Stochastic Dominance: Cases of Interval and P-Box Uncertainty

Kittawit Autchariyapanikul[1],
Olga Kosheleva[2], and Vladik Kreinovich[2]

[1]Faculty of Economics, Maijo University
Chiang Mai, Thailand, kittawit_a@mju.ac.th
[2]University of Texas at El Paso, El Paso, Texas 79968, USA
{olgak,vladik}@utep.edu

1. Decision making according to decision theory: a brief reminder

- Decision theory describes decisions of a *rational* decision maket.
- This decision maker's decisions satisfy commonsense conditions of rationality.
- For example, if a rational decision maker prefers A to B and prefers B to C, then this decision maker should prefer A to C.
- It is known that preferences of a rational decision maker can be described by a function $u(a)$ called *utility* such that:
  - the decision maker prefers a to b if and only if $u(a) > u(b)$, and
  - the utility of a situation in which we can get outcome a_i with probability p_i is equal to $p_1 \cdot u(a_1) + \ldots + p_n \cdot u(a_n)$.
- Stochastic dominance theory deals with the case when each outcome a_i is characterized by a numerical value – monetary gain.

2. Decision making according to decision theory: a brief reminder (cont-d)

- In this case, each alternative is described by a probabilistic uncertainty.
- So, we have a probability distribution on the set of all real numbers.
- For example, in the above case, we have a probability distribution:
  - that is located on the set $\{a_1, \ldots, a_n\}$ and
  - for which the probability of each value a_i is equal to p_i.
- In general, we can describe each probability distribution by a cumulative distribution function $F(x) = \text{Prob}(a \le x)$.
- In this case, the utility of this alternative is equal to the expected value $\int u(x)\, dF(x)$ of the utility function.

3. Stochastic dominance: a brief reminder

- In some practical situations, we do not know the decision maker's utility function.
- We only have *some* information about the utility function.
- In some such cases, we can sometimes conclude that one alternative is better than the other one.
- Such cases are known as cases of *stochastic dominance.*
- In some cases, all we know about the utility function is that this function is (non-strictly) increasing: if $a \le b$ then $u(a) \le u(b)$.
- In this case, for every two probability distributions $F(x)$ and $G(x)$, the following two conditions are equivalent:
  - for every non-strictly increasing function $u(x)$, the utility corresponding to $F(x)$ is $\ge$ the utility corresponding to $G(x)$;
  - for all x, we have $G(x) \le F(x)$.

4. Stochastic dominance: a brief reminder (cont-d)

- Sometimes, we also know that the decision maker is risk-averse, i.e.:
  - for each lottery in which the person gets amounts x_i with probabilities p_i,
  - the decision would prefer to receive the expected value $p_1 \cdot x_1 + \ldots + p_n \cdot x_n$ than to participate in this lottery.
- In view of the above-described relation between expected utility and decisions, this means that for all such cases, we have

$$u(p_1 \cdot x_1 + \ldots + p_n \cdot x_n) \geq p_1 \cdot u(x_1) + \ldots + p_n \cdot u(x_n).$$

- Such functions $u(x)$ are known as *concave*.

5. Stochastic dominance: a brief reminder (cont-d)

- In this case, for every two probability distributions $F(x)$ and $G(x)$, the following two conditions are equivalent:
  - for every non-strictly increasing concave function $u(x)$, the utility corresponding to $F(x)$ is $\ge$ the utility corresponding to $G(x)$;
  - for all x, we have

$$\int_{-\infty}^{x} G(t)\,dt \le \int_{-\infty}^{x} F(t)\,dt.$$

6. Need to consider interval and p-box uncertainty

- The above results assume that for each possible action, we know the probabilities of different outcomes.
- In practice, we often also have only partial information about these probabilities.
- Sometimes, we have no information at all about the probabilities, we only know the range $[\underline{x}, \overline{x}]$ of possible values.
- This case is known as the case of *interval uncertainty*.
- Sometimes, we have partial information about the cumulative distribution function $F(x)$.
- Uncertainty usually means that for each x:
  - instead of the exact value $F(x)$,
  - we only know the range $\left[\underline{F}(x), \overline{F}(x)\right]$ of possible values.

