A note on generalized inverses

Paul Embrechts[1], Marius Hofert[2]

2013-04-24

Abstract

Motivated by too restrictive or even incorrect statements about generalized inverses in the literature, properties about these functions are investigated and proven. Examples and counterexamples show the importance of generalized inverses in mathematical theory and its applications.

Keywords
Increasing function, generalized inverse, distribution function, quantile function.

MSC2010
60E05, 62E15, 26A48.

1 Introduction

It is well known that a real-valued, continuous, and strictly monotone function of a single variable possesses an inverse on its range. It is also known that one can drop the assumptions of continuity and strict monotonicity (even the assumption of considering points in the range) to obtain the notion of a generalized inverse. Generalized inverses play an important role in probability theory and statistics in terms of quantile functions, and in financial and insurance mathematics, for example, as Value-at-Risk or return period. Generalized inverses of increasing functions which are not necessarily distribution functions also frequently appear, for example, as transformations of random variables. In particular, proving the famous invariance principle of copulas under strictly increasing transformations on the ranges of the underlying random variables involves such transformations.

One can often work with generalized inverses as one does with ordinary inverses. To see this, one has to have several properties about generalized inverses at hand. Although these properties are often stated in the literature, one rarely finds detailed proofs of these results. Moreover, some of the statements found and often referred to are incorrect.

The main goal of this paper is therefore to state and prove important properties about generalized inverses of increasing functions. Furthermore, examples which stress their

[1]RiskLab, Department of Mathematics and Swiss Finance Institute, ETH Zurich, 8092 Zurich, Switzerland, embrechts@math.ethz.ch

[2]RiskLab, Department of Mathematics, ETH Zurich, 8092 Zurich, Switzerland, marius.hofert@math.ethz.ch. The author (Willis Research Fellow) thanks Willis Re for financial support while this work was being completed.

importance are presented and counterexamples for incorrect statements in the literature are given.

Before we begin, let us point out some references. Klement et al. (1999) introduce pseudo- and quasi-inverses and provide properties of such functions. Some properties of generalized inverses can also be found in Teschl and Falkner (2012). Note that we only consider univariate functions in this paper. For different notions of multivariate inverses in the context of quantile functions, see Chakak and Imlahi (2001), Serfling (2002), Serfling (2008), or Fraiman and Pateiro-López (2011).

2 Generalized inverses and quantile functions

Throughout this article, we understand *increasing* in the sense of non-decreasingness, that is, $T : \mathbb{R} \to \mathbb{R}$ is *increasing* if $T(x) \leq T(y)$ for all $x < y$. The following definition captures the notion of an inverse for such functions. Note that evaluating increasing functions at the symbols $-\infty$ or ∞ is always understood as the corresponding limit (possibly being $-\infty$ or ∞ itself).

Definition 2.1
For an increasing function $T : \mathbb{R} \to \mathbb{R}$ with $T(-\infty) = \lim_{x \downarrow -\infty} T(x)$ and $T(\infty) = \lim_{x \uparrow \infty} T(x)$, the *generalized inverse* $T^- : \mathbb{R} \to \bar{\mathbb{R}} = [-\infty, \infty]$ of T is defined by

$$T^-(y) = \inf\{x \in \mathbb{R} : T(x) \geq y\}, \ y \in \mathbb{R}, \tag{1}$$

with the convention that $\inf \emptyset = \infty$. If $T : \mathbb{R} \to [0, 1]$ is a distribution function, $T^- : [0, 1] \to \bar{\mathbb{R}}$ is also called the *quantile function* of T.

Remark 2.2

(1) If T is continuous and strictly increasing, T^- coincides with T^{-1}, the ordinary inverse of T on $\operatorname{ran} T = \{T(x) : x \in \mathbb{R}\}$, the range of T.

(2) Definition 2.1 for generalized inverses and quantile functions essentially appears throughout the stochastics literature; see for instance Resnick (1987, p. 3), Embrechts et al. (1997, p. 130, 554), and McNeil et al. (2005, p. 39). By this definition, the 0-quantile of a distribution function F is always $F^-(0) = -\infty$. For distribution functions F with $F(x) = 0$ for some $x \in \mathbb{R}$, this definition might be different from what one would expect. For example, for the distribution function $F(x) = 1 - \exp(-x)$ of the standard exponential distribution, the corresponding quantile function is $F^-(y) = -\log(1 - y)$ on $y \in (0, 1]$, which is often also considered as the quantile function on the whole unit interval $[0, 1]$ implying that $F^-(0) = 0$. To get this from the definition, Witting (1985), for example, specifically defines the 0-quantile as $F^-(0) = \sup\{x \in \mathbb{R} : F(x) = 0\}$. In the context of increasing functions in general, such a definition would require one to treat the case where T is constant to the left of some point as a special case, which makes statements and proofs involving generalized inverses more complicated. Another way to obtain $F^-(0) = 0$ for distribution functions with $F(x) = 0$ for some $x \in \mathbb{R}$ would be to restrict the

