Sklar’s Theorem revisited: an elaboration of the Rüschendorf transform approach

FRANK OERTEL
Deloitte & Touche GmbH
FSI Assurance
Quantitative Services & Valuation
D - 81669 Munich

Abstract

In many applications including financial risk measurement a certain class of multivariate distribution functions, copulas, has shown to be a powerful building block to reflect multivariate dependence between several random variables including the mapping of tail dependencies.

A key result in this field is Sklar’s Theorem which roughly states that any n-variate distribution function can be written as a composition of a suitable copula and an n-dimensional vector whose components are given by univariate marginal distribution functions, and that conversely the composition of an arbitrary copula and an arbitrary n-dimensional vector, consisting of n one-dimensional distribution functions (which need not be continuous), again is a n-variate distribution function whose i-th marginal is precisely the i-th component of the n-dimensional vector of the given distribution functions.

Meanwhile, in addition to the original sketch of a proof by Sklar himself, there exist several approaches to prove Sklar’s Theorem in its full generality, mostly under inclusion of probability theory and mathematical statistics but recently also rather technically under inclusion of non-trivial results from topology and functional analysis ([3]). An elegant probabilistic sketch of a proof was provided by L. Rüschendorf in [7].

We will revisit Rüschendorf’s - very short (and seemingly incomplete) - proof and elaborate important details to lighten the understanding of the basic underlying ideas of this proof including the major role of the so called “distributional transform”. Thereby, we will recognise that Rüschendorf’s proof mainly splits into two parts: a purely real-analytic one (without any assumption on randomness) and a probabilistic one.

To this end, we slightly generalise the approach in [7], allowing us to derive Theorem 2.10 and Lemma 2.14 - two results which might become very useful, in particular with respect to a simulation of random variables. To this end, we also provide a strict mathematical description of “flat pieces” of a right-continuous and non-decreasing real-valued function, leading to Corollary 2.7.

1. Introduction

The mathematical investigation of copulas started 1951, due to the following problem of M. Fréchet: suppose, one is given n random variables $X_1, X_2, \ldots, X_n$, all defined on the same probability space $(\Omega, \mathcal{F}, P)$, such that each random variable has a (non-necessarily continuous) distribution function $F_i$ ($i = 1, 2, \ldots, n$). What can then be said about the set of all possible n-dimensional distribution functions of the random vector $(X_1, X_2, \ldots, X_n)$ (cf. [6])? This question has an immediate answer if the random variables were assumed to be independent, since in this case there exists a unique n-dimensional distribution function of the random vector $(X_1, X_2, \ldots, X_n)$, which is given by the product $\Pi_{i=1}^n F_i$. However, if the

AMS 2010 subject classifications. 26A27, 60E05, 60A99, 62H05

Key Words and phrases. Copulas, distributional transform, generalised inverse functions, Sklar’s Theorem, tail dependence
random variables are not independent, there was no clear answer to M. Fréchet’s problem.

In [10], A. Sklar introduced the expression “copula” (referring to a grammatical term for a word that links a subject and predicate), and provided answers to some of the questions of M. Fréchet.

In the following 15 years, copulas (which are precisely finite dimensional distribution functions with uniformly distributed marginals), were mainly used in the framework of probabilistic metric spaces (cf. e.g. [8] and [9]). Later, probabilists and statisticians were interested in copulas, since copulas defined in a “natural way” nonparametric measures of dependence between random variables, allowing to include a mapping of tail dependencies. Since then, they began to play an important role in several areas of probability and statistics (including Markov processes and non-parametric statistics), in financial and actuarial mathematics (particularly with respect to the measurement of credit risk), and even in medicine and engineering.

One of the key results in the theory and applications of copulas, is Sklar’s Theorem (which actually has not been proven in [10]). It says:

**Sklar’s Theorem.** Let $F$ be a $n$-dimensional distribution function with marginals $F_1, \ldots, F_n$. Then there exists a copula $c_F$, such that for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$ we have

$$F(x_1, \ldots, x_n) = c_F(F_1(x_1), \ldots, F_n(x_n)).$$

Furthermore, if $F$ is continuous, the copula $c_F$ is unique. Conversely, for any univariate distribution functions $H_1, \ldots, H_n$, and any copula $c$, the composition $c \circ (H_1, \ldots, H_n)$ defines a $n$-dimensional distribution function with marginals $H_1, \ldots, H_n$.

Since the original proof of (the general non-continuous case of) Sklar’s Theorem is rather complicated and technical, there have been several attempts to provide different and more lucidly appearing proofs, involving mathematical techniques from probability theory and functional analysis.

Among those different proofs of Sklar’s Theorem, there is an elegant, yet very short proof, provided by L. Rüschendorf, originally published in [7]. He provided a very intuitive, and primarily probabilistic approach which allows to treat general distribution functions (including discrete parts and jumps) in a similar way as continuous distribution functions. To this end, he applied a generalised “distributional transform” which - according to [7] - has been used in statistics for a long time in relation to a construction of randomised tests. By making a consequent use of the properties of this generalised “distributional transform” together with Proposition 2.1 in [7], the proof of Sklar’s Theorem in fact follows immediately (cf. Theorem 2.2 in [7]). For the convenience of the reader we will complete our paper by providing the proof of Sklar’s Theorem again. All key inputs for the proof of Sklar’s Theorem clearly are provided by Proposition 2.1 in [7]. However, the very short proof of the latter result is rather difficult to reconstruct. It says:

**[7] - Proposition 2.1.** Let $X, V$ be two random variables, defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that $V \sim U(0, 1)$ and $V$ is independent of $X$. Let $F$ be the distribution function of the random variable $X$. Then $U := F_V(X) \sim U(0, 1)$, and $X = F^{-1}(U)$ $\mathbb{P}$-almost surely.
Here,
\[ F^-(\alpha) := q_F(\alpha) := \inf\{x \in \mathbb{R} : F(x) \geq \alpha\} \quad (0 < \alpha < 1) \]
defines the (left-continuous) lower quantile function of \( F \) which is a special case of a generalised inverse (cf. e.g. Definition A.20. in [5] and Definition 2.1 in [4]).