7. Need to consider interval and p-box uncertainty (cont-d)

- In this case:
  - for each tuple of values $(x_1, \ldots, x_n)$,
  - the set of possible values of the corresponding tuple $(F(x_1), \ldots, F(x_n))$ is a box

$$\left[\underline{F}(x_1), \overline{F}(x_1)\right] \times \ldots \times \left[\underline{F}(x_n), \overline{F}(x_n)\right].$$

- This box is called *probability box*, or *p-box*, for short.

8. How to make decisions under interval uncertainty

- According to decision theory:
  - to make decisions under interval uncertainty,
  - a decision maker has: to select his/her degree α_H of optimism-pessimism – a number from the interval $[0,1]$.
- Then, the utility of an interval $[\underline{x},\overline{x}]$ is determined as

$$\alpha_H \cdot u(\overline{x}) + (1-\alpha_H)\cdot u(\underline{x}).$$

- The name of this degree comes from the fact that when $\alpha_H = 1$, the utility is equal to $u(\overline{x})$.
- In this case, we only take into account the best-case scenario.
- We ignore the possibility that the outcome can be worse.
- This is clearly the case of pure optimism.

9. How to make decisions under interval uncertainty (cont-d)

- On the other hand, when $\alpha_H = 0$, the utility is equal to $u(\underline{x})$.
- In this case, we only take into account the worst-case scenario, and we ignore the possibility that the outcome can be better.
- This is clearly the case of pure pessimism.
- Of course, these two are extreme cases.
- In practice, decision makers take both the best-case and the worst-case scenarios into account, i.e., they select values $\alpha_H \in (0, 1)$.

10. How to make decisions under p-box uncertainty

- As we have mentioned, for a given probability distribution $F(x)$, the utility is equal to $\int u(x)\,dF(x)$.
- By using integration by part, we can reduce this integral to the form

$$-\int u'(x)\cdot F(x)\,dx.$$

- Since the function $u(x)$ is non-strictly increasing, we have $u'(x)\ge 0$.
- Thus, if $F(x)\le G(x)$, then we have

$$\int u(x)\,dF(x) = -\int u'(x)\cdot F(x)\,dx \ge -\int u'(x)\cdot G(x)\,dx = \int u(x)\,dG(x).$$

- As a result, when $\underline{F}(x)\le F(x)\le \overline{F}(x)$, we have

$$\int u(x)\,d\underline{F}(x) \ge \int u(x)\,dF(x) \ge \int u(x)\,d\overline{F}(x).$$

11. How to make decisions under p-box uncertainty (cont-d)

- Thus, in the p-box case, possible values of utility form an interval

$$\left[\int u(x)\,d\overline{F}(x), \int u(x)\,d\underline{F}(x)\right].$$

- So, according to decision theory, we need to select an alternative for which the following value is the largest:

$$\alpha_H \cdot \int u(x)\,d\underline{F}(x) + (1-\alpha_H) \cdot \int u(x)\,d\underline{F}(x).$$

- In this talk, we extend the known stochastic dominance results to the cases of interval and p-box uncertainty.
- Namely, we describe all the case with interval and p-box uncertainty when:
  - we can make a decision
  - without knowing the exact shape of the utility function.

12. Results: case of interval uncertainty

- Let $\alpha_H \in (0, 1)$ be given.
- Then, for every two intervals $\mathbf{x} = [\underline{x}, \overline{x}]$ and $\mathbf{y} = \left[\underline{y}, \overline{y}\right]$, the following two conditions are equivalent:
  - for every non-strictly increasing function $u(x)$, the utility corresponding to $\mathbf{x}$ is $\geq$ the utility corresponding to $\mathbf{y}$;
  - we have $\underline{x} \geq \underline{y}$ and $\overline{x} \geq \overline{y}$.
- For every two intervals $\mathbf{x} = [\underline{x}, \overline{x}]$ and $\mathbf{y} = \left[\underline{y}, \overline{y}\right]$, the following two conditions are equivalent:
  - for every non-strictly increasing concave function $u(x)$, the utility corresponding to $\mathbf{x}$ is $\geq$ the utility corresponding to $\mathbf{y}$;
  - we have $\underline{x} \geq \underline{y}$ and $\alpha_H \cdot \overline{x} + (1 - \alpha_H) \cdot \underline{x} \geq \alpha_H \cdot \overline{y} + (1 - \alpha_H) \cdot \underline{y}$.

