Transform orders and stochastic monotonicity of statistical functionals

Tommaso Lando[1,2] | **Idir Arab[3]** | **Paulo Eduardo Oliveira[3]**

[1]Department of Economics, University of Bergamo, Bergamo, Italy

[2]Department of Finance, VSB-TU Ostrava, Ostrava, Czech Republic

[3]CMUC, Department of Mathematics, University of Coimbra, Coimbra, Portugal

Correspondence
Tommaso Lando, Department of Economics, University of Bergamo, Bergamo, Italy.
Email: tommaso.lando@unibg.it

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 illustrate 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, hazard function, inequality, nonparametric test, stochastic dominance

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

cumulative distribution functions (CDFs). 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 CDF $\mathbb{F}_n$. 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 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, namely, we are interested in establishing relationships between their distribution functions that hold for any given sample, regardless of its size.

Let T be some functional, and let $\mathbb{F}_n$ and $\mathbb{G}_n$ be the empirical CDFs 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 2). This is a relevant information as it may be used to help characterizing finite sample properties of testing procedures, as we will be describing later. 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, specifically, functionals that are compatible with a general class of *transform orders*, generated by some family of reference functions (see Definition 1 below).

Our main result, Theorem 1, presented in Section 2, shows 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 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 stochastic order, the convex transform order, the star order, the superadditive and the dispersive orders (Shaked & Shanthikumar, 2007). Therefore, we are able to derive stochastic properties of several important families of functionals.

As an immediate application of Theorem 1, in Section 3 we obtain the stochastic monotonicity of many important measures of inequality, including the generalized entropy class (Shorrocks, 1980; Shorrocks & Slottje, 2002), the well-known Gini index (Gini, 1912) and some of its generalizations (Donaldson & Weymark, 1980; Mehran, 1976). Differently, in Section 4, we use Theorem 1 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).

2 | MAIN RESULTS

We begin by establishing some notations. The random variables X and Y have CDFs 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 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 $G \succ F$ do 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 $\succ$, that is,

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

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

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

2.1 | A general family of orders

We now introduce the following definition of a transform order, generated by a family of functions C, whose choice enables the generalization of some well-known stochastic orders. As usual, $f|_E$ denotes the restriction of a function f to the set E.

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

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 $\geq_C$ fulfills the basic properties of stochastic orders, and, moreover, $C_0 \subset C_1$ means that $F \geq_{C_0} G \Rightarrow F \geq_{C_1} G$. As limiting cases, if C is the class of nondecreasing functions, then $\geq_C$ is always verified, whereas if C contains just the identity function then $\geq_C$ coincides with equality in distribution, $=_d$. The usual stochastic order, generally defined by the relation $F(x) \leq G(x)$, for every $x \in \mathbb{R}$ (Shaked & Shanthikumar, 2007), is a transform order, obtained by choosing C as the class of functions ψ such that $\psi(x) \geq x$. Other well-known transform orders may be obtained by appropriately choosing C: the classes of convex, star-shaped, or superadditive functions, yield the *convex transform order*, *the star order*, or the *superadditive order*, respectively. Also the *dispersive order* can be obtained by choosing C as the class of functions ψ such that $\psi(x) - x$ is nondecreasing (Shaked & Shanthikumar, 2007). Since these orders are particularly important, we recall their definitions below (remember that $F^{-1} \circ G$ is always nondecreasing by construction). However, Definition 1 above is quite general and enables the possibility of defining new stochastic orders.

Definition 2. We say that X dominates Y with respect to

1. the usual stochastic order, denoted by $X \geq_{st} Y$, if $F^{-1} \circ G|_{S_G}(x) \geq x$, or equivalently, $F(x) \leq G(x)$, for every $x \in \mathbb{R}$;
2. the convex transform order, denoted by $X \geq_c Y$, if $F^{-1} \circ G|_{S_G}$ is convex;
3. the star order, denoted by $X \geq_* Y$, if $F^{-1} \circ G|_{S_G}$ is star-shaped;
4. the superadditive order, denoted by $X \geq_{su} Y$, if $F^{-1} \circ G|_{S_G}$ is superadditive;
5. the dispersive order, denoted by $X \geq_{disp} Y$, if $(F^{-1} \circ G(x) - x)|_{S_G}$ is increasing.

For nonnegative random variables, the following relations hold:

$$X \geq_c Y \;\Rightarrow\; X \geq_* Y \;\Leftrightarrow\; \log X \geq_{disp} \log Y \;\Rightarrow\; X \geq_{su} Y.$$

In the sequel, we will further extend Definition 2 to enable comparisons of (discrete) empirical distributions, which is necessary to obtain the desired stochastic monotonicity between statistical functionals.

2.2 | Stochastic monotonicity results

We are concerned with the problem of ranking statistical functionals in terms of the usual stochastic order. Therefore, we are interested in determining the conditions on F, G and T such that $P(T(\mathbb{F}_n) \leq x) \leq P(T(\mathbb{G}_n) \leq x)$ holds for every $x \in \mathbb{R}$. Results of this kind are not predictable in general. 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 $X > Y$ 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 result, Theorem 1, shows 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, some transform orders may be not defined for realizations, F_n and G_n, of the empirical distribution functions $\mathbb{F}_n$ and $\mathbb{G}_n$, as these realizations are discrete. For instance, in the empirical case, the composition $F_n^{-1} \circ G_n$ is a step function, hence neither convex nor star-shaped, so these CDFs are not comparable in terms of $\geq_c$ or $\geq_*$. Therefore, to prove our main results, we must *extend*, when needed, any transform order to the class of observed empirical distributions, according to the following general definition.

Definition 3. Let C be some family of nondecreasing functions. We say that X dominates Y with respect to the extended C-transform order, denoted as $X \geq_C^e Y$, if there exists a function $\phi \in C$ such that $\phi|_{S_G} = F^{-1} \circ G|_{S_G}$.

