Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 21–49

TRANSFORM ORDERS AND STOCHASTIC MONOTONICITY OF STATISTICAL FUNCTIONALS

TOMMASO LANDO, IDIR ARAB AND PAULO EDUARDO OLIVEIRA

ABSTRACT: In some inferential statistical methods, such as tests and confidence intervals, it is important to describe the stochastic behavior of statistical functionals, aside from their large sample properties. We study such a behavior in terms of the usual stochastic order. For this purpose, we introduce a generalized family of stochastic orders, which is referred to as transform orders, showing that it provides a flexible framework for deriving stochastic monotonicity results. Given that our general definition makes it possible to obtain some well known ordering relations as particular cases, we can easily apply our method to different families of functionals. These include some prominent inequality measures, such as the generalized entropy, the Gini index, and its generalizations. We also demonstrate the applicability of our approach by determining the least favorable distribution, and the behavior of some bootstrap statistics, in some goodness-of-fit testing procedures.

KEYWORDS: Gini index, inequality, L-statistic, Lorenz order, nonparametric test, stochastic dominance.
MATH. SUBJECT CLASSIFICATION (2020): 60E15, 60E05, 62G10, 62N05.

1. Introduction

In statistics, one is often interested in estimating an unknown characteristic of a given distribution F, rather than the distribution itself. In many such cases, these characteristics may be represented by some *probability functional* $T(F) : \mathcal{F} \to \mathbb{R}$, where $\mathcal{F}$ is the space of probability distributions. The most intuitive way of estimating $T(F)$ is the *plug-in* method, which simply consists in replacing the unknown F with its natural estimator, namely, the empirical distribution function $\mathbb{F}_n$. Correspondingly, we shall refer to $T(\mathbb{F}_n)$ as a *statistical functional*. However, since the empirical distribution is a random process, that is, a function of the random sample, then statistical functionals are random variables, and therefore it is important to study their behavior

Received December 03, 2021.
Tommaso Lando was supported by the Italian funds ex MURST 60% 2019, by the Czech Science Foundation (GACR) under project 20-16764S and moreover by SP2020/11, an SGS research project of VŠB-TU Ostrava. The support is greatly acknowledged.
This work was partially supported by the Centre for Mathematics of the University of Coimbra - UIDB/00324/2020, funded by the Portuguese Government through FCT/MCTES.

from a stochastic point of view. While most of the work in the literature focuses on large sample properties of statistical functionals, because of their applications to estimation theory, we are concerned with a different problem. In particular, let $\mathbb{F}_n$ and $\mathbb{G}_n$ be the empirical distribution functions that correspond to random samples from F and G, respectively. We are interested in determining the conditions on F, G and T such that $T(\mathbb{F}_n)$ *stochastically dominates* $T(\mathbb{G}_n)$, with respect to the *usual stochastic order* (recalled below in Definition 1), which is a relevant information for many testing and confidence interval procedures, for instance. Results of this kind may be rather complicated to derive, depending on the mathematical form of T, and on the initial conditions on the baseline distributions F and G. In this regard, Arendarczyk et al. (2021) recently studied the stochastic dominance behavior of the Greenwood statistic (Greenwood, 1946) when the baseline distributions are comparable in terms of the *star order*. We will show that this result is a particular case of a much more general one, making it possible to establish dominance relations within a wider family of statistical functionals. For this purpose, we introduce a generalized family of stochastic orders, generated by some reference class of functions $\mathcal{C}$, denoted as *transform orders*, which includes many relevant special cases. Our main results, Theorems 5 and 7, show that, if the baseline distributions are ordered with respect to some transform order, the statistical functionals are stochastically ordered, provided that the corresponding probability functional is *isotonic*, or order-preserving, with respect to the same transform order. This result has a wide range of applicability, as our definition of transform order includes the usual, the convex transform, the star, the superadditive and the dispersive order (Shaked and Shanthikumar, 2007). Therefore, we are able to derive stochastic properties of several important families of functionals. For example, in this paper we prove the stochastic monotonicity of the most relevant measures of inequality, including the generalized entropy class (Shorrocks, 1980; Shorrocks and Slottje, 2002), the established Gini index (Gini, 1912) and its generalizations (Mehran, 1976; Donaldson and Weymark, 1980). As an application, our result may be used to determine, in a simple way, the least favorable distribution, as well as the behavior of some bootstrap statistics, for some goodness-of-fit testing problems, related to the convex order or to the star order, which are relevant in areas such as survival analysis, reliability, and shape-constrained inference (Barlow et al., 1971; Deshpande, 1983;

Tenga and Santner, 1984; Hall and Van Keilegom, 2005; Groeneboom and Jongbloed, 2012; Lando, 2021).

2. Main result

We begin by defining some notations. The random variables X and Y have cumulative distribution functions F and G and supports S_F and S_G, respectively. Let us denote with F^{-1} the left continuous generalized inverse, namely, the quantile function, of a distribution F. We recall that a stochastic order is a binary relation $\succ$ over $\mathcal{F}$ that is reflexive and transitive. In particular, observe that $\succ$ does not generally satisfy the antisymmetry property, that is, $F \succ G$ and $G \succ F$ does not necessarily imply $X =_d Y$, and it is generally not total. A functional T is said to be *isotonic*, or order-preserving, with respect to $\succ$, whenever, for every pair $F, G \in \mathcal{F}$ such that $F \succ G$, it holds that $T(F) \geq T(G)$. Let us denote by $\mathcal{T}$ the class of functionals defined on $\mathcal{F}$. Then, we may represent by $\mathcal{I}(\succ)$ the class of functionals that are isotonic with respect to $\succ$:

$$\mathcal{I}(\succ) = \{T \in \mathcal{T} : F \succ G \implies T(F) \geq T(G), \forall F, G \in \mathcal{F}\}.$$

Likewise, the class of functionals that are *antitonic*, or order-reversing, with $\succ$, is denoted as $\mathcal{A}(\succ) = \{T \in \mathcal{T} : F \succ G \implies T(F) \leq T(G), \forall F, G \in \mathcal{F}\}$. It is worth noticing that, if $F \succ_0 G \implies F \succ_1 G$, then $\mathcal{I}(\succ_1) \subset \mathcal{I}(\succ_0)$.

Given that stochastic orders and probability functionals depend only the distribution functions of the random variables, we may write $F \succ G$ or $X \succ Y$, and $T(F)$ or $T(X)$, interchangeably.

We are concerned with the problem of ranking statistical functionals in terms of the usual stochastic order.

Definition 1. We say that $T(\mathbb{F}_n)$ stochastically dominates $T(\mathbb{G}_n)$, denoted by $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$, if $P(T(\mathbb{F}_n) \leq x) \leq P(T(\mathbb{G}_n) \leq x)$, for every $x \in \mathbb{R}$.