13. Results: case of p-box uncertainty

- Let $\alpha_H \in (0,1)$ be given.
- Then, for every two p-boxes $\mathbf{F}(x) = \left[\underline{F}(x), \overline{F}(x)\right]$ and $\mathbf{G}(x) = \left[\underline{G}(x), \overline{G}(x)\right]$, the following two conditions are equivalent:
  - for every non-strictly increasing function $u(x)$, the utility corresponding to $\mathbf{F}(x)$ is $\geq$ the utility corresponding to $\mathbf{G}(x)$;
  - for all x, we have $G(x) \leq F(x)$, where

$$F(x) \stackrel{\text{def}}{=} \alpha_H \cdot \underline{F}(x) + (1 - \alpha_H) \cdot \overline{F}(x) \text{ and}$$

$$G(x) \stackrel{\text{def}}{=} \alpha_H \cdot \underline{G}(x) + (1 - \alpha_H) \cdot \overline{G}(x).$$

14. Results: case of p-box uncertainty (cont-d)

- For every two p-boxes $\mathbf{F}(x) = \left[\underline{F}(x), \overline{F}(x)\right]$ and $\mathbf{G}(x) = \left[\underline{G}(x), \overline{G}(x)\right]$, the following two conditions are equivalent:
  - for every non-strictly increasing concave function $u(x)$, the utility corresponding to $\mathbf{F}(x)$ is $\geq$ the utility corresponding to $\mathbf{G}(x)$;
  - for all x, we have

$$\int_{-\infty}^{x} G(t)\,dt \leq \int_{-\infty}^{x} F(t)\,dt,$$

    where

$$F(x) \stackrel{\text{def}}{=} \alpha_H \cdot \underline{F}(x) + (1 - \alpha_H) \cdot \overline{F}(x) \text{ and}$$

$$G(x) \stackrel{\text{def}}{=} \alpha_H \cdot \underline{G}(x) + (1 - \alpha_H) \cdot \overline{G}(x).$$

15. Proof of the interval result for monotonic $u(x)$

- To prove the desired equivalence, we will prove the following two statements:
  - that if the second condition is satisfied, then the first condition is also satisfied, and
  - that if the second condition is not satisfied, then the first condition is also nor satisfied.
- Let us prove them one by one.
- Let us first prove that if the second condition is satisfied, then the first condition is also satisfied.
- Indeed, from $\underline{x} \geq \underline{y}$, we conclude that

$$(1 - \alpha_H) \cdot \underline{x} \geq (1 - \alpha_H) \cdot \underline{y}.$$

16. Proof of the interval result for monotonic $u(x)$ (cont-d)

- Similarly, from $\overline{x} \geq \overline{y}$, we conclude that

$$\alpha_H \cdot \overline{x} \geq \alpha_H \cdot \overline{y}.$$

- By adding the two centered inequalities, we conclude that

$$u(\mathbf{x}) = \alpha_H \cdot \overline{x} + (1 - \alpha_H) \cdot \underline{x} \geq \alpha_H \cdot \overline{y} + (1 - \alpha_H) \cdot \underline{y} = u(\mathbf{y}).$$

- Let us now prove that if the second condition is not satisfied, then the first condition is also not satisfied.
- The second condition consists of two inequalities.
- Thus, the fact that it is not satisfied means that one of these inequalities is false, i.e.:
  - either $\underline{x} < \underline{y}$
  - or $\overline{x} < \overline{y}$.
- We will consider these two cases one by one.

17. Proof of the interval result for monotonic $u(x)$ (cont-d)

- Let us first consider the case when $\underline{x} < \underline{y}$.
- In this case, we can consider the following non-strictly increasing utility function:
  - for $x < \underline{y}$, we have $u(x) = 0$, and
  - for $x \geq \underline{y}$, we take $u(x) = 1$.
- In this case, since $\overline{y} \geq \underline{y}$, we have $u(\overline{y}) = u\left(\underline{y}\right) = 1$ and thus,

$$u(\mathbf{y}) = \alpha_H \cdot u(\overline{y}) + (1 - \alpha_H) \cdot u\left(\underline{y}\right) = \alpha_H \cdot 1 + (1 - \alpha_H) \cdot 1 = 1.$$

