Article

Stochastic Order and Generalized Weighted Mean Invariance

Mateu Sbert [1,*] **, Jordi Poch** [1] **, Shuning Chen** [2] **and Víctor Elvira** [3]

[1] Graphics and Imaging Laboratory, University of Girona, 17003 Girona, Spain; jordi.poch@udg.edu
[2] Department of Civil Engineering, Urban Institute, School of Engineering, Kyushu University, Fukuoka 819-0395, Japan; njushuningchen@hotmail.com
[3] School of Mathematics, University of Edinburgh, Edinburgh EH9 3FD, UK; victor.elvira.arregui@gmail.com
[*] Correspondence: mateu@ima.udg.edu

Abstract: In this paper, we present order invariance theoretical results for weighted quasi-arithmetic means of a monotonic series of numbers. The quasi-arithmetic mean, or Kolmogorov–Nagumo mean, generalizes the classical mean and appears in many disciplines, from information theory to physics, from economics to traffic flow. Stochastic orders are defined on weights (or equivalently, discrete probability distributions). They were introduced to study risk in economics and decision theory, and recently have found utility in Monte Carlo techniques and in image processing. We show in this paper that, if two distributions of weights are ordered under first stochastic order, then for any monotonic series of numbers their weighted quasi-arithmetic means share the same order. This means for instance that arithmetic and harmonic mean for two different distributions of weights always have to be aligned if the weights are stochastically ordered, this is, either both means increase or both decrease. We explore the invariance properties when convex (concave) functions define both the quasi-arithmetic mean and the series of numbers, we show its relationship with increasing concave order and increasing convex order, and we observe the important role played by a new defined mirror property of stochastic orders. We also give some applications to entropy and cross-entropy and present an example of multiple importance sampling Monte Carlo technique that illustrates the usefulness and transversality of our approach. Invariance theorems are useful when a system is represented by a set of quasi-arithmetic means and we want to change the distribution of weights so that all means evolve in the same direction.

Keywords: stochastic order; weighted mean; arithmetic mean; harmonic mean; Kolmogorov mean; quasi-arithmetic mean; cross-entropy; Shannon entropy; Rényi entropy; Tsallis entropy; multiple importance sampling

1. Introduction and Motivation

Stochastic orders [1,2], are orders defined in probability theory and statistics, to quantify the concept of one random variable being *bigger* or *smaller* than another one. Discrete probability distributions, also called probability mass functions (pmf), are sequences of n-tuples of non-negative values that add up to 1, and can be thus interpreted in several ways, for instance, as weights in computation of moments of the discrete random variable described by the pmf, or as equivalence classes of compositional data [3]. Stochastic orders have found application in decision and risk theory [4], and in economics in general, among many other fields [2]. Some stochastic orders have been defined based on order invariance: two pmf's are ordered when the arithmetic means of any increasing sequence of real numbers weighted with the corresponding pmf's are ordered in the same direction. This raises the question whether this invariance might also hold for other kind of means beyond the arithmetic mean.

The quasi-arithmetic means, also called Kolmogorov or Kolmogorov–Nagumo means, are ubiquitous in many branches of science [5]. They have the expression $g^{-1}\left(\sum_{k=1}^{M} \alpha_k g(b_k)\right)$, where $g(x)$ is a real-valued strictly monotonous function, $\{b_k\}_{k=1}^{M}$ a sequence of reals, and

$\{\alpha_k\}_{k=1}^M$ a set of weights with $\sum_{k=1}^M \alpha_k = 1$. This family of means comprises the usual means: arithmetic $g(x) = x$, $\sum_{k=1}^M \alpha_k b_k$, harmonic $g(x) = 1/x$, $\left(\sum_{k=1}^M \frac{\alpha_k}{b_k} \right)^{-1}$, power mean $g(x) = x^p$, $\left(\sum_{k=1}^M \alpha_k b_k^p \right)^{\frac{1}{p}}$. For a long time, economists have discussed the best mean for a problem [6]. Harmonic mean is used for the price earning ratio, and power means are used to represent the aggregate labor demand and its corresponding wage [7], and the constant elasticity of substitution (CES) [8]. Yoshida [9,10] has studied the invariance under quasi-arithmetic means with function $g(x)$ increasing and for utility functions. In information theory, Alfred Rényi [11] defined axiomatically the entropy of a probability distribution as a Kolmogorov mean of the information $1/\log p_k$ conveyed by result k with probability p_k, and recently, Americo et al. [12] defined conditional entropy based on quasi-arithmetic means. In physics, the equivalent spring constant of springs combined in series is obtained as the harmonic mean of the individual spring constants, and in parallel, as their arithmetic mean [13], while the equivalent resistance of resistors combined in parallel is obtained by the harmonic mean of the individual resistances, and in series by their arithmetic mean [14]. In traffic flow [15], arithmetic and harmonic mean of speed distribution are used. In [16] both geometric and harmonic mean are used in addition to arithmetic mean to improve noise source maps.

In our recent work on inequalities for generalized, quasi-arithmetic weighted means [17], we found some invariance properties, that depended on the particular relationship considered between the sequences of weights. These relationships between weights define first stochastic order, and likelihood ratio order. Their application to multiple importance sampling (MIS), a Monte Carlo technique, has been presented in [18], its application to cross entropy in [19], and in [20] applications to image processing, traffic flow and income distribution have been shown. In [21], the invariance results on products of distributions of independent scalar r.v.'s [22] was generalized to any joint distribution of a 2-dimensional r.v.

In this paper, we show that the order invariance is a necessary and sufficient condition for first stochastic order, and that it holds under any quasi-arithmetic mean. We also study invariance under the second stochastic order, likelihood ratio, hazard-rate, and increasing convex stochastic orders. The fact that the invariance results hold for both increasing and decreasing monotonic functions allows us to use both utilities and liabilities, represented by efficiencies and expected error, respectively, to look for an optimal solution, where in liabilities we look for minimum expected error, while in utilities for maximum efficiency.

The rest of the paper is organized as follows. In Section 2, we introduce the stochastic order; in Section 3, the arithmetic mean and its relationship with stochastic order. In Section 4, we present the invariance theorems; in Section 5, we discuss the invariance for concave (convex) functions; in Section 6, its application to stochastic orders; in Section 7, we present an example based on the linear combination of Monte Carlo estimators. Finally, conclusions and future work are given in Section 8.

2. Stochastic Orders

Stochastic orders are pre-orders (i.e, binary relationships holding symmetric and transitive properties) defined on probability distributions with finite support. Note that equivalently, one can think of sequences (i.e., ordered sets) of non-negative weights/values that sum up to one. Observe that any sequence $\{\alpha_k\}$ of M positive numbers such that $\sum_{k=1}^M \alpha_k = 1$ can be considered a probability distribution. It can be seen too as an element of the (M-1)-Simplex, i.e., $\{\alpha_k\} \in \Delta_{M-1}$. While several interpretations hold and hence increase the range of applicability, in the remaining of this paper, we will talk of sequence without any loss of generality.

Notation. *We use the symbols $\prec, \succ$ to represent orders between two sequences $\{\alpha_k\}$ and $\{\alpha'_k\}$ of size M, e.g., and we write $\{\alpha_k\} \succ \{\alpha'_k\}$ or equivalently $\{\alpha'_k\} \prec \{\alpha_k\}$. We will denote the elements of the sequences without the curly brackets, e.g., the first element of the sequence $\{\alpha_k\}$ is denoted as α_1, while the last one is α_M. Moreover, the first and last elements of the sequence receive special attention, since when $\alpha_1 \leq \alpha'_1$, we can write this order as $\{\alpha_k\} \succ_F \{\alpha'_k\}$ (where $\succ_F$ stands for*

first) and whenever $\alpha_M \geq \alpha'_M$, we can write $\{\alpha_k\} \succ_L \{\alpha'_k\}$ (where $\succ_L$ stands for last*). The case with both $\alpha_1 \leq \alpha'_1$ and $\alpha_M \geq \alpha'_M$, we can denote as $\{\alpha_k\} \succ_{FL} \{\alpha'_k\}$. The orders $\succ_F, \succ_L, \succ_{FL}$ are superorders of first stochastic dominance order, $\{\alpha_k\} \succ_{FSD} \{\alpha'_k\}$, that will be studied in Section 6. We denote as $\{\alpha_{M-k+1}\}$ the sequence with the same elements of $\{\alpha_k\}$ but in reversed order.*

Example 1 (Toy example). *Given sequences $\{\alpha_k\} = \{0.5, 0.1, 0.2, 0.2\}$, and $\{\alpha'_k\} = \{0.4, 0.1, 0.2, 0.3\}$, we have both $\{\alpha_k\} \prec_F \{\alpha'_k\}$, $\{\alpha_k\} \prec_L \{\alpha'_k\}$, and thus $\{\alpha_k\} \prec_{FL} \{\alpha'_k\}$. On the other hand for sequences $\{\alpha_k\}$ and $\{\alpha''_k\} = \{0.4, 0.3, 0.2, 0.1\}$, we have $\{\alpha_k\} \prec_F \{\alpha''_k\}$, $\{\alpha_k\} \nprec_L \{\alpha''_k\}$, and thus $\{\alpha_k\} \nprec_{FL} \{\alpha''_k\}$.*

Property 1 (Mirror property). *One desirable property of stochastic orders is such that when reversing the ordering of the sequence elements, the stochastic order should be reversed as well, i.e., if $\{\alpha_k\} \succ \{\alpha'_k\}$, then $\{\alpha'_{M-k+1}\} \succ \{\alpha_{M-k+1}\}$. This is similar to the invariance of physical laws to the right and left hand. We call this property the* mirror *property.*

Definition 1 (Mirror property). *We say that a stochastic order has the mirror property if*

$$\{\alpha'_k\} \prec \{\alpha_k\} \Rightarrow \{\alpha_{M-k+1}\} \prec \{\alpha'_{M-k+1}\}. \tag{1}$$

Observe that the simple orders defined before, $\succ_F, \succ_L$, do not hold this property, but $\succ_{FL}$ holds it. We will see in Section 6 that usual stochastic orders do hold the mirror property. However, an order that is insensitive to the permutation of the elements of a sequence, like majorization or Lorentz order, does not hold the mirror property.

