

Tests for stochastic orders and mean order statistics

José R. Berrendero and Javier Cárcamo *

Departamento de Matemáticas, Universidad Autónoma de Madrid, 28049 Madrid

Abstract

The aim of this note is to emphasize the fact, not observed previously in the literature, that many discrepancy measures used in tests related to different stochastic orders can be expressed as expectations of order statistics. In this way, we provide a new meaning to the corresponding test statistics which allows us to understand better, and potentially improve, the testing procedures. As illustration, we consider tests to detect overdispersion with respect to a specific probability model. In this setting, a test for the Weibull distribution is discussed in detail.

Keywords: Convex order; Dilation order; Dispersive order; Total time on test transform order; HNBUE; NBUE; DMRL; Hypothesis testing; Order statistics; Exponential distribution; Weibull distribution.

MSC: 60E15; 60E99; 62E99; 62G10; 62E20; secondary 62N05

1 Introduction

As it was pointed out by Arnold *et al.* (2008), problems involving order statistics arise in the most varied settings. In the context of stochastic orders, it was shown that the convex-type orders are characterized by certain relations satisfied by the mean order statistics (see de la Cal and Cárcamo (2006, 2010)). This connection was exploited in Berrendero and Cárcamo (2009, 2010) and Baíllo *et al.* (2009) to obtain discrepancy measures for hypothesis tests related to stochastic dominance assumptions.

*This research was supported by the Spanish MEC, grants MTM2007-66632 and MTM2010-17366.

E-mail addresses: joser.berrendero@uam.es and javier.carcamo@uam.es

In this note, we revisit different procedures to solve two testing problems involving stochastic orders. We show that the discrepancies obtained for this purpose are functions of expectations of the order statistics. Surprisingly, this fact was not observed earlier.

We discuss two problems associated with a stochastic order \preceq .

Testing problem 1: Let (X_1, \dots, X_{n_1}) and (Y_1, \dots, Y_{n_2}) be two independent random samples from the variables X and Y , respectively. We consider the test

$$\begin{aligned} H_0 : X =_{\preceq} Y \text{ against } H_1 : X \preceq Y \text{ and } H_0 \text{ is not true,} \\ \text{where } X =_{\preceq} Y \text{ means } X \preceq Y \text{ and } Y \preceq X, \text{ simultaneously.} \end{aligned}$$

This problem has been extensively analyzed for different orderings. For instance, Aly (1990) and Marzec and Marzec (1991) considered the dispersive order, Belzunce *et al.* (2000, 2005b) the dilation order, Belzunce *et al.* (2001) the right spread order, Kochar *et al.* (2002) the new better than used in expectation (NBUE) order, Belzunce *et al.* (2005a) the total time on test transform order, Berrendero and Cárcamo (2010) the convex-type orders.

Testing problem 2: Let (X_1, \dots, X_n) be a random sample from X . We want to test

$$\begin{aligned} H_0 : X \in \mathcal{F} \text{ against } H_1 : Y \preceq X \text{ for some } Y \in \mathcal{F} \text{ and } X \notin \mathcal{F}, \\ \text{where } \mathcal{F} \text{ is a family of random variables.} \end{aligned}$$

This second test is relevant in several disciplines. In reliability theory, many classes of ageing distributions are defined by the comparison (according to different criteria) with the exponential family. Hence, \mathcal{F} is fixed as the class of exponential variables and, depending on the choice of the order \preceq , the alternative hypothesis includes different ageing distributions. For example, Hendi *et al.* (1998), Klar (2000), Belzunce *et al.* (2005b), and Berrendero and Cárcamo (2009) discussed tests for exponentiality against harmonic new better than used in expectation (HNBUE) alternatives. Belzunce *et al.* (2000, 2001) analyzed decreasing mean residual life (DMRL) and NBUE alternatives. In a different context, Baíllo *et al.* (2009) considered \mathcal{F} the family of Poisson variables and the convex order. In this case, the associated test can be viewed as a test for zero-inflation and overdispersion.

Most of the discrepancy measures can be expressed as expectations of order statistics (see Section 3). Moreover, these discrepancies are usually included in larger families of discrepancy measures also connected with mean order statistics. The goal of this note is to highlight this fact, which provides a deeper insight into the existing methodologies and gives a unified view of this subject. Further, the identification of a large family of discrepancies can be useful to improve the power of the tests under different alternatives as we point out in Section 4.

In Section 2, we survey the existing approaches to generate discrepancies for the aforementioned testing problems for convex-type orders. We also describe the link between the discrepancies and the expectation of order statistics. In Section 3, we exhibit a number of particular examples from the literature. The examples include the dilation, dispersive, total time on test transform and NBUE order. Tests for exponentiality against DMRL, NBUE and HNBUE alternatives are included. Section 4 includes an application of the main ideas of this paper. We propose a family of tests to detect overdispersion with respect to a specified distribution. An example related to Weibull distributions is discussed in detail and a simulation study is carried out to compare the power of the proposed tests in this case. Finally, Section 5 presents the main conclusions of this article.

