Likert-Scale Fuzzy Uncertainty from a Traditional Decision Making Viewpoint: Incorporating both Subjective Probabilities and Utility Information

Joe Lorkowski and Vladik Kreinovich

Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, Texas 79968, USA
lorkowski@computer.org, vladik@utep.edu

IFSA-NAFIPS’2013
Fuzzy Uncertainty: A Usual Description

- Fuzzy logic formalizes imprecise properties $P$ like “big” or “small” used in experts’ statements.
- It uses the degree $\mu_P(x)$ to which $x$ satisfies $P$:
  - $\mu_P(x) = 1$ means that we are confident that $x$ satisfies $P$;
  - $\mu_P(x) = 0$ means that we are confident that $x$ does not satisfy $P$;
  - $0 < \mu_P(x) < 1$ means that there is some confidence that $x$ satisfies $P$, and some confidence that it doesn’t.
- $\mu_P(x)$ is typically obtained by using a Likert scale:
  - the expert selects an integer $m$ on a scale from 0 to $n$;
  - then we take $\mu_P(x) := m/n$;
- This way, we get values $\mu_P(x) = 0, 1/n, 2/n, \ldots, n/n = 1$.
- To get a more detailed description, we can use a larger $n$. 
Need to Combine Fuzzy and Traditional Techniques

- Fuzzy tools are effectively used to handle imprecise (fuzzy) expert knowledge in control and decision making.
- On the other hand, traditional utility-based techniques have been useful in crisp decision making (e.g., in economics).
- It is therefore reasonable to combine fuzzy and utility-based techniques.
- One way to combine these techniques is to translate fuzzy techniques into utility terms.
- For that, we need to describe Leikert scale selection in utility terms.
- To the best of our knowledge, this was never done before.
- This is what we do in this talk.
Traditional Decision Theory: Reminder

- **Main assumption** – for any two alternatives $A$ and $A'$:
  - either $A$ is better (we will denote it $A' < A$),
  - or $A'$ is better (we will denote it $A < A'$),
  - or $A$ and $A'$ are of equal value (denoted $A \sim A'$).

- **Resulting scale** for describing the quality of different alternatives $A$:
  - to define a scale, we select a very bad alternative $A_0$ and a very good alternative $A_1$;
  - for each $p \in [0, 1]$, we can form a lottery $L(p)$ in which we get $A_1$ with probability $p$ and $A_0$ with probability $1 - p$;
  - for each reasonable alternative $A$, we have $A_0 = L(0) < A < L(1) = A_1$;
  - thus, for some $p$, we switch from $L(p) < A$ to $L(p) > A$, i.e., there exists a “switch” value $u(A)$ for which $L(u(A)) \equiv A$;
  - this value $u(A)$ is called the *utility* of the alternative $A$. 
Utility Scale

- We have a lottery $L(p)$ for every probability $p \in [0, 1]$:  
  - $p = 0$ corresponds to $A_0$, i.e., $L(0) = A_0$;  
  - $p = 1$ corresponds to $A_1$, i.e., $L(1) = A_1$;  
  - $0 < p < 1$ corresponds to $A_0 < L(p) < A_1$;  
  - $p < p'$ implies $L(p) < L(p')$.
- There is a continuous monotonic scale of alternatives: 
  \[
  L(0) = A_0 < \ldots < L(p) < \ldots < L(p') < \ldots < L(1) = A_1.
  \]
- This *utility scale* is used to gauge the attractiveness of each alternative.
How to Elicit the Utility Value: Bisection

▶ We know that $A \equiv L(u(A))$ for some $u(A) \in [0, 1]$.
▶ Suppose that we want to find $u(A)$ with accuracy $2^{-k}$.
▶ We start with $[u, \bar{u}] = [0, 1]$. Then, For $i = 1$ to $k$, we:
  ▶ compute the midpoint $u_{\text{mid}}$ of $[u, \bar{u}]$
  ▶ ask the expert to compare $A$ with the lottery $L(u_{\text{mid}})$
  ▶ if $A \leq L(u_{\text{mid}})$, then $u(A) \leq u_{\text{mid}}$, so we can take
    
    $$[u, \bar{u}] = [u, u_{\text{mid}}];$$

  ▶ if $A \geq L(u_{\text{mid}})$, then $u(A) \geq u_{\text{mid}}$, so we can take
    
    $$[u, \bar{u}] = [u_{\text{mid}}, \bar{u}].$$

▶ At each iteration, the width of $[u, \bar{u}]$ decreases by half.
▶ After $k$ iterations, we get an interval $[u, \bar{u}]$ of width $2^{-k}$ that contains $u(A)$.
▶ So, we get $u(A)$ with accuracy $2^{-k}$. 
Utility Theory and Human Decision Making