Therefore, we are interested in determining the conditions on F, G and T such that the above relation holds. To this aim, we introduce the following definition of transform orders, generated by a family of functions $\mathcal{C}$, whose choice enables the generalization of some well known orders. This new definition is important from a technical point of view, as we will discuss later. As usual, $f|_E$ denotes the restriction of a function f to the set E.

Definition 2. Let $\mathcal{C}$ be some family of non-decreasing functions. We say that X dominates Y with respect to the $\mathcal{C}$-transform order, denoted by $X \geq_{\mathcal{C}} Y$, if $F^{-1} \circ G|_{S_G} \in \mathcal{C}$.

Note that $X =_d F^{-1} \circ G(Y)$, so, basically, a transform order holds when the function that transforms (distributionally) one random variable into the other satisfies some properties of interest. Such properties fully characterize the dominance relations. It is easy to see that $\geq_{\mathcal{C}}$ fulfils the basic properties of stochastic orders, moreover, $\mathcal{C}_0 \subset \mathcal{C}_1$ entails that $F \geq_{\mathcal{C}_0} G \implies F \geq_{\mathcal{C}_1} G$. As limiting cases, if $\mathcal{C}$ is the class of non-decreasing functions, then $\geq_{\mathcal{C}}$ is always verified, whereas if $\mathcal{C}$ contains just the identity function then $\geq_{\mathcal{C}}$ coincides with equality in distribution, $=_d$. Choosing $\mathcal{C}$ as the class of convex, starshaped, or superadditive functions, we obtain the *convex transform order*, *the star order*, or the *superadditive order*, respectively. Also the *dispersive order* can be obtained by choosing $\mathcal{C}$ as the class of functions ψ such that $\psi(x)-x$ is nondecreasing (Shaked and Shanthikumar, 2007). Since these four orders are particularly important, we recall their definitions below (remember that $F^{-1}\circ G$ is always non-decreasing by construction). However, Definition 2 above is quite general and enables the possibility of defining new stochastic orders. Moreover, notice that, in the sense of Definition 2 the usual stochastic order is also a transform order, as it is obtained by choosing $\mathcal{C}$ as the class of functions ψ such that $\psi(x) \geq x$, that is $X \geq_{st} Y$ if and only if $F^{-1}\circ G(x) \geq x$ for every $x \in S_G$.

Definition 3. We say that X dominates Y with respect to

(1) the convex transform order, denoted by $X \geq_c Y$, if $F^{-1} \circ G|_{S_G}$ is convex;
(2) the star order, denoted by $X \geq_* Y$, if $F^{-1} \circ G|_{S_G}$ is starshaped;
(3) the superadditive order, denoted by $X \geq_{su} Y$, if $F^{-1} \circ G|_{S_G}$ is supperadditive;
(4) the dispersive order, denoted by $X \geq_{disp} Y$, if $(F^{-1} \circ G(x) - x)|_{S_G}$ is increasing.

For non-negative random variables, the following relations hold:

$$X \geq_c Y \implies X \geq_* Y \iff \log X \geq_{disp} \log Y \implies X \geq_{su} Y.$$

As discussed earlier, we are interested in determining conditions under which $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$. For instance, if $T(F) = \int_{\mathbb{R}} x dF(x)$ is the mean of X, it is easy to see that $T \in \mathcal{I}(\geq_{st})$ and, if $X \geq_{st} Y$, then $\frac{1}{n}\sum_i X_i \geq_{st} \frac{1}{n}\sum_i Y_i$, that is, the corresponding sample means are stochastically ordered. More generally, if X and Y are ranked in the usual stochastic order, this property holds for every functional which can be seen as an increasing function of the random sample (see Theorem 1.A.3 of Shaked and Shanthikumar (2007)). The stochastic monotonicity of the sample mean can be derived similarly for all stochastic orders that are closed under convolutions, such as the hazard rate order, the likelihood ratio order, the convolution order, the convex (concave) order and the increasing convex (concave) order (Shaked and Shanthikumar, 2007). However, results of this kind are not predictable in general, especially if one is interested in the behavior of statistical functionals that are not related to the stochastic orders discussed above, or which may have more general mathematical representations, not necessarily sums of random variables. In this regard, intuitively, one may wonder whether $T(F) \geq T(G)$ is sufficient for $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$, or whether $X \succ Y$ and $T \in \mathcal{I}(\succ)$ imply $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$, for a general order $\succ$. It is important to remark that, in general, this is not true (see Subsection 3.3 below). Our main results, Theorem 5 and Theorem7, show that transform orders provide a general framework for deriving stochastic monotonicity of statistical functionals of the form $T(\mathbb{F}_n)$, conditionally on the isotonicity properties of T. However, before proceeding, we need some further preliminary definitions and notations. Indeed, transform orders might be not defined for realizations, F_n and G_n, of the empirical distribution functions $\mathbb{F}_n$ and $\mathbb{G}_n$, as F_n and G_n are discrete. For instance, absolute continuity is necessary for comparisons in terms of $\geq_c$, and in the empirical case the composition $F_n^{-1} \circ G_n$ is a step function, hence not convex. Therefore, to prove our main result, we must *extend*, when needed, any transform order to the class of observed empirical distributions, according to the following general definition.

Definition 4. Let $\mathcal{C}$ be some family of non-decreasing functions. We say that X dominates Y with respect to the extended $\mathcal{C}$-transform order, denoted as $X \geq^e_{\mathcal{C}} Y$, if there exists a function $\phi \in \mathcal{C}$ such that $\phi|_{S_G} = F^{-1} \circ G|_{S_G}$.

Note the distinction between Definitions 2 and 4: the former requires that $F^{-1} \circ G$ satisfies some property, described by $\mathcal{C}$, on the support of G, while the latter assumes the existence of some function $\phi \in \mathcal{C}$, which coincides with the composition $F^{-1} \circ G$ on the support of G, formally, $\phi|_{S_G} = F^{-1} \circ G|_{S_G}$. This makes it possible to consider $\mathcal{C}$ as the class of convex or starshaped functions, even when F and G are empirical distributions, which is out of the scope of Definition 2. Indeed, with regard to the pair F_n, G_n, where $S_{G_n} = \{y_1, \ldots, y_n\}$, $F_n \geq^e_{\mathcal{C}} G_n$ holds, if there exists a function $\phi \in \mathcal{C}$ such that $\phi(y_i) = F_n^{-1} \circ G_n(y_i)$, $i = 1, \ldots, n$. Moreover, note that $\geq^e_{\mathcal{C}}$ is a weaker version of $\geq_{\mathcal{C}}$, that is, if $X \geq_{\mathcal{C}} Y$ then $X \geq^e_{\mathcal{C}} Y$ also holds.

