

A journey from multivariate normality to skewed data models

Tõnu Kollo

University of Tartu, Estonia, tonu.kollo@ut.ee

Keywords: copula, density approximation, elliptical distribution, multivariate normal distribution, skew elliptical distribution.

In "Introduction to the Multivariate Statistical Analysis", 1958, T. W. Anderson presented statistical models for the normal population in a compact matrix form. This became a dominating model in data analysis for years.

Elliptical distributions. As a generalization of the normal distribution Kelker (1970) introduced elliptical distributions. The first monographs on the topic appeared twenty years later. As in the normal case the distributions were elliptically contoured and axial symmetric.

Density approximations. Cornish & Fisher (1937) presented a complicated density function of Y through a simple density of X as a series expansion where moments of both distributions and derivatives of the density of X were involved. Finney (1963) carried over the idea to the multivariate case and gave a general expression between two densities in R^p . The approach made it possible to transform a symmetric (normal) density to skewed density approximation. I. Traat presented the multivariate density approximation in matrix representation in 1986. A general relation between two different dimensional densities was derived in 1998 by T. Kollo and D. von Rosen.

Skew elliptical distributions. In 1996 multivariate skew normal distribution was introduced by Adelchi Azzalini. The construction of transforming symmetric normal distribution to a skewed distribution was soon carried over to other continuous elliptical distributions. The distributions became quickly popular in modelling skewed data.

Copulas. Multivariate density approximations and skew elliptical distributions have marginals from the same distribution families (normal, t -distributed, for instance). Copula is a multivariate distribution function with uniform marginals and given dependence structure. Copula models enable to construct skewed data models with different type of marginals and given correlation structure. Nelsen (1998) gives the first monographic overview of the topic.

References

- [1] Cornish, E.A. & Fisher, R.A. (1937). Moments and cumulants in the specification of distributions. *International Statistical Review* **5**, 307-322.
- [2] Finney, D.J. (1963). Some properties of a distribution specified by its cumulants. *Technometrics* **5**, 63-69. C
- [3] Nelsen, R. (1998). *An Introduction to Copulas*. Springer, New York.