

## How to Compare Different Range Estimations: A Symmetry-Based Approach

O. Kosheleva and V. Kreinovich

University of Texas at El Paso, 500 W. University, El Paso, Texas 79968, USA;  
PH 1-915-747-6951; FAX 1-915-747-5030; emails: olgak@utep.edu, vladik@utep.edu

### ABSTRACT

To estimate the future value of a quantity  $y$  or a present value of a difficult-to-measure quantity  $y$ , we measure easier-to-measure quantities  $x_1, \dots, x_n$  related to  $y$  by a known dependence  $y = f(x_1, \dots, x_n)$ , and then use this dependence to compute an estimate  $\tilde{y}$  for  $y$ . Measurements are never absolute accurate: each measurement result  $\tilde{x}_i$  is, in general, different from the actual (unknown) value of the quantity  $x_i$ . Because of this, the estimate  $\tilde{y} = f(\tilde{x}_1, \dots, \tilde{x}_n)$  is, in general, different from the actual value  $y$ . In many practical situations, we only the upper bound  $\Delta_i$  on the measurement error  $\Delta x_i \stackrel{\text{def}}{=} \tilde{x}_i - x_i$ :  $|\Delta x_i| \leq \Delta_i$ . In this case, possible values of  $x_i$  form an interval  $[\tilde{x}_i - \Delta_i, \tilde{x}_i + \Delta_i]$ . Different values  $x_i$  from these intervals lead, in general, to different values  $y = f(x_1, \dots, x_n)$ ; thus, we need to find the range  $\mathbf{y}$  of all such values.

Computing this range is, in general, NP-hard (computationally intractable), so it is only feasibly possible to compute an *enclosure*  $\mathbf{Y} \supseteq \mathbf{y}$ . In the dynamic case, when we want to predict the values  $y(t)$  for different moments of time  $t$ , we have enclosures  $\mathbf{Y}(t)$  corresponding to different moments of time. Different methods produce different enclosures  $\mathbf{Y}(t)$ ,  $\mathbf{Y}'(t)$ , etc.; which one should we choose? In this paper, we show how symmetry-based ideas can help in making this choice.

### FORMULATION OF THE PROBLEM

**Need for data processing.** In many practical situations, we are interested in the future value of a quantity  $y$  or in its current value which is not easy to measure directly. In such situations, we find easier-to-measure quantities  $x_1, \dots, x_n$  which are related to  $y$  in a known way  $y = f(x_1, \dots, x_n)$ , and then use the measured values  $\tilde{x}_i$  of the auxiliary quantities  $x_i$  to estimate  $y$  as  $\tilde{y} = f(\tilde{x}_1, \dots, \tilde{x}_n)$ . This is usually called *data processing*.

**Need to take measurement uncertainty into account.** Measurements are never absolutely accurate; thus, the measured values  $\tilde{x}_i$  are somewhat different from the actual (unknown) values  $x_i$  and therefore, even when the model is exact, the estimate  $\tilde{y} = f(\tilde{x}_1, \dots, \tilde{x}_n)$  is, in general, different from the desired value  $y = f(x_1, \dots, x_n)$ . It is important to gauge this difference.

**Interval uncertainty: important case of measurement uncertainty.** Often, the only information that we have about each measurement error  $\Delta x_i \stackrel{\text{def}}{=} \tilde{x}_i - x_i$  is the upper bound  $\Delta_i$ , for which  $|\Delta x_i| \leq \Delta_i$ ; see, e.g., (8). In this case, the only information that we know about each  $x_i$  is that  $x_i$  belongs to the interval  $\mathbf{x}_i = [\tilde{x}_i - \Delta_i, \tilde{x}_i + \Delta_i]$ . We therefore need to find the range  $\mathbf{y}$  of  $f(x_1, \dots, x_n)$  when  $x_i \in \mathbf{x}_i$ .

The exact computation of this range is known to be NP-hard (5), so, in general, we estimate the *enclosure*  $\mathbf{Y} \supseteq \mathbf{y}$ . Techniques for computing such enclosures have been developed as part of *interval computations*; see, e.g., (7).