domain of F in (1) to a subset of $\mathbb{R}$. For the distribution function of the standard exponential distribution, for example, one could define the quantile function as $F^-(y) = \inf\{x \in [0, \infty) : F(x) \geq y\}$.
There are several reasons for not treating 0-quantiles any differently. First, as mentioned above, such definitions make it more complicated to work with these functions. Second, from the definition of a distribution function F of a random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as $F(x) = \mathbb{P}(X \leq x)$, it makes perfect sense to ask for the value of F at any $x \in \mathbb{R}$, even outside the range of X. So generally, the domain of a univariate distribution function should always be $\mathbb{R}$ and Definition 2.1 respects that. Finally, let us remark that from a statistical or practical point of view, the 0-quantile is irrelevant anyway.

Generalized inverses T^- are best thought of in terms of Figure 1. It highlights the two major differences to ordinary inverses. First, T is allowed to be flat. Flat parts of T precisely correspond to jumps in T^-. Second, T is allowed to be non-continuous and the jumps of T precisely correspond to flat parts of T^-.

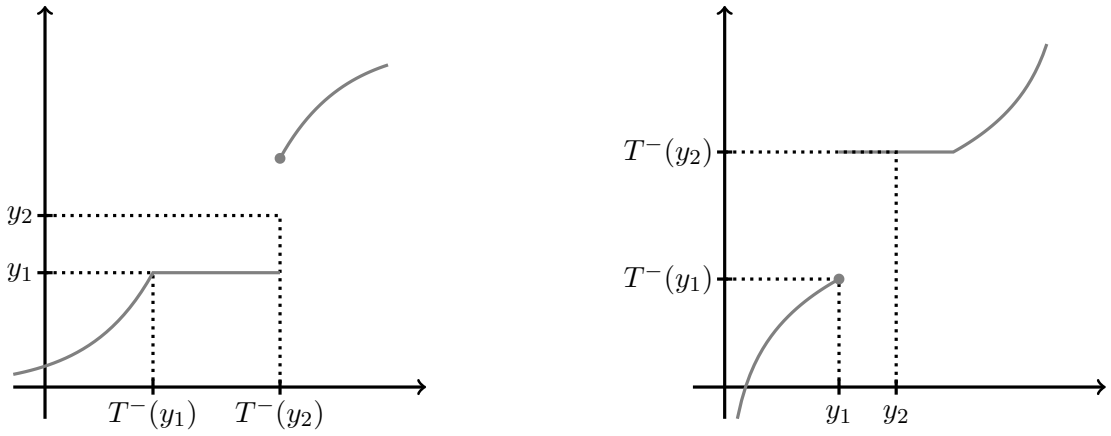

Figure 1 An increasing function (left) and its corresponding generalized inverse (right).

When working with generalized inverses, the following properties are often useful.

Proposition 2.3
Let $T : \mathbb{R} \to \mathbb{R}$ be increasing with $T(-\infty) = \lim_{x \downarrow -\infty} T(x)$ and $T(\infty) = \lim_{x \uparrow \infty} T(x)$, and let $x, y \in \mathbb{R}$. Then,

(1) $T^-(y) = -\infty$ if and only if $T(x) \geq y$ for all $x \in \mathbb{R}$. Similarly, $T^-(y) = \infty$ if and only if $T(x) < y$ for all $x \in \mathbb{R}$.

(2) T^- is increasing. If $T^-(y) \in (-\infty, \infty)$, T^- is left-continuous at y and admits a limit from the right at y.

(3) $T^-(T(x)) \leq x$. If T is strictly increasing, $T^-(T(x)) = x$.

(4) Let T be right-continuous. Then $T^-(y) < \infty$ implies $T(T^-(y)) \geq y$. Furthermore, $y \in \operatorname{ran} T \cup \{\inf \operatorname{ran} T, \sup \operatorname{ran} T\}$ implies $T(T^-(y)) = y$. Moreover, if $y < \inf \operatorname{ran} T$

then $T(T^-(y)) > y$ and if $y > \sup \operatorname{ran} T$ then $T(T^-(y)) < y$.

(5) $T(x) \geq y$ implies $x \geq T^-(y)$. The other implication holds if T is right-continuous. Furthermore, $T(x) < y$ implies $x \leq T^-(y)$.

(6) $(T^-(y-), T^-(y+)) \subseteq \{x \in \mathbb{R} : T(x) = y\} \subseteq [T^-(y-), T^-(y+)]$, where $T^-(y-) = \lim_{z \uparrow y} T^-(z)$ and $T^-(y+) = \lim_{z \downarrow y} T^-(z)$.