- Decision based on utility values
  - Which of the utilities $u(A'), u(A'')$, \ldots, of the alternatives $A', A'', \ldots$ should we choose?
  - By definition of utility, $A'$ is preferable to $A''$ if and only if $u(A') > u(A'')$.
  - We should always select an alternative with the largest possible value of utility.
  - So, to find the best solution, we must solve the corresponding optimization problem.

- Our claim is that when people make definite and consistent choices, these choices can be described by probabilities.
  - We are not claiming that people always make rational decisions.
  - We are not claiming that people estimate probabilities when they make rational decisions.
Estimating the Utility of an Action $a$

- We know possible outcome situations $S_1, \ldots, S_n$.
- We often know the probabilities $p_i = p(S_i)$.
- Each situation $S_i$ is equivalent to the lottery $L(u(S_i))$ in which we get:
  - $A_1$ with probability $u(S_i)$ and
  - $A_0$ with probability $1 - u(S_i)$.
- So, $a$ is equivalent to a complex lottery in which:
  - we select one of the situations $S_i$ with prob. $p_i = P(S_i)$;
  - depending on $S_i$, we get $A_1$ with prob. $P(A_1|S_i) = u(S_i)$.
- The probability of getting $A_1$ is
  $$P(A_1) = \sum_{i=1}^{n} P(A_1|S_i) \cdot P(S_i), \text{ i.e., } u(a) = \sum_{i=1}^{n} u(S_i) \cdot p_i.$$
- The sum defining $u(a)$ is the expected value of the outcome’s utility.
- So, we should select the action with the largest value of expected utility $u(a) = \sum p_i \cdot u(S_i)$. 
Subjective Probabilities

- Sometimes, we do not know the probabilities $p_i$ of different outcomes.
- In this case, we can gauge the subjective impressions about the probabilities.
- Let’s fix a prize (e.g., $1). For each event $E$, we compare:
  - a lottery $\ell_E$ in which we get the fixed prize if the event $E$ occurs and 0 if it does not occur, with
  - a lottery $\ell(p)$ in which we get the same amount with probability $p$.
- Here, $\ell(0) < \ell_E < \ell(1)$; so for some $p$, we switch from $\ell(p) < \ell_E$ to $\ell_E > \ell(p)$.
- This threshold value $ps(E)$ is called the subjective probability of the event $E$: $\ell_E \equiv \ell(ps(E))$.
- The utility of an action $a$ with possible outcomes $S_1, \ldots, S_n$ is thus equal to $u(a) = \sum_{i=1}^{n} ps(E_i) \cdot u(S_i)$. 
Traditional Approach Summarized

- We assume that
  - we know possible actions, and
  - we know the exact consequences of each action.
- Then, we should select an action with the largest value of expected utility.
Likert Scale in Terms of Traditional Decision Making

- Suppose that we have a Likert scale with $n + 1$ labels $0, 1, 2, \ldots, n$, ranging from the smallest to the largest.
  - We mark the smallest end of the scale with $x_0$ and begin to traverse.
  - As $x$ increases, we find a value belonging to label 1 and mark this threshold point by $x_1$.
  - This continues to the largest end of the scale which is marked by $x_{n+1}$
- As a result, we divide the range $[\underline{X}, \overline{X}]$ of the original variable into $n + 1$ intervals $[x_0, x_1], \ldots, [x_n, x_{n+1}]$:
  - values from the first interval $[x_0, x_1]$ are marked with label 0;
  - $\ldots$
  - values from the $(n + 1)$-st interval $[x_n, x_{n+1}]$ are marked with label $n$.
- Then, decisions are based only on the label, i.e., only on the interval to which $x$ belongs:

$$[x_0, x_1] \text{ or } [x_1, x_2] \text{ or } \ldots \text{ or } [x_n, x_{n+1}]$$
Which Decision To Choose Within Each Label?

