Flexible-dimensional EVR-OWA as mean estimator for symmetric distributions*

Juan Baz[1][0000−0003−1142−6077], Diego García-Zamora[2][0000−0002−0843−4714], Irene Díaz[3][0000−0002−3024−6605], Susana Montes[1][0000−0002−4701−2207], and Luis Martínez[2][0000−0003−4245−8813]

[1] Department of Statistics and O.R. and Mathematics Didactics, University of Oviedo, Oviedo, Spain bazjuan,montes@uniovi.es
[2] Department of Computer Science, University of Jaén, Jaén, Spain dgzamora,martin@ujaen.es
[3] Department of Computer Science, University of Oviedo, Oviedo, Spain sirene@uniovi.es

Abstract. In the field of statistics, linear combinations of order statistics, also known as L-statistics, have been widely used for the estimation of the mean of a population, which is equivalent to considering Ordered Weighted Averaging (OWA) operators over simple random samples. If previous data are available or the distribution of the deviation from the mean is known, it is possible to compute optimal OWA weights that minimize the Mean Squared Error of the estimation. However, the optimal weights can only be used for a specific sample size, while in real Statistics the number of values that must be aggregated may change. In order to overcome this limitation, this contribution proposes a method based on the use of the recently defined Extreme Value Reductions (EVRs) to fit the cumulative optimal OWA weights and then use these EVRs to compute new weights for a different sample size. In addition, theoretical and simulated results are provided to show that, if sample sizes that are similar to the original one are considered, the weights generated by using EVRs are also similar to the optimal ones.

Keywords: Mean Estimation · EVR-OWA operator · Extreme Values Reduction · Flexible Sample Size

1 Introduction

Estimating the mean of a population is a classical problem in statistics [21]. One of the approaches that has been considered in the literature consists of using linear combinations of order statistics, or L-statistics, in which the values of the random sample are multiplied by a weight depending on their position when

* This research has been partially supported by the Spanish Ministry of Science and Technology (TIN-2017-87600-P and PGC2018-098623-B-I00), the Spanish Ministry of Economy and Competitiveness (PGC2018-099402-B-I00) and by the Spanish Ministry of Universities (FPU2019/01203).

sorted from the lower to the higher one, and then added. This topic of statistics has been developed from 1952 [17], to the most recent contributions [9,16]. From the aggregation theory point of view, this linear combination of order statistics is equivalent to applying an Ordered Weighted Averaging (OWA) operator, see [4], to the random sample.

In some cases, the distribution of the deviation from the mean is known, or there are real data available that can be used to compute the optimal weights to minimize the Mean Squared Error (MSE) when estimating the mean using order statistics. In these situations, this optimal weighting is expected to be combined with a new random sample to obtain the best possible estimation. However, in some cases, the size of the new random sample could be different from the size of the former sample. One of the most notable examples is the case of censored samples [2,3,18], which are commonly applied in survival analysis [15], where the sample size can be reduced due to external factors. However, it is also possible that the sample size grows because of an increase in, for example, the frequency of the measure or the number of experts. For all of these cases, it is no longer possible to use the optimal weights determined for a specific sample.

In order to overcome this drawback, this contribution proposes a method for estimating the mean of a population with symmetric distribution based on the EVR-OWA operator introduced by García-Zamora et al. [10] which allow generating OWA weights for different values of the sample size. In particular, starting with optimal weights for a sample size, the cumulative weights are fitted using a family of Extreme Values Reductions (EVR) [11,10]. Subsequently, the weights for other sample sizes are computed by using the fitted EVR. Theoretical results that endorse this procedure are provided and the behavior of the method is explored by using simulated data from logistic and hyperbolic secant distributions.

The remainder of the paper is structured as follows. In Section 2, the main concepts and basic results involving mean estimation and the EVR-OWA operator are introduced. The use of the OWA operator for mean estimation is discussed in Section 3. The theoretical aspects regarding the convergence of the cumulative weights are included in Section 4. Section 5 is devoted to the definition and study of the behavior of the proposed procedure. Finally, the conclusions and some comments about future work are discussed in Section 6.

2 Preliminaries

In this section, we introduce the general concepts needed for understanding the contribution. In particular, we will show some basic definitions and results concerning mean estimation and the EVR-OWA operator.

2.1 Mean estimation

First, let us recall the basic concepts about mean estimation based on order statistics. Rohatgi et al. [21] has been used as the main reference.

Consider a random variable X. We denote its density and cumulative distribution functions as f and F, respectively. The support of the variable, that is, $\{x \in \mathbb{R} \mid f(x) > 0\}$, is denoted as S. Now, consider a random sample, that is, a (finite) sequence of random variables $X_1, \ldots, X_n$ such that they are all independent and have the same distribution as X.