Note the distinction between Definitions 1 and 3: the former requires that $F^{-1} \circ G$ satisfies some property, described by C, on the support of G, while the latter assumes the existence of some function $\phi \in C$, 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 is useful because, for instance, notions such as convexity and star-shapedness are generally not defined for functions whose support is a discrete set of points. In such cases, Definition 3 makes it possible to consider C as the class of convex or star-shaped functions, even when F and G are empirical distributions, which is out of the scope of Definition 1. Indeed, with regard to the pair F_n, G_n, where $S_{G_n} = \{y_1, \ldots, y_n\}$, the ordering $F_n \geq_C^e G_n$ holds if there exists a function $\phi \in C$ such that $\phi(y_i) = F_n^{-1} \circ G_n(y_i)$, $i = 1, \ldots, n$. Moreover, note that $\geq_C^e$ is a weaker version of $\geq_C$, that is, $X \geq_C Y$ implies $X \geq_C^e Y$.

We may now establish our main result. 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 $x_1^*,\ldots,x_n^*$. Then, we obtain the following sufficient conditions for $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$.

Theorem 1.

1. If $X \geq_C^e Y$ and $T \in \mathcal{I}(\geq_C^e)$, then $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$.
2. If $X \geq_C^e Y$ and $T(F_n^*) \geq T(G_n)$, then $T(\mathbb{F}_n) \geq_{st} T(\mathbb{G}_n)$.

Proof. 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 $X \geq_C^e Y$, then there exists some $\phi \in C$ satisfying Definition 3. Let Z be a random variable with distribution H and support S_H included in S_G. $X \geq_C^e Y$ 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 $F^{-1} \circ G(Z) \geq_C^e Z$, for every random variable Z with support included in that of Y. Given that the support of $\tilde{Y}_n$ is always included in the support of Y, this yields $\tilde{X}_n =_d F^{-1} \circ G(\tilde{Y}_n) \geq_C^e \tilde{Y}_n$. Now, in case 1), as we are assuming that $T \in \mathcal{I}(\geq_C^e)$, it follows that $T(F_n^*) \geq T(G_n)$, while, in case 2), this holds by assumption. 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 pair 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 to the probability integral transform, see theorem 1.A.1 of Shaked & Shanthikumar, 2007). ∎

Theorem 1 demonstrates the practical usefulness of our Definition 3. In fact, transform orders may sometimes not be meaningful, especially when choosing an unusual reference class C. However, they may be employed merely as tools for deriving stochastic properties of statistical functionals. For instance, to apply part 1 of Theorem 1, it is enough to show that the functional of interest T belongs to some isotonic family, $\mathcal{I}(\geq_C^e)$, and to do so, one may use existing results which ensure that $T \in \mathcal{I}(\geq_C^e)$. In particular, as we discuss in the next section, this approach works when C is the class of functions that defines the star order or the convex transform order, enabling the determination of the stochastic monotonicity of the main measures of inequality and skewness. Obviously, if T is isotonic with more than one of the orders in Definition 3, 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 $T \in \mathcal{I}(\geq_C^e)$, for instance by choosing a wider class C or by defining C according to the properties of T. Alternatively, we may apply part 2 of Theorem 1, replacing the isotonicity assumption, $T \in \mathcal{I}(\geq_C^e)$, with the weaker order-preserving property $T(F_n^*) \geq T(G_n)$.

Both approaches may be useful, depending on the situation. Several results available in the literature assert the isotonicity of many popular probability functionals, with respect to some of the transform orders mentioned above, hence part 1 immediately provides the ordering between the corresponding statistical functionals, as discussed in the following section. When such isotonicity results are not available, part 2 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, economics, 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 & 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 a skewness measure (Eberl & Klar, 2021; Oja, 1981; van Zwet, 1964), therefore our method may be also applied to skewness measures, but this is beyond the purpose of this paper.

Hereafter, we focus on nonnegative random variables with finite mean. Let $L_F(p) = \int_0^p F^{-1}(t)\, dt/\mu_X$, $p \in [0,1]$ be the *Lorenz curve* of F (as usual, μ_X represents the mean of X). 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 nonintersecting Lorenz curves shows less inequality. This gives rise to the definition of the *Lorenz order*.

Definition 4. 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 in 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 & 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.

Lemma 1. $\mathcal{I}(\geq_L) \subset \mathcal{I}(\geq_*^e)$.

Proof. Choose $T \in \mathcal{I}(\geq_L)$. The inclusion follows if we prove that $F_n \geq_*^e G_n \Rightarrow F_n \geq_L G_n$, as this latter implies that $T(F_n) \geq T(G_n)$. Let $\tilde{F}_n$ and $\tilde{G}_n$ be the linear interpolators of the jump points of F_n and 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 3. The function ϕ coincides with the linear interpolator of $F_n^{-1} \circ G_n$, which is star-shaped by construction. Then, since $\tilde{F}_n, \tilde{G}_n$ are continuous, $\tilde{F}_n \geq_* \tilde{G}_n$ and the ratio $\tilde{F}_n^{-1}/\tilde{G}_n^{-1}$ is nondecreasing (Shaked & Shanthikumar, 2007, p. 214), the sequence $F_n^{-1}(i/n)/G_n^{-1}(i/n)$ is nondecreasing for $i = 1,\ldots,n$, whereas $F_n^{-1}(p)/G_n^{-1}(p)$ is constant between $(i-1)/n$ and i/n, for $i = 2,\ldots,n$.

Hence, the function $R = F_n^{-1}/G_n^{-1}$ 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 1 and Lemma 1.

Corollary 1. *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 | Expectation-type transformations

It is well known that $X \geq_L Y$ if and only if $\mathbb{E}(\phi(X/\mu_X)) \geq \mathbb{E}(\phi(Y/\mu_Y))$, 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 & Bertoli-Barsotti, 2016). Now, the corresponding statistical functional is $T_\phi(\mathbb{F}_n) = \sum \phi(X_i/\overline{X}_n)/n$, where $X_1, \ldots, X_n$ is a random sample from X and $\overline{X}_n$ 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 & Slottje, 2002), obtained by setting $\phi_r(x) = x^r/(r(r-1))$, $r \neq 0, 1$, $\phi_0(x) = -\log x$, or $\phi_1(x) = x \log x$, respectively, which yield

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

This class gives, for $r = 1$, the Theil index, that is, a shifted version of the Shannon's entropy measure, applied to the quantities $X_i/(n\overline{X}_n)$ 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, $(2I_2)^{1/2} = CV_n$. Moreover, when $\phi(x) = |x - 1|$, we obtain the relative mean absolute deviation $T_\phi(\mathbb{F}_n) = \sum |X_i - \overline{X}_n|/(n\overline{X}_n)$. Then, the following result is an immediate consequence of Corollary 1.