- Since we only know the label $k$ to which $x$ belongs, we select $\tilde{x}_k \in [x_k, x_{k+1}]$ and make a decision based on $\tilde{x}_k$.
- Then, for all $x$ from the interval $[x_k, x_{k+1}]$, we use the decision $d(\tilde{x}_k)$ based on the value $\tilde{x}_k$.
- We should select intervals $[x_k, x_{k+1}]$ and values $\tilde{x}_k$ for which the expected utility is the largest.
Which Value $\tilde{x}_k$ Should We Choose

- To find this expected utility, we need to know two things:
  - the probability of different values of $x$; described by the probability density function $\rho(x)$;
  - for each pair of values $x'$ and $x$, the utility $u(x', x)$ of using a decision $d(x')$ when the actual value is $x$.

- In these terms, the expected utility of selecting a value $\tilde{x}_k$ can be described as
  \[
  \int_{x_k}^{x_{k+1}} \rho(x) \cdot u(\tilde{x}_k, x) \, dx.
  \]

- Thus, for each interval $[x_k, x_{k+1}]$, we need to select a decision $d(\tilde{x}_k)$ such that the above expression is maximized.

- Since the actual value $x$ can be in any of the $n + 1$ intervals, the overall expected utility is found by
  \[
  \sum_{k=0}^{n} \max \int_{x_k}^{x_{k+1}} \rho(x) \cdot u(\tilde{x}_k, x) \, dx.
  \]
Equivalent Reformulation In Terms of Disutility

- In the ideal case, for each value $x$, we should use a decision $d(x)$, and gain utility $u(x, x)$.
- In practice, we have to use decisions $d(x')$, and thus, get slightly worse utility values $u(x', x)$.
- The corresponding decrease in utility $U(x', x) \overset{\text{def}}{=} u(x, x) - u(x', x)$ is usually called disutility.
- In terms of disutility, the function $u(x', x)$ has the form
  \[ u(x', x) = u(x, x) - U(x', x), \]
- So, to maximize utility, we select the values $x_1, \ldots, x_n$ for which the disutility attains its smallest possible value:
  \[ \sum_{k=0}^{n} \min_{\tilde{x}_k} \int_{x_k}^{x_{k+1}} \rho(x) \cdot U(\tilde{x}_k, x) \, dx \to \min. \]
We focus on the use of Likert scales to elicit the values of the membership function \( \mu(x) \).

In our \( n \)-valued Likert scale:

- label 0 = \([x_0, x_1]\) corresponds to \( \mu(x) = 0/n \),
- label 1 = \([x_1, x_2]\) corresponds to \( \mu(x) = 1/n \),
- \ldots
- label \( n = [x_n, x_{n+1}] \) corresponds to \( \mu(x) = n/n = 1 \).

When \( n \) is huge, \( \mu(x) \) corresponds to the limit \( n \to \infty \) so the width of each interval is very small.

We can use the fact that each interval is narrow to simplify \( U(x', x) \) since \( x' \) and \( x \) belong to the same narrow interval.

Thus, the difference \( \Delta x \overset{\text{def}}{=} x' - x \) is small.
Using the Fact that Each Interval Is Narrow

Thus, we can expand $U(x + \Delta x, x)$ into Taylor series in $\Delta x$, and keep only the first non-zero term in this expansion.

$$U(x + \Delta x, x) = U_0(x) + U_1(x) \cdot \Delta x + U_2(x) \cdot \Delta x^2 + \ldots,$$

By definition of disutility,
$$U_0(x) = U(x, x) = u(x, x) - u(x, x) = 0$$

Similarly, since disutility is smallest when $x + \Delta x = x$, the first derivative is also zero.

