How to Deal with Inconsistent Intervals: Utility-Based Approach Can Overcome the Limitations of the Purely Probability-Based Approach

Kittawit Autchariyapanitkul[1], Tomoe Entani[2]
Olga Kosheleva[3], and Vladik Kreinovich[3]

[1]Maejo University, Chiang Mai, Thailand, kittawit_a@mju.ac.th

[2]Graduate School of Information Science, University of Hyogo
Kobe, Japan, entani@gsis.u-hyogo.ac.jp

[2]University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA, olgak@utep.edu, vladik@utep.edu

1. Expert knowledge is important

- In many application areas, we rely on expert estimates.
- Economics and finance are a good example of such reliance.
- We often rely on experts to predict the future value of inflation, the economic growth rate, etc.

2. Expert estimates often come in interval form

- Sometimes expert provide exact numerical values of the corresponding quantities.
- However, expert estimates are usually very approximate.
- The estimation errors cannot be ignored.
- It is therefore reasonable to make sure that the experts provide us:
  - not only with the numerical estimates $\widetilde{x}$ for the quantity of interest x,
  - but also with the estimated upper bound Δ on the estimation error.
- This means, in effect, that the expert estimates that the actual value x should be somewhere in the interval

$$[\underline{x}, \overline{x}] \stackrel{\text{def}}{=} [\widetilde{x} - \Delta, \widetilde{x} + \Delta].$$

3. Sometimes, intervals provided by several experts are consistent

- Often, we have several experts estimating the same quantity.
- In many practical situations, intervals $[\underline{x}_i, \overline{x}_i]$ provided by different experts are consistent – i.e., all these intervals have common value.
- In such situations, if we trust all these experts:
  - it is reasonable to conclude that the actual value belongs to all these intervals,
  - i.e., equivalently, belongs to their intersection.
- For example, suppose one reliable witness whose height is 170 cm:
  - saw that the suspect is 5 to 15 cm taller than him,
  - i.e., that the suspects's height is in the interval $[175, 185]$.

4. Sometimes, intervals provided by several experts are consistent (cont-d)

- Suppose also that another reliable witness whose height is 185 cm:
  - saw that the suspect was 5 to 15 cm shorter then him,
  - i.e., that the suspect's height is somewhere in the interval $[170, 180]$.
- It is reasonable to conclude that the actual height of the suspect lies in the intersection $[175, 185] \cap [170, 180] = [175, 180]$ of these intervals.
- Similarly, suppose that:
  - one reliable expert predicts that the economy's growth rate will be between 3 and 5 percent, and
  - another one predicts that it will be between 4 and 6 percent.

5. Sometimes, intervals provided by several experts are consistent (cont-d)

- Then it is reasonable to conclude that the actual growth rate will belong to the intersection of these two intervals:
  - i.e., to the interval $[3,5]\cap[4,6]=[4,5]$,
  - meaning that the growth rate will be between 4 and 5 percent.

6. But what if the intervals are inconsistent?

- In many other cases, the intervals provided by different experts are inconsistent.
- For example:
  - one expert may be sure that the growth rate will be between 3 and 4 percents, while
  - another expert is sure that it will be between 5 and 6 percents.
- The corresponding intervals $[3, 4]$ and $[5, 6]$ do not have a common point, so the experts cannot be both right.
- In other words, at least one of the experts underestimates his/her approximation error.
- If we enlarge the intervals to take the actual approximation error into account, the intervals will intersect.
- There are many ways to extend the two intervals so that the extended intervals start intersecting; which one should we choose?

7. What we do in this talk

- Our problem is related to uncertainty.
- So, to solve this problem, it seems reasonable to apply techniques:
  - that have been designed to deal with uncertainty and that have been successfully used for several centuries,
  - namely, probability-based techniques.
- So, first, we explain how probability-based approach can be applied to our problem.
- It turns out, however, that when we apply these techniques to the above problem, we get counter-intuitive results.
- The counter-intuitive character of these results is easy to understand:
  - probability-based techniques take care of how frequent different outcomes are, but
  - they do not take into account our preference between different outcomes.

8. What we do in this talk (cont-d)

- To consistently take these preferences into account when making decisions, researchers developed a special *decision theory*.
- In this theory, preferences are described by numerical values – known as *utilities* assigned to different outcomes.
- We show if we take utilities into account, then we indeed get a meaningful way to deal with inconsistent intervals.