Even though the expression for the density and cumulative distribution function of X are known, it depends on one or more unknown parameters. If Θ denotes the set of possible values for the unknown parameter θ, an estimator of θ is a function of the random sample whose image is Θ.

Definition 1. *Let $X_1, \ldots, X_n$ be a sequence of independent and identically distributed (iid for short) random variables with density function f_θ depending on some unknown parameters $\theta \in \Theta$. An estimator is a measurable function $f : \mathbb{R}^n \to \Theta$ that does not depend on the value of the unknown parameters.*

In classical statistics, scholars and researchers have defined and studied the desirable properties that an estimator for a certain parameter should satisfy. Here, we are going to focus on unbiasedness and efficiency.

Definition 2. *Let $X_1, \ldots, X_n$ be a sequence of random variables with the same density function f_θ depending on some unknown parameter $\theta \in \Theta$. An estimator T is called unbiased if $E[T] = \theta$ for any $\theta \in \Theta$.*

The efficiency regards on the Mean Squared Error (MSE) between two estimators for a parameter.

Definition 3. *Let $X_1, \ldots, X_n$ be a sequence of iid random variables with density function f_θ depending on the unknown parameter $\theta \in \Theta$ and T_1, T_2 two estimators of θ. It is said that T_1 is more efficient than T_2 if $MSE(T_1) \leq MSE(T_2)$ for any $\theta \in \Theta$ and exists $\theta_0 \in \Theta$ such that $MSE(T_1) < MSE(T_2)$.*

A relation between the bias and the efficiency of an estimator can be done by using the well-known Fréchet-Cramér-Rao inequality [6,8,20]. Since the EVR-OWA, and any other OWA operator, relies on the order of the aggregated values, when used over a random sample, we need to use the concept of order statistic.

Definition 4. *[21] Let $X_1, \ldots, X_n$ be a sequence of random variables. The function $X_{(k)}$ of $(X_1, \ldots, X_n)$ that takes the value k-th smaller value in each possible observation $(x_1, \ldots, x_n)$ of $(X_1, ..., X_n)$ is known as the k-th order statistic or the statistic of order k (of the sequence $X_1, \ldots, X_n$).*

The use of order statistics in estimation has been a classic research line in statistics [17,22,23,24] and continues to be an important topic today [1,7,9,13,16].

2.2 The EVR-OWA operator

This section provides a brief introduction to OWA operators [27,28] based on EVRs [10], which are essential to provide aggregation whose weights are positive, symmetric and prioritize the intermediate information.

Ordered Weighted Averaging Operators OWA operators were proposed to ensure that the importance of the aggregated values depends on their position with respect to the median value [27]. Formally:

Definition 5. *Let $w \in [0,1]^n$ be a weighting vector such that $\sum_{i=1}^{m} w = 1$. The OWA Operator $\Psi_w : [0,1]^n \to [0,1]$ associated to w is defined by:*

$$\Psi_w(\vec{x}) = \sum_{k=1}^{n} w_k x_{\sigma(k)} \ \forall \ \vec{x} \in [0,1]^n$$

where σ is a permutation of the n-tuple $(1,2,...,n)$ such that $x_{\sigma(1)} \geq x_{\sigma(2)} \geq ... \geq x_{\sigma(n)}$.

OWA operators generalize other aggregation functions [4]. For example, the weighting vector $w = (\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}) \in [0,1]^n$, produces the arithmetic mean, whereas the vectors $w = (1,0,...,0) \in [0,1]^n$ and $w = (0,...,0,1) \in [0,1]^n$ produce the maximum and the minimum operators, respectively.

It should be highlighted that OWA operators decreasingly order the elements to be aggregated, while order statistics are defined through an increasing order. However, when symmetric distributions are considered, these two definitions are equivalent.

Linear RIM quantifiers to compute OWA weights In order to define weights for OWA operators Yager [28] proposed the use of Fuzzy Linguistic Quantifiers [29]. Specifically, given a Regular Increasing Monotonous (RIM) quantifier, namely an increasing function $Q : [0,1] \to [0,1]$ such that $Q(0) = 0$ and $Q(1) = 1$, the weights for an OWA operator to aggregate $n \in \mathbb{N}$ elements were computed as follows:

$$w_k = Q\left(\frac{k}{n}\right) - Q\left(\frac{k-1}{n}\right) \quad \text{for } k = 1,2,...,n.$$

Note that the final values of the weights strongly depend on the choice of a suitable linguistic quantifier. One of the most widely extended choices [14,19] is the linear RIM quantifier $Q_{\alpha,\beta} : [0,1] \to [0,1]$, $0 \leq \alpha < \beta \leq 1$ defined by:

$$Q_{\alpha,\beta}(x) = \begin{cases} 0 & 0 \leq x < \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \alpha \leq x \leq \beta \\ 1 & x \geq \beta \end{cases},$$

which allow modifying the importance of the intermediate values by changing the values of α and β.