**Dynamic case.** Often, instead of a single quantity  $y$ , we are interested in the values  $y(t)$  corresponding to different values of the parameter  $t$ ; for example, we may be interested in the value of the quantity  $y$  at different future moments of time, or in the values of  $y$  at different locations. In this case, interval computations helps us find an enclosure  $\mathbf{Y}(t)$  for each  $t$ .

**Need to select the best technique.** There are many different interval techniques. By applying different techniques, we get different enclosures  $\mathbf{Y}(t)$ ,  $\mathbf{Y}'(t)$ , ... Each technique is better for different problems.

Theoretically, we can apply all the techniques and get an intersection of the resulting enclosures. However, in practice, this would be too time-consuming. So, for each class of problem, we need to select the most appropriate technique – the most appropriate within a given limitation on computation time.

To make this selection, it is reasonable to apply different techniques to several sample problems from the given class, and then select a technique based on the results. How can we compare these results? Of course, if one the techniques leads to narrower bounds for all  $t$  (e.g., if  $\mathbf{Y}(t) \subseteq \mathbf{Y}'(t)$  for all  $t$  and  $\mathbf{Y}(t) \neq \mathbf{Y}'(t)$  for some  $t$ ), this means that the technique that leads to  $\mathbf{Y}(t)$  is clearly better. However, often, one technique is better for some  $t$  while another technique is better for other  $t$ .

**What we do in this paper.** In this paper, we propose a symmetry-based approach to comparing two methods.

The mathematics behind this approach is similar to the mathematics used in (3) to explain the efficiency of the use of entropy (and generalized entropy) as criteria in selecting the best image among all images consistent with observations.

## ANALYSIS OF THE PROBLEM

**How preferences are described in decision theory: the notion of utility.** To describe preferences in our specific case, let us recall how preferences are described in the general case – as covered by decision theory (1; 4; 6; 9). In general, to describe preferences in numerical terms, we can select a very bad alternative  $A_0$  and a very good alternative  $A_1$ . Then, for each probability  $p \in [0, 1]$ , we can form a lottery  $L(p)$  in which we get  $A_1$  with probability  $p$  and  $A_0$  with the remaining probability  $1 - p$ . When  $p = 1$ , the lottery coincides with the alternative  $A_1$ ; when  $p = 0$ , the lottery coincides with the alternative  $A_0$ . The larger  $p$ , the better the lottery  $L(p)$ .

For each alternative  $A$  between  $A_0$  and  $A_1$ , for small  $p$ , the lottery  $L(p)$  is worse than  $A$ , while for  $p \approx 1$ , the lottery  $L(p)$  is better than the alternative  $A$ .

- If  $L(p) < A$  for some  $p$ , then  $L(p') < A$  for all smaller probabilities  $p'$ .
- Similarly, if  $A < L(p)$ , then  $A < L(p')$  for all larger probabilities  $p' > p$ .

Thus, there exists a threshold value  $u$  such that:

- for all probabilities  $p < u$ , we have  $L(p) < A$ , and
- for all probabilities  $u > p$ , we have  $L(p) > A$ .

This threshold value can be described, e.g., as  $\sup\{p : L(p) < A\}$ . The corresponding threshold value is called the *utility* of the alternative  $A$  and denoted by  $u(A)$ .

One can easily check that if  $u(A) > u(A')$ , then  $A$  is better than  $A'$ . Thus, in terms of the utility function, selecting the best alternative means selecting the alternative with the largest possible value of utility. In these terms, describing the user's preference means describing the user's utility function.

**How unique is utility.** The numerical value of the utility depends on the selection of the alternatives  $A_0$  and  $A_1$ . One can show that if we select a different pair of alternatives  $A'_0 < A'_1$ , then the new utility values  $u'(A)$  are related to the original utility values  $u(A)$  by an appropriate linear transformation:  $u'(A) = a \cdot u(A) + b$  for some real numbers  $a > 0$  and  $b$  which do not depend on  $A$ .

**Application to our problem.** Usually, we only describe the bounds  $\mathbf{Y}(t) = [\underline{Y}(t), \bar{Y}(t)]$  for the values  $t$  from a grid, i.e., for values of the type  $t_i = t_0 + i \cdot \Delta t$ , for some small time quantum  $\Delta t$ . In this case, the utility describing the quality of each estimation has the form  $u(\mathbf{Y}(t_0), \mathbf{Y}(t_1), \dots)$ .

