Some Investigations on Higher Mathematics Scores for Chinese University Students

Xun Ge and Jingyu Qian

Abstract—To investigate some relations between higher mathematics scores in Chinese graduate student entrance examination and calculus (resp. linear algebra, probability statistics) scores in subject’s completion examination of Chinese university, we select 20 students as a sample, to take higher mathematics score as a decision attribute and take calculus score, linear algebra score, probability statistics score as condition attributes. In this paper, we are based on rough-set theory (rough-set theory is a logic-mathematical method proposed by Z. Pawlak. In recent years, this theory has been widely implemented in the many fields of natural science and societal science.) to investigate importance of condition attributes with respective to decision attribute and strength of condition attributes supporting decision attribute. Results of this investigation will be helpful for university students to raise higher mathematics scores in Chinese graduate student entrance examination.

Keywords—Rough set, higher mathematics scores, decision attribute, condition attribute.

I. INTRODUCTION

In Chinese university higher education, higher mathematics is an important subject for students of science-engineering department, which includes mainly calculus, linear algebra and probability statistics. Many Chinese university students can not go to graduate school every year, because they can not to pass Chinese graduate student entrance examination of higher mathematics. It is natural to consider the following question.

Question 1.1: What are relations between higher mathematics scores in Chinese graduate student entrance examination and calculus (resp. linear algebra, probability statistics) scores in subject’s completion examination of Chinese university?

It is an interesting work to investigate the above question. For this purpose, the second author of this paper selected 20 students from science-engineering department of Yancheng Teachers University at random as a sample, and collected their calculus scores, linear algebra scores, probability statistics scores in Yancheng Teachers University subject’s completion examinations. In addition, he held a higher mathematics examination to these 20 students, which simulated Chinese graduate student entrance examination, and obtained their examination scores. In his investigation, he used traditional analytic methods (e.g. synthesis, appraisal, stratification and estimate of probability). However, he felt it appropriate unlikely to use these methods because these examination scores are uncertain. This leads us to give a further investigation by some new methods. Rough-set theory, which is a logic-mathematical method proposed by Z. Pawlak, has shown to be an effective tool in analyzing this type of issues [7], [8], [9]. In recent years, this theory has been widely implemented in the many fields of natural science and societal science [1], [2], [4], [5], [6], [10], [11], [13], [14], [15], [16], [17].

In this paper, we extract useful information hidden these data selected by the above and use rough-set theory to give a further investigation for Question 1.1. More precisely, we take higher mathematics score as a decision attribute and take calculus score, linear algebra score and probability statistics score as condition attributes. Based on rough-set theory, we investigate importance of condition attributes with respective to decision attribute and strength of condition attributes supporting decision attribute. Results of this investigation will be helpful for Chinese university students to raise higher mathematics scores in Chinese graduate student entrance examination.

II. PROPADEUTICS

Propaedeutics in this section belongs to Z. Pawlak (see [7], [8], [9], for example).

Notation 2.1: (1) For a set \( B \), \( |B| \) denotes the cardinal of \( B \).

(2) For a family of sets \( \mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_k \), \( \{ \mathcal{F}_i : i = 1, 2, \ldots, k \} = \{ \{ F_i : i = 1, 2, \ldots, k \} : F_i \in \mathcal{F}_i, i = 1, 2, \ldots, k \} \).

(3) Let \( R \) be an equivalence relation on a set \( U \). \( U/R \) denotes the family consisting of all equivalence classes with respect to \( R \) and \( [u] \) denotes the equivalence class with respect to \( R \) containing \( u \in U \).

(4) Let \( \mathcal{R} \) be a family of equivalence relations on \( U \). Then \( \{ U/R : R \in \mathcal{R} \} \) is a partition of \( U \) and is denoted by \( U/\mathcal{R} \).

The equivalence relation induced by \( U/\mathcal{R} \) is also denoted by \( \mathcal{R} \).

Definition 2.2: \( S = (U, A, V, f) \) is called an information system.