(7) T is continuous if and only if T^- is strictly increasing on $[\inf \operatorname{ran} T, \sup \operatorname{ran} T]$. T is strictly increasing if and only if T^- is continuous on $\operatorname{ran} T$.

(8) If T_1 and T_2 are right-continuous transformations with properties as T, then $(T_1 \circ T_2)^- = T_2^- \circ T_1^-$.

Proof

(1) This statement directly follows from Definition 2.1.

(2) T^- is increasing since $\{x \in \mathbb{R} : T(x) \geq y_2\} \subseteq \{x \in \mathbb{R} : T(x) \geq y_1\}$ for all $y_1, y_2 \in \mathbb{R} : y_1 < y_2$. Now let $T^-(y) \in (-\infty, \infty)$ and for convenience, let $y_0 = y$. To show left-continuity in y_0, let $(y_n)_{n \in \mathbb{N}} \subseteq \mathbb{R} : y_n \uparrow y_0$. Then $x_n := T^-(y_n) \leq x_0 := T^-(y_0)$, thus $x_n \nearrow x \leq x_0$ for $n \to \infty$ for some $x \in \mathbb{R}$ (where "$x_n \nearrow x$" is used to denote that x_n converges monotonically (increasing) to x). By definition of T^-, $T(x_n - \varepsilon) < y_n \leq T(x_n + \varepsilon)$ for all $\varepsilon > 0$ and $n \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$. If $x < x_0$, then $\varepsilon = (x_0 - x)/2$ implies $y_n \leq T(x_n + \varepsilon) \leq T(x_0 - \varepsilon) < y_0$, thus $y_0 = \lim_{n \uparrow \infty} y_n \leq T(x_0 - \varepsilon) < y_0$, a contradiction. To show that T^- admits a limit from the right at y, let $y_n \downarrow y \in \mathbb{R} : T^-(y) > -\infty$ and note that $(T^-(y_n))_{n \in \mathbb{N}}$ is decreasing and bounded from below by $T^-(y)$.

(3) The first part follows by definition of T^-. For the second part, note that T being strictly increasing implies that there is no $z < x$ with $T(z) \geq T(x)$, thus $T^-(T(x)) \geq x$.

(4) For the first part, $T^-(y) < \infty$ implies that $A = \{x \in \mathbb{R} : T(x) \geq y\} \neq \emptyset$; thus, there exists $(x_n)_{n \in \mathbb{N}} \subseteq A$ with $x_n \downarrow \inf A = T^-(y)$ for $n \to \infty$. By right-continuity of T, $T(T^-(y)) \swarrow T(x_n) \geq y$, so $T(T^-(y)) \geq y$; here "$T(T^-(y)) \swarrow T(x_n)$" is used to denote that $T(x_n)$ converges monotonically (decreasing) to $T(T^-(y))$.
Now consider the second part. First let $y \in \operatorname{ran} T$ and define $A = \{x \in \mathbb{R} : T(x) = y\} \neq \emptyset$. Note that $\inf A = T^-(y)$ and conclude that $T(T^-(y)) \swarrow T(x_n) = y$, thus $T(T^-(y)) = y$. Now let $y = \inf \operatorname{ran} T$ and without loss of generality, assume $y \notin \operatorname{ran} T$ (otherwise the previous part applies). This implies $T^-(y) = -\infty$. Since T is increasing, we obtain $T(T^-(y)) = T(-\infty) = \inf\{T(x) : x \in \mathbb{R}\} = \inf \operatorname{ran} T = y$. Similarly for $y = \sup \operatorname{ran} T$.
For the first part of the last statement, note that $y < \inf \operatorname{ran} T$ implies $T^-(y) = -\infty$, thus $T(T^-(y)) = T(-\infty) = \inf \operatorname{ran} T > y$. Similarly for $y > \sup \operatorname{ran} T$.

(5) The first statement follows by definition of T^-. For the second statement, note that $T^-(y) \leq x$ implies $y \leq T(T^-(y)) \leq T(x)$, where $y \leq T(T^-(y))$ follows from (4) since T is right-continuous. For the last part, let $x \in \mathbb{R}$ be such that $T(x) < y$ and z be any real number with $T(z) \geq y$. Since T is increasing, $z \geq x$. This implies $T^-(y) \geq x$.

