Factor copulas constructed from stochastic processes

Matthias Scherer

(joint work with G. Bernhart, J.-F. Mai, S. Schenk)

Randomizing the parameter(s) of a given parametric family of univariate random variables is a popular technique to enrich the distribution in concern with additional stochastic properties and to create new probability laws. On a multivariate level, another motivation is to introduce dependence to originally independent objects by means of a joint mixture variable affecting multiple random variables in a similar way. A well-known example are extendible Archimedean copulas that can be interpreted as the survival copulas arising from a two step experiment: Firstly, a positive random variable M is simulated. Secondly, a sequence of exponential random variables with rate parameter M is drawn independently. Other examples are credit-risk models where a joint (random) default probability $p \in (0,1)$ is used as mixture variable in a sequence of Bernoulli(p) experiments, or loss models for insurance claims based on Poisson-distributed count variables with joint (random) intensities.

Factor models created in this way are popular due to, among others, the following facts: They enjoy a great level of interpretability, they are straightforward to simulate in large dimensions, the dimension of the considered problem is flexible, convenient limit results for large-dimensional random vectors $(X_1, \ldots, X_d)$ for $(d \nearrow \infty)$ are often computable, parametric families of mixture variables imply parametric families of copulas, and extensions beyond conditional iid (i.e. homogeneous one-factor models) are typically easy to find by, e.g., using multiple factors or inhomogeneous marginal laws. Moreover, hierarchical constructions are immediate in many cases, see [10].

Most factor models currently considered – in theory as well as in practice – are based on the aforementioned idea of using random parameters. This ansatz, however, can be extended to more involved random objects, e.g. stochastic processes. Providing more mathematical structure, a famous result by Bruno de Finetti (see [2]) shows that an infinite sequence of random variables $\{X_k\}_{k\in\mathbb{N}}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is exchangeable if and only if it is conditionally iid, i.e. there exists a sub-σ-algebra $\mathcal{G} \subset \mathcal{F}$ s.t. for all $d \geq 2$:

$$\mathbb{P}(X_1 \leq x_1, \ldots, X_d \leq x_d \,|\, \mathcal{G}) = \prod_{k=1}^{d} \mathbb{P}(X_1 \leq x_k \,|\, \mathcal{G}), \quad x_1, \ldots, x_d \in \mathbb{R}.$$

This is equivalent to the existence of a random distribution function $t \mapsto F_t$ such that conditioned on $\mathcal{G} := \sigma(\{F_t\}_{t\in\mathbb{R}})$, the components $\{X_k\}_{k\in\mathbb{N}}$ are iid and can be represented as

$$X_k := \inf\{t \in \mathbb{R} \,:\, F_t \geq U_k\}, \quad \{U_k\}_{k\in\mathbb{N}} \overset{\text{iid}}{\sim} \mathcal{U}[0,1],$$

where the sequence $\{U_k\}_{k\in\mathbb{N}}$ is independent of $\{F_t\}_{t\in\mathbb{R}}$.

In the case of randomized parameters, the stochastic process $\{F_t\}_{t\in\mathbb{R}}$ is an ordinary distribution function with random parameters. Opposed to this quite simplistic approach, we will need "true" stochastic processes in the following. Motivated especially by applications in portfolio credit-risk modeling and insurance, our research group has considered the questions:

- Given an interesting class of multivariate probability distributions, can we identify the conditionally iid subfamily, i.e. the subfamily representable by a one-factor construction?
- How can we create new probability laws starting from one-factor models?

These questions are flanked by the search for stochastic representations, efficient sampling strategies, extensions to multi-factor models, and real-world applications. It is quite difficult to give a general description on how the above questions can be solved, since the answer heavily depends on the family one has in mind. Generally speaking, one first has to identify the exchangeable subfamily of the distribution in concern in dimension d (since this is a necessary condition for extendibility), then one has to let the dimension go to infinity, and finally one has to guess a stochastic model for $\{X_k\}_{k\in\mathbb{N}}$ that ultimately reveals the nature of $\mathcal{G} := \sigma(\{F_t\}_{t\in\mathbb{R}})$.