Our paper should bridge this gap. We will see later that Proposition 2.1 in [7] appears as a special case of Lemma 2.14 which itself does not require the assumption of an existing distribution function.

2. The proof

We are now going to give a detailed and complete proof of Proposition 2.1 in [7]. To this end, we will separate the proof into a purely real-analytic part (which does not involve any probability theory at all) and a measure theoretic and probabilistic part (which involves the use of Lebesgue-Stieltjes measures).

2.1. The non-random part

To reveal the main ingredients of Rüschendorf’s proof (cf. [7]) let us completely ignore any involved randomness and probability theory for the moment. We “only” are working within a subclass of real-valued functions, all defined on the real line, and with suitable subsets of the real line.

Let \( F : \mathbb{R} \to \mathbb{R} \) be an arbitrary right-continuous and non-decreasing function. Let \( x \in \mathbb{R} \). Since \( F \) is non-decreasing, it is well-known that both, the left-hand limit
\[ F(x-) := \lim_{z \uparrow x} F(z) = \sup \{F(z) : z \leq x\} , \]
and the right-hand limit
\[ F(x+) := \lim_{z \downarrow x} F(z) = \inf \{F(z) : z \geq x\} \]
are well-defined real numbers, satisfying \( F(x-) \leq F(x) \leq F(x+) \). Moreover, due to the assumed right-continuity of \( F \), it follows that \( F(x) = F(x+) \) for all \( x \in \mathbb{R} \). \( 0 \leq \Delta F(x) := F(x+) - F(x-) = F(x) - F(x-) \) denotes the (left-hand) “jump” of \( F \) at \( x \). We consider the following important transform of \( F \):

**Definition 2.1.** Let \( 0 \leq \lambda \leq 1 \). Put
\[ F_\lambda(x) := F(x, \lambda) := F(x-) + \lambda \Delta F(x) , \]
where \( x \in \mathbb{R} \). We call the real-valued function \( F_\lambda : \mathbb{R} \to \mathbb{R} \) Rüschendorf \( \lambda \)-transform.

Clearly, we have the following equivalent representation of the Rüschendorf \( \lambda \)-transform \( F_\lambda \):
\[ F_\lambda(x) = (1 - \lambda) F(x-) + \lambda F(x) \text{ for all } x \in \mathbb{R} . \]
In particular, for all \( (x, \lambda) \in \mathbb{R} \times [0, 1] \) the following inequality holds:
\[ F(x-) \leq F_\lambda(x) \leq F(x) . \quad (2.1) \]
Moreover, \( F_0(x) = F(x-) \) and \( F_1(x) = F(x) \).
**Assumption 2.2.** In the following we assume throughout that $F$ is bounded on $\mathbb{R}$ (i.e., the range $F(\mathbb{R})$ is a bounded subset of $\mathbb{R}$), implying that $F(\mathbb{R}) \subseteq [c_*, c^*]$ for some real numbers $c_* < c^*$. Moreover, let us assume that for any $\alpha \in (c_*, c^*)$ the set $\{x \in \mathbb{R} : F(x) \geq \alpha\}$ is non-empty and bounded from below.

Then the generalised inverse function $q_F^- : (c_*, c^*) \to \mathbb{R}$, given by

$$q_F^-(\alpha) := \inf\{x \in \mathbb{R} : F(x) \geq \alpha\},$$

is well-defined (cf. [4], [5]), implying in particular that

$$-\infty < q_F^- (\alpha) \leq q_F^-(\alpha +) = \inf\{x \in \mathbb{R} : F(x) > \alpha\} = \sup\{x \in \mathbb{R} : F(x) \leq \alpha\} =: q_F^+(\alpha) < \infty$$

for any $\alpha \in (c_*, c^*)$. Actually, since $F$ is assumed to be right-continuous, it follows that

$$q_F^-(\alpha) = \min\{x \in \mathbb{R} : F(x) \geq \alpha\}$$

for all $\alpha \in (c_*, c^*)$ (cf. the proof of [4], Proposition 2.3, (4)). Moreover, the following important inequality is satisfied:

$$F(q_F^-(\alpha) - \delta) < \alpha \leq F(q_F^-(\alpha) + \varepsilon)$$

(2.2)

for all $\alpha \in (c_*, c^*)$, $\delta > 0$, and for all $\varepsilon > 0$. Hence,

$$F(q_F^-(\alpha) -) \leq \alpha \leq F(q_F^-(\alpha) +) = F(q_F^+(\alpha))$$

(2.3)

for all $\alpha \in (c_*, c^*)$. Also recall that $\{x \in \mathbb{R} : F(x) \geq \alpha\} = [q_F^-(\alpha), \infty)$, respectively $\{x \in \mathbb{R} : F(x) < \alpha\} = (-\infty, q_F^-(\alpha))$ for any $\alpha \in (c_*, c^*)$.

By taking a closer look at $q_F^-(F_\lambda(x))$, we firstly obtain the following two statements:

**Remark 2.3.** Let $0 \leq \lambda < 1$ and $c_* < F_\lambda(x) < c^*$. Then

$q_F^-(F_\lambda(x)) \leq x$.

**Proof.** Fix $x \in \mathbb{R}$ and $0 \leq \lambda < 1$. Put $\alpha := F_\lambda(x)$. By assumption, $c_* < \alpha < c^*$. Thus, $q_F^-(\alpha)$ is well-defined. Since $F(x) \geq F_\lambda(x) = \alpha$, the claim follows. \hfill $\square$

**Remark 2.4.** Let $x \in \mathbb{R}$. Then

$$\left( F(x^-), F(x) \right) \subseteq \left\{ \alpha \in (c_*, c^*) : x = q_F^-(\alpha) \right\} \subseteq \left[ F(x^-), F(x) \right].$$

In particular,

$q_F^-(F_\lambda(x)) = x$

if $\Delta F(x) > 0$ and $0 < \lambda < 1$.