3. Quasi-Arithmetic Mean

Stochastic orders are usually defined by invariance to arithmetic mean [1,2] (see Section 6), and we want to investigate in this paper invariance to more general means. We define here the kind of means we are interested in.

Definition 2 (Quasi-arithmetic or Kolmogorov or Kolmogorov–Nagumo mean). *A quasi-arithmetic weighted mean (or Kolmogorov mean) $\mathcal{M}(\{b_k\}, \{\alpha_k\})$ of a sequence of real numbers $\{b_k\}_{k=1}^M$, is of the form $g^{-1}\left(\sum_{k=1}^M \alpha_k g(b_k)\right)$, where $g(x)$ is a real valued, invertible strictly monotonic function, with inverse function $g^{-1}(x)$ and $\{\alpha_k\}_{k=1}^M$ are positive weights such that $\sum_{k=1}^M \alpha_k = 1$.*

Examples of such mean are arithmetic weighted mean ($g(x) = x$), harmonic weighted mean ($g(x) = 1/x$), geometric weighted mean ($g(x) = \log x$) and in general, the weighted power means ($g(x) = x^r$).

Given a distribution $\{p_k\}$, Shannon entropy, $\mathcal{H}(\{p_k\}) = -\sum_k p_k \log p_k$, and Rényi entropy, $\mathcal{R}_\beta(\{p_k\}) = \frac{1}{1-\beta} \log \sum_k p_k^\beta$, can be considered as quasi-arithmetic means of the sequence $\{\frac{1}{\log p_k}\}$ with weights $\{p_k\}$: Shannon entropy with $g(x) = x$ (arithmetic mean or expected value), and Rényi entropy with $g(x) = 2^{(\beta-1)x}$ [11]. Tsallis entropy, $\mathcal{S}_q(\{p_k\}) = \frac{1-\sum_{i=1}^M p_i^q}{q-1}$, can be considered the weighted arithmetic mean of the sequence $\{\frac{1}{\ln_q p_k}\}$ with weights $\{p_k\}$, where $\ln_q x = \frac{x^{1-q}-1}{1-q}$ is the q-logarithm function [23]. Without loss of generality, we consider from now on that $\{b_k = f(k)\}$ where $f(x)$ is a real valued function.

Lemma 1. *Consider the sequences of M positive weights $\{\alpha_k\}_{k=1}^M$ and $\{\alpha'_k\}_{k=1}^M$, $\sum_{k=1}^M \alpha_k = 1, \sum_{k=1}^M \alpha'_k = 1$, $g(x)$ a strictly monotonic function. The following conditions (a), (b), (a'), (b'), (c)*

and (d) are equivalent.

(a) for $g(x)$ increasing, for any increasing function $f(x)$ the following inequality holds:

$$\sum_{k=1}^{M} \alpha'_k g(f(k)) \leq \sum_{k=1}^{M} \alpha_k g(f(k)) \tag{2}$$

(b) for $g(x)$ increasing, for any increasing function $f(x)$ the following inequality holds:

$$\sum_{k=1}^{M} \alpha_{M-k+1} g(f(k)) \leq \sum_{k=1}^{M} \alpha'_{M-k+1} g(f(k)) \tag{3}$$

(a') for $g(x)$ decreasing, for any increasing function $f(x)$ the following inequality holds:

$$\sum_{k=1}^{M} \alpha_k g(f(k)) \leq \sum_{k=1}^{M} \alpha'_k g(f(k)) \tag{4}$$

(b') for $g(x)$ decreasing, for any increasing function $f(x)$ the following inequality holds:

$$\sum_{k=1}^{M} \alpha'_{M-k+1} g(f(k)) \leq \sum_{k=1}^{M} \alpha_{M-k+1} g(f(k)) \tag{5}$$

(c) the following inequalities hold:

$$
\begin{aligned}
\alpha'_1 &\geq \alpha_1 \\
\alpha'_1 + \alpha'_2 &\geq \alpha_1 + \alpha_2 \\
&\vdots \\
\alpha'_1 + \ldots + \alpha'_{M-1} &\geq \alpha_1 + \ldots + \alpha_{M-1} \\
\alpha'_1 + \ldots + \alpha'_{M-1} + \alpha'_M &= \alpha_1 + \ldots + \alpha_{M-1} + \alpha_M
\end{aligned} \tag{6}
$$

(d) the following inequalities hold:

$$
\begin{aligned}
\alpha'_M &\leq \alpha_M \\
\alpha'_M + \alpha'_{M-1} &\leq \alpha_M + \alpha_{M-1} \\
&\vdots \\
\alpha'_M + \ldots + \alpha'_2 &\leq \alpha_M + \ldots + \alpha_2 \\
\alpha'_M + \alpha'_{M-1} + \ldots + \alpha'_1 &= \alpha_M + \alpha_{M-1} + \ldots + \alpha_1
\end{aligned} \tag{7}
$$

If $f(x)$ is decreasing, the inequalities in a) through b') are reversed.

Proof. Indirect or partial proofs can be found in [17,20]. We provide a complete proof in the Appendix A. $\square$

Note. *Observe that in Lemma 1, it is sufficient to consider the monotonicity of the sequences $\{f(k)\}_{k=1}^{M}$. Furthermore, Lemma 1 can be extended to any real sequences $\{y_k\}$ and $\{x_k\}$, such that $\sum_{k=1}^{M} y_k = \sum_{k=1}^{M} x_k$. It is enough to observe that the order of all the inequalities are unchanged by adding a positive constant, so that $\{y_k\}$ and $\{x_k\}$ can be made positive, and also are unchanged by the multiplication of a positive constant, so that the resulting $\{y_k\}$ and $\{x_k\}$ sequences can be normalized.*

Theorem 1. *Given a mean $\mathcal{M}$ with strictly monotonic function $g(x)$ and two distributions $\{\alpha'_k\}, \{\alpha_k\}$, the following propositions are equivalent:*
(a) for all increasing functions $f(x)$

$$\mathcal{M}(\{f(k)\}, \{\alpha'_k\}) \leq \mathcal{M}(\{f(k)\}, \{\alpha_k\}) \tag{8}$$

(b) for all increasing functions $f(x)$

$$\mathcal{M}(\{f(k)\}, \{\alpha_{M-k+1}\}) \leq \mathcal{M}(\{f(k)\}, \{\alpha'_{M-k+1}\}) \tag{9}$$

(a') for all decreasing functions $f(x)$

$$\mathcal{M}(\{f(k)\}, \{\alpha'_k\}) \geq \mathcal{M}(\{f(k)\}, \{\alpha_k\}) \tag{10}$$

(b') for all decreasing functions $f(x)$

$$\mathcal{M}(\{f(k)\}, \{\alpha_{M-k+1}\}) \geq \mathcal{M}(\{f(k)\}, \{\alpha'_{M-k+1}\}) \tag{11}$$

(c) Condition (c) of Lemma 1 holds

Proof. It is a direct consequence of Lemma 1 and the definition of quasi-arithmetic mean, observing that the inverse $g^{-1}(x)$ of a strictly monotonic increasing (respectively decreasing) function $g(x)$ is also increasing (respectively decreasing). □

4. Invariance

Theorem 2 (Invariance). *Given two distributions $\{\alpha'_k\}, \{\alpha_k\}$, and two quasi-arithmetic means $\mathcal{M}, \mathcal{M}^\star$, the following propositions are equivalent:*
(a) for all increasing functions $f(x)$

$$\mathcal{M}(\{f(k)\}, \{\alpha'_k\}) \leq \mathcal{M}(\{f(k)\}, \{\alpha_k\}) \tag{12}$$

(b) for all increasing functions $f(x)$

$$\mathcal{M}^\star(\{f(k)\}, \{\alpha'_k\}) \leq \mathcal{M}^\star(\{f(k)\}, \{\alpha_k\}) \tag{13}$$

(a') for all decreasing functions $f(x)$

$$\mathcal{M}(\{f(k)\}, \{\alpha'_k\}) \geq \mathcal{M}(\{f(k)\}, \{\alpha_k\}) \tag{14}$$

(b') for all decreasing functions $f(x)$

$$\mathcal{M}^\star(\{f(k)\}, \{\alpha'_k\}) \geq \mathcal{M}^\star(\{f(k)\}, \{\alpha_k\}) \tag{15}$$

Proof. It is a direct consequence of the observation that conditions (c) and (d) in Lemma 1 do not depend on any particular $g(x)$ function considered, and thus the order of inequalities does not change with the mean considered as long as $\{\alpha_k\}$ and $\{\alpha'_k\}$ are kept fixed. □

The following properties relate stochastic order with the quasi-arithmetic mean. Let $\mathcal{I}$ be the set of monotonous functions, $\mathcal{I}^<$ the set of increasing functions, $\mathcal{I}^>$ the set of decreasing functions.