9. Where do we get probabilities?

- To apply the probability-based approach, we need to have some information about the probabilities.
- However, in our case, we do not have any such information.
- The only information that each expert provides is an interval of possible values.
- We do not know which values from this interval are more probable and which are less probable.
- Situations in which we have several alternatives, and we do not know which ones are more probable and which are less probable, are ubiquitous.
- In such situations, a natural idea is:
  - to assign equal probabilities to all these alternatives,
  - i.e., in mathematical terms, to assume that the probability distribution is uniform.

10. Where do we get probabilities (cont-d)

- This idea makes perfect sense.
- For example:
  - if we have three suspects and we have no reasons to consider one of them more probable,
  - then it is reasonable to assign probability $1/3$ to each.
- This idea goes back to the early applications of probabilities and is thus known as *Laplace Indeterminacy Principle*.
- In particular:
  - in situations when all we know is that some quantity is located on an interval,
  - a natural idea is to assign equal probability to all the values from this interval,
  - i.e., to assume that the corresponding quantity is uniformly distributed on this interval.

11. Comment

- In our case, we know only that the probability distribution is located on the given interval.
- Out of all such distributions, we need to select a one that best reflects the given information.
- Our case is a particular example of a more general situation, when:
  - we know a family of possible probability distributions, and
  - out of all these distributions, we need to select a one that best reflects the related uncertainty.
- In the interval case, we could select a distribution that is located at one of the points with probability 1.
- However, by making this selection, we would have eliminated all uncertainty.
- A more reasonable approach is to select the distribution that preserves the original uncertainty as much as possible.

12. Comment (cont-d)

- A natural measure of uncertainty is:
  - the average number of binary ("yes"-"no") questions
  - that we need to ask to determine the actual value with a given accuracy.
- It is known that this average number is proportional to *entropy* $S \stackrel{\text{def}}{=} -\int f(x) \cdot \ln(f(x))\, dx$ of the probability distribution.
- Here, $f(x)$ is the distribution's density.
- Thus, a natural idea is to select the distribution with largest possible entropy.
- This idea is known as the *Maximum Entropy* approach.
- When all we know is the interval of possible values, the Maximum Entropy approach leads exactly to the uniform distribution.

13. Let us formulate our problem in precise terms

- We have several intervals $[\underline{x}_i, \overline{x}_i]$ which are inconsistent: $\cap[\underline{x}_i, \overline{x}_i] = \emptyset$.
- This means that we need to enlarge some of these intervals, i.e.:
  - come up with larger intervals $[\underline{X}_i, \overline{X}_i] \supseteq [\underline{x}_i, \overline{x}_i]$
  - so that the enlarged intervals will have a non-empty intersection $\cap[\underline{X}_i, \overline{X}_i] \neq \emptyset$.
- There are many way to perform such an enlargement, which one should we choose?
- A reasonable idea is to select *the most probable* one.
- This idea is known as the *Maximum Likelihood* approach.
- We assumed that for each interval, the distribution is uniform on this interval.

14. Let us formulate our problem in precise terms (cont-d)

- Thus, for each expert, the probability density corresponding to the interval $[\underline{X}_i, \overline{X}_i]$ is equal to $\dfrac{1}{\overline{X}_i - \underline{X}_i}$.
- Different experts are usually independent:
  - if the opinion of an expert is strongly determined by the opinions of others,
  - there is no need to ask this expert anything: his/her opinion is determined from the opinions of others.
- The experts are independent.
- So, the overall probability of selecting the enlargements is equal to the product of the probabilities $1/(\overline{X}_i - \underline{X}_i)$ corr. to experts.
- We want to select the extension whose probability is the largest.
- We thus arrive at the following precise formulation.

15. Probability-based formulation of the problem

- *Given:* n intervals $[\underline{x}_i, \overline{x}_i]$ whose intersection is empty $\cap[\underline{x}_i, \overline{x}_i] = \emptyset$.
- *Find:*
  - among all *extensions* $[\underline{X}_i, \overline{X}_i] \supseteq [\underline{x}_i, \overline{x}_i]$ for which the intersection is non-empty $\cap[\underline{X}_i, \overline{X}_i] \neq \emptyset$,
  - we need to select an extension with the largest value of the expression $\prod\limits_{i=1}^{n} \dfrac{1}{\overline{X}_i - \underline{X}_i}$.

16. Limitations of the Probability-Based Approach

- We want to show that the above formulation – while seemingly reasonable – leads to counter-intuitive results.
- Let us consider the simplest possible case when we have $n = 2$ non-intersecting intervals.
- For simplicity, let us assume that both intervals have the same width $w = \overline{x}_i - \underline{x}_i$.

