

UNIVERSIDADE FEDERAL FLUMINENSE
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

ROBERTO MALHEIROS MOREIRA FILHO

REDUÇÃO DE VALORES NO CRITÉRIO DE DECISÃO EM APLICAÇÕES DE ROUGH
SETS COM DOMINÂNCIA E SEUS IMPACTOS NA QUALIDADE DA APROXIMAÇÃO

NITERÓI

2012

ROBERTO MALHEIROS MOREIRA FILHO

REDUÇÃO DE VALORES NO CRITÉRIO DE DECISÃO EM APLICAÇÕES DE ROUGH
SETS COM DOMINÂNCIA E SEUS IMPACTOS NA QUALIDADE DA APROXIMAÇÃO

Tese apresentada ao Curso de Doutorado em Engenharia de Produção da Universidade Federal Fluminense como requisito parcial para obtenção de grau de Doutor. Área de Concentração: Sistemas, Apoio à Decisão e Logística.

Orientador: Prof. ANNIBAL PARRACHO SANT'ANNA, PhD.

Niterói

2012

ROBERTO MALHEIROS MOREIRA FILHO

REDUÇÃO DE VALORES NO CRITÉRIO DE DECISÃO EM APLICAÇÕES DE ROUGH
SETS COM DOMINÂNCIA E SEUS IMPACTOS NA QUALIDADE DA APROXIMAÇÃO

Tese apresentada ao Curso de Doutorado em Engenharia de Produção da Universidade Federal Fluminense como requisito parcial para obtenção de grau de Doutor. Área de Concentração: Sistemas, Apoio à Decisão e Logística.

Aprovada em: 30 de julho de 2012.

BANCA EXAMINADORA

Prof. ANNIBAL PARRACHO SANT'ANNA – Orientador
Universidade Federal Fluminense

Prof. CARLOS FRANCISCO SIMÕES GOMES
Universidade Federal Fluminense

Prof. LUIS ALBERTO DUNCAN RANGEL
Universidade Federal Fluminense

Prof. JOSÉ FABIANO SERRA COSTA
Universidade do Estado do Rio de Janeiro

Prof. FERNANDO MARQUES DE ALMEIDA NOGUEIRA
Universidade Federal de Juiz de Fora

Niterói

2012

À minha esposa Sybila e ao meu filho Robertinho pela atenção e compreensão nos momentos mais difíceis. Sem vocês o caminho seria mais árduo.

AGRADECIMENTO

Ao professor Annibal pela competência, capacidade e dedicação que permitiram a realização deste trabalho.

RESUMO

A criação de regras de apoio à decisão com base em sistemas de informação é o objeto central da Teoria dos Conjuntos Aproximativos - TCA (*Rough Sets Theory*). Um trabalho apresentado por Pawlak em 1982 deu início a diversos estudos com o objetivo de criar regras de decisão baseadas em sistemas de informações com múltiplos atributos condicionantes e um ou mais atributos de decisão. Ao longo do tempo, os atributos com relação de dominância, onde há uma escala de valores, vêm recebendo destaque. Para lidar com este tipo de dados foi desenvolvido a DRSA (Dominance-based Rough Sets Approach). O excessivo rigor exigido para a criação de regras pela proposta básica do DRSA fez com que novas propostas surgissem. Além das regras determinísticas, com 100% de certeza, surgiram regras probabilísticas, com algum percentual controlado de incerteza. Existem algumas vertentes de estudo de aplicações de DRSA e a abordagem aqui proposta explora a possibilidade de aumento na qualidade da aproximação e, conseqüentemente, na qualidade das regras geradas, considerando a possibilidade de união de algumas classes do atributo de decisão com princípio de dominância. Com isto, são preservados os princípios do uso de DRSA. De acordo com a necessidade do pesquisador, a redução de classes pode ser utilizada em conjunto com as outras alternativas de DRSA apresentadas até o momento. Duas novas propostas de união de classes do atributo do critério de decisão são apresentadas, comentadas e criticadas nesta tese, uma baseada em função densidade de probabilidades e outra baseada em transformações probabilísticas.

Palavras-chave: Conjuntos Aproximativos. Dominância. Redução de Classes. Transformações Probabilísticas.

ABSTRACT

Creating rules for the support of the decision process is the main subject of Rough Sets Theory. The study first published by Pawlak in 1982 was a catalyst of several studies focusing on creating rules for the support of the decision making process based on multiple attributes conditioning one or more decision attributes. As the studies evolved, attributes with the feature of dominance – attributes measured in some scale – have caught attention. This gave rise to DRSA (Dominance-based Rough Sets Approach). The excessively strict guidelines of DRSA original studies led to the creation of new guidelines that consider not only deterministic rules – 100% certainty – but also probabilistic rules, which account for a certain degree of uncertainty. There are other developments of DRSA and the approach here taken evaluates alternatives for enhancing the quality of the approximation evaluation, therefore enhancing the quality of the rules, by clustering classes of values of decision attributes without compromising the guiding principles of DRSA. According to the need of researcher, the reduction of classes can be used in conjunction with other alternatives of development of DRSA. Two different proposals for the clustering of attributes are presented and evaluated in this study, one based on density functions and the other based on probabilistic transformations.

Keywords: Rough Sets. Dominance. Reduction of Classes. Probabilistic Transformations.

SUMÁRIO

1 INTRODUÇÃO, p. 17

1.1 INTRODUÇÃO, p.17

1.2 JUSTIFICATIVA, p. 19

1.3 FORMULAÇÃO DE HIPÓTESES, p. 20

1.4 ESTRUTURA DO TRABALHO, p. 22

2 TEORIA DOS CONJUNTOS APROXIMATIVOS, p. 23

2.1 INDISCERNIBILIDADE, p. 24

2.2 REDUÇÃO DE ATRIBUTOS, p. 29

2.3 FUNÇÃO DE DISCERNIBILIDADE, p. 30

2.4 SIGNIFICÂNCIA DE ATRIBUTOS, p. 31

2.5 EXEMPLOS DE USO DE CONJUNTOS APROXIMATIVOS COM PRINCÍPIOS DE DOMINÂNCIA, p. 32

3 DRSA, p. 34

3.1 INTRODUÇÃO, p. 34

3.2 VC-DRSA, p. 37

4 PROPOSTAS ALTERNATIVAS AO USO DIRETO DE VC-DRSA, p. 41

4.1 INTRODUÇÃO, p. 41

4.2 MONOTONICIDADE DE PERTINÊNCIA À APROXIMAÇÃO INFERIOR, p. 46

4.3 NOVAS PROPOSTAS DE INCLUSÃO À APROXIMAÇÃO INFERIOR, p. 47

4.4 OUTRAS PROPOSTAS (VP-DRSA e ISVP-DRSA), p. 51

5 REDUÇÃO DE VALORES PARA O ATRIBUTO DE DECISÃO, p. 62

5.1 INTRODUÇÃO, p. 62

5.2 UNIÃO DE ATRIBUTOS NO CRITÉRIO DE DECISÃO, p. 64

5.3 IMPACTOS DA UNIÃO DE VALORES NO ATRIBUTO DE DECISÃO, p. 65

5.4 DISTRIBUIÇÃO CONTÍNUA A SER UTILIZADA NO CRITÉRIO DE DECISÃO, p. 70

5.4.1 Distribuição exponencial, p. 70

5.4.2 Distribuição normal, p. 75

5.4.3 Distribuição beta, p. 79

5.4.4 Outras distribuições, p. 86

6 CONJUNTOS APROXIMATIVOS COM REDUÇÃO DE CLASSES NO ATRIBUTO DE DECISÃO, p. 87

6.1 PROPOSTA 1 – UNIÃO POR FUNÇÃO DE DENSIDADE DE PROBABILIDADE, p. 87

6.2 EXEMPLO COM USO DA DISTRIBUIÇÃO NORMAL – PROPOSTA 1, p. 90

6.3 EXEMPLO COM USO DA DISTRIBUIÇÃO EXPONENCIAL – PROPOSTA 1, p. 94

6.4 EXEMPLO COM USO DA DISTRIBUIÇÃO BETA – PROPOSTA 1, p. 97

6.5 DIFERENÇAS ENCONTRADAS NAS CLASSES UNIDAS – PROPOSTA 1, p. 102

6.5.1 Exemplo usando distribuição normal, p. 102

6.5.2 Exemplo usando distribuição exponencial, p. 106

6.5.3 Exemplo usando distribuição beta, p. 110

6.6 PROPOSTA 2 – UNIÃO POR TRANSFORMAÇÃO PROBABILÍSTICA, p. 114

6.6.1 Exemplo usando a probabilidade de ser o maior valor, p. 116

6.6.2 Exemplo usando a probabilidade de ser o menor valor, p. 118

6.7 DIFERENÇAS ENCONTRADAS NAS CLASSES UNIDAS – PROPOSTA 2, p. 121

6.7.1 Exemplo usando a probabilidade de ser o maior valor, p. 121

6.7.2 Exemplo usando a probabilidade de ser o menor valor, p. 124

7 CONCLUSÃO, p. 127

8 REFERÊNCIAS, p. 130

ANEXOS, p. 134

ANEXO 1, p. 134

ANEXO 2, p. 137

ANEXO 3, p. 140

ANEXO 4, p. 141

ANEXO 5, p. 142

ANEXO 6, p. 144

ANEXO 7, p. 146

ANEXO 8, p. 149

ANEXO 9, p. 152

LISTA DE FIGURAS

- Figura 1 - Aproximações em TCA, p. 27
- Figura 2 - Atributos com dominância, p. 33
- Figura 3 - Inconsistência de atributos, p. 37
- Figura 4 - Visualização gráfica dos objetos, p. 42
- Figura 5 - Cálculo das aproximações, p. 45
- Figura 6 - Visualização gráfica após índice l, p. 46
- Figura 7 - Objetos de uma classe de dominância, p. 50
- Figura 8 - Objetos de uma classe de dominância, p. 50
- Figura 9 - Regiões P-positivas (VP_DRSA), p. 57
- Figura 10 - Regiões p-positivas (ISVP-DRSA), p. 61
- Figura 11 - Região de fronteira, p. 67
- Figura 12 - Região de fronteira (VC-DRSA), p. 69
- Figura 13 - Região de fronteira - classes unidas, p. 69
- Figura 14 - Aproximação para uma distribuição exponencial, p. 71
- Figura 15 - Classes dos dados e da distribuição exponencial, p. 75
- Figura 16 - Aproximação para uma distribuição normal, p. 76
- Figura 17 - Classes dos dados e da distribuição normal, p. 79
- Figura 18 - Aproximação para uma distribuição beta, p. 80
- Figura 19 - Classes dos dados e da distribuição beta, p. 84
- Figura 20 - Função acumulada beta utilizando Matlab, p. 85
- Figura 21 - Aproximação inicial. Distr. Normal, p. 91
- Figura 22 - Aproximação com índice de consistência. Distr. Normal, p. 92
- Figura 23 - Qualidade da aproximação com dados alterados. Distr. Normal, p. 93
- Figura 24 - Qualidade da aproximação inicial. Distr. Exponencial, p. 94
- Figura 25 - Aproximação com índice de consistência. Distr. Exponencial, p. 95
- Figura 26 - Qualidade da aproximação com dados alterados. Distr. Exponencial, p. 97
- Figura 27 - Qualidade da aproximação inicial. Distr. Beta, p. 98
- Figura 28 - Aproximação com índice de consistência. Distr. Beta, p. 99
- Figura 29 - Qualidade da aproximação com dados alterados. Distr. Beta, p. 101

Figura 30 - Regiões de Fronteira - Distr. Normal - Sistema de Informações Original, p. 103

Figura 31 - Regiões de Fronteira - Distr. Normal - Sistema de Informações Transformado,
p. 105

Figura 32 - Regiões de Fronteira - Distr. Exponencial - Sistema de Informações Original,
p. 107

Figura 33 - Regiões de Fronteira - Distr. Exponencial - Sistema de Informações
Transformado, p. 109

Figura 34 - Regiões de Fronteira - Distr. Beta - Sistema de Informações Original, p. 111

Figura 35 - Regiões de Fronteira - Distr. Beta - Sistema de Informações Transformado, p. 113

Figura 36 - Qualidade da aproximação com dados alterados - probab. de maxim, p. 118

Figura 37 - Qualidade da aproximação com dados alterados - prob. de minim, p. 120