Let us describe common sense properties of preference in terms of utility functions.

**Localness property.** Let us start with the following “localness” property. Let us assume that the two enclosures  $\mathbf{Y}(t)$  and  $\mathbf{Y}'(t)$  differ only for moments  $t$  from an interval  $[\underline{t}, \bar{t}]$ , and that  $\mathbf{Y}$  is preferred to  $\mathbf{Y}'$  ( $\mathbf{Y} \geq \mathbf{Y}'$ ). Since for all the values  $t \notin [\underline{t}, \bar{t}]$ , the bounds  $\mathbf{Y}(t)$  and  $\mathbf{Y}'(t)$  coincide, this means that, for the user, on the interval  $[\underline{t}, \bar{t}]$ , the bounds  $\mathbf{Y}(t)$  are better than the bounds  $\mathbf{Y}'(t)$ . From the common sense viewpoint, this preference should not depend on whatever common bounds  $\mathbf{Y}(t) = \mathbf{Y}'(t)$  we have for  $t \notin [\underline{t}, \bar{t}]$ .

Formally, if we have bounds  $\mathbf{Z}(t)$  and  $\mathbf{Z}'(t)$  for which:

- $\mathbf{Z}(t) = \mathbf{Z}'(t)$  for all  $t \notin [\underline{t}, \bar{t}]$ , and
- $\mathbf{Z}(t) = \mathbf{Y}(t)$  and  $\mathbf{Z}'(t) = \mathbf{Y}'(t)$  for all  $t \in [\underline{t}, \bar{t}]$ ,

then  $\mathbf{Z}$  should be preferable to  $\mathbf{Z}'$ :  $\mathbf{Z} \geq \mathbf{Z}'$ .

**How to describe localness property in terms of utility.** The “localness” property (also known as “independence” property) is a frequent feature in practical problems. It has been shown that when alternatives are characterized by  $n$  groups of parameters

$x_1, \dots, x_n$ , then the localness of the preference is equivalent to the utility function  $u(x_1, \dots, x_n)$  being of one of the two types (2):

- *additive*  $u(x_1, \dots, x_n) = u_1(x_1) + \dots + u_n(x_n)$  for some functions  $u_i(x_i)$ ; or
- *multiplicative*  $u(x_1, \dots, x_n) = u_1(x_1) \cdot \dots \cdot u_n(x_n)$  for some functions  $u_i(x_i)$ .

In our case, this means that we either maximize the sum  $\sum_i u(\mathbf{Y}(t_i), t_i)$  or the product  $\prod_i u(\mathbf{Y}(t_i), t_i)$ , where we denoted  $u_i(\mathbf{Y}(t_i))$  by  $u(\mathbf{Y}(t_i), t_i)$ .

Maximizing the product is equivalent to maximizing its logarithm  $\sum_i \ell(\mathbf{Y}(t_i), t_i)$ , where  $\ell(\mathbf{Y}(t_i), t_i) \stackrel{\text{def}}{=} \ln(u(\mathbf{Y}(t_i), t_i))$ . Thus, in both cases, we maximize the sum  $\sum_i u(\mathbf{Y}(t_i), t_i)$  for some function  $u(\mathbf{Y}, t)$ .

In the limit when  $\Delta t \rightarrow 0$ , the sum tends to an integral  $u(\mathbf{Y}) \stackrel{\text{def}}{=} \int u(\mathbf{Y}(t), t) dt$ . Thus, due to localness, we must maximize this integral.

So, to specify preferences, we must find an appropriate function  $u(\mathbf{Y}, t)$ . Let us show that other common-sense properties enable us to describe this function.

**Smoothness.** Small changes in time  $t$  and enclosure  $\mathbf{Y}(t)$  should lead to small changes in utility. It is therefore reasonable to require that the function  $u(\mathbf{Y}, t)$  is smooth. Specifically, we require that this function is at least twice differentiable.

**Using different time units and different starting points for measuring time.** Preferences should not change if we simply change the unit for measuring time (e.g., from minutes to seconds), or change the starting point.