Definition 3 (preserve mean order property). *We say that a stochastic order preserves mean order for a given mean $\mathcal{M}$ and a set $I^< \subset \mathcal{I}^<$ of increasing functions when, for all functions $f(x) \in I^<$ and any distributions $\{\alpha'_k\}, \{\alpha_k\}$,*

$$\{\alpha_k\} \prec \{\alpha'_k\} \Rightarrow \mathcal{M}(\{f(k)\}, \{\alpha'_k\}) \leq \mathcal{M}(\{f(k)\}, \{\alpha_k\}). \tag{16}$$

Definition 4 (preserve inverse mean order property)**.** *We say that a stochastic order preserves inverse mean order for a given mean $\mathcal{M}$ and a set $I^> \subset \mathcal{I}^>$ of decreasing functions when, for all functions $f(x) \in I^>$ and any distributions $\{\alpha'_k\}, \{\alpha_k\}$,*

$$\{\alpha_k\} \prec \{\alpha'_k\} \Rightarrow \mathcal{M}(\{f(k)\}, \{\alpha'_k\}) \geq \mathcal{M}(\{f(k)\}, \{\alpha_k\}). \tag{17}$$

Theorem 2 together with the preserve mean order properties allows us to state the following invariance property:

Theorem 3 (preserve mean order invariance)**.** *Given a stochastic order that preserves mean order (respectively preserves inverse mean order) for a given mean and for $\mathcal{I}^<$ (respectively for $\mathcal{I}^>$), then for any mean it preserves both mean order for $\mathcal{I}^<$ and inverse mean order for $\mathcal{I}^>$. In other words, the preserve mean order properties are invariant with respect to the mean considered.*

Observe that from Lemma 1 and Theorems 1 and 3 , we have that a necessary and sufficient condition for an order to preserve mean order for $\mathcal{I}^<$ (or preserve inverse mean order for $\mathcal{I}^>$) is the holding of Equation (6), independently of the mean considered. We will see in Section 6 that this corresponds to first stochastic dominance order.

5. Concavity and Convexity

Let us consider now $\mathcal{I}_v^< \subset \mathcal{I}^<$, the set of all increasing concave functions, $\mathcal{I}_x^< \subset \mathcal{I}^<$, the set of all increasing convex functions, $\mathcal{I}_v^> \subset \mathcal{I}^>$, the set of all decreasing concave functions, and $\mathcal{I}_x^> \subset \mathcal{I}^>$, the set of all decreasing convex functions. The following theorem relates the preserve mean order properties with the mirror property.

Theorem 4. *If an order holds the mirror property, then holding preserve mean order for $\mathcal{I}_v^<$ (respectively $\mathcal{I}_x^<$) implies holding preserve inverse mean order for $\mathcal{I}_v^>$ (respectively $\mathcal{I}_x^>$) and viceversa.*

Proof. Suppose $\{\alpha'_k\} \prec \{\alpha_k\}$ and $f(x)$ decreasing and concave (respectively convex). Then, by the mirror property $\{\alpha_{M-k+1}\} \prec \{\alpha'_{M-k+1}\}$, and by the hypothesis of the theorem

$$\begin{aligned} \mathcal{M}(\{f(k)\}, \{\alpha_k\}) &= \mathcal{M}(\{f(M-k+1)\}, \{\alpha_{M-k+1}\}) = \mathcal{M}(\{f(m(k))\}, \{\alpha_{M-k+1}\}) \quad (18) \\ &\leq \mathcal{M}(\{f(m(k))\}, \{\alpha'_{M-k+1}\}) = \mathcal{M}(\{f(k)\}, \{\alpha'_k\}), \end{aligned}$$

where $m(x) = M - x + 1$, and because if $f(x)$ is decreasing and concave (respectively convex) then $f(m(x))$ is increasing and concave (respectively convex). $\square$

The following result is necessary to prove Lemma 2.

Theorem 5. *Given $f(x)$ and weights $\{\alpha'_k\}, \{\alpha_k\}$, if inequality $\mathcal{M}(\{f(k)\}, \{\alpha_k\}) \leq \mathcal{M}(\{f(k)\}, \{\alpha_k\})$ holds for any $g(x)$ strictly increasing and convex (respectively concave) then it holds for any $g(x)$ strictly decreasing and concave (respectively convex). If it holds for any $g(x)$ strictly decreasing and convex (respectively concave) then it holds for any $g(x)$ strictly increasing and concave (respectively convex).*

Proof. Consider the quasi-arithmetic mean with Kolmogorov function $g^\star = -g(x)$ (with inverse $g^{\star-1} = g^{-1}(-x)$). When $g(x)$ is increasing, $g^\star(x)$ is decreasing and viceversa. When $g(x)$ is convex, $g^\star(x)$ is concave and viceversa. We have

$$\begin{aligned} g^{\star-1}\Big(\sum_i \alpha'_i g^\star(f(i))\Big) &= g^{-1}\Big(-\Big(\sum_i \alpha'_i(-g(f(i)))\Big)\Big) = g^{-1}\Big(\sum_i \alpha'_i g(f(i))\Big) \quad (19) \\ &\leq g^{-1}\Big(\sum_i \alpha_i g(f(i))\Big) = g^{-1}\Big(-\Big(\sum_i \alpha_i(-g(f(i)))\Big)\Big) = g^{\star-1}\Big(\sum_i \alpha_i g^\star(f(i))\Big) \end{aligned}$$

$\square$

The following Lemma is needed to prove how the invariance properties for one mean extends to other means.

Lemma 2. *Given two distributions $\{\alpha'_k\}, \{\alpha_k\}$, and a quasi-arithmetic mean $\mathcal{M}(\{f(k)\}, \{\alpha_k\})$ with function $g(x)$. Consider the following Equations:*

$$\mathcal{M}(\{f(k)\}, \{\alpha'_k\}) \leq \mathcal{M}(\{f(k)\}, \{\alpha_k\}), \tag{20}$$

and

$$\mathcal{M}(\{f(k)\}, \{\alpha_{M-k+1}\}) \leq \mathcal{M}(\{f(k)\}, \{\alpha'_{M-k+1}\}). \tag{21}$$

Then, for each line in Table 1, and for $g(x), f(x)$ filling the conditions in first and second column in Table 1 and holding Equations (20) and (21), Equations (20) and (21) hold too for $g^\star(x), f^\star(x)$ filling the conditions in columns three and four.

Additionally, for each line in Table 1, changing $f^\star(x)$ in column four from increasing to decreasing, Equations (20) and (21) hold with the inequalities reversed.

Table 1. For each line, for $g(x), f(x)$ filling the conditions in first and second column, then Equations (20) and (21) hold for $g^\star, f^\star$ filling the conditions in columns three and four. By changing $f^\star$ from increasing to decreasing, the reverse of Equations (20) and (21) hold. ICX: convex and increasing, ICV: concave and increasing, DCX: convex and decreasing, DCV: concave and decreasing.

	g	f	$g^\star$	$f^\star$
1	ICX	ICV	ICV	ICV
1'	"	"	DCX	ICV
2	DCX	ICX	DCV	ICX
2'	"	"	ICX	ICX
3	ICV	ICX	ICX	ICX
3'	"	"	DCV	ICX
4	DCV	ICV	DCX	ICV
4'	"	"	ICV	ICV
5	ICX	ICV	DCV	ICX
5'	"	"	ICX	ICX
6	DCX	ICX	ICV	ICV
6'	"	"	DCX	ICV
7	ICV	ICX	DCX	ICV
7'	"	"	ICV	ICV
8	DCV	ICV	ICX	ICX
8'	"	"	DCV	ICX

Proof. The proof of lines 1–8 in Table 1 is in the Appendix B. Lines 1'-8' in Table 1 are a direct consequence of Theorem 5 applied to lines 1–8. □

Theorem 6. *Given a quasi-arithmetic mean $\mathcal{M}(\{f(k)\}, \{\alpha_k\})$ with function $g(x)$, we have that if an order holds the mirror property then for each line in Table 1, and for $g(x)$ filling the condition in the first column in Table 1 and preserving order for $f(x)$ in the second column in Table 1, the mean for $g^\star(x)$ filling the condition in column three in Table 1 preserves order for $f^\star(x)$ filling the condition in column four in Table 1.*

Proof. It is enough to apply Lemma 2, taking into account that if Equation (20) holds then Equation (21) holds too by the mirror property. □

Corollary 1. *Consider the weighted arithmetic mean $\mathcal{A}(\{f(k)\}, \{\alpha_k\})$. Given an order that holds the mirror property and preserves the order for mean $\mathcal{A}$, then*

(a) If order is preserved for mean $\mathcal{A}$ and for $\mathcal{I}_v^<$ then order is preserved for any mean with concave-increasing/convex-decreasing (respectively concave-decreasing/convex-increasing) function $g(x)$ and for $\mathcal{I}_v^<$ (respectively for $\mathcal{I}_x^<$).

(b) If order is preserved for mean $\mathcal{A}$ and for $\mathcal{I}_x^<$, then order is preserved for any mean with convex-increasing/concave-decreasing (respectively convex-decreasing/concave-increasing) function $g(x)$ and for $\mathcal{I}_x^<$ (respectively for $\mathcal{I}_v^<$).

Proof. Arithmetic mean is a quasi-arithmetic mean with increasing function $g(x) = x$, which is both concave-increasing and convex-increasing, and thus Table 1 collapses to Table 2.

Table 2. For each line, $g(x) = x$, $f(x)$ filling the conditions in second column and Equations (20) and (21) holding, then Equations (20) and (21) hold too for $g^\star$, $f^\star$ filling the conditions in columns three and four. ICX: convex and increasing, ICV: concave and increasing, DCX: convex and decreasing, DCV: concave and decreasing.

f	$g^\star$	$f^\star$
ICV	ICV	ICV
ICV	DCX	ICV
ICX	ICX	ICX
ICX	DCV	ICX
ICV	DCV	ICX
ICV	ICX	ICX
ICX	DCX	ICV
ICX	ICV	ICV

$\square$

Functions x^p, for $p \geq 1$ are convex-increasing and with $p \leq 0$ are convex-decreasing over $\mathcal{R}_{++}$, e^x convex-increasing over $\mathcal{R}$, while $\log x$, x^p, for $0 < p \leq 1$ are concave-increasing over $\mathcal{R}_{++}$. Affine functions are both concave and convex over $\mathcal{R}$. If $g(x)$ is convex, $-g(x)$ is concave and viceversa, and the composition of concave-increasing and concave is concave, and convex-increasing and convex is convex. We will see in the next section that preserving the order for mean $\mathcal{A}$ for $\mathcal{I}_v^<$ is defined as second-order stochastic dominance or increasing concave order, and preserving the order for mean $\mathcal{A}$ for $\mathcal{I}_x^<$ is defined as increase convex order stochastic dominance. Both orders hold the mirror property and thus Corollary 1 applies to both.