So, the first nontrivial term is the second derivative
$$U_2(x) \cdot \Delta x^2 \approx U_2(x) \cdot (\tilde{x}_k - x)^2$$

So, we need to minimize the integral
$$\int_{x_k}^{x_{k+1}} \rho(x) \cdot U_2(x) \cdot (\tilde{x}_k - x)^2 \, dx.$$
The membership function $\mu(x)$ obtained by using Likert-scale elicitation is equal to

$$\mu(x) = \frac{\int_{X}^{x} (\rho(t) \cdot U_2(t))^{1/3} \, dt}{\int_{X}^{x} (\rho(t) \cdot U_2(t))^{1/3} \, dt},$$

where $\rho(x)$ is the probability density describing the probabilities of different values of $x$,

$$U_2(x) \overset{\text{def}}{=} \frac{1}{2} \cdot \frac{\partial^2 U(x + \Delta x, x)}{\partial^2 (\Delta x)},$$

$U(x', x) \overset{\text{def}}{=} u(x, x) - u(x', x)$, and

$u(x', x)$ is the utility of using a decision $d(x')$ corresponding to the value $x'$ in the situation in which the actual value is $x$. 
Comment:

- The resulting formula only applies to membership functions like “large” whose values monotonically increase with $x$.
- We can use a similar formula for membership functions like “small” which decrease with $x$.
- For “Approximately 0,” we separately apply these formulas to both increasing and decreasing parts.

The resulting membership degrees incorporate both probability and utility information.

This explains why fuzzy techniques often work better than probabilistic techniques without utility information.
Conclusion

- We have considered an ideal situation in which
  - we have full information about the probabilities $\rho(x)$, and
  - the user can always definitely decide between every two alternatives.
- In practice, we often only have intervals of possible values of $\rho(x)$.
- It is natural to assume that
  - $\rho(x) = \text{const}$ and $U_2(x) = \text{const}$ and
  - the resulting formula leads to a linear membership function (0 to 1 or 1 to 0) on the corresponding interval.
- This may explain why *triangular membership functions* (two such linear segments) are used in many fuzzy techniques.
- In the future, it is desirable to extend our formulas to the general interval-valued case.
Questions
Appendix 1: Utility Value

Let $A$ be any alternative such that $A_0 < A < A_1$; then:

- as $p$ increases from 0, $L(p) < A$;
- then, at some point, $L(p) > A$;
- so, there is a threshold separating values for which $L(p) < A$ from the values for which $L(p) > A$;
- this threshold is called the utility of alternative $A$:

$$u(A) \overset{\text{def}}{=} \sup\{p : L(p) < A\} = \inf\{p : L(p) > A\}$$

- Here, for every $\varepsilon > 0$, we have

$$L(u(A) - \varepsilon) < A < L(u(A) - \varepsilon).$$

- In this sense, the alternative $A$ is (almost) equivalent to $L(u(A))$; we will denote this almost equivalence by

$$A \equiv L(u(A)).$$
Appendix 2: Almost Uniqueness of Utility

- The definition of utility $u$ depends on the selection of two fixed alternatives $A_0$ and $A_1$.
- What if we use different alternatives $A'_0$ and $A'_1$?
- By definition of utility, every alternative $A$ is equivalent to a lottery $L(u(A))$ in which we get $A_1$ with probability $u(A)$ and $A_0$ with probability $1 - u(A)$.
- For simplicity, let us assume that $A'_0 < A_0 < A_1 < A'_1$. Then, for the utility $u'$, we get $A_0 \equiv L'(u'(A_0))$ and $A_1 \equiv L'(u'(A_1))$. 
Appendix 2: Almost Uniqueness of Utility

- So, the alternative $A$ is equivalent to a complex lottery in which:
  - we select $A_1$ with probability $u(A)$ and $A_0$ with probability $1 - u(A)$;
  - depending on which of the two alternatives $A_i$ we get, we get $A'_1$ with probability $u'(A_i)$ and $A'_0$ with probability $1 - u'(A_i)$.

- In this complex lottery, we get $A'_1$ with probability $u'(A) = u(A) \cdot (u'(A_1) - u'(A_0)) + u'(A_0)$.

- Thus, the utility $u'(A)$ is related with the utility $u(A)$ by a linear transformation $u' = a \cdot u + b$, with $a > 0$. 
Appendix 3: Reformulation In Terms of Disutility

► In the ideal case, for each value $x$, we should use a decision $d(x)$, and gain utility $u(x, x)$.

► In practice, we have to use decisions $d(x')$, and get slightly worse utility values $u(x', x)$.

► The corresponding decrease in utility $U(x', x) \overset{\text{def}}{=} u(x, x) - u(x', x)$ is usually called disutility.