(1) \( U \), a nonempty finite set, is called the universe of discourse.

(2) \( A = C \cup D \) is a finite set of attributes, where \( C \) and \( D \) are disjoint nonempty sets of condition attributes and decision attributes respectively.

(3) \( f : U \times A \rightarrow V \) is an information function.

(4) \( V = \{ V_\alpha : \alpha \in \mathcal{A} \} \), where \( V_\alpha = \{ f(u, \alpha) : u \in U \} \).

Remark 2.3: An information system \( S = (U, A, V, f) \) can be expressed a date table, which is called decision table, whose columns are labeled by elements of \( A \), rows are labeled by elements of \( U \), and \( f(u, \alpha) \) lies in the cross of the row labeled by \( u \) and the column labeled by \( \alpha \).
Notation 2.4: Let $S = (U, C, D, V, f)$ be an information system.

(1) For $a \in C$, $D$, we define an equivalence relation $\sim$ on $U$ as follows:
$$u_i \sim u_j \iff f(u_i, a) = f(u_j, a).$$
$U/a$ denotes the family consisting of all equivalence classes with respect to $\sim$.

(2) For $B \subset C$, $D$, \{ $\{u/b : b \in B\}$ is a partition of $U$, which is denoted $U/B$. The equivalence relation induced by $U/B$ is also denoted by $B$.

Definition 2.5: Let $R$ be an equivalence relation on an universe $U$ of discourse, and $X \subset U$. Put $R(X) = \{ [u] | [u] \in U/R, [u] \subset X \}$. $R(X)$ is called lower approximation of $X$.

III. DECISION TABLE

In this section, we establish an information system for our investigation, which is expressed by the following decision table. This information system includes all information we need in this investigation, where higher mathematics scores were obtained from higher mathematics examination simulated Chinese graduate student entrance examination and calculus scores, linear algebra scores and probability statistics scores were obtained from Yancheng Teachers University subject’s completion examination.

<table>
<thead>
<tr>
<th>Decision Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U$</td>
</tr>
<tr>
<td>$u_1$</td>
</tr>
<tr>
<td>$u_2$</td>
</tr>
<tr>
<td>$u_3$</td>
</tr>
<tr>
<td>$u_4$</td>
</tr>
<tr>
<td>$u_5$</td>
</tr>
<tr>
<td>$u_6$</td>
</tr>
<tr>
<td>$u_7$</td>
</tr>
<tr>
<td>$u_8$</td>
</tr>
<tr>
<td>$u_9$</td>
</tr>
<tr>
<td>$u_{10}$</td>
</tr>
<tr>
<td>$u_{11}$</td>
</tr>
<tr>
<td>$u_{12}$</td>
</tr>
<tr>
<td>$u_{13}$</td>
</tr>
<tr>
<td>$u_{14}$</td>
</tr>
<tr>
<td>$u_{15}$</td>
</tr>
<tr>
<td>$u_{16}$</td>
</tr>
<tr>
<td>$u_{17}$</td>
</tr>
<tr>
<td>$u_{18}$</td>
</tr>
<tr>
<td>$u_{19}$</td>
</tr>
<tr>
<td>$u_{20}$</td>
</tr>
</tbody>
</table>

The above decision table gives an information system $S = (U, C, D, V, f)$, where $U = \{ u_1, u_2, \ldots, u_{20} \}$, $C = \{ c_1, c_2, c_3 \}$, $D = \{ d \}$, $f$ and $V$ are given as Definition 2.2 and Remark 2.3. For some explanations of the above decision table, it is necessary to give the following remarks.

Remark 3.1: $U$ is the sample of 20 students.

Remark 3.2: $c_1, c_2, c_3$ are three condition attributes in the information system, i.e., $c_1, c_2, c_3$ denote linear algebra score, calculus score and probability statistics score respectively. $d$ is the decision attribute in the information system, i.e., $d$ denotes higher mathematics score.

Remark 3.3: Calculus for 20 students come from Yancheng Teachers University subject’s completion examination.