**Proof.** Let $\alpha \in (c_*, c^*)$ such that $x = q_F^-(\alpha)$. Due to (2.3) it follows that

$$F(x^-) \leq \alpha \leq F(x),$$

which gives the second inclusion. To prove the first inclusion, we may assume without loss of generality that $F$ is not continuous in $x$. So, let $F(x^-) < \alpha < F(x)$. Then $c_* < \alpha < c^*$ (else we would obtain the contradiction $\alpha \leq c_* \leq F(x^-)$, respectively $F(x) \leq c^* \leq \alpha$) and $F(x - \frac{1}{n}) < \alpha \leq F(x)$ for all $n \in \mathbb{N}$. Hence, $q_F^-(\alpha) \leq x < q_F^-(\alpha) + \frac{1}{n}$ for all $n \in \mathbb{N}$ (cf. [4], Proposition 2.3, (5)), implying the first inclusion. \hfill $\square$
Let us fix the function $F : \mathbb{R} \to [c_\ast, c^\ast]$, defined above. Throughout the remaining part of our paper, we follow the notation of [7] and put $\xi := q_F(\alpha)$.

**Definition 2.5.** Let $\alpha \in (c_\ast, c^\ast)$ and $0 \leq \lambda \leq 1$. Put:

$$A_{\lambda,\alpha} := \{ x \in \mathbb{R} : F_\lambda(x) \leq \alpha \}.$$ 

Firstly note that $A_{\lambda,\alpha}$ is non-empty. To see this, consider any $x < \xi = q_F(\alpha)$. Then $x \leq \xi - \delta$ for some $\delta > 0$. Hence, $F_\lambda(x) \leq F(x) < \alpha$. To motivate the following representation of the set $A_{\lambda,\alpha}$, let us assume for the moment that $F$ is continuous at $\xi$. Due to (2.3), it follows that $F(\xi) = \alpha$. Hence, in this case, $F_\lambda(\xi) = F(\xi) = \alpha$, implying that $\xi = q_F(\alpha) \in A_{\lambda,\alpha}$.

However, in the general (non-continuous) case, $\xi = q_F(\alpha)$ need not be an element of the set $A_{\lambda,\alpha}$. Therefore (by fixing $\alpha \in (c_\ast, c^\ast)$ and $0 \leq \lambda \leq 1$), we are going to represent the set $A_{\lambda,\alpha}$ as a disjoint union of the following three subsets of the real line:

$$A^+_{\lambda,\alpha} := A_{\lambda,\alpha} \cap \{ x \in \mathbb{R} : x > \xi \},$$

$$A^-_{\lambda,\alpha} := A_{\lambda,\alpha} \cap \{ x \in \mathbb{R} : x = \xi \},$$

and

$$A^-_{\lambda,\alpha} := A_{\lambda,\alpha} \cap \{ x \in \mathbb{R} : x < \xi \}.$$ 

Thus,

$$A_{\lambda,\alpha} = A^+_{\lambda,\alpha} \cup A^-_{\lambda,\alpha} \cup A^-_{\lambda,\alpha}. $$

Next, we are going to simplify the sets $A^+_{\lambda,\alpha}, A^-_{\lambda,\alpha}$ and $A^-_{\lambda,\alpha}$ as far as possible. To this end, we have to analyse carefully the jump $\Delta q_F(\alpha)$, implying that we have to check $\xi = q_F(\alpha) = q_F(\alpha)$ against the real number

$$\eta := q^+_F(\alpha) = \inf \{ x \in \mathbb{R} : F(x) > \alpha \} = \sup \{ x \in \mathbb{R} : F(x) \leq \alpha \} = q^-_F(\alpha).$$

The inequality (2.3) is also satisfied for $\eta$ (cf. [5], Lemma A.15): 

$$F(\eta) \leq \alpha \leq F(\eta).$$

Note that if $F$ were a distribution function, $\eta$ (respectively $\xi$) would be precisely the right (respectively left) $\alpha$-quantile of $F$.

Clearly, $\{ x \in \mathbb{R} : x > \xi \}$ and $F(x) = \alpha \subseteq A^+_{\lambda,\alpha}$ for every $\lambda \in [0, 1]$. However, if $0 < \lambda \leq 1$, we even obtain equality of both sets - since:

**Lemma 2.6.** Let $0 < \lambda \leq 1$ and $\alpha \in (c_\ast, c^\ast)$. Put $\xi := q^-_F(\alpha)$ and $\eta := q^+_F(\alpha)$.

(i) If $\xi < \eta$, then $F(\xi) = \alpha = F(\eta)$ and $\emptyset \neq \{ x \in \mathbb{R} : x > \xi \}$ and $F(x) = \alpha$. Moreover, the restricted function $F|_{A^+_{\lambda,\alpha}} : A^+_{\lambda,\alpha} \to \mathbb{R}$ is continuous, and

$$A^+_{\lambda,\alpha} = \{ x \in \mathbb{R} : x > \xi \} \text{ and } F(x) = \alpha = \begin{cases} (\xi, \eta) & \text{if } F(\eta) > \alpha \\ (\xi, \eta) & \text{if } F(\eta) = \alpha \end{cases}$$

\footnote{In [5], $q^+_F$ is called right-continuous inverse function of $F$.}
(ii) If $\xi = \eta$, then $A_{\lambda,\alpha}^+ = \emptyset$.

In particular, the following statements are equivalent:

(a) $0 < \Delta q_F'(\alpha) = \eta - \xi$;

(b) \{ $x \in \mathbb{R} : x > \xi$ and $F(x) = \alpha$ \} $\neq \emptyset$.

Proof. Put $B := \{ x \in \mathbb{R} : x > \xi$ and $F(x) = \alpha$ \}. Clearly, we always have $B \subseteq A_{\lambda,\alpha}^+$.

To verify (i), let $\xi < \eta$. Then $\xi < z_0 < \eta = \inf\{ x \in \mathbb{R} : F(x) > \alpha \}$ for some $z_0 \in \mathbb{R}$. Thus, $F(\xi) \leq F(z_0) \leq \alpha \leq F(\xi)$, implying that $z_0 \in B$ and $F(\xi) = \alpha$. Assume by contradiction that $F(\eta-) < \alpha$. Then $F(\eta - \varepsilon) < F(\xi)$ for all $\varepsilon > 0$, implying the contradiction $\eta \leq \xi$. Hence, $F(\eta-) = \alpha$.