17. What is reasonable to expect in this case

- We have two similar experts.
- So it is reasonable to expect that:
  - they will be treated similarly, i.e., that
  - both expert's intervals will be extended in the same way, i.e., by the same amount.
- We will show, however, that this is *not* what the probability-based approach recommends.
- To show this, let us introduce some notations.

18. What the probability-based approach recommends for this case

- **Proposition 1.**
  - *Let us assume that we have two non-intersecting intervals of equal width w.*
  - *In this case, the probability-based approach means keeping one of the two intervals intact, thus extending only one of the intervals.*
- All the proofs are placed at the end of this deck, we will cover them if we have time.
- This result means that:
  - either the first expert's interval is not extended at all,
  - or the second expert's interval is not extended at all.
- In both cases, one of the experts is assumed to be absolutely right – which is not what our intuition tells us is reasonable.

19. Utility-based approach

- Probability-based approach takes into account what is more probable and what is less probable.
- However, it does not take into account what is best for us.
- To decide which extension is best for us, we need to take into account our preferences.
- Decision theory shows how to describe preferences:
  - of a rational person,
  - i.e., e.g., a person who, if preferring A to B and B to C, also prefers A to C.

20. Utility-based approach (cont-d)

- These preferences can be described by assigning, to each possible alternative A a numerical value $u(A)$ in such a way that:
  - an alternative A is preferred to an alternative B if and only if $u(A) > u(B)$, and
  - for a situation S in which several outcomes $A_1, \ldots, A_n$ are possible, with corresponding probabilities $p_1, \ldots, p_n$, the utility $u(S)$ is:

$$u(S) = p_1 \cdot u(A_1) + \ldots + p_n \cdot u(A_n).$$

- Utility can be defined as follows
- We select a very bad alternative A_- that is worse that anything that we may actually encounter.
- We also select a very good alternative A_+ that is better than anything that we may actually encounter.

21. Utility-based approach (cont-d)

- Then, for each alternative A, we ask the user to compare this alternative with a lottery $L(p)$ in which the user:
  - will get A_+ with probability p, and
  - will get A_- with remaining probability $1 - p$.
- For small probabilities p, we have $L(p) \approx A_-$.
- Thus, due to our choice of A_-, the lottery $L(p)$ is worse than A; we will denote it by $L(p) < A$.
- For probabilities p close to 1, we have $L(p) \approx A_+$.
- Thus, due to our choice of A_+, the lottery A is worse than $L(p)$; $A < L(p)$.
- Clearly, if $p < p'$, then $L(p) < L(p')$. Thus:
  - if $p < p'$ and $L(p') < A$, then $L(p) < A$, and
  - if $p < p'$ and $A < L(p)$, then $A < L(p')$.

22. Utility-based approach (cont-d)

- Thus, as one can show, there exists a threshold value u such that:
  - for $p < u$, we have $L(p) < A$, and
  - for $p > u$, we have $A < L(p)$.
- This threshold value u is exactly the utility $u(A)$ of the alternative A.
- The numerical value $u(A)$ of the utility depends on the selection of the two alternatives A_- and A_+.
- It turns out that if we select a different pair of alternatives, then:
  - the new utility function $U(A)$ is related to the original one $u(A)$
  - by a linear expression: $U(A) = a_0 + a_1 \cdot u(A)$ for some constants a_0 and $a_1 > 0$.

23. Let us apply this idea to our case

- In general, the narrower the interval, the more information it carries about the quantity.
- So, for a single interval, utility u depends on the width w of the interval: the larger the width w, the smaller the utility $u(w)$.
- As we have mentioned, experts are independent.
- For independent events, utility is equal to the sum of the utilities.
- Thus, it makes sense to describe the utility of an extension as the sum of the utilities $u(W_i)$ corresponding to intervals $[\underline{X}_i, \overline{X}_i]$ with widths
$$W_i = \overline{X}_i - \underline{X}_i.$$
- By definition of utility, the best alternative is the one for which utility is the largest.
- Thus, we arrive at the following precise formulation of the problem:

24. Preliminary utility-based formulation of the problem

- *Given:* n intervals $[\underline{x}_i, \overline{x}_i]$ whose intersection is empty $\cap[\underline{x}_i, \overline{x}_i] = \emptyset$.
- *Find:*
  - among all *extensions* $[\underline{X}_i, \overline{X}_i] \supseteq [\underline{x}_i, \overline{x}_i]$ for which the intersection is non-empty $\cap[\underline{X}_i, \overline{X}_i] \neq \emptyset$,
  - we need to select an extension with the largest value of the expression $\sum\limits_{i=1}^{n} u(\overline{X}_i - \underline{X}_i)$.