- On the other hand, since $\underline{x} < \underline{y}$, we have $u(\underline{x}) = 0$, thus

$$(1 - \alpha_H) \cdot u(\underline{x}) = 0.$$

- For our utility function, we have $u(x) \leq 1$ for all x, in particular, we have $u(\overline{x}) \leq 1$, thus

$$\alpha_H \cdot u(\overline{x}) \leq \alpha_H < 1.$$

18. Proof of the interval result for monotonic $u(x)$ (cont-d)

- By adding the two centered inequalities, we conclude that

$$u(\mathbf{x}) = \alpha_H \cdot u(\overline{x}) + (1 - \alpha_H) \cdot u(\underline{x}) \le \alpha_H < 1 = u(\mathbf{y}).$$

- So, in this case, the first condition is also not satisfied.
- Let us now consider the case when $\overline{x} < \overline{y}$.
- In this case, we can consider the following non-strictly increasing utility function:
  - for $x < \overline{y}$, we have $u(x) = 0$, and
  - for $x \ge \overline{y}$, we take $u(x) = 1$.
- In this case, we have $u(\overline{y}) = 1$ and thus,

$$\alpha_H \cdot u(\overline{y}) = \alpha_H > 0.$$

- For this utility function, $u(x) \ge 0$ for all x, in particular, $u\left(\underline{y}\right) \ge 0$, thus

$$(1 - \alpha_H) \cdot u\left(\underline{y}\right) \ge 0.$$

19. Proof of the interval result for monotonic $u(x)$ (cont-d)

- By adding the two last centered inequalities, we conclude that

$$u(\mathbf{y}) = \alpha_H \cdot u(\overline{y}) + (1 - \alpha_H) \cdot u\left(\underline{y}\right) \geq \alpha_H > 0.$$

- On the other hand, since $\underline{x} \leq \overline{x} < \overline{y}$, we have $u(\underline{x}) = u(\overline{x}) = 0$, thus

$$u(\mathbf{x}) = \alpha_H \cdot u(\overline{x}) + (1 - \alpha_H) \cdot u(\underline{x}) = 0 < \alpha_H \leq u(\mathbf{y}).$$

- So, in this case, the first condition is also not satisfied.

20. Proof of the interval result for concave $u(x)$

- To prove the desired equivalence, we will prove the following two statements:
  - that if the second condition is satisfied, then the first condition is also satisfied, and
  - that if the second condition is not satisfied, then the first condition is also nor satisfied.
- Let us prove them one by one.
- Let us first prove that if the second condition is satisfied, then the first condition is also satisfied.
- For this purpose, we will consider two possible cases:
  - when $\overline{x} \geq \overline{y}$ and
  - when $\overline{x} < \overline{y}$.
- We will consider these two cases one by one.

21. Proof of the interval result for concave $u(x)$ (cont-d)

- If $\overline{x} \geq \overline{y}$, then:
  - taking into account that $\underline{x} \geq \underline{y}$ and that the utility function $u(x)$ is non-strictly increasing,
  - we conclude that $u(\overline{x}) \geq u(\overline{y})$ and $u(\underline{x}) \geq u\left(\underline{y}\right)$.
- Multiplying the first inequality by $\alpha_H > 0$ and the second one by $1 - \alpha_H > 0$ and adding the resulting inequalities, we conclude that

$$\alpha_H \cdot \overline{x} + (1 - \alpha_H) \cdot \underline{x} \geq \alpha_H \cdot \overline{y} + (1 - \alpha_H) \cdot \underline{.}$$

- So indeed, that the utility of the interval $\mathbf{x}$ is larger than or equal to the utility of the interval $\mathbf{y}$.
- Let us now consider the case when $\overline{x} < \overline{y}$, i.e., when $\underline{y} \leq \underline{x} \leq \overline{x} < \overline{y}$ and when

$$\alpha_H \cdot \overline{x} + (1 - \alpha_H) \cdot \underline{x} \geq \alpha_H \cdot \overline{y} + (1 - \alpha_H) \cdot \underline{y}.$$

