New results and tests for stochastic dominance between linear combinations

Tommaso Lando*[1] and Paulo Eduardo Oliveira†‡[2]

[1]Department of Economics, University of Bergamo, via dei Caniana 2, 24127, Bergamo, Italy
[2]CMUC, Department of Mathematics, University of Coimbra, Portugal

Abstract

Convex combinations of i.i.d. random variables without a finite mean can behave in a strikingly different way from the finite-mean case: as the weight vector becomes more balanced, the resulting combination may become stochastically larger, rather than less dispersed. Existing results establish stochastic dominance between pairs of linear combinations—or between a convex combination and the underlying variable—under shape restrictions on the distribution and structural assumptions on the weights. We expand the class for which the general result can be derived. Nonetheless, two practical limitations remain: (i) the sufficient conditions vary across results, and (ii) being non-necessary, they exclude many relevant configurations. Moreover, under a statistical perspective, where the true distribution of the data is assumed to be unknown, these conditions cannot be checked. Motivated by this gap, we develop nonparametric procedures to test whether two linear combinations are stochastically ordered. We propose two complementary approaches: a least-favorable calibration and a bootstrap-based method. We show that both tests control size asymptotically under the null of stochastic dominance and are consistent against alternatives of non-dominance. Monte Carlo experiments illustrate the finite-sample performance of the proposed procedures across a range of models and weight configurations.

Key words and phrases: Bootstrap, Empirical process, Infinite mean, Majorization, Nonparametric test, Sample mean, Stochastic order

AMS 2010 subject classification: 60E15, 62E20

1 Introduction

It is a generally accepted fact that the sample mean shrinks towards the theoretical mean when the sample size increases. When the sample size grows to infinity, the Strong Law of Large Numbers establishes the almost sure convergence of the sample mean to the mathematical expectation of the underlying random variable, assuming, of course, that the latter exists and is finite. More generally, for a fixed sample size, a well-known result in ordering theory (Shaked and Shantikumar, 2007, Theorem 3.A.35) shows that convex combinations of an independent and identically distributed (i.i.d.) sample become less

*Email: tommaso.lando@unibg.it, ORCID: 0000-0003-4288-0264
†Email: paulo@mat.uc.pt, ORCID: 0000-0001-7217-5705
‡P.E.O. acknowledges financial support by the Centre for Mathematics of the University of Coimbra (CMUC, https://doi.org/10.54499/UID/00324/2025) under the Portuguese Foundation for Science and Technology (FCT), Grants UID/00324/2025 and UID/PRR/00324/2025.

dispersed, in terms of the *convex order*, as the weight vector becomes more balanced, in terms of *majorization*. This implies, for instance, that the sample mean of size k is less dispersed than the sample mean of size $k-1$ (see Example 3.A.29 and, more generally, Corollary 3.A.31 in Shaked and Shantikumar (2007)). Again, the existence and finiteness of the mathematical expectation of the underlying random variable is crucial for these characterisations.

However, when the mean is not finite, the behaviour is less predictable, and often unexpected. In fact, Chen et al. (2024) proved that, for a specific family of distributions obtained by convex transformations of Pareto random variables, named *super-Pareto*, any convex combination of the sample stochastically dominates the generating random variable. This result has many important applications, including the surprising fact that the sample mean does not always shrink towards a constant when the sample size increases, but, on the contrary, it can get stochastically larger. More recently, a number of contributions enlarged the family of distributions for which such a result holds (Chen and Shneer, 2025; Arab et al., 2025; Müller, 2025; Vincent, 2025). In particular, Müller (2025) considered the class of variables that are convex transformations of the Cauchy distribution, and Arab et al. (2025) proved that the result holds for the class of random variables whose inverted distribution is subadditive, called InvSub. These two latter classes seem to be largest known families of distributions for which convex combinations of a sample stochastically dominate the underlying variable, and there is no inclusion between them.

Taken together, these results naturally raise the question of how stochastic dominance behaves when comparing different linear combinations of a sample, possibly involving distinct weight vectors and sample sizes. The comparison between different linear combinations was recently addressed by Chen et al. (2025), who proved that the stochastic dominance holds when the weights vectors satisfy a majorization relation, provided that the inverted distribution of the underlying random variable is concave. This class is strictly contained in the InvSub class of Arab et al. (2025). The result of Chen et al. (2025) implies many important results, including–within their class of distributions– stochastic monotonicity of the sample mean in the sample size.

Overall, the recent literature proposed different classes, not always included among them, yielding stochastic dominance properties at different levels of generality. In fact, all the aforementioned results only provide sufficient conditions for the dominance property. It is not difficult to find examples of distributions that are not included in any of these classes but for which, at least for specific choices of the combination coefficients, the dominance result still holds. Moreover, in real world applications, we are not able to verify whether a distribution is in some of the classes above, as we just observe realizations (data) from it. Therefore, from a statistical perspective, it is crucially important to be able to infer from the data if the underlying distribution fulfills some dominance property of this type. This could be done with two different approaches, either testing whether the underlying distribution belongs to one of the aforementioned classes, or testing the property directly. With regard to the first approach, testing whether a distribution belongs to the super-Pareto, or to the super-Cauchy class, could be relatively straightforward using existing methods, such as the one recently proposed by Lando and Benjrada (2025). Testing the condition proposed by Chen et al. (2025), which is based on concavity, would be also quite simple. On the other hand, developing tests for the subadditivity-based conditions considered in Chen and Shneer (2025) and Arab et al. (2025) appears to be an open and potentially challenging problem. Overall, the first approach offers no unique prescription: the choice of which condition to test is not obvious and may strongly affect the resulting inference. Moreover, testing these sufficient conditions would still leave out many interesting cases, as discussed earlier. Therefore, the second approach seems to

be more direct and effective. One can directly develop tests that detect the dominance between a linear combination of a sample and the underlying variable, or, more generally, between two linear combinations of interest.

This paper has a twofold objective: i) deriving new distributional conditions for comparing linear combinations of random variables, and; ii) in absence of any distributional assumption, proposing a statistical test to detect dominance between pairs of linear combinations. In Section 3, we review some relevant results and show that the recent one of Chen et al. (2025), which compares general linear combinations, holds under a weaker assumption. On the other hand, counterexamples show that substantial further generalizations may not hold. Since relying on classes of distributions is not always possible, in Section 4 we focus on those properties that can be obtained without any distributional assumptions. Namely, we consider a dominance between a pair of linear combinations as the starting point for obtaining infinitely many other dominance relations, just by relying on the mathematical properties of stochastic dominance. Our results appear particularly simple when the linear combinations reduce to the sample mean. In Section 5 we study inference for the distribution of a linear combination. We establish the uniform almost sure convergence of the plugin estimator and the weak convergence of the corresponding empirical process. As simpler special cases, we consider the case when the underlying distribution is a Cauchy, as any convex combination of i.i.d. standard Cauchy random variables is again standard Cauchy; and the case when the combination of interest is a sample mean. In Section 6, we propose two versions of a test for the general null hypothesis of dominance between a pair of linear combinations, versus the alternative of non-dominance, with two alternative methods to obtain critical values. The first approach is based on the already mentioned closure property of the Cauchy distribution, which can be seen as the least favourable case under the null. The second approach relies on a bootstrap procedure. We show that, in both cases, the test has asymptotically bounded size under the null hypothesis and it is consistent under the alternative. In Section 7, we examine and compare the two testing approaches by simulating under different scenarios. All proofs are presented in the Appendix.

2 Notation and preliminaries

We now introduce general concepts and terminology to be used throughout. Let X be a random variable with cumulative distribution function (CDF) F. The convolution product is denoted as $*$. As usual, the sample mean of an i.i.d. sample $X_1, \ldots, X_s$ is denoted as $\overline{X}_s$. Recall that, given a nondecreasing and nonnegative function g such that $g(0) = 0$, we say that g is anti-starshaped (at the origin) if, for every $x \geq 0$ and every $\lambda \in [0, 1]$, $g(\lambda x) \geq \lambda\, g(x)$, and that g is subadditive if, for every $x, y \geq 0$, $g(x + y) \leq g(x) + g(y)$. It is well known that concavity implies anti-starshapedness, which, in turn, implies subadditivity. Note that the terms "increasing" and "decreasing" are taken as "non-decreasing" and "non-increasing". Still, with respect to notations, we denote with $\to_p$, $\to_{a.s.}$, and $\to_d$ convergence in probability, almost surely, and in distribution of sequences of random variables, respectively. When referring to sequences of stochastic processes, weak convergence is denoted by $\rightsquigarrow$, as in Van der Vaart and Wellner (1996). Finally, we denote by $\ell^\infty(\mathbb{R})$ the space of all bounded real-valued functions on $\mathbb{R}$, equipped with the supremum norm, denoted as usual by $\|\cdot\|_\infty$.

Our main focus is on stochastic dominance. In particular, we study the conditions under which linear combinations satisfy this relation, as well as how to test whether stochastic dominance holds. Since this concept is central to our analysis, we recall its definition.

Definition 2.1 (Shaked and Shantikumar (2007), p. 3)**.** *Given two random variables X and Y, we say that Y stochastically dominates X, denoted as $Y \geq_{st} X$, if $P(X \leq x) \geq P(Y \leq x)$ for every $x \in \mathbb{R}$.*

Now, to be more specific, given $X_1, X_2, \ldots$ independent and identically distributed with X, and vectors of nonnegative weights $\overline{\theta} = (\theta_1, \ldots, \theta_s)$ and $\overline{\eta} = (\eta_1, \ldots, \eta_t)$, we will be interested in conditions under which

$$\theta_1 X_1 + \cdots + \theta_s X_s \ \geq_{st} \ \eta_1 X_1 + \cdots + \eta_t X_t. \tag{1}$$

As shown later, this result can be obtained based on suitable assumptions on the weight vectors and on the underlying distribution. Intuitively, a linear combination in which the weight vector is more concentrated in a single component is expected to be closer, in distribution, to the underlying variable, X. On the other hand, if the weight vector is more balanced, the linear combination may behave differently, for example, shrinking towards a constant, when the mean is finite, or becoming stochastically larger, when the mean is not finite, as discussed above. The translation of this intuitive behaviour of a vector's components into a mathematical formulation is possible through the notion of majorization (Marshall et al., 2011).

As a particularly relevant special case of (1), the dominance

$$\theta_1 X_1 + \cdots + \theta_s X_s \geq_{st} X, \tag{2}$$

for every integer $n \geq 2$ and every choice of nonnegative real numbers $\theta_1, \ldots, \theta_n$ that sum up to 1, has been studied in Chen et al. (2024), Müller (2025), Chen and Shneer (2025) and Arab et al. (2025). These authors identify different classes of distributions for which the dominance (2) holds. Among these classes, the largest ones are those proposed by Müller (2025) and Arab et al. (2025). There is no inclusion between these two classes, in fact, the former one includes distributions on the entire real line, whereas the latter allows for more flexibility in terms of shape, such as discontinuities. In particular, the class of Arab et al. (2025) is characterised by the subadditivity of the inverted distribution, i.e., the CDF $1 - F(\frac{1}{x})$.

Recently, Chen and Shneer (2025) have shown that the general dominance result (1) holds under some intuitively reasonable conditions on the weight vectors, and under a stronger distributional assumption. Namely, instead of subadditivity, they assume concavity of $1 - F(\frac{1}{x})$. We shall show, by example, that the stochastic dominance (1) does not follow from either the subadditivity or the anti-starshapedness of $1 - F(\frac{1}{x})$, and prove a relaxation of the concavity assumption used in Chen et al. (2025).

3 Dominance between linear combinations

The stochastic dominance (1) needs some control on the weight coefficients, as mentioned above. We recall the relevant majorization notion between vectors in $\mathbb{R}^s$.

