



Bootstrap-based tests for the total time on test and the excess wealth orders

Tommaso Lando ^{*}, Sirio Legramanti

Department of Economics, University of Bergamo, Italy

ARTICLE INFO

Keywords:

Expected shortfall
New better than used in expectation
Nonparametric test
Right spread
Stochastic order

ABSTRACT

Given a pair of non-negative random variables X and Y , we introduce a class of nonparametric tests for the null hypothesis that X dominates Y in the total time on test order. Critical values are determined using bootstrap-based inference, and the tests are shown to be consistent. The same approach is used to construct tests for the excess wealth order. As a byproduct, we also obtain a class of goodness-of-fit tests for the NBUE (New Better than Used in Expectation) family of distributions.

1. Introduction

The *total time on test* (TTT) and the *excess wealth* orders are related and established concepts in the field of stochastic orders (Shaked and Shantikumar, 2007), with applications in survival analysis, reliability, and economics, among others. However, despite their wide applicability, there is a lack of statistical tests for them. The few exceptions (Belzunce et al., 2001, 2005) consider the null hypothesis of equality in distribution versus the alternative of strict dominance. In this paper, we follow a complementary approach, which seems to be still unexplored for these orders, and consists in testing the null hypothesis of dominance versus the alternative of non-dominance.

We begin with some notations. Let H be a cumulative distribution function (CDF) defined on the nonnegative half-line, with finite mean μ_H and quantile function $H^{-1}(p) = \inf\{u : H(u) \geq p\}$. In this paper, we consider the following transforms:

$$T_H(p) = \int_0^{H^{-1}(p)} (1 - H(x))dx, \quad W_H(p) = \int_{H^{-1}(p)}^{\infty} (1 - H(x))dx = \mu_H - T_H(p).$$

T_H is referred to as the TTT transform of H , and is a primary tool in reliability and survival analysis (Marshall and Olkin, 2007). In particular, when H is the lifetime distribution of a population, $T_H(p)$ gives the expected lifetime of the items in the population which are smaller than the p th quantile. In fact, if $Z \sim H$, the TTT transform may be expressed as $T_H(p) = \mathbb{E}(\min(Z, H^{-1}(p)))$. On the other hand, W_H is known as the excess wealth transform, the *right spread* function, or *expected shortfall* at level p . In particular, when H is an income distribution, $W_H(p)$ can be interpreted as the expected additional wealth of the individuals whose wealth exceeds the p th quantile. More formally, $W_H(p) = \mathbb{E}(\max(Z, H^{-1}(p)))$. This transform is commonly employed in insurance mathematics (Dhaene et al., 2006), but it also has applications in reliability (Belzunce et al., 2001). Note that, unlike T_H , the function W_H can be defined even without the nonnegativity assumption. However, here we focus on the nonnegative case to simplify the exposition.

Given a pair of non-negative random variables, X and Y , with CDFs F and G , respectively, TTT and excess wealth transforms may be used to define two complementary stochastic orders, recalled in the following definition.

^{*} Corresponding author.

E-mail address: tommaso.lando@unibg.it (T. Lando).

Definition 1.

1. X dominates Y in the TTT order, abbreviated as $X \succeq_{\text{TTT}} Y$, if $T_F(p) \geq T_G(p)$, $\forall p \in [0, 1]$.
2. X dominates Y in the excess wealth order, abbreviated as $X \succeq_{\text{ew}} Y$, if $W_F(p) \geq W_G(p)$, $\forall p \in [0, 1]$.

The TTT order has been studied, for instance, by [Kochar et al. \(2002\)](#), [Li and Shaked \(2007\)](#), [Belzunce et al. \(2014\)](#) and [Belzunce and Martínez-Riquelme \(2015\)](#). It compares random variables in terms of size and dispersion. The TTT order is weaker than the usual stochastic order, but it implies the well-known *increasing concave order*, also known as *second-order stochastic dominance* ([Shaked and Shantikumar, 2007](#)). This means that $X \succeq_{\text{TTT}} Y$ implies $\mathbb{E}u(X) \geq \mathbb{E}u(Y)$ for every increasing concave function u . One of the main applications of the TTT order is comparing lifetime distributions. In particular, if X and Y represent the random lifetime in two populations, $X \succeq_{\text{TTT}} Y$ means that, for every proportion $p \in [0, 1]$, the expected lifetime of the items in X , up to its p th quantile $F^{-1}(p)$, is greater than or equal to the expected lifetime in Y , up to the corresponding quantile $G^{-1}(p)$.

On the other hand, the excess wealth order is commonly applied in insurance mathematics, where the main focus is comparing risks ([Denuit and Vermandele, 1999](#); [Sordo, 2009](#); [Barmalzan and Najafabadi, 2015](#)). Moreover, this order, also referred to as the *right spread order* ([Fernandez-Ponce et al., 1998](#)), also has a natural application in survival analysis and in reliability theory, as it compares expected residual lifetime beyond quantile thresholds; see [Shaked and Shantikumar \(1998\)](#), [Belzunce et al. \(2001\)](#) and [Kochar et al. \(2007\)](#). For example, consider two systems in which items, with random lifetime X and Y , respectively, are tested until the 100% of them fail. One may be interested in comparing the expected lifetime of the remaining items, which boils down to comparing $W_F(p)$ and $W_G(p)$. In this case, the first system has a larger expected residual lifetime, for every p , if $X \succeq_{\text{ew}} Y$. The properties of the excess wealth order are complementary to those of the TTT order: it is implied by the usual stochastic order and it implies the *increasing convex order* ([Shaked and Shantikumar, 2007](#)). The excess wealth order is also closely related to the *location independent risk order*, introduced by [Jewitt \(1989\)](#) and studied in the economics literature.

In spite of the wide literature about these orders, there are few statistical methods available to test whether $X \succeq_{\text{TTT}} Y$ or $X \succeq_{\text{ew}} Y$ hold. [Belzunce et al. \(2001, 2005\)](#) proposed tests for the null hypothesis of equality in distribution versus the alternatives of strict dominance (excluding the case of equal distributions). These tests have the disadvantage that we ignore how they behave when $X \not\succeq_{\text{TTT}} Y$. Hence, in this paper we follow a complementary approach, namely, we test $H_0^{\text{TTT}} : X \succeq_{\text{TTT}} Y$ versus the alternative $H_1^{\text{TTT}} : X \not\succeq_{\text{TTT}} Y$. In this case, it is difficult to determine the distribution of the test statistic under H_0^{TTT} , as this contains infinitely many pairs of distributions. Using a similar approach as in [Barrett et al. \(2014\)](#) (see also [Barrett and Donald, 2003](#)), we propose a family of conservative tests for the TTT order, in which critical values are approximated via bootstrap procedures. Given the relation between the TTT and the excess wealth transforms, the same approach may be used to test $H_0^{\text{ew}} : X \succeq_{\text{ew}} Y$ versus $H_1^{\text{ew}} : X \not\succeq_{\text{ew}} Y$. Consistency of these tests is established under relatively mild distributional assumptions, which allow for both dependent and independent sampling schemes.