Let $x \in A_{\lambda,\alpha}^+ \supseteq B$. Assume by contradiction that $F|_{A_{\lambda,\alpha}^+}$ is not continuous at $x$. Then $F(x-) < F(x-) + \lambda \Delta F(x) = F_\alpha(x) \leq \alpha$ (since $\lambda > 0$). Since $x > \xi$, we have $\xi \leq x - \frac{1}{n}$ for some $n \in \mathbb{N}$. Thus,

$$\alpha \leq F(\xi) \leq F\left(\xi + \frac{1}{2n}\right) \leq F\left(x - \frac{1}{2n}\right).$$

Hence, $\alpha \leq F(x-) < \alpha$, which is a contradiction. Thus, the restricted function $F|_{A_{\lambda,\alpha}^+}$ is continuous on $A_{\lambda,\alpha}^+$. Let $u \in A_{\lambda,\alpha}^+$. Since $F$ is continuous at $u$, it follows that

$$\alpha \leq F(\xi) \leq F(u) = F_\alpha(u) \leq \alpha.$$

Thus, $\emptyset \neq A_{\lambda,\alpha}^+ = B$.

To prove (ii), suppose that $A_{\lambda,\alpha}^+$ is non-empty. The previous calculations show that the existence of an element $u_0 \in A_{\lambda,\alpha}^+$ already implies $F(\xi) = F(u_0) = \alpha$. Consequently, \( \eta = q_F^+(\alpha) = \sup\{ x \in \mathbb{R} : F(x) \leq \alpha \} \) cannot coincide with $\xi = q_F^{-}(\alpha)$ (since $\xi < u_0 \leq \eta$), implying that $\xi < \eta$.

To finish the proof of (i), we have to verify (2.5). To this end, let $\xi < \eta$ and $x \in (\xi, \eta)$. Then there exists $\delta > 0$ such that $\xi < x - \delta < x < x + \delta < \eta = q_F^{+}(\alpha) = \inf\{ u \in \mathbb{R} : F(u) > \alpha \}$.

Consequently, $\alpha \leq F(\xi) \leq F(x - \delta) \leq F(x) \leq F(x + \delta) \leq \alpha$. Thus,

$$\{ (\xi, \eta) \} \subseteq \{ x \in \mathbb{R} : x > \xi$ and $F(x) = \alpha \} = B.$$

Moreover, [4], Proposition 2.3, (6) implies that

$$B = \{ x \in \mathbb{R} : x > \xi$ and $F(x) = \alpha \} \subseteq (\xi, \eta).$$

Hence,

$$(\xi, \eta) \subseteq B \subseteq (\xi, \eta).$$

If $F(\eta) > \alpha$, then $\eta \notin B$ and hence $B = (\xi, \eta)$. If $F(\eta) = \alpha$, then $\xi < \eta \in B$ and hence $B = (\xi, \eta)$. □

Regarding a visualisation of Lemma 2.6 consider the set $M_\alpha := \{ x \in \mathbb{R} : x \leq \xi$ and $F(x) = \alpha \}$ \( \overset{\text{(2.2)}}{\subseteq} \{ x \in \mathbb{R} : x = \xi$ and $F(x) = \alpha \} \in \{ \emptyset, \{ \xi \} \} \). Since

$$\{ x \in \mathbb{R} : F(x) = \alpha \} = \{ x \in \mathbb{R} : x > \xi$ and $F(x) = \alpha \} \cup M_\alpha,$$

Lemma 2.6 immediately implies the following precise mathematical description of the (preimages of) “flat pieces” of $F$:
Corollary 2.7. Let $0 < \lambda \leq 1$ and $\alpha \in (c_{s}, c^{*})$. Put $\xi := q_{F}^{-}(\alpha)$ and $\eta := q_{F}^{+}(\alpha)$.

(i) If $\xi < \eta$, then
\[
\emptyset \neq \{x \in \mathbb{R} : F(x) = \alpha\} = A_{\lambda, \alpha}^{+} \cup \{\xi\} = \begin{cases} 
[\xi, \eta) & \text{if } F(\eta) > \alpha \\
[\xi, \eta] & \text{if } F(\eta) = \alpha
\end{cases}
\]

(ii) If $\xi = \eta$, then
\[
\{x \in \mathbb{R} : F(x) = \alpha\} = \begin{cases} 
\emptyset & \text{if } \eta > F(\eta) > \alpha \\
\{\xi\} & \text{if } \eta = F(\eta) = \alpha
\end{cases}
\]

In particular, $F(\xi) = \alpha$ if and only if $\{x \in \mathbb{R} : F(x) = \alpha\} \neq \emptyset$. Moreover, $\eta - \xi = \Delta q_{F}^{-}(\alpha) > 0$ if and only if $\{\xi\} \subsetneq \{x \in \mathbb{R} : F(x) = \alpha\}$, and if $\eta > \xi$, then $\Delta F(\eta) = 0$ if and only if $F(\eta) = \alpha$.

Let $\mathcal{B}(\mathbb{R})$ denote the set of all Borel subsets of $\mathbb{R}$. In the following, let $\mu_{F} : \mathcal{B}(\mathbb{R}) \rightarrow [0, \infty]$ be the Lebesgue-Stieltjes measure of $F$. For a detailed description of the construction and properties of the Lebesgue-Stieltjes measure (including Lebesgue-Stieltjes integration), we refer the reader to e. g. [1] and [2]. For the convenience of the reader, we recall the following fundamental result (cf. [2], Theorem 12.4):

Theorem 2.8 (Lebesgue-Stieltjes measure). If $G : \mathbb{R} \rightarrow \mathbb{R}$ is an non-decreasing and right-continuous function, then there exists a unique Borel measure $\mu_{G}$ satisfying
\[
\mu_{G}((x, y]) = G(y) - G(x)
\]
for all $x, y \in \mathbb{R}$.

Clearly, this crucial result implies that $\mu_{G}((x, y]) = G(y-) - G(x)$ and hence
\[
\mu_{G}(\{y\}) = \mu_{G}((x, y]) - \mu_{G}((x, y]) = G(y) - G(y-) = \Delta G(y)
\]
for all $y \in \mathbb{R}$.