In precise terms, changing a starting point means replacing the original value  $t$  with  $t + t_0$ , and changing the measuring unit means changing  $t$  to  $\lambda \cdot t$ .

**Using different units and different starting points for the quantity  $y$ .** Similarly, we can use different starting points and different units for describing the quantity  $y$ . It is reasonable to require that the resulting preference relation should not change if we replace  $y$  with  $y + y_0$  or with  $\lambda \cdot y$ .

In contrast to time whose direction is fixed, for many quantities  $y$ , the direction is relative: for example, the fact that some electric charges are considered to be positive and some negative does not have physical meaning, we could as well switch signs, and all formulas will remain the same. As a result, for re-scaling  $y$ , we can use negative values  $\lambda$ .

Let us formalize these conditions and see what we can deduce from them.

## MAIN RESULT

### Definition 1.

- *By an additive utility function, we mean an expression of the type  $u(\mathbf{Y}) = \int u(\mathbf{Y}(t), t) dt$ , where  $u(\mathbf{Y}, t)$  is a twice differentiable function.*

- By a multiplicative utility function, we mean an expression of the type  $u(\mathbf{Y}) = \exp(\int u(\mathbf{Y}(t), t) dt)$ , where  $u(\mathbf{Y}, t)$  is a twice differentiable function.
- By a utility function, we mean either an additive utility function or a multiplicative utility function.
- We say that the functionals  $u(\mathbf{Y})$  and  $u'(\mathbf{Y})$  are equivalent if there exist real numbers  $a > 0$  and  $b$  for which  $u'(\mathbf{Y}) = a \cdot u(\mathbf{Y}) + b$  for all functions  $\mathbf{Y}(t)$ .

**Definition 2.**

- By a  $t$ -rescaling, we mean a transformation  $T_{\lambda, t_0}(t) = \lambda \cdot t + t_0$  for some  $\lambda > 0$  and  $t_0$ .
- For each function  $\mathbf{Y}(t)$ , by the result  $\mathbf{Z} = T_{\lambda, t_0}(\mathbf{Y})$  of applying the transformation  $T_{\lambda, t_0}(t)$  to this function, we mean a function  $\mathbf{Z}(t) \stackrel{\text{def}}{=} \mathbf{Y}(\lambda \cdot t + t_0)$ .
- We say that a utility function  $u(\mathbf{Y})$  is  $t$ -invariant if for every  $\lambda > 0$  and  $t_0$ , the functional  $u(T_{\lambda, t_0}(\mathbf{Y}))$  is equivalent to  $u(\mathbf{Y})$ .

**Definition 3.**

- By a  $y$ -rescaling, we mean a transformation  $Y_{\lambda, y_0}(y) = \lambda \cdot y + y_0$  for some  $\lambda > 0$  and  $y_0$ .
- For each function  $\mathbf{Y}(t) = [\underline{Y}(t), \overline{Y}(t)]$ , by the result  $\mathbf{Z} = Y_{\lambda, y_0}(\mathbf{Y})$  of applying the transformation  $Y_{\lambda, y_0}(t)$  to this function, we mean a function

$$\mathbf{Z}(t) \stackrel{\text{def}}{=} [\lambda \cdot \underline{Y}(t) + y_0, \lambda \cdot \overline{Y}(t) + y_0].$$

- We say that a utility function  $u(\mathbf{Y})$  is  $y$ -invariant if for every  $\lambda \neq 0$  and  $y_0$ , the functional  $u(Y_{\lambda, y_0}(\mathbf{Y}))$  is equivalent to  $u(\mathbf{Y})$ .

**Proposition 1.** Every  $t$ -invariant and  $y$ -invariant utility function is equivalent either to a functional  $\int (\overline{Y}(t) - \underline{Y}(t))^p dt$  for some real number  $p > 0$ , or to the functional  $\int \ln(\overline{Y}(t) - \underline{Y}(t)) dt$ .