2 Measures of discrepancy and the expectations of order statistics

The usual approach to generate discrepancies for the tests described earlier is to obtain results of stochastic equality under dominance. We illustrate this point by considering the (increasing) convex order since the examples discussed in the next section are related to this order.

We recall that given two integrable random variables X and Y , it is said that X is dominated by Y in the *convex order*, written $X \leq_{\text{cx}} Y$, if $E\phi(X) \leq E\phi(Y)$ for all convex functions ϕ for which the previous expectations are well defined. There are several related stochastic orders. For example, the *increasing convex*, the *concave* and the *increasing concave orders*, denoted \leq_{icx} , \leq_{cv} and \leq_{icv} , are defined analogously, but respectively changing the set of “convex functions” to “increasing and convex”, “concave” and “increasing and concave functions”. The equality for the (increasing) convex order is the equality in distribution ($=_{\text{st}}$). The Testing problem 1 becomes $H_0 : X =_{\text{st}} Y$ against $H_1 : X \leq_{\text{cx}} Y$ and $X \neq_{\text{st}} Y$. That is, it is a test for stochastic equality versus *strict* convex domination.

One way to construct discrepancy measures for this test is the following result (see Sordo and Ramos (2007) or Berrendero and Cárcamo (2010)) based on a signed and weighted Wasserstein distance (between X and Y)

$$\Delta_{\omega}(X, Y) := \int_0^1 (G^{-1}(t) - F^{-1}(t)) \omega(t) dt, \quad (1)$$

where F^{-1} and G^{-1} are the quantile functions of X and Y , respectively, and ω is a real function on $[0, 1]$ for which the above integral makes sense.

Result 1. *Let \mathcal{I} denote the class of strictly increasing functions on $[0, 1]$ and \mathcal{I}_0 the subset $\omega \in \mathcal{I}$ with the property $\omega(0) \geq 0$. If $X \leq_{\text{cx}} Y$, then $\Delta_{\omega}(X, Y) \geq 0$, for all $\omega \in \mathcal{I}$. If additionally, $\Delta_{\omega}(X, Y) = 0$ for some $\omega \in \mathcal{I}$, then $X =_{\text{st}} Y$.*

The result still holds if “ \leq_{cx} ” and “ \mathcal{I} ” are replaced by “ \leq_{icx} ” and “ \mathcal{I}_0 ”.

Under H_0 , X and Y are equally distributed and $\Delta_\omega(X, Y) = 0$, while, under H_1 , $X \leq_{\text{cx}} Y$ and $\Delta_\omega(X, Y) > 0$, for each $\omega \in \mathcal{I}$. In other words, for each increasing function ω , $\Delta_\omega(X, Y)$ is a discrepancy measure for the test. Once the function ω is fixed, the test is carried out by estimating the discrepancy and rejecting H_0 when the estimate is too large.

Alternatively, other authors implicitly used the following well-known result (see for instance Denuit *et al.* (2000), Bhattacharjee and Bhattacharya (2000), or Shaked and Shanthikumar (2006, Theorem 4.A.49)).

Result 2. *If $X \leq_{\text{cx}} Y$ and $E\phi(X) = E\phi(Y)$, for some strictly convex function ϕ such that the previous expectations exist, then $X =_{\text{st}} Y$. The result still holds if “ \leq_{cx} ” and “convex” are replaced by “ \leq_{icx} ” and “increasing and convex”, respectively.*

Therefore, provided the expectations exist, for any strictly convex function ϕ

$$\Lambda_\phi(X, Y) := E\phi(Y) - E\phi(X) \quad (2)$$

is also a discrepancy measure for the test.

The nice thing about the distances in (1) is that they are closely connected with the expected value of order statistics. Let $X_{i:k}$ denote the i -th order statistic of a random sample of size k from X . It is well known (see Arnold *et al.* (2008, eq. (5.2.8))) that

$$\Delta_{\omega_{i:k}}(X, Y) = EY_{i:k} - EX_{i:k}, \quad (3)$$

where

$$\omega_{i:k}(t) := \frac{k!}{(i-1)!(k-i)!} t^{i-1} (1-t)^{k-i}, \quad t \in [0, 1]. \quad (4)$$