Finally, note that the TTT order can be used to define the *new better than used in expectation* (NBUE) family of distributions, which also has applications in survival analysis and reliability. For example, it can be employed in age replacement policies ([Barlow and Proschan, 1975](#)), applied to shock models ([Block and Savits, 1978](#)), or leveraged to derive bounds for the survival function ([Cheng and Lam, 2002](#); [Brown, 2013](#)). The NBUE condition expresses the notion of positive ageing – meaning that ageing has an adverse effect on lifetime – in a weaker sense compared to other popular properties, such as the *increasing hazard rate*, the *increasing hazard rate on average* and the *decreasing mean residual life* ([Marshall and Olkin, 2007](#)). In particular, a lifetime random variable X is said to be NBUE if $X/\mu_F \succeq_{\text{TTT}} E$, where E is a unit exponential random variable. By exploiting this relation, the approach considered in this paper can also be used to define a class of bootstrap-based tests for the null hypothesis $H_0^{\text{NBUE}} : “X \text{ is NBUE}”$ versus the alternative $H_1^{\text{NBUE}} : “X \text{ is not NBUE}”$. This class of tests may be used as a complement to other popular tests (see for instance [Hollander and Proschan, 1975](#); [Anis and Mitra, 2011](#); [Anis and Basu, 2014](#)), which consider the null hypothesis of exponentiality, namely, $H_0^E : X/\mu_F =_d E$, versus the alternative that F is strictly NBUE, namely, $X/\mu_F \succ_{\text{TTT}} E$. Such tests are not exhaustive, in other words, they are not supposed to work under H_1^{NBUE} . However, if for a given dataset H_0^E is rejected whereas H_0^{NBUE} is not, one may conclude that the underlying distribution is strictly NBUE. Recently, another class of tests for H_0^{NBUE} versus H_1^{NBUE} has been proposed by [Lando \(2024\)](#), with a different approach which does not rely on bootstrap procedures.

The rest of this paper is organised as follows. In Section 2, we introduce a class of test statistics that measure departures from the null hypothesis H_0^{TTT} . In Section 3, we establish the consistency properties of our class of tests. In Section 4, we obtain tests for the NBUE null hypothesis using the relation between the TTT order and the NBUE class, while in Section 5 we extend our approach to the excess wealth order. Finite sample properties of our tests are investigated by simulation studies in Section 6. Finally, Section 7 illustrates the usefulness of the proposed tests through an application. All proofs are reported in [Appendix](#).

2. Construction of the tests

Throughout this paper, let $C[0, 1]$ denote the space of bounded continuous functions over $[0, 1]$. Given some $v \in C[0, 1]$ and $p \geq 1$, $\|v\|_p$ denotes the L^p norm of v , that is, for p finite, $\|v\|_p = (\int_0^1 |v(t)|^p dt)^{1/p}$, while $\|v\|_\infty = \sup_{t \in [0, 1]} |v(t)|$ is the uniform norm. The symbol $\rightharpoonup$ denotes weak convergence as in [Van der Vaart and Wellner \(1996\)](#), while $\rightarrow_p$ represents convergence in probability.

By [Definition 1](#), H_0^{TTT} is false if and only if the difference $\delta(u) = T_G(p) - T_F(p)$ is positive at some point $p \in [0, 1]$. Therefore, we can measure the positive part of δ , namely, $\delta_+ = \max(0, \delta)$ with the following family of functionals

$$\Psi_r(\delta) = \|\delta_+\|_r,$$

for $r \geq 1$, including $r = \infty$. In this paper, we focus just on the extreme cases $r = 1$ and $r = \infty$. However, the properties of the tests hold for every choice of $r \geq 1$, owing to Proposition 1 in [Lando and Legramanti \(2025\)](#).

Clearly, $\mathcal{H}_0^{\text{tr}}$ is false if and only if $\Psi_r(\delta) > 0$. Therefore, we reject the null hypothesis if an empirical version of $\Psi_r(\delta)$, namely $\Psi_r(\hat{\delta})$, is significantly large. We will now discuss how to estimate δ .

2.1. Estimation

Let $\mathcal{X} = \{X_1, \dots, X_n\}$ be an i.i.d. random sample from F , and $\mathcal{Y} = \{Y_1, \dots, Y_m\}$ an i.i.d. random sample from G . In this paper, we deal with two different sampling schemes: independent sampling and matched pairs. Under independent sampling, the samples $\mathcal{X}$ and $\mathcal{Y}$ are independent of each other and the sample sizes n and m may differ. Differently, in the matched-pairs scheme, $n = m$ and we have n i.i.d. pairs $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$ drawn from a bivariate distribution with marginal CDFs F and G , respectively. For our asymptotic analysis, we also assume that, as $n \rightarrow \infty$, also $m \rightarrow \infty$, and that $\lim_{n,m \rightarrow \infty} nm/(n+m) = \infty$, and $\lim_{n,m \rightarrow \infty} n/(n+m) = \lambda \in [0, 1]$.

The random samples $\mathcal{X}$ and $\mathcal{Y}$ yield the empirical CDFs

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x), \quad G_m(x) = \frac{1}{m} \sum_{j=1}^m \mathbb{1}(Y_j \leq x).$$

We will denote the order statistics of rank k from $\mathcal{X}$ and $\mathcal{Y}$ with $X_{(k)}$ and $Y_{(k)}$, and the sample means by $\bar{X}_n$ and $\bar{Y}_m$, respectively. The empirical TTT transform of F , denoted by T_{F_n} , is defined as the piecewise linear function which interpolates the jump points of the step function $\int_0^{F_n^{-1}(p)} (1 - F_n(x)) dx$. In particular, for $k = 0, \dots, n$, we have that

$$T_{F_n}\left(\frac{k}{n}\right) = \int_0^{X_{(k)}} (1 - F_n(t)) dt = \frac{1}{n} \sum_{i=1}^k (n - i + 1)(X_{(i)} - X_{(i-1)}) = \frac{n-k}{n} X_{(k)} + \frac{1}{n} \sum_{i=1}^k X_{(i)},$$

where $X_{(0)} := 0$. The sample mean is obtained as $T_{F_n}(1) = 1/n \sum_{i=1}^n X_{(i)} = \bar{X}_n$. Note that, when testing the lifetime of n items, $nT_{F_n}\left(\frac{k}{n}\right)$ gives the total time on test up to the k th observation. T_{G_m} is defined similarly. Now, an empirical version of δ is just

$$\hat{\delta} = T_{G_m} - T_{F_n}.$$

It can be shown that, as n and m tend to infinity, T_{F_n} and T_{G_m} converge strongly and uniformly to T_F and T_G , respectively, in $[0, 1]$ ([Barlow and Van Zwet, 1969](#), Theorem 2.1). This implies that $\hat{\delta}$ also converges strongly and uniformly to δ in $[0, 1]$, under our asymptotic regime.

3. Consistency

3.1. Asymptotic properties of the test statistic

Let us define the empirical TTT processes associated with F and G as

$$t_n^F = \sqrt{n}(T_{F_n} - T_F), \quad t_m^G = \sqrt{m}(T_{G_m} - T_G),$$

respectively. Note that such processes have trajectories in $C[0, 1]$. Correspondingly, consider the process

$$\Delta_{n,m} = \sqrt{r_{n,m}}(\hat{\delta} - \delta) = \sqrt{\frac{n}{n+m}} t_m^G - \sqrt{\frac{m}{n+m}} t_n^F,$$

where $r_{n,m} = nm/(n+m)$. Recall that, under our assumptions, $r_{n,m} \rightarrow \infty$ as n and m diverge. To establish the asymptotic properties of our class of tests, we determine the limit behaviour of $\Delta_{n,m}$ via the functional delta method. This requires the derivation of the Hadamard derivative of the map $H \rightarrow T_H$ with respect to the L^1 norm, which can be achieved under the following assumption.

Assumption 1. Both F and G are continuously differentiable with strictly positive densities f and g , respectively, and have a finite moment of order $2 + \epsilon$ for some $\epsilon > 0$. Moreover, $F(0) = G(0) = 0$.

Let B be a centred Gaussian element of $C[0, 1] \times C[0, 1]$, with covariance function

$$\text{Cov}(B(x_1, y_1), B(x_2, y_2)) = C(x_1 \wedge x_2, y_1 \wedge y_2) - C(x_1, y_1)C(x_2, y_2).$$

Under the independent-sampling scheme, $C(x_1, y_1) = x_1 y_1$ is the product copula, whereas, under the matched-pairs scheme, C is the copula associated with the pair (X_i, Y_i) , $i = 1, \dots, n$. Accordingly, $B_1(x_1) = B(x_1, 1)$ and $B_2(x_2) = B(1, x_2)$ are independent Brownian bridge processes under independent-sampling, but they may be dependent under paired sampling.