Definition 3.1. *Let $\overline{\theta} = (\theta_1, \ldots, \theta_s)$, $\overline{\eta} = (\eta_1, \ldots, \eta_s) \in \mathbb{R}^s$.*

1. *(Marshall et al. (2011), Definition 1.A.1) We say that $\overline{\theta}$ is dominated by $\overline{\eta}$ in majoriazation order, denoted $\overline{\theta} \prec \overline{\eta}$ if $\sum_{i=1}^{k} \theta_{i:s} \geq \sum_{i=1}^{k} \eta_{i:s}$, $k = 1, \ldots, s-1$, and $\sum_{i=1}^{s} \theta_i = \sum_{i=1}^{s} \eta_i$, where $\theta_{i:s}$ and $\eta_{i:s}$ represent the coordinates of $\overline{\theta}$ and $\overline{\eta}$, respectively, ordered increasingly.*

2. *(Marshall et al. (2011), Section 1.A.8) We say that $\overline{\theta}$ is a T-transform of $\overline{\eta}$ if there exist $\lambda \in [0,1]$ and distinct i and j in $\{1, \ldots, n\}$ such that $\theta_i = \lambda\eta_i + (1-\lambda)\eta_j$, $\theta_j = (1-\lambda)\eta_i + \lambda\eta_j$, and $\theta_k = \eta_k$, for $k \neq i, j$.*

The relation $\overline{\theta} \prec \overline{\eta}$ generally means that the components of $\overline{\theta}$ are more spread out compared to those of $\overline{\eta}$.

Remark 3.2. *Note that majorization assumes that $\overline{\theta}$ and $\overline{\eta}$ have the same number of components. This restriction is easily overcome. Indeed, every pair of vectors can be compared by filling with zeros the missing components of the shorter one. Take $\overline{\theta} \in \mathbb{R}^s$ and $\overline{\eta} \in \mathbb{R}^t$, where $s > t$. Then, define the vector $\widetilde{\eta} = (\eta_1, \ldots, \eta_t, \widetilde{0}_{s-t}) \in \mathbb{R}^s$, where $\widetilde{0}_{s-t}$ is a vector with $s - t$ coordinates all equal to 0. In this case, for example, we can compare the vector $(\frac{1}{2}, \frac{1}{2})$ with the s-dimensional vector $(\frac{1}{s}, \ldots, \frac{1}{s})$, for $s > 2$, as $(\frac{1}{2}, \frac{1}{2}, \widetilde{0}_{s-2}) \prec (\frac{1}{s}, \ldots, \frac{1}{s})$.*

The second notion in Definition 3.1, the T-transforms allow for a simple manipulation of the vectors keeping them ordered through majorization. This is useful for proving the general results.

Lemma 3.3 (Marshall et al. (2011), Section 1.A.3)**.** *Given $\overline{\theta}, \overline{\eta} \in \mathbb{R}^n$, $\overline{\theta} \prec \overline{\eta}$ if and only if $\overline{\theta}$ can be obtained from $\overline{\eta}$ by a finite number of T-transforms.*

The main result in Chen et al. (2025), quoted next, proves sufficient conditions for (1) to hold.

Theorem 3.4 (Chen et al. (2025), Theorem 1)**.** *Let X be a nonnegative random variable with CDF F such that $1-F(\frac{1}{x})$ is concave, and $X_1, \ldots, X_s$ i.i.d. with X. If $\overline{\theta}, \overline{\eta} \in (0, +\infty)^s$ are such that $\overline{\theta} \prec \overline{\eta}$, then $\sum_{i=1}^s \theta_i X_i \geq_{st} \sum_{i=1}^s \eta_i X_i$.*

Note that Theorem 3.4 implies (2) taking $\overline{\eta} = (1, 0, \ldots, 0)$. This same conclusion holds under the weaker assumption of subadditivity of $1 - F(\frac{1}{x})$, by Theorem 3.1 in Arab et al. (2025). Hence, one may naturally wonder whether this latter and weaker assumption suffices to obtain (1). The answer is negative. The following example shows that not even anti-starshapedness is enough.

Example 3.5. *Consider the CDF defined as*

$$F(x) = \begin{cases} 0, & \text{if } x \leq \frac{1}{4}, \\ 0.8 - \frac{0.2}{x}, & \text{if } \frac{1}{4} \leq x < \frac{1}{2}, \\ 0.4, & \text{if } \frac{1}{2} \leq x < 1, \\ 1 - \frac{0.6}{x}, & \text{if } x \geq 1. \end{cases}$$

It is easily seen that the inverted CDF $H(x) = 1-F(\frac{1}{x})$ is anti-starshaped, but not concave. Considering, for example, the weight vectors $(0.2, 0.8) \prec (0.3, 0.7)$, one can verify that the corresponding CDFs cross, so stochastic dominance does not hold.

However, we show that the conclusion of Theorem 3.4 still holds under weaker assumptions. For this purpose, consider the following class of distributions.

Definition 3.6. *We say that a cumulative distribution function F is of class $\mathcal{L}$ if $F(0) = 0$ and, given $0 \leq y \leq x \in \mathbb{R}$, $V(z) = F(x + yz) + F\left(x + \frac{y}{z}\right)$ is decreasing with respect to $z \in (0, 1]$.*

Our general stochastic comparison result may now be stated.

Theorem 3.7. *Let X be a random variable with CDF $F \in \mathcal{L}$, and $X_1, \ldots, X_s$ i.i.d. with X. If $\overline{\theta}, \overline{\eta} \in (0, +\infty)^n$ are such that $\overline{\theta} \prec \overline{\eta}$ then $\sum_{i=1}^s \theta_i X_i \geq_{st} \sum_{i=1}^s \eta_i X_i$.*

According to Remark 3.2, we may also compare linear combinations where the weight vectors have a different number of coordinates. This is particulary interesting as it allows to stochastically compare sample means of different sizes. Specifically, this means that under the conditions of Theorem 3.7, the sample mean is stochastically monotone with respect to the sample size. In particular, among these sample means, the case $s = 2$ is stochastically the smallest; hence it is the most stringent benchmark and a natural candidate for statistical testing.

The following proposition shows that the class considered by Chen et al. (2025), characterised by concavity of the inverted distribution, is included in $\mathcal{L}$. The inclusion is strict, as shown by Example 3.9.

Proposition 3.8. *Let F be a CDF such that $F(0) = 0$ and $H(x) = 1 - F(\frac{1}{x})$ is concave. Then $F \in \mathcal{L}$.*

Example 3.9. *We provide an explicit distribution of class $\mathcal{L}$ such that the corresponding H is not concave, thus showing that our class is wider than the one considered in Chen et al. (2025). Take Z with CDF $F_Z(x) = 1 - \frac{1}{(x+1)^\alpha}$, for $x \geq 0$, that is, a Pareto distribution starting at 0, the transformation*

$$v(x) = \begin{cases} x^a, & \text{if } x < 1, \\ x^b, & \text{if } x \geq 1, \end{cases}$$

where $b > a \geq 1$, and define $X = v(Z)$, so that

$$F(x) = \begin{cases} F_Z(x^{1/a}), & \text{if } x < 1, \\ F_Z(x^{1/b}), & \text{if } x \geq 1, \end{cases} \quad \text{and} \quad H(x) = \begin{cases} 1 - F_Z(x^{-1/a}), & \text{if } x > 1, \\ 1 - F_Z(x^{-1/b}), & \text{if } x \leq 1. \end{cases}$$

It is now a lengthy straightforward verification that one may choose the parameters α, a, $b > 0$ such that $F \in \mathcal{L}$, while H is not concave (take, for example, $a = 2$, $b = 3$ and $\alpha \in (0, 1]$).

We conclude the section with a remark.

Remark 3.10. *The sufficient condition in Theorem 3.7 is formulated in terms of majorization, which is only a partial order. When $s = t = 2$, majorization becomes a total order (for vectors with the same sum). In higher dimensions, incomparability is common—many pairs of vectors cannot be ordered—so the practical scope of the theorem is more limited. This does not mean that stochastic dominance cannot hold when weights are incomparable; rather, our sufficient condition is then silent and does not indicate whether dominance holds, nor in which direction. This provides additional motivation for a direct statistical verification of condition (1), which we pursue in Section 6.*

4 Stochastic dominance without distributional assumptions

We have been interested in finding conditions on the distribution of the underlying variables and the weights such that (1) holds. However, this stochastic dominance can also be seen as a starting point to obtain infinitely many other dominance relations, using somewhat standard arguments in ordering theory. In this subsection, we will focus on the implications of (1), with a special focus on the most common family of linear combinations, namely, sample means. We show that sample means satisfy a hierarchical property, whenever $\overline{X}_s \geq_{st} \overline{X}_t$. Namely, the latter condition implies $\overline{X}_{ks} \geq_{st} \overline{X}_{kt}$ for every positive integer k. This result can be extended using general linear combinations instead of just sample means. Moreover, we search for a result that can be seen as a special case of

Theorem 3.7, but does not rely on distributional assumptions on the underlying random variables. Such a result is possible if we assume $\overline{X}_s \geq_{st} X$ plus a stronger type of ordering between the weight vectors, which can be seen as a special case of majorization, defined as follows.

Definition 4.1. *Let $\overline{\theta} = (\theta_1, \ldots, \theta_k)$ be a vector of nonnegative weights such that $\sum_{i=1}^k \theta_i = 1$ and $h \geq 2$ an integer.*

1. *An h-split of $\overline{\theta}$ is a $(k+h-1)$-dimensional vector $(\theta_1, \ldots, \theta_j^{(h)}, \ldots, \theta_k)$ where, for some $j \in \{1, \ldots, k\}$, the coordinate θ_j is replaced by the h-dimensional vector $\theta_j^{(h)} = (\frac{\theta_j}{h}, \cdots, \frac{\theta_j}{h})$.*
2. *Given vectors $\overline{\theta}$ and $\overline{\eta}$ with nonnegative components we say that $\overline{\theta}$ is h-split majorized by $\overline{\eta}$, represented by $\overline{\theta} \prec_s \overline{\eta}$, if $\overline{\theta}$ can be obtained from $\overline{\eta}$ by a finite number of h-splits.*

It is immediate that h-split majorization implies the usual majorization, as given in Definition 3.1, after filling with zeros the missing components of the shorter vector. Indeed, if $\overline{\theta}$ is obtained from $\overline{\eta} \in \mathbb{R}^k$ by a finite number of h-splits, then $\overline{\theta} \in \mathbb{R}^\ell$ for some $\ell > k$. Then $\overline{\theta} \prec_h \overline{\eta}$ implies $\overline{\theta} \prec \widetilde{\eta}$, where $\widetilde{\eta} = (\eta_1, \ldots, \eta_k, \widetilde{0}_{\ell-k}) \in \mathbb{R}^\ell$ is defined as in Remark 3.2.

We are now ready to state the main result of this section. Let $\otimes$ denote the Kronecker product between vectors.

Proposition 4.2.

1. *Consider weight vectors $\overline{\theta} \in \mathbb{R}^s$ and $\overline{\eta} \in \mathbb{R}^t$. If $\sum_i \theta_i X_i \geq_{st} \sum_j \eta_j X_j$ then, for every vector $\overline{\pi} \in \mathbb{R}^k$,*
$$\sum_{i=1}^{sk} (\overline{\theta} \otimes \overline{\pi})_i X_i \geq_{st} \sum_{j=1}^{tk} (\overline{\eta} \otimes \overline{\pi})_j X_j.$$
2. *If $\overline{X}_s \geq_{st} \overline{X}_t$, then, for every positive integer k, $\overline{X}_{ks} \geq_{st} \overline{X}_{kt}$.*
3. *If $\overline{X}_s \geq_{st} X$ and $\overline{\theta} \prec_s \overline{\eta}$ with nonnegative coordinates, then $\sum_i \theta_i X_i \geq_{st} \sum_j \eta_j X_j$.*
4. *If $\overline{X}_s \geq_{st} \overline{X}_t$, then, for every positive integer k,*
$$\overline{X}_{s+k} \geq_{st} \frac{s}{s+k} \overline{X}_t + \frac{k}{s+k} \overline{X}_k.$$

A description of some dominance relations that can be derived from Proposition 4.2, given some starting points, is given in Figure 1.

It is worth noting that Proposition 4.2 only proves strict implications. Even in the case of the sample mean, in general, $\overline{X}_s \geq_{st} \overline{X}_t$ does not imply that $\overline{X}_k \geq_{st} \overline{X}_t$, for $k > s$. From Example 2.1 in Vincent (2025), it can be seen that the distribution of the St. Petersbourg lottery, represented as $X \overset{d}{=} 2^Y$, where Y is a geometric random variable, satisfies $\overline{X}_2 \geq_{st} X$ but $\overline{X}_3 \not\geq_{st} X$. This result can be proved applying Theorem 4.2 of Vincent (2025).

5 Estimating the distribution of a linear combination

We start now paving the way to tackle the problem of testing if condition (1) holds, which will be done in Section 6. To perform inference on this stochastic dominance relation, we first need an estimator of the distribution of weighted sums of i.i.d. random variables. This estimator can be obtained by applying the plug-in method to an iteration of the weighted convolution operator, that extends the usual convolution, defined as follows.

