On the distributional transform, Sklar's Theorem, and the empirical copula process

Ludger Rüschendorf[1]

[1] *University of Freiburg, Germany,*
email: ruschen@stochastik.uni-freiburg.de

We review the distributional transform of a random variable, some of its applications, and some related multivariate distributional transformations. The distributional transform is a useful tool, which allows in many respects to deal with general distributions in the same way as with continuous distributions. In particular it allows to give a simple proof of Sklar's Theorem in the general case. It has been used in the literature for stochastic ordering results. It is also useful for an adequate definition of the conditional value at risk measure and for many further purposes. We also discuss the multivariate quantile transform as well as the multivariate extension of the distributional transform and some of their applications. In the final section we consider an application to an extension of a limit theorem for the empirical copula process, also called empirical dependence function, to general not necessarily continuous distributions. This is useful for constructing and analyzing tests of dependence properties for general distributions.

Keywords: distributional transform, empirical copula process, empirical dependence function, Sklar's Theorem

1 Introduction

With the emergence of a need for multivariate dependence modelling in mathematical finance Sklar's Theorem from 1959 got a revival in the last ten years or so, in particular in the literature on risk management and more generally in mathematical economics and in mathematical finance modelling. The idea of Sklar's Theorem is to represent an n-dimensional distribution function F in two parts, the marginal distribution functions F_i and the copula C, describing the dependence part of the distribution. Both of them are connected by the formula

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)). \tag{1}$$

In the first part of this paper we give a simple proof of Sklar's Theorem. The proof is based on the distributional transform of real random variables which is discussed in Section 2. This transform allows to treat general distributions including discrete parts in much the same way as continuous distributions. It has been used in statistics for a long time for the construction of randomized tests. An early source of this transform is the statistics book by Ferguson (1967). This transform and its extensions were used a lot since the early eighties in connection with stochastic ordering, and in several papers and talks the connection to Sklar's Theorem was indicated (see e.g. Rüschendorf (1981) or the recent survey paper Rüschendorf (2005). It seems however useful to present this connection in more detail and explicitness. We also review some applications of the distributional transform to stochastic ordering and to an adequate definition of the conditional value at risk measure.

In the second part of the paper we introduce and discuss the multivariate extension of the distributional transform and of the quantile transform. These lead to regression type representations of random vectors and are useful for identification problems and various statistical test problems. For the detection of dependence properties, both transforms serve different purposes.

In the final section we discuss an application of the distributional transform to extend limit theorems for the empirical copula process, also called empirical dependence function or multivariate rank order process, to general distributions. Several test statistics of dependence properties can be represented as functionals of this empirical copula process and thus their limiting distribution can be obtained from this convergence theorem.

2 The distributional transform and Sklar's theorem

On a probability space $(\Omega, \mathfrak{A}, P)$ let X be a real random variable with distribution function F and let $V \sim U(0,1)$ be uniformly distributed on $(0,1)$ and independent of X. The *modified distribution function* $F(x, \lambda)$ is defined by

$$F(x, \lambda) := P(X < x) + \lambda P(X = x). \tag{2}$$

We define the (generalized) *distributional transform* of X by

$$U := F(X, V). \tag{3}$$

An equivalent representation of the distributional transform is

$$U = F(X-) + V(F(X) - F(X-)). \tag{4}$$

An early source for this transform is the statistics book of Ferguson (1967). For continuous d.f.s F, $F(x, \lambda)$ is identical to $F(x)$ and it is well-known that $U = F(X) \stackrel{d}{=} U(0,1)$. This property holds true for the distributional transform in general and the quantile transform is exactly the inverse of the distributional transform. Here the inverse of a distribution function F is defined as usual by

$$F^{-1}(u) = \inf\{x \in \mathbb{R}^1 : F(x) \geq u\}, \quad u \in (0,1).$$

For the sake of completeness we give a proof of this simple but interesting result.