Figura 38 - Regiões de Fronteira – classes unidas pela probab. de maxim, p. 122

Figura 39 - Regiões de Fronteira - Classes unidas pela probab. de minim, p. 125

LISTA DE TABELAS

- Tabela 1 - Sistema de Informações Original, p. 24
- Tabela 2 - Matriz de Discernibilidade, p. 31
- Tabela 3 - Sistema de Informações Original, p. 41
- Tabela 4 - Conjuntos de Objetos Dominantes e Dominados, p. 43
- Tabela 5 - Índice de Consistência para os Objetos, p. 45
- Tabela 6 - Interseção de Dominância (Dominado) com as Classes, p. 54
- Tabela 7 - Interseção de Dominância (Dominado) com as Classes, p. 55
- Tabela 8 - Índice para Aproximação Inferior (ISVP-DRSA), p. 60
- Tabela 9 - Índice para Aproximação Inferior (ISVP-DRSA), p. 60
- Tabela 10 - Inconsistência entre dois atributos, p. 62
- Tabela 11 - Sistema de Informação Original, p. 65
- Tabela 12 - Conjuntos Dominantes e Dominados, p. 66
- Tabela 13 - Aproximação de dados discretos para contínuos - Distr. Exponencial, p. 72
- Tabela 14 - Probabilidades das classes geradas pela distribuição exponencial, p. 74
- Tabela 15 - Frequência dos dados originais (discretos) – Distr. Exponencial, p. 75
- Tabela 16 - Aproximação de dados discretos para contínuos - Distr. Normal, p. 77
- Tabela 17 - Probabilidades das classes geradas pela distribuição normal, p. 78
- Tabela 18 - Frequência dos dados originais (discretos). Distr. Normal, p. 79
- Tabela 19 - Aproximação de dados discretos para contínuos - Distr. Beta, p. 81
- Tabela 20 - Probabilidades das classes geradas pela distribuição beta, p. 83
- Tabela 21 - Frequência dos dados originais (discretos). Distr. Beta, p. 84
- Tabela 22 - Classes unidas por probabilidade de maxim, p. 117
- Tabela 23 - Classes unidas por probabilidade de minim, p. 119

LISTA DE EQUAÇÕES

- Equação 1 - Indiscernibilidade, p. 25
- Equação 2 - Aproximação Inferior, p. 25
- Equação 3 - Aproximação Superior, p. 26
- Equação 4 - Região de Fronteira, p. 26
- Equação 5 - Acurácia na Aproximação, p. 27
- Equação 6 - Qualidade da Aproximação, p. 28
- Equação 7 - Diferença entre Qualidades Calculadas, p. 32
- Equação 8 - Definição de P-Domínio, p. 35
- Equação 9 - Definição de Conjunto P-Dominante, p. 35
- Equação 10 - Definição de Conjunto P-Dominado, p. 35
- Equação 11 - Aproximação Inferior com P-Dominância, p. 36
- Equação 12 - Aproximação Superior com P-Dominância, p. 36
- Equação 13 - Regiões de Fronteira com Dominância, p. 37
- Equação 14 - Aproximação Inferior com VC-DRSA, p. 38
- Equação 15 - Aproximação Superior com VC-DRSA, p. 38
- Equação 16 - Qualidade da Aproximação com Dominância, p. 38
- Equação 17 - Aproximações Superior e Inferior em VC-DRSA, p. 43
- Equação 18 - Coeficiente de Consistência Robusta Para Classes Pelo Menos t , p. 47
- Equação 19 - Coeficiente de Consistência Robusta Para Classes Até t , p. 47
- Equação 20 - Aproximação Inferior com Coeficiente Robusto, p. 47
- Equação 21 - Coeficiente de Consistência Robusto Alterado, p. 48
- Equação 22 - Índice Beta em VP-DRSA, p. 51
- Equação 23 - Região P-Positiva em VP-DRSA, p. 52
- Equação 24 - Região Não-Negativa em VP-DRSA, p. 53
- Equação 25 - Aproximação Inferior em VP-DRSA Alterada, p. 58
- Equação 26 - Aproximação Superior em VP-DRSA Alterada, p. 58
- Equação 27 - Coeficiente Beta em ISVP-DRSA, p. 58
- Equação 28 - Aproximação Inferior em ISVP-DRSA, p. 59
- Equação 29 - Aproximação Superior em ISVP-DRSA, p. 59
- Equação 30 - Divisão de Intervalos - Variável Quantitativa, p. 63

Equação 31 - Limite para União de Classes Abaixo de s , p. 88

Equação 32 - Limites para União de Classes Acima de s , p. 88

Equação 33 - Limites para União de Classes Abaixo de s_1 e Acima de s_2 , p. 89

Equação 34 - Média e Variância do Universo, p. 115

LISTA DE ABREVIATURAS

DRSA	Dominance-based Rough Sets Approach
ISVP-DRSA	Inclusion degree and Supported degree for DRSA
RST	Rough Sets Theory
TCA	Teoria dos conjuntos aproximativos
TFS	Teoria de Fuzzy Sets
VC-DRSA	Variable Consistency Dominance-based Rough Sets Approach
VP- DRSA	Variable-Precision Dominance-based Rough Sets Approach

1 INTRODUÇÃO

1.1 CONSIDERAÇÕES INICIAIS

O Auxílio à Tomada de Decisão é uma área de pesquisa voltada para o desenvolvimento e a análise das ferramentas usadas por especialistas para que apoiem à decisão de uma forma consistente. Uma ferramenta de *Data Mining* busca extrair algum padrão dentro de um conjunto de dados (MARAKAS, 1998), aqui apropriadamente chamado de sistema de informações. Diversas teorias têm sido construídas para fundamentar o *Data Mining* no Auxílio à Tomada de Decisão e uma, que nesta tese se destaca, foi introduzida por Zdzislaw Pawlak, em 1982 e chamada de *Rough Sets Theory (RST)*. Em português, encontra-se em Gomes, Gomes e Almeida (2006) a tradução “Teoria dos Conjuntos Aproximativos (TCA)”. Neste trabalho, o termo TCA será usado.

Gomes, L., Gomes, C. e Almeida (2006) destacam que a importância da TCA é focada em:

- Avaliação da importância de um critério;
- Eliminação de redundâncias;
- Determinação de regras que diminuam uma tabela de decisão;
- Avaliação de conflitos de opiniões.

A TCA baseia-se na ideia de indiscernibilidade. Dois objetos são considerados indiscerníveis se, pelos critérios condicionantes adotados, eles não podem ser separados, ou seja, são idênticos pelos critérios estabelecidos. Porém, nem todos os objetos indiscerníveis

são classificados, no atributo de decisão, da mesma maneira, é desta ideia que surgem os conflitos que a TCA busca estudar.

Sassi (2010) e Pila (2001) incluem nas vantagens da Teoria de *Rough Sets* a ausência da necessidade de informações preliminares ou adicionais sobre os dados ou de estipular uma distribuição probabilística. De fato, a TCA utiliza, única e exclusivamente, a estrutura interna dos dados analisados para modelar o conhecimento (ZIARKO, 1994). Cabral (2005) resume as vantagens da TCA em sua utilização para diminuir as dificuldades na transformação automática de dados em conhecimento.

Pode-se nomear conflitos em relação à indiscernibilidade na Teoria de *Rough Sets* - TCA - (PAWLAK, 1991) como *roughness*. A caracterização da *roughness* na identificação das unidades observadas tem-se concentrado no exame do vetor de avaliações segundo cada variável, mas envolve também a precisão na medição de cada variável. Este último aspecto tem recebido pouca atenção, mas esta tese mostra como influencia nos índices obtidos pela TCA para mensurar o nível de qualidade nas regras de decisão criadas.

Uma etapa importante da aplicação da TCA é a identificação dos *reducts* (definido como reduções (COUTO; GOMES, 2010)), conjuntos mínimos de variáveis de condição suficientes para identificar os conjuntos de observações indiscerníveis pelas variáveis de decisão. Dependendo da *roughness* na medição de cada variável, o número de reduções será maior ou menor, assim como o número de variáveis em cada redução.

Evoluções da TCA são apresentadas por diversos autores em artigos especializados. Um desenvolvimento em particular será objeto de atenção aqui. Diz respeito às situações em que os atributos, incluindo os de decisão (que posteriormente serão definidos formalmente), possuem, em suas classificações uma ordenação natural como, por exemplo, a ordenação em uma escala. Essa escala pode ser classificada como uma escala de preferências ou escala de dominância (PAWLAK, 1991).

Em escalas de preferência, uma avaliação com uma escala menor está contida em uma avaliação com escala superior. Quando se estudam critérios com escalas de preferências, tem-se o que Greco *et al.* (2001) nomearam como “*Dominance-based Rough Sets Approach*” (DRSA). Os mesmos autores denominam o estudo de TCA sem escalas de preferências como abordagem clássica.

Em DRSA, o princípio da indiscernibilidade continua, porém com algumas alterações de definições, que serão apresentadas adiante. Os conflitos destas possíveis indiscernibilidades no sistema de informações original dificultam a criação de regras de suporte à tomada de decisão. Aqui entra o conceito de VC-DRSA (*Variable Consistency*

Dominance-based Rough Sets Approach), apresentado por Greco *et al.* (2001), que permite uma entrada relaxada de objetos para a criação de regras de decisão. Novas propostas de análise para DRSA serão temas de estudos neste trabalho, assim como novas propostas para união de atributos com o intuito de minimizar possíveis conflitos entre os objetos e suas classificações.

1.2 JUSTIFICATIVA

Um sistema de informações pode ter objetos que “degradam” a classificação de um conjunto devido a uma incerteza associada aos mesmos. Objetos com atributos iguais no seu conjunto de atributos dominantes devem ter a mesma classificação no atributo de decisão (PESSOA, 2004). Caso isso não ocorra, uma inconsistência no sistema de informações original é percebida. A TCA não distingue as inconsistências do seu sistema de informações pela possibilidade de uma classificação em um valor mais comum ou mais afastado do centro de interesse. Toda inconsistência tem o mesmo peso para o cálculo da qualidade da aproximação.

O foco principal deste estudo será com atributos com relação de dominância, classificados como DRSA por Greco *et al.* (2001). Estes autores, pensando em maneiras de corrigir inconsistências que prejudicam a criação de regras para tomada de decisão, propuseram novos critérios de consideração de inconsistências, chamado pelos autores de VC-DRSA.

Após a apresentação formal da técnica de VC-DRSA, novas propostas como, por exemplo, o refinamento de VC-DRSA, VP-DRSA (*variable precision dominance-based approach*) e ISVP-DRSA (*inclusion degree and supported degree for DRSA*) foram apresentadas para minimizar impactos negativos da aplicação direta da primeira proposta de VC-DRSA.

A quantidade de valores disponíveis para o atributo de decisão pode ser uma das causas da imprecisão. A análise da quantidade de valores para o atributo de decisão pode se assemelhar ao estudo do número de categorias da escala de Likert, onde o número de categorias pode levar a interpretações distintas e alterar o resultado da pesquisa (ALEXANDRE *et al.*, 2003).

O objeto central desta tese é uma nova proposta de inclusão de objetos do sistema de informações original nas aproximações inferiores, conceito que será detalhado em capítulos

posteriores. A quantidade de objetos nas aproximações inferiores faz com que mais regras de decisão possam ser retiradas do sistema original aumentando, assim, a quantidade de objetos bem descritos por regras de decisão. Mesmo que essas regras criadas não tenham 100% de acurácia, os relaxamentos de entrada podem fazer com que novas regras possam ser geradas.

1.3 FORMULAÇÃO DE HIPÓTESES

Propostas para lidar com as possíveis inconsistências, principalmente em DRSA, relevam pequenas inconsistências, quantificadas por um índice l . Greco *et al.* (2001) nomearam uma abordagem baseada nesse princípio como VC-DRSA (*Variable Consistency Dominance-based Rough Sets Approach*). O estudo de VC-DRSA será desenvolvido em capítulos posteriores. A ideia básica é classificar inconsistências em uma escala numérica de 0 até 1. Quanto mais próximo de 1, mais próximo da consistência está o atributo.

Então, um “valor de corte” é pré-estabelecido e “pequenas inconsistências” são desprezadas. O objetivo é aumentar o número de regras de apoio à decisão baseadas no sistema de informações original. Apesar de não alterar os dados contidos no sistema original, o uso de VC-DRSA relaxa o rigor da teoria clássica de indiscernibilidade e seus critérios de entrada de objetos na aproximação inferior.

Com a proposta de VC-DRSA apresentada, novos autores procuraram aperfeiçoá-la, como Blaszczynski *et al.* (2006), com novas propostas para relaxamento de entrada na aproximação inferior, Inuiguchi e Yoshioka (2006), com uma abordagem nomeada pelos autores de VP-DRSA (*variable-precision dominance-based approach*), e Deng, Wang e Hu (2011), com a inclusão de um índice β e uma técnica denominada ISVP-DRSA (*Inclusion degree and Supported degree for DRSA*). As propostas apresentadas não alteram o sistema de informações original, apenas sugerem novas entradas de objetos nas aproximações inferiores.