► In terms of disutility, the function $u(x', x)$ has the form

$$u(x', x) = u(x, x) - U(x', x),$$
Appendix 3: Reformulation In Terms of Disutility

Thus, the optimized expression takes the form

\[
\int_{x_k}^{x_{k+1}} \rho(x) \cdot u(x, x) \, dx - \int_{x_k}^{x_{k+1}} \rho(x) \cdot U(\tilde{x}_k, x) \, dx.
\]

The first integral does not depend on \( \tilde{x}_k \); thus, the expression attains its maximum if and only if the second integral attains its minimum.

The resulting maximum thus takes the form

\[
\int_{x_k}^{x_{k+1}} \rho(x) \cdot u(x, x) \, dx - \min_{\tilde{x}_k} \int_{x_k}^{x_{k+1}} \rho(x) \cdot U(\tilde{x}_k, x) \, dx.
\]
Appendix 3: Reformulation In Terms of Disutility

Thus, we get the form

\[ \sum_{k=0}^{n} \int_{x_k}^{x_{k+1}} \rho(x) \cdot u(x, x) \, dx - \sum_{k=0}^{n} \min_{\tilde{x}_k} \int_{x_k}^{x_{k+1}} \rho(x) \cdot U(\tilde{x}_k, x) \, dx. \]

The first sum does not depend on selecting the thresholds.

Thus, to maximize utility, we should select the values \( x_1, \ldots, x_n \) for which the second sum attains its smallest possible value:

\[ \sum_{k=0}^{n} \min_{\tilde{x}_k} \int_{x_k}^{x_{k+1}} \rho(x) \cdot U(\tilde{x}_k, x) \, dx \to \min. \]
Appx 4: Derivations Leading to Membership Function

- In an $n$-valued scale:
  - the smallest label 0 corresponds to the value $\mu(x) = 0/n$,
  - the next label 1 corresponds to the value $\mu(x) = 1/n$,
  - ... 
  - the last label $n$ corresponds to the value $\mu(x) = n/n = 1$.

- Thus, for each $n$:
  - values from the interval $[x_0, x_1]$ correspond to the value $\mu(x) = 0/n$;
  - values from the interval $[x_1, x_2]$ correspond to the value $\mu(x) = 1/n$;
  - ... 
  - values from the interval $[x_n, x_{n+1}]$ correspond to the value $\mu(x) = n/n = 1$.

- The actual value of the membership function $\mu(x)$ corresponds to the limit $n \to \infty$, i.e., in effect, to very large values of $n$.

- Thus, in our analysis, we will assume that the number $n$ of labels is huge – and thus, that the width of each of $n + 1$ intervals $[x_k, x_{k+1}]$ is very small.
Appx 4: Derivations Leading to Membership Function

- The fact that each interval is narrow allows simplification of the expression $U(x', x)$.

- In the expression $U(x', x)$, both values $x'$ and $x$ belong to the same narrow interval.

- Thus, the difference $\Delta x \overset{\text{def}}{=} x' - x$ is small.

- So, we can expand $U(x', x) = U(x + \Delta x, x)$ into Taylor series in $\Delta x$, and keep only the first non-zero term.

- In general, we have

$$U(x + \Delta, x) = U_0(x) + U_1 \cdot \Delta x + U_2(x) \cdot \Delta x^2 + \ldots,$$

where

$$U_0(x) = U(x, x), \quad U_1(x) = \frac{\partial U(x + \Delta x, x)}{\partial (\Delta x)},$$

$$U_2(x) = \frac{1}{2} \cdot \frac{\partial^2 U(x + \Delta x, x)}{\partial^2 (\Delta x)}.$$ (7)
Here, by definition of disutility, we get
\[ U_0(x) = U(x, x) = u(x, x) - u(x, x) = 0. \]

Since the utility is the largest (and thus, disutility is the smallest) when \( x' = x \), i.e., when \( \Delta x = 0 \), the derivative \( U_1(x) \) is also equal to 0.

Thus, the first non-trivial term corresponds to the second derivative:
\[ U(x + \Delta x, x) \approx U_2(x) \cdot \Delta x^2, \]

reformulated in terms of \( \tilde{x}_k \) (that needs to be minimized)
\[ U(\tilde{x}_k, x) \approx U_2(x) \cdot (\tilde{x}_k - x)^2, \]

is substituted into the expression
\[ \int_{x_k}^{x_{k+1}} \rho(x) \cdot U(\tilde{x}_k, x) \, dx \]
We need to minimize the integral
\[ \int_{x_k}^{x_{k+1}} \rho(x) \cdot U_2(x) \cdot (\tilde{x}_k - x)^2 \, dx \]

by differentiating with respect to the unknown \( \tilde{x}_k \) and equating the derivative to 0.