Theorem 5. *If $X \geq^e_{\mathcal{C}} Y$ and $T \in \mathcal{I}(\geq^e_{\mathcal{C}})$, then $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$*

Proof: Consider a sample $y_1, \ldots, y_n$ from Y. As $X =_d F^{-1} \circ G(Y)$, the values $x_i^* = F^{-1} \circ G(y_i)$, $i = 1, \ldots, n$, may be seen as observations from X. Now, let us denote with G_n the realization of $\mathbb{G}_n$ corresponding to the sample $y_1, \ldots, y_n$, and with F_n^* the particular realization of $\mathbb{F}_n$ corresponding to the sample $x_1^*, \ldots, x_n^*$. Moreover, let $\tilde{X}_n$ and $\tilde{Y}_n$ be two random variables whose distributions are F_n^* and G_n, respectively. Obviously, $F^{-1} \circ G(\tilde{Y}_n) =_d \tilde{X}_n$. If $X \geq^e_{\mathcal{C}} Y$, then there exists some $\phi \in \mathcal{C}$ satisfying Definition 4. Let Z be a random variable with distribution H and support S_H included in S_G. $X \geq^e_{\mathcal{C}} Y$ implies

$$\phi|_{S_H} = (\phi|_{S_G})|_{S_H} = (F^{-1} \circ G|_{S_G})|_{S_H} = F^{-1} \circ G \circ H^{-1} \circ H|_{S_H}.$$

Note that the quantile function of $F^{-1} \circ G(Z)$ is $F^{-1} \circ G \circ H^{-1}$, hence the latter relation is equivalent to $F^{-1} \circ G(Z) \geq^e_{\mathcal{C}} Z$, for every random variable Z with support included in that of Y. Because the support of $\tilde{Y}_n$ is always included in the support of Y, this yields $\tilde{X}_n =_d F^{-1} \circ G(\tilde{Y}_n) \geq^e_{\mathcal{C}} \tilde{Y}_n$. Now, as we are assuming that $T \in \mathcal{I}(\geq^e_{\mathcal{C}})$, it follows that $T(F_n^*) \geq T(G_n)$. Note that we are not interested in comparing G_n and F_n^*, however, this result contains information about the distributional behavior of $T(\mathbb{F}_n)$ and $T(\mathbb{G}_n)$. In fact, since $T(F_n^*) \geq T(G_n)$ holds for every possible pairs of realizations $x_1^*, \ldots, x_n^*$ and $y_1, \ldots, y_n$, obtained as above, the latter relation can be equivalently expressed as $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$ (again, this characterization of the usual stochastic order is due the probability integral transform, see Theorem 1.A.1 of Shaked and Shanthikumar (2007)). ■

Remark 6. With regard to Definition 4, if $\phi = F_n^{-1} \circ G_n$, then $\geq_{\mathcal{C}}$ is already defined on the support of G_n by Definition 2. However, the extended Definition 4 is useful because, for instance, notions as convexity and starshapedness are generally not defined for functions whose support is a discrete set of points. Moreover, it seems that, although theoretically $\mathcal{I}(\geq^e_{\mathcal{C}}) \subset \mathcal{I}(\geq_{\mathcal{C}})$, in practice $T \in \mathcal{I}(\geq_{\mathcal{C}})$ is often equivalent to $T \in \mathcal{I}(\geq^e_{\mathcal{C}})$. However, we cannot guarantee this in general.

Theorem 5 demonstrates the practical usefulness of our Definition 4. In fact, transform orders may not be meaningful sometimes, especially when choosing an unusual reference class $\mathcal{C}$, but they may be employed merely as tools for deriving stochastic properties of statistical functionals. To apply Theorem 5, it is enough to show that the functional of interest T belongs to some isotonic family, $\mathcal{I}(\geq_{\mathcal{C}}^{e})$. To do so, one may use existing results which ensure that $T \in \mathcal{I}(\geq_{\mathcal{C}}^{e})$. In particular, as we discuss in the next section, this approach works when $\mathcal{C}$ is the star order or the convex transform order, enabling the determination of the stochastic monotonicity of the main measures of inequality and skewness. Of course, if T is isotonic with more than one of the orders in Definition 4, we should use the weakest one, in order to enlarge the range of applicability of the result. On the other hand, it might happen that our functional T is not isotonic with known orders. In this case, we may try to define a new ad hoc transform order such that $T \in \mathcal{I}(\geq_{\mathcal{C}}^{e})$, for instance by choosing a wider class $\mathcal{C}$ or by defining $\mathcal{C}$ according to the properties of T.

Before closing this section, it is worth noticing that the assumptions of Theorem 5 may be relaxed. In fact, one may replace the isotonicity assumption, $T \in \mathcal{I}(\geq_{\mathcal{C}}^{e})$, with a weaker order-preserving property. In particular, let G_n and F_n^* be a pair of empirical distribution functions, where F_n^* is obtained from the sample $F^{-1} \circ G(y_i)$, $i = 1, \ldots, n$ (as in the proof of Theorem 5). Going back to the proof of Theorem 5, it is easy to see that the following result is also true.

Theorem 7. *If $X \geq_{\mathcal{C}}^{e} Y$ and $T(F_n^*) \geq T(G_n)$, then $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$.*

The isotonicity assumption in Theorem 5 is stronger than that in Theorem 7. Nevertheless, several results available in literature assert the isotonicity of popular probability functionals, with respect to some of the transform orders mentioned above, hence Theorem 5 immediately provides the ordering between the corresponding statistical functionals, as discussed in the following section. When such isotonicity results are not available, Theorem 7 offers a weaker order-preserving assumption, allowing for the same conclusion. This latter approach will be explored in Section 4.

3. Stochastic behavior of inequality measures

We now consider some important families of statistical functionals that are commonly employed in several fields, including statistics, economic and finance. It should be noted that functionals that are isotonic with the usual stochastic order are generally seen as *location* functionals. Since every function of the random sample, which is nondecreasing in each argument, has a stochastically increasing behavior (Shaked and Shanthikumar, 2007, Theorem 1.A.3), then the stochastic monotonicity of *location estimators* may be easily derived. Therefore, we focus on some other families of functionals, that are generally not isotonic with the usual stochastic order, namely inequality measures. As discussed later in this section, inequality measures must be isotonic with the star order, so they fit perfectly to our framework. Similarly, notice that isotonicity with the convex transform order is a basic condition for skewness measure (van Zwet, 1964; Oja, 1981; Eberl and Klar, 2021), therefore our method may be also applied to skewness measures, but this is beyond the purpose of this paper.

Hereafter we focus on non-negative random variables with finite mean. Let $L_F(p) = \frac{1}{\mu_X}\int_0^p F^{-1}(t)\,dt$, $p \in [0,1]$ be the *Lorenz curve* of F (as usual, μ_X and σ_X represent the mean and the standard deviation of X, respectively). The Lorenz curve is a primary tool for representation of inequality (e.g. income inequality), as it is typically understood that the higher of two non-intersecting Lorenz curves shows less inequality. This gives rise to the definition of the *Lorenz order*.