Returning to our function $F$, a direct application of $\mu_{F}$ leads to another important implication of Lemma 2.6:

Corollary 2.9. Let $0 < \lambda \leq 1$ and $\alpha \in (c_{s}, c^{*})$. Then $A_{\lambda, \alpha}^{+} \in \mathcal{B}(\mathbb{R})$, and
\[
\mu_{F}(A_{\lambda, \alpha}^{+}) = 0.
\]

In particular, if $\xi < \eta$, then
\[
\mu_{F}(\{x \in \mathbb{R} : F(x) = \alpha\}) = \Delta F(\xi) = \alpha - F(\xi-).
\]

Proof. Nothing is to prove if $A_{\lambda, \alpha}^{+} = \emptyset$. So, let $A_{\lambda, \alpha}^{+} \neq \emptyset$. Then $\eta - \xi = \Delta q_{F}^{-}(\alpha) > 0$. Suppose first that $F(\eta) > \alpha$. Then
\[
A_{\lambda, \alpha}^{+} = (\xi, \eta) = \bigcup_{n=1}^{\infty} (\xi, \eta - \frac{1}{n})
\]
Consequently, since in general $F(x) = \alpha = F(\xi)$ for all $x \in (\xi, \eta)$, it follows that

$$
\mu_F(A_{\lambda,\alpha}^+) = \lim_{n \to \infty} \mu_F\left(\left(\xi, \eta - \frac{1}{n}\right)\right) = \lim_{n \to \infty} \left(F(\eta - \frac{1}{n}) - F(\xi)\right) = \alpha - \alpha = 0.
$$

Now suppose that $F(\eta) = \alpha$. Then $\eta \in A_{\lambda,\alpha}^+$, and it follows that $F$ is continuous at $\eta$. Thus, $\mu_F(\{\eta\}) = \Delta F(\eta) = 0$. Since in this case

$$
A_{\lambda,\alpha}^+ = (\xi, \eta) \cup \{\eta\},
$$

it consequently follows that

$$
\mu_F(A_{\lambda,\alpha}^+) = \lim_{n \to \infty} \left(F(\eta - \frac{1}{n}) - F(\xi)\right) + \mu_F(\{\eta\}) = \alpha - \alpha + 0 = 0.
$$

\[\square\]

Next, we are going to reveal in detail that the function $F$ is almost “invertible” at every $x \in \mathbb{R}$ which does not belong to the preimage $A_{\lambda,\alpha}^+$ of a “flat piece” of $F$. More precisely:

**Theorem 2.10.** Let $0 < \lambda \leq 1$. Assume that $\mu_F(\mathbb{R}) > 0$, $\mu_F(\{x \in \mathbb{R} : F_\lambda(x) = c_*\}) = 0$ and $\mu_F(\{x \in \mathbb{R} : F_\lambda(x) = c^*\}) = 0$, then

$$
x = q^-_F(F_\lambda(x)) \quad \mu_F\text{-almost everywhere}.
$$

**Proof.** Let $0 < \lambda \leq 1$. Consider the Borel set

$$
N_\lambda := \{x \in \mathbb{R} : F_\lambda(x) = c_*\} \cup \{x \in \mathbb{R} : F_\lambda(x) = c^*\} \cup \bigcup_{\alpha \in J_{q^-_F}} A_{\lambda,\alpha}^+,
$$

where $J_{q^-_F} := \{\alpha \in (c_*, c^*) : \Delta q^-_F(\alpha) > 0\}$ denotes the set of all jumps of the function $q^-_F$. Since the (left-continuous) function $q^-_F : (c_*, c^*) \to \mathbb{R}$ is non-decreasing, $J_{q^-_F}$ is at most countable. Hence, if $J_{q^-_F} \neq \emptyset$, there exists a subset $M$ of $\mathbb{N}$, and a sequence $(\alpha_n)_{n \in M}$, consisting of pairwise distinct elements $\alpha_n \in J_{q^-_F}$, such that $J_{q^-_F} = \{\alpha_n : n \in \mathbb{N}\}$. Thus,

$$
\bigcup_{\alpha \in J_{q^-_F}} A_{\lambda,\alpha}^+ = \bigcup_{n \in M} A_{\lambda,\alpha_n}.
$$

Corollary 2.9 therefore implies that - in any case - $\mu_F(N^\lambda) = 0$ and hence $\mathbb{R} \setminus N^\lambda \neq \emptyset$ (since $\mu_F(\mathbb{R}) > 0$ by assumption).

Let $x \in \mathbb{R} \setminus N^\lambda$. Put $\alpha(x) := F_\lambda(x)$. Then $c_* < \alpha(x) < c^*$, and $x \in \bigcap_{\alpha \in J_{q^-_F}} (\mathbb{R} \setminus A_{\lambda,\alpha}^+)$. Thus, $\xi(x) := q^-_F(\alpha(x))$ is well-defined. Consider $\eta(x) := q^+_F(\alpha(x)) = q^-_F(\alpha(x) +)$. First, let $J_{q^-_F} = \emptyset$. Then $\xi(x) = \eta(x)$. Lemma 2.6 therefore implies that $A_{\lambda,\alpha(x)}^+ = \emptyset$. In particular, $x \notin A_{\lambda,\alpha(x)}^+$. Hence, since $F_\lambda(x) = \alpha(x) \neq \alpha(x)$, it consequently follows that

$$
x \leq \xi(x) = q^-_F(\alpha(x)) = q^-_F(F_\lambda(x)),
$$

and hence $x = q^-_F(F_\lambda(x))$.