(6) Let $A_y = \{x \in \mathbb{R} : T(x) = y\}$. First, consider $(T^-(y-), T^-(y+)) \subseteq A_y$. Assume there exists an $x \in (T^-(y-), T^-(y+)) \backslash A_y$. This implies that (i) $T^-(y-) < x < T^-(y+)$ and (ii) $T(x) \neq y$. By (i), there exist $\varepsilon_1, \varepsilon_2 > 0$ such that $T^-(z_1) < x - \varepsilon_1 < x + \varepsilon_2 < T^-(z_2)$ for all $z_1 < y < z_2$. By (2), T^- is increasing, thus $T^-(y) \in [x - \varepsilon_1, x + \varepsilon_2]$. By (ii) one has either $T(x) < y$ or $T(x) > y$. If $T(x) < y$, let $z_1 = (T(x)+y)/2 \in (T(x), y)$. Since $T(x) < z_1$, it follows from (5) that $x \leq T^-(z_1)$ which contradicts the fact that $T^-(z_1) < x - \varepsilon_1$ for all $z_1 < y$. If $T(x) > y$, let $z_2 = T(x) > y$. By definition of T^-, $T^-(z_2) \leq x$, which contradicts $T^-(z_2) > x + \varepsilon_2$ for all $z_2 > y$.
Now consider $A_y \subseteq [T^-(y-), T^-(y+)]$. Without loss of generality, assume $A_y \neq \emptyset$. We show $T^-(y-) \leq \inf A_y$ and $\sup A_y \leq T^-(y+)$. For the former, $A_y \subseteq B_y := \{x \in \mathbb{R} : T(x) \geq y\}$ implies that $T^-(y-) \leq T^-(y) = \inf B_y \leq \inf A_y$. For the latter, let $z > y$ and $x \in A_y$. Then $T(x) = y < z$ and (5) implies that $T^-(z) \geq x$. It follows that for all $z > y$, $T^-(z) \geq x$ for all $x \in A_y$. This implies that for all $z > y$, $T^-(z) \geq \sup A_y$, which in turn implies that $T^-(y+) \geq \sup A_y$.

(7) Consider the first statement. We show that T is discontinuous if and only if T^- is not strictly increasing on $[\inf \operatorname{ran} T, \sup \operatorname{ran} T]$. For the only if part, let T be discontinuous at $x_0 \in \mathbb{R}$. Since T is increasing, this implies that $y_1 := T(x_0-) := \lim_{x \uparrow x_0} T(x)$ and $y_2 := T(x_0+) := \lim_{x \downarrow x_0} T(x)$ exist, $y_1 < y_2$, and $y_0 := T(x_0) \in [y_1, y_2]$. Now there exist $y_3, y_4 \in \mathbb{R}$ such that either $y_1 \leq y_0 < y_3 < y_4 < y_2$ or $y_1 < y_3 < y_4 < y_0 \leq y_2$, without loss of generality assume the latter. Note that for all $y \in [y_3, y_4] \subseteq [\inf \operatorname{ran} T, \sup \operatorname{ran} T]$, $y \notin \operatorname{ran} T$. By definition of T^-, this implies that T^- is constant on $[y_3, y_4]$, that is, T^- is not strictly increasing on $[\inf \operatorname{ran} T, \sup \operatorname{ran} T]$. For the if part, let T^- be not strictly increasing on $[\inf \operatorname{ran} T, \sup \operatorname{ran} T]$, that is, there exist $y_1, y_2 \in \mathbb{R}$ with $\inf \operatorname{ran} T \leq y_1 < y_2 \leq \sup \operatorname{ran} T$ such that $T^-(y) = x$ for all $y \in [y_1, y_2]$ and an $x \in \mathbb{R}$. By definition of T^-, this implies that $T(x - \varepsilon) < y_1 < y_2 \leq T(x + \varepsilon)$ for all $\varepsilon > 0$. Letting $\varepsilon \downarrow 0$, we obtain $T(x-) \leq y_1 < y_2 \leq T(x+)$, that is T is discontinuous at x.
Now consider the second statement. We show that T is not strictly increasing if and only if T^- is discontinuous on $\operatorname{ran} T$. For this we apply (6). For the only if part, there exists a $y \in \mathbb{R}$ such that A_y contains an open interval. It follows from the second inclusion in (6) that $T^-(y-) < T^-(y+)$, thus T^- is discontinuous at y; note that by definition of A_y, $y \in \operatorname{ran} T$. For the if part, there exists a $y \in \operatorname{ran} T$ such that $T^-(-y) < T^-(y+)$. It follows from the first inclusion in (6) that A_y contains the (non-empty) open interval $(T^-(-y), T^-(y+))$, thus T is not strictly increasing.

(8) Applying (5) to T_1 and T_2 leads to $(T_1 \circ T_2)^-(y) = \inf\{x \in \mathbb{R} : T_1(T_2(x)) \geq y\} = \inf\{x \in \mathbb{R} : x \geq T_2^-(T_1^-(y))\} = T_2^-(T_1^-(y))$.

□

Remark 2.4

(1) Many of the properties listed in Proposition 2.3 can be found in the literature. However, they are often stated under stronger conditions. For example, Embrechts et al. (1997, p. 555, Proposition A1.6 (a)) state that if $T : \mathbb{R} \to \mathbb{R}$ is increasing and right-continuous, then $T(x) \geq y$ if and only if $x \geq T^-(y)$. According to Proposition

2.3 (5), right-continuity is not needed for the only if part of the statement.