6. Application to Stochastic Orders and Cross-Entropy

6.1. First-Order Stochastic Dominance

Originating in the economics risk literature [24], first-order stochastic dominance [1,2] FSD, between two probability distributions, $\{\alpha_k'\}$, $\{\alpha_k\}$, $\{\alpha_k'\} \prec_{FSD} \{\alpha_k\}$, is defined as:

Definition 5. $\{\alpha_k'\} \prec_{FSD} \{\alpha_k\} \iff$ *for any increasing function $f(x)$,*

$$E[f(k)]_{\alpha_k'} \leq E[f(k)]_{\alpha_k}.$$

Remember that the expected value is the arithmetic weighted mean.

A necessary and sufficient condition for first-order stochastic dominance is defined by the condition $c)$ of Lemma 1, which is equivalent to condition $d)$ of Lemma 1, thus $\{\alpha_k'\} \prec_{FSD} \{\alpha_k\} \Rightarrow \{\alpha_{M-k+1}\} \prec_{FSD} \{\alpha_{M-k+1}'\}$, i.e., the mirror property holds. From the definition of FSD, and from Theorem 3, we can redefine FSD, $\{\alpha_k'\} \prec_{FSD} \{\alpha_k\}$ as there exists a mean $\mathcal{M}$ such that, for any increasing function $f(x)$, Equation (16) holds. This is, the definition of FSD is independent of the mean considered, while the original definition relies on the expected value (arithmetic mean). The mean considered can be arithmetic, harmonic, geometric or any other quasi-arithmetic mean.

Let us consider now a strictly monotonous function $g(x)$, and define a generalized cross-entropy $\mathcal{GCE}(\{p_i\},\{q_i\}) = g^{-1}(\sum_i p_i g(-\log q_i))$. Observe that it is a quasi-arithmetic mean, and for $g(x) = x$, we get the cross-entropy $CE(\{p_i\},\{q_i\}) = -\sum p_i \log q_i$. Other functions that generalize cross entropy have been defined in the context of training deep neural networks [25]. We can state the following theorem:

Theorem 7. *Given distributions $\{p'_i\}, \{p_i\}, \{q_i\}$, increasing, and $\{p'_k\} \prec_{FSD} \{p_k\}$, then for any mean $\mathcal{GCE}$ with function $g(x)$, $\mathcal{GCE}(\{p_i\},\{q_i\}) \leq \mathcal{GCE}(\{p'_i\},\{q_i\})$.*

Proof. Observe first that $f(k) = -\log q_k$ is a decreasing sequence. From the hypothesis, condition $c)$ of Lemma 1 holds, thus we can then apply Theorem 1 to the mean with function $g(x)$ and for $\{f(k)\}$ decreasing. $\square$

The following result relates the entropies of two distributions with the first stochastic order.

Theorem 8. *Given $\{p_k\}, \{q_k\}$ increasing, and $\{q_k\} \prec_{FSD} \{p_k\}$, then $\mathcal{H}(\{p_k\}) \leq \mathcal{H}(\{q_k\})$, where $\mathcal{H}$ stands for Shannon entropy.*

Proof. The Kullback–Leibler distance is always positive [26], $\sum_i p_i \log \frac{p_i}{q_i} \geq 0$, and thus we have that $\mathcal{H}(\{p_k\}) = -\sum_i p_i \log p_i \leq -\sum p_i \log q_i$. Applying Theorem 7 for $g(x) = x$, $-\sum p_i \log q_i \leq -\sum q_i \log q_i = \mathcal{H}(\{q_k\})$. $\square$

6.2. Second-Order Stochastic Dominance and Increase Convex Ordering

Definition 6. *Second-order stochastic dominance between two probability distributions, $\{\alpha_k\}$, $\{\alpha'_k\}$, $\{\alpha'_k\} \prec_{SSD} \{\alpha_k\}$, occurs when for any increasing concave function $f(x)$,*

$$E[f(k)]_{\alpha'_k} \leq E[f(k)]_{\alpha_k}.$$

Trivially

$$\{\alpha'_k\} \prec_{FSD} \{\alpha_k\} \Rightarrow \{\alpha'_k\} \prec_{SSD} \{\alpha_k\}.$$

Let us consider the cumulative distribution function, $F_i = \alpha_1 + \alpha_2 + \ldots + \alpha_i$, and the survival function $\bar{F}_i = 1 - F_i = \alpha_M + \alpha_{M-1} + \ldots + \alpha_{i+1}$. A necessary and sufficient condition for second-order stochastic dominance [1,2] is the following:

$$
\begin{aligned}
F'_1 &\geq F_1 \\
F'_1 + F'_2 &\geq F_1 + F_2 \\
&\vdots \\
F'_1 + \ldots + F'_{M-1} &\geq F_1 + \ldots + F_{M-1} \\
F'_1 + \ldots + F'_{M-1} + F'_M &\geq F_1 + \ldots + F_{M-1} + F_M,
\end{aligned}
\tag{22}
$$

or equivalently,

$$
\begin{aligned}
\bar{F}'_1 &\leq \bar{F}_1 \\
\bar{F}'_1 + \bar{F}'_2 &\leq \bar{F}_1 + \bar{F}_2 \\
&\vdots \\
\bar{F}'_1 + \ldots + \bar{F}'_{M-1} &\leq \bar{F}_1 + \ldots + \bar{F}_{M-1} \\
\bar{F}'_1 + \ldots + \bar{F}'_{M-1} + \bar{F}'_M &\leq \bar{F}_1 + \ldots + \bar{F}_{M-1} + \bar{F}_M.
\end{aligned}
\tag{23}
$$

Let us see first that $\{\alpha'_k\} \prec_{SSD} \{\alpha_k\} \Rightarrow \{\alpha_{M-k+1}\} \prec_{SSD} \{\alpha'_{M-k+1}\}$, i.e., SSD holds the mirror property. Effectively, define $G_i = \alpha_M + \alpha_{M-1} + \ldots + \alpha_{i+1}$, and the survival

function $\bar{G}_i = 1 - G_i = \alpha_1 + \alpha_2 + \ldots + \alpha_i$. However, $G_i = \bar{F}_i$, $\bar{G}_i = F_i$, $G'_i = \bar{F}'_i$, $\bar{G}'_i = F'_i$, and substituting in Equation (22) or Equation (23), we obtain the desired result.

From Corollary 1, second-order stochastic dominance preserves mean order for all means $\mathcal{M}$ defined by a concave-increasing/convex-decreasing (respectively concave-decreasing/convex-increasing) function $g(x)$, and for the set of all concave-increasing functions $\mathcal{I}_v^<$ (respectively convex-increasing functions $\mathcal{I}_x^<$). For instance, order is preserved for the geometric mean, or any mean with $g(x) = x^p$, $0 < p \leq 1$ and for $\mathcal{I}_v^<$, or for or any mean with $g(x) = x^p$ with $p < 0$, as the harmonic mean, for $\mathcal{I}_x^<$. In particular, we can state the following theorem about cross-entropy of two distributions,

Theorem 9. *Given distributions $\{p'_i\}, \{p_i\}, \{q_i = f(i)\}, f(x)$ concave-increasing, and $\{p'_i\} \prec_{SSD} \{p_i\}$, then $CE(\{p_i\}, \{q_i\}) \leq CE(\{p'_i\}, \{q_i\})$.*

Proof. We have that $CE(\{p_i\}, \{q_i\}) = -\log \mathcal{G}(\{q_i\}, \{p_i\}) = -\log(\Pi_i q_i^{p_i}) = -\log(\exp(\sum_i p_i \log q_i))$, where $\mathcal{G}$ stands for geometric mean. The geometric mean is a quasi-arithmetic mean with function $g(x) = \log x$, concave-increasing. Using Definition 6 and applying Corollary 1(a), we obtain the inequality $\mathcal{G}(\{q_i\}, \{p'_i\}) \leq \mathcal{G}(\{q_i\}, \{p_i\})$ and we apply the function $-\log x$ to both members of this inequality. $\square$

When we consider $f(x)$, a convex instead of a concave function, we talk of increasing convex order, ICX. This is,

Definition 7. *Second-order stochastic dominance between two probability distributions, $\{\alpha_k\}$, $\{\alpha'_k\}$, $\{\alpha'_k\} \prec_{ICX} \{\alpha_k\}$, occurs when for any increasing convex function $f(x)$,*

$$E[f(k)]_{\alpha'_k} \leq E[f(k)]_{\alpha_k}.$$

$\{\alpha_k\}$ is greater in increasing convex order than $\{\alpha'_k\}$, $\{\alpha'_k\} \prec_{ICX} \{\alpha_k\}$, if and only if the following inequalities hold:

$$
\begin{aligned}
F'_M &\geq F_M \\
F'_M + F'_{M-1} &\geq F_M + F_{M-1} \\
&\vdots \\
F'_M + \ldots + F'_2 &\geq F_M + \ldots + F_2 \\
F'_M + \ldots + F'_2 + F'_1 &\geq F_M + \ldots + F_2 + F_1
\end{aligned}
\tag{24}
$$

or equivalently,

$$
\begin{aligned}
\bar{F}'_M &\leq \bar{F}_M \\
\bar{F}'_M + \bar{F}'_{M-1} &\leq \bar{F}_M + \bar{F}_{M-1} \\
&\vdots \\
\bar{F}'_M + \ldots + \bar{F}'_2 &\leq \bar{F}_M + \ldots + \bar{F}_2 \\
\bar{F}'_M + \ldots + \bar{F}'_2 + \bar{F}'_1 &\leq \bar{F}_M + \ldots + \bar{F}_2 + \bar{F}_1
\end{aligned}
\tag{25}
$$

Trivially

$$\{\alpha'_k\} \prec_{FSD} \{\alpha_k\} \Rightarrow \{\alpha'_k\} \prec_{ICX} \{\alpha_k\}.$$