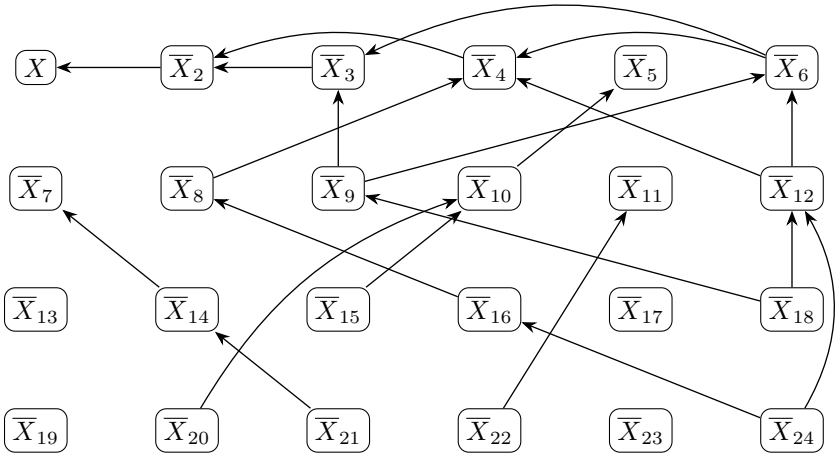

Figure 1: Dominance relations implied by $\overline{X}_3 \geq_{st} \overline{X}_2 \geq_{st} X$, with sample means of size $n \leq 24$.

Definition 5.1. *Let θ_1 and θ_2 be nonnegative real numbers and F and G two CDFs. We define the weighted convolution operator as*

$$L_{\theta_1,\theta_2,F}(G)(x) = \int G(\tfrac{x-\theta_2 t}{\theta_1})\, dF(t), \qquad x \in \mathbb{R}. \tag{3}$$

This operator reduces to the usual convolution when taking $\theta_1 = \theta_2 = 1$, hence it defines an extension of the convolution product. Moreover, it is easily seen that, if X_1 and X_2 are i.i.d. width CDF F, $L_{\theta_1,\theta_2,F}(F)(x) = P(\theta_1 X_1 + \theta_2 X_2 \leq x)$. For higher dimensional sums, a suitable iterated application of the operator gives a characterisation of the distribution function, as it is easily verified that, given $X_1, \ldots, X_s$ i.i.d. with CDF F and weights $\overline{\theta} = (\theta_1, \ldots \theta_s)$,

$$\begin{aligned} F_{\overline{\theta}}(x) &= P(\theta_1 X_1 + \cdots + \theta_s X_s \leq x) \\ &= L_{1,\theta_s,F} \circ L_{1,\theta_{s-1},F} \circ \cdots \circ L_{1,\theta_3,F} \circ L_{\theta_1,\theta_2,F}(F)(x) =: L_{\overline{\theta}}(F). \end{aligned} \tag{4}$$

Given a random sample $\widetilde{X}_1, \ldots, \widetilde{X}_n$ of size n from X, the empirical CDF is denoted, as usual, with $\mathbb{F}_n(x) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}(\widetilde{X}_i \leq x)$, hence the plug-in estimator of $F_{\overline{\theta}}$ is

$$\mathbb{F}_{n,\overline{\theta}}(x) = L_{1,\theta_s,\mathbb{F}_n} \circ L_{1,\theta_{s-1},\mathbb{F}_n} \circ \cdots \circ L_{1,\theta_3,\mathbb{F}_n} \circ L_{\theta_1,\theta_2,\mathbb{F}_n}(\mathbb{F}_n)(x) = L_{\overline{\theta}}(\mathbb{F}_n)(x). \tag{5}$$

We now address the asymptotic properties of the estimator $\mathbb{F}_{n,\overline{\theta}}$. The following result establishes the almost sure uniform convergence of $\mathbb{F}_{n,\overline{\theta}}$ to $F_{\overline{\theta}}$.

Proposition 5.2. *For every $\overline{\theta} = (\theta_1, \ldots, \theta_s) \in (0, +\infty)^s$, $\lim_{n\to+\infty} \left\| \mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{\overline{\theta}} \right\|_\infty = 0$ almost surely.*

Besides the almost sure convergence of $F_{n,\overline{\theta}}$, we need to describe the weak convergence of the associated empirical process. For this purpose, the Hadamard differentiability of $L_{\overline{\theta}}$ is an essential tool.

Proposition 5.3. *Given a vector of weights $\overline{\theta} = (\theta_1, \ldots, \theta_s)$, a CDF F, and $h \in \ell^\infty(\mathbb{R})$, the Hadamard derivative of $L_{\overline{\theta}}$ at F in the direction h is given by*

$$L'_{\overline{\theta}}(F; h) = \sum_{(a_1,\ldots,a_s)\in D} L_{1,\theta_s,a_s} \circ L_{1,\theta_{s-1},a_{s-1}} \circ \cdots \circ L_{1,\theta_3,a_3} \circ L_{\theta_1,\theta_2,a_1}, \tag{6}$$

where $D = \{(h, F, \ldots, F), (F, h, F, \ldots, F), \ldots, (F, \ldots, F, h)\}$.

Next, we prove the weak convergence of the empirical process associated with $\mathbb{F}_{n,\overline{\theta}}$, namely, $\sqrt{n}\left(\mathbb{F}_{n,\overline{\theta}} - F_{\overline{\theta}}\right)$, and characterise the main properties of the limiting process.

Theorem 5.4. *Let F be a continuous CDF. Then, for every $\overline{\theta} = (\theta_1, \ldots, \theta_s) \in (0, +\infty)^s$, there exists a centred Gaussian process $\mathbb{H}^F_{\overline{\theta}}$ such that*

1. *When $n \to +\infty$, $\sqrt{n}\left(\mathbb{F}_{n,\overline{\theta}} - F_{\overline{\theta}}\right) \rightsquigarrow \mathbb{H}^F_{\overline{\theta}} = L'_{\overline{\theta}}(F; \mathbb{B}_F)$, where $\mathbb{B}$ is a Brownian bridge, $\mathbb{B}_F = \mathbb{B} \circ F$, and $L'_{\overline{\theta}}(F; \mathbb{B}_F)$ is the Hadamard derivative of $L_{\overline{\theta}}$ at F in the direction of $\mathbb{B}_F$, given by (6).*
2. *Let $\overline{\theta}_{-j} = (\theta_1, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_s)$ be the $(s-1)$-dimensional vector obtained by removing the j-th coordinate, then*

$$\mathbb{H}^F_{\overline{\theta}} = \sum_{j=1}^{s} \int \mathbb{B}_F\left(\tfrac{x-t}{\theta_j}\right) dF_{\overline{\theta}_{-j}}(t) \; = \sum_{j=1}^{s} \mathbb{B}_F\left(\tfrac{\cdot}{\theta_j}\right) * F_{\overline{\theta}_{-j}}. \tag{7}$$

3. *The covariance operator of $\mathbb{H}^F_{\overline{\theta}}$ is given by $\mathrm{Cov}\left(\mathbb{H}^F_{\overline{\theta}}(x), \mathbb{H}^F_{\overline{\theta}}(y)\right) = \sum_{j,k} \Omega_{j,k}(x,y)$, where*

$$\Omega_{j,k}(x,y) = \begin{cases} \displaystyle\int F_{\overline{\theta}_{-j}}(x - \theta_j u) F_{\overline{\theta}_{-k}}(y - \theta_k u)\, dF(u) - F_{\overline{\theta}}(x) F_{\overline{\theta}}(y), & \text{if } j \neq k, \\ \displaystyle\int F_{\overline{\theta}_{-j}}(x \wedge y - \theta_j u)\, dF(u) - F_{\overline{\theta}}(x) F_{\overline{\theta}}(y), & \text{if } j = k. \end{cases}$$

5.1 The case of Cauchy distributed variables

Let us denote by $\mathcal{C}(x) = \frac{1}{\pi}\arctan(x) + \frac{1}{2}$, $x \in \mathbb{R}$, the standard Cauchy CDF. This special case is of particular interest, as linear combinations of i.i.d. such variables follow a scaled Cauchy distribution. In particular, when the components of the weight vector sum up to 1, the resulting (convex) combination still has distribution $\mathcal{C}$. Hence, we may find a simplified representation of the limiting process $\mathcal{H}^{\mathcal{C}}_{\overline{\theta}}$. Assume then that $X_1, \ldots, X_s$ are i.i.d. with CDF $\mathcal{C}$, and the weights $\overline{\theta} = (\theta_1, \ldots, \theta_s) \in (0, +\infty)^s$ are such that $\|\overline{\theta}\| = 1$, for simplicity. Therefore, $\mathcal{C}_{\overline{\theta}_{-j}}(x) = P\left(\sum_{i \neq j} \theta_i X_i \leq x\right) = \mathcal{C}\left(\frac{x}{1-\theta_j}\right)$. We may state a specific version of Theorem 5.4, whose proof follows immediately from the relation between $\mathcal{C}_{\overline{\theta}_{-j}}$ and $\mathcal{C}$.

Corollary 5.5. *Let $F = \mathcal{C}$. Then:*

1. *The limit process derived in Theorem 5.4 reduces to*

$$\mathbb{H}^{\mathcal{C}}_{\overline{\theta}} = \sum_{j=1}^{s} \int \mathbb{B}_{\mathcal{C}}\left(\tfrac{x-(1-\theta_j)t}{\theta_j}\right) d\mathcal{C}(t) \; = \sum_{j=1}^{s} \mathbb{B}_{\mathcal{C}}\left(\tfrac{\cdot}{\theta_j}\right) * \mathcal{C}(\tfrac{\cdot}{1-\theta_j}). \tag{8}$$

2. *The covariance operator of $\mathbb{H}^{\mathcal{C}}_{\overline{\theta}}$ is given by $\mathrm{Cov}\left(\mathbb{H}^{\mathcal{C}}_{\overline{\theta}}(x), \mathbb{H}^{\mathcal{C}}_{\overline{\theta}}(y)\right) = \sum_{j,k} \Omega^{\mathcal{C}}_{j,k}(x,y)$, where*

$$\Omega^{\mathcal{C}}_{j,k}(x,y) = \begin{cases} \displaystyle\int \mathcal{C}\left(\tfrac{x-\theta_j u}{1-\theta_j}\right) \mathcal{C}\left(\tfrac{y-\theta_k u}{1-\theta_k}\right) d\mathcal{C}(u) - \mathcal{C}(x)\mathcal{C}(y), & \text{if } j \neq k, \\ \displaystyle\int \mathcal{C}\left(\tfrac{x \wedge y - \theta_j u}{1-\theta_j}\right) d\mathcal{C}(u) - \mathcal{C}(x)\mathcal{C}(y), & \text{if } j = k. \end{cases}$$

5.2 A further special case: sample means

Sample means play an important role in statistics, so we give a closer look to the representations corresponding to this case, where $\overline{\theta} = (\frac{1}{s}, \ldots, \frac{1}{s})$. As the weights are now all equal, it is natural to seek expressions where integration is taken with respect to the usual convolution powers. We introduce some specific notation. The convolution power, as mentioned after Definition 5.1, is $F^{*2} = F * F = L_{1,1,F}(F)$, and it is easily seen that $F^{*s} = L_{1,1,F}(F^{*(s-1)})$. As for the Cauchy case, we start be characterising the distributions of the form $\mathrm{F}_{\overline{\theta}_{-j}}$, to see how this reflects on the further integrations. Taking into account that the coordinates of $\overline{\theta}$ are equal to $\frac{1}{s}$, considering $X_1, \ldots, X_s$ i.i.d. with CDF F, we have

$$F_{\overline{\theta}_{-j}}(x) = F_{\overline{X}_{s-1}}\left(\tfrac{sx}{s-1}\right).$$

Rewrite

$$\int \mathbb{B}_F\left(\tfrac{x-t}{\theta_j}\right) dF_{\overline{\theta}_{-j}}(t) = \int \mathbb{B}_F\left(\tfrac{x-t}{1/s}\right) dF_{\overline{\theta}_{-j}}(t) = \int \mathbb{B}_F\left(sx-u\right) dF_{\overline{\theta}_{-j},s}(u),$$

where $F_{\overline{\theta}_{-j},s}(u) = F_{\overline{\theta}_{-j}}(\frac{u}{s}) = F_{\overline{X}_{s-1}}(\frac{x}{s-1}) = F^{*(s-1)}(x)$. Hence, recalling (7), the first assertion in following statement is an immediate consequence of this arguing. The second assertion is a consequence of the fact that power convolutions of a Cauchy distribution coincide with the Cauchy distribution. Finally, a direct translation of the representation of the covariance operator described in Theorem 5.4 (assertion 3.), taking into account the rewriting of $F_{\overline{\theta}_{-j}}$ as above, yields the third assertion.