Starting from (3) and (4), it is not hard to show that many functions of the difference between the vectors of mean order statistics of the two variables can be expressed in the form $\Delta_\omega(X, Y)$, for some $\omega \in \mathcal{I}$ (or \mathcal{I}_0). For example, for $k \geq 2$, the function $\omega_1 := \omega_{k:k}$ belongs to \mathcal{I}_0 . Also, for $k \geq 2$, $\omega_2 := \omega_{k:k} - \omega_{1:k} \in \mathcal{I}$ and $\omega_3 := k + \omega_2 \in \mathcal{I}_0$. The generated discrepancies are

$$\begin{aligned} \Delta_{\omega_1}(X, Y) &= EY_{k:k} - EX_{k:k}, \\ \Delta_{\omega_2}(X, Y) &= E(Y_{k:k} - Y_{1:k}) - E(X_{k:k} - X_{1:k}), \\ \Delta_{\omega_3}(X, Y) &= E(Y_{k:k} - Y_{1:k}) - E(X_{k:k} - X_{1:k}) + k(EY - EX). \end{aligned} \quad (5)$$

This means that the comparison of the expected maxima or ranges of the two variables allows us to distinguish H_0 from H_1 .

3 Examples

In this section, we illustrate the previous ideas with a collection of examples from the literature. It should be remarked that in many cases only a small number of discrepancies are detected and no link with the mean of order statistics is given. First, we consider the Testing problem 1 for the dilation, dispersive, total time on test transform, and NBUE order. Afterwards, we examine the second problem with examples of tests of exponentiality against DMRL, NBUE and HNBUE alternatives. For each case, we show that the proposed discrepancy measures, although derived using diverse lines of reasoning, can be written in one of the two forms (1) and (2), and can accordingly be interpreted in terms of expected order statistics.

3.1 Examples of the Testing problem 1

Dilation order. Given two integrable random variables X and Y , the *dilation order*, \leq_{dil} , is defined by the condition $X - EX \leq_{\text{cx}} Y - EY$. Hence, any distance (between $X - EX$ and $Y - EY$) of the form (1) with ω increasing is a discrepancy for the corresponding test. Actually, it is readily checked that the quantity used in Belzunce *et al.* (2000, eq. (3)) is

$$E(Y - EY)_{2:2} - E(X - EX)_{2:2}. \quad (6)$$

It is also easy to see that the proposed family of measures given in Belzunce *et al.* (2005b) for this order is the distance $\Delta_{\omega_\alpha}(X - EX, Y - EY)$, with $\omega_\alpha \in \mathcal{I}_0$ ($0 < \alpha < 1$) defined by

$$\omega_\alpha(t) := \begin{cases} \frac{1}{2} \left(\frac{1}{\alpha} - 1\right) t^2 & \text{if } 0 \leq t \leq \alpha, \\ \frac{1}{2} (2t - t^2 - \alpha) & \text{if } \alpha < t \leq 1. \end{cases} \quad (7)$$

The limiting case $\alpha = 0$ (not considered in that article) is still a measure of discrepancy. In this case, using (3)-(4), it is easy to check that $\omega_0(t) := t - t^2/2$ generate the distance

$$\Delta_{\omega_0}(X - EX, Y - EY) = \frac{1}{6} (E(X - EX)_{1:3} - E(Y - EY)_{1:3}).$$

When α is close to 1, $\omega_\alpha(t)$ is a constant times t^2 . Thus, the associated measure is basically $E(Y - EY)_{3:3} - E(X - EX)_{3:3}$. In this way, $\Delta_{\omega_\alpha}(X, Y)$ ($0 \leq \alpha < 1$) is a measure between $EY_{3:3} - EX_{3:3}$ and $EX_{1:3} - EY_{1:3}$.

Dispersive order. X is said to be smaller than Y in the *dispersive order* ($X \leq_{\text{disp}} Y$), if $G^{-1} - F^{-1}$ is nondecreasing. According to Shaked and Shanthikumar (2006,

Theorem 3.B.16), $X \leq_{\text{disp}} Y$ implies $X \leq_{\text{dil}} Y$. Therefore, any measure $\Delta_\omega(X - EX, Y - EY)$, with $\omega \in \mathcal{I}$ works for the dispersive order.

The measure of discrepancy for the dispersive order in Aly (1990, eq. (2.3)) is $\Delta_\omega(X - EX, Y - EY)$, with $\omega(t) = 2t$, so it is

$$E(Y - EY)_{2:2} - E(X - EX)_{2:2}. \quad (8)$$

In other works, the considered discrepancy is of the form $\Delta_\omega(X, Y)$, with ω *not* increasing. This is the case of Marzec and Marzec (1991, eq. (2.1)), who used the measure corresponding to the weight function

$$b_\alpha(t) := \begin{cases} t \left(\frac{t}{2\alpha(1-\alpha)} - 1 \right) & \text{if } 0 \leq t \leq \alpha, \\ (2t - 1) \frac{\alpha}{2(1-\alpha)} & \text{if } \alpha < t \leq 1 - \alpha, \\ (1 - t) \left(\frac{t-1}{2\alpha(1-\alpha)} + 1 \right) & \text{if } 1 - \alpha \leq t \leq 1, \end{cases} \quad (0 < \alpha \leq 1/2).$$