Uma hipótese a ser testada neste trabalho é a seguinte: é possível, através de um método quantitativo, unir valores da escala de preferência e com isso aumentar o número de regras de decisão baseadas no sistema de informação?

As novas escalas produzidas garantirão um aumento nas regras de decisão e, conseqüentemente, maior qualidade da aproximação?

Ao contrário da VC-DRSA (e, por consequência VP-DRSA e ISVP-DRSA), onde os dados não são alterados, mas há um relaxamento na inconsistência, neste trabalho se propõe uma alteração nos dados originais com um aumento da *roughness* das medidas.

A proposição deste novo método de tratar inconsistências é baseada em uma investigação de possibilidades de diminuir a quantidade de valores disponíveis do atributo de decisão para que a qualidade da aproximação segundo a TCA seja superior à do sistema original. Serão propostos dois métodos: um baseado em distribuições de probabilidades que crie um algoritmo que permita a criação de regras para unir valores de escalas de preferência e outro com transformações probabilísticas baseadas na probabilidade de cada valor tornar-se o maior (ou menor) em uma replicação da medição do atributo de decisão.

A primeira proposta envolve encontrar uma distribuição de probabilidades contínua que se adeque aos valores do atributo de decisão. Assume-se que o atributo de decisão tem relação de dominância, isto é as classes de valores são ordenadas de modo que cada classe t é superior à classe $t-1$ e inferior à classe $t+1$. A ideia básica é encontrar classes de valores pouco frequentes e alocá-las a uma classe imediatamente superior ou inferior, diminuindo assim os valores disponíveis para o atributo de decisão e, conseqüentemente, propiciando o aumento da qualidade da aproximação da TCA.

A segunda proposta envolve atribuir um erro de medida do atributo de decisão para cada objeto. Levando em conta a distribuição de probabilidades deste erro, se transformam os valores discretos do atributo de decisão em distribuições de probabilidades. Como, agora, os valores são funções de probabilidade, é possível calcular a probabilidade de determinado valor ser o maior (ou o menor) em uma replicação qualquer.

A combinação de TCA com transformações probabilísticas já foi realizada em Sant'Anna (2004). É difícil definir qual a precisão adequada para medir cada variável, visto que suas escalas naturais podem ser muito diferentes. No caso de variáveis orientadas (dominância), a transformação probabilística torna possível reduzir todas as variáveis a uma mesma escala natural, qual seja a escala da probabilidade de atingir uma posição extrema, isto é, de a unidade de observação apresentar o maior ou o menor valor em replicações aleatórias da mesma medição.

Assim como Greco *et al.* (2001) desenvolveram a VC-DRSA relaxando condições estreitas para a aproximação inferior dos *rough sets* de modo que mais objetos possam pertencer a esta classe e, conseqüentemente, aumente a qualidade da aproximação toma-se como referência a DRSA para desenvolver e avaliar as alternativas exploradas nesta tese. A diferença é que, em vez de relaxar os critérios de indiscernibilidade, faz-se a união dos valores em atributos com dominância.

1.4 ESTRUTURA DO TRABALHO

O trabalho tem a seguinte estrutura:

- No capítulo 1 o problema é apresentado, assim como suas justificativas e hipóteses a serem explicadas ao longo da tese;
- No capítulo 2, é apresentada a TCA. A proposta aqui é mostrar detalhadamente como a TCA pode auxiliar a criação de regras e a tomada de decisão usando um sistema de informações e aplicando, ao mesmo, fundamentos de TCA;
- No capítulo 3, a técnica DRSA é apresentada, assim como a proposta de VC-DRSA que é usada, como será detalhado, para aumentar *o número de* regras de decisão no sistema de informações. Esse aumento pode ser creditado às “violações controladas” na definição clássica de inserção de objetos nas aproximações inferiores;
- No capítulo 4, as novas propostas de alteração do uso de DRSA são apresentadas e discutidas;
- No capítulo 5 é apresentada a ideia da redução de valores para o atributo de decisão através da análise da frequência dos valores. Duas propostas são apresentadas: pela distribuição contínua que mais se adequar aos dados do atributo de decisão e pela transformação probabilística no atributo de decisão na qual é calculada a probabilidade de o próprio valor do atributo de decisão do objeto ser o valor mais alto em uma replicação qualquer da avaliação;
- No capítulo 6 serão apresentados exemplos da mudança na qualidade da aproximação com o uso reduzido de valores para o atributo de decisão e possíveis mudanças nas reduções, além de comparação com os resultados obtidos com VC-DRSA;
- O capítulo 7 apresenta a conclusão.

2 TEORIA DOS CONJUNTOS APROXIMATIVOS

A TCA foi proposta por Pawlak (1982) e busca extrair de um banco de dados os atributos fundamentais para que se possa orientar a regras de decisão. Pawlak e Skowron (2007) destacam que a vantagem principal da TCA é a não necessidade de informações preliminares ou adicionais aos dados.

A TCA é aplicada em muitos casos reais, tais como: engenharia, medicina, análise de mercado, análise financeira e outros. Polkowski (1998) relata estudos de TCA em diversas áreas do conhecimento humano como, por exemplo: aprendizado de regras de decisão; mineração de dados; sistemas especialistas de suporte à decisão; agrupamento (*clustering*), etc. Pereira, Gómez e López (2008) apresentam um estudo com sucesso de aplicação da Teoria de *Rough Sets* na previsão de fracasso empresarial. Sant'Anna, A. e Sant'Anna, P. (2008) aplicam a teoria para localização de usinas hidroelétricas. Couto e Gomes (2010) utilizam TCA para estudar dados de recursos humanos em uma instituição financeira.

Pila (2001) compara a Teoria de *Fuzzy Sets* – Teoria dos Conjuntos Nebulosos (ZADEH, 1965) e TCA destacando que a Teoria dos Conjuntos Nebulosos trata da incerteza de informações e que a TCA trata da imprecisão e ambiguidade de dados. TCA e Teoria dos Conjuntos Nebulosos são duas abordagens independentes para o tratamento de um conhecimento imperfeito onde ambiguidades encontradas impedem um conhecimento perfeito e, conseqüentemente, uma regra perfeita para determinados objetos pertencentes ao sistema de informações original.

O início da aplicação da TCA dá-se após a obtenção dos dados iniciais. Estes dados, tabulados apropriadamente, são chamados de sistema de informação e devem conter: nas linhas os objetos individualizados do sistema original e nas colunas os diversos atributos condicionantes e o atributo ou atributos de decisão que deve ou devem ser explicados pelos atributos condicionantes para cada objeto do sistema de informações. Os objetos do sistema

de informação original podem ser agrupados. Esse agrupamento representa os possíveis objetos indiscerníveis quanto aos atributos condicionantes, ou seja, se dois objetos distintos possuem os mesmos valores para todos os atributos condicionantes, eles são da mesma classe e são considerados indiscerníveis. O que gera inconsistência em TCA é justamente a possibilidade destes objetos classificados como indiscerníveis possuírem classificações diferentes no atributo de decisão.

O sistema de informação apresentado na tabela 1, com informações para a identificação de um conjunto de dez animais, será utilizado como base para o entendimento dos fundamentos de *Rough Sets*. Chama-se de U o conjunto de objetos (animais), C os atributos condicionais e D o atributo de decisão. O sistema de informação é denotado por $\mathfrak{R} = (U, C \cup D)$.

A tabela 1 mostra o sistema de informação \mathfrak{R} .

U	atributos condicionais (c)			Decisão (d)
Animais	porte	pelos	habitat predom.	Ferocidade
1	pequeno	sim	Terrestre	Baixa
2	médio	não	Aquático	Média
3	grande	não	Terrestre	Alta
4	médio	sim	Aquático	Baixa
5	grande	não	Terrestre	Baixa
6	médio	sim	Aquático	Média
7	Médio	não	Terrestre	Média
8	Grande	sim	Terrestre	Alta
9	Médio	não	Aquático	Baixa
10	Pequeno	sim	Aquático	Média

Tabela 1 - Sistema de Informações Original

Fonte: autor

Neste sistema de informações, baseado em três atributos condicionais (porte, pelos e habitat predominante) estabelece-se a decisão quanto à ferocidade do animal.

Resumidamente, é possível notar que os atributos caracterizam os objetos com atribuições de valores a cada atributo presente no sistema de informações original. O atributo de decisão é estudado levando em consideração os valores dos atributos condicionantes.

2.1 INDISCERNIBILIDADE

Hein e Kroenke (2010) definem que indiscernir significa não conseguir distinguir uma coisa de outra por meio dos sentidos ou da inteligência humana. É a partir do conceito de

indiscernibilidade que são geradas as bases matemáticas da TCA. Os objetos do sistema classificados como indiscerníveis são grânulos de conhecimento e, a princípio, não devem se distinguir em nenhum atributo, nem mesmo os atributos de decisão.

Para qualquer subconjunto B de atributos condicionais $B \subseteq C$, se os objetos x e y em U não podem ser diferenciados, ou seja, são idênticos, em todos os atributos condicionais contidos no subconjunto de atributos B , então x e y são ditos indiscerníveis segundo B . A relação de indiscernibilidade classifica objetos com as mesmas características segundo Couto e Gomes (2010). Cabral (2005) afirma que a relação de indiscernibilidade está fortemente associada aos valores dos atributos que compõem o banco de dados original e que este banco será redefinido para melhor representar os dados após estabelecidas as relações de indiscernibilidade.

Denota-se por $c(x)$ o conjunto dos objetos em U com a mesma avaliação objeto x segundo o atributo c e por $B(x)$ o conjunto dos objetos em U com a mesma avaliação $c(x)$ segundo todos os atributos c em B .

Pode-se definir a relação de indiscernibilidade através da equação 1:

$$IND(B) = \{(x, y) \in U / \forall c \in B, c(x) = c(y)\}$$

Equação 1 - Indiscernibilidade

O conjunto de classes equivalentes determinadas por $IND(B)$ é representado por $U / IND(B)$. Então, por exemplo, para o sistema de informações mostrado na tabela 1, para alguns possíveis valores de B tem-se (utilizando a equação 1):

$$IND(\{porte\}) = \{(1,10), (2,4,6,7,9), (5,8)\}$$

$$IND(\{pelos, habitat\}) = \{(1,8), (2,9), (3,5,7), (4,6,10)\}$$

$$IND(\{porte, pelos, habitat\}) = \{(1), (2,9), (3,5), (4,6), (7), (8), (10)\}$$

Objetos indiscerníveis são tratados pela TCA como “grãos”. A união destes grãos ou conjuntos elementares constituirá o conceito de conhecimento sobre algo de interesse. Possíveis inconsistências em objetos indiscerníveis, ou seja, atributos de decisão diferentes

para objetos indiscerníveis incluem as inconsistências e, também, ações distintas de especialistas sobre os objetos.

O uso das aproximações inferior e superior permite classificar os objetos indiscerníveis de acordo com os possíveis atributos de decisão. Para todo X , subconjunto de U , são definidas duas aproximações: inferior, $\underline{P}(X)$, e superior, $\overline{P}(X)$, definidas como:

$$\underline{P}(X) = \{x \in U \mid P(x) \subseteq X\}$$

Equação 2 - Aproximação Inferior

$$\overline{P}(X) = \{X \in U \mid P(x) \cap X \neq \emptyset\}$$

Equação 3 - Aproximação Superior

Couto e Gomes (2010) definem que aproximação inferior é a união de todos os objetos elementares que estão inclusos em uma classe. A aproximação superior é a união de todos os conceitos elementares que têm alguma interseção (não vazia) com o objeto em questão.

Assim, os objetos incluídos na aproximação inferior são necessariamente membros de X . Os objetos classificados na aproximação superior são possíveis membros de X . Uma região de fronteira pode ser definida pela equação 4:

$$RF(X) = \overline{P}(X) - \underline{P}(X)$$

Equação 4 - Região de Fronteira

Ou seja, a diferença entre as aproximações superiores e inferiores (é importante notar que, por definição, $\underline{P}(X) \subseteq \overline{P}(X)$).

Pawlak (1982) define um conjunto X como áspero (*rough*) se a sua região de fronteira for não vazia e de preciso (*crisp*) se a sua região de fronteira for um conjunto vazio. Segundo Pawlak (2000) se a região de fronteira for não vazia (aproximações superior e inferior distintas) é preciso aplicar o conceito “aproximativo”.