The EVR-OWA operator In order to overcome some limitations of the linear RIM quantifier, García-Zamora et al. proposed the use of Extreme Values Reductions (EVRs) [10] as RIM linguistic quantifiers:

Definition 6. *[11] Let* $\hat{D} : [0,1] \to [0,1]$ *be a function satisfying:*

1. $\hat{D}$ *is an automorphism in the interval* $[0,1]$,
2. $\hat{D}$ *is a function of class* $\mathcal{C}^1$,
3. $\hat{D}$ *satisfies* $\hat{D}(x) = 1 - \hat{D}(1-x)\ \forall\ x \in [0,1]$,
4. $\hat{D}'(0) < 1$ *and* $\hat{D}'(1) < 1$,
5. $\hat{D}$ *is convex in a neighborhood of* 0 *and concave in a neighborhood of* 1,

then $\hat{D}$ *will be called Extreme Values Reduction (EVR) in the interval* $[0,1]$.

The main property of such functions is the fact that they reduce distances between the most extreme values of the interval $[0,1]$ whereas increase the distances between the intermediate values [11] (see Fig. 1).

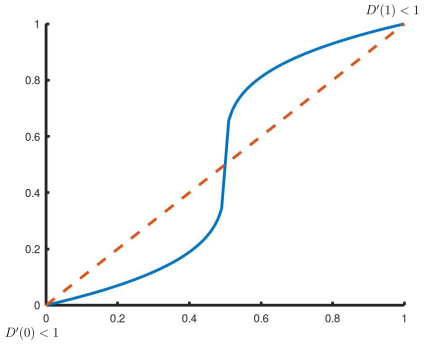

Fig. 1. Scheme of an EVR

For instance, some examples of EVRs are the functions $s_\alpha : [0,1] \to [0,1]$, $\alpha \in]0, \frac{1}{2\pi}]$ defined by

$$\hat{s}_\alpha(x) = x + \alpha \cdot \sin(2\pi x - \pi)\ \forall\ x \in [0,1]$$

and the polynomial functions $p_\alpha : [0,1] \to [0,1]$, $\alpha \in]0,1]$ defined as

$$p_\alpha(x) = (1-\alpha)x + 3\alpha x^2 - 2\alpha x^3\ \forall\ x \in [0,1].$$

Consequently, the EVR-OWA operator was defined as an OWA operator whose weights were computed by using an EVR [10]:

Definition 7. *Let* $\hat{D}$ *be an Extreme Values Reduction and consider* $n \in \mathbb{N}$. *Then, the family* $W = \{w_1, w_2, ..., w_n\}$, *where*

$$w_k = \hat{D}\left(\frac{k}{n}\right) - \hat{D}\left(\frac{k-1}{n}\right)\ \forall\ k \in \{1, 2, ..., n\},$$

receives the name of order n *weights associated with the EVR* $\hat{D}$, *and the OWA operator* $\Psi_{\hat{D}}$ *defined with respect to these weights.*

The latter procedure for computing the weights of the OWA operator can be seen as a particular example of an Extended Ordered Weighted Averaging (EOWA), we refer to [5], for which the weighting triangles are computed using an EVR.

3 OWA operator for mean estimation

Regarding the concepts of the previous section, applying an OWA operator to a random sample is equivalent to making a weighted average of the order statistics, i.e. an L-statistic. In this section, we will explore the use of the OWA operator as an estimator when there is symmetric noise.

Suppose that a quantity of interest takes the value μ. When measuring this quantity, a symmetric noise, i.e., a random variable with mean 0 such as $f(x) = f(-x)$ for any $x \in \mathbb{R}$, is added to the measure. Repeating the same measure gives us a random sample $X_1, \ldots, X_n$ in which all the variables have mean μ and are symmetric.

In this context, we may want to use an OWA operator to estimate the value of μ. However, we must choose the weighting vector. A common criterion in statistics is to minimize the Mean Squared Error (MSE). Let us consider the order statistics vector $\vec{Z} = \left(X_{(1)}, \ldots, X_{(n)}\right)$ and denote as $\Sigma = \mathrm{Var}\left[\vec{Z}\right]$ the covariance matrix of $\vec{Z}$ and as $\vec{\Delta} = E\left[\vec{Z}\right] - \mu\vec{1}$ the mean drift from μ of the components of $\vec{Z}$.

By using the basic properties of linear combinations of random variables (see [21]), the MSE to estimate μ has the following expression

$$E\left[\left(\mu - \Psi_w\left(\vec{X}\right)\right)^2\right] = \vec{w}'\left(\Sigma + \vec{\Delta}\vec{\Delta}'\right)\vec{w}.$$

From this expression, computing the optimal weights is equivalent to solving an optimization problem, for instance using Lagrange's multipliers procedure.