In the limiting case $\alpha \downarrow 0$, $b_0(t)$ is proportional to $2t - 1$. Hence, the generated discrepancy is again (8). The authors noted that the test with $\alpha = 0.5$ performs better under the models considered in that paper. The function $b_{0.5}$ is a polynomial of degree 2 on $[0, 0.5]$ and $[0.5, 1]$, so it is related to the vector $(EY_{1:3} - EX_{1:3}, EY_{2:3} - EX_{2:3}, EY_{3:3} - EX_{3:3})$.

Total time on test transform and NBUE order. Let X and Y be two nonnegative random variables. X is less than Y in the *total time on test transform order*, written $X \leq_{\text{ttt}} Y$, if

$$\int_0^{F^{-1}(p)} (1 - F(x)) dx \leq \int_0^{G^{-1}(p)} (1 - G(x)) dx, \quad 0 < p < 1.$$

It follows from Kochar *et al.* (2002, Corollary 3.1) (see also Shaked and Shanthikumar (2006, Corollary 4.B.27)) that, under certain simple conditions, the total time on test transform order implies the increasing concave order. In other words, in some situations $X \leq_{\text{ttt}} Y$ implies $-Y \leq_{\text{icx}} -X$. Hence, in these cases, the discrepancies $\Delta_\omega(-Y, -X)$ (with $\omega \in \mathcal{I}_0$) work for \leq_{ttt} . The discrepancy obtained in Belzunce *et al.* (2005a, Section 2.1) for this order can be rewritten as $EY_{1:2} - EX_{1:2}$.

For integrable and positive random variables, the NBUE order can be characterized by $Y/EY \leq_{\text{ttt}} X/EX$ (see Kochar *et al.* (2002, eq. (3.2))). The quantity S used in Kochar *et al.* (2002, Application 5.3) for the test corresponding to this order is

$$\frac{EY_{1:2}}{EY} - \frac{EX_{1:2}}{EX} = G(X) - G(Y),$$

where $G(X)$ stands for the popular *Gini index* of the variable X .

Remark 1. Different discrepancies can lead to very different results regarding the power of the test. For instance, when comparing the discrepancies $EY_{k:k} - EX_{k:k}$ ($k \geq 2$), it was shown in Berrendero and Cárcamo (2010) that, under various alternatives, relatively large values of k (around 10) give good power results. In most of the considered situations, the lowest power is achieved when $k = 2$ or 3.

3.2 Examples of the Testing problem 2

Testing exponentiality against DMRL alternatives. In Belzunce *et al.* (2000, Section 3.1) the problem of testing exponentiality against DMRL alternatives was considered. Since a nonnegative random variable X is DMRL if and only if $X_t \leq_{\text{dil}} X_s$ for all $0 \leq s < t$, where $X_r := \{X - r | X > r\}$ is the remaining life at age $r \geq 0$, the authors used the discrepancy obtained for the dilation order to derive an appropriate one in this case. The plug-in estimator for the expected maxima is

$$E_{F_n}(X_{k:k}) = \sum_{i=1}^n \left[\left(\frac{i}{n} \right)^k - \left(\frac{i-1}{n} \right)^k \right] X_{i:n} \asymp \frac{k}{n} \sum_{i=1}^n \left(\frac{i}{n} \right)^{k-1} X_{i:n}, \quad (9)$$

where F_n is the empirical distribution function of the sample and “ \asymp ” means that the estimators are asymptotically equivalent (when $n \rightarrow \infty$). Therefore, the proposed test statistics, $\Delta_{\text{DMRL}}(n)$, is an estimator of the quantity

$$\frac{1}{30} (-5EX_{2:2} + 15EX_{3:3} - 15EX_{4:4} + 6EX_{5:5} - EX_{6:6}).$$

Testing exponentiality against NBUE alternatives. A nonnegative random variable X is NBUE if and only if $Y \leq_{\text{dil}} X_t$ for all $t \geq 0$, where Y is an exponential random variable of mean EX and X_t is the remaining life at age t . Using (9), it can be checked that the discrepancy $\Delta_{\text{NBUE}}(n)$ proposed by Belzunce *et al.* (2000, Section 3.2) is an estimator of

$$\frac{1}{6} (EX + 3EX_{2:2} - 3EX_{3:3}).$$

Testing exponentiality against HNBUE alternatives. A nonnegative random variable X is HNBUE if $X \leq_{\text{icx}} Y$, where Y is an exponential variable with mean EX . It follows that X is HNBUE if and only if $X/EX \leq_{\text{icx}} Y^1$, where Y^1 is an exponential variable with mean 1. Therefore, we can use the distances (1) (between X/EX and Y^1) with ω increasing to generate a test for exponentiality against HNBUE alternatives. This idea was exploited in Berrendero and Cárcamo (2009). Also, the tests proposed by Belzunce *et al.* (2005b, Section 3) based on the weights (7) are also connected with the expectation of the order statistics.