Definition 8. We say that X dominates Y with respect to the Lorenz order, denoted by $X \geq_L Y$, if $L_F(p) \leq L_G(p)$ for every $p \in [0,1]$.

Note that the reverse relation ($L_F(p) \geq L_G(p)$) is sometimes used to define the Lorenz order. We use the same notation as Marshall et al. (2011), that is, X dominates Y if it exhibits higher inequality, as measured by the Lorenz curve. Using standard arguments in the ordering theory, it is possible to derive several classes of probability functionals that are isotonic with $\geq_L$. In particular, any functional $I(X)$ satisfying the following basic properties (see Shorrocks (1980)), may be seen as an inequality measure:

(1) $I \in \mathcal{I}(\geq_L)$;
(2) $I(X) \geq I(\mu_X)$;
(3) $I(aX) = I(X)$, $a > 0$;
(4) $I(X + b) \leq I(X)$, $b > 0$.

Actually, properties (2)–(4) are redundant, as they are implied by (1), that is, the Lorenz isotonicity of the inequality measure I is the crucial property.

It is well known that the star order implies the Lorenz order (Shaked and Shanthikumar, 2007), so that $\mathcal{I}(\geq_L) \subset \mathcal{I}(\geq_*)$. However, since the star order is usually defined only in the continuous case, we need the following extension.

Proposition 9. $\mathcal{I}(\geq_L) \subset \mathcal{I}(\geq_*^e)$.

Proof: Choose $T \in \mathcal{I}(\geq_L)$. The inclusion follows if we prove that $F_n \geq_*^e G_n \implies F_n \geq_L G_n$, as this latter implies that $T(F_n) \geq T(G_n)$. Let $\tilde{F}_n$, $\tilde{G}_n$ be the linear interpolators of the jump points of F_n, G_n, respectively. If $F_n \geq_*^e G_n$, then we can take $\phi = \tilde{F}_n^{-1} \circ \tilde{G}_n$ in Definition 4. The function ϕ coincides with the linear interpolator of $F_n^{-1} \circ G_n$, which is starshaped by construction. Then, since $\tilde{F}_n$, $\tilde{G}_n$ are continuous, $\tilde{F}_n \geq_* \tilde{G}_n$ and the ratio $\frac{\tilde{F}_n^{-1}}{\tilde{G}_n^{-1}}$ is nondecreasing (Shaked and Shanthikumar, 2007, p. 214), the sequence $\frac{F_n^{-1}(\frac{i}{n})}{G_n^{-1}(\frac{i}{n})}$ is nondecreasing for $i = 1, \ldots, n$, whereas $\frac{F_n^{-1}(p)}{G_n^{-1}(p)}$ is constant between $\frac{i-1}{n}$ and $\frac{i}{n}$, for $i = 2, \ldots, n$. Hence, the function $R = \frac{F_n^{-1}}{G_n^{-1}}$ is a nondecreasing step function. Without loss of generality, let $\mu_X = \mu_Y = 1$. Then $L_{F_n}(p) - L_{G_n}(p) = \int_0^p (F_n^{-1}(u) - G_n^{-1}(u))\, du$. If R is nondecreasing, and the means are equal, the quantile functions must cross, and $R(\cdot) - 1$ must have one sign change. However, $R(\cdot) - 1$ has the same sign as $F_n^{-1} - G_n^{-1}$, therefore the argument of Theorem 4.B.4 of Shaked and Shanthikumar (2007) implies the result. ∎

Now, the following result is an immediate consequence of Theorem 5.

Corollary 10. *If $I \in \mathcal{I}(\geq_L)$ and $X \geq_* Y$, then $I(\mathbb{F}_n) \geq_{st} I(\mathbb{G}_n)$.*

In the next subsections, we will focus on some particularly relevant families of inequality measures.

3.1. Expected transformations. It is well known that $X \geq_L Y$ if and only if $\mathbb{E}(\phi(\frac{X}{\mu_X})) \geq \mathbb{E}(\phi(\frac{Y}{\mu_Y}))$, for every convex function ϕ (Marshall et al., 2011). Therefore, any functional of the form

$$T_\phi(F) = \mathbb{E}\left(\phi\left(\frac{X}{\mu_X}\right)\right) = \int_0^\infty \phi\left(\frac{x}{\mu_X}\right) dF(x) = \int_0^1 \phi\left(\frac{F^{-1}(p)}{\mu_X}\right) dp,$$

where ϕ is convex, is isotonic with $\geq_L$, as it is easily seen to satisfy the properties of inequality measures (Lando and Bertoli-Barsotti, 2016). Now, the corresponding statistical functional is $T_\phi(\mathbb{F}_n) = \frac{1}{n}\sum \phi(\frac{X_i}{\overline{X}_n})$, where $\overline{X}_n$ is the sample mean. Several well-known indices belong to this general family, among which we may note the class of *generalized entropy*, or *additively decomposable* measures of inequality (Shorrocks, 1980; Shorrocks and Slottje, 2002), obtained by setting $\phi_r(x) = \frac{1}{r(r-1)}x^r$, $r \neq 0, 1$, $\phi_0(x) = -\log x$, or $\phi_1(x) = x \log x$, respectively, which yield

$$I_r(\mathbb{F}_n) = \begin{cases} \frac{1}{r(r-1)n} \sum (\frac{X_i}{\overline{X}_n})^r & r \neq 0, 1 \\ \frac{1}{n} \sum \log(\frac{\overline{X}_n}{X_i}) & r = 0 \\ \frac{1}{n} \sum \frac{X_i}{\overline{X}_n} \log(\frac{X_i}{\overline{X}_n}) & r = 1. \end{cases}$$

This class gives, for $r = 1$, the Theil index, that is, a shifted version of the Shannon's entropy measure, applied to the quantities $\frac{X_i}{n\overline{X}_n}$ instead of probabilities; for $r \in (0, 1]$, a monotonic transformation of the Atkinson's class (Atkinson, 1970); and, for $r = 2$, a simple transformation of the coefficient of variation CV_n, that is, $(2I_r)^{\frac{1}{2}} = CV_n$. Moreover, when $\phi(x) = |x - 1|$, we obtain the relative mean absolute deviation $T_\phi(\mathbb{F}_n) = \frac{1}{n\overline{X}_n} \sum |X_i - \overline{X}_n|$. Then, the following result is an immediate consequence of Corollary 10.