\[\text{Note that by construction } N^\lambda = \{x \in \mathbb{R} : F_\lambda(x) = c_*\} \cup \{x \in \mathbb{R} : F_\lambda(x) = c^*\} \text{ if } J_{q^-_F} = \emptyset.\]
Now let $J_{q_F} \neq \emptyset$. If $\alpha(x) \notin J_{q_F}$, it follows again that $\xi(x) = \eta(x)$ and hence
\[ x \leq \xi(x) = q_F^{-}(\alpha(x)) = \bar{q}_F^{-}(F\lambda(x)) \leq x, \]
as above. So, let $\alpha(x) \in J_{q_F}$. Then $\alpha(x) = \alpha_m$ for some $m \in \mathbb{M}$, and hence $A_{\alpha_m}^+ = A_{\alpha_m}^+$. Since $x \in \mathbb{R} \setminus N_\lambda \subseteq \mathbb{R} \setminus A_{\alpha_m}^{\pm}$, it follows once more again that $x \leq \xi(x) = q_F^{-}(\alpha(x))$, and hence
\[ x = \xi(x) = q_F^{-}(\alpha(x)) = \bar{q}_F^{-}(F\lambda(x)). \]

Next, we consider the set $A_{\lambda,\alpha}^-$. Again, in line with [7], we put $q := F(\xi-) = F(\xi) \geq \lambda \geq 0$. Then
\[ q + \beta = F(\xi) \geq \alpha \geq q. \]
Obviously, we may write:

**Remark 2.11.** $A_{\lambda,\alpha}^- = \{x \in \mathbb{R} : x = \xi and \beta \lambda \leq \alpha - q\}.$

Moreover, by using a similar argument like that one which has shown us that the set $A_{\lambda,\alpha}$ is non-empty, we further obtain

**Remark 2.12.** $A_{\lambda,\alpha}^- = (-\infty, \xi) = \{x \in \mathbb{R} : x < \xi\}.$

Observe that only the subset $A_{\lambda,\alpha}^-$ of $A_{\lambda,\alpha}$ does depend on the choice of $0 \leq \lambda \leq 1$.

### 2.2. The inclusion of randomness

In addition to our assumptions above, we now fix a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $X : \Omega \to \mathbb{R}$ and $V : \Omega \to \mathbb{R}$ be two given random variables (on this probability space) such that $V \sim U(0,1)$ is uniformly distributed over $(0,1)$ and independent of $X$. Let $0 < u \leq 1$ be given. Possibly, we might have to substitute $V$ canonically through the random variable $V_u$ defined on $\Omega$ via
\[
V_u(\omega) := \begin{cases} 
    u & \text{if } V(\omega) \leq 0 \\
    V(\omega) & \text{if } 0 < V(\omega) < 1 \\
    1 & \text{if } V(\omega) \geq 1
\end{cases}
\]
$V_u$ clearly satisfies $V_u(\Omega) \subseteq (0,1]$ and $V_u \sim U(0,1)$. Moreover, $V_u$ is also independent of $X$. Hence, without loss of generality, we may assume that already $V(\Omega) \subseteq (0,1]$. Let us consider the (now well-defined) random variable $F_V(X)$, defined on $\Omega$ via
\[
F_V(X)(\omega) := F_V(\omega)(X(\omega)) = F_\lambda(x),
\]
where here $\omega \in \Omega$, $\lambda := V(\omega)$ and $x := X(\omega)$. Next, we have to evaluate $\mathbb{P}(F_V(X) \leq \alpha)$; i.e., we wish to calculate
\[
\mathbb{P}(F_V(X) \leq \alpha) = \mathbb{P}\left(\{\omega \in \Omega : X(\omega) \in A_{V(\omega),\alpha}^+\}\right).
\]
Due to our previous observations, we have
\[
A_{V(\omega),\alpha}^+ = A_{V(\omega),\alpha}^+ \cup A_{V(\omega),\alpha}^- \cup A_{V(\omega),\alpha}^-
\]
for all \( \omega \in \Omega \). Consequently, given the assumed independence of \( V \) and \( X \), Lemma 2.6 implies that\(^5\):

\[
\mathbb{P}(F_V(X) \leq \alpha) = \mathbb{P}(X \in A^+_{V,\alpha}) + \mathbb{P}(X \in A^{-}_{V,\alpha}) + \mathbb{P}(X \in A^*_{V,\alpha}) = \mathbb{P}(X \in A^+_{V,\alpha}) + \mathbb{P}(X = \xi) \cdot \beta \mathbb{P}(\beta V \leq \alpha - q) + \mathbb{P}(X \in A^{-}_{V,\alpha}) = \mathbb{P}(X > \xi \text{ and } F(X) = \alpha) + \mathbb{P}(X = \xi) \cdot \beta \mathbb{P}(\beta V \leq \alpha - q) + \mathbb{P}(X < \xi).
\]

Remark 2.13. Observe, that we have not yet assumed that \( F \) is a distribution function.

Apparently, to continue with the calculation of the respective probabilities, we have consider the following two possible cases: \( \beta = 0 \) and \( \beta > 0 \):

(i) Let \( \beta = 0 \). Thus, since \( \alpha - q \geq 0 \), it follows that

\[
\mathbb{P}(F_V(X) \leq \alpha) = \mathbb{P}(X > \xi \text{ and } F(X) = \alpha) + \mathbb{P}(X \leq \xi)
\]

(ii) Let \( \beta > 0 \). Since \( V \sim U(0, 1) \) is uniformly distributed over \((0, 1)\), we have \( \mathbb{P}(\beta V \leq \alpha - q) = \frac{\alpha - q}{\beta} \). Hence, since \( \frac{\alpha - q}{\beta} - 1 = \frac{\alpha - F(\xi)}{\beta} \), it follows that

\[
\mathbb{P}(F_V(X) \leq \alpha) = \mathbb{P}(X > \xi \text{ and } F(X) = \alpha) + \left(\frac{\alpha - F(\xi)}{\beta}\right) \mathbb{P}(X = \xi) + \mathbb{P}(X \leq \xi).
\]

Moreover, by taking into account that \( F(\xi) = \alpha \) in case (i) (since \( F \) is continuous at \( \xi \) if \( \beta = 0 \)), we have arrived at the following important

Lemma 2.14. Suppose that \( F : \mathbb{R} \rightarrow \mathbb{R} \) satisfies Assumption 2.2. Let \( \alpha \in (c_*, c^*) \). Put \( \xi := q_F(\alpha) \) and \( \beta := \Delta F(\xi) \). Let \( X, V \) be two random variables, both defined on the same probability space \((\Omega, F, \mathbb{P})\), such that \( V \sim U(0, 1) \) and \( V \) is independent of \( X \). Then

\[
\mathbb{P}(F_V(X) \leq \alpha) - \alpha = \mathbb{P}(X > \xi \text{ and } F(X) = \alpha) + c_\beta(\mathbb{P}(X = \xi) - \beta) + (\mathbb{P}(X \leq \xi) - F(\xi)),
\]

where \( c_\beta := 0 \) if \( \beta = 0 \) and \( c_\beta := \frac{\alpha - F(\xi)}{\beta} \) if \( \beta \neq 0 \).