Mirror property is also immediate. From Corollary 1, second-order stochastic dominance preserves mean order for all means $\mathcal{M}$ defined by a convex-increasing/concave-decreasing (respectively convex-decreasing/concave-increasing) function $g(x)$ and for the set of all convex-increasing functions $\mathcal{I}_x^<$ (respectively concave-increasing functions $\mathcal{I}_v^<$). For instance, any mean with $g(x) = x^p$, $p \geq 1$ as the weighted quadratic mean with $\mathcal{I}_x^<$, or

$g(x) = x^p$, $p < 0$ as the harmonic mean with $\mathcal{I}_v^<$, or the mean with $g(x) = -\log x$ with $\mathcal{I}_v^<$. In particular, we can state the following result,

Theorem 10. *Given distributions $\{p_i'\}, \{p_i\}, \{q_i = f(i)\}, f(x)$ concave-increasing, and $\{p_i'\} \prec_{ICX} \{p_i\}$, then $CE(\{p_i\}, \{q_i\}) \leq CE(\{p_i'\}, \{q_i\})$.*

Proof. Consider the quasi-arithmetic mean $\mathcal{G}'(\{q_i\}, \{p_i\})$ with $g(x) = -\log x$, convex-decreasing. We have that $CE(\{p_i\}, \{q_i\}) = -\log \mathcal{G}'(\{q_i\}, \{p_i\})$. Using Definition 7 and applying Corollary 1(b), we have that $\mathcal{G}'(\{q_i\}, \{p_i'\}) \leq \mathcal{G}'(\{q_i\}, \{p_i\})$, and we apply to each member of the inequality the function $-\log x$. $\square$

6.3. Likelihood Ratio Dominance

Definition 8. *Likelihood ratio dominance, LR, $\{\alpha_k'\} \prec_{LR} \{\alpha_k\}$, is defined as*

$$\{\alpha_k'\} \prec_{LR} \{\alpha_k\} \iff \forall(i,j), 1 \leq i, j \leq M, i \leq j \Rightarrow \alpha_i'/\alpha_j' \geq \alpha_i/\alpha_j. \tag{26}$$

As $\alpha_i'/\alpha_j' \geq \alpha_i/\alpha_j \Rightarrow \alpha_j/\alpha_i \geq \alpha_j'/\alpha_i'$, we have that LR holds the mirror property.
It can be shown that LR order implies FSD order,
and then LR order holds Theorems 3 and 7 and Equation (16) holds for any mean $\mathcal{M}$,

Theorem 11. *Likelihood-ratio order implies first stochastic order, i.e.,*

$$\{\alpha_k'\} \prec_{LR} \{\alpha_k\} \Rightarrow \{\alpha_k'\} \prec_{FSD} \{\alpha_k\}. \tag{27}$$

Proof. There are indirect proofs in the literature [1,17]. A direct proof is in the Appendix C. $\square$

As the condition for LR order, Equation (26), is easy to check, this order comes very handy in proving sufficient condition for FSD order. Additionally, for the uniform distribution $\{1/N\}$ and any increasing distribution $\{\alpha_k\}$, we have that $\{1/M\} \prec_{LR} \{\alpha_k\}$, while for any decreasing distribution $\{\alpha_k\}$ we have that $\{\alpha_k\} \prec_{LR} \{1/M\}$.

Consider Shannon entropy, $\mathcal{H}(\{p_k\})$, and Rényi entropy $\mathcal{R}_\beta(\{p_k\})$ of a distribution $\{p_k\}$, that as seen in Section 3 are quasi-arithmetic means of the sequence $\{1/\log p_k\}$ with weights $\{p_k\}$, and with $g(x) = x$ and $g(x) = 2^{(\beta-1)x}$, respectively. Without loss of generality, we can consider $\{p_k\}$ in increasing order, then the sequence $\{1/\log p_k\}$ will be decreasing. We have that $\{1/M\} \prec \{p_k\}$, and then we can apply Theorem 3, and obtain for Shannon entropy

$$\mathcal{H}(\{p_k\}) \leq \mathcal{A}(\{1/\log p_k\}),$$

where $\mathcal{A}(\{1/\log p_k\})$ is the arithmetic mean of $\{1/\log p_k\}$, and for Rényi entropy, observing that $g^{-1}(x) = \frac{1}{1-\beta}\log x$,

$$\mathcal{R}_\beta(\{p_k\}) \leq \frac{1}{1-\beta}\log\sum_k \frac{1}{M}p_k^{(\beta-1)} = \frac{1}{1-\beta}\log\mathcal{A}(\{p_k^{(\beta-1)}\}).$$

For Tsallis entropy, which can be considered the weighted arithmetic mean of the sequence $\{1/\ln_q p_k\}$ with weights $\{p_k\}$, as for $\{p_k\}$ in increasing order the sequence $\{1/\ln_q p_k\}$ will be decreasing (the derivative of $\ln_q x$ is positive), we have similarly to Shannon entropy,

$$\mathcal{S}_q(\{p_k\}) \leq \mathcal{A}(\{1/\ln_q p_k\}).$$

We can also state the following result for the generalized cross entropy $\mathcal{GCE}$. We say that $\{p_i\}$ is comonotonic [27,28] with $\{q_i\}$ when, for all $1 \leq i, j \leq M$, $q_i < q_j \implies p_i \leq p_j$.

Theorem 12. *Given distributions $\{p_i\}, \{q_i\}$ comonotonic, then $\mathcal{GCE}(\{p_i\}, \{q_i\}) \leq \mathcal{GCE}(\{1/M\}, \{q_i\})$.*

Proof. Without loss of generality, we can reorder $\{p_i\}, \{q_i\}$ so that both are increasing. We know that $\{1/M\} \prec_{LR} \{p_k\}$ and thus $\{1/M\} \prec_{FSD} \{p_k\}$. From the definition of $\prec_{FSD}$, preserve mean order holds for increasing functions and for arithmetic mean, thus we can then apply Theorem 3 for $\{-\log q_i\}$ decreasing. □

From Theorem 12, when $\{p_i\}, \{q_i\}$ are comonotonic, then

$$CE(\{p_i\}, \{q_i\}) \leq \frac{1}{M} \sum_{k=1}^{M} -\log q_k. \tag{28}$$

6.4. Hazard Rate

Definition 9. *A probability distribution $\{\alpha_k\}$ is greater in hazard rate, HR, to $\{\alpha_k'\}$, $\{\alpha_k'\} \prec_{HR} \{\alpha_k\}$, if and only if for all $j \geq i$, the following condition is filled*

$$\frac{\bar{F}'_i}{\bar{F}_i} \geq \frac{\bar{F}'_j}{\bar{F}_j}.$$

Observe that this can be written as

$$\frac{\bar{F}_j}{\bar{F}'_j} \geq \frac{\bar{F}_i}{\bar{F}'_i},$$

or

$$\frac{1 - F_j}{1 - F'_j} \geq \frac{1 - F_i}{1 - F'_i},$$

If we consider now the sequences written in inverse order, i.e., $\alpha_{M-i+1}, \alpha'_{M-i+1}$, it is clear that if $i \leq j$ then $M - j + 1 \leq M - i + 1$, and from $1 - \bar{F}_j = \bar{F}_{M-j+1}$, we get

$$\frac{\bar{F}_{M-j+1}}{\bar{F}'_{M-j+1}} \geq \frac{\bar{F}_{M-i+1}}{\bar{F}'_{M-i+1}},$$

which means $\{\alpha_{M-k+1}\} \prec_{HR} \{\alpha'_{M-k+1}\}$, and HR order holds the mirror property. It can be shown that $\{\alpha_k'\} \prec_{HR} \{\alpha_k\} \Rightarrow \{\alpha_k'\} \prec_{FSD} \{\alpha_k\}$ and then HR order holds Theorems 3 and 7, and Equation (16) holds for any mean $\mathcal{M}$.

7. Example: Linear Combination of Monte Carlo Techniques

When we want to estimate an integral $\mu = \int h(x)dx$ using MIS (multiple importance sampling) Monte Carlo methods, we have several choices of techniques, each of them with a given pdf $\{p_k(x)\}$, $1 \leq k \leq n$, which provide the primary estimators $\{I_{k,1} = \frac{h(x)}{p_k(x)}\}$. If for $h(x) > 0$, we have that $p_k(x) > 0$, then the technique is unbiased. We are interested in optimal ways of combining the techniques. One option is linearly combining the different estimators $\{I_{i,n_i}\}$, $\sum_{i=1}^{M} w_i I_{i,n_i}$, with weights $\{w_k\}$ and sampling proportions $\{n_i = \alpha_i N\}$, with $N = \sum_{i=1}^{M} n_i$. If all techniques are unbiased, the resulting combination is also unbiased. The variance is given by

$$V_N = \sum_{i=1}^{M} w_i^2 \frac{v_i}{n_i} = \frac{1}{N} \sum_{i=1}^{M} w_i^2 \frac{v_i}{\alpha_i} = \frac{1}{N} V, \tag{29}$$

where v_i are the variances for the primary estimators of each technique and V is the variance for the primary MIS estimator.

The optimal combination of weights, i.e. the one that leads to minimum variance, has been studied in [18,29,30].

The variance value will depend on two sets of weights, $\{w_i\}$ and $\{\alpha_i\}$, but there are cases where we can reduce it to a single set of weights. We present the following examples, where variances are taken for $N = 1$:

- when $\alpha_i = w_i$, then $V = \sum_{i=1}^{M} \alpha_i v_i$, the weighted arithmetic mean ($g(x) = x$) of $\{v_i\}$.
- when the sampling proportions are fixed, the optimal variance is given by $\mathcal{H}(\{\frac{v_k}{\alpha_k}\})$, the weighted harmonic mean ($g(x) = 1/x$) of $\{v_i\}$.
- when weights $\{w_i\}$ are fixed, the optimal variance is given by $(\sum_{k=1}^{M} w_k \sqrt{v_k})^2$, which is the weighted power mean of $\{v_i\}$ with exponent $r = 1/2$ ($g(x) = x^{1/2}$).