(2) Concerning Proposition 2.3 (2), note that the notions of left-continuity and limits from the right do not exist in their classical definitions if $T^-(y) = \infty$. As an example, consider left-continuity and take a sequence $(y_n)_{n\in\mathbb{N}} \subseteq \mathbb{R} : y_n \uparrow y$. Since T^- is increasing and $T^-(y) = \infty$ (that is $T^-(y) = \infty$ if and only if $T(x) < y$ for all $x \in \mathbb{R}$), it is clear that $T^-(y_n) \nearrow \infty$ for $n \to \infty$. But one can not talk about left-continuity here, since quantities such as $|T^-(y_n) - T^-(y)|$ do not make sense for the epsilon-delta definition of left-continuity if $T^-(y) = \infty$ (if $T^-(y_n)$ is finite for all n, this absolute distance is ∞ for all n; if $T^-(y_n) = \infty$ from some n on, then it is not even defined). Similarly for the limit from the right (which can only be ∞ since $T^-(y_n) = \infty$; but as before, showing that $T^-(y_n) \searrow T^-(y)$ with the definition of convergence is not possible). The same reasoning applies to the case where $T^-(y) = -\infty$.

(3) Consider the second statement in Proposition 2.3 (4). Note that the assumption $y \in \operatorname{ran} T$ can not be replaced by $y \in [\inf \operatorname{ran} T, \sup \operatorname{ran} T]$ in general, which is clear from Figure 1 if one considers the point $y = y_2$ for example.

(4) To see that the other implication of the first statement in Proposition 2.3 (5) does not hold in general, consider $T(x) = \mathbb{1}_{(0,\infty)}(x)$ (the indicator function of the positive real numbers), $x = 0$, and $y = 1/2$. Then $T(x) = T(0) = 0 < 1/2 = y$ although $x = 0 \geq 0 = T^-(1/2) = T^-(y)$.

(5) The first statement of Proposition 2.3 (7) is not correct anymore if T^- is only strictly increasing on $\operatorname{ran} T$. As a counterexample, consider $T(x) = \mathbb{1}_{[0,\infty)}(x)$, $x \in \mathbb{R}$, the indicator function of the non-negative real numbers. Then $T^-(y) = -\infty$ for $y \in (-\infty, 0]$, $T^-(y) = 0$ for $y \in (0, 1]$, and $T^-(y) = \infty$ for $y \in (1, \infty)$. Thus, T^- is strictly increasing on $\operatorname{ran} T = \{0, 1\}$ but T is not continuous. Let us remark that T is also a distribution function.

3 Examples and counterexamples

Generalized inverses appear at various points in the literature. In probability and statistics, they mainly appear as quantile functions, for example, when building confidence intervals or in terms of quantile-quantile plots for goodness-of-fit tests. Also, the median and interquartile range are defined in terms of quantile functions. Many of the basic results in extreme value theory involve generalized inverses; see for instance Embrechts et al. (1997) and the references therein.

We now prove some important results involving generalized inverses and quantile functions based on the results of Proposition 2.3. The first establishes the relation between any univariate distribution function and the uniform distribution on the unit interval. This is important for sampling and goodness-of-fit testing of univariate distributions. The second result shows the invariance principle of copulas under strictly increasing transformations on the ranges of the underlying random variables. This is an important result from the theory of dependence modeling between random variables via copulas. It

allows one to study the dependence structure independent of the marginal distribution functions. Finally, we discuss some incorrect statements which can be found in the literature.

3.1 Examples

An important application of quantile functions is the *inversion method* for generating random variables from univariate distributions in general. This is often applied in Monte Carlo simulations. The other way around, it can be applied as a goodness-of-fit test for univariate distributions. Both applications are contained in the following proposition.

Proposition 3.1
Let F be a distribution function and $X \sim F$.

(1) If F is continuous, $F(X) \sim \mathrm{U}[0,1]$.

(2) If $U \sim \mathrm{U}[0,1]$, $F^-(U) \sim F$.

Proof

(1) By Proposition 2.3 (7), F^- is strictly increasing on $[0,1]$. Therefore, $\mathbb{P}(F(X) \leq x) = \mathbb{P}(F^-(F(X)) \leq F^-(x))$. Although F may not be strictly increasing, it is so on $\operatorname{ran} X$; here and in what follows, the range $\operatorname{ran} X$ of a random variable X is defined as $\operatorname{ran} X = \{x \in \mathbb{R} : \mathbb{P}(X \in N_x) > 0 \text{ for all neighborhoods } N_x \in \mathcal{B}(\mathbb{R}) \text{ of } x\}$. By Proposition 2.3 (3), it therefore follows that $\mathbb{P}(F^-(F(X)) \leq F^-(x)) = \mathbb{P}(X \leq F^-(x)) = F(F^-(x))$ for all $x \in \mathbb{R}$. By Proposition 2.3 (4) this equals x for all $x \in (0,1) \subseteq \operatorname{ran} F$. Thus $\mathbb{P}(F(X) \leq x) = x$, $x \in (0,1)$, which implies $F(X) \sim \mathrm{U}[0,1]$.

