Abstracts of the Annual Conference of Finnish Statisticians 2007, "Multivariate Methods".
Eds. J. Niemelä, S. Puntanen, and E. P. Liski. Dept. of Mathematics, Statistics and Philosophy,
University of Tampere, 2007, ISBN 978-951-44-6957-2, pages 18–22

Skewed multivariate distributions

Tõnu Kollo
Institute of Mathematical Statistics
University of Tartu
J. Liivi 2
50409 Tartu, ESTONIA
Email: tonu.kollo@ut.ee

In last 10 years big changes have happened in the area of multivariate distributions. Until middle of the 1990s beside multivariate normal model only several families of symmetric elliptically contoured distributions were available for practical use. Different robustness studies were based on elliptical distributions, also there were several attempts to model data with heavier than normal tail area with elliptical distributions. Several books appeared on the topic in the beginning of the 1990s: Fang & Zhang (1990) and Fang, Kotz & Ng (1990) and Gupta & Varga (1993) can be mentioned as most relevant.

Skew-normal distribution

The development started with papers Azzalini & Dalla Valle (1996), Azzalini & Capitanio (1998) where multivariate skew normal distribution was introduced and examined. The idea of construction of the skew normal distribution was later on successfully carried over to other elliptical distributions. The multivariate skew normal distribution is defined by its density function in the following way:

$$f(\mathbf{x}) = 2 f_{N_p(\mathbf{0},\boldsymbol{\Sigma})}(\mathbf{x}) \Phi(\boldsymbol{\alpha}'\mathbf{x}) \tag{1}$$

where p-vector $\boldsymbol{\alpha}$ is the shape parameter and positive definite $p \times p$ -matrix $\boldsymbol{\Sigma}$ is the scale parameter while Φ denotes the distribution function of the standard normal distribution $N(0,1)$. Random vector $\mathbf{X}$ with the density (1) is denoted by $\mathbf{X} \sim SN_p(\boldsymbol{\Sigma}, \boldsymbol{\alpha})$. So the density of the multivariate normal distribution is transformed to a skewed density by the distribution function of the standard normal distribution. A useful fact about the distribution is that its moment generating function has a simple form:

$$M(\mathbf{t}) = 2e^{\frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}} \Phi\Big(\frac{\boldsymbol{\alpha}'\boldsymbol{\Sigma}\mathbf{t}}{\sqrt{1+\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}}}\Big).$$

This makes it possible to derive the first moments of the distribution by taking matrix derivatives from $M(\mathbf{t})$. Expectation and dispersion matrix are of the form:

$$E\mathbf{X} = \sqrt{\frac{2}{\pi}} \frac{\boldsymbol{\Sigma}\boldsymbol{\alpha}}{\sqrt{1 + \boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}}}, \qquad D\mathbf{X} = \boldsymbol{\Sigma} - E\mathbf{X}E\mathbf{X}'.$$

Other skew-elliptical distributions

It comes out that it is possible to form skewed families from elliptical distributions by the same scheme: the multivariate density is multiplied to a univarite distribution function where in argument we have a parameter vector which regulates skewness and is called the shape parameter. Results in this direction are presented in the collective monograph Genton (2004), Chapter 3. The idea has been generalized to different directions. In so-called Branco and Day approach (Branco & Day, 2001) an elliptical density is multiplied to the distribution function of a related elliptical distribution. In Arnold and Beaver approach (Arnold & Beaver, 2002) conditional approach is used. This brought Arnold to another type of generalization of the skew-normal distribution: in (1) the argument $\boldsymbol{\alpha}'\mathbf{x}$ of the distribution function was changed to a polynomial in $\mathbf{x}$ with several parameters involved (Arnold, Castillo & Sarabia, 2002). This change results with multimodality of the distribution what gives much bigger variability in shapes of distributions. In Wang-Boyer-Genton approach (Wang, Boyer & Genton, 2004) an elliptical density is multiplied to a general skewing function $\pi(\mathbf{x} - \boldsymbol{\theta})$ where $\pi\colon \mathbb{R}^p \to [0, 1]$ with $\pi(-\mathbf{x}) = 1 - \pi(\mathbf{x})$ and $\boldsymbol{\theta} \in \mathbb{R}^p$.

Skew *t* distribution

Beside skew-normal distributions another class of special interest among multivariate elliptical distributions has been multivariate t distribution. In Kotz & Nadarajah (2004) a systematic overview is given on multivariate t distribution as well as its skewed variants. The distribution is of interest in applications because of heavier tails than normal (skew-normal) distribution. Different constructions of skewed t distributions are based on the density function:

$$f(\mathbf{x}) = \frac{\Gamma((\nu + p)/2)}{(\pi\nu)^{p/2}\Gamma(\nu/2)|R|^{1/2}} \Big[1 + \frac{1}{\nu}(\mathbf{x} - \boldsymbol{\mu})'\mathbf{R}^{-1}(\mathbf{x} - \boldsymbol{\mu})\Big]^{-(\nu+p)/2},$$

where $\boldsymbol{\mu}$ and $\mathbf{R}$ denote the mean vector and the correlation matrix respectively and ν is the number of degrees of freedom.