Klefsjö (1983) examined the statistics

$$Q_{1,\nu} = \frac{1}{n\bar{X}} \sum_{i=1}^n \left(\nu \left(1 - \frac{i}{n}\right)^{\nu-1} - \frac{1}{\nu} \right) X_{i:n}, \quad \nu \geq 2,$$

$$Q_{2,\nu} = \frac{1}{n\bar{X}} \sum_{i=1}^n \left(\sum_{k=1}^{\nu} \frac{1}{k} - \nu \left(\frac{i}{n}\right)^{\nu-1} \right) X_{i:n}, \quad \nu \geq 2,$$

where $\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i$ is the sample mean. Taking into account (9), $Q_{2,\nu}$ is an estimator of $EX_{\nu:\nu}/EX - EY_{\nu:\nu}^1$. Analogously, $Q_{1,\nu}$ is the empirical counterpart of $EY_{1:\nu}^1 - EX_{1:\nu}/EX$. The unbiased version of $Q_{1,2} = Q_{2,2}$ was analyzed in Hendi *et al.* (1998).

On the other hand, the test statistics $T_{n,a}$ in Klar (2000, eq. (2)–(4)) can be rewritten as estimators of the discrepancies $\Lambda_{\phi_a}(X/EX, Y_1)$, where Λ_{ϕ} is defined in (2) and, for $t > 0$, the convex function ϕ_a is

$$\phi_0(t) := \frac{1}{2}t^2 \quad \text{and} \quad \phi_a(t) := \frac{1}{a^2}e^{-at}, \quad \text{if } a > 0.$$

In other words,

$$\Lambda_{\phi_0}(X/EX, Y^1) = \frac{1}{2} \left(E(X/EX)^2 - E(Y^1)^2 \right),$$

$$\Lambda_{\phi_a}(X/EX, Y^1) = \frac{1}{a^2} \left(Ee^{-aX/EX} - Ee^{-aY^1} \right), \quad \text{if } a > 0.$$

Therefore, for $a > 0$, they compared the Laplace transform of the variable X/EX with the corresponding to that of a mean 1 exponential variable. When $a = 0$, $T_{n,0}$ is based on the comparison of the second moments of X/EX and Y_1 . It was pointed out by Klar (2000) that $T_{n,0}$ and $T_{n,1}$ appear in different tests related to the exponential distribution in various references of the literature.

4 An application: testing a specified model against overdispersion

The links between order statistics and stochastic orders discussed in the previous sections provide a large class of potentially useful discrepancies for a variety of testing problems. The interpretation of the discrepancies may serve as a guide to devise large families of test statistics. Different statistics may lead to tests with different power performances and, therefore, a richer family of discrepancies increases the opportunity to select a powerful test for the problem at hand. In this section we provide an application to illustrate these facts.

4.1 General description of the problem

Let X_0 be a random variable with a specified known distribution function F_0 . Given a random sample X_1, \dots, X_n , it may be of interest to know if F_0 appropriately ac-

counts for the variability observed in the sample or if, otherwise, an alternative model is preferred on this ground. In this case, we say that the sample shows overdispersion with respect to F_0 . The notion of overdispersion can be formalized in terms of dominance with respect to an appropriate variability stochastic order. Thus, we consider a testing problem such that under the null hypothesis the underlying distribution of the sample is F_0 , but under the alternative hypothesis the distribution underlying the data, F , dominates F_0 in the increasing convex order. In the framework described in Section 1, this problem corresponds to the Testing problem 2 with \mathcal{F} only including the random variable X_0 , and \preceq being the increasing convex order.

Our approach is to use the family of discrepancies given by Equation (5) with $k = 2, 3, \dots$. This amounts to considering differences between the expected maxima under H_0 and H_1 . The expected maximum under H_1 has to be estimated from the sample. A natural estimator arises by computing the expected maxima under the empirical distribution of the sample X_1, \dots, X_n (see Equation (9)). As for the expected maximum under H_0 , there are instances where it is explicitly known (see the Weibull example below). If not, we can always resort to numerical methods or simulations in order to obtain a suitable approximation. We end up with the following class of test statistics:

$$\Delta_k := E_{F_n}(X_{k:k}) - E_{F_0}(X_{k:k}) = \sum_{i=1}^n \left[\left(\frac{i}{n} \right)^k - \left(\frac{i-1}{n} \right)^k \right] X_{i:n} - E_{F_0}(X_{k:k}). \quad (10)$$