Observe that the variance in these three cases is a quasi-arithmetic mean.

Let us now order $\{v_k\}$ in increasing order, and let us take $\alpha'_k = h(v_k)$, $h(x)$ decreasing, and $\alpha_k = 1/M$. We have that $\alpha'_k = h(v_k) \prec_{LR} \{\alpha_k = 1/M\}$ and thus, by Theorem 11, $\alpha'_k = h(v_k) \prec_{FSD} \{\alpha_k = 1/M\}$, and we can apply Theorem 2 with $f(k) = v_k$ to the quasi-arithmetic means above. Thus, the variance is less when taking sampling proportions or coefficients decreasing in v_k than for equal sampling or equal weighting.

The same would be the case when considering a different measure of error, whenever the error of the combined techniques can also be expressed as a weighted quasi-arithmetic mean of the values of this measure for all techniques. For instance, suppose we use as measure of error the standard deviation, $\sigma_i^2 = v_i$, $\sigma^2 = V$, the above cases become

- when $\alpha_i = w_i$, then $\sigma = \sqrt{V} = \sqrt{\sum_{i=1}^{M} \alpha_i \sigma_i^2}$, the weighted root mean square or quadratic mean of $\{\sigma_i\}$, or weighted power mean with $r = 2$ ($g(x) = x^2$).
- when the weights $\{w_i\}$ are fixed, the optimal standard deviation is given by $\sigma = \sqrt{V} = \sqrt{H(\{\frac{\sigma_k^2}{\alpha_k}\})}$, the power mean of $\{\sigma_i\}$ with $r = -2$ ($g(x) = x^{-2}$).
- when the sampling proportions are fixed, the optimal standard deviation is given by $\sigma = \sqrt{V} = \sum_{k=1}^{M} w_k \sigma_k$, the arithmetic mean of $\{\sigma_i\}$ ($g(x) = x$).

For more details see [18].

7.1. Liabilities vs. Utilities

From the above example, we can study the relationship between utilities, which we try to maximize, and liabilities, that we try to minimize. As a liability can always be considered as the inverse of a utility, we can see from Theorem 1 that establishing invariance properties on the order between means of utilities is equivalent to doing this with liabilities, except that the order is inverted. We can define for instance the efficiency $E = 1/V$ as the inverse of the variance, and obtain it as a weighted mean of individual technique's efficiencies, $e_i = 1/v_i$. Consider, for instance, the first case in the example above, where $V = \sum_{i=1}^{M} \alpha_i v_i$. We have

$$E = \frac{1}{V} = \frac{1}{\sum_{i=1}^{M} \alpha_i v_i} = \frac{1}{\sum_{i=1}^{M} \alpha_i e_i^{-1}} = \mathcal{H}(\{\frac{e_i}{\alpha_i}\}),$$

where $\mathcal{H}$ denotes the weighted harmonic mean.
Considering now the second case

$$E = \frac{1}{V} = \frac{1}{\mathcal{H}(\{\frac{v_k}{\alpha_k}\})} = \sum_{i=1}^{M} \alpha_i v_i^{-1} = \sum_{i=1}^{M} \alpha_i e_i,$$

or weighted arithmetic mean. For the third case,

$$E = \frac{1}{V} = \frac{1}{(\sum_{k=1}^{M} w_k \sqrt{v_k})^2} = (\sum_{k=1}^{M} w_k e_k^{-1/2})^{-2},$$

which is the weighted power mean with $r = -1/2$.

8. Conclusions and Future Work

We have presented in this paper the relationship between stochastic orders and quasi-arithmetic means. We have proved several ordering invariance theorems, that show that given two distributions under a certain stochastic order, the ordering of the means is preserved for any quasi-arithmetic mean we might consider, this is, not only for the arithmetic mean (or expected value). We have shown how the results apply to first order, second order, likelihood ratio, hazard-rate, and increasing convex stochastic orders, and its application to cross-entropy. We have also presented an application example based on the linear combination of Monte Carlo estimators, and shown that the invariance allows costs or liabilities to be considered as the symmetric case of utilities.

In the future, we want to generalize our results to spatial weight matrices [31]. The rows in a spatial weight matrix are weights that give the influence of n entities over each other. Different weighted means such as arithmetic, harmonic, or geometric [32] can be used to compute this influence. We can thus apply our invariance results to each row. We will also investigate which of our results for Shannon entropy extend to Tsallis entropy too. Both Shannon and Tsallis entropy are weighted arithmetic means [23], and given $\{p_k\}$ monotonic, both $\{1/\log p_k\}$ and $\{1/\ln_q p_k\}$ are monotonic too. Finally, we will investigate the invariance of the different stochastic orders under the operations of compositional data [3].

Author Contributions: Conceptualization, M.S. and S.C.; methodology, M.S. and J.P.; validation, J.P. and V.E.; writing—original draft preparation, M.S.; writing—review and editing, all authors. All authors have read and agreed to the published version of the manuscript.

Appendix A. Proof of Lemma 1

Proof. Subtracting 1 from both sides of each inequality proves $(c) \Leftrightarrow (d)$. To prove $(a) \Rightarrow (c)$, we proceed in the following way. Consider the increasing sequence $\{g(f(1)), \ldots, g(f(1)), g(f(M)), \ldots, g(f(M))\}$, $f(1) < f(M)$ (and thus $g(f(1)) < g(f(M))$ by the strict monotonicity of $g(x)$), and where $g(f(1))$ is written l times, denote $\mathbf{L} = a_1 + \ldots + a_l$, $\mathbf{L}' = a_1' + \ldots + a_l'$. Since $a_{l+1} + \ldots + a_M = 1 - \mathbf{L}$, $a_{l+1}' + \ldots + a_M' = 1 - \mathbf{L}'$, then (a) gives

$$\mathbf{L}'g(f(1)) + (1 - \mathbf{L}')g(f(M)) - \mathbf{L}g(f(1)) - (1 - \mathbf{L})g(f(M)) \leq 0,$$

i.e.,

$$(\mathbf{L}' - \mathbf{L})(g(f(1)) - g(f(M))) \leq 0 \Rightarrow \mathbf{L}' \geq \mathbf{L}.$$

This proves the first $M - 1$ inequalities in c, thus $(a) \Rightarrow (c)$. Observe now that

$$\mathbf{L}' \geq \mathbf{L} \Rightarrow \mathbf{1}\text{-}\mathbf{L} \geq \mathbf{1}\text{-}\mathbf{L}',$$

and thus $(a) \Rightarrow (d)$ (although this can be seen directly from $(a) \Rightarrow (c)$ and $(c) \Leftrightarrow (d)$. To prove that (b) implies (c) and (d), consider the sequence,

$$\{g(f(1)), \ldots, g(f(1)), g(f(M)), \ldots, g(f(M))\},$$

$f(1) < f(M)$, $g(f(M))$ is written l times, and the same definitions as before for $\mathbf{L}, \mathbf{L}'$, then (b) gives

$$(1 - \mathbf{L})g(f(1)) + \mathbf{L}g(f(M)) - (1 - \mathbf{L}')g(f(1)) - \mathbf{L}'g(f(M)) \leq 0,$$

and we proceed as above. Thus, $(a) \Rightarrow (c)$.

Let us see now that (c) implies (a).

Define for $1 \leq k \leq M$, $A_k = \sum_{j=1}^{k} \alpha_k$, $A'_k = \sum_{j=1}^{k} \alpha'_k$, and $A_0 = A'_0 = 0$. Then,

$$\sum_{k=1}^{M} \alpha_k g(f(k)) - \sum_{k=1}^{M} \alpha'_k g(f(k)) = \sum_{k=1}^{M} (\alpha_k - \alpha'_k) g(f(k)) \tag{A1}$$

$$= \sum_{k=1}^{M} (A_k - A_{k-1} - A'_k + A'_{k-1}) g(f(k))$$

$$= \sum_{k=1}^{M} (A_k - A'_k) g(f(k)) - \sum_{k=1}^{M} (A_{k-1} - A'_{k-1}) g(f(k))$$

$$= \sum_{k=1}^{M-1} (A_k - A'_k) g(f(k)) - \sum_{k=0}^{M-1} (A_k - A'_k) g(f(k+1))$$

$$= \sum_{k=1}^{M-1} (A_k - A'_k) g(f(k)) - \sum_{k=1}^{M-1} (A_k - A'_k) g(f(k+1))$$

$$= \sum_{k=1}^{M-1} (A_k - A'_k)(g(f(k)) - g(f(k+1))) \geq 0,$$

as (c) implies that for all k, $A'_k - A_k \geq 0$, and $\{g(f(k))\}$ is an increasing sequence. Thus, $(c) \Rightarrow (a)$.

Repeating the proof for the sequences $\{\alpha_{M-k+1}\}$ and $\{\alpha'_{M-k+1}\}$, we obtain that $(d) \Rightarrow (b)$.

Let $g(x)$ be decreasing. Observe that

$$\sum_{k=1}^{M} \alpha_{M-k+1} g(f(k)) = \sum_{k=1}^{M} \alpha_{M-k+1} g(f(m(M-k+1))) = \sum_{k=1}^{M} \alpha_k g(f(m(k))),$$

where $m(k) = M - k + 1$, and thus $(a) \Rightarrow (b')$ because $\{g(f(m(k)))\}$ is an increasing sequence. Now suppose $g(x)$ is increasing, from

$$\sum_{k=1}^{M} \alpha_k g(f(k)) = \sum_{k=1}^{M} \alpha_{M-k+1} g(f(M-k+1)) = \sum_{k=1}^{M} \alpha_{M-k+1} g(f(m(k))),$$

we have that $(b') \Rightarrow (a)$ because $\{g(f(m(k)))\}$ is a decreasing sequence. We can show similarly that $(b) \Leftrightarrow (a')$.