Theorem 1. Under [Assumption 1](#), for the sampling schemes discussed in [Section 2.1](#) we have

$$\sqrt{r_{n,m}}(\hat{\delta} - \delta) \rightsquigarrow \tilde{\mathcal{T}} = \sqrt{\lambda} T_G - \sqrt{1 - \lambda} T_F \quad \text{in } C[0, 1],$$

where

$$T_F(p) = \frac{(1-p)B_1(p)}{f \circ F^{-1}(p)} + \int_0^p B_1(y) dF^{-1}(y), \quad T_G(p) = \frac{(1-p)B_2(p)}{g \circ G^{-1}(p)} + \int_0^p B_2(y) dG^{-1}(y).$$

We measure departures from the null hypothesis with the test statistic $\hat{\Psi}_r = \sqrt{r_{n,m}}\Psi_r(\hat{\delta})$. In particular, “large” values of $\hat{\Psi}_r$ lead to rejection of $\mathcal{H}_0^{\text{itt}}$. In fact, under $\mathcal{H}_0^{\text{itt}} : \delta \leq 0$, $\sqrt{r_{n,m}}\Psi_r(\hat{\delta}) \leq \sqrt{r_{n,m}}\Psi_r(\hat{\delta} - \delta)$ by properties of Ψ_r . Moreover, it can be shown that Ψ_r satisfies some continuity properties (Lando and Legramanti, 2025, Proposition 1), hence, the continuous mapping theorem implies that $\sqrt{r_{n,m}}\Psi_r(\hat{\delta} - \delta) \rightsquigarrow \Psi_r(\tilde{\mathcal{T}})$. In other words, the test statistic $\hat{\Psi}_r$ is dominated under the null hypothesis by a statistic that has the same asymptotic behaviour as $\Psi_r(\tilde{\mathcal{T}})$. The same approach was used by Barrett et al. (2014) in testing Lorenz dominance.

Now, a critical value c_α is determined as the $(1 - \alpha)$ quantile of $\Psi_r(\tilde{\mathcal{T}})$, that is, $P(\Psi_r(\tilde{\mathcal{T}}) > c_\alpha) = \alpha$. For any $\alpha < 1/2$, such a quantile is positive, finite, and unique because $\tilde{\mathcal{T}}$ is a mean-zero Gaussian process. Since, under $\mathcal{H}_1^{\text{itt}} : X \not\leq_{\text{itt}} Y$, the statistic $\hat{\Psi}_r/\sqrt{r_{n,m}} = \Psi_r(\hat{\delta})$ converges in probability to the strictly positive number $\Psi_r(\delta)$, then it is clear that $\hat{\Psi}_r$ will diverge and exceed the threshold value c_α . These properties are summarised in the following lemma. The proof is omitted as it can be obtained using similar arguments as in Lemma 4 of Barrett et al. (2014).

Lemma 1.

1. Under $\mathcal{H}_0^{\text{itt}}$, $\sqrt{r_{n,m}}\Psi_r(\hat{\delta}) \leq \sqrt{r_{n,m}}\Psi_r(\hat{\delta} - \delta) \rightsquigarrow \Psi_r(\tilde{\mathcal{T}})$. Moreover, for any $\alpha < 1/2$, the $(1 - \alpha)$ quantile of the distribution of $\Psi_r(\tilde{\mathcal{T}})$ is positive, finite, and unique.
2. Under $\mathcal{H}_1^{\text{itt}}$, $\sqrt{r_{n,m}}\hat{\Psi}_r \rightarrow_p \infty$.

Now, critical values may be approximated by simulating the distribution of $\Psi_r(\tilde{\mathcal{T}})$ via bootstrap procedures.

3.2. Bootstrap-based decision rule

Let us denote the bootstrap estimators of the empirical CDFs F_n and G_m , respectively, as F_n^* and G_m^* . These may be written as

$$F_n^*(x) = \frac{1}{n} \sum_{i=1}^n M_i^{(1)} \mathbb{1}(x \leq X_i), \quad G_m^*(x) = \frac{1}{m} \sum_{i=1}^m M_i^{(2)} \mathbb{1}(x \leq Y_i),$$

where $M^{(1)} = (M_1^{(1)}, \dots, M_n^{(1)})$ and $M^{(2)} = (M_1^{(2)}, \dots, M_m^{(2)})$ are independent of the data and are drawn from a multinomial distribution according to the chosen sampling scheme. In particular, under the independent-sampling scheme, $M^{(1)}$ and $M^{(2)}$ are independently drawn from multinomial distributions with uniform probabilities over n and m trials, respectively. Under the matched-pairs scheme, we have $M^{(1)} = M^{(2)}$, drawn from the multinomial distribution with uniform probabilities over $n = m$ trials, which means that we sample (with replacement) pairs of data, from the n pairs $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$. Correspondingly, the bootstrap estimators of T_{F_n} and T_{G_m} are given, respectively, by $T_{F_n^*}$ and $T_{G_m^*}$. Accordingly, we have $\hat{\delta}^* = T_{G_n^*} - T_{F_n^*}$. By applying the functional delta method for the bootstrap (see 3.9.3 of Van der Vaart and Wellner (1996)), one may show that $\sqrt{r_{n,m}}\Psi_r(\hat{\delta}^* - \hat{\delta})$ has the same limiting distribution as $\sqrt{r_{n,m}}\Psi_r(\tilde{\mathcal{T}})$. This allows the approximation of the p -values $p = P\{\sqrt{r_{n,m}}\Psi_r(\hat{\delta}^* - \hat{\delta}) > \sqrt{r_{n,m}}\Psi_r(\hat{\delta})\}$, using K bootstrap replications:

$$p \approx \frac{1}{K} \sum_{k=1}^K \mathbb{1}(\sqrt{r_{n,m}}\Psi_r(\hat{\delta}_k^* - \hat{\delta}) > \sqrt{r_{n,m}}\Psi_r(\hat{\delta})),$$

where $\hat{\delta}_k^*$ is the k th resampled realisation of $\hat{\delta}^*$. We reject $\mathcal{H}_0^{\text{itt}}$ whenever $p < \alpha$.

As proved in the following proposition, the bootstrap approximation of the critical values is consistent. This allows us to exploit the results of Lemma 1 and establish the asymptotic behaviour of our class of tests.

Proposition 1. Under Assumption 1 and the sampling schemes in Section 2.1,

1. If $\mathcal{H}_0^{\text{itt}}$ is true, $\lim_{n \rightarrow \infty} P\{\text{reject } \mathcal{H}_0^{\text{itt}}\} \leq \alpha$;
2. If $\mathcal{H}_0^{\text{itt}}$ is false, $\lim_{n \rightarrow \infty} P\{\text{reject } \mathcal{H}_0^{\text{itt}}\} = 1$.

The results obtained in this section may be extended to the case of randomly right-censored data, which is common in life testing problems. In particular, the proposed approach still works if we replace the empirical CDF with the Kaplan–Meier estimator (Kaplan and Meier, 1958), tailored to deal with this kind of data. The weak convergence of the Kaplan–Meier process has been derived by Gill (1983). More recently, Dobler (2019) proved that the bootstrap version of such a process converges weakly to the same limit. To establish the asymptotic properties of the tests in this case, we need to assume that the CDFs under analysis have compact supports, with strictly positive and continuous densities. This allows us to leverage Lemma 3.10.24 in Van der Vaart and Wellner (1996), obtain the Hadamard differentiability of the inverse map with respect to the uniform norm, and derive that of the TTT transform as in Theorem 1. Finally, the properties of the tests follow again by the functional delta method.