(2) By Proposition 2.3 (5), $\mathbb{P}(F^-(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x)$ for all $x \in \mathbb{R}$.

□

As we can see from Proposition 3.1, transforming a random variable by its continuous distribution function always leads to the same distribution, the standard uniform distribution. Transforming a random vector componentwise in this way, however, may lead to different multivariate distributions than the multivariate standard uniform distribution. This distribution depends on the dependence structure of the transformed random variables and is captured by the underlying copula.

Copulas are distribution functions with standard uniform univariate margins. They play an important role in modeling dependencies between random variables. By Sklar's Theorem (see, for example Sklar (1996) or Rüschendorf (2009)), any multivariate distribution function H with marginals F_j, $j \in \{1, \ldots, d\}$, can be written as

$$H(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)), \ \boldsymbol{x} \in \mathbb{R}^d, \tag{2}$$

for a copula C. This decomposition is uniquely defined on $\prod_{j=1}^d \operatorname{ran} F_j$. Given $\boldsymbol{X} \sim H$, we call any C which fulfills (2) *a copula of* H (or $\boldsymbol{X}$); similarly, we say that H (or $\boldsymbol{X}$) *has copula* C.

We now prove some important statistical results about copulas. Here, increasing transformations T (and their generalized inverses) as well as distribution functions (and

their quantile functions) naturally appear. Once more, these results are well-known; see for instance McNeil et al. (2005, p. 188) or Nelsen (2006, p. 25).

Proposition 3.2
Let $\boldsymbol{X} = (X_1, \dots, X_d)^\mathsf{T}$ have joint distribution function H with continuous marginals F_j, $j \in \{1, \dots, d\}$. Then $\boldsymbol{X}$ has copula C if and only if $(F_1(X_1), \dots, F_d(X_d))^\mathsf{T} \sim C$.

Proof
First consider the only if part. Proposition 3.1 (1) implies that $F_j(X_j)$ is continuously distributed for all $j \in \{1, \dots, d\}$. By Proposition 2.3 (5), and since X_j is continuously distributed for all $j \in \{1, \dots, d\}$, we have $\mathbb{P}(F_1(X_1) \le u_1, \dots, F_d(X_d) \le u_d) = \mathbb{P}(F_1(X_1) < u_1, \dots, F_d(X_d) < u_d) = \mathbb{P}(X_1 < F_1^-(u_1), \dots, X_d < F_d^-(u_d)) = \mathbb{P}(X_1 \le F_1^-(u_1), \dots, X_d \le F_d^-(u_d)) = H(F_1^-(u_1), \dots, F_d^-(u_d)) = C(\boldsymbol{u})$, where H denotes the distribution function of $\boldsymbol{X}$ and the last equality follows from Sklar's Theorem.
For the if part, note that F_j is strictly increasing on $\operatorname{ran} X_j$ for all $j \in \{1, \dots, d\}$. Applying Proposition 2.3 (3) and (5) leads to $\mathbb{P}(\boldsymbol{X} \le \boldsymbol{x}) = \mathbb{P}(F_1^-(F_1(X_1)) \le x_1, \dots, F_d^-(F_d(X_d)) \le x_d) = \mathbb{P}(F_1(X_1) \le F_1(x_1), \dots, F_d(X_d) \le F_d(x_d))$, which, by assumption, equals $C(F_1(x_1), \dots, F_d(x_d))$. By Sklar's Theorem, this means that $\boldsymbol{X}$ has copula C. □

The following proposition addresses two versions of the famous invariance principle of copulas which involves increasing transformations of random variables.

Proposition 3.3 (Invariance principle)

(1) Let $\boldsymbol{X} = (X_1, \dots, X_d)^\mathsf{T}$ have joint distribution function H with continuous marginals F_j, $j \in \{1, \dots, d\}$, and copula C. Let $T_j : \mathbb{R} \to \mathbb{R}$ be strictly increasing on $\operatorname{ran} X_j$, $j \in \{1, \dots, d\}$. Then $(T_1(X_1), \dots, T_d(X_d))^\mathsf{T}$ (also) has copula C.

(2) Let $\boldsymbol{X} = (X_1, \dots, X_d)^\mathsf{T}$ have joint distribution function H with marginals F_j, $j \in \{1, \dots, d\}$, and copula C. Let $T_j : \mathbb{R} \to \mathbb{R}$ be right-continuous and strictly increasing on $\operatorname{ran} X_j$, $j \in \{1, \dots, d\}$. Then $(T_1(X_1), \dots, T_d(X_d))^\mathsf{T}$ (also) has copula C.