(1) $c_{11}$ indicates score lower than 60.
(2) $c_{21}$ indicates score between 60 and 80.
(3) $c_{31}$ indicates score between 81 and 100.

Remark 3.4: Linear algebra scores for 20 students come from Yancheng Teachers University subject’s completion examination.

(1) $c_{12}$ indicates score lower than 60.
(2) $c_{22}$ indicates score between 60 and 80.
(3) $c_{32}$ indicates score between 81 and 100.

Remark 3.5: probability statistics scores for 20 students come from Yancheng Teachers University subject’s completion examination.

(1) $c_{13}$ indicates score lower than 60.
(2) $c_{23}$ indicates score between 60 and 80.
(3) $c_{33}$ indicates score between 81 and 100.

Remark 3.6: Higher mathematics scores for 20 students come from higher mathematics examination simulated Chinese graduate student entrance examination.

(1) $d_1$ indicates score lower than 90.
(2) $d_2$ indicates score between 90 and 120.
(3) $d_3$ indicates score between 120 and 150.

Proposition 3.7: The following are some related partitions of $U$.

(1) $U/c_1 = \{ \{ u_1, u_2, u_3, u_4, u_5, u_13, u_14, u_15 \}, \{ u_6, u_16, u_17, u_{18}, u_{19}, u_{20} \} \}$.

(2) $U/c_2 = \{ \{ u_1, u_3, u_4, u_5, u_6, u_7, u_{18}, u_{19}, u_{20}, u_{21}, u_{22}, u_{23}, u_{24}, u_{25}, u_{26}, u_{27}, u_{28}, u_{29}, u_{30} \} \}$.

(3) $U/c_3 = \{ \{ u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_{18}, u_{19}, u_{20}, u_{21}, u_{22}, u_{23}, u_{24}, u_{25}, u_{26}, u_{27}, u_{28}, u_{29}, u_{30} \} \}$.

(4) $U/d = \{ \{ u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}, u_{13}, u_{14}, u_{15}, u_{16}, u_{17}, u_{18}, u_{19}, u_{20} \} \}$.

(5) $U/C = \{ \{ u_1 \}, \{ u_3 \}, \{ u_4 \}, \{ u_{12} \}, \{ u_{13} \}, \{ u_{14} \}, \{ u_{15} \}, \{ u_{16} \}, \{ u_{17} \}, \{ u_{18} \}, \{ u_{19} \}, \{ u_{20} \} \}$.

(6) $U/c_2, c_3 = \{ \{ u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}, u_{13}, u_{14}, u_{15}, u_{16}, u_{17}, u_{18}, u_{19}, u_{20} \} \}$.

IV. IMPORTANCE

Definition 4.1: Let $\mathcal{R}$ be a family of equivalence relations on $U$ and $R' \subset \mathcal{R}$. Let $Q$ be an equivalence relation on $U$. (1) Put $\text{pos}_{\mathcal{R}}(Q) = \{ R(X) : X \in U/Q \}$.  

$\text{pos}_{\mathcal{R}}(Q)$ is called positive region of $Q$ with respect to $\mathcal{R}$.

(2) Put $\gamma_{\mathcal{R}}(Q) = \frac{\text{pos}_{\mathcal{R}}(Q)}{|U|}$.  

$\gamma_{\mathcal{R}}(Q)$ is called positive degree of $Q$ with respect to $\mathcal{R}$.  

(3) Put $\sigma_{\mathcal{R}}(Q') = \gamma_{\mathcal{R}}(Q) - \gamma_{\mathcal{R}}(Q')$.  

$\sigma_{\mathcal{R}}(Q)$ is called importance of $R'$ with respect to $Q$.

Remark 4.2: For information system $S = (U, C, D, V, f)$, let $c \in C$. Put  

$$\sigma_{\mathcal{R}}(c) = \gamma_{\mathcal{R}}(D) - \gamma_{\mathcal{R}}(c \downarrow c)(D).$$
Then $\sigma_{CD}(c)$ is the importance of condition attribute $c$ with respect to decision attribute $d$.