4. Goodness-of-fit tests for the NBUE family

The classic way of expressing the NBUE condition is requiring that, for every $t > 0$, $\mathbb{E}(X - t | X > t) \leq \mathbb{E}(X)$. This means that an item of any age t has a mean residual life smaller than or equal to the expected lifetime of a new item in the same population. Generally, ageing properties can also be defined in terms of a suitable stochastic order between the random variable of interest X and the unit exponential E , which indeed expresses the concept of “no ageing” thanks to its memoryless property. In this regard, as

discussed in Section 1, X is NBUE if $X/\mu_F \geq_{\text{itt}} E$. This stochastic inequality boils down to $S_F(p) \geq p, \forall p \in [0, 1]$, where $S_F = T_F/\mu_F$ is the scaled TTT transform of F , while it is easy to see that the TTT transform of E is just the identity function.

We now introduce a family of tests for $\mathcal{H}_0^{NBUE} : X/\mu_F \geq_{\text{itt}} E$ versus $\mathcal{H}_1^{NBUE} : X/\mu_F \not\geq_{\text{itt}} E$ based on a random sample of size n from F . Goodness-of-fit tests for ageing properties of this type are especially relevant in life testing problems and in shape-constrained statistical inference. There are many examples of tests of this type for the increasing hazard rate family, including the one of [Tenga and Santner \(1984\)](#), also based on the TTT transform. Our goodness-of-fit approach to NBUE is complementary to that of the articles cited in Section 1, which consists in testing the null hypothesis of exponentiality versus strictly NBUE alternatives. However, in this setting, computing critical values requires either a least favourable distribution, as in [Lando \(2024\)](#), or resorting to bootstrap, as in the present paper.

Let now $S_{F_n} = T_{F_n}/\bar{X}$ be the empirical version of S_F . We can measure deviations from the null hypothesis by measuring the positive part of the function $\hat{\delta}(p) := p - S_{F_n}(p)$, which is the empirical counterpart of $\delta(p) = p - S_F(p)$. Under $\mathcal{H}_0^{NBUE}$, $p - S_F(p) \leq 0, \forall p \in [0, 1]$. Therefore $\hat{\delta}(p) \leq S_F(p) - S_{F_n}(p), \forall p \in [0, 1]$. The following result, based on [Theorem 1](#), establishes the weak convergence of the process $s_n^F = \sqrt{n}(S_{F_n} - S_F)$.

Corollary 1. Under [Assumption 1](#), for $n \rightarrow \infty$, $s_n^F \rightsquigarrow S_F$ in $C[0, 1]$, where

$$S_F(p) = \frac{T_F(p)}{\mu_F} - \frac{T_F(p)T_F(1)}{\mu_F^2}.$$

Since S_F is a centred Gaussian process, we can establish that, under $\mathcal{H}_0^{NBUE}$, $\sqrt{n}\Psi_r(\hat{\delta}) \leq \sqrt{n}\Psi_r(S_{F_n} - S_F) \rightsquigarrow \Psi_r(S_F)$, where, for any $\alpha < 1/2$, the $(1 - \alpha)$ quantile of $\Psi_r(S_F)$ is positive, finite, and unique. On the other hand, under $\mathcal{H}_1^{NBUE}$, the test statistic $\sqrt{n}\Psi_r(\hat{\delta})$ diverges in probability. Now, we may approximate the distribution of $\sqrt{n}\Psi_r(S_{F_n} - S_F)$ using the same bootstrap procedure discussed earlier. In particular, let $S_{F_n}^*$ be the bootstrap version of S_{F_n} . The p -values may be approximated as $p \approx 1/K \sum_{k=1}^K \mathbb{1}(\sqrt{r_{n,m}}\Psi_r((S_{F_n} - S_{F_n}^*)_k) > \sqrt{r_{n,m}}\Psi_r(\hat{\delta}))$, where $(S_{F_n} - S_{F_n}^*)_k$ is the k th realisation of the simulated process $S_{F_n} - S_{F_n}^*$. Accordingly, we may establish the following asymptotic properties.

Proposition 2. Under [Assumption 1](#),

1. If $\mathcal{H}_0^{NBUE}$ is true, $\lim_{n \rightarrow \infty} P\{\text{reject } \mathcal{H}_0^{NBUE}\} \leq \alpha$;
2. If $\mathcal{H}_0^{NBUE}$ is false, $\lim_{n \rightarrow \infty} P\{\text{reject } \mathcal{H}_0^{NBUE}\} = 1$.

5. Tests for the excess wealth order

Similarly to the case of the TTT order, $\mathcal{H}_0^{ew}$ is false if and only if the difference $\delta(p) = W_G(p) - W_F(p)$ is positive at some point $p \in [0, 1]$. Letting $W_{F_n}(p) = \bar{X}_n - T_{F_n}$ and $W_{G_m}(p) = \bar{Y}_m - T_{G_m}$, we reject the null hypothesis if $\Psi_r(W_{G_m} - W_{F_n})$ is large enough. As n and m tend to infinity, $\hat{\delta} := W_{G_m} - W_{F_n}$ converges strongly and uniformly to δ in $[0, 1]$. The empirical excess wealth processes associated with F and G may be defined as

$$w_n^F = \sqrt{n}(W_{F_n} - W_F), \quad w_m^G = \sqrt{m}(W_{G_m} - W_G),$$

respectively. These processes yield

$$\Delta_{n,m} := \sqrt{r_{n,m}}(\hat{\delta} - \delta) = \sqrt{\frac{n}{n+m}}w_m^G - \sqrt{\frac{m}{n+m}}w_n^F.$$

Weak convergence of $\Delta_{n,m}$ is established in the following corollary, using the same arguments of [Theorem 1](#) (the proof is omitted).

Corollary 2. Under [Assumption 1](#), for the sampling schemes discussed in Section 2.1 we have

$$\sqrt{r_{n,m}}(\hat{\delta} - \delta) \rightsquigarrow \tilde{W} = \sqrt{\lambda}\mathcal{W}_G - \sqrt{1 - \lambda}\mathcal{W}_F \quad \text{in } C[0, 1],$$

where

$$\mathcal{W}_F(p) = -\frac{(1-p)B_1(p)}{f \circ F^{-1}(p)} + \int_p^1 \frac{B_1(y)}{f \circ F^{-1}(y)} dy, \quad \mathcal{W}_G(p) = -\frac{(1-p)B_2(p)}{g \circ G^{-1}(p)} + \int_p^1 \frac{B_2(y)}{g \circ G^{-1}(y)} dy.$$

Using the same arguments of Section 3, we may establish an equivalent result to [Lemma 1](#). Denoting the bootstrap estimators of W_{F_n} and W_{G_m} as $W_{F_n}^*$ and $W_{G_m}^*$, respectively, and letting $\hat{\delta}^* = W_{G_m}^* - W_{F_n}^*$, the functional delta method for the bootstrap implies that $\sqrt{r_{n,m}}\Psi_r(\hat{\delta}^* - \hat{\delta})$ has the same limiting distribution of $\sqrt{r_{n,m}}\Psi_r(\tilde{W})$. This allows the approximation of p -values using bootstrap replications as in Section 3.2. Accordingly, by the same arguments used in the proof of [Proposition 1](#), we show consistency of the corresponding bootstrap-based tests.

Proposition 3. Under [Assumption 1](#) and the sampling schemes in Section 3.1,

1. If $\mathcal{H}_0^{ew}$ is true, $\lim_{n \rightarrow \infty} P\{\text{reject } \mathcal{H}_0^{ew}\} \leq \alpha$;
2. If $\mathcal{H}_0^{ew}$ is false, $\lim_{n \rightarrow \infty} P\{\text{reject } \mathcal{H}_0^{ew}\} = 1$.