Proposition 1. *Let $X_1, \ldots, X_n$ a random sample in which any variable has mean μ. Then, the weighting vector $\vec{w}$ (verifying that $\sum_{i=1}^{n} w_i = 1$, $w_i \geq 0, i = 1, 2, ..., n$) which minimize $E\left[\left(\mu - \Psi_w(\vec{x})\right)^2\right]$ is*

$$\vec{w} = \frac{\left(\Sigma + \vec{\Delta}\vec{\Delta}'\right)^{-1}\vec{1}}{\vec{1}'\left(\Sigma + \vec{\Delta}\vec{\Delta}'\right)^{-1}\vec{1}}.$$

Notice that we allow the weights to have a negative value. Although this does not coincide with the classical definition of the OWA weights, for our approach, this is a desirable flexibilization in the definition of the operator. Firstly, we are making greater the feasible region, thus the result is at least as good as in the positive weights case. Secondly, the closed expression of Proposition 1, only

achievable by allowing negative weights, eases the computations in the main result of the next section. We also remark here that, even though it is possible to construct examples where negative weights appear, in most cases, among which are the most widely used distributions, all the weights are positive.

Therefore, we have optimal weights that depend on the distribution of the noise and the size of the random sample considered. Notice that multiplying the noise by a factor α does not change the optimal weights, since the only change would be a factor α^2 multiplying $\Sigma + \vec{\Delta}\vec{\Delta}'$. In Figure 2, the simulated optimal weights for the Logistic and Hyperbolic secant distributions, when $n = 20$, are presented. The density functions, respectively f_{L} and f_{Hs}, of these distributions are as follows:

$$f_{\mathrm{L}}(x) = \frac{1}{4\sigma}\mathrm{sech}^2\left(\frac{x-\mu}{2\sigma}\right) \ \left(\mu \in \mathbb{R}, \sigma \in \mathbb{R}^+\right), \qquad f_{\mathrm{Hs}}(x) = \frac{1}{2}\mathrm{sech}\left(\frac{\pi x}{2}\right),$$

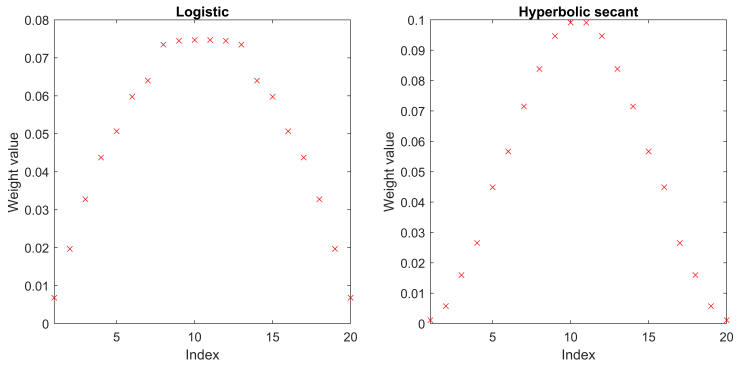

Fig. 2. Optimal weights for the Logistic and Hyperbolic secant distribution when $n = 20$.

As we indicated in the Introduction, many real-life problems require dealing with a non-fixed random sample size. In these cases, even we have an expression of our optimal weights, which could be computed from previous data, we cannot apply the OWA operator to our new data because the number of aggregated values changes. In this direction, one may wonder if there exists any relation between optimal weights when having the same distribution but different sample sizes. If we can find a connection, we can use the optimal weights initially calculated for a specific sample size to calculate a suitable weighting vector for a different value of n.

However, it is difficult to compare weighting vectors with different length. To fix that, we can follow the same idea that is used in the generation of weights presented in Subsection 2.2, but in the other direction. Given a weighting vector $\vec{w}$, let us define a cumulative weight function $W : \left\{0, \frac{1}{n}, \ldots, \frac{n-1}{n}, 1\right\} \to \mathbb{R}$ such

that:

$$W\left(\frac{k}{n}\right) = \sum_{i=1}^{k} w_i$$

Surprisingly, when we represent the cumulative weight functions for different but close values of n, using the distributions considered in Figure 2, we can see that the points seem to distribute in a common line (see Figure 3).