25. What are reasonable utility functions: mathematical description of the problem

- To use the above criterion, we need to find out what are the reasonable utility functions $u(w)$.
- To find this out, let us take into account that we are talking expert estimates of different quantities such as height.
- For most of these quantities, the numerical value depends of the choice of a measuring unit.
- For example, if we replace centimeters with meters, then all numerical values are divided by 100: 175 cm becomes 1.75 m.
- Usually, there is no preferable measuring unit; which measuring unit to choose is a matter of convenience.
- In this case, it is reasonable to require that our preferences do not depend on this choice.

26. What are reasonable utility functions (cont-d)

- In mathematical terms:
  - if we replace the original measuring unit with a new unit which is λ times smaller,
  - then all numerical values – including numerical values of each width w – are multiplied by λ: $w \mapsto \lambda \cdot w$.
- So, if in the original units, we had width w, the same width in the new units is described as $\lambda \cdot w$.
- So, the same preferences that in the original units were described by the utility function $u(w)$ are now described by a new function $u(\lambda \cdot w)$.
- These two utility functions should describe the preference relation.
- As we have mentioned, this means that they should be linearly related.

27. What are reasonable utility functions (cont-d)

- In other words, for every $\lambda > 0$, there should be some values $a_0(\lambda)$ and $a_1(\lambda)$ for which, for every $w > 0$, we have

$$u(\lambda \cdot w) = a_0(\lambda) + a_1(\lambda) \cdot u(w).$$

- We will call this *scaling equation.*

28. What are reasonable utility functions: from mathematical description to explicit formulas

- It is reasonable to assume that:
  - small changes of width lead to equally small changes in the utility,
  - i.e., that the dependence $u(w)$ is smooth (differentiable).
- We say that two utility functions $u(w)$ and $U(w)$ are *equivalent* if there exist constant a_0 and $a_1 > 0$ for which, for all w, we have

$$U(w) = a_0 + a_1 \cdot u(w).$$

- **Proposition 2.** *Every decreasing differentiable function $u(w)$ that satisfies the scaling equation for all $w > 0$ and $\lambda > 0$ is equivalent*
  - *either to* $-\ln(w)$
  - *or to* $-\text{sign}(b) \cdot w^b$ *for some* $b \neq 0$.
- **Proposition 3.** *For $u(w) = -\ln(w)$, the utility-based formulation is equivalent to the probability-based formulation.*

29. What are reasonable utility functions: from mathematical description to explicit formulas (cont-d)

- We have shown that the probability-based formulation leads to counter-intuitive results.
- Hence, we should only consider the power-law utility functions.
- Thus, the above formulation takes the following form.

30. Final utility-based formulation of the problem

- *Given:* n intervals $[\underline{x}_i, \overline{x}_i]$ whose intersection is empty $\cap[\underline{x}_i, \overline{x}_i] = \emptyset$.
- *Find:*
  - among all *extensions* $[\underline{X}_i, \overline{X}_i] \supseteq [\underline{x}_i, \overline{x}_i]$ for which the intersection is non-empty $\cap[\underline{X}_i, \overline{X}_i] \neq \emptyset$,
  - we need to select an extension with the largest value of the expression $-\text{sign}(b) \cdot \sum\limits_{i=1}^{n} (\overline{X}_i - \underline{X}_i)^b$.

31. The resulting utility-based description is in perfect accordance with our intuition

- **Proposition 4.** *For the case when* $b > 1$*:*
  - *when we have two non-intersecting intervals of equal width* w*,*
  - *the utility-based approach means that we extend both intervals by the same additional width.*
- In this talk, we focused on *direct* inconsistency, when several expert estimates of the same quantity are inconsistent.
- The same approach can be applied to the case of *indirect* inconsistency, when:
  - we know the relation between several quantities, and
  - expert estimates of these quantities are inconsistent with this relation.

32. The resulting utility-based description is in perfect accordance with our intuition (cont-d)

- For example, values of current I, voltage V, and resistance R must satisfy Ohm's law $V = I \cdot R$.
- In this case, interval estimates $[0.9, 1.1]$ for I and R and $[0.7, 0.8]$ for V are inconsistent.
- Indeed, in this case, possible values of the product $I \cdot R$ range from $0.9 \cdot 0.9 = 0.81$ to $1.1 \cdot 1.1 = 1.21$.
- Thus, they cannot be smaller than or equal to 0.8.
- In such situations, we can use the same approach to select the best extension of intervals.
- The only difference from above formulation is that the consistency condition will be more complicated.