Thus, we conclude that
\[ \int_{x_k}^{x_{k+1}} \rho(x) \cdot U_2(x) \cdot (\tilde{x}_k - x) \, dx = 0, \]
i.e., that
\[ \tilde{x}_k \cdot \int_{x_k}^{x_{k+1}} \rho(x) \cdot U_2(x) \, dx = \int_{x_k}^{x_{k+1}} x \cdot \rho(x) \cdot U_2(x) \, dx, \]
and thus, that
\[ \tilde{x}_k = \frac{\int_{x_k}^{x_{k+1}} x \cdot \rho(x) \cdot U_2(x) \, dx}{\int_{x_k}^{x_{k+1}} \rho(x) \cdot U_2(x) \, dx} \]
which can be simplified because the intervals are narrow.
We denote the midpoint of the interval \([x_k, x_{k+1}]\) by 
\[
\bar{x}_k \overset{\text{def}}{=} \frac{x_k + x_{k+1}}{2},
\]
and denote \(\Delta x \overset{\text{def}}{=} x - \bar{x}_k\),

then we have \(x = \bar{x}_k + \Delta x\).

Expanding into Taylor series in terms of a small value \(\Delta x\) and keeping only main terms, we get

\[
\rho(x) = \rho(\bar{x}_k + \Delta x) = \rho(\bar{x}_k) + \rho'(\bar{x}_k) \cdot \Delta x \approx \rho(\bar{x}_k),
\]

where \(f'(x)\) denoted the derivative of a function \(f(x)\), and

\[
U_2(x) = U_2(\bar{x}_k + \Delta x) = U_2(\bar{x}_k) + U_2'(\bar{x}_k) \cdot \Delta x \approx U_2(\bar{x}_k).
\]
Using these new $\rho(\bar{x}_k)$ and $U_2(\bar{x}_k)$, we get

\[
\tilde{x}_k = \frac{\rho(\bar{x}_k) \cdot U_2(\bar{x}_k) \cdot \int_{x_k}^{x_{k+1}} x \, dx}{\rho(\bar{x}_k) \cdot U_2(\bar{x}_k) \cdot \int_{x_k}^{x_{k+1}} dx} = \frac{\int_{x_k}^{x_{k+1}} x \, dx}{\int_{x_k}^{x_{k+1}} dx} = \int_{x_k}^{x_{k+1}} \frac{1}{2} \cdot \frac{(x_{k+1}^2 - x_k^2)}{x_{k+1} - x_k} = \frac{x_{k+1} + x_k}{2} = \bar{x}_k.
\]

Substituting $\tilde{x}_k = \bar{x}_k$ into the integral and understanding that, on the $k$-th interval, we have $\rho(x) \approx \rho(\bar{x}_k)$ and $U_2(x) \approx U_2(\bar{x}_k)$,

we conclude that the integral takes the form

\[
\int_{x_k}^{x_{k+1}} \rho(\bar{x}_k) \cdot U_2(\bar{x}_k) \cdot (\bar{x}_k - x)^2 \, dx = \rho(\bar{x}_k) \cdot U_2(\bar{x}_k) \cdot \int_{x_k}^{x_{k+1}} (\bar{x}_k - x)^2 \, dx.
\]
Appx 4: Derivations Leading to Membership Function

- When $x$ goes from $x_k$ to $x_{k+1}$, the difference $\Delta x$ between $x$ and the interval’s midpoint $\bar{x}_k$ ranges from $-\Delta_k$ to $\Delta_k$, where $\Delta_k$ is the interval’s half-width:

\[
\Delta_k \overset{\text{def}}{=} \frac{x_{k+1} - x_k}{2}.
\]

- In terms of the new variable $\Delta x$, the right-hand side of the integral has the form

\[
\int_{\bar{x}_k}^{x_{k+1}} (\bar{x}_k - x)^2 \, dx = \int_{-\Delta_k}^{\Delta_k} (\Delta x)^2 \, d(\Delta x) = \frac{2}{3} \cdot \Delta_k^3.
\]