Asymmetric Laplace distribution

The next class under consideration is asymmetric multivariate Laplace distribution. It was described by Tomasz Kozubowski and co-authors in a series of papers in the second half of the 1990s and the results with applications were presented in Kotz, Kozubowski & Podgorski (2001). Laplace distributed random p-vector $\mathbf{X}$ can be presented as a random sum (with geometric distribution) of multivariate normal vectors and belongs thus to the class of geometric stable distributions. This is the only distribution with finite moments in this class what gives a possibility to apply moment method in parameter estimation. The distribution is defined through the characteristic function:

$$\varphi(\mathbf{t}) = \frac{1}{1 - i\mathbf{t}'\boldsymbol{\theta} + \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}},$$

where p-vector $\boldsymbol{\theta}$ is the shape parameter (regulates location and skewness) and positive definite $p \times p$ -matrix $\boldsymbol{\Sigma}$ is considered as the scale parameter. Differentiation of the characteristic function gives us the first moments:

$$E\mathbf{X} = \boldsymbol{\theta}, \qquad D\mathbf{X} = \boldsymbol{\Sigma} + \boldsymbol{\theta}\boldsymbol{\theta}'.$$

As one can see, the Laplace distribution has heavier tails than skew-normal when they both stem from the normal distribution $N_p(\mathbf{0}, \boldsymbol{\Sigma})$. This makes the distribution specially attractive for financial applications. Unfortunately the distribution does not have explicit expression for its density function. Another disadvantage is that if $\mathbf{X}$ has asymmetric multivariate Laplace distribution, the sum $\mathbf{X} + \mathbf{a}$ does not belong to the same family, where $\mathbf{a}$ is a constant vector.

Multivariate extreme value distributions

Interest to rare events and their estimation has brought people to extreme value distributions. The book Kotz & Nadarajah (2005) is the third edition of the overview where Chapter 3 is dedicated to multivariate extreme value distributions (first edition in 2000). We can refer also to Coles (2001), Ch. 8 and Chapter 7 by A. L. Fougères in collective monograph Finkenstädt & Rootzen (2004).

Pros and cons

As we have have seen there is wide range of skewed multivariate distributions available for use. What are pluses and minuses of this situation? For most classes we can point out the following advantages.

(i) For modelling skewed data rich distribution families of different shapes are available.

(ii) As a rule, these distributions are easy to simulate.

(iii) Often the first moments have simple form and are easy to find.

We can point out also several problems and complications. One circle of questions is related to parameter estimation. Already in first papers by Azzalini et al it was shown that maximum likelihood method can give wrong answers in parameter estimation for the skew-normal distribution. The same is true for more general skew-elliptical distributions. Classical moment method works but the estimate of the shape parameter is biased. Bias correction brings us to quite complicated formulae (see Dunajeva, Kollo & Traat, 2003, for example). In practical fitting problems we have additionally to deal with constant shift. As the shape parameter also influences location , we have big problems with separation of these two parameters. No good universal solution exists so far. In practical examples the model is often found after empirical experimentation. In some situations reparametrization helps to solve the problem (Kollo & Srivastava, 2004). It seems that higher order moments (third and fourth) are needed to apply the moment method in three and four parameter cases. Another problem is related to hypothesis testing. Can we trust the found estimates? Not much is known about their distributions. As higher order moments are involved in estimation process, the results are sensitive to outliers. Simulation experiments show that convergence to asymptotic distributions can be very slow.

References

Arnold, B. C. & Beaver, R. J. (2002). Skew multivariate models related to hidden truncation and/or selective reporting. *Test* **11** 7–54.

Arnold, B. C., Castillo, E. & Sabia, J. M. (2002). Conditionally specified multivariate skewed distributions. *Sankhyā, Ser. A* **64** 206–226.

Azzalini, A. & Capitanio, A. (1999). Statistical applications of the multivariate skew normal distribution. *J. R. Statist. Soc. Ser. B* **61** 579–602.

Azzalini, A. & Dalla Valle, A. (1996). The multivariate skew normal distribution. *Biometrika* (1996) **83** 715–726.

Branco, M. D. & Day, D. K. (2001). A general class of multivariate skew-elliptical distributions. *J. Multivariate Anal.* **79** 99–113.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer, London.

Dunajeva, O., Kollo, T. & Traat, I. (2003). Bias correction for the shape paprameter of the skew normal distribution. *Tatra Mt. Math. Publ.* **26** 281–289.

Fang, K.-T., Kotz, S. & Ng, K.W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London.

Fang, K.-T. & Zhang, Y.-T. (1990). *Generalized Multivariate Analysis*. Science Press, Beijing, Springer-Verlag, Berlin.

Finkenstädt, B. & Rootzen, H. (Eds.) (2004). *Extreme Values in Finance, Telecommunications, and the Environment.* Chapman & Hall/CRC, Boca Raton.

Genton, M. G. (Ed.) (2004). *Skew-Elliptical Distributions and Their Applications. A Journey Beyond Normality*. Chapman & Hall/CRC, Boca Raton.

Gupta, A. K. & Varga, T. (1993). *Elliptically Contoured Models in Statistics*. Kluwer, Dordrecht.

Kollo, T. & Srivastava, M. S. (2004). Estimation and testing of parameters in multivariate Laplace distribution. *Comm. Statist., Theory and Methods* **33** 263–2387.

Kotz, S., Kozubowski, T. J. & Podgorski, K. (2001). *The Laplace Distribution, and Generalizations: A Revisit with Applications to Communications, Economics, Engineering and Finance*, Birkhäuser, Boston.

Kotz, S., & Nadarajah, S. (2004). *Multivariate t Distributions and Their Applications*, Cambridge University Press, Cambridge.

Kotz, S., & Nadarajah, S. (2005). *Extreme Value Distributions. Theory and Applications.* Imperial College Press, London.

Wang, J., Boyer, J. & Genton, M. G. (2004). A skew-symmetric representation of multivariate distributions. *Statist. Sinica* **14** 1259–1270.