**Discussion.**

- For  $p = 1$ , we select a method with the smallest average width.
- For  $p = 2$ , we select a method based on the mean square width.
- For  $p \rightarrow \infty$ , we select a method with the smallest worst-case width

$$\max_t (\overline{Y}(t) - \underline{Y}(t)).$$

**Proof: main ideas.** As we have mentioned, the mathematics of this proof is similar to the mathematics from (3). Because of this similarity (and due to page limitations), we will concentrate on the main ideas behind our proof, with a main emphasis on the differences between our problem and the problem solved in (3).

For an additive utility function, invariance relative to a time shift  $t \rightarrow t + t_0$  means

that for every  $t_0$ , there exist values  $a(t_0)$  and  $b(t_0)$  for which

$$\int u(\mathbf{Y}(t), t + t_0) dt = a(t_0) \cdot \int u(\mathbf{Y}(t), t) dt + b(t_0).$$

Taking a variation derivative of both sides of this equality over  $\mathbf{Y}(t)$ , we conclude that for every  $t, t_0$ , and  $\mathbf{Y}$ , we have  $D(t + t_0) = a(t_0) \cdot D(t)$ , where  $D(t) \stackrel{\text{def}}{=} \frac{\partial u}{\partial \underline{\mathbf{Y}}}(\mathbf{Y}, t)$  (a similar functional equation holds for a partial derivative relative to  $\bar{\mathbf{Y}}$ ).

Since we assumed that the function  $u(\mathbf{Y}, t)$  is twice differentiable, its derivative  $D(t)$  is differentiable. Thus,  $a(t_0) = \frac{D(t + t_0)}{D(t)}$  is also differentiable, as a ratio of two differentiable functions. Differentiating the equality  $D(t + t_0) = a(t_0) \cdot D(t)$  relative to  $t_0$  and setting  $t_0 = 0$ , we conclude that  $\frac{dD}{dt} = a \cdot D$ , where  $a \stackrel{\text{def}}{=} \frac{da}{dt_0}|_{t_0=0}$ ; hence,  $D(t) = C \cdot \exp(a \cdot t)$ .

On the other hand, invariance relative to  $t \rightarrow \lambda \cdot t$  means that for every  $\lambda > 0$ , there exist values  $A(\lambda)$  and  $B(\lambda)$  for which

$$\int u(\mathbf{Y}(t), \lambda \cdot t) dt = A(\lambda) \cdot \int u(\mathbf{Y}(t), t) dt + B(\lambda).$$

Taking a variation derivative of both sides of this equality over  $\mathbf{Y}(t)$ , we conclude that for every  $t, \lambda$ , and  $\mathbf{Y}$ , we have  $D(\lambda \cdot t) = A(\lambda) \cdot D(t)$ . Since the derivative  $D(t)$  is differentiable, the function  $A(\lambda)$  is also differentiable, as a ratio of two differentiable functions. Differentiating the equality  $D(\lambda \cdot t) = A(\lambda) \cdot D(t)$  relative to  $\lambda$  and setting  $\lambda = 1$ , we conclude that  $t \cdot \frac{dD}{dt} = A \cdot D$ , where  $A \stackrel{\text{def}}{=} \frac{dA}{d\lambda}|_{\lambda=1}$ . For  $D(t) = C \cdot \exp(a \cdot t)$ ,

this is only possible when  $a = 0$ , i.e., when the partial derivative  $D(t) = \frac{\partial u}{\partial \underline{\mathbf{Y}}}(\mathbf{Y}, t)$  does not depend on  $t$  at all.

A similar statement holds for the partial derivative relative to  $\bar{\mathbf{Y}}(t)$ . Thus, we can conclude that the utility function is equivalent to an expression  $\int u(\mathbf{Y}(t)) dt$  in which the integrated function  $u(\mathbf{Y}(t))$  does not have any explicit dependence on time  $t$ .

Instead of a function  $u(\underline{\mathbf{Y}}, \bar{\mathbf{Y}})$ , it is convenient to consider an equivalent function  $v(s, w)$ , where:

- $m \stackrel{\text{def}}{=} \frac{\underline{\mathbf{Y}} + \bar{\mathbf{Y}}}{2}$  is the midpoint of the interval  $\mathbf{Y}$ ,
- $w \stackrel{\text{def}}{=} \bar{\mathbf{Y}} - \underline{\mathbf{y}}$  is the width of this interval, and
- $v(s, m) \stackrel{\text{def}}{=} u\left(m - \frac{w}{2}, m + \frac{w}{2}\right)$ .