Proposition 2.1 (Distributional transform) *Let U be the distributional transform of X as defined in* (3)*. Then*

$$U \stackrel{d}{=} U(0,1) \quad \textit{and} \quad X = F^{-1}(U) \ \textit{a.s.} \tag{5}$$

Proof: For $0 < \alpha < 1$ let $q_\alpha^-(X)$ denote the lower α-quantile, that is $q_\alpha^-(X) = \sup\{x : P(X \le x) < \alpha\}$. Then $F(X, V) \le \alpha$ if and only if

$$(X, V) \in \{(x, \lambda) : P(X < x) + \lambda P(X = x) \le \alpha\}.$$

If $\beta := P(X = q_\alpha^-(X)) > 0$ and with $q := P(X < q_\alpha^-(X))$ this is equivalent to $\{X < q_\alpha^-(X)\} \cup \{X = q_\alpha^-(X), q + V\beta \le \alpha\}$. Thus we obtain

$$P(U \le \alpha) = P(F(X, V) \le \alpha) = q + \beta P(V \le \frac{\alpha - q}{\beta}) = q + \beta \frac{\alpha - q}{\beta} = \alpha.$$

If $\beta = 0$, then

$$P(F(X, V) \le \alpha) = P(X < q_\alpha^-(X)) = P(X \le q_\alpha^-(X)) = \alpha.$$

By definition of U,

$$F(X-) \le U \le F(X). \tag{6}$$

For any $u \in (F(x-), F(x)]$ it holds that $F^{-1}(u) = x$. Thus by (6) we obtain that $F^{-1}(U) = X$ a.s. □

The distributional transform has a lot of interesting consequences. It implies that in many respects the case of discrete or mixed type distributions does not need some extra consideration compared to the case of continuous distributions. In particular it implies a simple proof of Sklar's Theorem.

Theorem 2.2 (Sklar's Theorem) *Let $F \in \mathcal{F}(F_1, \dots, F_n)$ be an n-dimensional distribution function with marginals $F_1, \dots, F_n$. Then there exists a copula C (i.e. an n-dimensional distribution function on $[0,1]^n$ with uniform marginals) such that*

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)). \tag{7}$$

Proof: Let $X = (X_1, \dots, X_n)$ be a random vector on a probability space $(\Omega, \mathfrak{A}, P)$ with distribution function F and let V be independent of X and uniformly distributed on $(0,1)$, $V \sim U(0,1)$. Considering the distributional transforms $U_i := F_i(X_i, V)$, $1 \le i \le n$, we have by Proposition 2.1 $U_i \sim U(0,1)$, and $X_i = F_i^{-1}(U_i)$ a.s., $1 \le i \le n$. Thus defining C to be the distribution function of $U = (U_1, \dots, U_n)$ we obtain

$$\begin{aligned} F(x) &= P(X \le x) = P(F_i^{-1}(U_i) \le x_i, 1 \le i \le n) \\ &= P(U_i \le F_i(x_i), 1 \le i \le n) = C(F_1(x_1), \dots, F_n(x_n)), \end{aligned}$$

i.e. C is a copula of F. □

Remark 2.3 *a)* ***Copula and dependence.*** *From the construction of the distributional transform it is clear that the distributional transform is not unique in the case when the distribution has discrete parts. Different choices of the randomizations V at the jumps or in the components, i.e. choosing $U_i = F_i(X_i, V_i)$, may introduce* artificial *local dependence between the components of a random vector on the level of the copula. From the copula alone one does not see whether some local positive or negative dependence is a real one or just comes from a choice of the copula. For dimension $n = 2$ the copula in Figure 1 could mean a real switch of local positive and negative dependence for the original distribution, but it might also be an artefact resulting from the randomization in case the marginals are e.g. both two point distributions while the joint distribution in this case could be even comonotone. Thus the copula information alone is not sufficient to decide dependence properties.*

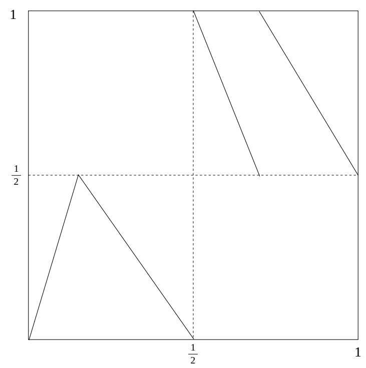

Figure 1: Copula of uniform distribution on line segments

b) ***Stochastic ordering.*** *The construction of copulas based on the distributional transform as in the proof of Sklar's theorem above has been used in early papers on stochastic ordering. The following typical example of this type of results is from Rüschendorf (1981, Proposition 7).*

Let F_i, G_i be one-dimensional d.f.s with $F_i \leq G_i$, $1 \leq i \leq n$. Then for any $F \in \mathcal{F}(F_1, \dots, F_n)$ there exists an element $G \in \mathcal{F}(G_1, \dots, G_n)$ with $G \leq_{\mathrm{st}} F$ and conversely for any $G \in \mathcal{F}(G_1, \dots, G_n)$ there exists an $F \in \mathcal{F}(F_1, \dots, F_n)$ with $F \leq_{\mathrm{st}} G$. Here $\leq_{\mathrm{st}}$ denotes multivariate stochastic ordering.