A direct application of Theorem 2.1 in Li et al. (2001) allows us to derive the asymptotic distribution of Δ_k ($k \geq 2$) under H_0 :

Proposition 1. *Let U be a random variable uniformly distributed in $[0, 1]$, let $N(\mu, \sigma^2)$ denote a normal distribution with mean μ and variance σ^2 , and denote convergence in distribution as \rightarrow_d . Then, under $H_0 : F = F_0$,*

$$\sqrt{n}\Delta_k \rightarrow_d N(0, \sigma_k^2), \quad \text{as } n \rightarrow \infty,$$

where $\sigma_k^2 := \text{Var}(W_k)$, with

$$W_k := kU^{k-1}F_0^{-1}(U) + k(k-1) \int_U^1 t^{k-2}F_0^{-1}(U) dt. \quad (11)$$

Observe that it is always possible to approximate σ_k^2 with any desired degree of accuracy by generating a large enough number of uniform random variables. Proposition 1 yields the following family of critical regions with asymptotic significance level α :

$$\mathcal{R}_k := \left\{ \sqrt{n}\Delta_k / \sigma_k > z_\alpha \right\}, \quad k \geq 2,$$

where z_α is the appropriate percentile of the standard normal distribution. Another possibility, more expensive computationally, is to approximate via Monte Carlo (generating a large number of realizations of Δ_k under the null hypothesis) the critical values beyond which H_0 should be rejected for a given significance level.

4.2 Application to the Weibull model

The Weibull distribution enjoys a large number of applications in diverse fields, usually related with aging problems and reliability. See for example Rinne (2009) for a recent and thorough survey of applications and methods. In this subsection we apply the general test described above to this important particular case. Then, F_0 is a Weibull distribution, whose cumulative distribution function is given by

$$F_0(x) = 1 - \exp \left\{ - \left(\frac{x}{b_0} \right)^{c_0} \right\}, \quad x > 0,$$

where $b_0 > 0$ is a scale parameter and $c_0 > 0$ is a shape parameter. We denote this distribution as $W(c_0, b_0)$. Under the alternative hypothesis, the distribution underlying the data dominates $W(c_0, b_0)$ in the increasing convex order.

For our purposes, we consider a particular case of a formula, due to Lieblein (1955), for the moments of the order statistics of a Weibull distribution to obtain the following explicit expression for the expectation of $X_{k:k}$ under the Weibull model $W(c_0, b_0)$:

$$E_{F_0}(X_{k:k}) = b_0 k \Gamma \left(1 + \frac{1}{c_0} \right) \sum_{i=0}^{k-1} \frac{(-1)^i \binom{k-1}{i}}{(i+1)^{1+1/c_0}}. \quad (12)$$

Using (12), the family of test statistics (10) reduces to

$$\Delta_k = \sum_{i=1}^n \left[\left(\frac{i}{n} \right)^k - \left(\frac{i-1}{n} \right)^k \right] X_{i:n} - b_0 k \Gamma \left(1 + \frac{1}{c_0} \right) \sum_{i=0}^{k-1} \frac{(-1)^i \binom{k-1}{i}}{(i+1)^{1+1/c_0}}.$$

We can now apply Proposition 1 to derive the asymptotic distribution of Δ_k under H_0 . The asymptotic variance is given by $\sigma_k^2 = \text{Var}(W_k)$, where

$$W_k = b_0 k U^{k-1} (-\log(1-U))^{1/c_0} + b_0 k (k-1) \int_U^1 t^{k-2} (-\log(1-U))^{1/c_0} dt.$$

The last expression follows from (11) since $F_0^{-1}(u) = b_0 (-\log(1-u))^{1/c_0}$.

4.3 Simulations

We have carried out a small simulation study to compare and assess the performance of the tests given by \mathcal{R}_k for several values of $k \geq 2$, and both for the asymptotic and the Monte Carlo critical values. We have considered three different models related to (mixtures of) Weibull models:

(M0) The Weibull $W(2, 1)$ plays the role of the null hypothesis.

(M1) Mixture of $W(2, 1)$ and $W(1.5, 1)$ (with weights $1/2$).

(M2) Mixture of $W(2, 1)$ and $W(1, 1)$ (with weights $1/2$).

For some applications, mixtures of two Weibull distributions have been proposed as alternative to the Weibull model, so that it may be of interest to detect deviations from the Weibull distribution towards this type of alternatives. See Murthy et al. (2004, Chapter 8) for an overview of the applications and properties of Weibull mixtures.