To conclude, let us slightly point towards the fact that Lemma 2.14 could also be viewed as a building block of a probabilistic limit theorem (whose detailed discussion would then exceed the main goal of this paper, though).

2.3. The role of the distribution function of \( X \)

From now on, we assume that \( F := F_X = \mathbb{P}(X \leq -) \) already is the distribution function of the random variable \( X \). Consequently, we are going to work with \( c_* = 0 \) and \( c^* = 1 \). Moreover, since \( \lim_{x \to -\infty} F(x) = 0 = c_* \) and \( \lim_{x \to \infty} F(x) = 1 = c^* \), it follows that in fact for any \( 0 < \alpha < 1 \) the set \( \{ x \in \mathbb{R} : F(x) \geq \alpha \} \) is non-empty and bounded from below, implying that \( |q_F(\alpha)| < \infty \) for all \( \alpha \in (0, 1) \).

\(^5\)Here, \( \{ X \in A_{V,\alpha} \} := \{ \omega \in \Omega : X(\omega) \in A_{V(\omega),\alpha} \} \) and \( \{ X \in A^*_{V,\alpha} \} := \{ \omega \in \Omega : X(\omega) \in A^*_{V(\omega),\alpha} \} \), where \( i \in \{+,-,\sim\} \).
**Proposition 2.15.** Let $X, V$ be two random variables, both defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that $V \sim U(0, 1)$ and $V$ is independent of $X$. Let $F = F_X$ be the distribution function of $X$. Then $F_V(X) \sim U(0, 1)$ is a uniformly distributed random variable. Moreover,

$$\mathbb{P}(F(X) \leq \alpha) = \alpha = \mathbb{P}(X \leq q_F(\alpha))$$
onumber

on the set $\{ \alpha \in (0, 1) : q_F(\alpha) < q_F^+(\alpha) \}$.

**Proof.** Let $0 < \alpha < 1$. Lemma 2.14 - applied to $F = F_X$ - directly leads to

$$\mathbb{P}(F_V(X) \leq \alpha) - \alpha = \mathbb{P}(X > \xi \text{ and } F(X) = \alpha) .$$

Corollary 2.9 further implies that for any $0 < \lambda \leq 1$ we have

$$\mathbb{P}(X > \xi \text{ and } F(X) = \alpha) = \mathbb{P}(X \in A_{\lambda, \alpha}) = \mu_F(A_{\lambda, \alpha}) = 0 .$$

Thus, we have

$$\mathbb{P}(F_V(X) \leq \alpha) = \alpha \text{ for any } 0 < \alpha < 1 . \quad \text{(2.6)}$$

Consequently, $\sigma$-additivity of the probability measure $\mathbb{P}$ allows one to continuously extend (2.6) to the whole real line. Hence, $F_V(X) \sim U(0, 1)$ is uniformly distributed.

Now let $\alpha \in (0, 1)$ such that $\xi := q_F(\alpha) < q_F^+(\alpha) =: \eta$. Thanks to Corollary 2.9, it follows that

$$\mathbb{P}(F(X) = \alpha) = \mu_F(F = \alpha) = \alpha - F(\xi^{-}) = \mathbb{P}(X = \xi) .$$

Since always

$$\mathbb{P}(F(X) < \alpha) = \mathbb{P}(X < \xi) ,$$

it follows that

$$\mathbb{P}(F(X) \leq \alpha) = \mathbb{P}(X \leq \xi) = F(\xi) = \alpha ,$$

and we are done. \qed

In order to complete the proof of statement of Proposition 2.1 in [7], let us recall that the assumed independence of the random variables $X$ and $V$ implies that the bivariate distribution function $F_{(V,X)}$ of the random vector $(V, X)$ coincides with the product of the distribution functions $F_V$ and $F_X$. Moreover, since $V \sim U(0, 1)$, $F_V(C) = \mathbb{P}(V \in C)$ clearly coincides with the Lebesgue measure of $C$ for all Borel sets $C \in \mathcal{B}((0, 1])$. Hence, if $\Phi : (0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$ denotes an arbitrary non-negative (or bounded) Borel function on $((0, 1] \times \mathbb{R}, \mathcal{B}((0, 1]) \times \mathcal{B}(\mathbb{R}))$, an immediate application of the Fubini-Tonelli Theorem leads to

$$\mathbb{E}[\Phi(V, X)] = \int_0^1 \mathbb{E}[\Phi(\lambda, X)] \, d\lambda = \int_0^1 \left( \int_{\mathbb{R}} \Phi(\lambda, x) \mu_F(\text{d}x) \right) \, d\lambda \quad \text{(2.7)}$$