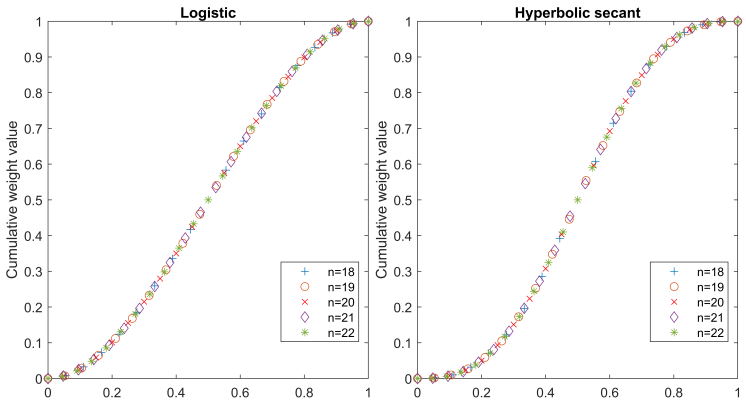

Fig. 3. Cumulative weights for the Logistic and Hyperbolic secant distribution when $n \in \{18, 19, 20, 21, 22\}$.

4 Convergence of cumulative weights

In this section, the behavior shown in Figure 3 is used as inspiration to define flexible OWA operators for the mean estimation. In particular, we fit the cumulative weights with a function f and then, if necessary, generate new weights as $w_i = f\left(\frac{i}{n}\right) - f(\frac{i-1}{n})$, $i \in \{1, \ldots, n\}$. However, it is necessary to state a theoretical result that sustain this procedure. In this section, we will give a result in this regard by proving that, if the distribution is sufficiently regular, then the cumulative weight points converge to a function on the unit interval. The most easy example is the Gaussian distribution. In this case, since the optimal weights are the balanced ones [17], then the cumulative weights are always over the graph of the identity function defined over the unit interval.

Before proving the main result, let us prove a useful lemma that allows to ease the computations when the distribution is symmetric.

Lemma 1. *Let $X_1, \ldots, X_n$ a sequence of idd random variables with symmetric distribution and mean μ. Then, the weighting vector $\vec{w}$ (verifying that $\sum_{i=1}^{n}$,*

$w_i = 1, w_i \geq 0\ i = 1, 2, ..., n)$ *which minimizes* $E\left[\left(\mu - \Psi_w(\vec{X})\right)^2\right]$ *also minimizes* $Var\left[\Psi_w(\vec{X})\right]$.

Proof. Since the distribution is symmetric, we have that $Cov\left[X_{(i)}, X_{(j)}\right] = Cov\left[X_{(n-i+1)}, X_{(n-j+1)}\right]$, $E[X_{(i)}] - \mu = \mu - E[X_{(n-i+1)}]$ and $E\left[X_{\left(\frac{n}{2}\right)}\right] = \mu$ (if n is even) for any $i, j \in \{1, \ldots, n\}$. Thus, Σ is a persymmetric matrix (see [12]) and $\vec{\Delta}$ holds $\Delta_i = -\Delta_{n-i+1}$ and $\Delta_{\frac{n}{2}} = 0$ (if n is even).

By performing the same procedure as in Proposition 1, the weights that minimize the variance are

$$\vec{w} = \frac{\Sigma^{-1}\vec{1}}{\vec{1}'\Sigma^{-1}\vec{1}},$$

and since the inverse of a persymmetric matrix is persymmetric [12], the resultant weights hold $w_i = w_{n-i+1}$ for any $i \in \{1, \ldots, n\}$. The result follows by noticing that $\vec{w}'\vec{\Delta} = 0$ ■

In conclusion, since we are considering symmetric distributions, the optimal weights depend only on Σ. Since we want to find an expression in the limit when $n \to \infty$, we should study the asymptotic behavior of Σ.

Lemma 2. *[25,26] Let* $X_1, \ldots, X_n$ *be a sequence of iid random variables with density function* f *and cumulative distribution* F *such that* f *is continuous and strictly positive in* $F^{-1}((0,1))$ *and there exists* $\epsilon > 0$ *such that*

$$\lim_{x \to \infty} |x|^{\epsilon} \left[1 - F(x) + F(-x)\right] = 0.$$

Then, for any $\delta > 0$ *and* $p, q \in [\delta, 1-\delta], p \leq q$*:*

$$\lim_{n \to \infty} (n+2) Cov\left(X_{(nq)}, X_{(np)}\right) = \frac{(1-p)q}{f\left(F^{-1}(p)\right) f\left(F^{-1}(q)\right)}$$

uniformly.

Remark 1. Note that Σ^{-1} may be heuristically approximated when n goes to infinity as $\Sigma^{-1} \sim (n+1)(n+2)DQD$ [25], where D is a diagonal matrix that satisfies $D_{i,i} = f\left(F^{-1}(\frac{i}{n+1})\right)$ for any $i \in \{1, \ldots, n\}$ and $Q = (q_{ij})$ is the matrix

$$Q = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & \cdots & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}$$

In the next result, we give sufficient conditions that ensure the convergence of the cumulative weights and also state the expression of the limit.