22. Proof of the interval result for concave $u(x)$ (cont-d)

- Here,

$$\underline{x} = \underline{y} + \frac{\underline{x} - \underline{y}}{\overline{y} - \underline{y}} \cdot (\overline{y} - \underline{y}) = \frac{\underline{x} - \underline{y}}{\overline{y} - \underline{y}} \cdot \overline{y} + \frac{\overline{y} - \underline{x}}{\overline{y} - \underline{y}} \cdot \underline{y}.$$

- Since the utility function is concave, we conclude that

$$u(\underline{x}) \geq \frac{\underline{x} - \underline{y}}{\overline{y} - \underline{y}} \cdot u(\overline{y}) + \frac{\overline{y} - \underline{x}}{\overline{y} - \underline{y}} \cdot u\left(\underline{y}\right), \text{ so}$$

$$u(\underline{x}) \geq u\left(\underline{y}\right) + \frac{\underline{x} - \underline{y}}{\overline{y} - \underline{y}} \cdot \left(u(\overline{y}) - u\left(\underline{y}\right)\right).$$

- Similarly,

$$\overline{x} = \underline{y} + \frac{\overline{x} - \underline{y}}{\overline{y} - \underline{y}} \cdot (\overline{y} - \underline{y}) = \frac{\overline{x} - \underline{y}}{\overline{y} - \underline{y}} \cdot \overline{y} + \frac{\overline{y} - \overline{x}}{\overline{y} - \underline{y}} \cdot \underline{y}.$$

23. Proof of the interval result for concave $u(x)$ (cont-d)

- Since the utility function is concave, we conclude that

$$u(\overline{x}) \geq \frac{\overline{x}-\underline{y}}{\overline{y}-\underline{y}} \cdot u(\overline{y}) + \frac{\overline{y}-\overline{x}}{\overline{y}-\underline{y}} \cdot u\left(\underline{y}\right), \text{ so}$$

$$u(\overline{x}) \geq u\left(\underline{y}\right) + \frac{\overline{x}-\underline{y}}{\overline{y}-\underline{y}} \cdot \left(u(\overline{y}) - u\left(\underline{y}\right)\right).$$

- Multiplying the second inequality by α_H and the first one by $1-\alpha_H$ and adding the resulting inequalities, we conclude that

$$u(\mathbf{x}) \geq u\left(\underline{y}\right) + \frac{x-\underline{y}}{\overline{y}-\underline{y}} \cdot \left(u(\overline{y}) - u\left(\underline{y}\right)\right).$$

- Here, we denoted $x \stackrel{\text{def}}{=} \alpha_H \cdot \overline{x} + (1-\alpha_H) \cdot \underline{x}$.
- We have assumed that

$$x \geq y \stackrel{\text{def}}{=} \alpha_H \cdot \overline{y} + (1-\alpha_H) \cdot \underline{y}.$$

24. Proof of the interval result for concave $u(x)$ (cont-d)

- Thus, we can conclude that

$$u(\mathbf{x}) \geq u\left(\underline{y}\right) + \frac{y - \underline{y}}{\overline{y} - \underline{y}} \cdot \left(u(\overline{y}) - u\left(\underline{y}\right)\right).$$

- Here, by definition of y, we have

$$\frac{y - \underline{y}}{\overline{y} - \underline{y}} = \alpha_H.$$

- Thus, the above inequality takes the form

$$u(\mathbf{x}) \geq u\left(\underline{y}\right) + \alpha_H \cdot (u(\overline{y}) - u\left(\underline{y}\right)) = \alpha_H \cdot u(\overline{y}) + (1 - \alpha_H) \cdot u\left(\underline{y}\right).$$

- So indeed $u(\mathbf{x}) \geq u(\mathbf{y})$.
- Let us now prove that if the second condition is not satisfied, then the first condition is also not satisfied.
- The second condition consists of two inequalities.

