多変量アルキメデスコピュラに基づく正規化スコアと判別

北海道大学　鈴川晶夫

北海道大学　安田佑喜

多変量確率分布の変量間従属性のモデル化において, コピュラは有用である. 代表的な多変量生存時間モデルである共有フレイルティモデルは, 変量間従属性をアルキメデスコピュラで表現する多変量分布モデルである. d 次元アルキメデスコピュラは

$$C(\boldsymbol{u}) = \psi\left(\sum_{i=1}^{d} \psi^{-1}(u_i)\right), \quad \boldsymbol{u} \in [0,1]^d \tag{1}$$

により与えられる. ここに, 関数 $\psi : [0,\infty) \to [0,1]$ は, d-単調な減少連続関数である (McNeil and Nešlehová 2009). また, d 次元アルキメデスコピュラ C に対するケンドール分布関数は

$$K_d(x) = \sum_{k=0}^{d-2} \frac{(-1)^k(\psi^{-1}(x))^k \psi^{(k)}(\psi^{-1}(x))}{k!} + \frac{(-1)^{d-1}(\psi^{-1}(x))^{d-1}\psi_{-}^{(d-1)}(\psi^{-1}(x))}{(d-1)!} \tag{2}$$

により与えられる (McNeil and Nešlehová 2009).

次の定理により, アルキメデスコピュラ C を分布関数にもつ d 次元確率変数ベクトル $\boldsymbol{U}$(周辺分布は $[0,1]$ 上の一様分布) の独立成分が与えられる.

定理. d 次元確率ベクトル $\boldsymbol{U}$ の分布関数が (1) 式のアルキメデスコピュラ C で与えられるとする.

$$S_i = \frac{\psi^{-1}(U_i)}{\sum_{j=i}^{d} \psi^{-1}(U_j)}, \quad i = 1, 2, \ldots, d, \quad T = C(\boldsymbol{U})$$

とおくとき, $S_1, \ldots, S_{d-1}, T$ は互いに独立である. また, $i = 1, 2, \ldots, d$ に対して, S_i はベータ分布 $\mathcal{B}eta(1, d-i)$ に従い, T の分布関数は (2) 式のケンドール分布関数 K_d により与えられる. さらに, $V_i = (1-S_i)^{d-i}$, $i = 1, 2, \ldots, d$, $K = K_d(T)$ とおくと, $V_1, V_2, \ldots, V_{d-1}$, K は互いに独立に $[0,1]$ 上の一様分布に従う.

本報告において, この定理を用いて独立正規化スコアを提案し, その独立正規化スコアに基づく判別分析法について議論する.

参考文献

1. Genest, C., Rivest, L.-P. (1993). Statistical Inference Procedures for Bivariate Archimedean Copula. *Journal of the American Statistical Association*, **88**, 1034-1043.

2. Genest, C., Rivest, L.-P. (2001). On the multivariate probability integral transformation. *Statistics & Probability Letters*, **53(4)**, 391–399.

3. Joe, H. (2014). *Dependence Modeling with Copulas.* Chapman and Hall, London.

4. McNeil, A.J., Nešlehová, J. (2009). Multivariate Archimedean copulas, d-monotone functions and ℓ1-norm symmetric distributions. *Ann. Statist*, **37**, 3059–3097.

5. McNeil, A.J., Quesada-Molina, J.J., Rodríguez-Lallena, J.A., Úbeda-Flores, M. (2001). Kendall distribution functions. *Statistics & Probability Letters*, **65(3)**, 263–268.