Corollary 5.6. *Let F be continouous, $s \geq 2$ be fixed, take $\overline{\theta} = (\frac{1}{s}, \ldots, \frac{1}{s})$, and represent the limit of the corresponding empirical process, according to Theorem 5.4, as $\mathbb{H}_s^F$.*

1. *Then $\mathbb{H}_s^F(x) = s\mathbb{B}_F * F^{*(s-1)}(sx)$.*
2. *In the particular case where X is Cauchy distributed, $\mathbb{H}_s^{\mathcal{C}}(x) = s\int \mathbb{B}_{\mathcal{C}}(sx - (s-1)t)\, d\mathcal{C}(t)$.*
3. *The covariance operator of the process is*

$$\mathrm{Cov}\left(\mathbb{H}_s^F(x), \mathbb{H}_s^F(y)\right) = s^2\left(\int F^{*(s-1)}(sx-u)\, F^{*(s-1)}(sy-u)\, dF(u) - F^{*s}(sx)F^{*s}(sy)\right).$$

6 Testing dominance between linear combinations

We consider a test for the null hypothesis $\mathcal{H}_0 : \sum_i \theta_i X_i \geq_{st} \sum_i \eta_i X_i$ versus the alternative $\mathcal{H}_1 : \sum_i \theta_i X_i \not\geq_{st} \sum_i \eta_i X_i$, for given weight vectors $\overline{\theta} = (\theta_1, \ldots, \theta_s) \in (0, +\infty)^s$ and $\overline{\eta} = (\eta_1, \ldots, \eta_t) \in (0, +\infty)^t$, where s and t are not necessarily equal. As is customary in the literature on stochastic dominance tests (see Chapter 2.2.2 of Whang (2019) for a review), we consider a Kolmogorov–Smirnov-type test statistic, based on the sup-norm of the difference between empirical versions of $F_{\overline{\theta}}$ and $F_{\overline{\eta}}$, namely,

$$T_{\overline{\theta},\overline{\eta}}(\mathbb{F}_n) = \left\| \mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}} \right\|_\infty, \tag{9}$$

where $\mathbb{F}_{n,\overline{\theta}} = L_{\overline{\theta}}(\mathbb{F}_n)$ and $\mathbb{F}_{n,\overline{\eta}} = L_{\overline{\eta}}(\mathbb{F}_n)$ are given by (5). Differently from the classic problem of testing stochastic dominance, in our case the test statistic is a functional of only one empirical CDF, $\mathbb{F}_n$. This test statistic is clearly location and scale invariant.

We propose two different tests both based on $T_{\overline{\theta},\overline{\eta}}$. The first one, based on the Cauchy distribution, applies to weight vectors such that $\|\overline{\theta}\| = \|\overline{\eta}\| = 1$, exploits the fact that convex combinations of Cauchy variables are also Cauchy distributed, making it possible to identify the asymptotic least favourable distribution of the test statistic under the null hypothesis. Actually, this test may be adapted to general weight vectors by suitable scaling, thus at the cost of some extra notational complexity. The second test is based on a bootstrap procedure, and does not depend on any restriction on the weight vectors. Both tests are consistent and have asymptotically bounded size under the null hypothesis. For convenience, we shall be writing $F \in \mathcal{H}_0$ whenever this hypothesis is satisfied for random variables with CDF F, and similarly for other hypothesis under consideration.

6.1 The test based on the Cauchy distribution

Recall that for this subsection we will be assuming that $\|\overline{\theta}\| = \|\overline{\eta}\| = 1$. We have denoted the standard Cauchy CDF as $\mathcal{C}$ (see Subsection 5.1), so its stability with respect to convex combinations implies that $L_{\overline{\theta}}(\mathcal{C}) = \mathcal{C}$. Moreover, it is clear that $C \in \mathcal{H}_0$. Therefore, taking into account the behaviour of the distribution of convex combinations of i.i.d. random variables, one could expect that, if $F = \mathcal{C}$, $\mathbb{F}_{n,\overline{\theta}}$ and $\mathbb{F}_{n,\overline{\eta}}$, the estimated distributions of $\sum_i \theta_i X_i$ and $\sum_i \eta_i X_i$, respectively, will be "close" to each other, and to $\mathcal{C}$, especially for large n, compared to the general case when $F \in \mathcal{H}_0$. In other words, we expect that the Cauchy case is the least favourable for the distribution of $T_{\overline{\theta},\overline{\eta}}$ under the null hypothesis. Focus now on the strict sub-null $\widetilde{\mathcal{H}}_0 = \{F : F_{\overline{\theta}}(x) < F_{\overline{\eta}}(x) \text{ a.e.}\}$, thus excluding boundary cases for which $F_{\overline{\theta}}$ and $F_{\overline{\eta}}$ may coincide on a set of non-null Lebesgue measure. Note that, if X has an almost everywhere analytic density function, which includes all the standard families of continuous distributions, the coincidence on an interval of the densities implies that they coincide in the whole domain, so this case is excluded from $\mathcal{H}_0 \setminus \widetilde{\mathcal{H}}_0$. Apart from the special case of $\mathcal{C}$ (or the trivial case of distributions which coincide to 0 outside their supports, which is irrelevant for our analysis), we are not aware of any further examples of distributions in $\mathcal{H}_0 \setminus \widetilde{\mathcal{H}}_0$.

Theorem 6.1. *Assume that F is continuous.*

1. *If $F \in \widetilde{\mathcal{H}}_0$, then $\sqrt{n}\left\|\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}}\right\|_\infty \to_p 0$.*
2. *If $F = \mathcal{C}$, then $\sqrt{n}\left\|\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}}\right\|_\infty \to_p \left\|\mathbb{H}^{\mathcal{C}}_{\overline{\theta}} - \mathbb{H}^{\mathcal{C}}_{\overline{\eta}}\right\|_\infty \geq 0$.*

We now describe the test based on this particular behaviour of the Cauchy distribution.

The Cauchy based test:

1. *Denote with $\mathbb{C}_n$ the empirical CDF obtained by sampling from the distribution $\mathcal{C}$.*
2. *Generate, via the Monte Carlo method, the critical value $c_{\alpha,n}$ as the $(1-\alpha)$-quantile of $\sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(\mathbb{C}_n)$.*
3. *Reject $\mathcal{H}_0$ if $\sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(F_n) \geq c_{\alpha,n}$, where F_n is a realization of $\mathbb{F}_n$.*
4. *The p-value is given by $P(T_{\overline{\theta},\overline{\eta}}(\mathbb{C}_n) \geq T_{\overline{\theta},\overline{\eta}}(F_n))$.*

This test has the appropriate properties, as shown in the next statement.

Proposition 6.2. *Assume the CDF F is continuous.*

1. *If $F = \mathcal{C}$, $\lim_{n\to\infty} P(\text{reject } \mathcal{H}_0) = \alpha$.*

2. *If* $F \in \widetilde{\mathcal{H}}_0$, $\lim_{n\to\infty} P(\text{reject } \mathcal{H}_0) = 0$.

3. *If* $F \in \mathcal{H}_1$, $\lim_{n\to\infty} P(\text{reject } \mathcal{H}_0) = 1$.

Although the boundary case when $F \in \mathcal{H}_0 \setminus \widetilde{\mathcal{H}}_0$ seems to be practically negligible, the behaviour of the test in this case could not be established. In fact, the analysis would require a much finer control of the covariance structure of the limiting Gaussian process, which we are currently unable to obtain. For this reason, we develop and study a second approach based on the bootstrap.

6.2 A bootstrap test

We now propose an alternative test, where critical values are computed via bootstrap, in the spirit of Barrett and Donald (2003) (see also Lando and Legramanti (2025); Barrett et al. (2014)). Under the null hypothesis, $F_{\overline{\theta}}(x) - F_{\overline{\eta}}(x) \leq 0$ for every $x \in \mathbb{R}$. Therefore,

$$\mathbb{F}_{n,\overline{\theta}}(x) - \mathbb{F}_{n,\overline{\eta}}(x) \leq \left(\mathbb{F}_{n,\overline{\theta}}(x) - \mathbb{F}_{n,\overline{\eta}}(x)\right) - \left(F_{\overline{\theta}}(x) - F_{\overline{\eta}}(x)\right), \tag{10}$$

which implies that

$$\sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(\mathbb{F}_n) \leq \sqrt{n} \left\| \left(\left(\mathbb{F}_{n,\overline{\theta}}(x) - \mathbb{F}_{n,\overline{\eta}}(x)\right) - \left(F_{\overline{\theta}}(x) - F_{\overline{\eta}}(x)\right) \right)_+ \right\|_\infty \to_d \left\| \mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}} \right\|_\infty . \tag{11}$$

The rejection region for this test will be based in the distribution of $\left\| \mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}} \right\|_\infty$, thus defining a conservative test. For practical matters, the distribution of $\left\| \mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}} \right\|_\infty$ can be approximated by bootstrap. Before describing the test, we justify the bootstrapping procedure. Let $\mathbb{F}^b_n(x) = \frac{1}{n}\sum_{i=1}^n M_i \mathbb{1}(\widetilde{X}_i \leq x)$, where M_i is the number of times $\widetilde{X}_i$ is resampled from the original sample, be the bootstrap estimator of $\mathbb{F}_n$. Correspondingly, we can compute $\mathbb{F}^b_{n,\overline{\theta}}$ and define the bootstrap version of the bounding process in (10)

$$\mathcal{G}_n = \sqrt{n}\left((\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}}) - (F_{\overline{\theta}} - F_{\overline{\eta}}) \right)$$

with

$$\mathcal{G}^b_n = \sqrt{n}\left(\left(\mathbb{F}^b_{n,\overline{\theta}} - \mathbb{F}^b_{n,\overline{\eta}}\right) - \left(\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}}\right) \right).$$

The functional delta method for the bootstrap implies that these two processes have the same limit: $\mathcal{G}^b_n \rightsquigarrow \mathbb{H}_{\overline{\theta}} - \mathbb{H}_{\overline{\eta}}$, which then implies that $\left\|\mathcal{G}^b_n\right\|_\infty \to_d \left\| \mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}} \right\|_\infty$.

The bootstrap test

1. *Build* $b = 1, \ldots, B$ *resamples from the empirical CDF* F_n *of the original sample, getting the corresponding empirical distribution functions* F^b_n*, as described above.*

2. *Get the bootstrapped versions of the distribution of the weighted sums* $F^b_{n,\overline{\theta}}$ *and* $F^b_{n,\overline{\eta}}$.

3. *The bootstrap p-value, defined as* $p = P\left(\left\|\mathcal{G}^b_n\right\|_\infty > \sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(F_n)\right)$*, is approximated by*

$$\widehat{p} = \frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\left(\left\|\mathcal{G}^b_n\right\|_\infty > \sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(F_n) \right).$$

4. *The rejection region of the test with significance level* $\alpha \in (0,1)$ *is described by* $\sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(F_n) > c_{\alpha,n}$, *where*

$$c_{\alpha,n} = \inf\left\{ y : P\left(\left\| \mathcal{G}_n^b \right\|_\infty > y | \widetilde{X}_1, \ldots, \widetilde{X}_n \right) \leq \alpha \right\}. \tag{12}$$

We may now prove the relevant asymptotic properties of the bootstrap test.

Proposition 6.3. *Assume* F *is continuous.*

1. *If* $F \in \mathcal{H}_0$, $\lim_{n\to\infty} P(\textit{reject } \mathcal{H}_0) \leq \alpha$.
2. *If* $F \in \mathcal{H}_1$, $\lim_{n\to\infty} P(\textit{reject } \mathcal{H}_0) = 1$.

7 Simulations

To give some insights on the performance of the tests, we provide simulated rejection rates, computed at significance level $\alpha = 0.1$. We consider sample sizes $n = 100$ or $n = 500$. All simulations reported considered 1000 replications, either for bootstrapping or Monte Carlo simulations. In Subsection 7.1 we focus on testing dominance between the sample mean and the underlying variable, as this example is particularly simple and relevant. Then, in Subsection 7.2, we extend our simulation study to a more general framework, to compare linear combinations with different weights and dimensions. This includes the case in which the weight vectors are not ordered in majorization, for which the theory discussed in Section 4 does not give any indication.