Proof

(1) Assume without loss of generality that T_j is right-continuous at its (at most countably many) discontinuities. Since X_j is continuously distributed and T_j is strictly increasing on $\operatorname{ran} X_j$, the distribution function G_j of $T_j(X_j)$ is given by $G_j(x) = \mathbb{P}(T_j(X_j) \le x) = \mathbb{P}(T_j(X_j) < x)$, which, by Proposition 2.3 (5), equals $\mathbb{P}(X_j < T_j^-(x)) = \mathbb{P}(X_j \le T_j^-(x)) = F_j(T_j^-(x))$. Since F_j is continuous and, by Proposition 2.3 (7), T_j^- is continuous on $\operatorname{ran} T_j(X_j)$, G_j is continuous on $\operatorname{ran} T_j(X_j)$. Since $T_j(X_j)$ does not put mass outside $\operatorname{ran} T_j(X_j)$, G_j is (even) continuous on $\mathbb{R}$, $j \in \{1, \dots, d\}$. As T_j is strictly increasing on $\operatorname{ran} X_j$, $j \in \{1, \dots, d\}$, it now follows from Proposition 2.3 (3) that $\mathbb{P}(G_j(T_j(X_j)) \le u_j,\ j \in \{1, \dots, d\}) = \mathbb{P}(F_j(T_j^-(T_j(X_j))) \le u_j,\ j \in \{1, \dots, d\}) = \mathbb{P}(F_j(X_j) \le u_j,\ j \in \{1, \dots, d\})$. Since $\boldsymbol{X}$ has copula C, the only if part of Proposition 3.2 implies that this equals $C(\boldsymbol{u})$. By continuity of G_j, $j \in \{1, \dots, d\}$, the claim then follows from the if part of Proposition 3.2.

(2) By Proposition 2.3 (4) and since T_j is strictly increasing on $\operatorname{ran} X_j$, the distribution function G_j of $T_j(X_j)$ is given by $G_j(x) = \mathbb{P}(T_j(X_j) \le x) = \mathbb{P}(T_j(X_j) \le$

$T_j(T_j^-(x))) = \mathbb{P}(X_j \leq T_j^-(x)) = F_j(T_j^-(x))$ for all $x \in \operatorname{ran} T_j$. Since $T_j(X_j)$ does not take on values outside $\operatorname{ran} T_j$ with non-zero probability, this implies that $G_j(x) = F_j(T_j^-(x))$ for all $x \in \mathbb{R}$, $j \in \{1, \ldots, d\}$. Applying the same logic implies $\mathbb{P}(T_j(X_j) \leq x_j,\ j \in \{1, \ldots, d\}) = \mathbb{P}(T_j(X_j) \leq T_j(T_j^-(x_j)),\ j \in \{1, \ldots, d\}) = \mathbb{P}(X_j \leq T_j^-(x_j),\ j \in \{1, \ldots, d\}) = H(T_1^-(x_1), \ldots, T_d^-(x_d))$. By Sklar's Theorem, this equals $C(F_1(T_1^-(x_1)), \ldots, F_d(T_d^-(\ x_d))) = C(G_1(x_1), \ldots, G_d(x_d))$. Again from Sklar's Theorem, it follows that $(T_1(X_1), \ldots, T_d(X_d))^\mathsf{T}$ has copula C.

□

3.2 Counterexamples

Some important properties of generalized inverses for increasing functions stated in the literature are not correct. We now address some popular statements found in textbooks and give counterexamples. We stress that these counterexamples specifically address increasing functions which are not distribution functions.

Statement 1
McNeil et al. (2005, p. 495, Proposition A.3 (vi)) states that if T is increasing, then $T(T^-(y)) \geq y$.

As a counterexample for Statement 1, consider $T(x) = \mathbb{1}_{(0,\infty)}(x)$, $x \in \mathbb{R}$, the indicator function of the positive real numbers. Then $T(T^-(1/2)) = T(0) = 0 < 1/2$, so Statement 1 is not correct in general. From this counterexample we see that T also has to be right-continuous for $T(T^-(y)) \geq y$ to hold. But even both increasingness and right-continuity do not suffice, as the following Statement 2 shows.

Statement 2
Both Resnick (1987, p. 3, Inequality (0.6b)) and Embrechts et al. (1997, p. 555, Proposition A1.6 (d)) state that if T is increasing and right-continuous, then $T(T^-(y)) \geq y$.