- Thus, the integral takes the form

\[
\frac{2}{3} \cdot \rho(\bar{x}_k) \cdot U_2(\bar{x}_k) \cdot \Delta_k^3.
\]
The problem of selecting the Likert scale thus becomes the problem of minimizing the sum

$$\frac{2}{3} \cdot \sum_{k=0}^{n} \rho(\bar{x}_k) \cdot U_2(\bar{x}_k) \cdot \Delta_k^3.$$ 

Here,

$$\bar{x}_{k+1} = x_{k+1} + \Delta_{k+1} = (\bar{x}_k + \Delta_k) + \Delta_{k+1} \approx \bar{x}_k + 2\Delta_k,$$

so

$$\Delta_k = (1/2) \cdot \Delta\bar{x}_k,$$

where $\Delta\bar{x}_k \overset{\text{def}}{=} \bar{x}_{k+1} - \bar{x}_k.$

Thus, we get the form

$$\frac{1}{3} \cdot \sum_{k=0}^{n} \rho(\bar{x}_k) \cdot U_2(\bar{x}_k) \cdot \Delta_k^2 \cdot \Delta\bar{x}_k. \quad (11)$$
In terms of the membership function, we have $\mu(x_k) = k/n$ and $\mu(x_{k+1}) = (k + 1)/n$.

Since the half-width $\Delta_k$ is small, we have

$$\frac{1}{n} = \mu(x_{k+1}) - \mu(x_k) = \mu(x_k + 2\Delta_k) - \mu(x_k) \approx \mu'(x_k) \cdot 2\Delta_k,$$

thus,

$$\Delta_k \approx \frac{1}{2n} \cdot \frac{1}{\mu'(x_k)}.$$

Substituting this expression into the sum, we get

$$I = \sum_{k=0}^{n} \frac{\rho(x_k) \cdot U_2(x_k)}{(\mu'(x_k))^2} \cdot \Delta x_k.$$

(12)
The expression $I$ is an integral sum, so when $n \to \infty$, this expression tends to the corresponding integral

$$I = \int \frac{\rho(x) \cdot U_2(x)}{(\mu'(x))^2} \, dx.$$  

With respect to the derivative $d(x) \overset{\text{def}}{=} \mu'(x)$, we need to minimize the objective function

$$I = \int \frac{\rho(x) \cdot U_2(x)}{d^2(x)} \, dx$$  

under the constraint that

$$\int_{\overline{X}}^X d(x) \, dx = \mu(\overline{X}) - \mu(X) = 1 - 0 = 1.$$  

(13)
By using the Lagrange multiplier method, we can reduce to the unconstrained problem of minimizing the functional

\[ I = \int \frac{\rho(x) \cdot U_2(x)}{d^2(x)} \, dx + \lambda \cdot \int d(x) \, dx. \]

Differentiating with respect to \( d(x) \) and equating the derivative to 0, we conclude that

\[ -2 \cdot \frac{\rho(x) \cdot U_2(x)}{d^3(x)} + \lambda = 0, \]

i.e., that \( d(x) = c \cdot (\rho(x) \cdot U_2(x))^{1/3} \) for some constant \( c \).

Thus, \( \mu(x) = \int_{X}^{X} d(t) \, dt = c \cdot \int_{X}^{X} (\rho(t) \cdot U_2(t))^{1/3} \, dt. \)

The constant \( c \) must be determined by the condition that \( \mu(\bar{X}) = 1. \)

Thus, we arrive at the resulting formula.
Appendix 4: Resulting Formula

The membership function $\mu(x)$ obtained by using Likert-scale elicitation is equal to

$$
\mu(x) = \frac{\int_{x}^{X}(\rho(t) \cdot U_2(t))^{1/3} \, dt}{\int_{x}^{X}(\rho(t) \cdot U_2(t))^{1/3} \, dt},
$$

where $\rho(x)$ is the probability density describing the probabilities of different values of $x$,

$U_2(x) \overset{\text{def}}{=} \frac{1}{2} \cdot \frac{\partial^2 U(x + \Delta x, x)}{\partial^2(\Delta x)}$, 

$U(x', x) \overset{\text{def}}{=} u(x, x) - u(x', x)$, and

$u(x', x)$ is the utility of using a decision $d(x')$ corresponding to the value $x'$ in the situation in which the actual value is $x$. 