Corollary 11. *If $X \geq_* Y$, then $\frac{1}{n} \sum \phi(\frac{X_i}{\overline{X}_n}) \geq_{st} \frac{1}{n} \sum \phi(\frac{Y_i}{\overline{Y}_n})$, for every convex function ϕ.*

Notice that Arendarczyk et al. (2021) proved that the Greenwood statistic has a stochastic increasing behavior with respect to the star order. However, since the Greenwood statistic is a transformation of the coefficient of variation, namely $(1 + CV_n^2)/n$, Theorem 1 of Arendarczyk et al. (2021) follows as a consequence of Corollary 11.

3.2. Distorted expectations. Let H be a *distortion function*, that is a non-decreasing function on the unit interval, such that $H(0) = 0$ and $H(1) = 1$, and let $\tilde{H}(p) = 1 - H(1 - p)$ be the corresponding dual distortion function. Probability functionals of the form

$$\begin{aligned} \mathbb{E}_H(F) &= \frac{1}{\mu_X} \int_0^\infty x \, dH \circ F(x) = \frac{1}{\mu_X} \int_0^1 F^{-1}(p) \, dH(p) \\ &= \frac{1}{\mu_X} \int_0^\infty \tilde{H}(1 - F(x)) \, dx, \end{aligned}$$

are generally referred to as distorted expectations, distortion risk measures, or Gini-type functionals (Wang and Young, 1998; Muliere and Scarsini, 1989; Lando and Bertoli-Barsotti, 2020). It can be seen that $X \geq_L Y$ if and only if $\mathbb{E}_H(F) \leq (\geq)\mathbb{E}_H(G)$, for every concave (convex) distortion H. Then, $\mathbb{E}_H$ is antitonic with respect to $\geq_L$, provided that H is concave. If we denote by $X_{(1)}, \ldots, X_{(n)}$ the order statistics of a random sample from X, the corresponding statistical functional is

$$\mathbb{E}_H(\mathbb{F}_n) = \frac{1}{\overline{X}_n} \sum X_{(i)} \int_{\frac{i-1}{n}}^{\frac{i}{n}} dH(p) = \frac{1}{\overline{X}_n} \sum X_{(i)} \left[H\left(\frac{i}{n}\right) - H\left(\frac{i-1}{n}\right)\right].$$

Such linear combinations of order statistics are generally referred to as *L-statistics* (Serfling, 1984). Several important indices belong to this family. For instance, by choosing $H(p) = 1 - (1-p)^k$, $k \geq 1$, we obtain the class of generalized Gini indices $\Gamma_H = 1 - \mathbb{E}_H$, introduced by Donaldson and Weymark (1980). In particular, the classic Gini coefficient of inequality (Gini, 1912) is given by $\Gamma = \Gamma_H$, with $H(p) = 1 - (1-p)^2$, that is

$$\begin{aligned} \Gamma(\mathbb{F}_n) &= 1 - \frac{1}{n^2 \overline{X}_n} \sum X_{(i)}(2n - 2i + 1) = 1 - 2 \int_0^1 L_{\mathbb{F}_n}(p)\, dp \\ &= \frac{\sum_{i=1}^n \sum_{j=1}^n |X_{(i)} - X_{(j)}|}{2n^2 \overline{X}_n}. \end{aligned}$$

As for the expressions above, we recall that there are several alternative ways to represent Γ (see Chapter 1 of Yitzhaki and Schechtman (2013)). For example, note that, if k is a positive integer, $\int_0^\infty (1 - F(x))^k\, dx = \mathbb{E}(\min(X_1, \ldots, X_k))$, so that $\Gamma(F) = 1 - \frac{\mathbb{E}(\min(X_1, X_2))}{\mu_X}$.

Similarly, the functional $\frac{1}{\mu_X} \int_0^\infty F^{-1}(p) h(p)\, dp$, where h is non-decreasing and such that $H(p) = \int_0^p h(t)\, dt$, $p \in [0,1]$, is isotonic with respect to $\geq_L$. If we set $w(p) = h(p) - 1$, $p \in [0,1]$, without loss of generality in terms of isotonicity, we obtain the family of *linear inequality measures*

$$\tilde{\Gamma}_w(F) = \frac{1}{\mu_X} \int_0^1 F^{-1}(p) w(p)\, dp,$$

studied by Mehran (1976). In particular, it can be shown that the weight function $w(p) = 2p - 1$ yields again the Gini coefficient, whereas $w(p) = 0$, $p \in (0,1)$, $w(0) = -1$, $w(1) = 1$ gives the relative range $\tilde{\Gamma}_w(\mathbb{F}_n) = \frac{1}{\overline{X}_n}(X_{(n)} - X_{(1)})$. It is easy to see that Γ_H and $\tilde{\Gamma}_w$, as probability functionals, satisfy the four properties of inequality measures discussed earlier. Again, Corollary 10 gives the following result.

Corollary 12. *If $X \geq_* Y$, then $\Gamma_H(\mathbb{F}_n) \geq_1 \Gamma_H(\mathbb{G}_n)$, for every concave distortion function H, and $\tilde{\Gamma}_w(\mathbb{F}_n) \geq_{st} \tilde{\Gamma}_w(\mathbb{G}_n)$, for every non-decreasing weight function w on $[0,1]$ such that $\int_0^1 w(p)\,dp = 0$.*

3.3. Some useful remarks. One may wonder if, in Theorem 5, a transform order may be replaced by a different kind of order $\succ$, that is, if, in general, $F \succ G$ and $T \in \mathcal{I}(\succ)$ imply $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$, where $\succ$ is not a transform order. For instance, can one replace $\geq_*$ by $\geq_L$ in Corollary 10? More generally, one may even wonder if the stochastic ordering assumption between the baseline distributions could be relaxed: for instance, would $T(F) \geq T(G)$ suffices to obtain $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$? (This would also mean that $T(F) = T(G)$ implies $T(\mathbb{F}_n) =_d T(\mathbb{G}_n)$). The following counterexample provides a negative answer to these conjectures.

