Approximate Stochastic Dominance Revisited

Chon Van Le, Olga Kosheleva, and Vladik Kreinovich

Abstract According to decision theory, in general, to recommend the best of possible actions, we need to know, for each possible action, the probabilities of different outcomes, and we also need to know the decision maker's utility function – that describes his/her preferences. For some pairs of probability distributions, however, we can make such a recommendation without knowing the exact form of the utility function – e.g., in financial applications, we only need to know that a larger amount is preferable to a smaller one. Such situations, when we can make decisions based only on the information about probabilities, are known as *stochastic dominance*. The usual analysis of such situations is based in the idealized assumption that any difference in utility, no matter how small, is important. In reality, very small changes in utility value are irrelevant. From this viewpoint, if the utility corresponding to the distribution $F_2(x)$ is always either larger or only slightly smaller than the utility corresponding to $F_1(x)$, then we can still conclude that the second action is better (or of the same quality) than the first action. In this paper, we show how to describe such approximate stochastic dominance in precise terms.

Chon Van Le
International University, Vietnam National University, Ho Chin Minh City, Vietnam
e-mail: lvchon@hcmiu.edu.vn

Olga Kosheleva
Department of Teacher Education, University of Texas at El Paso
500 W. University, El Paso, Texas 79968, USA, e-mail: olgak@utep.edu

Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso
500 W. University, El Paso, Texas 79968, USA, e-mail: vladik@utep.edu

1 Formulation of the Problem

How rational people make decisions? One of the main objectives of econometrics is to help people make decisions. For this purpose, it is necessary to know how people actually make decisions. This has been studied by *decision theory*; see, e.g., [2, 3, 6, 8, 9, 10, 11]. To be more precise, decision theory describes how *rational* people make decisions.

Rational means, e.g., that if a person prefers A to B, and prefers B to C, then this person should also prefer A to C. Real-life decision makers are not always rational in this sense; see, e.g., [5] and references therein. However, if we point out this inconsistency to the decision maker, he/she will correct it. From this viewpoint, it makes sense to restrict ourselves to rational decision makers.

How does decision theory describes decisions of such people? A decision means that the decision maker decides between several possible actions a. We cannot predict the outcome of each action with 100% accuracy – e.g., investing money in a certain stock:

- may increase the amount of money if this stock continues to rise,
- or it may lead to a loss of money if the corresponding company fails.

In general, each action has several possible outcomes $o_1, \ldots, o_n$. Based on prior experiences, we can usually estimate the probabilities $p_1, \ldots, p_n$ of different outcomes. In such situations, decisions of a rational decision maker can be described as follows:

- to each possible outcome o, we assign a numerical value $u(o)$ which is called the *utility* of o; the function that assigns, to each outcome, its utility, is (naturally) called *utility function*;
- to each action a in which we get outcome o_i with probability p_i, we assign the expected value $E[u]$ of the utility, i.e., the value

$$u(a) = E[u] \stackrel{\text{def}}{=} p_1 \cdot u(o_1) + \ldots + p_n \cdot u(o_n);$$

- then, out of several possible actions, the decision maker selects the action with the largest value of its utility $u(a)$.

The utility values can be elicited from the decision maker as follows. We select two ideal outcomes:

- an outcome o_- which is worse than anything that we will actually encounter, and
- an outcome o_+ which is better than anything that we will actually encounter.

Then, for each value p from the interval $[0, 1]$, we can consider a "lottery" – that we will denote by $L(p)$ – in which:

- we get the very good outcome o_+ with probability p, and
- we get the very bad outcome o_- with the remaining probability $1 - p$.

For any actual outcome o:

- when the probability p is small, the lottery $L(p)$ is close to the very bad outcome a_- and is, thus, clearly worse than o; we will denote this by $L(p) < o$;
- when the probability p is close to 1, the lottery $L(p)$ is close to the very good outcome a_+ and is, thus, clearly better than o: $o < L(p)$.

Of course, if $p < p'$, then the lottery $L(p)$ is worse than the lottery $L(p')$ in which the probability of the very good alternative is higher: $L(p) < L(p')$. Thus, if $p < p'$, then:

- $o < L(p)$ implies $o < L(p')$; and
- $L(p') < o$ implies $L(p) < o$.

