -n)$，$\phi(0, 0) = 1$ となることは明らかである．以下，$m \ge 1$ に対して，具体的な $\phi(m, n)$ の例を与える．

例１：Wehrly–Johnsonコピュラ

$$\phi(m, n) = \begin{cases} \tilde{\phi}(m), & n = -qm, \\ 0, & n \ne -qm, \end{cases}$$

とすると，Wehrly & Johnson (1980) のコピュラが得られる．ここに，$\tilde{\phi}$ は円周上の密度関数のフーリエ級数，$q \in \{-1, 1\}$．

このとき，密度関数 (1) は，次の形で表現できる．

$$c(\theta_1, \theta_2) = \frac{1}{2\pi} g(\theta_1 - q\theta_2).$$

ここに，$g(\theta) = (2\pi)^{-1}[1 + 2\mathrm{Re}\{\sum_{m=1}^{\infty} \tilde{\phi}(m) e^{-im\theta}\}]$.

例２：新たなコピュラ１

$$\phi(m, n) = \begin{cases} \check{\phi}(n), & m = a, \ qn \le -1, \\ 0, & \text{otherwise}. \end{cases}$$

ここに，$\check{\phi}$ はある条件を満たす関数，$a \in \mathbb{N}$，$q \in \{-1, 1\}$．

例えば，$\check{\phi}(n) = \gamma \rho^{|n|-1}$ となるとき，密度関数 (1) は，

$$c(\theta_1, \theta_2) = \frac{1}{4\pi^2} \left\{ 1 + 2\gamma \frac{\cos(a\theta_1 - q\theta_2) - \rho\cos(a\theta_1)}{1 + \rho^2 - 2\rho\cos\theta_2} \right\},$$

となる．ここで，$0 \le \gamma \le 0.5$，$0 \le \rho \le 1 - 2\gamma$．

例３：新たなコピュラ２

$$\phi(m, n) = \begin{cases} \gamma \rho_1^{m-1} \rho_2^{|n|-1}, & m, -qn \in \mathbb{N}, \\ 0, & \text{otherwise}. \end{cases}$$

ただし，$0 < \rho_1, \rho_2 < 1$，$q \in \{-1, 1\}$，また，γ はある不等式を満たす．

このとき，密度関数 (1) は，以下のように表すことができる．

$$c(\theta_1, \theta_2) = \frac{1}{4\pi^2} \left\{ 1 + 2\gamma \frac{\cos(\theta_1 - q\theta_2) - \rho_2\cos\theta_1 - \rho_1\cos\theta_2 + \rho_1\rho_2}{(1 + \rho_1^2 - 2\rho_1\cos\theta_1)(1 + \rho_2^2 - 2\rho_2\cos\theta_2)} \right\}.$$

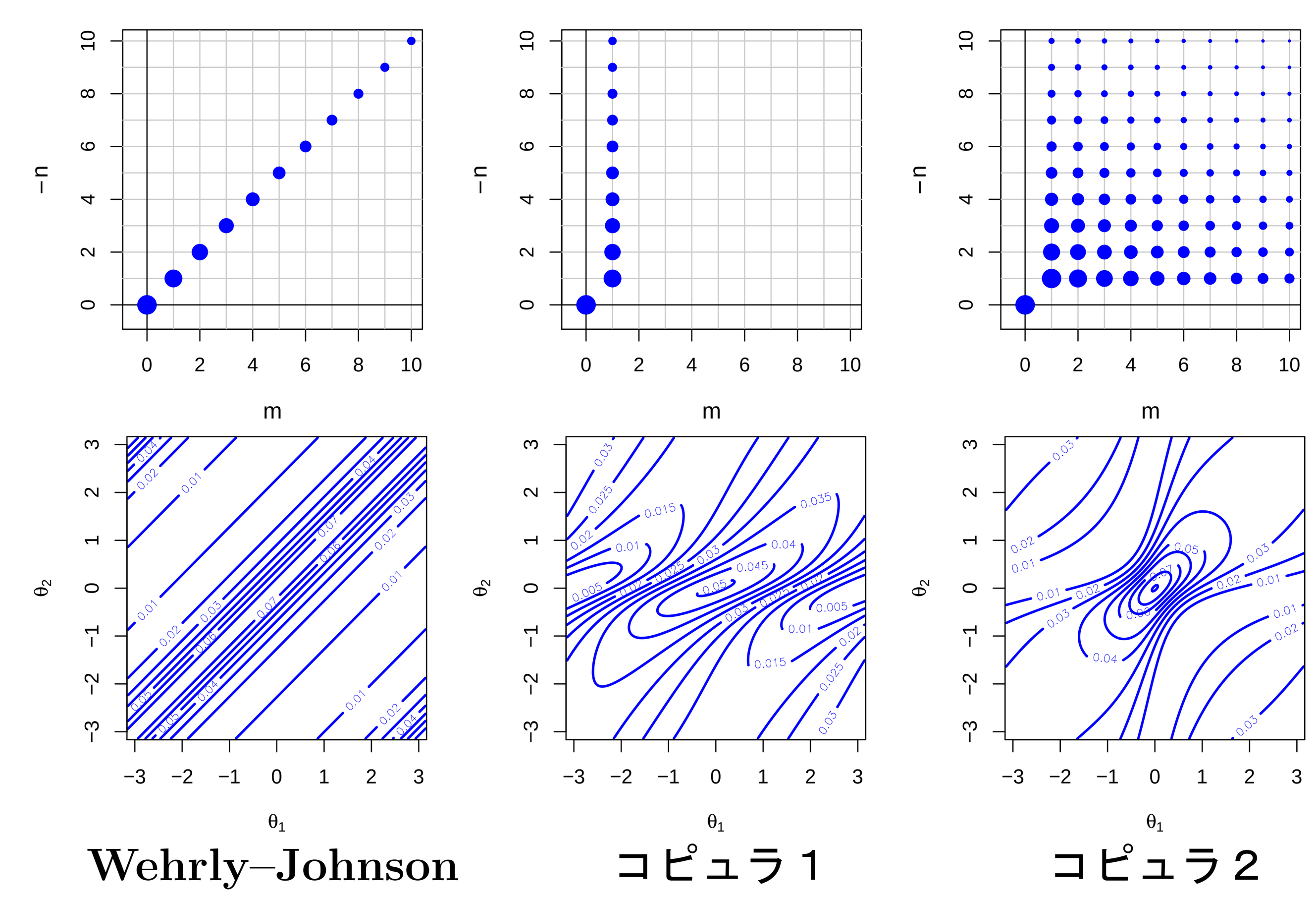

図２．【上段】$\{(m, n) \,|\, \phi(m, n) \ne 0\}$ のプロット（$q = 1$）．
【下段】密度関数の等高線プロット．ここに，$q = 1$，（左）$\tilde{\phi}(m) = 0.5^m$，（中）$\check{\phi}(n) = 0.25 \times 0.5^{|n|-1}$，$a = 1$，（右）$\gamma = 0.32$，$\rho_1 = \rho_2 = 0.5$．