Corollary 2. *If $X \geq_* Y$, then $\sum \phi(X_i/\overline{X}_n)/n \geq_{st} \sum \phi(Y_i/\overline{Y}_n)/n$, 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 2.

3.2 | Distorted expectations

Let H be a *distortion function*, that is a nondecreasing 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

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

are generally referred to as distorted expectations, distortion risk measures, or Gini-type functionals (Lando & Bertoli-Barsotti, 2020; Muliere & Scarsini, 1989; Wang & Young, 1998). 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

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

As for the expressions above, we recall that there are several alternative ways to represent Γ (see chapter 1 of Yitzhaki & 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 $\Gamma(F) = 1 - \mathbb{E}(\min(X_1, X_2))/\mu_X$.

Similarly, the functional $\int_0^\infty F^{-1}(p)h(p) \, dp/\mu_X$, where h is nondecreasing 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 $\tilde{\Gamma}_w(\mathbb{F}_n) = (X_{(n)} - X_{(1)})/\overline{X}_n$. 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 1 gives the following result.

Corollary 3. *If* $X \geq_* Y$, *then* $\Gamma_H(\mathbb{F}_n) \geq_{st} \Gamma_H(\mathbb{G}_n)$, *for every concave distortion function H, and* $\tilde{\Gamma}_w(\mathbb{F}_n) \geq_{st} \tilde{\Gamma}_w(\mathbb{G}_n)$, *for every nondecreasing 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 1, 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 1? 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 discrete random variables with uniform distributions 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 the Gini indices are such that $\Gamma(F) > \Gamma(G)$, but $F \not\geq_*^e G$. 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 $F \not\geq_*^e G$. 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 sizes, by using a large number of simulation runs. An approximate representation of the CDFs 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 $\Gamma(\mathbb{F}_n) \not\geq_{st} \Gamma(\mathbb{G}_n)$. 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 1. 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), statistical functionals converge to the constant value $T(F)$. Therefore, transform orders provide suitable tools for controlling the stochastic behavior of statistical functionals.
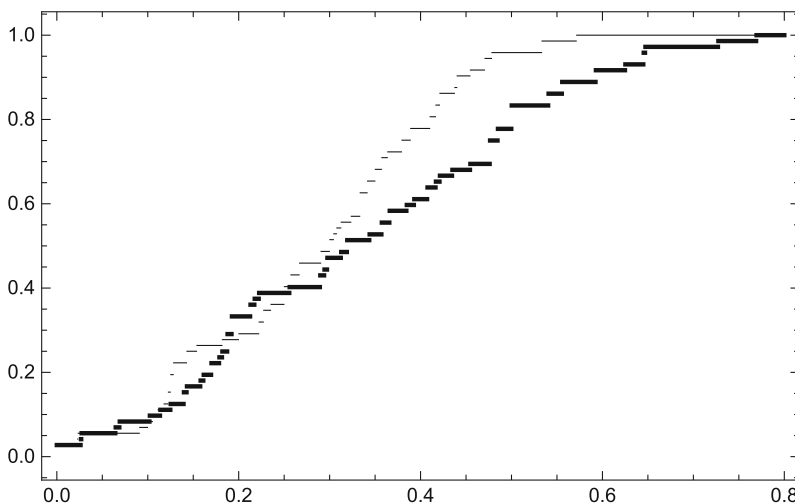

FIGURE 1   Cumulative distribution functions of the Gini coefficient, generated by 10^6 samples from F (thick) and G (thin)

4 | APPLICATION TO GOODNESS-OF-FIT TESTS

In the statistical literature, tests of the null hypothesis $\mathcal{H}_0^R : G^{-1} \circ F|_R \in \mathcal{C}$, versus the alternative $\mathcal{H}_1^R : G^{-1} \circ F|_R \notin \mathcal{C}$, where $R \subset S_F$ is some interval, 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. Given G, $G^{-1} \circ F$ is called the *generalized hazard function* of F (Barlow et al., 1971). Tests for $G \geq_c F$ and $G \geq_* F$ have several applications: for example, by choosing $G(x) = \mathcal{E}(x) := 1 - \exp(-x)$, $x \geq 0$, namely, the unit exponential distribution, we obtain tests for the *increasing hazard rate* (IHR) and the *increasing hazard rate average* (IHRA) properties, respectively, which are fundamental tools for survival analysis and reliability theory (Marshall & Olkin, 2007). Nonparametric tests of these properties have been studied extensively in the literature, most of which deal with the null hypothesis of exponentiality ($F = \mathcal{E}$) versus the alternative that F is IHR (Barlow & Proschan, 1969; Bickel, 1969; Bickel & Doksum, 1969; Proschan & Pyke, 1967) or IHRA (Ahmad, 1994; Deshpande, 1983; Klefsjö, 1983; Kochar, 1985; Link, 1989; Wells & Tiwari, 1991), but not exponential. In such cases, it is clear that the distribution of the test statistic can be obtained using $F = \mathcal{E}$. Differently, we study goodness-of-fit tests where the determination of the distribution of the test statistic under the (composite) null hypothesis is not obvious and it represents an important issue (Durot, 2007; Gijbels & Heckman, 2004; Groeneboom & Jongbloed, 2012; Hall & van Keilegom, 2005; Lando, 2022; Lando et al., 2021; Tenga & Santner, 1984). As shown in the sequel, part 2 of Theorem 1 may be used to establish stochastic ordering properties of test statistics under some circumstances of interest. These results make it possible to control the behavior of the test for a given sample size, which is an important aspect, comparable to asymptotic properties.

In the next subsections, we shall remove the restriction R and just consider $\mathcal{H}_0 : G^{-1} \circ F \in \mathcal{C}$ versus $\mathcal{H}_1 : G^{-1} \circ F \notin \mathcal{C}$, to simplify the notations. It is easily seen that the results in Sections 4.1 and 4.2 hold even in the restricted case. Differently, a restricted null hypothesis is considered in Section 4.2.1.