**Theorem 2.16.** Let $X, V$ be two random variables, defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that $V \sim U(0, 1)$ and $V$ is independent of $X$. Let $F = F_X$ be the distribution function of the random variable $X$. Then

$$X = q_F^-(F_V(X)) \quad \mathbb{P}\text{-almost surely.}$$

In particular, if in addition $\mathbb{P}(F(X) = 0) = 0$ and $\mathbb{P}(F(X) = 1) = 0$, then

$$X = q_F^-(F(X)) \quad \mathbb{P}\text{-almost surely.}$$
Proof. Since $F$ is the distribution function of $X$, it follows that $\mu_F = \mathbb{P}(X \in \cdot)$ and hence $\mu_F(\mathbb{R}) = 1 > 0$. Let $B_\lambda := \{x \in \mathbb{R} : F_\lambda(x) = 0\}$, where $0 < \lambda \leq 1$. Then
\[
\mathbb{P}(F_\lambda(X) = 0) = \mathbb{E}[\mathbb{1}_{B_\lambda}(X)].
\]
On the other hand, equality 2.7 clearly implies
\[
\mathbb{P}(F_\lambda(X) = 0) = \int_0^1 \mathbb{E}[\mathbb{1}_{B_\lambda}(X)] \, d\lambda.
\]
Hence, since $F_\lambda(X) \sim U(0,1)$, it follows that $\int_0^1 \mathbb{E}[\mathbb{1}_{B_\lambda}(X)] \, d\lambda = 0$, implying that for (Lebesgue-)almost all $\lambda \in (0,1)$ we have
\[
\mu_F(F_\lambda = 0) = \mathbb{P}(F_\lambda(X) = 0) = \mathbb{E}[\mathbb{1}_{B_\lambda}(X)] = 0.
\]
Similarly, we obtain
\[
\mu_F(F_\lambda = 1) = 0
\]
for (Lebesgue-)almost all $\lambda \in (0,1]$. Hence, outside of a Borel set $L \in \mathcal{B}(0,1]$ of Lebesgue-measure 0, $\mu_F = \mathbb{P}(X \in \cdot)$ satisfies all requirements of Theorem 2.10.

Thus, given the construction in the proof of Theorem 2.10, it follows that for all $\lambda \in \mathbb{R} \setminus L$ there exists a $\mu_F$-Borel null set $N_\lambda$, such that for any $x \in \mathbb{R} \setminus N_\lambda$ the value $q_F^{-1}(F_\lambda(x))$ is well-defined and satisfies $q_F^{-1}(F_\lambda(x)) = x$. Hence, since
\[
\mathbb{P}(X \in N_V) \overset{(2.7)}{=} \int_0^1 \mathbb{P}(X \in N_\lambda) \, d\lambda = \int_{\mathbb{R} \setminus L} \mathbb{P}(X \in N_\lambda) \, d\lambda = \int_{\mathbb{R} \setminus L} \mu_F(N_\lambda) \, d\lambda = 0,
\]
it consequently follows $X \equiv q_F^{-1}(F_\lambda(X))$ on the set $\Omega \setminus (\{V \in L\} \cup \{X \in N_V\})$. \hfill \square

For the convenience of the reader, we conclude our paper with a full and - from now on - very straightforward proof of the general version of Sklar’s Theorem, complemented with a nice (and seemingly novel) observation, induced by Lemma 2.6.

**Corollary 2.17** (Sklar’s Theorem). Let $n \in \mathbb{N}$ and $F_{(X_1,\ldots,X_n)}$ be a joint $n$-variate distribution function of a random vector $(X_1, X_2, \ldots, X_n) : \Omega \rightarrow \mathbb{R}^n$ with marginals $F_i := F_{X_i}$ ($i = 1,2,\ldots,n$). Then there exist a copula $c$ such that for all $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$
\[
F_{(X_1,\ldots,X_n)}(x_1, x_2, \ldots, x_n) = c(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)).
\]
If all $F_i$ are continuous, then the copula $c$ is unique. Otherwise, $c$ is uniquely determined on $\prod_{i=1}^n F_i(\mathbb{R})$. Conversely, if $c$ is a copula and $H_1, H_2, \ldots, H_n$ are distribution functions, then the function $F$ defined by
\[
F(x_1, x_2, \ldots, x_n) := c(H_1(x_1), H_2(x_2), \ldots, H_n(x_n))
\]
is a joint distribution function with marginals $H_1, H_2, \ldots, H_n$.  

12
Proof. Let \( i \in \{1, 2, \ldots, n\} \). Put \( U_i := V_i \Delta F_i + F_i^- \). According to Theorem 2.16 there exist null sets \( M_1, M_2, \ldots, M_n \in \mathcal{F} \), such that on \( \Omega \setminus \bigcup_{i=1}^n M_i := q_{F_i^+}(U_i) \) is well-defined and satisfies \( X_i \equiv Z_i \) for every \( i \in \{1, 2, \ldots, n\} \). Thus, \( \mathbb{P}(M) = 0 \), where \( M := \bigcup_{i=1}^n M_i \).

Let \( F(x_1, \ldots, x_n) := \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) \) denote the \( n \)-variate distribution function of the random vector \( (X_1, X_2, \ldots, X_n) \). Consider the copula

\[
c(\gamma_1, \gamma_2, \ldots, \gamma_n) := \mathbb{P}(U_1 \leq \gamma_1, U_1 \leq \gamma_2, \ldots, U_n \leq \gamma_n),
\]

where \( (\gamma_1, \gamma_2, \ldots, \gamma_n) \in [0, 1]^n \). Since

\[
\{u \in (0, 1) : q_{F_i}(u) \leq x_i\} = \{u \in (0, 1) : u \leq F_i(x_i)\}
\]

for all \( i \in \{1, 2, \ldots, n\} \) and \( \mathbb{P}(M) = 0 \), it consequently follows

\[
c(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) = \mathbb{P}\left(\{Z_1 \leq x_1, Z_2 \leq x_2, \ldots, Z_n \leq x_n\} \cap \mathbb{R} \setminus M\right)
\]

\[
= \mathbb{P}\left(\{X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n\} \cap \mathbb{R} \setminus M\right) = F(x_1, \ldots, x_n)(x_1, x_2, \ldots, x_n).
\]

\( \square \)

Combining Sklar’s Theorem with Lemma 2.6, we immediately obtain another interesting result:

**Remark 2.18.** Let \( (\alpha_1, \alpha_2, \ldots, \alpha_n) \in (0, 1)^n \), satisfying \( q_{F_i}^{-1}(\alpha_i) < q_{F_i}^+(\alpha_i) \) for all \( i \in \{1, 2, \ldots, n\} \). Then

\[
c(\alpha_1, \alpha_2, \ldots, \alpha_n) = F(x_1, x_2, \ldots, x_n)\left(q_{F_1}^-(\alpha_1), q_{F_2}^-(\alpha_2), \ldots, q_{F_n}^-(\alpha_n)\right).
\]

**References**