Let X and Y be the discrete random variables with uniform probabilities on the supports $\{1, 3.5, 6, 6.5, 9, 11\}$ and $\{2, 3, 5, 7, 7.5, 10\}$, respectively. It is easily seen that $F \geq_L G$ and $\Gamma(F) > \Gamma(G)$, but $F \ngeq_*^e G$. Note that we focus on discrete distributions, because, using continuous parametric models, it is difficult to find instances such that $F \geq_L G$ and $F \ngeq_*^e G$. In the discrete case, the distributions of $\Gamma(\mathbb{F}_n)$ and $\Gamma(\mathbb{G}_n)$ are also discrete, with finite support included in the unit interval. Therefore, we may obtain a precise approximation of these distributions, for small sample size, by using a large number of simulation runs. An approximate representation of the cumulative distribution functions of $\Gamma(\mathbb{F}_n)$ and $\Gamma(\mathbb{G}_n)$, based on one million random samples of size $n = 3$ from F and G, is shown in Figure 1. The functions are clearly crossing, hence $\Gamma(\mathbb{F}_n) \ngeq_{st} \Gamma(\mathbb{G}_n)$. In particular, we observe that these functions exhibit some "bumps" within different intervals, and this represents an obstruction for the dominance relation. When the distributions are star ordered ($F \geq_*^e G$), such bumps occur within some overlapping intervals, indeed the dominance relation is guaranteed by Corollary 10. We obtained a similar behavior for other small sample sizes, such as $n = 4, 5$. This result is quite unexpected, however, we may conclude that both conjectures above are false. In particular, the behavior of the functionals is not "under control" if the sample size is small, whereas we know that, for large sample sizes, under some conditions (which are satisfied for probability functionals of the form $\int t(x)\,dF(x)$, such as transformed and distorted expectations),

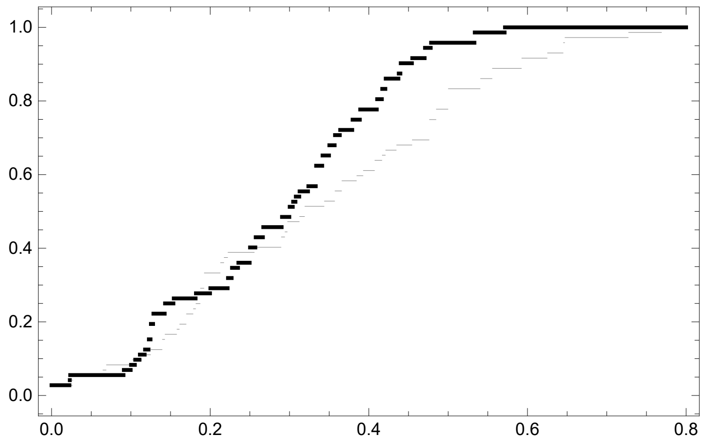

FIGURE 1. Cumulative distribution functions of the Gini coefficient, generated by 1000000 samples from F (thick) and G (thin).

statistical functionals converge to the constant value $T(F)$. Therefore, transform orders provide suitable tools for controlling the stochastic behavior of statistical functionals.

4. Application to goodness-of-fit tests

In the statistical literature, tests for the null hypothesis

$$\mathcal{H}_0 : G^{-1}F \in \mathcal{C}, \tag{1}$$

where F is the distribution of interest and G is known, may be particularly interesting. Hereafter we will focus on the case in which $G(0) = 0$, to avoid some technical issues. Tests for $\mathcal{H}_0 : G \geq_c F$ and $\mathcal{H}_0 : G \geq_* F$ have several applications: for example, by choosing G to be the unit exponential distribution, we obtain tests for the *increasing hazard rate* (the hazard rate is the derivative of $-\log(1-F)$) and the *increasing hazard rate average* properties, respectively, which are fundamental tools for survival analysis and reliability theory (Marshall and Olkin, 2007). Tests of this kind have been studied, for example, by Proschan and Pyke (1967), Barlow and Proschan (1969), Bickel (1969), Bickel and Doksum (1969), Deshpande (1983), Tenga and Santner (1984), Hall and Van Keilegom (2005), Durot (2007), Groeneboom and Jongbloed (2012), Gijbels and Heckman (2004), Lando et al. (2021), or Lando (2021), among many other authors.

4.1. Testing convexity. Let us define the *greatest convex minorant* of a function ϕ, as the largest convex function ϕ_c that does not exceed ϕ (see Tenga and Santner (1984) or Lando et al. (2021) for an explicit characterization of this function). An intuitive way of testing $\mathcal{H}_0 : G \geq_c F$ (Tenga and Santner, 1984; Durot, 2007; Groeneboom and Jongbloed, 2012; Lando et al., 2021) consists in measuring the distance between $G^{-1} \circ \mathbb{F}_n$ and its greatest convex minorant $(G^{-1} \circ \mathbb{F}_n)_c$, that is

$$T_c(\mathbb{F}_n) = \sup_{S_n} |G^{-1} \circ \mathbb{F}_n - (G^{-1} \circ \mathbb{F}_n)_c| = \sup_{S_n} (G^{-1} \circ \mathbb{F}_n - (G^{-1} \circ \mathbb{F}_n)_c),$$

where S_n is the set of points $X_{(2)}, \ldots, X_{(n-1)}$ ($X_{(1)}, X_{(n)}$ are neglected because the difference $G^{-1} \circ \mathbb{F}_n - (G^{-1} \circ \mathbb{F}_n)_c$ is always $\frac{1}{n}$ there). A weighted version of $T_c(\mathbb{F}_n)$ may also be considered without loss of generality. Of course, since $\mathcal{H}_0$ is a composite hypothesis, the determination of the *least favorable distribution* of $T_c(\mathbb{F}_n)$ under $\mathcal{H}_0$ is especially critical. With regard to the increasing hazard rate hypothesis, this distribution is indeed obtained by simulating from the unit exponential, as proved by Tenga and Santner (1984). We remark that a more general result may be achieved as an application of Theorem 7. As expected, the least favorable distribution for $T_c(\mathbb{F}_n)$ under $\mathcal{H}_0 : G \geq_c F$, is determined just by simulating from G. Let $\Phi_p^G(F) = (G^{-1} \circ F)_c(F^{-1}(p))$. The following lemma is a consequence of Theorem 2.2 of Tenga and Santner (1984), and it establishes the order-reversing behavior of Φ_p^G.

Lemma 13. *Assume that $G \geq_c F$. Then $\Phi^G_{\frac{i}{n}}(G_n) \leq \Phi^G_{\frac{i}{n}}(F_n^*)$, for $i = 1, \ldots, n$.*

Then, Theorem 7 gives the following result. Note that this has been already obtained by Lando et al. (2021), however, here we show it to illustrate our method, as a particular case of a more general family of testing procedures.

Theorem 14. *Under $\mathcal{H}_0 : G \geq_c F$, $T_c(\mathbb{G}_n) \geq_{st} T_c(\mathbb{F}_n)$.*

Proof: Clearly, by Lemma 13, $-\Phi^G_{\frac{i}{n}}(G_n) \geq -\Phi^G_{\frac{i}{n}}(F_n^*)$. Note that $T_c(\mathbb{F}_n) = \sup_{i=2,\ldots,n-1}(G^{-1}(\frac{i}{n}) - \Phi^G_{\frac{i}{n}}(\mathbb{F}_n))$, then Theorem 7 implies the result. ■

Therefore, we reject $\mathcal{H}_0$ when $T_c(F_n) \geq c_{\alpha,n}$, where $c_{\alpha,n}$ is the solution of $P(T_c(\mathbb{G}_n) \geq c_{\alpha,n}) \geq \alpha$, and α is the size of the test. Alternatively, we can compute the p-value of the test, that is, $p = P(T_c(\mathbb{G}_n) \geq T_c(F_n))$.