25. Proof of the interval result for concave $u(x)$ (cont-d)

- Thus, the fact that it is not satisfied means that one of these inequalities is false, i.e.:
  - either $\underline{x} < \underline{y}$
  - or $\alpha_H \cdot \overline{x} + (1 - \alpha_H) \cdot \underline{x} < \alpha_H \cdot \overline{y} + (1 - \alpha_H) \cdot \underline{y}$.
- We will consider these two cases one by one.
- Let us first consider the case when $\underline{x} < \underline{y}$.
- In this case, we can consider the following non-strictly increasing concave utility function:
  - for $x \le \underline{y}$, we have $u(x) = x$, and
  - for $x \ge \underline{y}$, we take $u(x) = \underline{y}$.
- In this case, since $\underline{y} \ge \overline{y}$, we have $u\left(\underline{y}\right) = u(\overline{y}) = \underline{y}$ and thus,

$$u(\mathbf{y}) = \alpha_H \cdot u(\overline{y}) + (1 - \alpha_H) \cdot u\left(\underline{y}\right) = \alpha_H \cdot \underline{y} + (1 - \alpha_H) \cdot \underline{y} = \underline{y}.$$

26. Proof of the interval result for concave $u(x)$ (cont-d)

- On the other hand, by definition of our utility function, since $\underline{x} < \underline{y}$, we have $u(\underline{x}) = \underline{x} < \underline{y}$.
- Since $\alpha_H \in (0, 1)$, we have $1 - \alpha_H > 0$ and thus,

$$(1 - \alpha_H) \cdot u(\underline{x}) < (1 - \alpha_H) \cdot \underline{y}.$$

- For our utility function, we have $u(x) \leq \underline{y}$ for all x, in particular, $u(\overline{x}) \leq \underline{y}$, thus

$$\alpha_H \cdot u(\overline{x}) \leq \alpha_H \cdot \underline{y}.$$

- By adding two centered inequalities, we conclude that

$$u(\mathbf{x}) = \alpha_H \cdot u(\overline{x}) + (1 - \alpha_H) \cdot u(\underline{x}) < \alpha_H \cdot \underline{y} + (1 - \alpha_H) \cdot \underline{y} = \underline{y} = u(\mathbf{y}).$$

- So indeed, the first condition is not satisfied.

27. Proof of the interval result for concave $u(x)$ (cont-d)

- Let us now consider the case when

$$\alpha_H \cdot \overline{x} + (1 - \alpha_H) \cdot \underline{x} < \alpha_H \cdot \overline{y} + (1 - \alpha_H) \cdot \underline{y}.$$

- In this case, we can consider the utility function $u(x) = x$.
- This function is increasing and concave.
- However, for this function, the above inequality means that

$$u(\mathbf{x}) < u(\mathbf{y}).$$

- Thus, in this case, the first condition is also not satisfied.
- The proposition is thus proven.

28. Proof of p-box results

- According to decision theory, the utility of a p-box $\left[\underline{F}(x), \overline{F}(x)\right]$ is equal to

$$\alpha_H \cdot \int u(x)\, d\underline{F}(x) + (1-\alpha_H) \cdot \int u(x)\, d\overline{F}(x).$$

- By applying integration by parts, we can transform each of these integrals into an equivalent form

$$-\int u(x) \cdot \underline{F}(x)\, dx \text{ and } -\int u(x) \cdot \overline{F}(x)\, dx.$$

- Thus, the utility of a p-box takes the form

$$\alpha_H \cdot \left(-\int u(x) \cdot \underline{F}(x)\, dx\right) + (1-\alpha_H) \cdot \left(-\int u(x) \cdot \overline{F}(x)\, dx\right).$$

- We can conclude that this expression is equal to

$$-\int u(x) \cdot (\alpha_H(x) \cdot \underline{F}(x) + (1-\alpha_H) \cdot \overline{F}(x))\, dx.$$

29. Proof of p-box results (cont-d)

- By definition of the combination $F(x)$, this means

$$-\int u(x) \cdot F(x)\,dx = \int u(x)\,dF(x).$$

- Thus, the utility of a p-box $\left[\underline{F}(x), \overline{F}(x)\right]$ is equal to the utility corresponding to the probability distribution

$$F(x) = \alpha_H \cdot \underline{F}(x) + (1 - \alpha_H) \cdot \overline{F}(x).$$

- Because of this, Propositions dealing with p-boxes follow from the propositions that deal with probability distributions.

30. Acknowledgments

This work was supported in part by:

- National Science Foundation grants 1623190, HRD-1834620, HRD-2034030, and EAR-2225395;
- AT&T Fellowship in Information Technology;
- program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and
- a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

The authors are greatly thankful to Arnab Bhattacharjee for valuable discussions.