4.1 | Testing convexity

Define the *greatest convex minorant* of a function ϕ as the largest convex function ϕ_c that does not exceed ϕ (see Tenga & Santner, 1984 for a formal definition of this function). An intuitive way of testing $\mathcal{H}_0 : G \geq_c F$ versus $\mathcal{H}_1 : G \not\geq_c F$ consists in measuring the distance between $G^{-1} \circ \mathbb{F}_n$ and its greatest convex minorant $(G^{-1} \circ \mathbb{F}_n)_c$, or, more generally,

$$
\begin{aligned}
T_c^K(\mathbb{F}_n) &= \sup_{i=2,\ldots,n-1} \left| K \circ G^{-1} \circ \mathbb{F}_n(X_{(i-1)}) - K \circ (G^{-1} \circ \mathbb{F}_n)_c(X_{(i)}) \right| \\
&= \sup_{i=2,\ldots,n-1} \left(K \circ G^{-1}\left(\frac{i-1}{n}\right) - K \circ (G^{-1} \circ \mathbb{F}_n)_c(X_{(i)}) \right),
\end{aligned}
$$

where K is some strictly increasing transformation defined on S_G. This family of tests is quite general. Some interesting examples are obtained by choosing K as the identity function (Durot, 2007; Lando et al., 2021; Tenga & Santner, 1984); or setting $K = G$ (Groeneboom & Jongbloed, 2012; Lando, 2022), in this latter case $T_c^K(\mathbb{F}_n)$ boils down to a Kolmogorov–Smirnov-type test statistic. It can be seen that T_c^K is scale invariant, which is a basic condition for tests of this kind. Note that the difference is evaluated among the shifted points $X_{(i)}$ and $X_{(i-1)}$ because, by construction,

$G^{-1} \circ \mathbb{F}_n(X_{(i-1)}) \geq (G^{-1} \circ \mathbb{F}_n)_c(X_{(i)})$. A weighted version of $T_c^K(\mathbb{F}_n)$ 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 $T_c^K(\mathbb{F}_n)$ under $\mathcal{H}_0$ is especially critical. With regard to the IHR 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 1. As expected, the least favorable distribution of $T_c^K(\mathbb{F}_n)$ 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 2. *Assume that* $G \geq_c F$. *Then* $\Phi_{i/n}^G(G_n) \leq \Phi_{i/n}^G(F_n^*)$, *for* $i = 1, \ldots, n$.

Theorem 1 and Lemma 2 imply the following result. Note that this is a slight generalization of existing results (Lando et al., 2021; Tenga & Santner, 1984), however, here we show it to illustrate our method, as an immediate application of part 2 of Theorem 1.

Proposition 1. *Under* $\mathcal{H}_0 : G \geq_c F$, $T_c^K(\mathbb{G}_n) \geq_{st} T_c^K(\mathbb{F}_n)$.

Proof. Clearly, by Lemma 2, $\Phi_{i/n}^G(G_n) \leq \Phi_{i/n}^G(F_n^*)$. Note that

$$T_c^K(\mathbb{F}_n) = \sup_{i=2,\ldots,n-1} \left(K \circ G^{-1}\left(\frac{i-1}{n}\right) - K \circ \Phi_{i/n}^G(\mathbb{F}_n) \right),$$

then Theorem 1 implies the result, since K is increasing. ∎

Therefore, we reject $\mathcal{H}_0$ when $T_c^K(F_n) \geq c_{\alpha,n}$, where $c_{\alpha,n}$ is the solution of $P(T_c^K(\mathbb{G}_n) \geq c_{\alpha,n}) = \alpha$. Proposition 1 ensures that the *size* of the test is α, that is, under $\mathcal{H}_0$, $P(T_c^K(\mathbb{F}_n) \geq c_{\alpha,n}) \leq \alpha$. Moreover, under the alternative "$G^{-1} \circ F$ is concave," reversing Proposition 1, we can similarly obtain $P(T_c^K(\mathbb{F}_n) \geq c_{\alpha,n}) \geq P(T_c^K(\mathbb{G}_n) \geq c_{\alpha,n}) = \alpha$, that is, the test is *unbiased* under concave alternatives. The p-value of the test is $p = P(T_c^K(\mathbb{G}_n) \geq T_c^K(F_n))$.

As well known, the determination of the least favorable distribution of T_c^K yields conservative tests. In order to improve the power of the test against nonconvex 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 an isotonic estimate of F, instead of sampling from G (Groeneboom & Jongbloed, 2012). In particular, denote with $\hat{\mathbb{F}}_n^G$ the (bootstrap) empirical distribution obtained by sampling from F_n^G. The following result holds.

Proposition 2. *Under* $\mathcal{H}_0 : G \geq_c F$, $T_c^K(\mathbb{G}_n) \geq_{st} T_c^K(\hat{\mathbb{F}}_n^G)$.

Proof. Let F_n^G be the estimate of F which corresponds to the empirical distribution F_n^*. F_n^G is a realization of $\hat{\mathbb{F}}_n^G$. Lemma 2 yields $\Phi_{i/n}^G(G_n) \leq \Phi_{i/n}^G(F_n^*)$, 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

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

Hence, as shown in the proof of Proposition 1, we obtain that $T_c^K(\mathbb{G}_n) \geq_{st} T_c^K(\hat{\mathbb{F}}_n^G)$. ∎

Since quantiles are isotonic with $\geq_{st}$, Proposition 2 ensures that, with this method, we obtain a smaller critical value and, correspondingly, a higher value of the power function under $\mathcal{H}_1$. On the other hand, a test obtained by this method does not satisfy anymore the properties implied by Proposition 1 (size and unbiasedness).

4.2 | Testing star-shapedness

Now, let us consider tests for $\mathcal{H}_0 : G \geq_* F$ versus $\mathcal{H}_1 : G \not\geq_* F$. Note that $G \geq_* F$ requires that $F(0) = 0$. Define the *greatest star-shaped minorant* of a function ϕ, as the largest star-shaped function ϕ_* that does not exceed ϕ (see section 2 of Wang, 1987 for a formal definition). The same arguments used earlier may determine the least favorable distribution, under $\mathcal{H}_0$, of the test statistic

$$T_*^K(\mathbb{F}_n) = \sup_{i=2,\ldots,n-1} \left(K \circ G^{-1}\left(\frac{i-1}{n}\right) - K \circ (G^{-1} \circ \mathbb{F}_n)_*(X_{(i)}) \right).$$

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 2.