7.1 Tests for $\overline{X}_s \geq_{st} X$

We first concentrate on testing the null hypothesis $\mathcal{H}_0 : \overline{X}_s \geq_{st} X$, for values of $s = 2, 3, 4$. We focus on distributions that, for suitable choice of parameters, are known to satisfy this dominance relation, namely, the Pareto, loglogistic or the Fréchet families of distributions, with CDFs $\mathcal{P}(x; sh) = 1 - \frac{1}{x^{sh}}$, $x \geq 1$, $\mathcal{F}(x; sh) = \exp(-x)^{-sh}$, $x \geq 0$ and $\mathcal{L}(x; sh) = \frac{x^{sh}}{1+x^{sh}}$, $x \geq 0$, respectively, where sh is a positive shape parameter in all three cases. Recall that $\mathcal{H}_0$ can only hold for distributions with infinite mean, which is the case of the families of distributions just mentioned for $sh \leq 1$. Differently, when the shape parameter exceeds one, the CDFs of the linear combinations cross, and it should be easier to detect departures from $\mathcal{H}_0$ as sh gets larger. We are interested in discovering this behaviour by increasing the value of sh, as well as the sample size. For all these cases, the null hypothesis is always true, or always false, for every choice of s, provided that $sh \leq 1$, or $sh > 1$, respectively. Then, we analyse a case for which $\mathcal{H}_0$ is true for some values of s and false for some others. In particular, let $X \overset{d}{=} 0.45Y + 0.55Z$, where Y Bernoulli $(\frac{1}{2})$ and Z is a Pareto with $sh = 1$. For this model, the null hypothesis is false for $s = 2, 3$ and true for $s = 4$ (see Figure 3). We also consider the Students' t distribution with df degrees of freedom. In this case, the CDFs of the linear combinations cross for every choice of df, except for the case $df = 1$, which coincides with the Cauchy distribution, where we expect to observe a power close to α.

As a general comment, not unexpectedly, the bootstrap test shows an overall better power behaviour. However, the running times are considerably smaller for the Cauchy-based test: when simulating a comparison between some sample mean and the underlying random variable, while the bootstrap test takes around 15 minutes running on a `i9` machine, the Cauchy-based test needs just a few seconds to find results when running on a `i7` machine.

	$sh=1$		$sh=2$		$sh=3$		$sh=4$		$sh=5$	
n	100	500	100	500	100	500	100	500	100	500
Bootstrap test										
$s=2$	0.000	0.000	0.051	0.104	0.133	0.358	0.208	0.585	0.271	0.709
$s=3$	0.001	0.000	0.078	0.153	0.189	0.559	0.307	0.808	0.395	0.924
$s=4$	0.000	0.000	0.067	0.208	0.206	0.684	0.329	0.889	0.433	0.965
Cauchy-based test										
$s=2$	0.002	0.001	0.030	0.019	0.066	0.114	0.093	0.178	0.117	0.247
$s=3$	0.000	0.000	0.017	0.034	0.052	0.206	0.121	0.438	0.176	0.542
$s=4$	0.000	0.000	0.021	0.055	0.066	0.337	0.121	0.598	0.134	0.754

Table 1: Test for $\overline{X}_s \geq_{st} X$. Simulated power with Pareto alternatives.

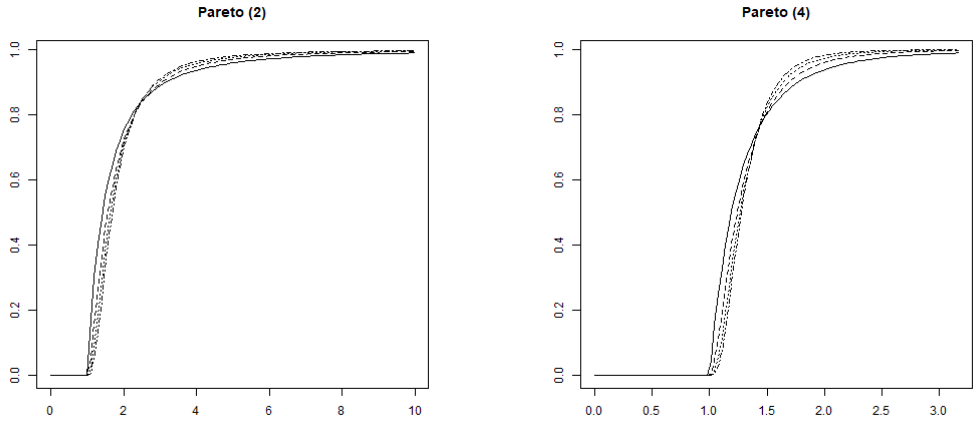

Figure 2: CDF for the Pareto distribution (solid) and the corresponding CDFs for the sample mean of size $s=2$ (dashed), $s=3$ (dotted) and $s=4$ (dashdotted).

The simulated power in the Pareto case is given in Table 1. For this model, the CDFs cross for $sh > 1$, but they are very close one to another, making it particularly difficult to detect non-dominance, especially for smaller values of sh and s (see Figure 2). The results show that the bootstrap test outperforms the Cauchy-based test. Moreover, the power increases with respect to n, sh and s.

The performance of the tests in the loglogistic and Fréchet cases shows the same trend as in the Pareto case, with respect to n, s and sh. In both these cases, the power behaviour of the test is better than the one observed for the Pareto alternatives, meaning that dominance of sample means for these distributions is easier to identify. The simulated powers are reported in Tables 2 and 3. It is worth noticing that the Cauchy-based test behaves almost as well as the bootstrap test for the loglogistic alternatives, while it lacks power when dealing with the Pareto or Fréchet alternatives.

The general behaviour is confirmed also in the case of the Student's t distribution, where the power gets larger whenever $df < 1$ or $df > 1$. As expected, when $df = 1$, the simulated power is close to the significance level $\alpha = 0.1$ we used for the numerical study. The simulated power values are reported in Table 4.

Finally, we consider the mixture case discussed earlier. The results are described in Table 5. First note that the crossing of the CDF for $s = 3$ is by a very slight margin, making it difficult to detect (see Figure 3). Both tests correctly identify the nondominance

	$sh=1$		$sh=2$		$sh=3$		$sh=4$		$sh=5$	
n	100	500	100	500	100	500	100	500	100	500
Bootstrap test										
$s=2$	0.002	0.000	0.109	0.206	0.304	0.678	0.451	0.888	0.557	0.956
$s=3$	0.000	0.000	0.130	0.323	0.393	0.862	0.586	0.985	0.739	0.999
$s=4$	0.000	0.000	0.145	0.423	0.402	0.941	0.637	0.998	0.767	1.000
Cauchy-based test										
$s=2$	0.001	0.000	0.048	0.037	0.127	0.368	0.330	0.628	0.333	0.750
$s=3$	0.002	0.000	0.039	0.118	0.185	0.661	0.358	0.895	0.444	0.979
$s=4$	0.000	0.000	0.059	0.188	0.220	0.807	0.336	0.969	0.528	0.998

Table 2: Test for $\overline{X}_s \geq_{st} X$. Simulated power with loglogistic alternatives.

	$sh=1$		$sh=2$		$sh=3$		$sh=4$		$sh=5$	
n	100	500	100	500	100	500	100	500	100	500
Bootstrap test										
$s=2$	0.001	0.000	0.087	0.156	0.212	0.507	0.322	0.768	0.403	0.856
$s=3$	0.001	0.000	0.097	0.254	0.279	0.748	0.432	0.940	0.558	0.987
$s=4$	0.002	0.000	0.114	0.321	0.296	0.865	0.478	0.975	0.612	0.996
Cauchy-based test										
$s=2$	0.001	0.000	0.053	0.041	0.096	0.203	0.181	0.379	0.228	0.557
$s=3$	0.002	0.000	0.035	0.072	0.114	0.442	0.136	0.706	0.272	0.841
$s=4$	0.001	0.000	0.029	0.132	0.143	0.554	0.186	0.866	0.258	0.948

Table 3: Test for $\overline{X}_s \geq_{st} X$. Simulated power with Fréchet alternatives.

	$df=.5$		$df=1$		$df=1.5$		$df=3$		$df=5$	
n	100	500	100	500	100	500	100	500	100	500
Bootstrap test										
$s=2$	0.444	0.903	0.118	0.117	0.247	0.483	0.691	0.984	0.883	1.000
$s=3$	0.513	0.953	0.122	0.102	0.290	0.696	0.805	1.000	0.954	1.000
$s=4$	0.540	0.974	0.108	0.111	0.287	0.726	0.866	1.000	0.986	1.000
Cauchy-based test										
$s=2$	0.409	0.736	0.119	0.095	0.193	0.307	0.544	0.914	0.744	0.993
$s=3$	0.506	0.932	0.071	0.103	0.253	0.546	0.639	0.994	0.846	1.000
$s=4$	0.482	0.959	0.086	0.107	0.221	0.582	0.691	1.000	0.905	1.000

Table 4: Test for $\overline{X}_s \geq_{st} X$. Simulated power with Student's t alternatives.

	$s=2$		$s=3$		$s=4$		$s=5$	
n	100	500	100	500	100	500	100	500
Bootstrap test	0.058	0.520	0.004	0.003	0.003	0.000	0.001	0.000
Cauchy-based test	0.105	0.521	0.028	0.091	0.006	0.002	0.002	0.000

Table 5: Test for $\overline{X}_s \geq_{st} X$. Simulated power with alternative $X \overset{d}{=} 0.45Y + 0.55Z$, Y Bernoulli $(\frac{1}{2})$, Z Pareto with $sh = 1$.

of $\overline{X}_2$, although they need a large sample to reach a reasonable power. For this mixture distribution the Cauchy-based test delivers a larger power than the bootstrap test.

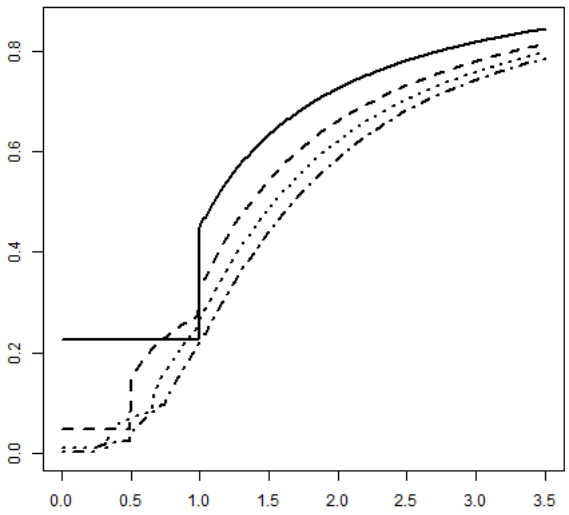

Figure 3: CDFs of X (solid), $\overline{X}_2$ (dashed), $\overline{X}_3$ (dotted), and $\overline{X}_4$ (dashdotted). $X \overset{d}{=} 0.45Y + 0.55Z$, Y Bernoulli $(\frac{1}{2})$, Z Pareto with $sh = 1$.

7.2 Testing dominance between different linear combinations

In this subsection, we compare linear combinations with different weight vectors. The sample size in this case is fixed, $n = 500$. We start by testing $\mathcal{H}_0 : \theta_1 X_1 + \theta_2 X_2 \geq_{st} X$. The results are reported in Tables 6 and 7, using the same set of distributions as alternatives as in Subsection 7.1. The results exhibit the same pattern as in the comparison of sample means. Moreover, power is clearly higher for weight vectors close to $(\frac{1}{2}, \frac{1}{2})$, i.e. those defining the sample mean, and lower when the weights concentrate on a single observation. This indicates that majorization influences non-dominance (CDF crossings), in line with the classical finite-mean theory linking majorization of weights to convex-order comparisons of weighted sums (Shaked and Shantikumar, 2007, Ch. 3.A). Expectedly, although the Cauchy-based test is much faster to run, it has a significantly lower power.

We now focus on the tests for $\mathcal{H}_0 : \theta_1 X_1 + \theta_2 X_2 \geq_{st} \overline{X}_2$. The results are reported in Tables 8 and 9. Even in this case, the bootstrap test consistently outperforms the Cauchy-based test.