The advantage of this representation is that with respect to  $y$ -shifts  $y \rightarrow y + y_0$ , only the midpoint  $m$  changes, and  $m \rightarrow m + y_0$ . Thus, invariance with respect to  $y$ -shifts

means that

$$\int v(m(t) + y_0, w(t)) dt = a(y_0) \cdot \int v(m(t), w(t)) dt + b(y_0)$$

for appropriate functions  $a(y_0)$  and  $b(y_0)$ . Differentiating both sides of this equality relative to  $m(t)$ , we conclude that, for the corresponding derivative  $D(m)$ , we have  $D(m + y_0) = a(y_0) \cdot D(m)$ . Similarly to the case of time shifts, this implies that  $D(m)$  exponentially depends on  $m$  – which, if we take into account scale-invariance, implies that there is no dependence on the midpoint  $m(t)$  at all.

Therefore, the utility function takes the form  $\int v(w(t)) dt$ , where  $w(t) = \bar{Y}(t) - \underline{Y}(t)$ . In terms of the width  $m(t)$ , the transformation  $y \rightarrow \lambda \cdot y$  leads to  $m(t) \rightarrow \lambda \cdot m(t)$ . Thus, invariance relative to this transformation implies that

$$\int v(\lambda \cdot w(t)) dt = A(\lambda) \cdot \int v(w(t)) dt + B(\lambda).$$

For the corresponding variational derivatives  $D(w)$ , we get  $\lambda \cdot D(\lambda \cdot w) = A(\lambda) \cdot D(w)$ . Hence  $D(\lambda \cdot w) = c(\lambda) \cdot D(w)$ , where we denoted  $c(\lambda) \stackrel{\text{def}}{=} \frac{A(\lambda)}{\lambda}$ . Here, the function  $c(\lambda)$  is differentiable as a ratio of two differentiable functions. Differentiating the equality  $D(\lambda \cdot w) = c(\lambda) \cdot D(w)$  with respect to  $\lambda$  and taking  $\lambda = 1$ , we get  $w \cdot \frac{dD}{dw} = c \cdot D$ , where  $c \stackrel{\text{def}}{=} \frac{dc}{d\lambda}|_{\lambda=1}$ . Moving all the terms related to  $D$  to one side and all the terms related to  $w$

to another side, we get  $\frac{dD}{D} = c \cdot \frac{dw}{w}$ . Integrating both sides, we get  $\ln(D) = c \cdot \ln(w) + c_0$  for some constant  $c_0$ . Thus, for  $D = \exp(\ln(D))$ , we get  $D(w) = \text{const} \cdot w^c$ .

Let us now recall that  $D(w) = \frac{\partial v}{\partial w}$ . Thus, to recover the expression for  $v(w)$ , we must integrate this derivative  $D(w)$  with respect to  $w$ .

- For  $c \neq -1$ , integration leads to the power dependence  $v(w) = w^p$  for  $p = c + 1$ .
- For  $c = -1$ , we get the logarithmic dependence.

Thus, for additive utility functions, the proposition is proven.

For multiplicative utility functions, the proof is similar.

## HOW TO COMPARE ESTIMATES REQUIRING DIFFERENT COMPUTATION TIME?

**Formulation of the problem.** In the previous section, we compared estimates  $\mathbf{Y}$  and  $\mathbf{Y}'$  produced by two estimation methods requiring the same computation time  $T$ . A similar approach can be used to compare estimates corresponding to different values of the computation time  $T \neq T'$ .

**Analysis of the problem.** In this case, we need to explicitly describe the dependence of the utility value on  $T$ . To deal with this dependence, we can take into account that the computation time  $T$  can also be described by using different units of time.