The proof in that paper uses the distributional transform $U_i = F_i(X_i, V)$ of a random vector $X \sim F \in \mathcal{F}(F_1, \dots, F_n)$. The distribution function of $U = (U_1, \dots, U_k)$ is a copula of F. Then the vector Y is defined by the quantile transforms of the components $Y = (G_1^{-1}(U_1), \dots, G_n^{-1}(U_n))$. By construction $Y \sim G \in \mathcal{F}(G_1, \dots, G_n)$ and from the assumption we obtain $Y \leq X$ pointwise and thus $G \leq_{\mathrm{st}} F$.

c) ***Conditional value at risk.*** *A recent application of the distributional transform is to risk measures. It is well-known that the conditional tail expectation*

$$TCE_\alpha(X) := -E(X \mid X \le q_\alpha), \tag{8}$$

where q_α *is the lower* α*-quantile of the risk* X*, does not define a coherent risk measure except when restricted to continuous distributions. This defect can be overcome by using the distributional transform* $U = F(X, V)$ *and defining the modified version, which we call conditional value at risk* (CVR_α)

$$\mathrm{CVR}_\alpha(X) = -E(X \mid U \le \alpha). \tag{9}$$

By some simple calculations (see Burgert and Rüschendorf (2006) one sees that

$$\mathrm{CVR}_\alpha(X) = -\frac{1}{\alpha}[EX1(X < q_\alpha) + q_\alpha(\alpha - P(X < q_\alpha)] = ES_\alpha(X). \tag{10}$$

Thus the more natural definition of CVR_α *coincides with the well established expected shortfall risk measure* $ES_\alpha(X)$ *which is a coherent risk measure. As consequence the expected shortfall is represented as conditional expectation and our definition in* (9) *of the conditional value at risk seems to be appropriate for this purpose.*

3 The multivariate distributional transform and the quantile transform

The distributional transform $F(X, V)$ as well as the inverse quantile transform $F^{-1}(U)$ have been generalized to the multivariate case. Let $X = (X_1, \dots, X_d)$ be a random vector with distribution function F and let $V_1, \dots, V_d$ be iid $U(0, 1)$-distributed random variables. Then the multivariate *quantile transform* $Y := \tau_F^{-1}(V)$ is defined recursively as

$$\begin{aligned} Y_1 &:= F_1^{-1}(V_1) \\ Y_k &:= F_{k|1,\dots,k-1}^{-1}(V_k \mid Y_1, \dots, Y_{k-1}), \quad 2 \le k \le d, \end{aligned} \tag{11}$$

where $F_{k|1,\dots,k-1}$ denote the conditional distribution functions.

The multivariate quantile transform was introduced in O'Brien (1975), Arjas and Lehtonen (1978), and Rüschendorf (1981). The basic result is that

$$Y = \tau_F^{-1}(V) \text{ is a random vector with d.f. } F. \tag{12}$$

This construction method has immediate applications to stochastic ordering results. If F and G are two d-dimensional d.f.s, then

$$\tau_F^{-1} \le \tau_G^{-1} \;\; [\lambda^d] \text{ implies } F \le_{\mathrm{st}} G. \tag{13}$$

(see Rüschendorf (1981). Pointwise ordering of the coupling construction by the multivariate quantile transform implies stochastic ordering. (13) implies many of the sufficient conditions known for stochastic ordering $\leq_{\mathrm{st}}$.