So, for sufficiently small p, we have $L(p) < o$, and for sufficiently large p, we have $o < L(p)$. There is a threshold separating these two groups of values. This threshold is what is called the utility $u(o)$ of the outcome o. In this case:

- if $p < u(o)$, then $L(p) < o$, and
- if $u(o) < p$, then $o < L(p)$.

The very bad event o_- is equivalent to the lottery $L(0)$, so $u(0_-) = 0$. Similarly, the very good event o_+ is equivalent to the lottery o_+, so $u(o_+) = 1$.

Comment. Thus defined utility is a probability, so it is always located on the interval $[0, 1]$. However, utility does not have to satisfy this restriction: one can easily check that if the function $u(o)$ describes the person's preferences, then, for each a_0 and $a_1 > 0$, the function $u'(o) \stackrel{\text{def}}{=} a_0 + a_1 \cdot u(o)$ describes the same preferences equally well – and for this new utility function, the values can be anywhere between a_0 and $a_0 + a_1$.

Decision making: from general case to financial case. As we have mentioned, in general, each action can be characterized by a probability distribution on the set of all possible outcomes. The decision maker prefers action 2 to action 1 if the expected value $E_2[u]$ of his/her utility under the action-2 distribution is larger than the expected value $E_1[u]$ under the action-1 distribution, i.e., when $E_1[u] \leq E_2[u]$.

In financial situations, the outcome is usually the resulting amount of money x. So, the set of all possible outcomes is the real line, and each action can be characterized by a probability distribution on a real line – as described, e.g., by a cumulative probability distribution function $F(x) \stackrel{\text{def}}{=} \text{Prob}(X \leq x)$.

What if we do not know the utility function? Often, we do not know the utility function $u(x)$ corresponding to a decision maker, all we know is that the more money, the better – i.e., that the utility function $u(x)$ is non-strictly increasing.

In such situations, sometimes, we can still decide which action is better. Thus situation – when one action is better than another based only on their probabilities, is known as *stochastic dominance.*

Comments.

- We will assume that the utility function $u(x)$ is *non-strictly* increasing, i.e., that $x < x'$ implies $u(x) \leq u(x')$. The reason for this is that if a person has a huge amount of money, he/she may not get any additional opportunities if he/she gets some more money, so we have $u(x) = u(x')$ for some $x < x'$.

- In practice, the overall amount of money is bounded. Let us denote the largest possible amount of money by M. Corresponding, the largest amount of debt is also M, which corresponds to the amount of money $-M$.
  - When the value x is equal to $-M$, the situation is as bad as it can be – which, as we have described, corresponds to the very bad outcome o_-, i.e., to utility 0.
  - When the amount x is equal to M, the situation is as good as it can be – which, as we have described, corresponds to the very good outcome o_+ for which the utility is 1.

  Thus, if we restrict ourselves to utilities whose values are between 0 and 1, then we must have $u(-M) = 0$ and $u(M) = 1$.

Stochastic dominance: known result. Here is the known result about stochastic dominance.

Proposition 1. *For every two probability distributions $F_1(x)$ and $F_2(x)$ on the real line, the following two conditions are equivalent to each other:*

- $F_1(x) \ge F_2(x)$ *for all x, and*
- $E_1[u] \le E_2[u]$ *for all possible non-strictly increasing functions* $u(x)$.

Comment. This is a known result, but we will nevertheless present its proof, since the proof of our result is based on this proof. For reader's convenience, all the proofs are placed in a special Proofs section.

Let us reformulate the above known result. In the above result, the inequalities in the two equivalent conditions are opposite to one another. We can make these inequalities going in the same direction if:

- instead of the cumulative distribution function $F(x) = \text{Prob}(X \le x)$,
- we consider the probabilities of the opposite event $S(x) \stackrel{\text{def}}{=} \text{Prob}(X > x)$.

The corresponding function has several names: it is known as *complementary distribution function*, *tail distribution*, *exceedance*, *survival function*, and *reliability function*. One can easily check that $S(x) = 1 - F(x)$, so $F_1(x) \ge F_2(x)$ is equivalent to $S_1(x) \le S_2(x)$. Thus, in terms of this function, Proposition 1 takes the following form:

Proposition 1′. *For every two survival functions $S_1(x) = 1 - F_1(x)$ and $S_2(x) = 1 - F_2(x)$ on the real line, the following two conditions are equivalent to each other:*

- $S_1(x) \le S_2(x)$ *for all x, and*
- $E_1[u] \le E_2[u]$ *for all possible non-strictly increasing functions* $u(x)$.