As a counterexample for Statement 2, consider the logistic function $T(x) = 1/(1 + \exp(-x))$, $x \in \mathbb{R}$, and $y = 2$. Then $T(T^-(2)) = T(\infty) = 1 < 2$, so Statement 2 is not correct in general. From this counterexample we see that $T(T^-(y)) \geq y$ does not have to hold if $T^-(y) = \infty$ (so if $T(x) < y$ for all $x \in \mathbb{R}$, which in particular also implies that $y \notin \operatorname{ran} T$). However, note that it does hold if $T^-(y) = \infty$ as long as $y = \sup \operatorname{ran} T$, in which case even $T(T^-(y)) = y$ is true; see Proposition 2.3 (4).

To see that the following statement is not correct in general, one may use the same counterexample as for Statement 2. Also note that by Proposition 2.3 (4), the continuity assumption can be relaxed to right-continuity.

Statement 3
Embrechts et al. (1997, p. 555, Proposition A1.6 (d)) state that if T is increasing and continuous, then $T(T^-(y)) = y$.

The following statement is the same as Statement 3, under one additional assumption. Again, the logistic function can be used as a counterexample, this time taking $y = -1$.

Statement 4
McNeil et al. (2005, p. 495, Proposition A.3 (viii)) state that if $T^-(y) < \infty$ and if T is increasing and continuous, then $T(T^-(y)) = y$.

Note that in contrast to Statement 4, the statement $T(T^-(y)) \geq y$ is correct (given that $T^-(y) < \infty$), even under weaker assumptions (only right-continuity); see Proposition 2.3 (4).

Finally, let us now consider the special case of a distribution function F. Under this assumption, Statements 1 and 2 become correct, by Proposition 2.3 (4). To see this, note that for $y \in [0, 1)$, $F^-(y) < \infty$, so the first part of Proposition 2.3 (4) applies; for $y = 1 = \sup \operatorname{ran} F$, the second part applies. Also Statements 3 and 4 become correct since continuity implies that $\operatorname{ran} F \supseteq (0, 1)$, so that $\operatorname{ran} F \cup \{\inf \operatorname{ran} F, \sup \operatorname{ran} F\} = [0, 1]$. Thus, the second part of Proposition 2.3 (4) implies $F(F^-(y)) = y$ for all $y \in [0, 1]$.

4 Conclusion

We stated and proved several properties about generalized inverses. Furthermore, we gave examples to stress their importance from both a theoretical perspective and in applications. Finally, counterexamples for statements found in the literature show that one has to be aware of the precise statements when working with generalized inverses, for which this article provides guidance. The latter statement particularly applies when going from probability distribution functions to more general increasing functions.

Acknowledgements
The authors would like to thank Jozef Doboš (Pavol Jozef Šafárik University in Košice), Fabrizio Durante (Free University of Bozen-Bolzano), and Bartosz Langowski for giving valuable feedback on an early version of the paper. Furthermore, the authors would like to thank two anonymous referees for carefully reading the paper and giving suggestions.

References

Chakak, A. and Imlahi, L. (2001), Multivariate Probability Integral Transformation: Application to maximum likelihood estimation, *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales, Serie A Matemáticas*, 95(2), 201–212.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997), Modelling Extremal Events for Insurance and Finance, Springer.

Fraiman, R. and Pateiro-López, B. (2011), Functional quantiles, *Recent advances in functional data analysis and related topics*, ed. by F. Ferraty, Springer, 123–129.

Klement, E. P., Mesiar, R., and Pap, E. (1999), Quasi- and pseudo-inverses of monotone functions, and the construction of t-norms, *Fuzzy Sets and Systems*, 104, 3–13.

McNeil, A. J., Frey, R., and Embrechts, P. (2005), Quantitative Risk Management: Concepts, Techniques, Tools, Princeton University Press.

Nelsen, R. B. (2006), An Introduction to Copulas, Springer.

Resnick, S. I. (1987), Extreme Values, Regular Variation, and Point Processes, Springer.

Rüschendorf, L. (2009), On the distributional transform, Sklar's Theorem, and the empirical copula process, *Journal of Statistical Planning and Inference*, 139(11), 3921–3927.

Serfling, R. (2002), Quantile Functions for Multivariate Analysis: Approaches and Applications, *Statistica Neerlandica*, 56, 214–232.

Serfling, R. (2008), A Mahalanobis Multivariate Quantile Function, `http://www.utdallas.edu/~serfling/papers/Mahalanobis.pdf` (2011-07-17).

Sklar, A. (1996), Random variables, distribution functions, and copulas – a personal look backward and forward, *Distributions with Fixed Marginals and Related Topics*, 28, 1–14.

Teschl, G. and Falkner, N. (2012), On the substitution rule for Lebesgue–Stieltjes integrals, *Expositiones Mathematicae*, `doi:10.1016/j.exmath.2012.09.002`.

Witting, H. (1985), Mathematische Statistik: Parametrische Verfahren bei festem Stichprobenumfang, Teubner.