Theorem 1. *Let $X_1, \ldots, X_n$ be a sequence of iid random variables with support S and symmetric distribution whose density function and cumulative distribution are respectively denoted as f and F. Let us assume that f is bounded, continuous, twice differentiable, and strictly positive on $F^{-1}((0,1))$. Suppose that:*

- *There exists a sequence $k(n)$ such that $\frac{k(n)}{n^3} \to \infty$ satisfying that for any $\delta > 0$ and $p, q \in [\delta, 1-\delta], p \leq q$*

$$\lim_{n\to\infty} k(n)(n+1)^2 \left((n+2) f\left(F^{-1}(p)\right) f\left(F^{-1}(q)\right) \Sigma_{np,nq} - p(1-q)\right) = 0$$

  uniformly,
- $\int_0^1 f\left(F^{-1}(x)\right) \left(\frac{d^2}{dx^2} f\left(F^{-1}(x)\right)\right) dx < \infty$.

Then, for any $q \in [0,1] \cup \mathbb{Q}$ with irreducible fraction $\frac{a}{b}$:

$$\lim_{n\to\infty} W^{(nb)}(q) = \begin{cases} \frac{L + \int_0^q f\left(F^{-1}(x)\right)\left(\frac{d^2}{dx^2} f\left(F^{-1}(x)\right)\right) dx}{2L + \int_0^1 f(F^{-1}(x))\left(\frac{d^2}{dx^2} f(F^{-1}(x))\right) dx} & \text{if } \lim\limits_{x \to \inf S} \frac{f(x)^2}{F(x)} = L < \infty \\ \frac{1}{2} & \text{otherwise} \end{cases}.$$

Proof. For the sake of simplicity, we provide here a sketch of the proof because developing the necessary computations would require several pages. Let us denote the inverse of the covariance matrix of the order statistics of dimension n as $\Sigma(n)^{-1}$. Consider the following sequence:

$$W^{(nb)}(q) = \frac{\sum_{i=1}^{na} \sum_{j=1}^{nb} \left(\Sigma(nb)^{-1}\right)_{i,j}}{\sum_{i=1}^{nb} \sum_{j=1}^{nb} \left(\Sigma(nb)^{-1}\right)_{i,j}}.$$

The imposed convergence conditions guarantee that the following equality holds:

$$\lim_{n\to\infty} \sum_{i=1}^{na} \sum_{j=1}^{nb} (DQD)_{i,j} = \lim_{n\to\infty} \sum_{i=1}^{na} \sum_{j=1}^{nb} \left(\Sigma(nb)^{-1}\right)_{i,j},$$

where D and Q are the matrices defined in Remark 1.

By substituting the expression of the elements of DQD and using the symmetry of the distribution, the following expression is obtained:

$$\lim_{n\to\infty} W^{(nb)}(q) = \lim_{n\to\infty} \frac{f\left(F^{-1}\left(\frac{1}{n}\right)\right)^2 + \frac{1}{n}\int_0^q f\left(F^{-1}(x)\right)\left(\frac{d^2}{dx^2} f\left(F^{-1}(x)\right)\right) dx}{2f\left(F^{-1}\left(\frac{1}{n}\right)\right)^2 + \frac{1}{n}\int_0^1 f\left(F^{-1}(x)\right)\left(\frac{d^2}{dx^2} f\left(F^{-1}(x)\right)\right) dx}.$$

Note that the limit of the first term of the numerator multiplied by n is equivalent to $\lim_{x\to\inf S} \frac{f(x)^2}{F(x)} = L$. If this limit converges, we have the first case

of the Theorem. If the limit diverges, the integral terms are negligible, and the limit is 0.5. Due to the continuity of f, the sequence must be convergent (it cannot be oscillatory). ■

Although these conditions may be too restrictive, and some of them, as the fast convergence of the moments of the ordered statistics, are hard to check, in our numerical experiments the convergence holds for all the considered distributions. Moreover, even if we simulate distributions that do not satisfy some condition (for instance, the density function of the Laplace distribution is not differentiable), the convergence still holds.

5 EVR-OWA operator as mean estimator for symmetric distributions

In this section, we define a method to fit the cumulative weights associated with symmetric distributions based on EVRs. We also consider here the limit case $\hat{D}(x) = x \ \forall \ x \in [0,1]$, which corresponds to the balanced weights associated with Gaussian distribution. In this case, we consider a family of EVRs consisting of functions of the form:

$$\hat{D}_{\alpha,\beta,\lambda} = \lambda s_\alpha + (1-\lambda) p_\beta$$

where $\alpha \in \left[0, \frac{1}{2\pi}\right]$, $\beta, \lambda \in [0,1]$.

