Problem set on Fuzzy Set Theory

A) Paper-based problems

A.1) Consider the following dataset:

<table>
<thead>
<tr>
<th>individuals</th>
<th>Calorie intake</th>
<th>Education</th>
<th>Housing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1500</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1800</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>2200</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>1200</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>1300</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>1700</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1000</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>2000</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>

Variable description:
calorie intake per day
educational attainment: 0=illiterate, 1-12 grades of edu (1-7 primary; 8-12 secondary)
housing (composite index): from 1 (worst) to 5 (best)

Calculate $\mu(x_i)$ according to the membership functions discussed (linear, trapezoidal and logistic) and compare the results (set the thresholds as you consider appropriate).

A.2) The table below shows md of a set of individuals to the fs “being educated” and “being nourished”

<table>
<thead>
<tr>
<th>individuals</th>
<th>Being educated</th>
<th>Being nourished</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1,00</td>
<td>1,00</td>
</tr>
<tr>
<td>2</td>
<td>0,75</td>
<td>1,00</td>
</tr>
<tr>
<td>3</td>
<td>0,50</td>
<td>0,25</td>
</tr>
<tr>
<td>4</td>
<td>0,00</td>
<td>0,50</td>
</tr>
<tr>
<td>5</td>
<td>0,25</td>
<td>0,50</td>
</tr>
<tr>
<td>6</td>
<td>1,00</td>
<td>0,00</td>
</tr>
<tr>
<td>7</td>
<td>0,00</td>
<td>0,00</td>
</tr>
<tr>
<td>8</td>
<td>0,50</td>
<td>0,50</td>
</tr>
</tbody>
</table>

Calculate the aggregation operators, including average operators, presented and compare the results. What do you consider most suitable?
B) Computer-based problems (using Stata) and ‘Indonesian_Data.dta’ (from Suman’s problem)

There are achievement levels of 24,920 adults (age > 15 years) for three dimensions: Per capita expenditure (Rupiah), Body mass index (BMI, Kg/m2), and years of education (see Suman’s problem set for the variable description)

b.1) Calculate \( \mu(x_{ij}) \) according to the membership functions discussed (linear, trapezoidal and logistic) and compare the results (you can normalize as suggested by Suman or as you consider most appropriate)

b.2) Calculate the aggregation operators, including averaging operators, and compare the results
Appendix

Membership functions

Linear function

\[ \mu(x) = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \]

Trapezoidal function

\[ \mu(x) = 1 \quad \text{if} \ x_{\text{min}} \leq x \leq z \]
\[ \mu(x) = 0 \quad \text{if} \ x_k \leq x \leq x_{\text{max}} \]
\[ \mu(x) = \frac{x_k - x}{x_k - z} \quad \text{if} \ z \leq x \leq x_k \]

Sigmoid (or logistic) function

\[ x_k = \mu(x) = 0 \quad x_h = \mu(x) = 0.5 \quad x_w = \mu(x) = 1 \]
\[ \mu(x) = 1 \quad \text{if} \ x_{\text{min}} \leq x \leq x_w \]
\[ \mu(x) = 0 \quad \text{if} \ x_k \leq x \leq x_{\text{max}} \]
\[ \mu(x) = 1 - \frac{1}{2} \left[ \frac{x_w - x}{x_w - x_k} \right]^2 \quad \text{if} \ x_w \leq x \leq x_h \]
\[ \mu(x) = \frac{1}{2} \left[ \frac{x_k - x}{x_k - x_h} \right]^2 \quad \text{if} \ x_h \leq x \leq x_k \]

Based on cumulative distribution functions (Cheli, Lemmi)
\[\mu(x^k) = \begin{cases} \mu(x_{k-1}) & \text{if } k = 1 \\ \frac{\mu(x_{k-1}) + F(x_k) - F(x_{k-1})}{1 - F(x_k)} & \text{if } k > 1 \end{cases}\]

or simply:
\[\mu(x^k) = \max[0, (F(x_k) - F(x_1))/(1 - F(x_1))]\]

Aggregation operators

standard intersection
\[\mu_{A \cap B} = \min[\mu_A, \mu_B]\]

standard union
\[\mu_{A \cup B} = \max[\mu_A, \mu_B]\]

weak intersection (or algebraic product)
\[\mu_{AB} = [\mu_A \cdot \mu_B]\]

weak union (or algebraic sum)
\[\mu_{A+B} = [\mu_A + \mu_B - \mu_A \cdot \mu_B]\]

bounded difference
\[\mu_{A \cap B} = \max[0, \mu_A + \mu_B - 1]\]

bounded sum
\[\mu_{A \cup B} = \min[1, \mu_A + \mu_B]\]

average
- simple
\[h^a = h(a_1, a_2, \ldots, a_n) = [a_1^a + a_2^a + \ldots + a_n^a]/n^{1/a}\]

- weighted
\[h^a = h(a_1, a_2, \ldots, a_n, w_1, w_2, \ldots, w_n) = [\Sigma w_i a_i^a]/n\]
with \(w_i \geq 0\) and \(\Sigma w_i = 1\)

- frequency-based
\[w_i = \ln 1/f_i\]
\[w_i = \ln[1/n\Sigma \mu_{ij}]\]