33. Proof of Proposition 1

- Without losing generality, we can assume that the first interval is to the left of the second one, i.e., that $\overline{x}_1 < \underline{x}_2$.
- Let us denote the distance between the given intervals by $d \stackrel{\text{def}}{=} \underline{x}_2 - \overline{x}_1$.
- Maximizing the probability is equivalent to minimizing the product of the widths $\overline{X}_i - \underline{X}_i$ of the enlarged intervals.
- Since the interval $[\underline{X}_1, \overline{X}_1]$ extends the first given interval, we must have $\underline{X}_1 \le \underline{x}_1$.
- If we have $\underline{X}_1 < \underline{x}_1$:
  - then we can replace $\underline{X}_1$ with $\underline{x}_1$ and
  - thus, get intersecting extensions with smaller width of the first interval – and
  - hence, smaller product of widths.
- Thus, the smallest product of widths is attained when $\underline{X}_1 = \underline{x}_1$.

34. Proof of Proposition 1 (cont-d)

- Similarly, we can conclude that the smallest product of widths is attained when $\overline{X}_2 = \overline{x}_1$.
- In this case, the fact that the intervals intersects means that $\overline{X}_1 \geq \underline{X}_2$.
- If we have $\overline{X}_1 > \underline{X}_2$, then:
  - we can shorten the first extended interval to the upper endpoint $\overline{X}_1 = \underline{X}_2$,
  - while keeping the two extended intervals intersecting.
- After this shortening:
  - the width of the first extended interval decreases, while
  - the width of the second one remains the same.
- Thus, the product of the widths decreases.
- Therefore, in situations when the product of widths is the smallest possible, we cannot have $\overline{X}_1 > \underline{X}_2$.

35. Proof of Proposition 1 (cont-d)

- So, we must have $\overline{X}_1 = \underline{X}_2$.
- So, in the most probable case, the two intervals divide the distance d between themselves:
  - the first interval has width $w+d_1$, where we denote $d_1 \stackrel{\text{def}}{=} \overline{X}_1 - \overline{x}_1$, and
  - the second interval has width $w + d - d_1$.
- The smallest product of the widths corresponding to the smallest value of the product $(w + d_1) \cdot (w + d - d_1)$.
- One can easily check that:
  - this quadratic function attains its maximum for $d_1 = d/2$, and
  - this expression is increasing for $d_1 < d/2$ and decreasing for $d_1 > d/2$.
- Thus, for values $d_1 \in [0, 1]$, the product attains the smallest value if either when $d_1 = 0$ or when $d_1 = d$.

36. Proof of Proposition 1 (cont-d)

- What does this mean?
- The case $d_1 = 0$ means that the first expert's interval is not extended at all.
- The case $d_1 = d$ means that the second expert's interval is not extended at all.
- The proposition is thus proven.

37. Proof of Proposition 2

- Let us first show that the functions $a_0(\lambda)$ and $a_1(\lambda)$ are also differentiable.
- Indeed, let us consider the scaling equation for two values w_1 and w_2:

$$u(\lambda \cdot w_1) = a_0(\lambda) + a_1(\lambda) \cdot u(w_1),$$

$$u(\lambda \cdot w_2) = a_0(\lambda) + a_1(\lambda) \cdot u(w_2).$$

- Subtracting the second equation from the first one, we get

$$u(\lambda \cdot w_2) - u(\lambda \cdot w_1) = a_1(\lambda) \cdot (u(w_2) - u(w_1)).$$

- Hence $a_1(\lambda) = \dfrac{u(\lambda \cdot w_2) - u(\lambda \cdot w_1)}{u(w_2) - u(w_1)}$.
- Since the function $u(w)$ is differentiable, the right-hand side of this formula is differentiable as well.
- Thus, its left-hand side – i.e., the function $a_1(\lambda)$ – is also differentiable.