Need for approximate stochastic dominance. In the above text, we assumed that any difference in utility, no matter how small, is important. In reality, very small changes in utility value are irrelevant. For example, if two cars cost, correspondingly,

\$30,000 and \$30,001, hardly anyone would take this difference seriously (although, as marketing shows, that difference between \$10.00 and \$9.99 does make a psychological difference for actual buyers – but we decide to focus on rational decisions makers.)

From this viewpoint, if the utility corresponding to the distribution $F_2(x)$ is always either larger then – or only slightly smaller than – the utility corresponding to $F_1(x)$, then we can still conclude that the second action is better – or of the same quality – as the first action. How can we describe such approximate stochastic dominance in precise terms?

What is known. Several results about approximate stochastic dominance are known; see, e.g., [1, 7] and references therein.

What we do in this paper. In this paper, we provide another general result about approximate stochastic dominance.

2 Main Result

Definition 1. *Let $M > 0$ be a real number. By a* 0-1-M-utility function*, we mean a non-strictly increasing function $u(x)$ from the interval $[-M, M]$ to the interval $[0, 1]$ for which $u(-M) = 0$ and $u(M) = 1$.*

Definition 2 (reminder)**.** *A function $f(z)$ is called* concave *if for all real numbers $z_1, \ldots, z_n$ and for all non-negative numbers $c_1, \ldots, c_n$ whose sum is 1, we have*

$$f(c_1 \cdot z_1 + \ldots + c_n \cdot z_n) \geq c_1 \cdot f(z_1) + \ldots + c_n \cdot f(z_n).$$

Examples. The following functions are concave:

- linear functions, e.g., function $f(z) = (1 + \varepsilon) \cdot z + \delta$ for some small $\varepsilon > 0$ and $\delta > 0$;
- a function $f(z) = z^a$ for $0 \leq a \leq 1$ as defined for non-negative x; e.g, the function $f(z) = z^{1-\varepsilon}$ for some small positive ε;
- the function $\ln(z)$, etc.

Proposition 2. *Let $M > 0$ be a real number, and let $f(z)$ be a concave function. Then, for every two survival functions $S_1(x) = 1 - F_1(x)$ and $S_2(x) = 1 - F_2(x)$ on the interval $[-M, M]$, the following two conditions are equivalent to each other:*

- *$S_1(x) \leq f(S_2(x))$ for all x, and*
- *$E_1[u] \leq f(E_2[u])$ for all possible 0-1-M-utility functions $u(x)$.*

Comments.

- For $f(z) = z$, we have Proposition 1 – restricted to 0-1-M-utility functions.

- For concave functions $f(z)$ corresponding to small $\varepsilon > 0$ and $\delta > 0$, we get exactly the desired result: that the utility values corresponding to the second distribution is always almost larger than the utility corresponding to F_1 if and only if the survival function corresponding to F_2 is always almost larger than the survival function correspond to $F_1(x)$.
- It is worth mentioning that some versions of stochastic dominance are related to Wasserstein metric and Optimal Transport Theory; see, e.g., [4]. It would be nice to extend our result to this case as well.

3 Proofs

Proof of Proposition 1. For convenience, we will prove the equivalent Proposition $1'$.

1°. First, let us prove that if $E_1[u] \le E_2[u]$ for all u, then $S_1(x) \le S_2(x)$ for all x. Indeed, if we take a function $u(y) = \theta_x(y)$ that is:

- equal to 0 for all $y \le x$ and
- equal to 1 for all $y > x$,

then for all i, the expected value $E_i[u]$ will be exactly equal to $S_i(x)$.

2°. Let us now prove that if $S_1(x) \le S_2(x)$ for all x, then $E_1[u] \le E_2[u]$ for all u.