This family consists of convex combinations of EVRs of the families s_α and p_α, defined in Section 2.2, and the limit case corresponding to the identity function. This family has been considered because it has a good behavior regarding the logistic and hyperbolic secant distributions, but it can be extended to a more wide family if needed.

We have applied the latter family to fit the cumulative weights, when $n = 20$, for the logistic and hyperbolic secant distributions. The results are shown in Figure 4, in addition to the cumulative weights for $n = 21$ and $n = 19$.

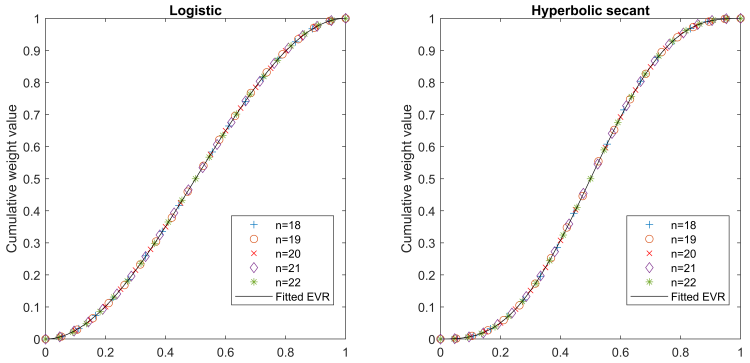

Fig. 4. EVRs fitted to the cumulative weights when $n = 20$ and cumulative weights for $n = \{18, 19, 20, 21, 22\}$ for the Logistic and Hyperbolic secant distributions.

In particular, the optimal parameters are, in the case of the logistic distribution, $\alpha = 0.1592, \beta = 1, \lambda = 0.0325$ and, in the case of the hyperbolic secant distribution, $\alpha = 0.1591, \beta = 0.7361, \lambda = 0.9730$. As it is shown in Figure 4, the fit seems reasonable not only for $n = 20$ but also for $n = 19$ and $n = 21$. We have also computed the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) of the fit for the three values of n, which can be consulted in Table 1.

Distribution	Sample size	RMSE $\left(\times 10^{-3}\right)$	MAE $\left(\times 10^{-3}\right)$
Logistic	18	0.94	0.72
	19	1.19	0.92
	20	0.80	0.62
	21	0.83	0.66
	22	0.94	0.73
Hyperbolic secant	18	1.59	1.19
	19	1.14	0.89
	20	0.60	0.37
	21	1.19	0.98
	22	0.98	0.66

Table 1. Root Squared Mean and Mean Absolute Errors of the EVR fitted for the case $n = 20$ when $n \in \{19, 20, 21\}$ for the logistic and hyperbolic secant distributions.

As we can see, the RMSE and MAE are lower when $n = 20$, as expected because we have fitted the EVRs using these points. However, the increase when the value of n is changed is not too high, and it seems reasonable to approximate the constructed EVR-OWA for $n \in \{18, 19, 21, 22\}$, for the considered distributions. Qualitatively, we expect that both RMSE and MAE increase as the difference between n and 20 increases.

Comparing both distributions, the fit for the hyperbolic secant distribution seems to be better than the one for the logistic distribution when $n = 20$, but the RMSE and MAE increase more when moving to $n = 19$ or $n = 21$.

6 Conclusions and future work

In this contribution, a method for constructing an EVR-OWA operator as a mean estimator for symmetric distributions that allow changes in the sample size has been discussed. First, optimal weights regarding Proposition 1 are computed, using simulated or real data with a fixed sample size n. Then, a family of EVRs is used to fit the cumulative weights. Finally, we use the fitted function to generate weights associated to another sample sized different to n.

In order to justify the use of this method, we have presented Theorem 1, which states that when n goes to infinity, under some particular conditions, the cumulative weights converge.

The method have been illustrated using the Logistic and Hyperbolic secant distributions for $n = 20$, considering the convex linear combination of sinusoidal and polynomial EVRs, see Figure 4. Keeping in mind the RMSE and the MAE of the fit (see Table 1), we conclude that we have to obtain reasonably adequate results when considering $n = 19$ or $n = 21$.

As future work, we want to extend the study to non-symmetric distributions and also to real data. In this regard, the EVR functions are too limited and we need a more general family of functions defined over the unit interval. From a theoretical point of view, we need to extend Theorem 1 for non-symmetric distributions, and we also wonder if some distribution requirements can be relaxed.