Quando o atributo de decisão tem a característica de dominância, as classes inferior e superior são definidas considerando a ordem estabelecida por esse atributo. Basta utilizar, para estabelecer estas definições, apenas os conjuntos X cujos objetos são os objetos de U com valor para esse atributo acima de um certo patamar. Por exemplo, para o atributo de decisão ‘ferocidade’, um conjunto a ser considerado será $X = \{3, 8\}$, o conjunto dos animais

considerados de ferocidade alta. Seja B o conjunto dos três atributos condicionais {porte, pelos, habitat predominante}, que como demonstrado, determina a partição $U|IND(B) = \{\{1\}, \{2,9\}, \{3,5\}, \{4,6\}, \{7\}, \{8\}, \{10\}\}$. As aproximações de X serão, então:

Aproximação inferior (equação 2): $\underline{P}(X) = \{8\}$ (regra de decisão: “alta”)

Aproximação superior (equação 3): $\overline{P}(X) = \{\{3,5\}, \{8\}\}$

Região de fronteira: $RF(X) = \{3,5\}$ (regra de decisão: “alta/média”)

Fora da região: $U - \overline{P}(X) = \{\{1\}, \{2,9\}, \{4,6\}, \{7\}, \{10\}\}$ (regra de decisão: “média/baixa”)

A figura 1 representa esquematicamente as aproximações:

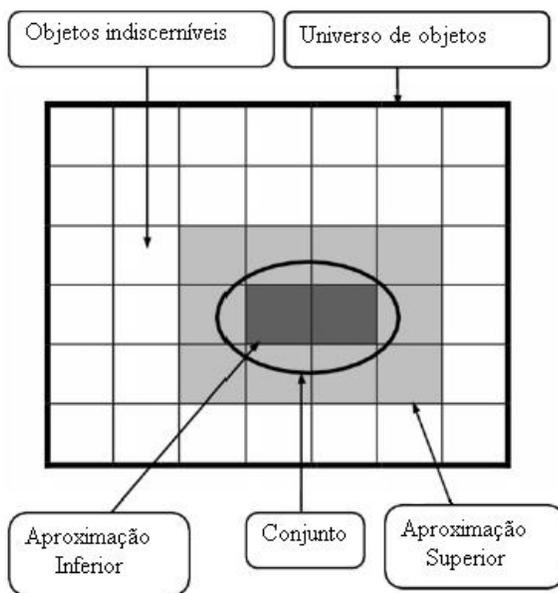


Figura 1 - Aproximações em TCA
Fonte: Pawlak e Skowron (2006).

As aproximações podem ser avaliadas numericamente. Associa-se o símbolo α à avaliação da acurácia na aproximação de conjuntos. Em relação às aproximações inferior e superior, Pawlak e Skowron (2007) define o coeficiente de imprecisão pela equação 5:

$$\alpha_B(X) = \frac{|\underline{P}(X)|}{|\overline{P}(X)|}$$

Equação 5 - Acurácia na Aproximação

O símbolo $||$ representa a cardinalidade (quantidade de objetos em cada um dos conjuntos representados). Como observado anteriormente, a aproximação inferior é um subconjunto da aproximação superior. Portanto, se o coeficiente de imprecisão for igual a 1 é porque as duas aproximações são iguais, não gerando região de fronteira. Neste caso, Pawlak (1982) define que X é preciso (*crisp*). Caso contrário, o coeficiente é menor do que 1 e X será impreciso (*rough*). O mesmo autor define que:

- Se $\underline{P}(X) \neq 0$ e $P(X) \neq U$ $\Rightarrow X$ é imprecisamente B-definível (*roughly B-definable*);
- Se $\underline{P}(X) = 0$ e $\overline{P}(X) \neq U$ $\Rightarrow X$ é internamente B-indefinível (*internally B-indefinable*);
- Se $\underline{P}(X) \neq 0$ e $\overline{P}(X) = U$ $\Rightarrow X$ é externamente B-indefinível (*externally B-indefinable*);
- Se $\underline{P}(X) = 0$ e $\overline{P}(X) = U$ $\Rightarrow X$ é totalmente B-indefinível (*totally B-indefinable*);

No exemplo dado, o coeficiente de imprecisão (equação 5) será dado por:

$$\alpha_B(X) = \frac{|P(X)|}{|\overline{P}(X)|} = \frac{1}{3}$$

Pawlak e Skowron (2007) definem também o índice γ para medir a qualidade da aproximação de uma classificação por outra. Considere-se uma partição $\mathfrak{S} = \{X_1, \dots, X_n\}$. Na prática, esta será a partição determinada pelo atributo de decisão. O índice γ é calculado dividindo a soma das cardinalidades das aproximações inferiores dos elementos da partição pelo número total de objetos em U . Isto é:

$$\gamma_B(\mathfrak{S}) = \sum_i \frac{|B(X_i)|}{|U|}$$

Equação 6 - Qualidade da Aproximação

Por exemplo, o índice de qualidade da aproximação da partição IND (ferocidade) = $\{\{1,4,5,9\}, \{2,6,7,10\}, \{3,8\}\}$ pelo conjunto de atributos B será (equação 6) :

$$\gamma_B(\mathcal{N}) = \sum_i \frac{|B(X_i)|}{|U|} = \frac{4}{10}.$$

O valor 4 do numerador representa a soma das cardinalidades das aproximações inferiores para os valores do atributo de decisão “baixa” (cardinalidade 1 – objeto (1)), “média” (cardinalidade 2 – objetos (7, 10)) e “alta” (cardinalidade 1 – objeto (8)). Como o sistema possui no total 10 objetos, o denominador é igual a 10.

A interpretação do índice γ encontrado acima (igual a 0,40) é que 40% dos objetos do sistema de informações geram regras bem definidas para o apoio à decisão.

2.2 REDUÇÃO DE ATRIBUTOS

Qualquer sistema de informações usado para explicar aproximadamente uma classificação que tenha a quantidade de atributos reduzida e preserve a qualidade da aproximação é mais adequado do que o sistema de informações original. No caso da redução, os atributos que saem do sistema de informações são considerados dispensáveis se, com a retirada desses atributos, as aproximações inferiores e superiores são as mesmas de antes (PAWLAK; SKOWRON, 2007).

Os subconjuntos de atributos que mantêm a classificação original são chamados de redução (*reducts*). Formalmente, um conjunto de atributos P é redução de A se $P \subseteq A$ mantém todas as relações de indiscernibilidade de A . Ou seja, se um conjunto P , com menos atributos que o conjunto A , mantém todas as relações de indiscernibilidade do conjunto A e possui uma cardinalidade (quantidade de atributos no seu conjunto) menor do que A é porque o conjunto original A possui atributos redundantes.

Pode haver mais de uma redução no sistema de informações original. Os atributos que aparecem em todas as reduções, ou seja, estão na interseção de todas as reduções, representam o núcleo (*core*).

Conforme destaca Cabral (2005), menos atributos fornecem uma forma mais simples para representar o mesmo conhecimento. Essa economia na quantidade de dados trabalhados facilita o auxílio à tomada de decisão, pois menos atributos são necessários na geração de regras que auxiliem à tomada de decisão.

A maneira de encontrar reduções e núcleos passa pela função de discernibilidade.

2.3 FUNÇÃO DE DISCERNIBILIDADE

A função de discernibilidade é uma função booleana que auxilia na obtenção de reduções e núcleos na TCA. Esta função é obtida através de uma matriz de discernibilidade que tem as seguintes características:

- É uma matriz quadrada $n \times n$, onde n representa o número de classes vindo de $n = |U|IND_A(B)$. As células desta matriz representam os atributos nos quais as classes possuem valores diferentes. Obviamente, a diagonal principal desta matriz será vazia, pois uma classe é sempre semelhante a ela mesma.
- $(c_{ij}) = \{a \in B : a(x_i) \neq a(x_j)\}$ para $i < j = 1, 2, 3, \dots, n$
- Os atributos que constituem o núcleo são os elementos únicos que constituem qualquer célula $(c_{ij}) \forall i < j$.
- Para constituir reduções, basta verificar para todas as células $i < j$ da matriz de discernibilidade quais atributos, além dos que pertencem ao núcleo, necessitam ser adicionados para que nenhuma célula desta matriz tenha interseção igual a zero.

Define-se como B' o conjunto de atributos mínimos que satisfaça a relação: $B' \cap c(i, j) \neq \emptyset \forall i < j$. Fica simples notar que os atributos únicos nas células da matriz de discernibilidade devem entrar no conjunto B' .

Para o sistema de informações apresentado na tabela 1, existem 7 classes indiscerníveis (considerando o conjunto com todos os atributos condicionantes). Sejam Porte, Pelos, Habitat predominante denotados como Po , Pe e H respectivamente. As classes indiscerníveis são (utilizando a equação 1):

$$U|IND_A(B) = \{\{1\}, \{2,9\}, \{3,5\}, \{4,6\}, \{7\}, \{8\}, \{10\}\}, c_1, c_2, c_3, c_4, c_5, c_6 \text{ e } c_7$$

respectivamente. A matriz de discernibilidade (7x7) é apresentada na tabela 2:

	c1	c2	c3	c4	c5	c6	c7
c1	-						
c2	Po,Pe,H	-					
c3	Po,Pe	Po,H	-				
c4	Po,H	Pe	Po,Pe,H	-			
c5	Po,Pe	H	Po	Pe,H	-		
c6	Po	Po,Pe,H	Pe	Po,H	Po,Pe	-	
c7	H	Po,Pe	Po,Pe,H	Po,H	Po,Pe,H	Po,H	-

Tabela 2 - Matriz de Discernibilidade

Fonte: autor.

Como os três atributos são encontrados isoladamente em pelo menos uma célula, todos os três fazem parte do núcleo e, portanto, não há um conjunto $B' \subset B$ que satisfaça as regras de decisão.

2.4 SIGNIFICÂNCIA DE ATRIBUTOS

Pawlak (1991) define a importância dos atributos empregando o índice γ de qualidade da aproximação. Qualquer retirada de atributo(s) que pertença a uma redução pode prejudicar os acertos dos atributos restantes em relação ao atributo condicional. Define-se:

- C subconjunto de atributos condicionais
 - D atributo de decisão
 - a atributo pertencente ao subconjunto de atributos condicionais
 - $\gamma(C, D) = \text{qualidade da aproximação de } U/D \text{ por } C$
 - $\gamma(C - \{a\}, D) = \text{qualidade da aproximação de } U/D \text{ por } C - \{a\}$
- o No exemplo abordado:
- C subconjunto de atributos condicionais {Po,Pe, H}
 - D atributo de decisão {ferocidade}
 - a atributo pertencente ao subconjunto de atributos condicionais {Po ou Pe ou H}
 - $\gamma(C, D) = \text{qualidade da aproximação de } U/D \text{ em } C = 4/10$
 - $\gamma(C - \{a\}, D) = \text{qualidade da aproximação de } U/D \text{ em } C - \{a\}$

- $a = Po \Rightarrow \gamma(C - \{Po\}, D) = 0$
- $a = Pe \Rightarrow \gamma(C - \{Pe\}, D) = 3/10$
- $a = H \Rightarrow \gamma(C - \{H\}, D) = 1/10$

A diferença entre as qualidades calculadas com e sem um determinado atributo a é dada por:

$$\delta_{(C,D)}(a) = \delta(a) = \frac{(\gamma(C,D) - \gamma(C - \{a\}, D))}{\gamma(C,D)} = 1 - \frac{\gamma(C - \{a\}, D)}{\gamma(C,D)}$$

Equação 7 - Diferença entre Qualidades Calculadas

Portanto para o exemplo, tem-se (utilizando a equação 7):

$$\delta_{(C,D)}(Po) = 1 - \frac{0}{4/10} = 1; \delta_{(C,D)}(Pe) = 1 - \frac{3/10}{4/10} = 0,25 \text{ e } \delta_{(C,D)}(H) = 1 - \frac{1/10}{4/10} = 0,75.$$

Quanto mais próximo de 1 for a precisão de aproximação, mais importante é o atributo condicional para a geração de regras. No exemplo anterior, como a redução é formada pelos três atributos condicionais, a retirada de qualquer um deles vai ocasionar em perdas na regra de decisão. Note-se que o atributo mais importante é o “Porte” porque, sem esse atributo, nenhuma regra de decisão seria possível. Em segundo, em termos de importância, vem o atributo Habitat Predominante (3 das 4 regras antes satisfeitas ficariam perdidas) e, por último, o atributo condicional Pelos (1 das 4 regras satisfeitas ficaria perdida).