In contrast to the one-dimensional case there are however many different coupling constructions which are natural and useful. The construction in (11) depends on the ordering of the coordinates. Choosing a different ordering according to some permutation yields a different stochastic ordering condition in (13). The quantile transform (11) implies in particular a *regression representation* of the random vector $Y \sim F$ as

$$Y_2 = f_2(Y_1, V_2), \quad Y_3 = f_3(Y_1, Y_2, V_3)), \quad \ldots, \quad Y_d = f_d(Y_1, \ldots, Y_{d-1}, V_d) \tag{14}$$

representing Y_k as function of the past $Y_1, \ldots, Y_{k-1}$ and of some innovation V_k. A classical subject of probability theory is to establish a *functional characterization* as in (14) of various subclasses of stochastic processes like Markov chains, m-dependent sequences etc. (see Rüschendorf and de Valk (1993) for some of these directions).

The multivariate distributional transform was introduced in the general case in Rüschendorf (1981). The special case of absolutely continuous conditional distribution functions was already established much earlier by Rosenblatt (1952). Let X be a d-dimensional random vector and let $V_1, \ldots, V_d$ be iid $U(0,1)$-distributed random variables, $V = (V_1, \ldots, V_d)$. Then define for $\lambda = (\lambda_1, \ldots, \lambda_d) \in [0,1]^d$

$$\tau_F(x, \lambda) := (F_1(x_1, \lambda_1), F_2(x_2, \lambda_2 \mid x_1), \ldots, F_d(x_d, \lambda_d \mid x_1, \ldots, x_{d-1})), \tag{15}$$

where

$$\begin{aligned} F_1(x_1, \lambda_1) &= P(X_1 < x_1) + \lambda_1 P(X_1 = x_1), \\ F_k(x_k, \lambda_k \mid x_1, \ldots, x_{k-1}) &= P(X_k < x_k \mid X_j = x_j, j \leq k-1) \\ &\quad + \lambda_k P(X_k = x_k \mid X_j = x_j, j \leq k-1) \end{aligned} \tag{16}$$

are the distributional transforms of the one-dimensional conditional distributions. Finally the *multivariate distributional transform* of X is defined as

$$U := \tau_F(X, V) \tag{17}$$

Its basic properties are:

a)

$$U \sim U\big((0,1)^d\big) \tag{18}$$

b) τ_F^{-1} is the *inverse* of $\tau_F(x, \lambda)$, i.e.

$$X = \tau_F^{-1}(U) \text{ a.s.} \tag{19}$$

Remark 3.1 *a)* ***Standard representation.*** *The multivariate distributional transform standardizes a random vector X by a random vector U uniformly distributed on $[0,1]^d$. In comparison to the copula it also normalizes the marginals but 'forgets' all dependence information. In consequence this transformation serves different purposes compared to the copula construction. It gives in* (19) *a regression representation called 'standard representation' or 'innovation representation' of a process X of the form*

$$X_k = f_k(U_1, \dots, U_k), \quad k = 1, 2, \dots \tag{20}$$

regressing on an independent innovation sequence U_1, U_2,...

At the same time (19) *implies a regression representation for the sequence (X_n) itself:*

$$X_k = f_k(X_1, \dots, X_{k-1}, U_k) \text{ a.s.}$$

while in (14) *we get only a distributional variant of this regression representation. In fact this type of regression representation was first given for $n = 2$ by Skorohod (1976).*

b) ***Identification and statistical tests.*** *For the construction of a goodness of fit test for the hypothesis $H_0 : F = F_0$ the multivariate distributional transform allows to construct simple test statistics by checking whether the transformed random vectors $Y_i = \tau_{F_0}(X_i, V^i)$, $1 \le i \le n$, are uniformly distributed on the unit cube $[0,1]^d$. The problem however for the practical application is the calculation of conditional d.f.s.*

This principle of standardization is also useful for various other kinds of identification problems and for statistical tests as for example for the test of the two sample problem $H_0 : F = G$. Here using the empirical version of the distributional transform based on the pooled sample, we have again to check whether the transformed sample is a realization of an $U([0,1]^d)$-distributed variety.

c) ***Temporal dependence of vectors.*** *If $X^1, \dots, X^n$ describes the development of a d-dimensional portfolio vector over time, then to describe the development of dependence over time it may be useful to transform $X^1, \dots, X^n$ to $U^1, \dots, U^n$ by the multivariate distributional transform applied to the X_i (or its empirical counterpart). The dependence of the time evolution is preserved by this transformation and should be easier detected in the transformed form. On the other hand if one wants to see the development of dependencies within the portfolio vector it is more useful to use the copula transformation $\widetilde{U}^1, \dots, \widetilde{U}^n$ of $X^1, \dots, X^n$ which retains important aspects of the dependence within the portfolio. Thus both transformations serve different aspects of statistical inference in this problem.*