Lemma 3. *Assume that* $G \geq_* F$. *Then* $\Psi_{i/n}^G(G_n) \leq \Psi_{i/n}^G(F_n^*)$, *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-star-shaped, and recall that $x_i^* = u(y_i)$ are the observations which determine F_n^*. Note that $G^{-1} \circ G_n(y_i) = G^{-1}(i/n)$. We now describe explicitly $(G^{-1} \circ G_n)_*$ (see Wang, 1988). Let

$$a_i = \frac{1}{y_i} G^{-1}\left(\frac{i-1}{n}\right), \quad i = 1,\ldots,n,$$

be the slope of the line connecting the origin to the point $(y_i, G^{-1}((i-1)/n))$. The analytical expression of the greatest star-shaped 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 $\alpha_1 = 0$, and, moreover, star-shapedness 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: $G^{-1} \circ F_n^*(x_i^*) = G^{-1} \circ G_n(y_i) = G^{-1}((i-1)/n)$. 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

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

for $k = i,\ldots,n$, which follows form the anti-star-shapedness of the function u, that is, $u(x)/x$ is decreasing. ∎

Using Lemma 3 and similar arguments as those used in the previous subsection, we can ensure that the least favorable distribution of T_*^K, under $\mathcal{H}_0$, is determined by simulating from G. Accordingly, the test determined by $P(T_*^K(\mathbb{G}_n) \geq c_{\alpha,n}) = \alpha$ has size α and it is unbiased against anti-star-shaped alternatives. Moreover, the bootstrap estimator which arises from the isotonic estimate $G \circ (G^{-1} \circ F_n)_*$, denoted as $\tilde{\mathbb{F}}_n^G$, satisfies an equivalent property to that established by Proposition 2. We summarize these results as follows (the proof is omitted).

Proposition 3. *Under* $\mathcal{H}_0 : G \geq_* F$,

1. $T_*^K(\mathbb{G}_n) \geq_{st} T_*^K(\mathbb{F}_n)$;
2. $T_*^K(\mathbb{G}_n) \geq_{st} T_*^K(\tilde{\mathbb{F}}_n^G)$.

Differently from the tests discussed in the Section 4.1, which have been studied extensively in the literature, tests of the null hypothesis $G \geq_* F$ are relatively unexplored, so it is worth discussing them further. In fact, our results can be useful for testing goodness-of-fit to the IHRA family, which is particularly relevant in reliability and survival analysis, or some generalizations of this property, as described by Barlow et al. (1971).

4.2.1 | An example: testing the IHRA property

A CDF F is said to be IHRA if the *hazard function* $H = -\ln(1 - F)$ is star-shaped, or equivalently, if the *hazard rate average* (HRA) function $H(x)/x$ is increasing (Marshall & Olkin, 2007). Tests of the IHRA property have already been considered in the literature, however, these are based on the null hypothesis that F is exponential, versus IHRA alternatives (Ahmad, 1994; Deshpande, 1983; Klefsjö, 1983; Kochar, 1985; Link, 1989; Wells & Tiwari, 1991), which is clearly different from our approach.

Plugging $G = K = \mathcal{E}$ into the expression of T_*^K, we obtain a Kolmogorov–Smirnov-type test statistic, which can be used to test goodness-of-fit to the IHRA family. More generally, we may consider the following restricted hypotheses $\mathcal{H}_0^\nu : H|_{S_\nu}$ is star-shaped versus $\mathcal{H}_1^\nu : H|_{S_\nu}$ is not star-shaped, where $S_\nu = \{x : x \leq F^{-1}(1 - \nu)\}$, $\nu \in [0, 1]$ (clearly $\mathcal{H}_0$ coincides with $\mathcal{H}_0^0$). Let us define n_ν by $X_{(n_\nu)} = \mathbb{F}_n^{-1}(1 - \nu)$, and $S_{n,\nu} = \{x \leq X_{(n_\nu)}\}$. Accordingly, we have the following restricted test statistic:

$$T_*^\nu(\mathbb{F}_n) = \sup_{i=2,\dots,n_\nu-1} \left(\frac{i-1}{n} - \mathcal{E} \circ \left(\mathbb{H}_n|_{S_{n,\nu}} \right)_* (X_{(i)}) \right),$$

where $\mathbb{H}_n = -\ln(1 - \mathbb{F}_n)$ is the empirical hazard function. It is easy to see that this restricted test still satisfies the properties established by Proposition 3, for every $\nu \in [0, 1]$. Therefore, the distribution of T_*^ν is obtained by simulating from $G = \mathcal{E}$. We reject $\mathcal{H}_0$ when $T_*^\nu(\mathbb{F}_n) \geq c_{\alpha,n,\nu}$, where $c_{\alpha,n,\nu}$ is the solution of $P(T_*^\nu(\mathbb{G}_n) \geq c_{\alpha,n,\nu}) = \alpha$, giving the size of the test. The test is unbiased against alternatives that exhibit a decreasing hazard rate average (DHRA), corresponding to an anti-star-shaped behavior of H, in S_ν. Moreover, for every $\nu \in (0, 1]$, it is possible to establish the following consistency property.

Theorem 2. *If* $\mathcal{H}_0^\nu$ *is false,* $\lim_{n\to\infty} P(T_*^\nu(\mathbb{F}_n) \geq c_{\alpha,n,\nu}) = 1$, *for every* $\nu > 0$.

Proof. Note that $T_*^\nu(\mathbb{F}_n) = \sup_{S_{n,\nu}} \left| \mathbb{F}_n(x) - \mathcal{E} \circ (\mathbb{H}_n|_{S_{n,\nu}})_*(x) \right| - 1/n$, so we may equivalently consider the "sup" on the right-hand side, which is easier to manipulate, instead of T_*^ν. By the Glivenko–Cantelli Theorem, $\mathbb{F}_n$ converges a.s. and uniformly to F, whereas $\mathcal{E}^{-1}(p) = -\ln(1 - p)$ is uniformly continuous in $[0, 1 - \nu]$ (for $\nu > 0$), then $\mathbb{H}_n = \mathcal{E}^{-1} \circ \mathbb{F}_n$ converges a.s. and uniformly to $H = \mathcal{E}^{-1} \circ F$ in S_ν. If $\mathcal{H}_0^\nu$ is true, we can apply lemma 3.1 of Wang (1988), which yields

$$\sup_{S_{n,\nu}} |\mathbb{H}_n(x) - H(x)| \geq \sup_{S_{n,\nu}} \left| (\mathbb{H}_n|_{S_{n,\nu}})_*(x) - H(x) \right|,$$

implying strong uniform consistency of $(\mathbb{H}_n|_{S_{n,\nu}})_*$ in S_ν (clearly, if $\nu > 0$, $\mathbb{F}_n^{-1}(1 - \nu)$ converges a.s. to $F^{-1}(1 - \nu)$). Since $\mathcal{E}$ is absolutely continuous, then $\mathbb{F}_{n,\nu}^{\mathcal{E}} := \mathcal{E} \circ (\mathbb{H}_n|_{S_{n,\nu}})_*$ converges a.s. and uniformly to F in S_ν, for every $\nu > 0$. Now, it can be easily seen that, under $\mathcal{H}_0^\nu$, $c_{\alpha,n,\nu} \to 0$ a.s. for $n \to \infty$.