6. Simulations

In this section, we investigate the finite sample properties of our tests through some numerical analyses. In particular, we focus on tests for $\mathcal{H}_0^{III}$ and $\mathcal{H}_0^{NBUE}$. We do not focus on tests for the excess wealth order, because their behaviour is similar to that of our tests for the TTT order. We consider the Weibull and the Singh–Maddala distributions. Namely, we denote with $W(a, b)$ the Weibull distribution with shape parameter $a > 0$, scale parameter $b > 0$, and CDF $F_W(x; a, b) = 1 - \exp(-(x/b)^a)$. Moreover, we denote with $SM(a, b, c)$ the Singh–Maddala distribution with shape parameters $a > 0, b > 0$, scale parameter $c > 0$, and CDF $F_{SM}(x; a, b, c) = 1 - ((x/c)^b + 1)^{-a}$. The TTT transform of $W(a, b)$ is

$$T_{FW}(p) = \frac{b}{a} \left(\Gamma\left(\frac{1}{a}, 0\right) - \Gamma\left(\frac{1}{a}, -\log(1-p)\right) \right),$$

where Γ is the gamma function, while that of $SM(a, b, c)$ exists just for $ab > 1$ and has the following expression

$$T_{FSM}(p) = c \left((1-p)^{-1/a} - 1 \right)^{1/b} {}_2F_1\left(a, \frac{1}{b}; 1 + \frac{1}{b}; 1 - (1-x)^{-1/a}\right),$$

where ${}_2F_1$ is the hypergeometric function.

We simulate different scenarios from the distributions above, under the independent sampling framework, for $n = 50, 100, 200, 500$, to investigate finite sample properties of the tests based on Ψ_∞ and Ψ_1 . For each scenario, we run 500 simulation experiments and sample 500 bootstrap replicates. The significance level of the test is set to $\alpha = 0.1$. To simplify computations, we use the step function $\tilde{T}_{F_n}(p) = \int_0^{F_n^{-1}(p)} (1 - F_n(x)) dx$ instead of T_{F_n} , since the difference among them is asymptotically negligible (recall that T_{F_n} interpolates the jumps points of $\tilde{T}_{F_n}$).

6.1. Tests for $\mathcal{H}_0^{III}$

We first let $X \sim W(a, 1)$ and $Y \sim W(2, 1)$. In this case, we have $X \geq_{III} Y$ when $a \geq 2$, and the dominance becomes more evident for larger values of a . Both tests in these cases provide rejection rates that are always smaller than the nominal level α and get closer to 0 when the sample size grows. For this reason, we do not report these results here. On the other hand, letting $X \sim W(2, 1)$ and $Y \sim W(a, 1)$ we clearly have $X \not\geq_{III} Y$ for $a > 1$. In this case, it is not difficult to detect deviations from the null hypothesis since T_G is always above T_F . This is confirmed by the rejection rates in Fig. 1. In this case, Ψ_1 tends to deliver larger power compared to Ψ_∞ . However, it may be harder to reject $\mathcal{H}_0^{III}$ when T_G and T_F cross. We consider the case in which $X \sim W(a, 1)$ and $Y \sim W(1, 1.2)$, for $a \geq 1$, and also the reverse case, namely, $X \sim W(1, 1.2)$ and $Y \sim W(a, 1)$. The latter case is the most critical to detect, in particular for smaller values of a . Our results show that Ψ_∞ often provides larger power compared to Ψ_1 in these scenarios. For $X \sim W(1, 1.2)$ and $Y \sim W(1.25, 1)$ both tests (especially Ψ_1) struggle to reject the null hypothesis, at least for the sample sizes considered (we obtain larger power increasing n to 2000). However, even in this critical case, the p -values tend to decrease as the sample size grows, coherently with the consistency property in Proposition 1. Overall, our results suggest that Ψ_∞ may be more reliable than Ψ_1 , especially when the TTT transforms cross.

6.2. Tests for $\mathcal{H}_0^{NBUE}$

First, recall that ageing properties are scale-independent, so there is no loss of generality in setting scale parameters equal to 1. The Weibull distribution $W(a, b)$ is NBUE for $a \geq 1$, while, for $a < 1$, it belongs to the *new worse than used in expectation* (NWUE) class, which is the family of distributions characterised by the reverse stochastic inequality, $E \geq_{III} X/\mu_F$. Letting $X \sim W(a, 1)$, where a varies within $[0.8, 1.1]$, we may observe the performance of the tests under the null and the alternative hypotheses. The results, reported in Fig. 2, show that Ψ_1 provides larger power under the alternative compared to Ψ_∞ . The latter seems to be more conservative, as it appears also by observing the behaviour of the tests under the null hypothesis. For both tests, the rejection rates tend to 1 under the alternative and tend to 0 under the null, confirming the asymptotic consistency properties. Moreover, the empirical power is always above the nominal level 0.1 when $a < 1$, showing an unbiased behaviour against NBUE alternatives.

We also generate samples from the $SM(a, b, c)$ distribution, which is neither NBUE nor NWUE for $a \in [1, 2]$ and $b = 1.5$. Actually, in this case, the scaled TTT transform S_{FSM} crosses the identity function, which is the TTT transform of the unit exponential, from above. In particular, it is more difficult to detect non-dominance for larger values of a . The simulation results show that, in this more critical case, Ψ_∞ yields larger rejection rates compared to Ψ_1 , which seems to be less sensitive in detecting violations of the NBUE property especially for $a = 2$.

7. An example of application to survival data

As an illustrative example, we apply our tests to the dataset in Hoel (1972), which reports the survival times of some mice which had received a radiation dose at an age of 5–6 weeks. The mice affected by the two main diseases, namely, thymic lymphoma (51 mice) and reticulum cell sarcoma (53 mice), were separated from the others (77 mice).

According to biologists (Hoel, 1972), these two diseases are lethal and they are independent of one another and of other causes of death. One may be interested in testing whether the lifetime of the mice affected by these major diseases is stochastically smaller or larger (in the usual sense) than that of the other mice. In particular, tests for the usual stochastic order (Barrett and Donald, 2003) suggest that thymic lymphoma shortens the lifespan significantly, compared to reticulum cell sarcoma and to the other causes of