We finally apply our tests to the case when the vectors are not ordered by majorization. In this case, the theory does not provide any indication. Dominance can still hold, but we do not know in which direction, that is, we don't know which linear combination dominates the other. With this motivation, we test $\mathcal{H}_0 : \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_3 \geq_{st} \eta_1 X_1 + \eta_2 X_2 + \eta_3 X_3$ as well the reversed null hypothesis $\mathcal{H}_0 : \eta_1 X_1 + \eta_2 X_2 + \eta_3 X_3 \geq_{st} \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_3$.

Pareto (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.4, 0.6)	0.000	0.100	0.372	0.570	0.708
(0.2, 0.8)	0.000	0.119	0.344	0.484	0.618
loglogistic (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.4, 0.6)	0.000	0.220	0.670	0.885	0.967
(0.2, 0.8)	0.000	0.197	0.587	0.780	0.859
Fréchet (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.4, 0.6)	0.000	0.162	0.514	0.761	0.869
(0.2, 0.8)	0.000	0.157	0.451	0.646	0.765
Student (df)	$df=0.5$	$df=1$	$df=1.5$	$df=3$	$df=5$
(0.4, 0.6)	0.899	0.121	0.490	0.984	1.000
(0.2, 0.8)	0.850	0.119	0.435	0.940	0.986

Table 6: Simulated power of the bootstrap test for weighted combinations.

Pareto (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.4, 0.6)	0.000	0.020	0.042	0.136	0.244
(0.2, 0.8)	0.000	0.006	0.069	0.054	0.104
loglogistic (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.4, 0.6)	0.000	0.056	0.348	0.594	0.740
(0.2, 0.8)	0.000	0.025	0.160	0.363	0.445
Fréchet (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.4, 0.6)	0.000	0.035	0.147	0.353	0.514
(0.2, 0.8)	0.000	0.022	0.107	0.161	0.264
Student (df)	$df=0.5$	$df=1$	$df=1.5$	$df=3$	$df=5$
(0.4, 0.6)	0.782	0.104	0.303	0.917	0.990
(0.2, 0.8)	0.661	0.102	0.233	0.637	0.836

Table 7: Simulated power of the Cauchy-based test for weighted combinations.

Pareto (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.1, 0.9) vs. (0.5, 0.5)	0.000	0.063	0.159	0.258	0.336
(0.25, 0.75) vs. (0.5, 0.5)	0.037	0.048	0.084	0.118	0.138
loglogistic (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.1, 0.9) vs. (0.5, 0.5)	0.004	0.137	0.356	0.530	0.663
(0.25, 0.75) vs. (0.5, 0.5)	0.039	0.103	0.165	0.243	0.301
Fréchet (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.1, 0.9) vs. (0.5, 0.5)	0.003	0.105	0.276	0.409	0.523
(0.25, 0.75) vs. (0.5, 0.5)	0.044	0.090	0.142	0.185	0.222
Student (df)	$df=0.5$	$df=1$	$df=1.5$	$df=3$	$df=5$
(0.1, 0.9) vs. (0.5, 0.5)	0.351	0.115	0.239	0.724	0.898
(0.25, 0.75) vs. (0.5, 0.5)	0.210	0.121	0.150	0.289	0.428

Table 8: Simulated power of the bootstrap test between different weighted combinations.

Pareto (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.1, 0.9) vs. (0.5, 0.5)	0.002	0.011	0.032	0.064	0.120
(0.25, 0.75) vs. (0.5, 0.5)	0.028	0.024	0.034	0.023	0.022
loglogistic (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.1, 0.9) vs. (0.5, 0.5)	0.001	0.024	0.100	0.239	0.301
(0.25, 0.75) vs. (0.5, 0.5)	0.045	0.053	0.052	0.089	0.101
Fréchet (sh)	$sh=1$	$sh=2$	$sh=3$	$sh=4$	$sh=5$
(0.1, 0.9) vs. (0.5, 0.5)	0.002	0.015	0.092	0.129	0.161
(0.25, 0.75) vs. (0.5, 0.5)	0.036	0.033	0.020	0.046	0.044
Student (df)	$df=0.5$	$df=1$	$df=1.5$	$df=3$	$df=5$
(0.1, 0.9) vs. (0.5, 0.5)	0.156	0.084	0.160	0.481	0.670
(0.25, 0.75) vs. (0.5, 0.5)	0.114	0.101	0.113	0.141	0.180

Table 9: Simulated power of the Cauchy-based test between different weighted combinations.

	$sh=1$	$sh=0.8$	$sh=0.6$
Bootstrap test			
$\overline{\theta}=(0.1, 0.35, 0.55)$ vs. $\overline{\eta}=(0.15, 0.25, 0.6)$	0.202	0.209	0.236
$\overline{\eta}=(0.15, 0.25, 0.6)$ vs. $\overline{\theta}=(0.1, 0.35, 0.55)$	0.109	0.104	0.115
Cauchy-based test			
$\overline{\theta}=(0.1, 0.35, 0.55)$ vs. $\overline{\eta}=(0.15, 0.25, 0.6)$	0.120	0.126	0.166
$\overline{\eta}=(0.15, 0.25, 0.6)$ vs. $\overline{\theta}=(0.1, 0.35, 0.55)$	0.093	0.069	0.089

Table 10: Simulated power to detect the direction of stochastic dominance in the presence of nonmajorized weights (underlying distribution is the loglogistic with shape parameter sh).

In all such cases X has a loglogistic distribution. Two choices of $\overline{\eta}$ and $\overline{\eta}$ such that $\overline{\eta} \not\succ \overline{\theta}$ and $\overline{\theta} \not\succ \overline{\eta}$ are considered in Table 10 and Table 11. In the first case, the two CDFs are very close and almost indistinguishable, similarly to the Cauchy case. The tests behave accordingly, showing power close to α in one direction and slightly larger than α in the other direction (where the null is false). In the latter case, one distribution dominates the other more clearly, especially for smaller values of the shape parameter. Both tests are able to detect this situation, showing power close to 0 when the null is true, and always larger than 0.7 when the null is false. We finally study a three-dimensional case where $\overline{\eta}$ and $\overline{\theta}$ are ordered in majorization. The results, reported in Table 12, show that, in all cases considered, the power is close to 0 when the null is true, and close to 1 when it is false.

A Proofs

Proof of Theorem 3.7. We follow the initial arguments of the proof of Theorem 1 in Chen et al. (2025), that we include here for sake of completeness. Given Lemma 3.3, it is enough to verify the result when $\overline{\theta}$ and $\overline{\eta}$, are related by a T-transform, and, as these transformations only change two coordinates and the X_i are i.i.d., the result will follow if we prove the result for $s=2$. Moreover, it is obvious that we may take the coordinates of $\overline{\theta}$ and $\overline{\eta}$ ordered and such that $\theta_1+\theta_2=\eta_1+\eta_2=1$, meaning that $\overline{\theta}=(\theta, 1-\theta)$, $\overline{\eta}=(\eta, 1-\eta)$, $0\le\eta\le\theta\le\frac{1}{2}$. Hence, we need to prove that, for this choice the coefficients

	$sh = 1$	$sh = 0.8$	$sh = 0.6$
Bootstrap test			
$\overline{\theta} = (0.09, 0.41, 0.5)$ vs. $\overline{\eta} = (0.1, 0.1, 0.8)$	0.05	0.002	0.001
$\overline{\eta} = (0.1, 0.1, 0.8)$ vs. $\overline{\theta} = (0.09, 0.41, 0.5)$	0.941	0.941	0.864
Cauchy-based test			
$\overline{\theta} = (0.09, 0.41, 0.5)$ vs. $\overline{\eta} = (0.1, 0.1, 0.8)$	0.003	0.006	0.005
$\overline{\eta} = (0.1, 0.1, 0.8)$ vs. $\overline{\theta} = (0.09, 0.41, 0.5)$	0.887	0.795	0.735

Table 11: Simulated power to detect the direction of stochastic dominance in the presence of nonmajorized weights (underlying distribution is the loglogistic with shape parameter sh).

	$sh = 1$	$sh = 0.8$	$sh = 0.6$
Bootstrap test			
$\overline{\theta} = (0.2, 0.3, 0.5)$ vs. $\overline{\eta} = (0.1, 0.1, 0.8)$	0.998	0.987	0.972
$\overline{\eta} = (0.1, 0.1, 0.8)$ vs. $\overline{\theta} = (0.2, 0.3, 0.5)$	0.003	0.001	0.000
Cauchy-based test			
$\overline{\theta} = (0.2, 0.3, 0.5)$ vs. $\overline{\eta} = (0.1, 0.1, 0.8)$	0.985	0.956	0.921
$\overline{\eta} = (0.1, 0.1, 0.8)$ vs. $\overline{\theta} = (0.2, 0.3, 0.5)$	0.001	0.000	0.001

Table 12: Simulated power with majorized weights (underlying distribution is the loglogistic with shape parameter sh).

θ and η,

$$P(\theta X_1 + (1-\theta)X_2 > x) > P(\eta X_1 + (1-\eta)X_2 > x),$$

holds for every $x \geq 0$. As in the proof of Theorem 1 in Chen et al. (2025), we may write

$$P(\theta X_1 + (1-\theta)X_2 > x) = P(X_1 > x, X_2 > x) + \int_0^x \overline{F}\left(\tfrac{x-\theta t}{1-\theta}\right) + \overline{F}\left(\tfrac{x-(1-\theta)t}{\theta}\right) dF(t),$$

which then follows if we prove that

$$V(\theta) = \overline{F}\left(\tfrac{x-\theta t}{1-\theta}\right) + \overline{F}\left(\tfrac{x-(1-\theta)t}{\theta}\right)$$

is increasing with respect to $\theta \in [0,1]$. Note now that $\frac{x-\theta t}{1-\theta} = x + (x-t)\frac{\theta}{1-\theta}$ and $\frac{x-(1-\theta)t}{\theta} = x + (x-t)\frac{1-\theta}{\theta}$. As $\frac{\theta}{1-\theta}$ is increasing in θ, ranging from 0 to 1 when $\theta \in [0,1]$, and $t \in [0,x]$, representing $y = x - t$ and $z = \frac{\theta}{1-\theta}$, the increasingness of V follows directly from the definition of the class $\mathcal{L}$, concluding the proof. □

Proof of Proposition 3.8. Obviously, it holds that $F(s) = 1 - H(\frac{1}{s})$, for every $s > 0$. As H is the CDF of X^{-1}, it is increasing. It is easily seen that the function introduced in Definition 3.6 may be represented as $V(z) = 2 - H(a(z)) - H(b(z))$, where $a(z) = \frac{1}{x+yz}$ and $b(z) = \frac{1}{x+\frac{y}{z}}$. As $z \in (0,1]$, it follows that $a(z) \geq b(z)$, which then implies that $H'(a(z)) \leq H'(b(z))$, due to the concavity of H. Moreover,

$$a'(z) + b'(z) = \frac{y(1-z^2)(x^2-y^2)}{(x+yz)(xz+y)} \geq 0,$$

hence $a(z) + b(z)$ is increasing. Finally,

$$\big(H(a(z)) + H(b(z))\big)' = H'(a(z))a'(z) + H'(b(z))b'(z) \geq H'(a(z))\big(a'(z) + b'(z)\big) \geq 0,$$

that is, $H(a(z)) + H(b(z))$ is increasing. Therefore, $V(z) = 2 - \big(H(a(z)) + H(b(z))\big)$ is decreasing. □

Proof of Proposition 4.2. It is well known that $U \geq_{st} V$ implies $\sum_k \pi_k U_k \geq_{st} \sum_k \pi_k V_k$, by closure of stochastic dominance under convolution and nonnegative scaling (see for instance Theorem 1.A.3 of Shaked and Shantikumar (2007)). Taking $U = \sum_i \theta_i X_i$ and $V = \sum_j \eta_j X_j$, the first assertion follows immediately since

$$\sum_{\ell=1}^{k} \pi_\ell U_\ell = \sum_{i}^{nk} (\overline{\theta} \otimes \overline{\pi})_i X_i \quad \text{and} \quad \sum_{\ell=1}^{k} \pi_\ell V_\ell = \sum_{j}^{mk} (\overline{\eta} \otimes \overline{\pi})_j X_j,$$

where U_ℓ and V_ℓ are i.i.d. copies of U and V, respectively.

The second assertion is an immediate consequence of assertion 1.