38. Proof of Proposition 2 (cont-d)

- From the scaling equation, we can now conclude that the function $a_0(\lambda) = u(\lambda \cdot w_1) - a_0(\lambda) + a_1(\lambda) \cdot u(w_1)$:
  - is the difference of two differentiable functions and
  - is, thus, differentiable as well.
- Since all three functions used in scaling equation are differentiable, we can differentiate both sides with respect to λ.
- This will lead us to the following expression:

$$w \cdot u'(\lambda \cdot w) = a_0'(\lambda) + a_1'(\lambda) \cdot u(w).$$

- In particular, for $\lambda = 1$, we get: $w \cdot u'(w) = c_0 + c_1 \cdot u(w)$, where we denoted $c_i \stackrel{\text{def}}{=} a_i'(1)$.
- If both c_0 and c_1 were equal to 0, we would get $u'(w) = 0$ and thus, $u(w)$ would be a constant.
- However, we assumed that the function $u(w)$ is a decreasing function of the width w.

39. Proof of Proposition 2 (cont-d)

- Thus, the expression $c_0 + c_1 \cdot u(w)$ is not identically 0.
- The above formula can be reformulated as $w \cdot \dfrac{du}{dw} = c_0 + c_1 \cdot u$.
- We can separate the variables in this equality if we:
  - multiply both sides by dw,
  - divide both sides by w, and
  - divide both sides by $c_0 + c_1 \cdot u$.
- Then, we get the following equality: $\dfrac{du}{c_0 + c_1 \cdot u} = \dfrac{dw}{w}$.
- If $c_1 = 0$, then integrating both sides of this formula leads to $\dfrac{u}{c_0} = \ln(w) + C$, where C denotes the integration constant.
- Thus in this case $u(w) = c_0 \cdot \ln(w) + c_0 \cdot C$.
- Since the function $u(w)$ should be decreasing, we must have $c_0 < 0$.

40. Proof of Proposition 2 (cont-d)

- So this function is equivalent to $-\ln(w)$.
- When $c_1 \neq 0$, we have $d(c_0 + c_1 \cdot u) = c_1 \cdot du$, so
$$du = \frac{d(c_0 + c_1 \cdot u)}{c_1}.$$
- Substituting this expression into the above formula, we get
$$\frac{1}{c_1} \cdot \frac{d(c_0 + c_1 \cdot u)}{c_0 + c_1 \cdot u} = \frac{dw}{w}.$$
- Integrating both sides, we get $\dfrac{1}{c_1} \cdot \ln(c_0 + c_1 \cdot u) = \ln(w) + C$.
- Hence $\ln(c_0 + c_1 \cdot u) = c_1 \cdot \ln(w) + c_1 \cdot C$.
- Applying $\exp(x)$ to both sides, we get
$$c_0 + c_1 \cdot u(w) = \exp(c_1 \cdot C) \cdot \cdot w^{c_1}.$$
- So $u(w) = \dfrac{\exp(c_1 \cdot C)}{c_1} \cdot w^{c_1} - \dfrac{c_0}{c_1}$.

41. Proof of Proposition 2 (cont-d)

- Thus, the utility function $u(w)$ is equivalent to the power law $u(w) = \pm w^b$ for some value b.
- This function should be decreasing, so:
  - for $b > 0$, we take the minus sign, and
  - for $b < 0$, we take the plus sign.
- In general, $u(w) = -\text{sign}(b) \cdot w^b$.
- The proposition is proven.

42. Proof of Proposition 3

- For the function $u(w) = -\ln(w)$, maximizing $\sum u(wW_i)$ is equivalent to maximizing the expression $E \stackrel{\text{def}}{=} -\sum_{i=1}^{n} \ln(W_i)$.
- Since the function $\exp(x)$ is increasing, this is equivalent to maximizing $\exp(E)$.
- One can show that $\exp(E)$ is exactly what we maximized in the maximum likelihood approach.
- The proposition is proven.

43. Proof of Proposition 4

- Similarly to the proof of Proposition 1, we can show that in this case, the utility solution leads to $W_1 = w + d_1$ and $W_2 = w + d - d_1$.
- In this case, the utility formulation requires us to maximize the expression $-(w + d_1)^b - (w + d - d_1)^b$.
- According to calculus, to find the largest value of this expression for $d_1 \in [0, d]$, we need to compare the its values:
  - at the endpoints of this interval, i.e., for $d_1 = 0$ and $d_1 = d$, and
  - at a point d_1 at which the derivative of this expression is equal to 0.
- Differentiating this expression with respect to d_1 and equating the derivative to 0, we get $d_1 = d/2$.
- For $b > 1$, the value corresponding to $d_1 = d/2$ is larger than the values corresponding to $d_1 = 0$ and $d_1 = d$.
- This is due to convexity of the function x^b.

44. Proof of Proposition 4 (cont-d)

- Thus, the maximum utility is indeed attained when both intervals are extended equally.
- The proposition is proven.

45. 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).