To give an idea of the difficulty of detecting deviations from (M0) towards (M1) and (M2), Figure 1 displays the cumulative distribution functions corresponding to the three models. The expectations of random variables distributed as (M0), (M1) and (M2) are 0.886, 0.894 and 0.943, respectively. These figures together with the fact that there is only one point at which each pair of distribution functions cross each other show that both (M1) and (M2) dominate (M0) in the increasing convex order (see Shaked and Shanthikumar (2006, Theorem 4.A.22)).

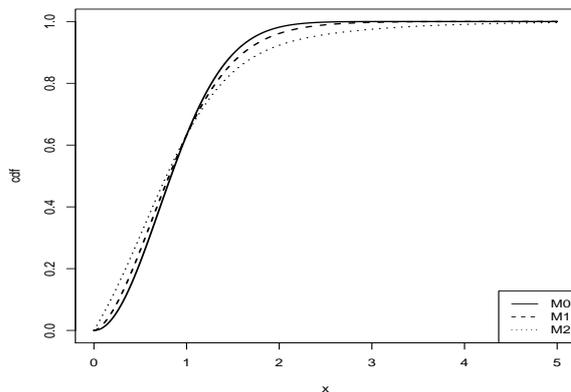


Figure 1: Cumulative distribution functions corresponding to (M0), (M1) and (M2).

We have considered three different sample sizes $n = 20, 100, 200$. For each model and sample size we have generated 10,000 samples. For each sample we have applied nine different tests for H_0 with nominal significance level $\alpha = 0.05$: the asymptotic tests based on Δ_k (labeled as Δ_k^a) with $k = 2, 4, 6, 8$; the four tests based on the same statistics but with critical values approximated via Monte Carlo (labeled as Δ_k^{mc}); and the Kolmogorov-Smirnov test (KS). It should be noted that the KS test does not take into account any information about the stochastic dominance under the alternative, so it is not fair to compare its power with that of the tests we propose here. Still, the KS test is informative to understand better the departures

Model	n	Δ_2^{mc}	Δ_4^{mc}	Δ_6^{mc}	Δ_8^{mc}	Δ_2^{a}	Δ_4^{a}	Δ_6^{a}	Δ_8^{a}	KS
M0	20	0.051	0.050	0.051	0.050	0.041	0.034	0.033	0.030	0.054
M0	100	0.049	0.048	0.051	0.049	0.045	0.042	0.041	0.040	0.044
M0	200	0.048	0.045	0.045	0.047	0.046	0.040	0.041	0.040	0.048
M1	20	0.136	0.189	0.204	0.215	0.121	0.154	0.165	0.169	0.064
M1	100	0.241	0.386	0.464	0.488	0.227	0.365	0.432	0.460	0.103
M1	200	0.334	0.559	0.650	0.699	0.325	0.545	0.638	0.681	0.176
M2	20	0.406	0.532	0.574	0.590	0.386	0.499	0.537	0.551	0.128
M2	100	0.770	0.926	0.958	0.967	0.761	0.919	0.955	0.964	0.450
M2	200	0.927	0.992	0.998	0.999	0.925	0.992	0.997	0.998	0.809

Table 1: Proportion of rejections across 10,000 samples for nine different tests. Nominal significance level is $\alpha = 0.05$.

of the alternatives (M1) and (M2) from the null (M0). In Table 1 we report the proportion of rejections across the 10,000 replications of the experiment.

Regarding the results under (M0), observe that the asymptotic tests are somewhat conservative, especially for small sample sizes ($n = 20$) and for large values of k . For $n = 100$ and $n = 200$, the empirical significance levels are already quite close to the nominal value 0.05 although consistently below it. The use of Monte Carlo to approximate the critical values allows us to attain significance levels closer to 0.05, which reflects in a slightly higher power under (M1) and (M2).

The most important conclusion that can be drawn from Table 1 is that as k gets larger the power also increases. The difference is significant when we compare $k = 2$ with $k = 8$. Therefore, the link between the discrepancies and the order statistics, which allows us to introduce a complete family of test statistics based on $X_{k:k}$ for $k = 2, 3, \dots$, is useful to find tests with higher power.

5 Conclusion

We emphasize the connection between the discrepancy measures considered in tests related to stochastic orders and the expectations of the order statistics. From our point of view, this connection has several advantages: (a) It provides a precise meaning to the discrepancies. For instance, the discrepancies can be interpreted as the expected lifetime of different systems (in reliability theory), as the mean wealth in ordered samples of the population (in welfare and inequality analysis), or as the expected loss of different loss events (in risk analysis or insurance). (b) Since the natural estimators of the expectation of the order statistics are L-statistics, the

widely developed theory for L-statistics can be used to obtain easily the asymptotic distribution of the test statistics (see, for instance, the books by Serfling (1980) and Shorack and Wellner (1986) or, more recently, Li *et al.* (2001)). (c) Finally, the interpretation of the discrepancies in terms of mean order statistics usually allows us to identify a larger family of potential discrepancies to perform the test. This is important since the testing problems allow for many potential deviations from the null, and therefore, different discrepancies can lead to very different powers depending on the distribution under the alternative. Thus, having a larger family of discrepancies makes possible the choice of tests with a good power behavior for a large set of deviations from the null, or the use of an adaptive approach.