For the third assertion, it is enough to prove the result for a single s-split. Let $\overline{\eta} = (\eta_1, \ldots, \eta_k)$ and fix $j \in \{1, \ldots, k\}$ so that

$$\overline{\theta} = (\eta_1, \ldots, \eta_{j-1}, \underbrace{\tfrac{\eta_j}{s}, \ldots, \tfrac{\eta_j}{s}}_{s \text{ times}}, \eta_{j+1}, \ldots, \eta_k).$$

Write

$$\sum_{i=1}^{k} \eta_i X_i = \eta_j X_j + \sum_{i \neq j} \eta_i X_i =: \eta_j X_j + Z,$$

where Z is independent of $(X_j, X_{k+1}, \ldots, X_{k+s-1})$. After splitting η_j into s equal parts we obtain

$$\sum_{i=1}^{k+s} \theta_i X_i = \frac{\eta_j}{s} \left(X_j + X_{k+1} + \cdots + X_{k+s-1}\right) + Z = \eta_j \overline{X}_s^{(j)} + Z,$$

where $\overline{X}_s^{(j)} := \frac{1}{s}(X_j + X_{k+1} + \cdots + X_{k+s-1})$ has the same distribution as $\overline{X}_s$. Multiplying by the nonnegative constant η_j preserves the dominance $\overline{X}_s^{(j)} \geq_{st} X_j$, hence $\eta_j \overline{X}_s^{(j)} \geq_{st} \eta_j X_j$. Finally, adding Z to both sides preserves the stochastic order, so

$$\sum_{i=1}^{k+s} \theta_i X_i = \eta_j \overline{X}_s^{(j)} + Z \geq_{st} \eta_j X_j + Z = \sum_{j=1}^{k} \eta_j X_j,$$

completing the proof of this assertion.

Concerning assertion 4., write $\overline{X}_{s+k}$ as the weighted average:

$$\overline{X}_{s+k} = \frac{s}{s+k} \left(\frac{1}{s} \sum_{i=1}^{s} X_i\right) + \frac{k}{s+k} \left(\frac{1}{k} \sum_{i=s+1}^{s+k} X_i\right) = \frac{s}{s+k} \overline{X}_s + \frac{k}{s+k} \overline{X}_k,$$

where the two block means are independent. Replacing $\overline{X}_s$ with $\overline{X}_t$, which is stochastically smaller by assumption, gives the result. □

Proof of Proposition 5.3. The weighted convolution operator is obviously linear with respect to each CDF argument (see (3)). Therefore, given $t \in \mathbb{R}$ and h_t, $h \in \ell^\infty(\mathbb{R})$, such that $\lim_{t \to 0} \|h_t - h\|_\infty = 0$, using this bilinearity,

$$L_{\theta_1,\theta_2,F+th_t}(F + th_t) - L_{\theta_1,\theta_2,F}(F) = t\left(L_{\theta_1,\theta_2,F}(h_t) + L_{\theta_1,\theta_2,h_t}(F)\right) + t^2 L_{\theta_1,\theta_2,h_t}(h_t),$$

it follows that, when $\overline{\theta} = (\theta_1, \theta_2)$, the Hadarmard derivative at F in the direction h of the operator is $L'_{(\theta_1,\theta_2)}(F;h) = L_{\theta_1,\theta_2,F}(h) + L_{\theta_1,\theta_2,h}(F)$. For a general vector $\overline{\theta} = (\theta_1, \ldots, \theta_s)$ note first that $L_{(\theta_1,\ldots,\theta_s)}(F) = L_{1,\theta_s,F}(L_{(\theta_1,\ldots,\theta_{s-1})}(F))$. Applying the chain rule to this representation, one gets

$$L'_{(\theta_1,\ldots,\theta_s)}(F;h) = L_{1,\theta_s,F}(L'_{(\theta_1,\ldots,\theta_{s-1})}(F;h)) + L_{1,\theta_s,h}(L_{(\theta_1,\ldots,\theta_{s-1})}(F)).$$

Iterating this expression concludes the proof of the representation (6). □

Proof of Proposition 5.2. It is enough to prove the claim for $\bar{\theta} = (\theta_1, \theta_2)$, since the general case follows by iterating the same argument. Write

$$\mathbb{F}_{n,\bar{\theta}}(x) - F_{\bar{\theta}}(x) = \int \mathbb{F}_n\left(\frac{x-\theta_2 t}{\theta_1}\right) d\mathbb{F}_n(t) - \int F\left(\frac{x-\theta_2 t}{\theta_1}\right) dF(t) = A_n(x) + B_n(x), \quad (13)$$

where

$$A_n(x) := \int \mathbb{F}_n\left(\frac{x-\theta_2 t}{\theta_1}\right) - F\left(\frac{x-\theta_2 t}{\theta_1}\right) dF(t),$$
$$B_n(x) := \int \mathbb{F}_n\left(\frac{x-\theta_2 t}{\theta_1}\right) d(\mathbb{F}_n - F)(t).$$

We prove that $\sup_{x\in\mathbb{R}} |A_n(x)| \to_{a.s.} 0$ and $\sup_{x\in\mathbb{R}} |B_n(x)| \to_{a.s.} 0$.
By the triangle inequality and the Glivenko-Cantelli theorem,

$$\begin{aligned}\sup_{x\in\mathbb{R}} |A_n(x)| &\le \sup_{x\in\mathbb{R}} \int \left|\mathbb{F}_n\left(\frac{x-\theta_2 t}{\theta_1}\right) - F\left(\frac{x-\theta_2 t}{\theta_1}\right)\right| dF(t) \\ &\le \int \sup_{x\in\mathbb{R}} \left|\mathbb{F}_n\left(\frac{x-\theta_2 t}{\theta_1}\right) - F\left(\frac{x-\theta_2 t}{\theta_1}\right)\right| dF(t) \\ &= \sup_{y\in\mathbb{R}} |\mathbb{F}_n(y) - F(y)| \to_{a.s.} 0.\end{aligned}$$

Denote $F_k(x) = F(\frac{x}{k})$. Then write

$$\int \mathbb{F}_n\left(\frac{x-\theta_2 t}{\theta_1}\right) dF(t) = \int \mathbb{F}_n\left(\frac{x}{\theta_1} - u\right) dF_{\frac{\theta_2}{\theta_1}}(u) = \mathbb{F}_n * F_{\frac{\theta_2}{\theta_1}}\left(\frac{x}{\theta_1}\right).$$

Proceeding similarly for the remaining term in $B_n(x)$ and using commutativity of convolution, we obtain

$$B_n(x) = \int \mathbb{F}_{\frac{\theta_2}{\theta_1},n}\left(\frac{x}{\theta_1} - u\right) - F_{\frac{\theta_2}{\theta_1}}\left(\frac{x}{\theta_1} - u\right) d\mathbb{F}_n(u),$$

where $\mathbb{F}_{\frac{\theta_2}{\theta_1},n}$ denotes the empirical CDF of the sample $\frac{\theta_2}{\theta_1}X_1, ..., \frac{\theta_2}{\theta_1}X_n$. Since $\mathbb{F}_n$ is a CDF,

$$|B_n(x)| \le \sup_{u\in\mathbb{R}} \left|\mathbb{F}_{\frac{\theta_2}{\theta_1},n}\left(\frac{x}{\theta_1} - u\right) - F_{\frac{\theta_2}{\theta_1}}\left(\frac{x}{\theta_1} - u\right)\right| = \sup_{y\in\mathbb{R}} \left|\mathbb{F}_{\frac{\theta_2}{\theta_1},n}(y) - F_{\frac{\theta_2}{\theta_1}}(y)\right|.$$

Taking the supremum over $\mathbb{R}$ yields

$$\sup_{x\in\mathbb{R}} |B_n(x)| \le \sup_{y\in\mathbb{R}} \left|\mathbb{F}_{\frac{\theta_2}{\theta_1},n}(y) - F_{\frac{\theta_2}{\theta_1}}(y)\right| \to_{a.s.} 0$$

again by the Glivenko-Cantelli theorem.

We conclude that $\sup_{x\in\mathbb{R}} \left|\mathbb{F}_{n,\bar{\theta}}(x) - \mathbb{F}_{\bar{\theta}}(x)\right| \to_{a.s.} 0$, which proves Proposition 5.2 for $\bar{\theta} = (\theta_1, \theta_2)$. □

Proof of Theorem 5.4. The continuity of F implies that $\sqrt{n}\,(\mathbb{F}_n - F) \rightsquigarrow \mathbb{B}_F$ in $\ell^\infty(\mathbb{R})$. Then, the convergence of the empirical process is an immediate consequence of the functional delta method as $\sqrt{n}\,\left(\mathbb{F}_{n,\overline{\theta}} - F_{\overline{\theta}}\right) = \sqrt{n}\,\left(L_{\overline{\theta}}(\mathbb{F}_n) - L_{\overline{\theta}}(F)\right) \rightsquigarrow L'_{\overline{\theta}}(F;\mathbb{B}_F) =: \mathbb{H}^F_{\overline{\theta}}$.

To prove the representation of $\mathbb{H}^F_{\overline{\theta}}$, first, note that, using (6), we know that

$$L'_{\overline{\theta}}(F;\mathbb{B}_F) = \sum_{(a_1,\ldots,a_s)\in D} L_{1,\theta_s,a_s} \circ L_{1,\theta_{s-1},a_{s-1}} \circ \cdots \circ L_{1,\theta_3,a_3} \circ L_{\theta_1,\theta_2,a_1}, \tag{14}$$

where $D = \{(\mathbb{B}_F, F, \ldots, F),\, (F, \mathbb{B}_F, F, \ldots, F), \ldots, (F, \ldots, F, \mathbb{B}_F)\}$, as follows from Proposition 5.3. Consider some fixed $(a_1, \ldots, a_s) \in D$, that is, the set of vectors that have $\mathbb{B}_F$ in its j-th coordinate and F in the others. Then,

$$\begin{aligned} & L_{1,\theta_s,a_s} \circ L_{1,\theta_{s-1},a_{s-1}} \circ \cdots \circ L_{1,\theta_3,a_3} \circ L_{\theta_1,\theta_2,a_1}(x) \\ &= \int\cdots\int \mathbb{1}_{\{\sum_{i=1}^s \theta_i u_i \le x\}}\, da_1(u_1)\cdots da_s(u_s) \\ &= \int\cdots\int \mathbb{1}_{\{\sum_{i=1}^s \theta_i u_i \le x\}}\, dF(u_1)\cdots d\mathbb{B}_F(u_j)\cdots dF(u_s). \end{aligned}$$

Recalling that, according to (4), $F_{\overline{\theta}_{-j}}$ is the CDF of $\sum_{i\neq j}\theta_i X_i$, integrating with respect to all variables except the j-th one, and bearing in mind that the condition of the indicator is equivalent to $\sum_{i\neq j}\theta_i u_i \le x - \theta_j u_j$, we can write

$$\int\cdots\int \mathbb{1}_{\{\sum_{i=1}^s \theta_i u_i \le x\}} \prod_{i\neq j} dF(u_i) = F_{\overline{\theta}_{-j}}(x - \theta_j u_j).$$

Summing over the set C we obtain

$$L'_{\overline{\theta}}(F;\mathbb{B}_F) = \sum_{j=1}^s \int F_{\overline{\theta}_{-j}}(x - \theta_j u)\, d\mathbb{B}_F(u).$$

Now, note that

$$F_{\overline{\theta}_{-j}}(x - \theta_j u) = \int \mathbb{1}_{\{t \le x - \theta_j u\}}\, dF_{\overline{\theta}_{-j}}(t) = \int \mathbb{1}_{\{u \le \frac{x-t}{\theta_j}\}}\, dF_{\overline{\theta}_{-j}}(t),$$

so that, by Fubini's Theorem,

$$\begin{aligned} \sum_{j=1}^s \int F_{\overline{\theta}_{-j}}(x - \theta_j u)\, d\mathbb{B}_F(u) &= \sum_{j=1}^s \int \left(\int \mathbb{1}_{\{u \le \frac{x-t}{\theta_j}\}}\, dF_{\overline{\theta}_{-j}}(t) \right) d\mathbb{B}_F(u) \\ &= \sum_{j=1}^s \int \left(\int \mathbb{1}_{\{u \le \frac{x-t}{\theta_j}\}}\, d\mathbb{B}_F(u) \right) dF_{\overline{\theta}_{-j}}(t) \\ &= \sum_{j=1}^s \int \mathbb{B}_F\left(\tfrac{x-t}{\theta_j}\right) dF_{\overline{\theta}_{-j}}(t). \end{aligned}$$

The weak limit $\mathbb{H}^F_{\overline{\theta}}$ is a centred Gaussian process, since it is obtained by linear transformations of the Gaussian process $\mathbb{B}_F$.