**Lemma 4.3:** The following hold.

1. $\text{pos}_{\text{C}}(D) = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}, u_{13}, u_{14}, u_{15}, u_{16}, u_{17}, u_{18}, u_{19}, u_{20}\}$

2. $\text{pos}_{c_1, c_2}(D) = \{u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}, u_{13}, u_{14}, u_{15}, u_{16}, u_{17}, u_{18}, u_{19}, u_{20}\}$

3. $\text{pos}_{c_1, c_2}(D) = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}, u_{13}, u_{14}, u_{15}, u_{16}, u_{17}, u_{18}, u_{19}, u_{20}\}$

4. $\text{pos}_{c_1, c_2}(D) = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}, u_{13}, u_{14}, u_{15}, u_{16}, u_{17}, u_{18}, u_{19}, u_{20}\}$

**Proof:** By Proposition 3.7(4), (6), (7) (8) and Definition 4.1(1), we obtain the following.

1. $\gamma(C) \geq 0.65$.

2. $\gamma_{c_1, c_2}(D) = 0.45$.

3. $\gamma_{c_1, c_2}(D) = 0.50$.

4. $\gamma_{c_1, c_2}(D) = 0.60$.

**V. SUPPORT DEGREES**

**Definition 5.1:** For information system $S = (U, C, D, V, f)$, let $W \subseteq U$ and $c \in C$.

1. $S_c(W)$ is called a support subset of $W$ with respect to condition attribute $c$, where $S_c(W) = \{u : u \in U \cap \{c \in W\}\}$.

2. $S_{c'}(W)$ is called a support degree of $W$ with respect to condition attribute $c$, where $S_{c'}(W) = \frac{|S_c(W)|}{|U|}$.

3. $S_{c}(d)$ is called a support subset of decision attribute $d$ with respect to condition attribute $c$, where $S_{c}(d) = \{S_c(W) : W \in U/d\}$

4. $S_{c}(d)$ is called a support degree of decision attribute $d$ with respect to condition attribute $c$, where $S_{c}(d) = \frac{|S_c(d)|}{|U|}$.

**Remark 5.2:** By rough-set theory, $S_{c}(d)$ denotes the strength of condition attribute $c$ supporting decision attribute $d$.

**Proposition 5.3:** Let $S = (U, C, D, V, f)$ be an information system, $c \in C$ and $D = \{d\}$. If $U \cap U = \{W_1, W_2, \ldots, W_k\}$, then the following hold.

1. $S_c(d) = S_{c_1}(W_1) \cap S_{c_2}(W_2) \cap \cdots \cap S_{c_k}(W_k)$.

2. $i, j = 1, 2, \ldots, k, i \neq j, S_c(W_i) \cap S_c(W_j) = \emptyset$.

3. $S_{c}(d) = S_{c_1}(W_1) + S_{c_2}(W_2) + \cdots + S_{c_k}(W_k)$.

**Proof:** (1) It holds by Definition 5.1(1) immediately.

(2) Since $U \cap U = \{W_1, W_2, \ldots, W_k\}$ is a partition of $U$, $W_1 \cap W_j = \emptyset$ for all $i, j = 1, 2, \ldots, k$ where $i \neq j$. Note that $S_c(W_i) \subseteq W_i, S_c(W_j) \subseteq W_j$. So $S_c(W_i) \cap S_c(W_j) = \emptyset$.

(3) By the above (1) and (2).

**Notation 5.4:** Put $W_1 = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}, u_{13}, u_{14}, u_{15}, u_{16}, u_{17}, u_{18}, u_{19}, u_{20}\}$. Then $U \cap U = \{W_1, W_2, W_3\}$.

**Lemma 5.5:** The following hold.

1. $S_{c_1}(W_1) = \emptyset, S_{c_2}(W_1) = \emptyset$.

2. $S_{c_2}(W_2) = \emptyset, S_{c_3}(W_2) = \emptyset$.

