Université catholique de Louvain
ISBA
Voie du Roman Pays 20
B-1348 Louvain-la-Neuve
Belgium

November 8, 2011

Université Claude Bernard Lyon I
Service des Etudes Doctorales
43, Boulevard du 11 Novembre 1918
69 622 Villeurbanne Cedex
France

Dear President,

It is my pleasure to send you hereby my report on the doctoral thesis

> "*Modélisation de la dépendance et mesures de risque multidimensionnelles*" by Elena Di Bernardino,

submitted for the PhD degree (Specialité Mathématiques Appliquées) at your university.

The central topic of the thesis concerns mathematical models for dependent random variables, seen as risks in a financial or actuarial context. The emphasis is on tail dependence, that is, on dependence between extreme events, occurring with low frequency but having a high impact, such as for instance large negative returns on the prices of different financial assets in a portfolio.

The core of the thesis, chapters 2–4 (in English), can be divided into two parts:

- chapter 2 deals with the estimation of tails of bivariate distributions;
- chapters 3 and 4 concern the definition, properties and estimation of tail-related risk measures for dependent pairs of risks, notably bivariate versions of value-at-risk and conditional tail expectation.

Furthermore, chapter 1 (in French), presents a review of the thesis. The thesis is concluded with a chapter on future perspectives as well as a series of technical appendices.

Chapters 2 and 3 are based upon papers that have been submitted for publication, whereas chapter 4 has already appeared as an article in the journal *ESAIM: Probability and Statistics*. I am confident that the first two papers will find their way to the scientific literature as well. Throughout, the material consists of a combination of rigorous mathematical theory and numerical experiments, involving Monte Carlo simulations as well as applications on data sets from insurance and hydrology.

Let me now review chapters 2–4 in some more detail. Afterwards, I will describe some research opportunities opened up by the thesis. The report concludes with a final appreciation.

First, **chapter 2** covers the paper *Estimating bivariate tails: a copula based approached*, co-authored by Véronique Maume-Deschamps and Clémentine Prieur, and currently submitted for publication. The goal of the paper is to construct and analyse an estimator for the joint upper tail of a bivariate distribution. There are two challenges: the tail region of interest may be so extreme that it contains no actual data points, and the joint tail may exhibit both asymptotic dependence and asymptotic independence.

The central idea of the paper is to propose a bivariate extension of the Pickands–Balkema–de Haan theorem providing asymptotic approximations to distribution tails. In the original, univariate result, the approximation only concerns the family of generalised Pareto distributions. In the bivariate case, however, the tail dependence structure needs to be dealt with as well. In the paper, this is done via the high-threshold limit of the upper tail dependence copula, defined as the copula of the conditional distribution of the random pair (X, Y) given that both X and Y exceed a threshold, the marginal exceedance probabilities tending to zero but being the same. Using bivariate regular variation theory, the limit copula can be parametrised via a parameter θ and a univariate function g.

The generalised Pareto limits and the limiting upper tail dependence copula together provide asymptotic approximations to the joint upper tail of the random pair (X, Y). Replacing population quantities by sample quantities at a penultimate threshold u yields a plug-in estimator for the joint tail. Interestingly, the estimators of the dependence structure are nonparametric and rank-based, guaranteeing maximum flexibility and offering protection against the risk of model misspecification.

The estimators of the objects appearing in the tail representation are shown to be (uniformly) consistent at certain rates. These results are transferred to consistency results for the joint tail itself. In the paper, the method is illustrated on four data sets, whereas in the appendix, the finite-sample performance of the method is assessed via simulation studies.

Chapter 3 is based on the paper *Some proposals about bivariate risk measures* co-authored by Areski Cousin, and currently submitted for publication. Extensions of the value-at-risk (VaR) and of the conditional tail expection (CTE) are presented for the case of a dependent pair of nonnegative random variables (X, Y), to be thought of as a pair of dependent, risky losses. Let $F(x, y) = \Pr(X \leq x, Y \leq y)$ be the joint cumulative distribution function of the pair (X, Y). The bivariate VaR at risk level

$0 < \alpha < 1$ is then defined as

$$\mathrm{VaR}(\alpha) = \big(\mathrm{E}[X \mid F(X,Y) = \alpha],\ \mathrm{E}[Y \mid F(X,Y) = \alpha]\big),$$

where the conditional expectation has to be interpreted via the bivariate densities of the pairs $(X, F(X,Y))$ and $(Y, F(X,Y))$. Similarly, the bivariate CTE is defined as

$$\mathrm{CTE}(\alpha) = \big(\mathrm{E}[X \mid F(X,Y) \geq \alpha],\ \mathrm{E}[Y \mid F(X,Y) \geq \alpha]\big).$$