2.5 EXEMPLOS DE USO DE CONJUNTOS APROXIMATIVOS COM PRINCÍPIOS DE DOMINÂNCIA

Liou *et al.* (2011) mostra o uso de conjuntos aproximativos com atributos com relação de dominância e sem relação de dominância para o gerenciamento de um aeroporto internacional. Os atributos utilizados pelos autores são listados na figura 2:

Especificações de Atributos/Critérios para qualidade de serviço em aeroportos

Attribute/criterion name	Attribute/criterion value	Value set	Preference
<i>Condition attributes</i>			
Gender	Female; male	{F, M}	None
Age	Below 40; above 40	{1, 2}	None
Occupation	Government employee; company employee; student; others	{G, P, S, O}	None
Education	High school or below; college or above	{H, C}	None
Income	NT\$ 30,000 or less; 30,001–60,000; 60,000 and above	{1, 2, 3}	None
User frequency	Airport usage 1–2; 3–4; 5 or more (times/in last 12 months)	{1, 2, 3}	None
Convenience	Poor; medium; satisfactory	{1, 2, 3}	Gain
Comfortable	Poor; medium; satisfactory	{1, 2, 3}	Gain
ICQ	Poor; medium; satisfactory	{1, 2, 3}	Gain
Transportation	Poor; medium; satisfactory	{1, 2, 3}	Gain
Courtesy	Poor; medium; satisfactory	{1, 2, 3}	Gain
Information	Poor; medium; satisfactory	{1, 2, 3}	Gain
Security	Poor; medium; satisfactory	{1, 2, 3}	Gain
Price of shops	Low; medium; high	{1, 2, 3}	Cost
<i>Decision criterion</i>			
Overall satisfaction	Poor; medium; satisfactory	{1, 2, 3}	Gain

Figura 2 - Atributos com dominância

Fonte: Liout et al. (2011).

Como a preferência (quarta coluna) indica, alguns critérios não possuem relação de dominância (*preference = none*) enquanto os demais possuem. O uso de atributos com ou sem ordenação de preferência não impede o uso da TCA. O artigo mostra o quanto a TCA pode auxiliar a tomada de decisões com exemplos reais e enfatiza que o uso da relação de dominância tem maiores aplicações no cotidiano do que as relações baseadas na teoria clássica da relação de indiscernibilidade.

Liou e Tzeng (2010) fazem uma análise sobre o tema: mercado aéreo. Usando relações de dominância, os autores propõem a utilidade de *Rough Sets* com relação de dominância (*DRSA- Dominance-based Rough Sets Approach*) demonstrando que a teoria auxilia a identificar clientes, determinar características dos clientes e facilitar o desenvolvimento de estratégias de marketing. Os autores ainda comparam o DRSA com a técnica de análise discriminante mostrando que DRSA tem uma melhor predição do que a análise discriminante.

Zhu, Cai e Lu (2008) utilizam a técnicas de conjuntos aproximativos com relação de dominância para o estudo da influência de fatores no mercado imobiliário de Wuhan (capital da província de Hubei).

3 DRSA

3.1 INTRODUÇÃO

O uso de critérios baseados em preferências tem uma participação importante nos estudos de TCA. Quando há esta relação de preferência, ao invés do uso da indiscernibilidade como apresentada na seção anterior, as aproximações levam em conta as relações de dominância estabelecidas pelos critérios.

Suponha um atributo de decisão D. Usando este atributo, divide-se o universo U de objetos em n classes possíveis (n representando a quantidade de valores distintos que o atributo de decisão apresenta. Se D apresenta n valores, ordenáveis por preferência de 1 a n, as classes podem ser definidas por:

$$Cl_1 = \{x_i \in U \mid D_i = 1\} \quad \forall i = 1, \dots, n$$

$$Cl_2 = \{x_i \in U \mid D_i = 2\} \quad \forall i = 1, \dots, n$$

...

$$Cl_n = \{x_i \in U \mid D_i = n\} \quad \forall i = 1, \dots, n$$

Portanto, $\cup Cl_t = U$ e todos os elementos do sistema de informação pertencem a uma e apenas uma classe. Como as classes respeitam uma preferência, pode-se afirmar que se $r > s$, os objetos pertencentes a Cl_r são preferíveis em relação aos objetos de Cl_s . Assim torna-se conveniente trabalhar com as classes acumuladas:

$$Cl_t^{\geq} = \cup_{s \geq t} Cl_s \quad ; \quad Cl_t^{\leq} = \cup_{s \leq t} Cl_s \quad ; \quad t = 1, \dots, n.$$

A leitura para $x \in Cl_t^{\geq}$ é que o objeto x pertence à classe Cl_t ou é preferível aos elementos desta classe. Enquanto que $x \in Cl_t^{\leq}$ representa que x pertence a, no máximo, a classe Cl_t . Obviamente, a união de todas as classes Cl_t é igual a U . Assim como:

$$Cl_{t-1}^{\leq} = U - Cl_t^{\geq} \text{ e } Cl_t^{\geq} = U - Cl_{t-1}^{\leq}.$$

Sejam x e y dois objetos de U . Afirma-se que x domina y ($x D_p y$), em relação a um conjunto de atributos P , se, segundo cada um destes atributos, a preferência por x for maior ou igual á preferência por y . Supondo a preferência por y segundo o atributo p representada por uma função $\rho(y, p)$.

$$x D_p y \Leftrightarrow \rho(x, p) \geq_p \rho(y, p) \quad \forall p \in P$$

Equação 8 - Definição de P-Domínio

Obviamente o próprio objeto sempre será dominado e dominará a si mesmo. Para relações de dominância, pode-se definir:

Um conjunto de objetos dominando x (conjunto P-dominante):

$$D_p^+(x) = \{y \in U : y D_p x\}$$

Equação 9 - Definição de Conjunto P-Dominante

Um conjunto de objetos dominados por x (conjunto P-dominado):

$$D_p^-(x) = \{y \in U : x D_p y\}$$

Equação 10 Definição de Conjunto P-Dominado

Uma vez definidas as classes e todas as relações de conjunto de objetos P-dominantes e P-dominados é possível definir as aproximações superior e inferior. Em relação à definição da seção anterior, a diferença é que, agora, são usadas relações de dominância ao invés de indiscernibilidade.

As aproximações inferior e superior podem ser definidas, respectivamente, por:

$$\underline{P}(Cl_i^{\geq}) = \{x \in U : D_p^+(x) \subseteq Cl_i^{\geq}\}$$

Equação 11 - Aproximação Inferior com P-Dominância

$$\overline{P}(Cl_i^{\geq}) = \{x \in U : D_p^-(x) \cap Cl_i^{\geq} \neq \emptyset\}$$

Equação 12 - Aproximação Superior com P-Dominância

Analogamente pode-se usar as equações 11 e 12 para definir as aproximações para classes no máximo até a classe t:

$$\underline{P}(Cl_i^{\leq}) = \{x \in U : D_p^-(x) \subseteq Cl_i^{\leq}\} \text{ e } \overline{P}(Cl_i^{\leq}) = \{x \in U : D_p^+(x) \cap Cl_i^{\leq} \neq \emptyset\}$$

Uma análise das equações mostradas acima permite observar que:

- Para que um objeto x_i pertença a $\underline{P}(Cl_i^{\geq})$ - os objetos que dominam x_i devem estar contidos na classe Cl_i^{\geq} .
- Para que um objeto x_i pertença a $\overline{P}(Cl_i^{\geq})$ - x_i deve dominar pelo menos um objeto de Cl_i^{\geq} , de modo que a aproximação superior conterà a própria classe Cl_i^{\geq} .
- Para que um objeto x_i pertença a $\underline{P}(Cl_i^{\leq})$ - os objetos dominados por x_i devem estar contidos em Cl_i^{\leq} ;
- Para que um objeto x_i pertença a $\overline{P}(Cl_i^{\leq})$ - pelo menos um objeto que domina x_i deve pertencer a Cl_i^{\leq} , de modo que a aproximação superior conterà a própria classe Cl_i^{\leq} .

Portanto, as relações abaixo podem ser verificadas:

$$\underline{P}(Cl_i^{\geq}) \subseteq Cl_i^{\geq} \subseteq \overline{P}(Cl_i^{\geq}) \quad , \quad \underline{P}(Cl_i^{\leq}) \subseteq Cl_i^{\leq} \subseteq \overline{P}(Cl_i^{\leq})$$

A região de fronteira (RF) é definida, como antes (equação 4), como a diferença entre as aproximações superior e inferior para a classe em questão.

$$RF(Cl_t^{\geq}) = \bar{P}(Cl_t^{\geq}) - \underline{P}(Cl_t^{\geq}) \quad , \quad RF(Cl_t^{\leq}) = \bar{P}(Cl_t^{\leq}) - \underline{P}(Cl_t^{\leq})$$

Equação 13 - Regiões de Fronteira com Dominância

O princípio da dominância afirma que objeto com classificações pelo menos iguais a outro objeto, em um determinado subconjunto de atributos condicionantes, deve ter, no seu atributo de decisão, classificação pelo menos igual ao mesmo objeto. Em outras palavras, um objeto r que P -domina (segundo os atributos do subconjunto de atributos P) outro objeto s , não pode pertencer uma classe Cl inferior à do objeto s . Caso isso ocorra, define-se os objetos como sendo inconsistentes quanto ao princípio da dominância.

Dembczynski *et al.* (2011) definem, na figura 3, visualmente a inconsistência do princípio a dominância:

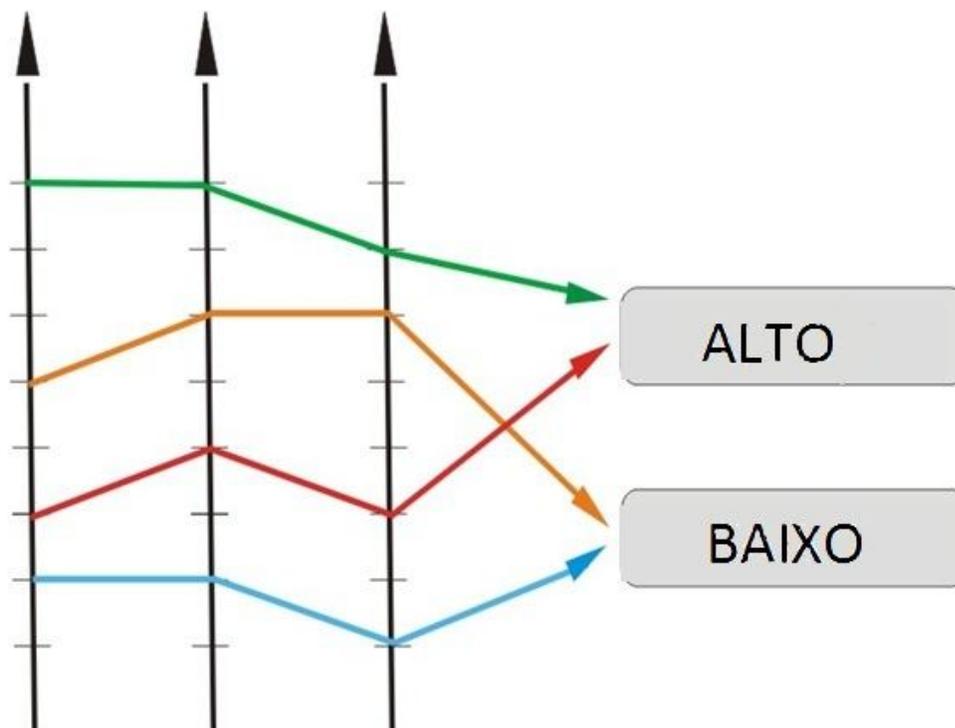


Figura 3 - Inconsistência de atributos
Fonte: Adaptado de Dembczynski *et al.* (2011).

3.2 VC-DRSA

O uso de VC-DRSA permite um relaxamento no critério de entrada da aproximação inferior, controlado por um nível, chamado de nível de consistência l , que pode variar entre

zero e um. Se o nível for igual a 1, o critério para a entrada será o mesmo da definição anterior, caso seja menor do que um, mais alguns objetos podem fazer parte da aproximação inferior.

Por VC-DRSA, para um objeto x pertencer à aproximação inferior, deve-se ter:

$$\underline{P}^l(Cl_i^{\geq}) = \left\{ x \in Cl_i^{\geq} : \frac{|D_p^+(x) \cap Cl_i^{\geq}|}{|D_p^+|} \geq l \right\}, \underline{P}^l(Cl_i^{\leq}) = \left\{ x \in Cl_i^{\leq} : \frac{|D_p^-(x) \cap Cl_i^{\leq}|}{|D_p^-(x)|} \geq l \right\}$$

Equação 14 - Aproximação Inferior com VC-DRSA

onde $| \cdot |$ representa a cardinalidade do conjunto.

O relaxamento aparece exatamente no nível de consistência l . Uma análise das aproximações leva a aceitar que a cardinalidade do numerador será igual a do denominador, ou seja, l igual a 1, quando $D_p^+(x)$ estiver contido em Cl_i^{\geq} (ou $D_p^-(x)$ estiver contido em Cl_i^{\leq}), o que remete à definição de aproximações (inferior e superior) com relações de dominância (DRSA).