**Definition 4.**

- By an additive utility function, we mean an expression of the type  $u(\mathbf{Y}, T) = \int u(\mathbf{Y}(t), t, T) dt$ , where  $u(\mathbf{Y}, t, T)$  is a twice differentiable function.
- By a multiplicative utility function, we mean an expression of the type  $u(\mathbf{Y}, T) = \exp(\int u(\mathbf{Y}(t), t, T) dt)$ , where  $u(\mathbf{Y}, t, T)$  is a twice differentiable function.
- By a utility function, we mean either an additive utility function or a multiplicative utility function.
- We say that the functionals  $u(\mathbf{Y}, T)$  and  $u'(\mathbf{Y}, T)$  are equivalent if there exist real numbers  $a > 0$  and  $b$  for which  $u'(\mathbf{Y}, T) = a \cdot u(\mathbf{Y}, T) + b$  for all functions  $\mathbf{Y}(t)$  and values  $T > 0$ .

**Definition 5.**

- We say that a utility function  $u(\mathbf{Y}, T)$  is  $t$ -invariant if for every  $\lambda > 0$  and  $t_0$ , the functional  $u(T_{\lambda, t_0}(\mathbf{Y}), T)$  is equivalent to  $u(\mathbf{Y}, T)$ .
- We say that a utility function  $u(\mathbf{Y}, T)$  is  $y$ -invariant if for every  $\lambda \neq 0$  and  $y_0$ , the functional  $u(Y_{\lambda, y_0}(\mathbf{Y}), T)$  is equivalent to  $u(\mathbf{Y}, T)$ .
- By a  $T$ -rescaling, we mean a transformation  $T_{\lambda, y_0}(T) = \lambda \cdot T + T_0$  for some  $\lambda > 0$  and  $T_0$ .
- We say that a utility function  $u(\mathbf{Y}, T)$  is  $T$ -invariant if for every  $\lambda > 0$  and  $T_0$ , the functional  $u(\mathbf{Y}, \lambda \cdot T + T_0)$  is equivalent to  $u(\mathbf{Y}, T)$ .

**Proposition 2.** Every  $t$ -invariant,  $y$ -invariant, and  $T$ -invariant utility function is equivalent either to a functional  $T^{-q} \cdot \int (\bar{Y}(t) - \underline{Y}(t))^p dt$  for some real number  $p > 0$  and  $q > 0$ , or to a functional  $T^{-q} \cdot \int \ln(\bar{Y}(t) - \underline{Y}(t)) dt$  for some  $q > 0$ .

*Comment.* The proof of this proposition (and of the following Proposition 3) is similar to the proof of Proposition 1.

**AUXILIARY RESULT: HOW TO COMPARE DIFFERENT ESTIMATIONS FOR THE PARETO-OPTIMAL FRONT**

**What is Pareto-optimal front: brief reminder.** In many practical problems, we have several objective functions  $f_1(x), \dots, f_n(x)$ . In this situation, it makes sense to dismiss a solution  $x$  if there exists a better (dominating) solution  $x'$  (i.e., the one for which  $f_i(x') \leq f_i(x)$  for all  $i$  and  $f_i(x') < f_i(x)$  for some  $i$ ), and keep only non-dominated solutions.

From this viewpoint, it is desirable to find the *Pareto optimal front*, i.e., the set of all the tuples  $(y_1, \dots, y_n) = (f_1(x), \dots, f_n(x))$  corresponding to non-dominated



alternatives  $x$ .

**How to estimate Pareto-optimal front.** A natural way to estimate the Pareto-optimal front is to compute the values of all the objective functions  $f_i(x)$  for several different alternatives  $x$ . Based on these computations, we then dismiss the dominant solutions, and keep the tuples  $(f_1(x), \dots, f_n(x))$  corresponding to non-dominant alternatives. The resulting surface  $y_n = f(y_1, \dots, y_{n-1})$  is then returned as an approximation to the actual Pareto-optimal front. (Due to the finiteness of the sample, this surface provides only an approximate description of the actual Pareto optimal front.)

There exist several different techniques for such an estimation. Which one should we select?

**What we do in this section.** We show that for  $n = 2$ , a symmetry-based approach (similar to the approach from the previous section) can help us in selecting the best estimation method – as a method which leads to a function  $y_2 = f(y_1)$  which is the best estimation for the Pareto optimal front.