2.1°. Let us start with the piece-wise constant functions $u(x)$ for which there exist points

$$x_0 = -\infty < x_1 < \ldots < x_n < x_{n+1} = +\infty$$

and values $u_j = u(x_j)$ such that:

- for all $x \in (x_0, x_1]$, we have $u(x) = u(x_0)$,
- for all $x \in (x_1, x_2]$, we have $u(x) = u(x_1)$,
- ...
- for all $x \in (x_j, x_{j+1}]$, we have $u(x) = u(x_j)$,
- ...
- for all $x \in (x_{n-1}, x_n]$, we have $u(x) = u(x_{n-1})$, and
- for all $x \in (x_n, x_{n+1}]$, we have $u(x) = u(x_n)$.

It is easy to check that the resulting function $u(x)$ is a linear combination of the functions $\theta_{x_j}(x)$:

$$u(x) = u(x_0) + (u(x_1) - u(x_0)) \cdot \theta_{x_1}(x) + \ldots + (u(x_j) - u(x_{j-1})) \cdot \theta_{x_j}(x) + \ldots +$$

$$(u(x_n) - u(x_{n-1})) \cdot \theta_{x_n}(x).$$

The mean value of a linear combination is equal to the linear combination of mean values, so

$$E_i[u] = u(x_0) + (u(x_1) - u(x_0)) \cdot E_i[\theta_{x_1}] + \ldots + (u(x_j) - u(x_{j-1})) \cdot E_i[\theta_{x_j}] + \ldots +$$

$$(u(x_n) - u(x_{n-1})) \cdot E_i[\theta_{x_n}].$$

Here, for any j, we have $E_j[\theta_{x_j}(x)] = S_j(x_n)$, thus we have

$$E_i[u] = u(x_0) + (u(x_1) - u(x_0)) \cdot S_i(x_1) + \ldots + (u(x_j) - u(x_{j-1})) \cdot S_i(x_j) + \ldots +$$

$$(u(x_n) - u(x_{n-1})) \cdot S_i(x_n). \quad (1)$$

For each j, we have

$$S_1(x_j) \le S_2(x_j). \quad (2)$$

Since the utility function is non-strictly increasing and $x_{j-1} < x_j$, the difference $u(x_j) - u(x_{j-1})$ is non-negative. Thus, multiplying both sides of the inequality (1) by the non-negative difference $u(x_j) - u(x_{j-1})$, we conclude that

$$(u(x_j) - u(x_{j-1})) \cdot S_1(x_j) \le (u(x_j) - u(x_{j-1})) \cdot S_2(x_j). \quad (3)$$

Adding up all these inequalities, adding $u(x_0)$ to both sides, and taking into account the formula (1), we conclude that indeed

$$E_1[u] \le E_2[u]. \quad (4)$$

2.2°. For each non-strictly increasing function $u(x)$ and for each selection of values $x_1 < \ldots < x_n$, we can form a piece-wise constant function as described in Part 2.1 of this proof. As we take denser and denser set of values x_j covering larger and larger domain, we will get functions that are closer and closer to $u(x)$. So in the limit, from the inequality (4) for these auxiliary functions, we get the same inequality for the original function $u(x)$.

The proposition is proven.

Proof of Proposition 2.

1°. Similarly to Part 1 of the previous proof, we can show that if $E_1[u] \le f(E_2[u])$ for all 0-1-M utility functions u, then $S_1(x) \le f(S_2(x))$ for all x.

2°. Let us now prove that if $S_1(x) \le f(S_2(x))$ for all x, then $E_1[u] \le f(E_2[u])$ for all 0-1-M utility functions u. Similarly to the previous proof, let us start with the case when the function $u(x)$ is piece-wise constant. In this case, we have $x_0 = -M$ and $x_{n+1} = M$. By definition of a 0-1-M utility function, we have $u(x_0) = 0$ and $u(x_{n+1}) = 1$. Then, the formula (1) takes a somewhat simplified form

$$E_i[u] = u(x_1) \cdot S_i(x_1) + \ldots + (u(x_j) - u(x_{j-1})) \cdot S_i(x_j) + \ldots +$$

$$(u(x_n) - u(x_{n-1})) \cdot S_i(x_n) + (1 - u(x_n)) \cdot S_i(x_{n+1}). \quad (5)$$