Portanto, em VC-DRSA a aproximação inferior será sempre maior ou igual ao conjunto formado pela definição sem o uso do nível de consistência.

A aproximação superior pode ser escrita como:

$$\overline{P}^l(Cl_i^{\geq}) = U - \underline{P}^l(Cl_{i-1}^{\leq}), \overline{P}^l(Cl_i^{\leq}) = U - \underline{P}^l(Cl_{i+1}^{\geq})$$

Equação 15 - Aproximação Superior com VC-DRSA

As regiões de fronteira (RF) são definidas da mesma forma que a definição anterior, ou seja, a diferença entre a aproximação superior e inferior.

A qualidade da aproximação em VC-DRSA pode ser definida por:

$$\gamma_p(Cl) = \frac{|U - (\cup RF(Cl_i^{\leq}))|}{|U|} = \frac{|U - (\cup RF(Cl_i^{\geq}))|}{|U|}$$

Equação 16 - Qualidade da Aproximação com Dominância

Dessa maneira, a qualidade da aproximação em VC-DRSA será, no mínimo, igual à qualidade da aproximação sem reduzir o nível de consistência l . Para Greco *et al.* (2001) não há uma regra para pré-estabelecer um valor para o nível de consistência l .

Blaszczynski, Greco e Slowinski (2006) discutem três propriedades da aproximação de conjuntos definidas por conjuntos aproximativos. Seja X um subconjunto qualquer gerado, por exemplo, pelo critério de decisão como $X = Cl_2^z$, ou outro qualquer:

- $\underline{P}(X) \subseteq X \subseteq \overline{P}(X)$: a aproximação inferior de um subconjunto X está contida no subconjunto X e a aproximação superior do conjunto X contém o próprio conjunto X – *rough inclusion*;
- $\forall_{X \subseteq U} \underline{P}(X) = U - \overline{P}(U - X)$: a aproximação inferior de um conjunto X é a diferença entre todos os objetos do sistema e a aproximação superior do complemento de X em relação ao universo de objetos – *complementarity*;
- $\forall_{X \subseteq U} \forall_{P, R \subseteq C, P \subseteq R} \underline{P}(X) \subseteq \underline{R}(X) \text{ e } \overline{R}(X) \subseteq \overline{P}(X)$: seja C o conjunto de critérios condicionantes. Dois subconjuntos de C são definidos como P e R . P é subconjunto de R . A aproximação inferior segundo P está contida na aproximação inferior segundo R . No caso das aproximações superiores a situação se inverte, a aproximação superior segundo R está contida na aproximação superior segundo P . Se P mantiver a mesma relação de indiscernibilidade de R (R sendo o conjunto de todos os atributos e P uma redução), as aproximações inferior e superior serão iguais. - monotonicidade da acurácia da aproximação.

O relaxamento e inclusão de alguns objetos na aproximação inferior e, conseqüentemente, o aumento na qualidade da aproximação é tema de diversos estudos. Na prática, a aplicação de VC-DRSA direta causa alguns problemas como, por exemplo, a inclusão de alguns objetos em aproximações inferiores e exclusão de um outro, pertencente a mesma classe que domina este objeto.

Com relação a eventuais falhas na aplicação direta de VC-DRSA (como será apresentada no capítulo 4) alguns autores sugerem a criação de outros índices além do índice de consistência l . Inuiguchi e Yoshioka (2006) propõem um tratamento diferenciado a respeito dos “*outliers*” (objetos que, supostamente, constam em uma classe de atributo de decisão), por exemplo, inferior, porém P -dominam diversos objetos de uma classe superior. O método é chamado de VP-DRSA. Uma combinação de VC-DRSA e VP-DRSA é proposta

por Hu, Wang e Deng (2011), chamada de ISVP-DRSA também gera um outro índice de entrada para aproximação inferior. Estas técnicas serão detalhadas no capítulo 4.

Todos os modelos propostos buscam adequações ao modelo original da TCA para permitir uma qualidade da aproximação maior do que a original e com isso extrair outras regras para suporte ao apoio à decisão.

4 PROPOSTAS ALTERNATIVAS AO USO DIRETO DE VC-DRSA

4.1 INTRODUÇÃO

Como abordado anteriormente, o uso de VC-DRSA como apresentado originalmente, pode levar a uma aproximação inferior de X_i^{\geq} ter, no seu conjunto, um objeto x enquanto que o objeto y que P-domina o objeto x , não faz parte desta aproximação (logicamente, neste caso $x, y \in X_i^{\geq}$). O mesmo raciocínio pode ser estendido a uma aproximação inferior de X_i^{\leq} com o objeto y fora desta aproximação quando y é P-dominado por x .

Blaszczynski (2006) classifica a situação descrita como: “*lack of monotonicity of membership to lower approximation*” (em tradução livre “falta de monotonicidade de pertinência à aproximação inferior”).

Um exemplo desta situação é apresentado:

Suponha que um sistema de informações tenha os seguintes dados apresentados na tabela 3 (as escalas dos atributos têm relação de dominância):

U	Atributos condicionais		Decisão
Objeto	A1	A2	
1	1	1	1
2	5	5	1
3	6	7	1
4	7	6	1
5	2	3	0
6	3	2	0
7	4	4	0
8	8	8	0

Tabela 3 - Sistema de Informações Original

Fonte: autor.

A visualização gráfica dos 8 objetos, de acordo com a sua decisão facilita a visão das dominâncias entre os objetos (figura 4):

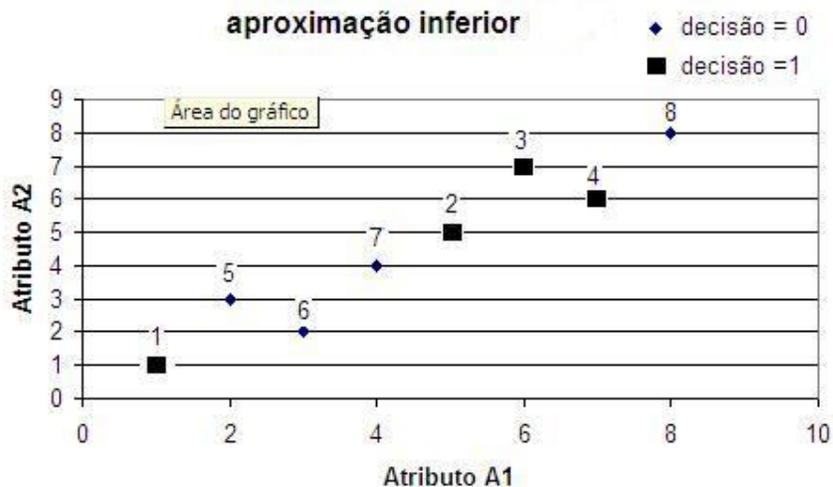


Figura 4 - Visualização gráfica dos objetos
Fonte: autor.

Da figura 4, pode-se estabelecer as duas classes para análise:

$Cl^{\geq 1} = \{1,2,3,4\}$ com cardinalidade igual a 4 e;

$Cl^{\leq 0} = \{5,6,7,8\}$ com cardinalidade igual a 4.

Pelo princípio da P-dominância e considerando os dois atributos A_1 e A_2 como o subconjunto de atributos condicionantes tem-se a seguinte apresentação dos objetos que dominam o objeto x (conjunto P-dominante) $D_p^+(x)$ e os objetos que são dominados pelo objeto x (conjunto P-dominado) $D_p^-(x)$.

A tabela 4 facilita a visualização dos conjuntos P-dominante e P-dominado (entre parênteses aparece a cardinalidade de cada dos conjuntos encontrados). Os conjuntos são encontrados com a utilização das equações 7 (dominantes) e 8 (dominados):

X	Dominantes $D_p^+(x)$	Dominados $D_p^-(x)$
1	{1,2,3,4,5,6,7,8} (8)	{1} (1)
2	{2,3,4,8} (4)	{1,2,5,6,7} (5)
3	{3,8} (2)	{1,2,3,5,6,7} (6)
4	{4,8} (2)	{1,2,4,5,6,7} (6)
5	{2,3,4,5,7,8} (6)	{1,5} (2)
6	{2,3,4,6,7,8} (5)	{1,6} (2)
7	{2,3,4,7,8} (5)	{1,5,6,7} (4)
8	{8} (1)	{1,2,3,4,5,6,7,8} (8)

Tabela 4 - Conjuntos de Objetos Dominantes e Dominados
Fonte: autor.

Nota-se que pelo critério de entrada em DRSA, onde as aproximações são estabelecidas por:

$$\underline{P}(Cl_i^{\geq}) = \{x \in U : D_p^+(x) \subseteq Cl_i^{\geq}\}, \overline{P}(Cl_i^{\geq}) = \{x \in U : D_p^-(x) \cap Cl_i^{\geq} \neq \emptyset\}$$

Equação 17 - Aproximações Superior e Inferior em VC-DRSA

Utilizando a equação 17 pode-se chegar as aproximações para classes “no máximo t” diretamente:

$$\underline{P}(Cl_i^{\leq}) = \{x \in U : D_p^-(x) \subseteq Cl_i^{\leq}\} \text{ e } \overline{P}(Cl_i^{\leq}) = \{x \in U : D_p^+(x) \cap Cl_i^{\leq} \neq \emptyset\}$$

Utilizando a equação 10, a aproximação inferior é estabelecida da seguinte forma:

$$\underline{P}(Cl^{\geq 1}) = \{x \in U : D_p^+(x) \subseteq Cl^{\geq 1}\} = \{ \}$$

Ou seja, nenhum objeto é dominado apenas por membros pertencentes a classe $Cl^{\geq 1}$. Basta verificar que o objeto “8” domina todos os objetos e pertence à classe $Cl^{\leq 0}$. Essa condição é suficiente para que a aproximação inferior seja, segundo P, um conjunto vazio.

Utilizando a equação 12, a aproximação superior é estabelecida da seguinte forma:

$$\overline{P}(Cl^{\geq 1}) = \{x \in U : D_p^-(x) \cap Cl^{\geq 1} \neq \emptyset\} = \{1,2,3,4,5,6,7,8\}$$

Ou seja, todos os objetos dominam pelo menos um membro de $CI^{\geq 1}$. Basta verificar que o objeto “1” é dominado por todos os objetos e pertence a classe $CI^{\geq 1}$. Essa condição é suficiente para que $\overline{P}(CI^{\geq 1}) = \{1,2,3,4,5,6,7,8\}$.

$\underline{P}(CI^{\leq 0}) = \{x \in U : D_p^-(x) \subseteq CI^{\leq 0}\} = \{ \}$, ou seja, nenhum objeto domina apenas os membros pertencentes a classe $CI^{\leq 0}$. Basta verificar que o objeto “1” é dominado por todos os objetos e pertence à classe $CI^{\geq 1}$. Essa condição é suficiente para que a aproximação inferior seja um conjunto vazio.

$\overline{P}(CI^{\leq 0}) = \{x \in U : D_p^+(x) \cap CI^{\leq 0} \neq \emptyset\} = \{1,2,3,4,5,6,7,8\}$, ou seja, todos os objetos são dominados por, pelo menos, um membro de $CI^{\leq 0}$. Basta verificar que o objeto “8” domina todos os objetos e pertence a classe $CI^{\leq 0}$. Essa condição é suficiente para que $\overline{P}(CI^{\leq 0}) = \{1,2,3,4,5,6,7,8\}$.

Como as duas aproximações inferiores são iguais a zero, a qualidade da aproximação medida (equação 16), estabelecida como:

$$\gamma_B(X) = \sum_i \frac{|P(X_i)|}{|U|} = \frac{0+0}{8} = 0$$

Ou seja, nenhuma regra pode ser estabelecida, pois não há objeto em nenhuma aproximação inferior.

A figura 5 retrata o uso do *software* jMAF[®] para o cálculo das aproximações (o código fonte que gera esta tabela para o *software* jMAF[®] está incluído no anexo 4 deste trabalho):

Attributes: 3 Examples: 8

No	#	Object Name	G1 (+)	G2 (+)	PS (+)
1		Example_1	1	1	1
2		Example_2	5	5	1
3		Example_3	6	7	1
4		Example_4	7	6	1
5		Example_5	2	3	0
6		Example_6	3	2	0
7		Example_7	4	4	0
8		Example_8	8	8	0

Console Standard Unions

Quality of approximation: 0.000

Union name	Accuracy	Cardinality
At most 0	0.000	4
Lower		0
Upper		8
Boundary		8
At least 1	0.000	4
Lower		0
Upper		8
Boundary		8

Figura 5 - Cálculo das aproximações
Fonte: autor.