We believe the presented perspective helps to unify and clarify the literature on this topic, and paves the way for further improvements of the existing methods.

Acknowledgements

The authors would like to thank the Editor and the reviewers for their comments which help to improve the first version of the manuscript.

References

- Aly, E-E. A.A. (1990). A simple test for dispersive ordering. *Statist. Probab. Lett.* **9**, 323–325.
- Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (2008). *A First Course in Order Statistics*. SIAM Classics in Applied Mathematics.
- Baíllo, A., Berrendero, J.R. and Cárcamo, J. (2009). Tests for zero-inflation and overdispersion: A new approach based on the stochastic convex order. *Comput. Statist. Data Anal.* **53**, 2628–2639.
- Bhattacharjee, M.C., Bhattacharya, R.N. (2000). Stochastic equivalence of convex ordered distributions. *Probab. Engrg. Inform. Sci.* **14**, 33–48.
- Belzunce, F., Candel, J. and Ruiz, J.M. (2000). Testing mean residual alternatives by dispersion of residual lives. *J. Statist. Plann. Inference* **86**, 113–127.
- Belzunce, F., Li, X., Pinar, J.F. and Ruiz, J.M. (2005a). Test for the total time on test transform order. *J. Statist. Plann. Inference* **133**, 111–121.
- Belzunce, F., Pinar, J.F., and Ruiz, J.M. (2001). A family of tests for the right spread order. *Statist. Probab. Lett.* **54**, 79–92.
- Belzunce, F., Pinar, J.F. and Ruiz, J.M. (2005b). On testing the dilation order and HNBUE alternatives. *Ann. Inst. Statist. Math.* **57**, 803–815.

- Berrendero, J.R. and Cárcamo, J. (2009). Characterizations of exponentiality within the HNBUE class and related tests. *J. Statist. Plann. Inference* **139**, 2399–2406.
- Berrendero, J.R. and Cárcamo, J. (2010). Tests for the second order stochastic dominance based on L-statistics. Accepted for publication in *J. Bus. Econom. Statist.* DOI: 10.1198/jbes.2010.07224
- de la Cal, J. and Cárcamo, J. (2006). Stochastic orders and majorization of mean order statistics. *J. Appl. Prob.* **43**, 704–712.
- de la Cal, J. and Cárcamo, J. (2010). Inverse stochastic dominance, majorization, and mean order statistics. *J. Appl. Prob.* **47**, 277–292.
- Denuit, M., Lefèvre, C., and Shaked, M. (2000). On the theory of high convexity stochastic orders. *Statist. Probab. Lett.* **47**, 287–293.
- Hendi, M.I., Al-Nachawati, H., Montasser, M., and Alwasel, I.A. (1998). An exact test for HNBUE class of life distributions, *J. Statist. Comput. Simul.* **60**, 161–275.
- Klefsjö, B. (1983). Testing exponentiality against HNBUE. *Scand. J. Statist.* **10**, 65–75.
- Klar, B. (2000). A class of test for exponentiality against HNBUE alternatives. *Statist. Probab. Lett.* **47**, 199–207.
- Kochar, S.C., Li, X. and Shaked, M. (2002). The total time on test transform and the excess wealth stochastic orders of distributions. *Adv. Appl. Probab.* **34**, 826–845.
- Li, D., Rao, M.B. and Tomkins, R.J. (2001), The law of the iterated logarithm and central limit theorem for L-statistics. *J. Multivariate Anal.* **78**, 191–217.
- Lieblein, J. (1955). On moments of order statistics from the Weibull distribution. *Annals of Mathematical Statistics*, **26**, 330–333.
- Marzec, L. and Marzec, P. (1991). On testing equality in dispersion of two probability distributions. *Biometrika* **78**, 923–925.
- Murthy, D.N.P., Xie, M. and Jiang, R. (2004). *Weibull Models*. Wiley, New York.
- Rinne, H. (2009). *The Weibull distribution. A handbook*. Chapman & Hall/CRC, Boca Raton.
- Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

Shaked, M. and Shanthikumar, J.G. (2006). *Stochastic Orders*. Springer Series in Statistics.

Shorack, G. R. and Wellner J.A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

Sordo, M.A. and Ramos, H.M. (2007). Characterization of stochastic orders by L-functionals. *Statist. Papers* **48**, 249–263.