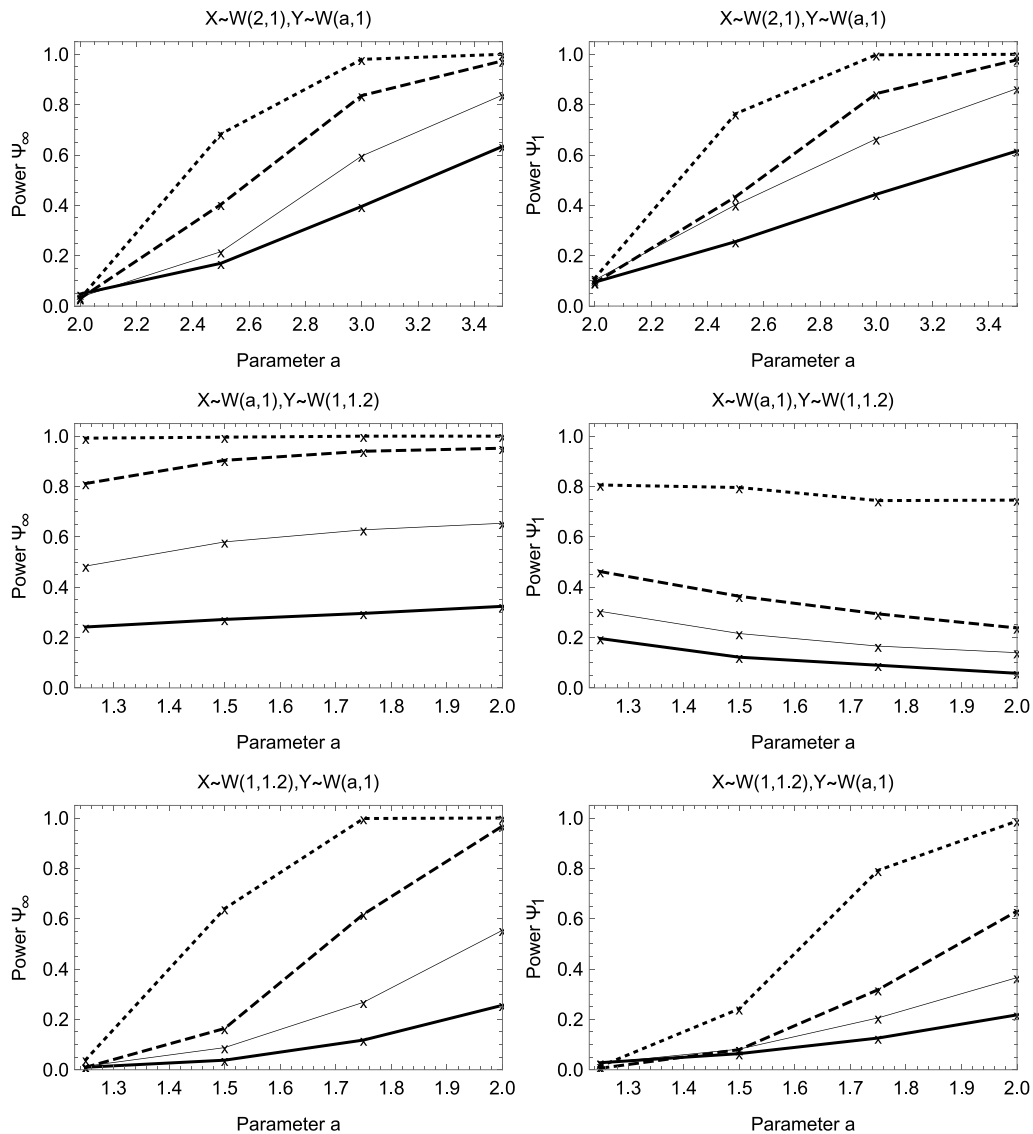


Fig. 1. Tests for H_0^{III} . Rejection rates for sample sizes $n = 50$ (solid), $n = 100$ (solid, thin), $n = 200$ (dashed), $n = 500$ (dotted).

death. On the other hand, these tests also reveal that the lifetime of the mice affected by reticulum cell sarcoma, denoted with X , is not stochastically smaller than that of the mice which died by other minor causes, denoted with Y (in this case, the hypothesis of dominance is rejected with a p -value close to 0). In particular, for small values of p , the p -quantiles of X are apparently larger than those of Y (see the left panel of Fig. 3), meaning that some other causes lead to death more rapidly than reticulum cell sarcoma. Moreover, sample means do not differ too much, but the sample mean of X (i.e., 635) is slightly larger compared to that of Y (i.e., 565), owing to one outlier in X (i.e., 986). However, the non-affected mice seem to live longer at larger quantile values.

It is then useful to understand whether weaker orders hold. It appears from Fig. 3 that we might have $X \leq_{ew} Y$ and $X \geq_{itt} Y$. By applying our test for the excess wealth order, we find a p -value around 0.5 for both Ψ_1 and Ψ_∞ . The tests of Belzunce et al. (2001), at least for some values of their parameter α , confirm that the excess wealth order may hold; moreover, in this case, the hypothesis $X =_d Y$ is easily rejected by the Kolmogorov–Smirnov test. Hence, the data suggest that $X \leq_{ew} Y$ (strictly). In other words, for every $p \in [0, 1]$, the mice that survive beyond the p -quantile seem to have a longer expected residual lifetime when they are not affected by this disease. Similarly, by applying our test for the TTT order, we obtain even larger p -values, indicating also that $X \geq_{itt} Y$, as confirmed by any of the tests in Belzunce et al. (2005). So, vice-versa, the expected lifetime of the mice that do not survive more than the p -quantile is smaller for the other (and possibly, more rapid) causes of death.

This example illustrates how the tests proposed in the present paper can assist in the analysis of survival data, in particular when stronger stochastic orders are not confirmed by tests, but the data suggest that the weaker orders considered in this paper

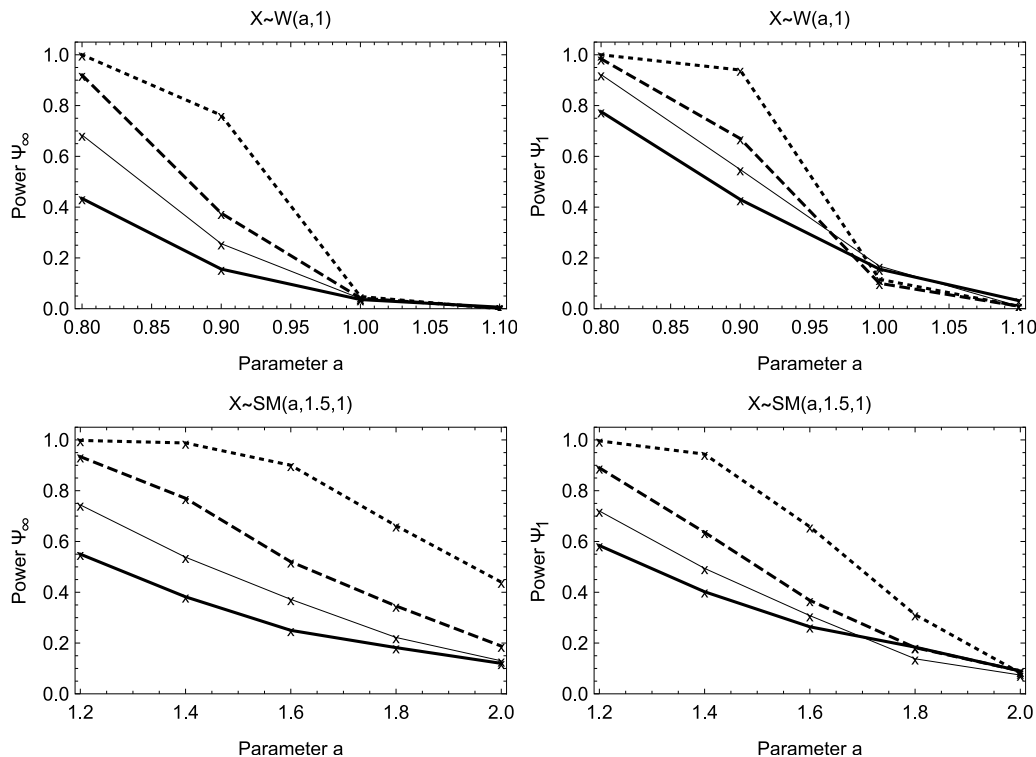


Fig. 2. Tests for $H_0^{NBU^E}$. Rejection rates for sample sizes $n = 50$ (solid), $n = 100$ (solid, thin), $n = 200$ (dashed), $n = 500$ (dotted).

may hold. Our tests can be used as an alternative, or complement, to those proposed by Belzunce et al. (2001, 2005), which in fact are not tailored to work under H_1^{ttt} or H_1^{ew} . This is especially useful when the hypothesis of equal distributions is rejected by a goodness-of-fit test, as in the example considered in this section.

Acknowledgements

We are grateful to the anonymous referees for their useful suggestions and constructive comments.