4.2. Testing starshapedness. Note that $G \geq_* F$ requires that $F(0) = 0$. Define the *greatest starshaped minorant* of a function ϕ, as the largest starshaped function ϕ_* that does not exceed ϕ. The same arguments used earlier may determine the least favorable distribution of the test statistic

$$T_*(\mathbb{F}_n) = \sup_{S_n} |G^{-1} \circ \mathbb{F}_n - (G^{-1} \circ \mathbb{F}_n)_*| = \sup_{S_n} (G^{-1} \circ \mathbb{F}_n - (G^{-1} \circ \mathbb{F}_n)_*),$$

under $\mathcal{H}_0 : G \geq_* F$.

Let $\Psi_p^G(F) = (G^{-1} \circ F)_*(F^{-1}(p))$. The following lemma establishes the order-reversing behavior of Ψ_p^G, corresponding to Lemma 13.

Lemma 15. *Assume that $G \geq_* F$. Then $\Psi_{\frac{i}{n}}^G(G_n) \leq \Psi_{\frac{i}{n}}^G(F_n^*)$, for $i = 1, \ldots, n$.*

Proof: To simplify notations, let $y_1, \ldots, y_n$ be ordered realizations from G. The function $u = F^{-1} \circ G$ is, by assumption, anti-starshaped, and recall that $x_i^* = u(y_i)$ are the observations which determine F_n^*. Note that $G^{-1} \circ G_n(y_i) = G^{-1}(\frac{i}{n})$. We now describe explicitly $(G^{-1} \circ G_n)_*$ (see Wang (1988)). Let $a_i = \frac{1}{y_i} G^{-1}(\frac{i-1}{n})$ be the slope of the line connecting the origin to the point $(y_i, G^{-1}(\frac{i-1}{n}))$. The analytical expression of the greatest starshaped minorant is given by $(G^{-1} \circ G_n)_*(x) = \alpha_i x$, for $x \in [y_{i-1}, y_i)$, where $y_0 = 0$ and $\alpha_i = \min\{a_j,\ j = i, \ldots, n\}$ (note that $\alpha = 0$, and, moreover, starshapedness requires that the slopes are increasing). $(G^{-1} \circ F_n^*)_*$ is defined similarly. Note that the two functions have different domains, but coincide at corresponding points: $G^{-1} \circ F_n^*((x_i^*)^-) = G^{-1} \circ G_n(y_i^-) = G^{-1}(\frac{i-1}{n})$. We want to prove that $(G^{-1} \circ F_n^*)_*(u(y_i)) \geq (G^{-1} \circ G_n)_*(y_i)$, for $i = 1, \ldots n$. It is sufficient to prove

$$\frac{G^{-1}(\frac{k-1}{n})}{u(y_k)} u(y_i) \geq \frac{G^{-1}(\frac{k-1}{n})}{y_k} y_i \iff \frac{u(y_i)}{y_i} \geq \frac{u(y_k)}{y_k},$$

for $k = i, \ldots, n$, which follows due to the anti-starshapedness of the function u, that is, $\frac{1}{x} u(x)$ is decreasing. ∎

Then, the following result is an immediate consequence of Theorem 7. The proof is omitted.

Theorem 16. *Under $\mathcal{H}_0 : G \geq_* F$, $T_*(\mathbb{G}_n) \geq_{st} T_*(\mathbb{F}_n)$.*

This is a novel result, which can be useful for testing goodness-of-fit to the increasing hazard rate average family (see, for instance, Deshpande (1983)) or some generalizations of such a property, as described by Barlow et al. (1971).

Remark 17. Note that the results obtained above may be further generalized. For instance we may consider other types of distances instead of the supremum norm, such as an empirical L_1 distance between the estimators, in the spirit of Groeneboom and Jongbloed (2012). Moreover, we may take, as a test statistic, any distance between $\mathbb{F}_n$ and $\mathbb{F}_n^G = G \circ (G^{-1} \circ \mathbb{F}_n)_c$ (Lando, 2021), that is the so-called *isotonic estimator* of F under $G \geq_c F$, as these are just monotonic transformations of $G^{-1} \circ \mathbb{F}_n$ and $(G \circ \mathbb{F}_n)_c$. The least favorable distribution of test statistics of these kinds may be similarly derived via the aforementioned method, and clearly the same arguments can be extended to tests for $G \geq_* F$.

4.3. A less conservative approach. As well known, the determination of the least favorable distribution of T_c yields conservative tests. In order to improve the power of the test against non-convex alternatives ($\mathcal{H}_1 : G \not\geq_c F$), a more modern approach consists in performing bootstrap resampling from the distribution $F_n^G = G \circ (G^{-1} \circ F_n)_c$, that is the isotonic estimate of F, instead of sampling from G (Hall and Van Keilegom, 2005; Groeneboom and Jongbloed, 2012). In particular, denote with $\hat{\mathbb{F}}_m^G$ the (bootstrap) empirical distribution obtained by a random sample of size m from F_n^G. The following result holds.

Theorem 18. *Under $\mathcal{H}_0 : G \geq_c F$, $T_c(\mathbb{G}_n) \geq_{st} T_c(\hat{\mathbb{F}}_m^G)$.*

Proof: Let F_n^G be the s estimate of F which corresponds to the empirical distribution F_n^*. F_n^G is a realization of $\hat{\mathbb{F}}_m^G$. Lemma 13 yields $\Phi^G_{\frac{i}{n}}(G_n) \leq \Phi^G_{\frac{i}{n}}(F_n^*)$, for $i = 1, \ldots, n$. Note that F_n^G is a minorant of F_n^* (equivalently, $(F_n^*)^{-1}$ is a minorant of $(F_n^G)^{-1}$). Then, since $G^{-1} \circ F_n^G = (G^{-1} \circ F_n^*)_c$, we obtain

$$\begin{aligned}
\Phi^G_{\frac{i}{n}}(G_n) \leq \Phi^G_{\frac{i}{n}}(F_n^*) &= (G^{-1} \circ F_n^*)_c((F_n^*)^{-1}(\frac{i}{n})) \\
&\leq (G^{-1} \circ F_n^*)_c((F_n^G)^{-1}(\frac{i}{n})) = (G^{-1} \circ F_n^G)_c((F_n^G)^{-1}(\frac{i}{n})) = \Phi^G_{\frac{i}{n}}(F_n^G).
\end{aligned}$$

Hence, as shown in the proof of Theorem 14, we obtain that $T_c(\mathbb{G}_n) \geq_{st} T_c(\hat{\mathbb{F}}_m^G)$. ■

Since quantiles are isotonic with $\geq_{st}$, Theorem 18 ensures that, with this method, we obtain a smaller critical value and, correspondingly, a higher value of the power function under $\mathcal{H}_1$. The same arguments apply to tests for $G \geq_* F$.