Os índices de entrada estabelecidos por VC-DRSA podem ser calculados para cada um dos objetos e o seu algoritmo de cálculo é estabelecido por:

A equação 14 estabelece os critérios de inclusão de objetos na aproximação inferior. Para cada objeto pode-se calcular qual seria o valor mínimo de l que permitiria a entrada do objeto em questão na aproximação inferior. Os valores são apresentados na tabela 5 (o numerador refere-se a $|D_p^+(x) \cap Cl^{\geq l}|$ se $x \in Cl^{\geq l}$ e $|D_p^-(x) \cap Cl^{\leq l}|$ se $x \in Cl^{\leq l}$, enquanto que o denominador refere-se a $|D_p^+(x)|$ se $x \in Cl^{\geq l}$ e $|D_p^-(x)|$ se $x \in Cl^{\leq l}$).

Objetos	nível
1	4/8
2	3/4
3	1/2
4	1/2
5	1/2
6	1/2
7	3/4
8	4/8

Tabela 5 - Índice de Consistência para os Objetos

Portanto, se $l = 0,75$, os objetos 2 e 7 farão parte das aproximações inferiores relativas às suas classes. Com isso a configuração das aproximações superiores será, também, alterada, pois, em VC-DRSA, as aproximações superiores são formadas por (utilizando a equação 14):

$$\overline{P}^{0.75}(CI^{\geq 1}) = U - \underline{P}^{0.75}(CI^{\leq 0}) = \{1,2,3,4,5,6,7,8\} - \{7\} = \{1,2,3,4,5,6,8\} \text{ e}$$

$$\overline{P}^{0.75}(CI^{\leq 0}) = U - \underline{P}^{0.75}(CI^{\geq 1}) = \{1,2,3,4,5,6,7,8\} - \{2\} = \{1,3,4,5,6,7,8\}$$

4.2 MONOTONICIDADE DE PERTINÊNCIA À APROXIMAÇÃO INFERIOR

Analisando a figura 6, duas regras podem ser estabelecidas com os objetos 2 e 7. Ocorre aqui o que Blasczynski, Greco, Slowinski e Szlag (2006) denominaram de falta de monotonicidade da pertinência à aproximação inferior. A análise gráfica da figura 6 auxilia a percepção:

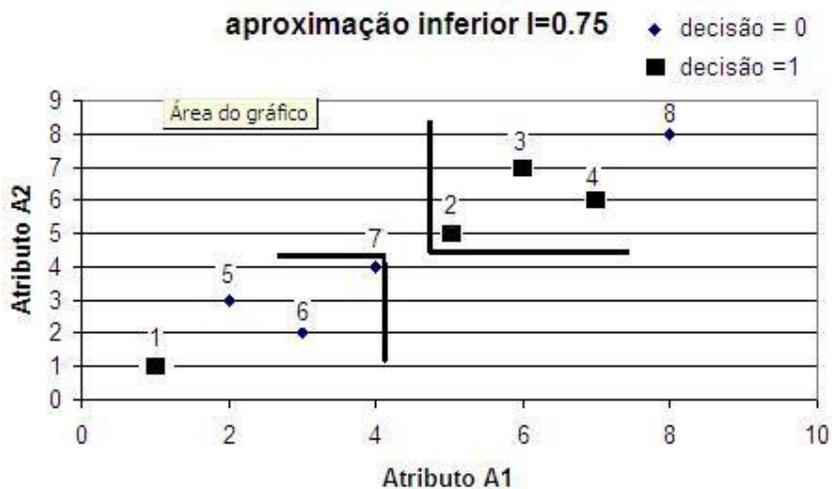


Figura 6 - Visualização gráfica após índice l
Fonte: autor.

Nota-se que os objetos 3 e 4 dominam o objeto 2 e não fazem parte de $\underline{P}^{0.75}(CI^{\geq 1})$. O mesmo ocorre com 5 e 6 que são dominados por 7 e não fazem parte de $\underline{P}^{0.75}(CI^{\leq 0})$. Esses

objetos não pertencem à aproximação inferior porque os seus níveis para entrada são todos iguais a $1/2$, portanto, são excluídos da nova aproximação inferior criada.

O que ocorre nesta abordagem é que a entrada de um objeto na aproximação inferior por VC-DRSA (2 e 7) requer que outros objetos auxiliem a entrada destes ($\{3, 4\}$ e $\{5,6\}$) fazendo com que os níveis aumentem até a entrada.

4.3 NOVAS PROPOSTAS DE INCLUSÃO À APROXIMAÇÃO INFERIOR

Para minimizar este problema, Slowinski e Greco (2004) propuseram o uso do chamado “*robust consistency coefficients*” (em tradução livre, coeficientes de consistência robusta). Este coeficiente é definido como, para cada união X_i^{\geq} e cada objeto $x \in X_i^{\geq}$, o coeficiente de consistência robusta:

$$\varepsilon_p^{X_i^{\geq}}(x) = \max_{y \in D_p^+(x)} \frac{|D_p^+(y) \cap X_i^{\geq}|}{|D_p^+(y)|}$$

Equação 18 - Coeficiente de Consistência Robusta Para Classes Pelo Menos t

Uma extensão do raciocínio leva a:

$$\varepsilon_p^{X_i^{\leq}}(x) = \max_{y \in D_p^-(x)} \frac{|D_p^-(y) \cap X_i^{\leq}|}{|D_p^-(y)|}$$

Equação 19 - Coeficiente de Consistência Robusta Para Classes Até t

De acordo com Slowinski e Greco (2004) as aproximações inferiores estendidas (ao nível de consistência l) são definidas por:

$$\underline{P}(X_i^{\geq}) = \{x \in X_i^{\geq} : \varepsilon_p^{X_i^{\geq}}(x) \geq l\}, \underline{P}(X_i^{\leq}) = \{x \in X_i^{\leq} : \varepsilon_p^{X_i^{\leq}}(x) \geq l\}$$

Equação 20 - Aproximação Inferior com Coeficiente Robusto

O critério de entrada na aproximação inferior garante a monotonicidade de entrada na aproximação inferior, o que significa que se um objeto x pertencer à aproximação inferior todo objeto que domina x e pertence à mesma classe que x também entrará na aproximação

$\underline{P}^l(X_i^{\geq})$. O mesmo conceito pode ser aplicado para $\underline{P}^l(X_i^{\leq})$, porém com objetos que são dominados por x .

Blaszczynski *et al.* (2006) observam que é possível que alguns objetos pertençam a $\underline{P}^l(X_i^{\geq})$ porque P-dominam objetos $y \in X_{i-1}^{\leq}$ e não são objetos da sua própria classe. Dessa maneira, os autores sugerem uma nova regra para $\varepsilon_p^{X_i^{\geq}}(x)$:

$$\varepsilon_p^{X_i^{\geq}}(x) = \max_{y \in D_p^+(x) \cap X_i^{\geq}} \frac{|D_p^+(y) \cap X_i^{\geq}|}{|D_p^+(y)|}, \varepsilon_p^{X_i^{\leq}}(x) = \max_{y \in D_p^-(x) \cap X_i^{\leq}} \frac{|D_p^-(y) \cap X_i^{\leq}|}{|D_p^-(y)|}$$

Equação 21 - Coeficiente de Consistência Robusto Alterado

Esse critério garante a monotonicidade de entrada na aproximação inferior e, ao mesmo tempo, garante que uma aproximação inferior de X_i^{\geq} ou de X_i^{\leq} conterá apenas objetos pertencentes às respectivas classes.

As propriedades básicas são mantidas com o uso deste critério de inclusão:

1. Inclusão aproximativa: $\underline{P}(X) \subseteq X \subseteq \overline{P}(X)$;
2. Complementaridade: $\underline{P}(X) = U - \overline{P}(U - X)$;
3. Monotonicidade da acurácia da aproximação:

$$\forall_{X \subseteq U} \forall_{P, R \subseteq C, P \subseteq R} \underline{P}(X) \subseteq \underline{R}(X) \text{ e } \overline{R}(X) \subseteq \overline{P}(X)$$

Utilizando o coeficiente de consistência robusta e o mesmo nível $l = 0,75$ como critério de entrada para a aproximação inferior, tem-se a seguinte memória de cálculo do coeficiente de consistência robusta para os objetos do exemplo:

Os objetos pertencem às classes:

$$CI^{\leq 0} = \{5,6,7,8\}; CI^{\geq 1} = \{1,2,3,4\}$$

Utilizando a equação 21, para os objetos pertencentes à classe $CI^{\geq 1}$:

$$\begin{aligned}\varepsilon^{X \geq 1}(1) &= \max_{y \in D_P^+(x) \cap X_i^{\geq}} \frac{|D_P^+(y) \cap X_i^{\geq}|}{|D_P^+(y)|} \Rightarrow \text{objeto 1} \Rightarrow \max\left[\frac{4}{8}\right] = \frac{4}{8}; \\ \varepsilon^{X \geq 1}(2) &= \max_{y \in D_P^+(x) \cap X_i^{\geq}} \frac{|D_P^+(y) \cap X_i^{\geq}|}{|D_P^+(y)|} \Rightarrow \text{objetos 1 e 2} \Rightarrow \max\left[\frac{4}{8}; \frac{3}{4}\right] = \frac{3}{4}; \\ \varepsilon^{X \geq 1}(3) &= \max_{y \in D_P^+(x) \cap X_i^{\geq}} \frac{|D_P^+(y) \cap X_i^{\geq}|}{|D_P^+(y)|} \Rightarrow \text{objetos 1, 2 e 3} \Rightarrow \max\left[\frac{4}{8}; \frac{3}{4}; \frac{1}{2}\right] = \frac{3}{4}; \\ \varepsilon^{X \geq 1}(4) &= \max_{y \in D_P^+(x) \cap X_i^{\geq}} \frac{|D_P^+(y) \cap X_i^{\geq}|}{|D_P^+(y)|} \Rightarrow \text{objetos 1, 2 e 4} \Rightarrow \max\left[\frac{4}{8}; \frac{3}{4}; \frac{1}{2}\right] = \frac{3}{4}.\end{aligned}$$

Utilizando a equação 21, para os objetos pertencentes à classe $CI^{\leq 0}$:

$$\begin{aligned}\varepsilon^{X \leq 0}(5) &= \max_{y \in D_P^-(x) \cap X_i^{\leq}} \frac{|D_P^-(y) \cap X_i^{\leq}|}{|D_P^-(y)|} \Rightarrow \text{objetos 5,7 e 8} \Rightarrow \max\left[\frac{1}{2}; \frac{3}{4}; \frac{4}{8}\right] = \frac{3}{4}; \\ \varepsilon^{X \leq 0}(6) &= \max_{y \in D_P^-(x) \cap X_i^{\leq}} \frac{|D_P^-(y) \cap X_i^{\leq}|}{|D_P^-(y)|} \Rightarrow \text{objetos 6,7 e 8} \Rightarrow \max\left[\frac{1}{2}; \frac{3}{4}; \frac{4}{8}\right] = \frac{3}{4}; \\ \varepsilon^{X \geq 1}(7) &= \max_{y \in D_P^+(x) \cap X_i^{\leq}} \frac{|D_P^-(y) \cap X_i^{\leq}|}{|D_P^-(y)|} \Rightarrow \text{objetos 7 e 8} \Rightarrow \max\left[\frac{3}{4}; \frac{4}{8}\right] = \frac{3}{4}; \\ \varepsilon^{X \geq 1}(8) &= \max_{y \in D_P^+(x) \cap X_i^{\leq}} \frac{|D_P^-(y) \cap X_i^{\leq}|}{|D_P^-(y)|} \Rightarrow \text{objetos 8} \Rightarrow \max\left[\frac{4}{8}\right] = \frac{4}{8}.\end{aligned}$$

Como o critério para entrada na aproximação inferior (equação 20) é:

$$\underline{P}(X_i^{\geq}) = \{x \in X_i^{\geq} : \varepsilon_p^{X_i^{\geq}}(x) \geq l\}, \underline{P}(X_i^{\leq}) = \{x \in X_i^{\leq} : \varepsilon_p^{X_i^{\leq}}(x) \geq l\}$$

Tem-se:

$$\underline{P}^{0,75}(CI^{\geq 1}) = \{2,3,4\}$$

$$\underline{P}^{0,75}(CI^{\leq 0}) = \{5,6,7\}$$

A formação das aproximações superiores é calculada da seguinte forma (equação 17):

$$\begin{aligned}\overline{P}^{0,75}(CI^{\geq 1}) &= U - \underline{P}^{0,75}(CI^{\leq 0}) = \{1,2,3,4,5,6,7,8\} - \{5,6,7\} = \{1,2,3,4,8\} \text{ e} \\ \overline{P}^{0,75}(CI^{\leq 0}) &= U - \underline{P}^{0,75}(CI^{\geq 1}) = \{1,2,3,4,5,6,7,8\} - \{2,3,4\} = \{1,5,6,7,8\}\end{aligned}$$

O software jMAF[®] utiliza a proposta de Blaszczynski *et al.* (2006). Ao calcular, através do software, as saídas obtidas são apresentadas nas figuras 7 e 8:

Os objetos pertencentes às aproximações da classe $CI^{\leq 0}$ são apresentados na figura 7:

Attributes: 3 Examples: 8

No	#	Object Name	G1 (+)	G2 (+)	PS (+)
1		Example_1	1	1	1
2		Example_2	5	5	1
3		Example_3	6	7	1
4		Example_4	7	6	1
5		Example_5	2	3	0
6		Example_6	3	2	0
7		Example_7	4	4	0
8		Example_8	8	8	0

Quality of approximation: 0.750 Track in Editor

Union name	Accuracy	Cardinality
At most 0	0.600	4
Lower		3
Example_5		
Example_6		
Example_7		
Upper		5
Example_1		
Example_5		
Example_6		
Example_7		
Example_8		
Boundary		2
Example_1		
Example_8		
At least 1	0.600	4

Figura 7 - Objetos de uma classe de dominância
Fonte: autor.