Appendix. Proofs

Proof of Theorem 1. Let $\mathbb{L}$ be the space of maps $z : [0, \infty) \rightarrow \mathbb{R}$ with $\lim_{x \rightarrow \infty} z(x) = 0$ and $\lim_{x \rightarrow 0} z(x) = 1$, and the norm $\|z\|_{\mathbb{L}} = \max\{\|z\|_{\infty}, \|1 - z\|_1\}$. Moreover, let $\mathbb{B}$ be the Banach space of measurable functions z from $(0, 1)$ to $\mathbb{R}$ equipped with the norm $\|z\|_{\mathbb{B}} = \int_0^1 |z(p)| dp$. As shown by Kaji (2018), under Assumption 1, the map $\theta(F) = F^{-1}$, from CDFs to quantile functions, is Hadamard differentiable at F , tangentially to the set $\mathbb{L}_0$ of continuous functions in $\mathbb{L}$, with derivative map $\theta'_F : \mathbb{L}_0 \rightarrow \mathbb{B}$ given by $\theta'_F(z) = -(z \circ F^{-1})(F^{-1})'$. Now, we need to establish the Hadamard differentiability of the map $\Phi : F^{-1} \rightarrow T_F$ as a map from $\mathbb{B}$ to $C[0, 1]$. By the change of variable $y = F(x)$, and then integrating by parts, we can write

$$T_F(p) = \int_0^{F^{-1}(p)} (1 - F(x)) dx = \int_0^p (1 - y) dF^{-1}(y) = F^{-1}(p)(1 - p) + \int_0^p F^{-1}(y) dy.$$

By linearity of the operator $\phi(F^{-1})$, the Hadamard derivative of ϕ at F^{-1} in the direction α is still ϕ , namely $\phi'_{F^{-1}}(\alpha) = \phi(\alpha)$. That is, letting $F_t^{-1} = F^{-1} + t\alpha_t$, where $\|\alpha_t - \alpha\|_1 \rightarrow 0$ and $t \rightarrow 0$,

$$\left\| \frac{\phi(F_t^{-1}) - \phi(F^{-1})}{t} - \phi'_{F^{-1}}(\alpha) \right\|_1 \leq \int_0^1 |(1 - p)(\alpha_t(p) - \alpha(p))| dp + \int_0^1 \int_0^p |\alpha_t(y) - \alpha(y)| dy dp \rightarrow 0.$$

By the chain rule for Hadamard differentiability (Van der Vaart and Wellner, 1996, Lemma 3.9.3), we have that the map $\phi \circ \theta : F \rightarrow T_F$ is Hadamard differentiable at F , tangentially to $\mathbb{L}_0$, with derivative

$$(\phi \circ \theta)'_F(z(p)) = \phi'_{\theta(F)} \circ \theta'_F(z(p)) = (1 - p)\theta'_F(z(p)) + \int_0^p \theta'_F(z(y)) dy = -(1 - p)(z \circ F^{-1}(p))(F^{-1}(p))' - \int_0^p (z \circ F^{-1}(y)) dF^{-1}(y).$$

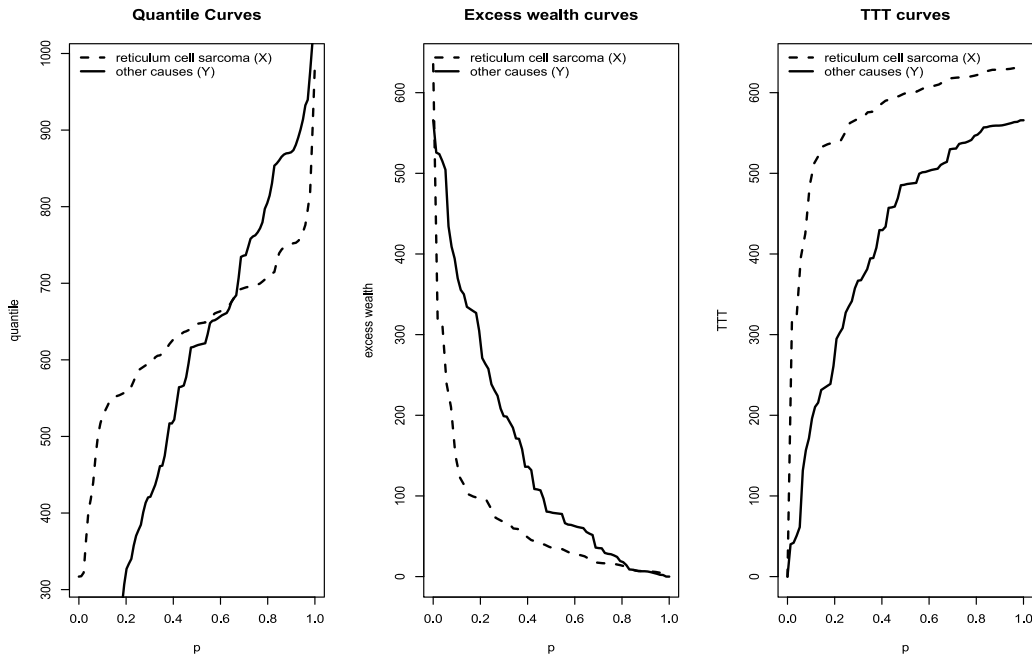


Fig. 3. Empirical quantile curves and excess wealth transforms of the data in Section 7. The lifetime of the mice affected by reticulum cell sarcoma (X) appears to be smaller in the excess wealth order, and larger in the TTT order.

Now, observe that

$$\left(\frac{\sqrt{n}(F_n - F)}{\sqrt{m}(G_m - G)} \right) \rightsquigarrow \left(\frac{B_1 \circ F}{B_2 \circ G} \right) \text{ in } \mathbb{L} \times \mathbb{L},$$

as shown in Lemma 5.1 of Sun and Beare (2021). Then, the functional delta method (Van der Vaart and Wellner, 1996, Theorem 3.9.4) implies the joint weak convergence

$$\begin{aligned} \left(\frac{\sqrt{n}(T_{F_n}(p) - T_F(p))}{\sqrt{m}(T_{G_m}(p) - T_G(p))} \right) &= \left(\frac{\sqrt{n}(\phi \circ \theta(F_n)(p) - \phi \circ \theta(F)(p))}{\sqrt{m}(\phi \circ \theta(G_m)(p) - \phi \circ \theta(G)(p))} \right) \rightsquigarrow \left(\frac{(\phi \circ \theta)'_F(B_1 \circ F)(p)}{(\phi \circ \theta)'_G(B_2 \circ G)(p)} \right) \\ &= \left(\frac{-(1-p)B_1(p)(F^{-1})'(p) - \int_0^p B_1(y)dF^{-1}(y)}{-(1-p)B_2(p)(G^{-1})'(p) - \int_0^p B_2(y)dG^{-1}(y)} \right) = \left(\frac{-\mathcal{T}_F(p)}{-\mathcal{T}_G(p)} \right) =_d \left(\frac{\mathcal{T}_F(p)}{\mathcal{T}_G(p)} \right) \text{ in } C[0, 1] \times C[0, 1]. \end{aligned} \quad (1)$$

The last relation ($=_d$) is because $\mathcal{T}_F$ and $\mathcal{T}_G$ are centred Gaussian processes. Then the results follow from the continuous mapping theorem, as in the proof of Lemma 2.1 of Sun and Beare (2021). $\square$