Summing up, Theorem 7 provides a flexible approach for determining the least favorable distribution, as well as the behavior of bootstrap statistics, in different kinds of testing procedures, especially when the test statistic, or a suitable transformation of it, may be shown to fulfil an order-preserving property with respect to a transform order.

References

[1] Arendarczyk, M., Kozubowski, T.J., Panorska, A.K., 2021. The Greenwood statistic, stochastic dominance, clustering and heavy tails. Scandinavian Journal of Statistics (early view). doi:10.1111/sjos.12520.

[2] Atkinson, A.B., 1970. On the measurement of inequality. Journal of Economic Theory 2, 244–263. doi:10.1016/0022-0531(70)90039-6.

[3] Barlow, R., Bartholomew, D., Bremner, J., Brunk, H., 1971. Statistical Inference Under Order Restrictions. Wiley.

[4] Barlow, R.E., Proschan, F., 1969. A note on tests for monotone failure rate based on incomplete data. Annals of Mathematical Statistics 40, 595–600. doi:10.1214/aoms/1177697727.

[5] Bickel, P., 1969. Tests for monotone failure rate II. Annals of Mathematical Statistics 40, 1250–1260. doi:10.1214/aoms/1177697500.

[6] Bickel, P.J., Doksum, K.A., 1969. Tests for monotone failure rate based on normalized spacings. Annals of Mathematical Statistics 40, 1216–1235. doi:10.1214/aoms/1177697498.

[7] Deshpande, J.V., 1983. A class of tests for exponentiality against increasing failure rate average alternatives. Biometrika 70, 514–518. doi:10.2307/2335571.

[8] Donaldson, D., Weymark, J.A., 1980. A single-parameter generalization of the gini indices of inequality. Journal of Economic Theory 22, 67–86. doi:10.1016/0022-0531(80)90065-4.

[9] Durot, C., 2007. On the $\mathbb{L}_p$-error of monotonicity constrained estimators. Annals of Statistics 35, 1080–1104. doi:10.1214/009053606000001497.

[10] Eberl, A., Klar, B., 2021. Expectile-based measures of skewness. Scandinavian Journal of Statistics doi:10.1111/sjos.12518.

[11] Gijbels, I., Heckman, N., 2004. Nonparametric testing for a monotone hazard function via normalized spacings. Journal of Nonparametric Statistics 16, 463–477. doi:10.1080/10485250310001622668.

[12] Gini, C., 1912. Variabilità e mutabilità. Reprinted in Memorie di metodologica statistica, Pizetti .

[13] Greenwood, M., 1946. The statistical study of infectious diseases. Journal of the Royal Statistical Society 109, 85–110. doi:10.2307/2981176.

[14] Groeneboom, P., Jongbloed, G., 2012. Isotonic L_2-projection test for local monotonicity of a hazard. Journal of Statistical Planning and Inference 142, 1644–1658. doi:10.1016/j.jspi.2012.02.004.

[15] Hall, P., Van Keilegom, I., 2005. Testing for monotone increasing hazard rate. Annals of Statistics 33, 1109–1137. doi:10.1214/009053605000000039.

[16] Lando, T., 2021. Testing convexity of the generalised hazard function. Statistical Papers (forthcoming). doi:10.1007/s00362-021-01273-w.

[17] Lando, T., Arab, I., Oliveira, P.E., 2021. Second-order stochastic comparisons of order statistics. Statistics 55. doi:10.1080/02331888.2021.1960527.

[18] Lando, T., Bertoli-Barsotti, L., 2016. Weak orderings for intersecting lorenz curves. Metron 74, 177–192. doi:10.1007/s40300-016-0087-6.

[19] Lando, T., Bertoli-Barsotti, L., 2020. Distorted stochastic dominance: a generalized family of stochastic orders. Journal of Mathematical Economics 90, 132–139. doi:10.1016/j.jmateco.2020.07.005.
[20] Marshall, A.W., Olkin, I., 2007. Life distributions. volume 13. Springer.
[21] Marshall, A.W., Olkin, I., Arnold, B.C., 2011. Inequalities: theory of majorization and its applications. Elsevier.
[22] Mehran, F., 1976. Linear measures of income inequality. Econometrica: Journal of the Econometric Society 44, 805–809. doi:10.2307/1913446.
[23] Muliere, P., Scarsini, M., 1989. A note on stochastic dominance and inequality measures. Journal of Economic Theory 49, 314–323. doi:10.1016/0022-0531(89)90084-7.
[24] Oja, H., 1981. On location, scale, skewness and kurtosis of univariate distributions. Scandinavian Journal of statistics 8, 154–168. doi:http://www.jstor.org/stable/4615828.
[25] Proschan, F., Pyke, R., 1967. Tests for monotone failure rate, in: Fifth Berkley Symposium, pp. 293–313.
[26] Serfling, R., 1984. Generalized L-, M-, and R-statistics. Annals of Statistics 12, 76–86. doi:10.1214/aos/1176346393.
[27] Shaked, M., Shanthikumar, J.G., 2007. Stochastic orders. Springer Series in Statistics.
[28] Shorrocks, A., Slottje, D., 2002. Approximating unanimity orderings: An application to Lorenz dominance. Journal of Economics 77, 91–117. doi:10.1007/BF03052501.
[29] Shorrocks, A.F., 1980. The class of additively decomposable inequality measures. Econometrica: Journal of the Econometric Society 48, 613–625. doi:10.2307/1913126.
[30] Tenga, R., Santner, T.J., 1984. Testing goodness of fit to the increasing failure rate family. Naval Research Logistics Quarterly 31, 617–630. doi:10.1002/nav.3800310411.
[31] Wang, J.L., 1988. Optimal estimations of starshaped distribution functions. Statistics & Decisions 6, 21–32. doi:10.1524/strm.1988.6.12.21.
[32] Wang, S.S., Young, V.R., 1998. Ordering risks: Expected utility theory versus yaari's dual theory of risk. Insurance: Mathematics and Economics 22, 145–161. doi:10.1016/S0167-6687(97)00036-X.
[33] Yitzhaki, S., Schechtman, E., 2013. The Gini methodology: A primer on a statistical methodology. Springer.
[34] van Zwet, W.R., 1964. Convex transformations of random variables. MC Tracts .

Tommaso Lando
SAEMQ, University of Bergamo (Italy) and VŠB-TU Ostrava (Czech Republic)
E-mail address: tommaso.lando@unibg.it

Idir Arab
CMUC, Department of Mathematics, University of Coimbra, Portugal
E-mail address: idir.bhh@gmail.com

Paulo Eduardo Oliveira
CMUC, Department of Mathematics, University of Coimbra, Portugal
E-mail address: paulo@mat.uc.pt