3. $S_{c_3}(W_3) = \emptyset, S_{c_3}(W_3) = \emptyset$.

**Remark 4.6:** By Remark 4.2 and Proposition 4.5, we have the following conclusions.
Proof: They are obtained by some simple computations and we omit them.

**Lemma 5.6:** The following hold.

1. \( \text{spt}_{c_1}(W_1) = 0.40, \text{spt}_{c_2}(W_2) = \text{spt}_{c_3}(W_3) = 0 \)
2. \( \text{spt}_{c_2}(W_1) = 0.25, \text{spt}_{c_2}(W_2) = \text{spt}_{c_2}(W_3) = 0 \)
3. \( \text{spt}_{c_2}(W_1) = 0.20, \text{spt}_{c_2}(W_2) = \text{spt}_{c_2}(W_3) = 0 \)

**Proof:** By Remark 4.2 and Lemma 4.4, we obtain the following.

1. \( \frac{|S_{c_1}(W_1)|}{|U|} = 0.40, \frac{|S_{c_2}(W_2)|}{|U|} = 0, \frac{|S_{c_3}(W_3)|}{|U|} = 0 \)
2. \( \frac{|S_{c_2}(W_1)|}{|U|} = 0.25, \frac{|S_{c_2}(W_2)|}{|U|} = 0, \frac{|S_{c_3}(W_3)|}{|U|} = 0 \)
3. \( \frac{|S_{c_2}(W_1)|}{|U|} = 0.20, \frac{|S_{c_2}(W_2)|}{|U|} = 0, \frac{|S_{c_3}(W_3)|}{|U|} = 0 \)

Now we give the strength of condition attribute \( c \in C \) supporting decision attribute \( d \).

**Proposition 5.7:** The following hold.

1. \( \text{spt}_{c_1}(d) = 0.40 \)
2. \( \text{spt}_{c_2}(d) = 0.25 \)
3. \( \text{spt}_{c_3}(d) = 0.20 \)

**Proof:** By Proposition 5.3(3), we obtain the following.

1. \( \text{spt}_{c_1}(d) = \text{spt}_{c_2}(W_1) + \text{spt}_{c_3}(W_2) + \text{spt}_{c_3}(W_3) = 0.40 + 0 + 0 = 0.40 \)
2. \( \text{spt}_{c_2}(d) = \text{spt}_{c_2}(W_1) + \text{spt}_{c_2}(W_2) + \text{spt}_{c_2}(W_3) = 0.25 + 0 + 0 = 0.25 \)
3. \( \text{spt}_{c_3}(d) = \text{spt}_{c_3}(W_1) + \text{spt}_{c_3}(W_2) + \text{spt}_{c_3}(W_3) = 0.20 + 0 + 0 = 0.20 \)

**Remark 5.8:** By Remark 5.2 and Proposition 5.7, we have the following conclusions.

1. The strength of calculus score supporting higher mathematics score is maximal (the strength is 0.40).
2. The strength of linear algebra score supporting higher mathematics score is between the strength of calculus score and the strength of probability statistics score (the strength is 0.25).
3. The strength of probability statistics score supporting higher mathematics score is minimal (the strength is 0.20).

**VI. Postscript**

1. The investigation in this paper is conducted with a sample of 20 students. The validity of the research conclusion and associated discussions is limited by the relatively small sample size. However, as stated earlier, results of this investigation will be helpful for Chinese university students to raise higher mathematics scores in Chinese graduate student entrance examination.
2. The investigation in this paper is based on partitions of the finite universe \( U \) of discourse, but by using these partitions we are not able to solve neighboring question in numerical representations for some factor attributes. In recent years, the Rough Set theory has been developed from partitions of the universe of discourse to covers of the universe of discourse (see [12], [18], for example), which may provide a satisfactory solution for this neighboring question. Further exploratory might be performed towards this direction.

**Acknowledgment**

This project is supported by NSFC (No. 10571151 and 10671173).

**References**