Proof of Proposition 1. As proved in Lemma 5.2 of Sun and Beare (2021),

$$\left(\frac{\sqrt{n}(F_n^* - F_n)}{\sqrt{m}(G_m^* - G_m)} \right) \overset{as*}{\rightsquigarrow} \left(\frac{B_1 \circ F}{B_2 \circ G} \right) \text{ in } \mathbb{L} \times \mathbb{L},$$

where $\overset{as*}{\rightsquigarrow}$ denotes weak convergence conditional on the data almost surely, see Kosorok (2008, p.20). The proof of Theorem 1 establishes the Hadamard differentiability of the map $\phi \circ \theta : H \rightarrow T_H$, so that the functional delta method for the bootstrap implies

$$\sqrt{r_{n,m}} \left(\frac{T_{F_n^*} - T_{F_n}}{T_{G_m^*} - T_{G_m}} \right) \overset{P}{\rightsquigarrow} \left(\frac{\lambda \mathcal{T}_F}{(1-\lambda)\mathcal{T}_G} \right) \text{ in } C[0, 1] \times C[0, 1],$$

where $\overset{P}{\rightsquigarrow}$ denotes weak convergence conditional on the data in probability (see Kosorok, 2008, p.20). This entails that $\Psi_r(\sqrt{r_{n,m}}(\hat{\delta}_k^* - \hat{\delta})) \overset{P}{\rightsquigarrow} \Psi_r(\tilde{\mathcal{T}})$ by the continuous mapping theorem. The test rejects the null hypothesis if the test statistic exceeds the bootstrap threshold value $c_n^*(\alpha) = \inf\{y : P(\sqrt{r_{n,m}} \Psi_r(\hat{\delta}_k^* - \hat{\delta}) > y | \mathcal{X}, \mathcal{Y}) \leq \alpha\}$. However, the weak convergence result implies $c_n^*(\alpha) \rightarrow c(\alpha) = \inf\{y : P(\Psi_r(\tilde{\mathcal{T}}) > y) \leq \alpha\}$ in probability, so Lemma 1 yields the result. $\square$

Proof of Corollary 1. Using the same notations for θ and Ψ as in the proof of Theorem 1, we now focus on the map $\eta \circ \Psi \circ \theta : F \rightarrow S_F$, where $\eta : C[0, 1] \rightarrow C[0, 1]$ is defined as $\eta(T_F) = T_F(\cdot)/T_F(1)$. The Hadamard derivative of η at T_F in the direction $u \in C[0, 1]$ is given

by

$$\eta'_{T_F}(u) = \frac{u(\cdot)}{T_F(1)} - \frac{u(1)T_F(\cdot)}{T_F(1)^2},$$

see the proof of Lemma 2.1 in Sun and Beare (2021). Therefore, the result follows from the chain rule and the functional delta method. $\square$

References

- Anis, M., Basu, K., 2014. Tests for exponentiality against NBUE alternatives: a Monte Carlo comparison. *J. Stat. Comput. Simul.* 84 (2), 231–247.
- Anis, M., Mitra, M., 2011. A generalized hollander–proschan type test for NBUE alternatives. *Stat. Probab. Lett.* 81 (1), 126–132.
- Barlow, R.E., Proschan, F., 1975. *Statistical Theory of Reliability and Life Testing: Probability Models*. Holt, Rinehart and Winston New York.
- Barlow, R.E., Van Zwet, W.R., 1969. Asymptotic properties of isotonic estimators for the generalized failure rate function. Part 1: strong consistency. Technical Report, California Univ Berkeley Operations Research center.
- Barmalzan, G., Najafabadi, A.T.P., 2015. On the convex transform and right-spread orders of smallest claim amounts. *Insurance Math. Econom.* 64, 380–384.
- Barrett, G.F., Donald, S.G., 2003. Consistent tests for stochastic dominance. *Econom.* 71 (1), 71–104.
- Barrett, G.F., Donald, S.G., Bhattacharya, D., 2014. Consistent nonparametric tests for Lorenz dominance. *J. Bus. Econom. Statist.* 32 (1), 1–13.
- Belzunce, F., Li, X., Pinar, J.-F., Ruiz, J.-M., 2005. Tests for the total time on test transform order. *J. Stat. Plan. Inf.* 133 (1), 111–121.
- Belzunce, F., Martínez-Riquelme, C., 2015. Some results for the comparison of generalized order statistics in the total time on test and excess wealth orders. *Statist. Papers* 56, 1175–1190.
- Belzunce, F., Martínez-Riquelme, C., Ruiz, J.M., 2014. A characterization and sufficient conditions for the total time on test transform order. *Test* 23, 72–85.
- Belzunce, F., Pinar, J., Ruiz, J., 2001. A family of tests for right spread order. *Statist. Probab. Lett.* 54 (1), 79–92.
- Block, H.W., Savits, T.H., 1978. Shock models with NBUE survival. *J. Appl. Probab.* 15 (3), 621–628.
- Brown, M., 2013. Sharp bounds for NBUE distributions. *Ann. Oper. Res.* 208 (1), 245–250.
- Cheng, K., Lam, Y., 2002. Reliability bounds for the NBUE distributions. *Naval Res. Logist.* 49 (8), 781–797.
- Denuit, M., Vermandele, C., 1999. Lorenz and excess wealth orders, with applications in reinsurance theory. *Scand. Actuar. J.* 1999 (2), 170–185.
- Dhaene, J., Vanduffel, S., Goovaerts, M.J., Kaas, R., Tang, Q., Vyncke, D., 2006. Risk measures and comonotonicity: a review. *Stoch. Models* 22 (4), 573–606.
- Dobler, D., 2019. Bootstrapping the kaplan–meier estimator on the whole line. *Ann. Inst. Statist. Math.* 71, 213–246.
- Fernandez-Ponce, J.M., Kochar, S.C., Muñoz-Perez, J., 1998. Partial orderings of distributions based on right-spread functions. *J. Appl. Probab.* 35 (1), 221–228.
- Gill, R., 1983. Large sample behaviour of the product-limit estimator on the whole line. *Ann. Statist.* 49–58.
- Hoel, D.G., 1972. A representation of mortality data by competing risks. *Biometrics* 475–488.
- Hollander, M., Proschan, F., 1975. Tests for the mean residual life. *Biometrika* 62 (3), 585–593.
- Jewitt, I., 1989. Choosing between risky prospects: the characterization of comparative statics results, and location independent risk. *Manag. Sci.* 35 (1), 60–70.
- Kaji, T., 2018. *Essays on Asymptotic Methods in Econometrics* (Ph.D. thesis). Massachusetts Institute of Technology.
- Kaplan, E.L., Meier, P., 1958. Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* 53 (282), 457–481.
- Kochar, S.C., Li, X., Shaked, M., 2002. The total time on test transform and the excess wealth stochastic orders of distributions. *Adv. in Appl. Probab.* 34 (4), 826–845.
- Kochar, S., Li, X., Xu, M., 2007. Excess wealth order and sample spacings. *Stat. Methodol.* 4 (4), 385–392.
- Kosorok, M.R., 2008. *Introduction to Empirical Processes and Semiparametric Inference*. Springer.
- Lando, T., 2024. Nonparametric inference for NBUE distributions based on the TTT transform. *Statist. Probab. Lett.* 110157.
- Lando, T., Legramanti, S., 2025. A new class of nonparametric tests for second-order stochastic dominance based on the Lorenz P–P plot. *Scand. J. Stat.* 52 (1), 480–512.
- Li, X., Shaked, M., 2007. A general family of univariate stochastic orders. *J. Stat. Plan. Inference* 137 (11), 3601–3610.
- Marshall, A.W., Olkin, I., 2007. *Life Distributions*. Springer, New York.
- Shaked, M., Shanthikumar, J.G., 1998. Two variability orders. *Probab. Engrg. Inform. Sci.* 12 (1), 1–23.
- Shaked, M., Shanthikumar, J.G., 2007. *Stochastic Orders*. Springer, New York.
- Sordo, M.A., 2009. Comparing tail variabilities of risks by means of the excess wealth order. *Insurance Math. Econom.* 45 (3), 466–469.
- Sun, Z., Beare, B.K., 2021. Improved nonparametric bootstrap tests of Lorenz dominance. *J. Bus. Econom. Statist.* 39 (1), 189–199.
- Tenga, R., Santner, T.J., 1984. Testing goodness of fit to the increasing failure rate family. *Nav. Res. Logist. Q.* 31 (4), 617–630.
- Van der Vaart, A., Wellner, J., 1996. *Weak Convergence and Empirical Processes: with Applications to Statistics*. Springer Science & Business Media.