These definitions are natural extensions of the univariate concepts. In case the variables X and Y exhibit perfect positive dependence (comonotonicity), the bivariate risk measures above are even shown to coincide with the univariate versions. Let $U = F_X(X)$ and $V = F_Y(Y)$, with F_X and F_Y the marginal distribution functions of X and Y, respectively, assumed to be continuous, and let C be the (unique) copula of F. It is shown that

$$\mathrm{VaR}(\alpha) = \big(\mathrm{E}[F_X^{-1}(U) \mid C(U,V) = \alpha],\ \mathrm{E}[F_Y^{-1}(V) \mid C(U,V) = \alpha]\big),$$
$$\mathrm{CTE}(\alpha) = \big(\mathrm{E}[F_X^{-1}(U) \mid C(U,V) \geq \alpha],\ \mathrm{E}[F_Y^{-1}(V) \mid C(U,V) \geq \alpha]\big),$$

demonstrating how the risk measures depend on the margins F_X and F_Y on the one hand and on the copula C of F on the other hand. Notably, these representations allow to analyse how the risk measures behave with respect to stochastic orderings and positive dependence concepts and how they vary in function of the risk level, α. Furthermore, the risk measures are shown to behave well with respect to axiomatic risk theory, imposing desirable properties that any risk measure should possess. In the special case that the copula C is Archimedean, analytical expressions are derived in terms of the Archimedean generator.

Geometrically, the risk measures are linked to the conditional distribution of the pair (X, Y) given its position relative to the level set $L(\alpha) = \{(x,y) \in \mathbb{R}_+^2 \mid F(X,y) \geq \alpha\}$ and its topological boundary. These level sets play a central role in **chapter 4**, based on the paper *Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory*, co-authored by Thomas Laloë, Véronique Maume-Deschamps and Clémentine Prieur, and to appear in the journal *ESAIM: Probability and Statistics*. Given an estimator F_n for F, a plug-in estimator for $L(\alpha)$ is proposed via $L_n(\alpha) = \{(x,y) \in \mathbb{R}_+^2 \mid F_n(x,y) \geq \alpha\}$; similarly for the boundary $\partial L(\alpha) = \{(x,y) \mid F(x,y) = \alpha\}$, which is estimated by replacing F by F_n. In order to derive consistency of the estimator, truncated versions $L(\alpha)^T = L(\alpha) \cap [0,T]^2$ are considered, where $T = T_n$ is allowed to grow to infinity at certain rates; similarly for the boundary $\partial L(\alpha)^T$ and for the versions based on F_n rather than on F. If F_n converges to F in a suitable sense (uniformly or in pth mean), then the plugin estimators of $L(\alpha)^{T_n}$ and $\partial L(\alpha)^{T_n}$ are shown to be consistent in terms of the Hausdorff distance and in terms of the Lebesgue measure of the symmetric difference. Each of these modes of convergence is interesting in its own respect. Moreover, rates of convergence are established in terms of the convergence rate of F_n to F.

Interestingly, the estimator F_n can be both a parametric or nonparametric. In case it is nonparametric, a natural estimator for the bivariate CTE(α) is constructed as the bivariate sample mean of the pairs (X_i, Y_i) that belong to the region $L_n(\alpha)^{T_n}$. The theoretical results are confirmed via simulation studies and they are illustrated by means of a case study in non-life insurance.

The thesis opens up quite a number of promising research perspectives, some of which are mentioned at the end of the thesis. Here is my own wish list:

- The contributions in the thesis are phrased in the bivariate case. How about extensions to the general, multivariate case?
- What are the relations between the bivariate VaR and CTE on the one hand and the univariate VaR and CTE of (linear) functions of the two variables on the other hand? This would be of obvious interest in analysing financial portfolios.
- Is the plug-in estimator $\widehat{\mathrm{CTE}}_\alpha(X, Y)^{T_n}$ also uniformly consistent as a function of α? Is it asymptotically normal? To answer these questions might require high-level empirical process theory.
- For the same estimator, can one get rid of the truncation at level T_n? And what if $\alpha = \alpha_n$ tends to 1, the risk level becoming more and more extreme?

In conclusion, the contribution of the doctoral thesis is to provide a number of new probabilistic concepts with relevance to risk theory and to make these available for practical use via novel statistical methods. The mathematical theory is worked out in detail and the asymptotic results are illustrated with extensive numerical studies. The results in the thesis provide solid ground upon which future scientific research can be inspired. Therefore, I am most happy to recommend Elena Di Bernardino for the degree of PhD in the Spécialité Mathématiques Appliquées.

Sincerely,

Johan Segers