**Definition 6.**

- By an additive utility function, we mean an expression of the type  $u(f) = \int u(f(y_1), y_1) dy_1$ , where  $u(y_2, y_1)$  is a twice differentiable function.
- By a multiplicative utility function, we mean an expression of the type  $u(f) = \exp(\int u(f(y_1), y_1) dy_1)$ , where  $u(y_2, y_1)$  is a twice differentiable function.
- By a utility function, we mean either an additive utility function or a multiplicative utility function.
- We say that the functionals  $u(f)$  and  $u'(f)$  are equivalent if there exist real numbers  $a > 0$  and  $b$  for which  $u'(f) = a \cdot u(f) + b$  for all functions  $f(y_1)$ .

**Definition 7.**

- By a  $y_1$ -rescaling, we mean a transformation  $F_{\lambda, y_0}(y_1) = \lambda \cdot y_1 + y_0$  for some  $\lambda > 0$  and  $y_0$ .
- For each function  $f(y_1)$ , by the result  $g = F_{\lambda, y_0}(f)$  of applying the transformation  $F_{\lambda, y_0}(y_1)$  to this function, we mean a function  $g(y_1) \stackrel{\text{def}}{=} f(\lambda \cdot y_1 + y_0)$ .
- We say that a utility function  $u(f)$  is  $y_1$ -invariant if for every  $\lambda > 0$  and  $y_0$ , the functional  $u(F_{\lambda, y_0}(f))$  is equivalent to  $u(f)$ .

**Definition 8.**

- By a  $y_2$ -rescaling, we mean a transformation  $S_{\lambda, y_0}(y_2) = \lambda \cdot y_2 + y_0$  for some  $\lambda > 0$  and  $y_0$ .
- For each function  $f(y_1)$ , by the result  $g = D_{\lambda, y_0}(f)$  of applying the transformation  $S_{\lambda, y_0}(y_2)$  to this function, we mean a function  $h(y_1) \stackrel{\text{def}}{=} \lambda \cdot f(y_1) + y_0$ .
- We say that a utility function  $u(f)$  is  $y_2$ -invariant if for every  $\lambda > 0$  and  $y_0$ , the functional  $u(S_{\lambda, y_0}(f))$  is equivalent to  $u(f)$ .

**Proposition 3.** Every  $y_1$ -invariant and  $y_2$ -invariant utility function is equivalent to the functional  $\int f(y_1) dy_1$ .

**Discussion.** So, we should select a method for which the area under the curve  $y_2 = g(y_1)$  is the smallest.

**Acknowledgments.** This work was supported in part by the National Science Foundation grants HRD-0734825, HRD-1242122, and DUE-0926721.

We are thankful to the anonymous referees for valuable suggestions.

## REFERENCES

- Fishburn, P. C. (1969) *Utility Theory for Decision Making*, John Wiley & Sons Inc., New York.
- Fishburn, P. C. (1988) *Nonlinear Preference and Utility Theory*, John Hopkins Press, Baltimore, Maryland.
- Kosheleva, O. (1998) “Symmetry-group justification of maximum entropy method and generalized maximum entropy methods in image processing”, In: Erickson, G. J., Rychert, J. T., and Smith, S. R. (eds.), *Maximum Entropy and Bayesian Methods*, Kluwer, Dordrecht, pp. 101–113. 1998).
- Kreinovich, V. (2014) “Decision Making under Interval Uncertainty (and beyond)”, In: Guo, P., and Pedrycz, W. (eds.), *Human-Centric Decision-Making Models for Social Sciences*, Springer Verlag, pp. 163–193.
- Kreinovich, V., Lakeyev, A., Rohn, J., and Kahl, P. (1997) *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Kluwer, Dordrecht.
- Luce, R. D., and Raiffa, R. (1989) *Games and Decisions: Introduction and Critical Survey*, Dover, New York.
- Moore, R. E., Kearfott, R. B., and Cloud, M. J. (2009) *Introduction to Interval Analysis*, SIAM Press, Philadelphia, Pennsylvania.
- Rabinovich, S. (2005) *Measurement Errors and Uncertainties: Theory and Practice*, American Institute of Physics, New York.
- Raiffa, H. (1997) *Decision Analysis*, McGraw-Hill, Columbus, Ohio.