Define $F_\nu^{\mathcal{E}} = \mathcal{E} \circ (\mathcal{E}^{-1} \circ F|_{S_\nu})_*$. Suppose that $\mathcal{H}_0^\nu$ is false, that is, $\mathcal{E} \circ F$ is not star-shaped in S_ν, and let $d = \sup_{S_\nu} |F - F_\nu^{\mathcal{E}}| = \sup_{S_\nu}(F - F_\nu^{\mathcal{E}})$. Then, F_n converges uniformly to F, whereas $\mathbb{F}_{n,\nu}^{\mathcal{E}}$ converges uniformly to $F_\nu^{\mathcal{E}}$ on S_ν, with probability 1. Moreover, by assumption, $d > 0$. Therefore, given some $\epsilon \in (0, d)$, there exists some n_0 such that, for $n > n_0$, $\sup_{S_{n,\nu}} |F_n - F| < \epsilon/2$ and $\sup_{S_{n,\nu}} |\mathbb{F}_{n,\nu}^{\mathcal{E}} - F_\nu^{\mathcal{E}}| < \epsilon/2$, with probability 1. Then, for $n > n_0$,

$$\mathbb{F}_n(x) - \mathbb{F}_{n,\nu}^{\mathcal{E}}(x) > F(x) - \frac{\epsilon}{2} - \left(F_\nu^{\mathcal{E}}(x) + \frac{\epsilon}{2}\right) = F(x) - F_\nu^{\mathcal{E}}(x) - \epsilon,$$

almost surely, for every $x \in S_{n,\nu}$, which implies

$$\sup_{S_{n,\nu}} \left|\mathbb{F}_n(x) - \mathbb{F}_{n,\nu}^{\mathcal{E}}(x)\right| > \sup_{S_{n,\nu}} \left|F(x) - \mathbb{F}_{n,\nu}^{\mathcal{E}}(x) - \epsilon\right| = \sup_{S_{n,\nu}} \left(F(x) - \mathbb{F}_{n,\nu}^{\mathcal{E}}(x)\right) - \epsilon = d - \epsilon > 0.$$

Therefore, since ϵ can be arbitrarily small, $P(\sup_{S_{n,\nu}} |\mathbb{F}_n(x) - \mathbb{F}_{n,\nu}^{\mathcal{E}}(x)| \geq d) \to 1$, for $n \to \infty$. But since $c_{\alpha,n,\nu} \to 0$, then $P(T_*^\nu(\mathbb{F}_n) \geq c_{\alpha,n,\nu}) \to 1$. ∎

The value of ν should be chosen as small as possible, for example, $\nu = 0.01, 0.05$. One might even choose ν as a function of n, for example $\nu = n^{-1/2}$. Yet, smaller ν's yield slower convergence, as we will need more observations in order to get $c_{\alpha,n,\nu}$ close enough to 0. In the limit case $\nu = 0$, we cannot ensure uniform a.s. consistency of $\mathbb{F}_{n,0}^{\mathcal{E}}$ in S_0, accordingly, the consistency of the test T_*^0 cannot be established.

4.2.2 | Simulations

The numerical performance of the test T_*^ν can be depicted by a simulation study. We focus on some critical alternatives, obtained using popular parametric families. By scale invariance, the scale parameters are always set to 1. We consider DHRA, bathtub-shaped (decreasing-increasing) HRA, and bell-shaped (increasing-decreasing) HRA models. The Weibull distribution with CDF $F(x) = 1 - e^{-x^a}$, $a, x > 0$, is DHRA for $a < 1$ and IHRA for $a \geq 1$; the beta distribution, with CDF $F(x) = \int_0^x t^{a-1}(1 - t)^{b-1} \, dt/B(a, b)$, $a, b > 0$, $x \in [0, 1]$, is bathtub-shaped HRA for $a < 1 < b$ and IHRA for $a, b \geq 1$; the Burr distribution with CDF $F(x) = 1 - (1 + x^b)^{-a}$, $x, a, b > 0$, is DHRA for $b \leq 1$ and bell-shaped HRA for $b > 1$.