Consider now $f(x)$ decreasing. Reversed (a) is

$$\sum_{k=1}^{M} \alpha_k g(f(k)) \leq \sum_{k=1}^{M} \alpha'_k g(f(k)), \tag{A2}$$

but it can be written as

$$\sum_{k=1}^{M} \alpha_{M-k+1} g(f(M-k+1)) \leq \sum_{k=1}^{M} \alpha'_{M-k+1} g(f(M-k+1)) \tag{A3}$$

$$= \sum_{k=1}^{M} \alpha_{M-k+1} g(f(m(k))) \leq \sum_{k=1}^{M} \alpha'_{M-k+1} g(f(m(k))),$$

where $\{g(f(m(k)))\}$ is an increasing sequence, and thus $(c) \Rightarrow (a)$ when $f(x)$ is decreasing and for order of inequality in (a) reversed. The other cases for $f(x)$ decreasing can be analogously proven. $\square$

Appendix B. Proof of Lines 1–8 from Table 1 in Lemma 2

Proof. Let us prove first Lines 1–4 from Table 1.

Let $g(x)$ be the strictly monotonic function associated with mean $\mathcal{M}$. Consider the following inequalities. For $g(x)$ increasing, $g^\star$ increasing:

$$
\begin{aligned}
\sum_{k=1}^{M} \alpha'_k g^\star(f^\star(k)) &= \sum_{k=1}^{M} \alpha'_k g(g^{-1}(g^\star(f^\star(k)))) \\
&\leq \sum_{k=1}^{M} \alpha_k g(g^{-1}(g^\star(f^\star(k)))) = \sum_{k=1}^{M} \alpha_k g^\star(f^\star(k)).
\end{aligned}
\tag{A4}
$$

For $g(x)$ decreasing, $g^\star$ decreasing:

$$
\begin{aligned}
\sum_{k=1}^{M} \alpha_k g^\star(f^\star(k)) &= \sum_{k=1}^{M} \alpha_k g(g^{-1}(g^\star(f^\star(k)))) \\
&\leq \sum_{k=1}^{M} \alpha'_k g(g^{-1}(g^\star(f^\star(k)))) \leq \sum_{k=1}^{M} \alpha'_k g^\star(f^\star(k)).
\end{aligned}
\tag{A5}
$$

Observe that the central inequalities in Equations (A4) and (A5) are true whenever Equation (20) holds and viceversa. We are interested in considering all possibilities for it to hold, when $f^\star$ and $f(x) = g^{-1}(h(x))$ are increasing, with $h(x) = g^\star(f^\star(x))$. We need first to study the convexity/concavity of the function $g^{-1}(h(x))$. Let us suppose all functions are in $\mathcal{C}^2$, then $h'(x) = g^{\star\prime}(f^\star(x))f^{\star\prime}(x)$, and $h''(x) = g^{\star\prime\prime}(f^\star(x))(f^{\star\prime}(x))^2 + g^{\star\prime}(f^\star(x))f^{\star\prime\prime}(x)$. Dropping the arguments for convenience, we have $h'' = g^{\star\prime\prime}(f^{\star\prime})^2 + g^{\star\prime}f^{\star\prime\prime}$. For h'' to be positive (convex) $g^{\star\prime\prime}$ has to be positive (convex) and $g^{\star\prime}, f^{\star\prime\prime}$ have to be both positive or negative (thus $g^\star$ convex-increasing, f convex, or $g^\star$ convex-decreasing, f concave). For h'' to be negative, $g^{\star\prime\prime}$ has to be negative (concave) and $g^{\star\prime}, f^{\star\prime\prime}$ have to have a different sign (thus $g^\star$ concave-increasing, f concave, or $g^\star$ concave-decreasing, f convex). For $g^{-1}(h(x))$, we can apply the same rule. Observe also that the composition of two increasing or two decreasing functions is increasing, and increasing and decreasing or viceversa is decreasing. For the inverse function, as $(g^{-1}(g(x)) = x$ we have that $g^{-1\prime}(g(x))g'(x) = 1$, $g^{-1\prime\prime}(g(x))(g'(x))^2 + g^{-1\prime}(g(x))g''(x) = 0$, which isolating $g^{-1\prime\prime}(g(x))$ and dropping arguments becomes $g^{-1\prime\prime} = -g^{-1\prime}g''$, and thus for g^{-1} convex-increasing (concave-increasing), the inverse is concave-increasing (convex-increasing), while for g^{-1} convex-decreasing (concave-decreasing), the inverse is convex-decreasing (concave-decreasing). In Table A1, we show the different possibilities for $f(x)$ to be increasing when $f^\star$ is increasing. □

Table A1. Different possible combinations where the concavity/convexity of $g^{-1}(h(x))$ can be predicted for $g(x)$ and $f^\star, f(x)$ increasing. ICX: convex and increasing, ICV: concave and increasing, DCX: convex and decreasing, DCV: concave and decreasing.

	g	g^{-1}	$g^\star$	$f^\star$	$h = g^\star \circ f^\star$	$f = g^{-1} \circ h$
1	ICX	ICV	ICV	ICV	ICV	ICV
2	DCX	DCX	DCV	ICX	DCV	ICX
3	ICV	ICX	ICX	ICX	ICX	ICX
4	DCV	DCV	DCX	ICV	DCX	ICV

Let us prove now Lines 5–8 from Table 1.

Consider the following inequalities, where $m(x) = M - x + 1$, for $g(x)$ increasing, $g^\star$ decreasing,

$$
\begin{aligned}
\sum_{k=1}^{M} \alpha_k g^\star(f^\star(k)) &= \sum_{k=1}^{M} \alpha_{M-k+1} g^\star(f^\star(M-k+1)) \\
&= \sum_{k=1}^{M} \alpha_{M-k+1} g(g^{-1}(g^\star(f^\star(m(k))))) \le \sum_{k=1}^{M} \alpha'_{M-k+1} g(g^{-1}(g^\star(f^\star(m(k))))) \\
&= \sum_{k=1}^{M} \alpha'_{M-k+1} g^\star(f^\star(M-k+1)) = \sum_{k=1}^{M} \alpha'_k g^\star(f^\star(k)),
\end{aligned}
\tag{A6}
$$

and for $g(x)$ decreasing, $g^\star$ increasing

$$
\begin{aligned}
\sum_{k=1}^{M} \alpha'_k g^\star(f^\star(k)) &= \sum_{k=1}^{M} \alpha'_{M-k+1} g^\star(f^\star(M-k+1)) \\
&= \sum_{k=1}^{M} \alpha'_{M-k+1} g(g^{-1}(g^\star(f^\star(m(k))))) \le \sum_{k=1}^{M} \alpha_{M-k+1} g(g^{-1}(g^\star(f^\star(m(k))))) \\
&= \sum_{k=1}^{M} \alpha_{M-k+1} g^\star(f^\star(M-k+1)) = \sum_{k=1}^{M} \alpha_k g^\star(f^\star(k)).
\end{aligned}
\tag{A7}
$$

Observe that the central inequalities in Equations (A6) and (A7) are true whenever Equation (21) holds and viceversa. We are interested in considering all possibilities for it to hold, when $f^\star$ and $f(x) = g^{-1}(h(x))$ are increasing, with $h(x) = g^\star(f^\star(m(x)))$. As before, we need first to study the convexity/concavity of the function $g^{-1}(h(x))$.

Consider the function $h(x) = g^\star(f^\star(m(x)))$. The second derivative is, as $m'(x) = -1$ and $m''(x) = 0$, $h''(x) = g^{\star\prime\prime}(f^\star(m(x))f^{\star\prime 2}(m(x)) + g^{\star\prime}(f^\star(m(x))f^{\star\prime\prime}(m(x))$. Let us drop the arguments for convenience, then $h'' = g^{\star\prime\prime}f^{\star\prime 2} + g^{\star\prime}f^{\star\prime\prime}$. Observe that it is the same result obtained above. In Table A2, we see the possible combinations for $f(x)$ to be increasing when $f^\star(x)$ is increasing. The difference with Table A1 is that we have to consider here $m(x)$.

Table A2. Different possible combinations where the concavity/convexity of $g^{-1}(h(x))$ can be predicted and $f(x)$ is increasing when $g(x)$ is increasing and $g^\star(x)$ decreasing (or viceversa) when $f^\star$ is increasing. ICX: convex and increasing, ICV: concave and increasing, DCX: convex and decreasing, DCV: concave and decreasing.

	g	g^{-1}	$g^\star$	$f^\star$	$f^\star(m)$	$h = g^\star \circ f^\star \circ m$	$f = g^{-1} \circ h$
5	ICX	ICV	DCV	ICX	DCX	ICV	ICV
6	DCX	DCX	ICV	ICV	DCV	DCV	ICX
7	ICV	ICX	DCX	ICV	DCV	ICX	ICX
8	DCV	DCV	ICX	ICX	DCX	DCX	ICV

Then, we can summarize the results from Tables A1 and A2 in Table 1. This will do to prove Equation (20). The results for Equation (21) are obtained as above by substituting $\{\alpha_k\}, \{\alpha'_k\}$ by $\{\alpha'_{M-k+1}\}, \{\alpha_{M-k+1}\}$.