4 Empirical copula process and empirical dependence function

In this section we consider the problem of testing or describing dependence properties of multivariate distributions based on a sequence of observations. Some basic tools for the construction of test statistics which typically are based on some classical dependence measures like Kendall's τ or Spearman's ϱ (see Nelsen (1999)) or related dependence functionals are the reduced empirical process, the empirical copula function and the normalized empirical copula process. The distributional transform allows to extend some limit theorems known for the case of continuous distributions to more general distribution classes.

Let $X_j = (X_{j,1}, \dots, X_{j,k})$, $1 \le j \le n$ be k-dimensional random vectors with d.f. $F \in \mathcal{F}(F_1, \dots, F_k)$. For the statistical analysis of dependence properties of F a useful tool is the *reduced empirical process* (which one might also call *copula process*) defined for $t \in [0,1]^k$ by

$$V_n(t) := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (I(U_{j,1} \le t_1, \dots, U_{j,k} \le t_k) - C(t)), \tag{21}$$

where $U_j = (U_{j,1}, \dots, U_{j,k})$ is the distributional transform of X_j, $U_{j,i} = F_i(X_{j,i}, V^j)$, and C is the corresponding copula $C(t) = P(U_j \le t)$.

The construction of the distributional transforms $U_{j,i}$ is based on knowing the marginal d.f.s F_i. If F_i are not known it is natural to use empirical versions of them. Let

$$\widehat{F}_i(x_i) = \frac{1}{n} \sum_{j=1}^{n} 1_{(-\infty, x_i]}(X_{j,i}) \tag{22}$$

denote the empirical d.f.s of the i-th components of $X_1, \dots, X_n$. Then in the case of a continuous d.f. F the empirical counterparts of the distributional transforms are

$$\widehat{U}_{j,i} := \widehat{F}_i(X_{j,i}), \quad \widehat{U}_j = (\widehat{U}_{j,1}, \dots, \widehat{U}_{j,k}). \tag{23}$$

For continuous d.f. F_i we have that

$$n\widehat{U}_{j,i} = n\widehat{F}_i(X_{j,i}) = R^n_{j,i} \tag{24}$$

are the ranks of $X_{j,i}$ in the n-tuple of i-th components $X_{1,i}, \dots, X_{n,i}$ of $X_1, \dots, X_n$ and the ranks $R_{1,i}, \dots, R_{n,i}$ are a.s. a permutation of $1, \dots, n$. The *empirical copula function* is then given by

$$\widehat{C}_n(t) = \frac{1}{n} \sum_{j=1}^{n} I(\widehat{U}_j \le t), \quad t \in [0,1]^k \tag{25}$$

which induces the *normalized empirical copula process*

$$\begin{aligned} L_n(t) &:= \sqrt{n}(\widehat{C}_n(t) - C(t)) \\ &= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \{I(R_{j,1}^n \le nt_1, \dots, R_{j,k}^n \le nt_k) - C(t)\}, \quad t \in [0,1]^k. \end{aligned} \tag{26}$$

This normalized empirical copula process was introduced in Rüschendorf (1976, Section 3) (see also Rüschendorf (1974)) under the name *multivariate rank order process*. In fact in that paper more generally the sequential version of the process

$$L_n(s,t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{[ns]} \{I(\widehat{U}_j \le t) - C(t)\}, \quad s \in [0,1], t \in [0,1]^k \tag{27}$$

was introduced and analyzed for nonstationary and mixing random variables.

The empirical copula function $\widehat{C}_n$ was also introduced somewhat later by Deheuvels (1979) under the name *empirical dependence function*. Based on limit theory for the reduced empirical process it is shown in Rüschendorf (1974, 1976) and also in a series of papers of Deheuvels starting with Deheuvels (1979) that the normalized empirical copula process converges to a Gaussian process. Several nonparametric measures of dependence like Spearman's ϱho or Kendall's τau have corresponding empirical versions which can be represented as functionals of L_n. As consequence one obtains asymptotic distributions for these test statistics for testing dependence properties.