The results are reported in Figure 2, which contains the plots of the rejection probabilities at level $\alpha = 0.1$, estimated using 2000 simulation runs. To cover a large number of alternatives, for each distribution considered we fixed a grid of parameters within some intervals of interest. We considered sample sizes $n = 10, 50, 100, 200, 300$, setting $\nu = 0.05$, so the restriction has an effect just for $n \geq 50$. The figures show that the test has a remarkable performance in terms of empirical power, coherently with the behavior of the HRA. In fact, some alternatives considered are particularly difficult to detect, as the HRA may exhibit a very slight decrease (Weibull with $a \in [0.8, 1)$) or a subtle nonmonotonic behavior (beta with $a \geq 0.7$; Burr with $b \geq 1.5$). In the first case, the power is always larger than α, independently of the sample size, confirming the unbiasedness property implied by Proposition 3; in the latter case, the power may be smaller than α, for small
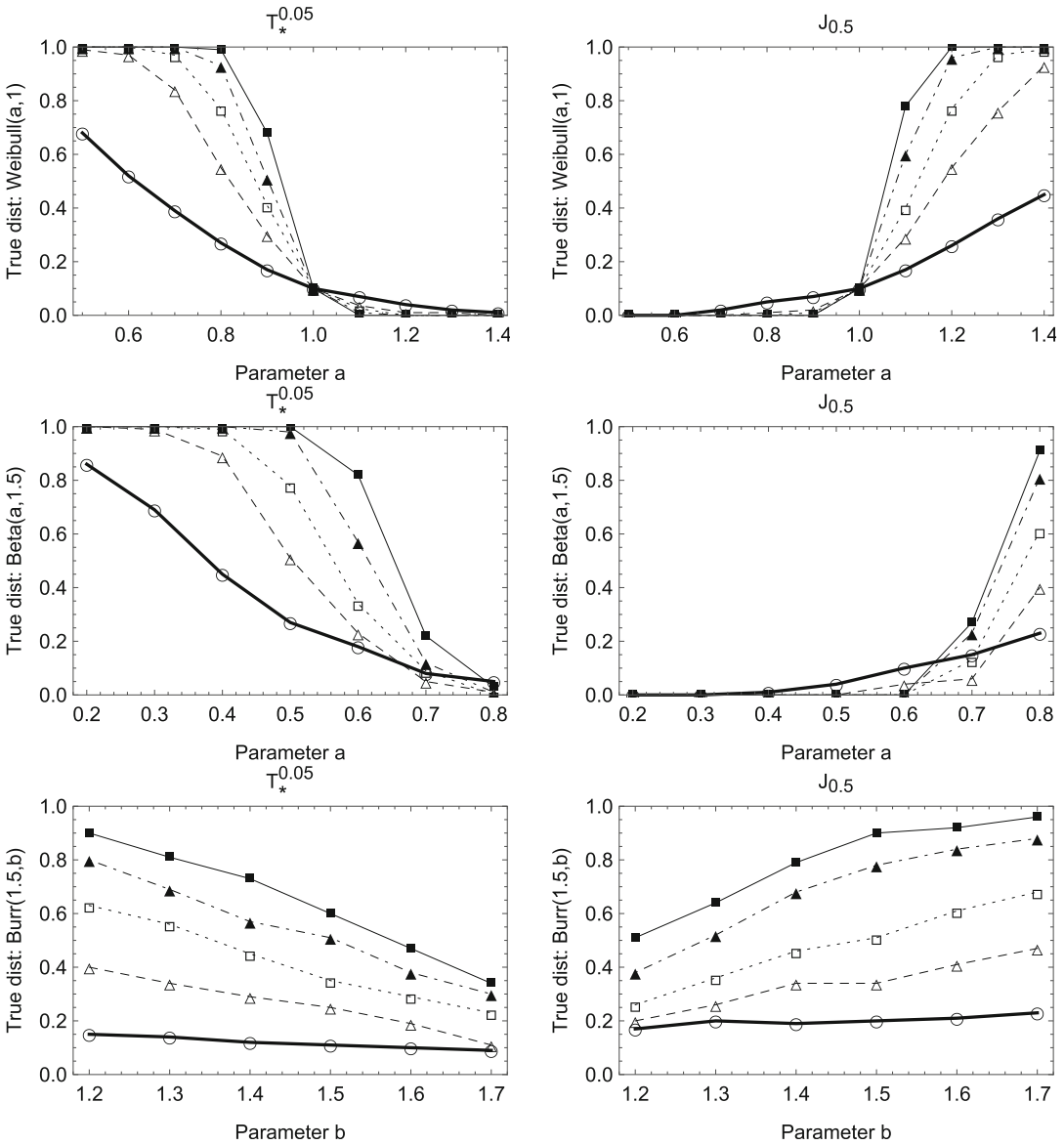
FIGURE 2 Rejection rates of $T_*^{0.05}$ and $J_{0.5}$ for increasing hazard rate average (Weibull $a \geq 1$), decreasing hazard rate average (Weibull $a < 1$), bathtub-shaped hazard rate average (HRA) (beta) and bell-shaped HRA (Burr) distributions. Sample size: $n = 10$ (○), $n = 50$ (△), $n = 100$ (□), $n = 200$ (▲), $n = 300$ (■)

values of n (beta with $a \geq 0.7$), confirming that unbiasedness of tests of this type holds just against DHRA alternatives, but when the HRA has a nonmonotonic behavior, this is not guaranteed. Under the null hypothesis, the size is always bounded by α, even when the HRA is close to the boundary of DHRA (Weibull with $a > 1$). Clearly, as the test is consistent, the empirical power increases with sample size, when $\mathcal{H}_0^{0.05}$ is false, even under the most critical alternatives (this should be evaluated just for $n \geq 50$, as remarked earlier; the beta with $a \geq 0.7$ is the most critical case and the power may be increased by increasing the sample size, or by choosing a larger value of ν).

The test is compared with the one proposed by Deshpande (1983) and later studied by Bandyopadhyay and Basu (1989), aimed at testing the null hypothesis $\widetilde{\mathcal{H}}_0 : F = \mathcal{E}$ versus $\widetilde{\mathcal{H}}_1$: "$F$ is IHRA and not exponential", as in Klefsjö (1983), Kochar (1985), Link (1989), Wells and Tiwari (1991), Ahmad (1994). This approach is quite different from ours, in fact, such tests are supposed to reject the IHRA hypothesis more easily, when it is false, compared to T_*^ν. However, our simulation results reveal some possible issues of these kinds of tests. Actually, Deshpande (1983) introduced a family of rank-based test statistics, denoted as J_p, where the parameter p ranges in $(0, 1)$. As proved by Bandyopadhyay and Basu (1989), the choice for p does not have consequences on the properties of the test, therefore, in the plots we show the results of $J_{0.5}$, which is the test recommended by Deshpande (1983). This test is unbiased and consistent, for every p, under IHRA alternatives. However, deviations from exponentiality should not necessarily imply that $\widetilde{\mathcal{H}}_1$ is true, as tests of this type are not intended to deal with bathtub or bell-shaped HRA models; indeed J_p may be misleading when F has a nonmonotone HRA, as it can be seen in the case of the beta distribution ($a \geq 0.7$) or Burr distribution. In such cases, by increasing the sample size we also increase the power, that is, the probability of rejecting $\widetilde{\mathcal{H}}_0$, suggesting that the (false) strict IHRA alternative may be true (we obtained similar results for several values of p). In other words, in some critical bathtub and bell-shaped HRA situations, J_p might lead to the wrong decision. So, given that nonmonotone HRA distributions may be quite common, we recommend using T_*^ν. Looking at the plots, it is important to remember that T_*^ν and J_p cannot be compared in terms of power, since they deal with reversed hypotheses, hence one may only compare their overall performance. This example also shows why it is important to study and understand the behavior of the test statistics using ordering constraints.