The bilinearity of the covariance implies that

$$\begin{aligned}\operatorname{Cov}\left(\mathbb{H}^F_{\overline{\theta}}(x), \mathbb{H}^F_{\overline{\theta}}(y)\right) &= \sum_{j,k} \Omega_{j,k}(x,y) \\ &= \sum_{j,k} \iint \operatorname{Cov}\left(\mathbb{B}_F\left(\tfrac{x-t}{\theta_j}\right), \mathbb{B}_F\left(\tfrac{x-s}{\theta_k}\right)\right) dF_{\overline{\theta}_{-j}}(t) dF_{\overline{\theta}_{-k}}(s), \\ &= \sum_{j,k} \iint F\left(\tfrac{x-t}{\theta_j} \wedge \tfrac{x-s}{\theta_k}\right) - F\left(\tfrac{x-t}{\theta_j}\right) F\left(\tfrac{x-s}{\theta_k}\right) dF_{\overline{\theta}_{-j}}(t) dF_{\overline{\theta}_{-k}}(s).\end{aligned}$$

Using now the representation $F(z) = \int \mathbb{1}_{\{u \leq z\}} \, dF(u)$ for the integrands, and Fubini's Theorem, the conclusion follows. □

Proof of Theorem 6.1. For a general CDF F, the process $\mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}}$ is Gaussian, so it has trajectories that are almost surely continuous. Therefore, with probability 1 we have

$$\sup_{\{x:\, F_{\overline{\theta}}(x) < F_{\overline{\eta}}(x)\}} \left(\mathbb{H}^F_{\overline{\theta}}(x) - \mathbb{H}^F_{\overline{\eta}}(x)\right) = \sup_{\{x:\, F_{\overline{\theta}}(x) < F_{\overline{\eta}}(x) \text{ a.e.}\}} \left(\mathbb{H}^F_{\overline{\theta}}(x) - \mathbb{H}^F_{\overline{\eta}}(x)\right).$$

So, without loss of generality, we may assume that $\widetilde{\mathcal{H}}_0 = \{F : F_{\overline{\theta}}(x) < F_{\overline{\eta}}(x),\, x \in \mathbb{R}\}$. As in the proof of Proposition 5.3, the operator $L_{\overline{\theta},\overline{\eta}} = (L_{\overline{\theta}}, L_{\overline{\eta}}) : \ell^\infty(\mathbb{R}) \times \ell^\infty(\mathbb{R}) \to \ell^\infty(\mathbb{R}) \times \ell^\infty(\mathbb{R})$ is Hadamard differentiable at the point (F, F) on the direction of $(h_1, h_2) \in \ell^\infty(\mathbb{R}) \times \ell^\infty(\mathbb{R})$, with derivative given by $L'_{\overline{\theta},\overline{\eta}}((F,F);(h_1,h_2)) = (L'_{\overline{\theta}}(F;h_1), L'_{\overline{\eta}}(F;h_2))$. Then, the functional delta method gives

$$\sqrt{n}\left(\mathbb{F}_{n,\overline{\theta}} - F_{\overline{\theta}}, \mathbb{F}_{n,\overline{\eta}} - F_{\overline{\eta}}\right) \rightsquigarrow (\mathbb{H}^F_{\overline{\theta}}, \mathbb{H}^F_{\overline{\eta}})$$

in $\ell^\infty(\mathbb{R}) \times \ell^\infty(\mathbb{R})$. Since the difference operator is linear and continuous, it follows that

$$\sqrt{n}\,\left((\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}}) - (F_{\overline{\theta}} - F_{\overline{\eta}})\right) \rightsquigarrow \mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}}.$$

Define the map $\mathcal{M} : \ell^\infty(\mathbb{R}) \to \ell^\infty(\mathbb{R})$ by $\mathcal{M}(h) = h_+ = \max(0, h)$. Clearly, taking into account that we are referring to CDFs, $\sup(\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}}) = \sup \mathcal{M}(\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}})$. Now, when $F \in \mathcal{H}_0$, $\mathcal{M}(F_{\overline{\theta}} - F_{\overline{\eta}}) = 0$, so we have that $\mathcal{M}(\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}}) = \mathcal{M}(\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}}) - \mathcal{M}(F_{\overline{\theta}} - F_{\overline{\eta}})$. Following Fang and Santos (2019, p. 384), the map $\mathcal{M}$ is Hadamard directionally differentiable at $\delta = F_{\overline{\theta}} - F_{\overline{\eta}}$, with derivative

$$\mathcal{M}'(\delta; h)(x) = \begin{cases} h(x), & \text{if } \delta(x) > 0 \text{ a.e.}, \\ \max(0, h(x)), & \text{if } \delta(x) = 0 \text{ a.e.}, \\ 0, & \text{if } \delta(x) < 0 \text{ a.e.} \end{cases}$$

Therefore, we may apply the functional delta method to the mapping $\mathcal{M}$, to obtain

$$\sqrt{n}\,\left(\mathcal{M}(\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}}) - \mathcal{M}(F_{\overline{\theta}} - F_{\overline{\eta}})\right) \rightsquigarrow \mathcal{M}'(\delta; \mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}}),$$

in $\ell^\infty(\mathbb{R})$. Finally, $F \in \widetilde{\mathcal{H}}_0$ if and only if $\delta < 0$, so $\mathcal{M}'(\delta; \mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}}) = 0$. Accordingly, the continuous mapping theorem implies that $\sqrt{n} \sup \mathcal{M}(\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}}) \to_p 0$. Differently, for $F = \mathcal{C}$, $\delta(x) = 0$ for every $x \in \mathbb{R}$, hence, the Hadamard directional derivative coincides with $\mathcal{M}$, and so, we have that $\sqrt{n} \sup(\mathbb{F}_{n,\overline{\theta}} - \mathbb{F}_{n,\overline{\eta}}) \to_d \sup(\mathbb{H}^{\mathcal{C}}_{\overline{\theta}} - \mathbb{H}^{\mathcal{C}}_{\overline{\eta}})$, which is a nonnegative random variable. □

Proof of Proposition 6.2. The first assertion is obvious. Note that, as $\mathbb{H}^{\mathcal{C}}_{\overline{\theta}} - \mathbb{H}^{\mathcal{C}}_{\overline{\eta}}$ is a centred Gaussian process, $\sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(\mathbb{C}_n)$ converges in distribution to a nonnegative random variable with strictly positive and finite quantile, namely, $c_{\alpha,n} \to c_\alpha < \infty$. If $F \in \widetilde{\mathcal{H}}_0$, $\sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(\mathbb{F}_n) \to_p 0$, therefore the second assertion follows immediately. On the other hand, if $F \in \mathcal{H}_1$, $T_{\overline{\theta},\overline{\eta}}(\mathbb{F}_n) \to_p t_0 > 0$, therefore, $\sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(\mathbb{F}_n) \to_p +\infty$, proving the third assertion. □

Proof of Proposition 6.3. Assume $F \in \mathcal{H}_0$, then, as stated in (11), $\sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(\mathbb{F}_n)$ is bounded above by a sequence of random variables which converges in distribution to $\left\| \mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}} \right\|_\infty$. Since $\mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}}$ is a centred Gaussian process, then the $(1-\alpha)$-quantile of $\left\| \mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}} \right\|_\infty$, denoted as c_α, is finite, positive and unique. Therefore, under $\mathcal{H}_1$, $T_{\overline{\theta},\overline{\eta}}(\mathbb{F}_n,) \to_p t_0 > 0$, so, $\sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(\mathbb{F}_n) \to_p \infty$. The bootstrap process $\sqrt{n}\,(\mathbb{F}^b_n - \mathbb{F}_n)$ converges weakly almost surely to $\mathbb{B}_F$, conditionally on the data. Then, the functional delta method for the bootstrap implies that, for each weight vector $\overline{\theta}$, $\sqrt{n}\,\left(T_{\overline{\theta}}(\mathbb{F}^b_n) - T_{\overline{\theta}}(\mathbb{F}_n)\right)$ converges weakly in probability to $\mathbb{H}^F_{\overline{\theta}}$, conditionally on the data (see Theorem 2.9 in Kosorok (2008)). Therefore, we can conclude that $\mathcal{G}^b_n$ converges weakly in probability to $\mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}}$, conditionally on the data. Finally, by the continuous mapping theorem, $\left\|\mathcal{G}^b_n\right\|_\infty \to_d \left\| \mathbb{H}^F_{\overline{\theta}} - \mathbb{H}^F_{\overline{\eta}} \right\|_\infty$ conditionally on the data. The test rejects $\mathcal{H}_0$ if $\sqrt{n}\, T_{\overline{\theta},\overline{\eta}}(F_n) > c_{\alpha,n}$ given by (12), and we now have that $c_{\alpha,n} \to_p c_\alpha$. Therefore, under the null hypothesis $\mathcal{H}_0$, $\lim_{n\to\infty} P\left(T_{\overline{\theta},\overline{\eta}}(\mathbb{F}_n) > c_{\alpha,n}\right) \leq \alpha$. Differently, if $F \in \mathcal{H}_1$, $\lim_{n\to\infty} P\left(T_{\overline{\theta},\overline{\eta}}(\mathbb{F}_n) > c_{\alpha,n}\right) = 1$ because c_α is finite. □

References

Idir Arab, Tommaso Lando, and Paulo Eduardo Oliveira. Convex combinations of random variables stochastically dominate the parent for a new class of heavy tailed distributions. *Electron. Commun. Probab.*, 30:1–11, 2025. doi: 10.1214/25-ECP714.

Garry F Barrett and Stephen G Donald. Consistent tests for stochastic dominance. *Econometrica*, 71(1):71–104, 2003.

Garry F Barrett, Stephen G Donald, and Debopam Bhattacharya. Consistent nonparametric tests for lorenz dominance. *Journal of Business & Economic Statistics*, 32(1): 1–13, 2014.

Yuyu Chen and Seva Shneer. Risk aggregation and stochastic dominance for super heavy-tailed random variables. *ASTIN Bulletin*, Published online:1–14, 2025. doi: 10.1017/asb.2025.10053.

Yuyu Chen, Paul Embrechts, and Ruodu Wang. Technical note—an unexpected stochastic dominance: Pareto distributions, dependence, and diversification. *Oper. Res.*, 73:1336–1344, 2024. doi: 10.1287/opre.2022.0505.

Yuyu Chen, Taizhong Hu, Seva Shneer, and Zhenfeng Zou. Stochastic dominance for linear combinations of infinite-mean risks, 2025. URL https://arxiv.org/abs/2505.01739.

Zheng Fang and Andres Santos. Inference on directionally differentiable functions. *Rev. Econ. Stud.*, 86(1):377–412, 2019. doi: 10.1093/restud/rdy049.

Michael R Kosorok. *Introduction to empirical processes and semiparametric inference*. Springer, 2008.

Tommaso Lando and Mohammed Es-Salih Benjrada. A new class of tests for convex-ordered families based on expected order statistics. *Electron. J. Stat.*, 19(1):2780–2802, 2025. doi: 10.1214/25-EJS2398.

Tommaso Lando and Sirio Legramanti. A new class of nonparametric tests for second-order stochastic dominance based on the lorenz p–p plot. *Scandinavian Journal of Statistics*, 52(1):480–512, 2025.

Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: theory of majorization and its applications*. Springer Series in Statistics. Springer, New York, second edition, 2011. ISBN 978-0-387-40087-7. doi: 10.1007/978-0-387-68276-1. URL https://doi.org/10.1007/978-0-387-68276-1.

Alfred Müller. Some remarks on the effect of risk sharing and diversification for infinite mean risks. *ASTIN Bulletin*, Published online 2025:1–10, 2025. doi: 10.1017/asb.2025.10054.

Moshe Shaked and J. George Shantikumar. *Stochastic Orders*. Springer, New York, 2007.

Aad Van der Vaart and Jon Wellner. *Weak convergence and empirical processes: with applications to statistics*. Springer Science & Business Media, 1996.

Léonard Vincent. Diversification and stochastic dominance: When all eggs are better put in one basket. 2025. doi: 10.48550/arXiv.2507.16265. URL https://arxiv.org/abs/2507.16265.

Yoon-Jae Whang. *Econometric analysis of stochastic dominance: Concepts, methods, tools, and applications*. Cambridge University Press, 2019.