Suppose now $f^\star(x)$ is decreasing and that Equations (20) and Equation (21) hold for $g^\star(x)$ and for $f^\star(x)$ increasing. Then, the following inequalities, for $g^\star(x)$ increasing,

$$
\begin{aligned}
\sum_k \alpha_k g^\star(f^\star(k)) &= \sum_k \alpha_{M-k+1} g^\star(f^\star(m(k))) \\
&\le \sum_k \alpha'_{M-k+1} g^\star(f^\star(m(k))) = \sum_k \alpha'_k g^\star(f^\star(k)),
\end{aligned}
\tag{A8}
$$

$$\sum_k \alpha'_{M-k+1} g^\star(f^\star(k)) \quad = \quad \sum_k \alpha'_k g^\star(f^\star(m(k))) \tag{A9}$$
$$\leq \quad \sum_k \alpha_k g^\star(f^\star(m(k))) = \sum_k \alpha_{M-k+1} g^\star(f^\star(k)),$$

and for $g^\star(x)$ decreasing,

$$\sum_k \alpha'_k g^\star(f^\star(k)) \quad = \quad \sum_k \alpha'_{M-k+1} g^\star(f^\star(m(k))) \tag{A10}$$
$$\leq \quad \sum_k \alpha_{M-k+1} g^\star(f^\star(m(k))) = \sum_k \alpha_k g^\star(f^\star(k)),$$

$$\sum_k \alpha_{M-k+1} g^\star(f^\star(k)) \quad = \quad \sum_k \alpha_k g^\star(f^\star(m(k))) \tag{A11}$$
$$\leq \quad \sum_k \alpha'_k g^\star(f^\star(m(k))) = \sum_k \alpha'_{M-k+1} g^\star(f^\star(k)),$$

hold because $f^\star(m(x))$ is an increasing function. Thus, Equations (20) and (21) hold with the direction of inequality reversed. Observe also that $f^\star(m(x))$ will have the same character of convexity/concavity than $f^\star(x)$. Applying Equations (A8)–(A11), according to whether $g^\star$ is increasing or decreasing, to each line of Table 1, we obtain the desired result.

Appendix C. Proof of Theorem 11

Proof. First, note that if $\frac{\alpha'_i}{\alpha'_j} \geq \frac{\alpha_i}{\alpha_j}$ for $i \leq j$, then $\alpha'_1 \geq \alpha_1$ and $\alpha'_M \leq \alpha_M$, see [17]. Let us proceed now by induction. For $M = 2$, the condition $c)$ of Lemma 1 is true, because $\alpha'_1 \geq \alpha_1$, and $\alpha'_1 + \alpha'_2 = \alpha_1 + \alpha_2$. Let us suppose that it is true for $M - 1$. If $\{\alpha'_k\}_{k=1}^M \prec_{LR} \{\alpha_k\}_{k=1}^M$, then it is immediate that $\{\alpha'_k / \sum_{k=1}^{M-1} \alpha'_k\}_{k=1}^{M-1} \prec_{LR} \{\alpha_k / \sum_{k=1}^{M-1} \alpha_k\}_{k=1}^{M-1}$, and by the induction hypothesis, the condition $c)$ of Lemma 1 holds, and thus

$$\alpha'_1 / \sum_{k=1}^{M-1} \alpha'_k \quad \geq \quad \alpha_1 / \sum_{k=1}^{M-1} \alpha_k \tag{A12}$$
$$(\alpha'_1 + \alpha'_2) / \sum_{k=1}^{M-1} \alpha'_k \quad \geq \quad (\alpha_1 + \alpha_2) / \sum_{k=1}^{M-1} \alpha_k$$
$$\ldots \quad \geq \quad \ldots$$
$$(\alpha'_1 + \alpha'_2 + \ldots + \alpha'_{M-1}) / \sum_{k=1}^{M-1} \alpha'_k \quad = \quad (\alpha_1 + \alpha_2 + \ldots + \alpha_{n-1}) / \sum_{k=1}^{M-1} \alpha_k.$$

However, we know that $\alpha'_M \leq \alpha_M$, and thus $\sum_{k=1}^{M-1} \alpha'_k \geq \sum_{k=1}^{M-1} \alpha_k$ because $\sum_{k=1}^{M-1} \alpha'_k = \sum_{k=1}^{M-1} \alpha_k = 1$, and thus we obtain condition $c)$ of Lemma 1. □

References

1. Belzunce, F.; Martinez-Riquelme, C.; Mulero, J. *An Introduction to Stochastic Orders*; Academic Press: Cambridge, MA, USA, 2016; pp. i–ii. doi:10.1016/B978-0-12-803768-3.09977-4. [CrossRef]
2. Shaked, M.; Shanthikumar, G. *Stochastic Orders*; Springer: Berlin/Heidelberg, Germany, 2007; 474p. [CrossRef]
3. Pawlowsky-Glahn, V.; Egozcue, J.; Tolosana-Delgado, R. *Modeling and Analysis of Compositional Data*; J. Wiley & Sons: Hoboken, NJ, USA, 2015.
4. Levy, H. Stochastic Dominance and Expected Utility: Survey and Analysis. *Manag. Sci.* **1992**, *38*, 555–593. [CrossRef]
5. Bullen, P. *Handbook of Means and Their Inequalities*; Springer Science+Business Media: Berlin/Heidelberg, Germany, 2003.
6. Coggeshall, F. The Arithmetic, Geometric, and Harmonic Means. *Q. J. Econ.* **1886**, *1*, 83–86. [CrossRef]
7. Fernández-Villaverde, J. The Econometrics of DSGE Models. *SERIEs* **2010**, *1*, 3–49. [CrossRef]
8. Wikipedia Contributors. Constant Elasticity of Substitution. Available online: https://en.wikipedia.org/wiki/Constant_elasticity_of_substitution (accessed on 14 April 2021).

9.   Yoshida, Y. Weighted Quasi-Arithmetic Means and a Risk Index for Stochastic Environments. *Int. J. Uncertainty, Fuzziness -Knowl.-Based Syst.* **2011**, *19*, 1–16. [CrossRef]
10.   Yoshida, Y. Weighted Quasi-Arithmetic Means: Utility Functions and Weighting Functions. In *Modeling Decisions for Artificial Intelligence, MDAI 2013*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 25–36.
11.   Rényi, A. On Measures of Entropy and Information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*; University of California Press: Berkeley, CA, USA, 1961; Volume 1, pp. 547–561.
12.   Américo, A.; Khouzani, M.; Malacaria, P. Conditional Entropy and Data Processing: An Axiomatic Approach Based on Core-Concavity. *IEEE Trans. Inf. Theory* **2020**, *66*, 5537–5547. [CrossRef]
13.   Wikipedia Contributors. Series and Parallel Springs. Available online: https://en.wikipedia.org/wiki/Series_and_parallel_springs (accessed on 2 April 2021).
14.   Wikipedia Contributors. Resistor. Available online: https://en.wikipedia.org/wiki/Resistor (accessed on 2 April 2021).
15.   Wikibooks Contributors. Fundamentals of Transportation/Traffic Flow, 2017. Available online: https://en.wikibooks.org/wiki/Fundamentals_of_Transportation/Traffic_Flow (accessed on 18 January 2018).
16.   Padois, T.; Doutres, O.; Sgard, F.; Berry, A. On the use of geometric and harmonic means with the generalized cross-correlation in the time domain to improve noise source maps. *J. Acoust. Soc. Am.* **2016**, *140*. [CrossRef]
17.   Sbert, M.; Poch, J. A necessary and sufficient condition for the inequality of generalized weighted means. *J. Inequalities Appl.* **2016**, *2016*, 292. [CrossRef]
18.   Sbert, M.; Havran, V.; Szirmay-Kalos, L.; Elvira, V. Multiple importance sampling characterization by weighted mean invariance. *Vis. Comput.* **2018**, *34*, 843–852. [CrossRef]
19.   M.Sbert.; J.Poch.; M.Chen.; A.Bardera. Some Order Preserving Inequalities for Cross Entropy and Kullback-Leibler Divergence. *Entropy* **2018**, *20*, 959. [CrossRef]
20.   Sbert, M.; Ancuti, C.; Ancuti, C.O.; Poch, J.; Chen, S.; Vila, M. Histogram Ordering. *IEEE Access* **2021**, *9*, 28785–28796. [CrossRef]
21.   Sbert, M.; Yoshida, Y. Stochastics orders on two-dimensional space: Application to cross entropy. In *Modeling Decisions for Artificial Intelligence - MDAI 2020*; Torra, V., Narukawa, Y., Nin, J., Agell, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2020.
22.   Yoshida, Y. Weighted quasi-arithmetic means on two-dimensional regions: An independent case. In *Modeling Decisions for Artificial Intelligence - MDAI 2016*; Torra, V., Narukawa, Y., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 82–93.
23.   Tsallis, C.; Baldovin, F.; Cerbino, R.; Pierobon, P. Introduction to Nonextensive Statistical Mechanics and Thermodynamics. Available online: http://arXiv:cond-mat/0309093v1 (accessed on 13 May 2021).
24.   Hadar, J.; Russell, W. Rules for Ordering Uncertain Prospects. *Am. Econ. Rev.* **1969**, *59*, 25–34.
25.   Zhang, Z.; Sabuncu, M.R. Generalized Cross Entropy Loss for Training Deep Neural Networks with Noisy Labels. *arXiv* **2018**, arXiv:1805.07836.
26.   Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; John Wiley & Sons: New York, NY, USA, 2006.
27.   Dellacherie, C. Quelques commentarires sur les prolongements de capacités. Available online: http://www.numdam.org/article/SPS_1971__5__77_0.pdf (accessed on 22 May 2021).
28.   Renneberg, D. *Non Additive Measure and Integral*; Kluwer Academic Publ.: Dordrecht, The Netherlands, 1994.
29.   Havran, V.; Sbert, M. Optimal Combination of Techniques in Multiple Importance Sampling. In *Proceedings of the 13th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry, VRCAI '14*; ACM: New York, NY, USA, 2014; pp. 141–150. [CrossRef]
30.   Sbert, M.; Havran, V. Adaptive multiple importance sampling for general functions. *Vis. Comput.* **2017**, *33*, 845–855. [CrossRef]
31.   Zhou, X.; Lin, H. Spatial Weights Matrix. In *Encyclopedia of GIS*; Springer: Boston, MA, USA, 2008; p. 1113. [CrossRef]
32.   Smith, M.J.D.; Goodchild, M.F.; Longley, P. *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*; Troubador Publishing Ltd: Leicester, UK, 2015.