From the inequalities $S_1(x_j) \le f(S_2(x_j))$, multiplying each of them by $u(x_j) - u(x_{j-1}) \ge 0$ and adding up the results, we conclude that

$$E_1[u] = u(x_1) \cdot S_1(x_1) + \ldots + (u(x_j) - u(x_{j-1})) \cdot S_1(x_j) + \ldots +$$

$$(u(x_n) - u(x_{n-1})) \cdot S_1(x_n) + (1 - u(x_n)) \cdot S_1(x_{n+1}) \le$$
$$u(x_1) \cdot f(S_2(x_1)) + \ldots + (u(x_j) - u(x_{j-1})) \cdot f(S_2(x_j)) + \ldots +$$
$$(u(x_n) - u(x_{n-1})) \cdot f(S_2(x_n)) + (1 - u(x_n)) \cdot f(S_2(x_{n+1})). \quad (6)$$

Since the function $f(z)$ is concave, and the non-negative coefficients $u(x_j)-u(x_{j-1})$ add up to 1, we conclude that

$$u(x_1) \cdot f(S_2(x_1)) + \ldots + (u(x_j) - u(x_{j-1})) \cdot f(S_2(x_j)) + \ldots +$$
$$(u(x_n) - u(x_{n-1})) \cdot f(S_2(x_n)) + (1 - u(x_n)) \cdot f(S_2(x_{n+1})) \le$$
$$f(u(x_1) \cdot S_2(x_1) + \ldots + (u(x_j) - u(x_{j-1})) \cdot S_2(x_j) + \ldots +$$
$$(u(x_n) - u(x_{n-1})) \cdot S_2(x_n) + (1 - u(x_n)) \cdot S_2(x_{n+1})) = f(E_2[u]). \quad (7)$$

From the inequalities (6) and (7), we conclude that $E_1[u] \le f(E_2[u])$ for piece-wise constant utility functions.

Similarly to the previous proof, we can then take the limit and thus prove that the desired inequality $E_1[u] \le f(E_2[u])$ holds for all 0-1-M utility functions. The proposition is proven.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), HRD-1834620 and HRD-2034030 (CAHSI Includes), EAR-2225395 (Center for Collective Impact in Earthquake Science C-CIES), and by the AT&T Fellowship in Information Technology.

It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, by a grant from the Hungarian National Research, Development and Innovation Office (NRDI), and by the Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Thailand.

The authors are thankful to Hung T. Nguyen for his encouragement.

References

1. R. Bruni, F. Cesarone, A. Scozzari, and F. Tardella, "On exact and approximate stochastic dominance strategies for portfolio selection", *European Journal of Operational Research*, 2017, Vol. 259, pp. 322–329.
2. P. C. Fishburn, *Utility Theory for Decision Making*, John Wiley & Sons Inc., New York, 1969.
3. P. C. Fishburn, *Nonlinear Preference and Utility Theory*, The John Hopkins Press, Baltimore, Maryland, 1988.

4. Z. Hosseini-Nodeh, R. Khanjani-Shiraz, and P. M. Pardalos, “Distributionally robust portfolio optimization with second-order stochastic dominance based on Wasserstein metric”, *Information Sciences*, 2022, Vol. 613, pp. 828–852.
5. D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus, and Giroux, New York, 2011.
6. V. Kreinovich, “Decision making under interval uncertainty (and beyond)", In: P. Guo and W. Pedrycz (eds.), *Human-Centric Decision-Making Models for Social Sciences*, Springer Verlag, 2014, pp. 163–193.
7. V. Kreinovich, H. T. Nguyen, and S. Sriboonchitta, “What if we only have approximate stochastic dominance?", In: V.-N. Huynh, V. Kreinovich, S. Sriboonchitta, and K. Suriya (eds.), *Econometrics of Risk*, Springer Verlag, Berlin, Heidelberg, 2015, pp. 53–61.
8. R. D. Luce and R. Raiffa, *Games and Decisions: Introduction and Critical Survey*, Dover, New York, 1989.
9. H. T. Nguyen, O. Kosheleva, and V. Kreinovich, “Decision making beyond Arrow's ‘impossibility theorem’, with the analysis of effects of collusion and mutual attraction", *International Journal of Intelligent Systems*, 2009, Vol. 24, No. 1, pp. 27–47.
10. H. T. Nguyen, V. Kreinovich, B. Wu, and G. Xiang, *Computing Statistics under Interval and Fuzzy Uncertainty*, Springer Verlag, Berlin, Heidelberg, 2012.
11. H. Raiffa, *Decision Analysis*, McGraw-Hill, Columbus, Ohio, 1997.