References

1. Ahsanullah, M., Alzaatreh, A.: Parameter estimation for the log-logistic distribution based on order statistics. REVSTAT **16**(4), 429–443 (2018)
2. Almongy, H.M., Almetwally, E.M., Alharbi, R., Alnagar, D., Hafez, E.H., Mohie El-Din, M.M.: The weibull generalized exponential distribution with censored sample: estimation and application on real data. Complexity **2021** (2021)
3. Alzeley, O., Almetwally, E.M., Gemeay, A.M., Alshanbari, H.M., Hafez, E., Abu-Moussa, M.: Statistical inference under censored data for the new exponential-x fréchet distribution: Simulation and application to leukemia data. Computational Intelligence and Neuroscience **2021** (2021)
4. Beliakov, G., Bustince, H., Calvo, T.: A practical guide to averaging functions. Springer (01 2016)
5. Calvo, T., Mayor, G., Torrens, J., Suñer, J., Mas, M., Carbonell, M.: Generation of weighting triangles associated with aggregation functions. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **8**(04), 417–451 (2000)
6. Cramér, H.: A contribution to the theory of statistical estimation. Scandinavian Actuarial Journal **1946**(1), 85–94 (1946)
7. Dytso, A., Cardone, M., Veedu, M.S., Poor, H.V.: On estimation under noisy order statistics. In: 2019 IEEE International Symposium on Information Theory (ISIT). pp. 36–40. IEEE (2019)
8. Fréchet, M.: Sur l'extension de certaines évaluations statistiques au cas de petits échantillons. Revue de l'Institut International de Statistique pp. 182–205 (1943)
9. Gao, W., Zhang, T., Yang, B.B., Zhou, Z.H.: On the noise estimation statistics. Artificial Intelligence **293**, 103451 (2021)
10. García-Zamora, D., Labella, Á., Rodríguez, R.M., Martínez, L.: Symmetric weights for owa operators prioritizing intermediate values. the evr-owa operator. Information Sciences **584**, 583–602 (2022)
11. García-Zamora, D., Labella, Á., Rodríguez, R.M., Martínez, L.: Non Linear Preferences in Group Decision Making. Extreme Values Amplifications and Extreme Values Reductions. International Journal of Intelligent Systems (2021)
12. Golub, G.H., Van Loan, C.F.: Matrix computations. Johns Hopkins studies in the mathematical sciences. Johns Hopkins University Press, Baltimore (1996)
13. Hassan, A.S., Abd-Allah, M.: Exponentiated weibull-lomax distribution: properties and estimation. Journal of Data Science **16**(2), 277–298 (2018)
14. Herrera-Viedma, E., Herrera, F., Chiclana, F.: A consensus model for multiperson decision making with different preference structures. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on **32**, 394 – 402 (06 2002)

15. Klein, J.P., Moeschberger, M.L.: Survival analysis: techniques for censored and truncated data, vol. 1230. Springer, New York (2003)
16. Kumar, D., Kumar, M., Joorel, J.S.: Estimation with modified power function distribution based on order statistics with application to evaporation data. Annals of Data Science pp. 1–26 (2020)
17. Lloyd, E.: Least-squares estimation of location and scale parameters using order statistics. Biometrika **39**(1/2), 88–95 (1952)
18. Narisetty, N., Koenker, R.: Censored quantile regression survival models with a cure proportion. Journal of Econometrics **226**(1), 192–203 (2022)
19. Palomares, I., Estrella, F.J., Martínez, L., Herrera, F.: Consensus under a fuzzy context: Taxonomy, analysis framework afryca and experimental case of study. Information Fusion **20**, 252–271 (2014)
20. Radhakrishna Rao, C.: Information and accuracy attainable in the estimation of statistical parameters. Bulletin of the Calcutta Mathematical Society **37**(3), 81–91 (1945)
21. Rohatgi, V.K., Saleh, A.M.E.: An introduction to probability and statistics. John Wiley & Sons, Hoboken, New Jersey (2015)
22. Sarhan, A.: Estimation of the mean and standard deviation by order statistics. part ii. The Annals of Mathematical Statistics **26**(3), 505–511 (1955)
23. Sarhan, A.: Estimation of the mean and standard deviation by order statistics. part iii. The Annals of Mathematical Statistics pp. 576–592 (1955)
24. Sarhan, A.E.: Estimation of the mean and standard deviation by order statistics. The Annals of Mathematical Statistics pp. 317–328 (1954)
25. Stephens, M.: Asymptotic calculations of functions of expected values and covariances of order statistics. Canadian Journal of Statistics **18**(3), 265–270 (1990)
26. Stigler, S.M.: Linear functions of order statistics. The Annals of Mathematical Statistics **40**(3), 770–788 (1969)
27. Yager, R.: Families of OWA operators. Fuzzy Sets and Systems **59**(2), 125 – 148 (1993)
28. Yager, R.: Quantifier guided aggregation using OWA operators. International Journal of Intelligent Systems **11**(1), 49–73 (1996)
29. Zadeh, L.: A computational approach to fuzzy quantifiers in natural languages. Computers & Mathematics with Applications **9**, 149–184 (12 1983)