Since the applications we have in mind involve the modeling of default times or the arrival times of insurance claims, it is not surprising that we consider fatal shock models on the one hand, and various multivariate extensions of the exponential law on the other hand. Fatal shock models have been thoroughly studied and applied in different fields, e.g., finance, hydrology, insurance, and reliability theory. In their classical stochastic representation, distinct shocks E_I hit combinations of components of a d-dimensional vector, i.e. the vector $(X_1, \dots, X_d)$ is defined as

$$X_k := \min_{\emptyset \neq I \subset \{1,\dots,d\}} \{E_I : k \in I\}, \quad k = 1, \dots, d. \tag{1}$$

The seminal example for both – multivariate exponential laws and fatal shock constructions as in (1) – is the Marshall–Olkin law (see [12]), in which the shocks E_I are independent and exponentially distributed. While this is convenient to interpret and use in low dimensions, the number of involved shocks increases exponentially in d, preventing such models from being applicable even in moderate dimensions. This problem is overcome as soon as a one-factor subfamily is identified. For the Marshall–Olkin law, one can show that the exchangeable subfamily is parameterized by d-monotone vectors, which for $(d \nearrow \infty)$ provides a link to completely monotone sequences. These are linked to Bernstein functions by a result in [4], which are finally in a one-to-one relation with Lévy subordinators (see [9]). The model for $\{X_k\}_{k\in\mathbb{N}}$, called Lévy-frailty construction in [8], can then be formulated as

$$X_k := \inf\{t \geq 0 : \Lambda_t \geq \epsilon_k\}, \quad k \in \mathbb{N}, \tag{2}$$

with a sequence of iid unit exponentials $\{\epsilon_k\}_{k\in\mathbb{N}}$ and $\{\Lambda_t\}_{t\geq 0}$ a Lévy subordinator. On a theoretical basis, this provides a marvelous link between classes of multivariate probability laws and the corresponding families of stochastic processes.

Generalizations to arbitrary exchangeable fatal shock models, extending the Marshall–Olkin model beyond the embedded exponential law, are considered in

[5]. Among others, all copulas of the functional form

$$C(u_1, \ldots, u_d) = \prod_{k=1}^{d} g_k(u_{(k)}), \quad d \geq 2, \tag{3}$$

are characterized analytically and by means of a fatal shock model. Here, $u_{(1)} \leq \ldots \leq u_{(d)}$ denotes the ordered arguments of C. Moreover, a link to additive processes, serving as stochastic factor, is revealed, and the specific cases of Sato processes (see [6]) and the Dirichlet process (see [7]) are discussed in quite some detail.

Turning to alternative multivariate definitions of the exponential law (apart from the Marshall–Olkin distribution), MSMVE distributions and distributions with exponential minima are natural candidates, see [3]. For both classes, the extendible subfamilies are characterized in [11] and a stochastic model based on strong (respectively weak) IDT subordinators is given. Specific families of MSMVE distributions (respectively their associated extreme-value copula) are constructed in [1]. This provides new (low-parametric) families of extreme-value copulas that might be interesting for applications, among others, due to their convenient stochastic representation.

References

[1] Bernhart, G.; Mai, J.-F.; Scherer, M.: *Constructing min-stable multivariate exponential distributions from Bernstein functions.* Working paper, (2015).

[2] De Finetti, B.: *La prévision: ses lois logiques, ses sources subjectives.* Annales de l'Insitut Henri Poincaré, **7** (1937) 1–68.

[3] Esary, J.D.; Marshall, A.W.: *Multivariate distributions with exponential minimums.* Annals of Statistics, **2**(1), (1974) 84–98.

[4] Gnedin, A.; Pitman, A.: *Moments of convex distribution functions and completely alternating sequences.* IMS Collections, Probability and Statistics: Essays in Honor of David A. Freedman, **2**, (2008) 30–41.

[5] Mai, J.-F.; Schenk, S.; Scherer, M.: *Exchangeable exogenous shock models.* Forthcoming in Bernoulli, (2015).

[6] Mai, J.-F.; Schenk, S.; Scherer, M.: *Two novel characterizations of self-decomposability on the half-line.* Working paper, (2015).

[7] Mai, J.-F.; Schenk, S.; Scherer, M.: *Analyzing model uncertainty via a distortion of the stochastic root: a Dirichlet prior approach.* Working paper, (2015).

[8] Mai, J.-F.; Scherer, M.: *Lévy-frailty copulas*, Journal of Multivariate Analysis **100**(7), (2009) 1567–1585.

[9] Mai, J.-F.; Scherer, M.: *Reparameterizing Marshall–Olkin copulas with applications to sampling.* Journal of Statistical Computation and Simulation, **81**(1) (2011) 59–78.

[10] Mai, J.-F.; Scherer, M.: *H-extendible copulas*, Journal of Multivariate Analysis **110**, (2012) 151–160.

[11] Mai, J.-F.; Scherer, M.: *Characterization of extendible distributions with exponential minima via processes that are infinitely divisible with respect to time*, Extremes, **17**(1) (2014), 77–95.

[12] Marshall, A.W.; Olkin, I.: *A multivariate exponential distribution.* Journal of the American Statistical Association, **62** (1967) 30–44.