The distributional transform in Section 2 suggests to consider an extension of the emprical copula process to the case of general d.f.s F. The empirical versions of the $U_{j,i}$ are now defined as

$$\widehat{U}_{j,i} = \widehat{F}_i(X_{j,i}, V^j) \tag{28}$$

which are exactly $U(0,1)$ distributed. In order to avoid artificial dependence as described in Section 2 it is natural to let the copula $C_j(t) = P(U_j \le t)$, $t \in [0,1]^k$, be based on the same randomization V^j in all components of the j-th random vector such that $C_j(t) = C(t)$, $1 \le j \le n$. We define the normalized empirical copula process by

$$L_n(t) = \sqrt{n}(\widehat{C}_n(t) - C(t)), \quad t \in [0,1]^k. \tag{29}$$

The copula C has bounded nondecreasing partial derivatives a.s. on $[0,1]^k$ (see Nelsen (1999, p. 11)). Now the proof of Theorem 3.3 in Rüschendorf (1976) extends to the case of general distribution.

The basic assumption of this theorem is convergence of the reduced sequential empirical process, the sequential version of V_n in (21) (defined as in (27) for L_n). This assumption has been established for various classes of independent and mixing sequences of random vectors.

A) Assume that the reduced sequential process $V_n(s,t)$ converges weakly to an a.s. continuous Gaussian process V_0 in the Skorohod space D_{k+1}.

Note that the additional assumptions on V_0 in Rüschendorf (1976) served there to obtain stronger convergence results or to deal with more general assumptions on the distributions.

Theorem 4.1 *Under condition A) the sequential version $L_n(s,t)$ of the normalized empirical copula process converges weakly to the a.s. continuous Gaussian process L_0 given by*

$$L_0(s,t) = V_0(s,t) - s\sum_{i=1}^{n} \frac{\partial C(t)}{\partial t_i} V_0(1,\dots,1,t_i,\dots,1) \tag{30}$$

Based on this convergence result asymptotic distributions of test statistics testing dependence properties can be derived as in the continuous case. The proofs are based on representations or approximations of these statistics by functionals of L_n. For examples of this type see Rüschendorf (1974, 1976) and Deheuvels (1979, 1981).

References

Arjas, E. and T. Lehtonen (1978). Approximating many server queues by means of single server queues. *Math. Oper. Res.* **3**, 205–223.

Burgert, C. and L. Rüschendorf (2006). On the optimal risk allocation problem. *Stat. Decis.* **24**(1), 153–171.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. Un test non paramètrique d'independance. *Bulletin de la Classe des Sciences, V. Série, Académie Royale de Belgique* **65**, 274–292.

Deheuvels, P. (1981). A Kolmogorov–Smirnov type test for independence and multivariate samples. *Rev. Roum. Math. Pures Appl.* **26**, 213–226.

Ferguson, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach.* Probability and Mathematical Statistics, Vol. 1. New York: Academic Press.

Nelsen, R. B. (1999). *An Introduction to Copulas. Properties and Applications.* Lecture Notes in Statistic 139. Springer.

O'Brien, G. L. (1975). The comparison method for stochastic processes. *Ann. Probability* **3**, 80–88.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Annals of Mathematical Statistics* **23**, 470–472.

Rüschendorf, L. (1974). *Verteilungskonvergenz in ϕ-mischenden Prozessen mit Anwendungen für Order- und Rangstatistiken.* Ph. D. thesis, University of Hamburg.

Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Ann. Statist.* **4**(5), 912–923.

Rüschendorf, L. (1981). Stochastically ordered distributions and monotonicity of the OC-function of sequential probability ratio tests. *Mathematische Operationsforschung und Statistik Series Statistics* **12**(3), 327–338.

Rüschendorf, L. (2005). Stochastic ordering of risks, influence of dependence and a.s. constructions. In N. Balakrishnan, I. G. Bairamov, and O. L. Gebizlioglu (Eds.), *Advances on Models, Characterizations and Applications*, pp. 19–56. Chapman & Hall/CRC Press.

Rüschendorf, L. and V. de Valk (1993). On regression representations of stochastic processes. *Stochastic Processes and Their Applications* **46**(2), 183–198.

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.

Skorohod, A. V. (1976). On a representation of random variables. *Theory Probab. Appl.* **21**(3), 645–648.