ACKNOWLEDGMENTS

ORCID

Tommaso Lando https://orcid.org/0000-0003-4288-0264
Idir Arab https://orcid.org/0000-0002-9408-684X
Paulo Eduardo Oliveira https://orcid.org/0000-0001-7217-5705

REFERENCES

Ahmad, I. A. (1994). A class of statistics useful in testing increasing failure rate average and new better than used life distributions. *Journal of Statistical Planning and Inference*, *41*(2), 141–149.

Arendarczyk, M., Kozubowski, T. J., & Panorska, A. K. (2021). The Greenwood statistic, stochastic dominance, clustering and heavy tails. *Scandinavian Journal of Statistics*, *49*(1), 331–352.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*(3), 244–263.

Bandyopadhyay, D., & Basu, A. P. (1989). A note on tests for exponentiality by Deshpande. *Biometrika*, *76*(2), 403–405.

Barlow, R., Bartholomew, D., Bremner, J., & Brunk, H. (1971). *Statistical inference under order restrictions*. Wiley.

Barlow, R. E., & Proschan, F. (1969). A note on tests for monotone failure rate based on incomplete data. *Annals of Mathematical Statistics*, *40*(2), 595–600.

Bickel, P. (1969). Tests for monotone failure rate II. *Annals of Mathematical Statistics*, *40*(4), 1250–1260.

Bickel, P. J., & Doksum, K. A. (1969). Tests for monotone failure rate based on normalized spacings. *Annals of Mathematical Statistics*, *40*(4), 1216–1235.

Deshpande, J. V. (1983). A class of tests for exponentiality against increasing failure rate average alternatives. *Biometrika*, *70*(2), 514–518.

Donaldson, D., & Weymark, J. A. (1980). A single-parameter generalization of the Gini indices of inequality. *Journal of Economic Theory*, *22*(1), 67–86.

Durot, C. (2007). On the $\mathbb{L}_p$-error of monotonicity constrained estimators. *Annals of Statistics*, *35*(3), 1080–1104.

Eberl, A., & Klar, B. (2021). Expectile-based measures of skewness. *Scandinavian Journal of Statistics*, *49*(1), 373–399.

Gijbels, I., & Heckman, N. (2004). Nonparametric testing for a monotone hazard function via normalized spacings. *Journal of Nonparametric Statistics*, *16*(3-4), 463–477.

Gini, C. (1912). *Variabilità e mutabilità*. Reprinted in Memorie di metodologica statistica, Pizetti.

Greenwood, M. (1946). The statistical study of infectious diseases. *Journal of the Royal Statistical Society*, *109*(2), 85–110.

Groeneboom, P., & Jongbloed, G. (2012). Isotonic L_2-projection test for local monotonicity of a hazard. *Journal of Statistical Planning and Inference*, *142*(7), 1644–1658.

Hall, P., & van Keilegom, I. (2005). Testing for monotone increasing hazard rate. *Annals of Statistics*, *33*(3), 1109–1137.

Klefsjö, B. (1983). Some tests against aging based on the total time on test transform. *Communications in Statistics - Theory and Methods*, *12*(8), 907–927.

Kochar, S. C. (1985). Testing exponentiality against monotone failure rate average. *Communications in Statistics - Theory and Methods*, *14*(2), 381–392.

Lando, T. (2022). Testing convexity of the generalised hazard function. *Statistical Papers*, *63*, 1271–1289.

Lando, T., Arab, I., & Oliveira, P. E. (2021). Second-order stochastic comparisons of order statistics. *Statistics*, *55*(3), 561–579.

Lando, T., & Bertoli-Barsotti, L. (2016). Weak orderings for intersecting Lorenz curves. *Metron*, *74*(2), 177–192.

Lando, T., & Bertoli-Barsotti, L. (2020). Distorted stochastic dominance: A generalized family of stochastic orders. *Journal of Mathematical Economics*, *90*, 132–139.

Link, W. A. (1989). Testing for exponentiality against monotone failure rate average alternatives. *Communications in Statistics - Theory and Methods*, *18*(8), 3009–3017.

Marshall, A. W., & Olkin, I. (2007). *Life distributions* (Vol. *13*). Springer.

Marshall, A. W., Olkin, I., & Arnold, B. C. (2011). *Inequalities: Theory of majorization and its applications*. Elsevier.

Mehran, F. (1976). Linear measures of income inequality. *Econometrica*, *44*, 805–809.

Muliere, P., & Scarsini, M. (1989). A note on stochastic dominance and inequality measures. *Journal of Economic Theory*, *49*(2), 314–323.

Oja, H. (1981). On location, scale, skewness and kurtosis of univariate distributions. *Scandinavian Journal of Statistics*, *8*, 154–168.

Proschan, F., & Pyke, R. (1967). *Tests for monotone failure rate*. Proceedings of the 5th Berkley Symposium (Vol. 3, pp. 293–313).

Serfling, R. (1984). Generalized L-, M-, and R-statistics. *Annals of Statistics*, *12*, 76–86.

Shaked, M., & Shanthikumar, J. G. (2007). *Stochastic orders Springer Series in Statistics*. Springer.

Shorrocks, A., & Slottje, D. (2002). Approximating unanimity orderings: An application to Lorenz dominance. *Journal of Econometrics*, *77*(1), 91–117.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica*, *48*, 613–625.

Tenga, R., & Santner, T. J. (1984). Testing goodness of fit to the increasing failure rate family. *Naval Research Logistics*, *31*(4), 617–630.

van Zwet, W. R. (1964). *Convex transformations of random variables*. MC Tracts.

Wang, J.-L. (1987). Estimating IFRA and NBU survival curves based on censored data. *Scandinavian Journal of Statistics*, *14*, 199–210.

Wang, J.-L. (1988). Optimal estimations of starshaped distribution functions. *Statistics & Decisions*, *6*(1-2), 21–32.

Wang, S. S., & Young, V. R. (1998). Ordering risks: Expected utility theory versus Yaari's dual theory of risk. *Insurance: Mathematics & Economics*, *22*(2), 145–161.

Wells, M. T., & Tiwari, R. C. (1991). A class of tests for testing an increasing failure-rate-average distribution with randomly right-censored data. *IEEE Transactions on Reliability*, *40*(2), 152–156.

Yitzhaki, S., & Schechtman, E. (2013). *The Gini methodology: A primer on a statistical methodology*. Springer.