Os objetos pertencentes às aproximações da classe $CI^{\geq 1}$ são apresentados na figura 8:

Attributes: 3 Examples: 8

No	#	Object Name	G1 (+)	G2 (+)	PS (+)
1		Example_1	1	1	1
2		Example_2	5	5	1
3		Example_3	6	7	1
4		Example_4	7	6	1
5		Example_5	2	3	0
6		Example_6	3	2	0
7		Example_7	4	4	0
8		Example_8	8	8	0

Quality of approximation: 0.750 Track in Editor

Union name	Accuracy	Cardinality
At most 0	0.600	4
At least 1	0.600	4
Lower		3
Example_2		
Example_3		
Example_4		
Upper		5
Example_1		
Example_2		
Example_3		
Example_4		
Example_8		
Boundary		2
Example_1		
Example_8		

Figura 8 - Objetos de uma classe de dominância
Fonte: autor.

A qualidade da aproximação refere-se à soma das cardinalidades das aproximações inferiores em relação à cardinalidade do total de objetos, ou seja, utilizando a equação 16:

$$\gamma_B(X) = \sum_i \frac{|B(X_i)|}{|U|} = \frac{3+3}{8} = \frac{6}{8} = 0,75.$$

4.4 OUTRAS PROPOSTAS (VP-DRSA e ISVP-DRSA)

Outras propostas de formação de aproximações inferiores foram apresentadas por alguns autores. INUIGUCHI & ISHIOKA (2006) sugerem a criação de um índice β , que é apresentado por (supondo que $x \in Cl_i^{\geq}$):

$$\beta = \frac{|D_p^-(x) \cap Cl_i^{\geq}|}{|D_p^-(x) \cap Cl_i^{\geq}| + |D_p^+(x) \cap Cl_{i-1}^{\leq}|}$$

Equação 22 - Índice Beta em VP-DRSA

A interpretação do índice β pode ser feita da seguinte maneira:

No numerador ($|D_p^-(x) \cap Cl_i^{\geq}|$), está o total de objetos que um objeto x domina e pertence à classe Cl_i^{\geq} , ou seja, é a quantidade de objetos que endossa a presença de x na classe Cl_i^{\geq} , pois o objeto x P-domina uma quantidade igual à cardinalidade do conjunto de objetos que pertencem à classe Cl_i^{\geq} . Por outro lado o denominador é formado pela soma do numerador acrescida da cardinalidade do termo ($|D_p^+(x) \cap Cl_{i-1}^{\leq}|$) que representa a quantidade de objetos que P-dominam o objeto x e fazem parte da classe Cl_{i-1}^{\leq} , ou seja, que não endossam a participação do objeto x na classe Cl_i^{\geq} .

Como as cardinalidades, por definição o total de objetos em um subconjunto qualquer, nunca serão negativas, o índice β será sempre um valor entre zero e um. O denominador nunca será igual a zero, pois um objeto x faz parte dos grupos $D_p^+(x)$ e $D_p^-(x)$ (um objeto P-domina ele próprio e também é P-dominado por ele mesmo). Como a união dos objetos da tabela original pode sempre ser escrita como forma de $U = Cl_{i-1}^{\leq} + Cl_i^{\geq}$, infere-se que uma das

duas formações do denominador terá cardinalidade mínima igual a 1, portanto o denominador de β nunca será zero.

Então, para um objeto x qualquer, β é a razão entre a quantidade de objetos que endossam a presença de x em Cl_t^{\geq} pela quantidade de objetos que endossam a presença em Cl_t^{\geq} somada quantidade de objetos que endossam a presença de x em Cl_{t-1}^{\leq} .

Um raciocínio equivalente pode levar a criação do índice β para $x \in Cl_t^{\leq}$.

$$\beta = \frac{|D_p^+(x) \cap Cl_t^{\leq}|}{|D_p^+(x) \cap Cl_t^{\leq}| + |D_p^-(x) \cap Cl_{t+1}^{\geq}|}, \text{ a interpretação deste índice pode ser resumida}$$

por:

O numerador representa a quantidade de objetos que dominam x e pertencem à classe Cl_t^{\leq} , ou seja, endossam a presença de x na classe Cl_t^{\leq} . O denominador, além desta parcela, tem somado a quantidade de objetos que são dominados por x e fazem parte de Cl_{t+1}^{\geq} , ou seja, não endossam a presença de x em Cl_t^{\leq} .

Assim como o índice β anterior, o valor desta divisão está limitada à região entre zero e um e o denominador nunca será zero, pois x domina e é dominado por ele próprio e, também, pertence a Cl_t^{\leq} ou Cl_{t+1}^{\geq} . Portanto uma das duas cardinalidades será, pelo menos, igual a um. Por definição não há cardinalidade negativa.

Inuiguchi e Ishioka (2006) definiram a aproximação inferior de Cl_t^{\geq} como sendo uma “região P-positiva” de Cl_t^{\geq} ($POS_p^l(Cl_t^{\geq})$), sendo P o subconjunto de atributos condicionantes, (originalmente o nome sugerido foi *P-positive region of Cl_t^{\geq}*). O valor de l é assumido da mesma maneira que em VC-DRSA.

$$POS_p^l(Cl_t^{\geq}) = \left\{ x \in U : \frac{|D_p^-(x) \cap Cl_t^{\geq}|}{|D_p^-(x) \cap Cl_t^{\geq}| + |D_p^+(x) \cap Cl_{t-1}^{\leq}|} \geq l \right\}$$

$$POS_p^l(Cl_t^{\leq}) = \left\{ x \in U : \frac{|D_p^+(x) \cap Cl_t^{\leq}|}{|D_p^+(x) \cap Cl_t^{\leq}| + |D_p^-(x) \cap Cl_{t+1}^{\geq}|} \geq l \right\}$$

Equação 23 - Região P-Positiva em VP-DRSA

A aproximação superior de CI_i^{\geq} , $NNG_p^l(CI_i^{\geq})$, correspondente “P-região não-negativa de CI_i^{\geq} (*P-non-negative regions of CI_i^{\geq}*) é definida como complemento, em relação ao total de objetos da região positiva de CI_{i-1}^{\leq} . Um raciocínio equivalente leva a $POS_p^l(CI_i^{\geq})$, ou seja:

$$NNG_p^l(CI_i^{\geq}) = U - POS_p^l(CI_{i-1}^{\leq}), NNG_p^l(CI_i^{\leq}) = U - POS_p^l(CI_{i+1}^{\geq})$$

Equação 24 - Região Não-Negativa em VP-DRSA

Inuiguchi e Ishioka (2006) provam que:

$$POS_p^l(CI_i^{\geq}) \subseteq NNG_p^l(CI_i^{\geq}) \text{ e } POS_p^l(CI_i^{\leq}) \subseteq NNG_p^l(CI_i^{\leq})$$

Inuiguchi, Yoshioka e Kusunoki (2009) mostram, porém, que não é possível provar que $POS_p^l(CI_i^{\geq}) \subseteq CI_i^{\geq}$ e $POS_p^l(CI_i^{\leq}) \subseteq CI_i^{\leq}$, ou seja, é possível que a região P-positiva de CI_i^{\geq} (equivalente a aproximação inferior em VC-DRSA) possua objetos que não pertencem a CI_i^{\geq} . O mesmo raciocínio serve para CI_i^{\leq} . Essa importante propriedade para criação de regras de apoio à decisão não é satisfeita em VP-DRSA.

Utilizando o exemplo anterior para aplicar a definição de VP-DRSA, em uma tabela com 8 objetos, tem-se os cálculos dos valores para $POS_p^l(CI_i^{\geq}); POS_p^l(CI_i^{\leq}); NNG_p^l(CI_i^{\geq}); NNG_p^l(CI_i^{\leq})$. O valor utilizado como referência de entrada de objetos na aproximação inferior utilizando VC-DRSA foi $l = 0,75$. O conjunto P será formado pelos dois critérios condicionantes A1 e A2. O mesmo valor para l será usado para VP-DRSA. Para os oito objetos do sistema original, tem-se os seguintes valores calculados:

Para o cálculo de $POS_p^{0,75}(CI_i^{\geq})$ a tabela 6 facilita o uso do algoritmo de classificação em VP-DRSA:

x	$D_p^-(x) \cap CI^{\geq 1}$	$D_p^+(x) \cap CI^{\leq 0}$
1	{1}	{5,6,7,8}
2	{1,2}	{8}
3	{1,2,3}	{8}
4	{1,2,4}	{8}
5	{1}	{5,7,8}
6	{1}	{6,7,8}
7	{1}	{7,8}
8	{1,2,3,4}	{8}

Tabela 6 - Interseção de Dominância (Dominado) com as Classes
Fonte: autor.

A utilização equação 23 leva aos seguintes cálculos:

$$POS_p^{0,75}(CI^{\geq 1}) = \left\{ 1 \in U : \frac{|D_p^-(1) \cap CI^{\geq 1}|}{|D_p^-(1) \cap CI^{\geq 1}| + |D_p^+(1) \cap CI^{\leq 0}|} = \frac{1}{(1+4)} \geq 0,75 \right\} (falso)$$

$$POS_p^{0,75}(CI^{\geq 1}) = \left\{ 2 \in U : \frac{|D_p^-(2) \cap CI^{\geq 1}|}{|D_p^-(2) \cap CI^{\geq 1}| + |D_p^+(2) \cap CI^{\leq 0}|} = \frac{2}{(2+1)} \geq 0,75 \right\} (falso)$$

$$POS_p^{0,75}(CI^{\geq 1}) = \left\{ 3 \in U : \frac{|D_p^-(3) \cap CI^{\geq 1}|}{|D_p^-(3) \cap CI^{\geq 1}| + |D_p^+(3) \cap CI^{\leq 0}|} = \frac{3}{(3+1)} \geq 0,75 \right\} (verdadeiro)$$

$$POS_p^{0,75}(CI^{\geq 1}) = \left\{ 4 \in U : \frac{|D_p^-(4) \cap CI^{\geq 1}|}{|D_p^-(4) \cap CI^{\geq 1}| + |D_p^+(4) \cap CI^{\leq 0}|} = \frac{3}{(3+1)} \geq 0,75 \right\} (verdadeiro)$$

$$POS_p^{0,75}(CI^{\geq 1}) = \left\{ 5 \in U : \frac{|D_p^-(5) \cap CI^{\geq 1}|}{|D_p^-(5) \cap CI^{\geq 1}| + |D_p^+(5) \cap CI^{\leq 0}|} = \frac{1}{(1+3)} \geq 0,75 \right\} (falso)$$

$$POS_p^{0,75}(CI^{\geq 1}) = \left\{ 6 \in U : \frac{|D_p^-(6) \cap CI^{\geq 1}|}{|D_p^-(6) \cap CI^{\geq 1}| + |D_p^+(6) \cap CI^{\leq 0}|} = \frac{1}{(1+3)} \geq 0,75 \right\} (falso)$$

$$POS_p^{0,75}(CI^{\geq 1}) = \left\{ 7 \in U : \frac{|D_p^-(7) \cap CI^{\geq 1}|}{|D_p^-(7) \cap CI^{\geq 1}| + |D_p^+(7) \cap CI^{\leq 0}|} = \frac{1}{(1+2)} \geq 0,75 \right\} (falso)$$

$$POS_p^{0,75}(CI^{\geq 1}) = \left\{ 8 \in U : \frac{|D_p^-(8) \cap CI^{\geq 1}|}{|D_p^-(8) \cap CI^{\geq 1}| + |D_p^+(8) \cap CI^{\leq 0}|} = \frac{4}{(4+1)} \geq 0,75 \right\} (verdadeiro)$$

Portanto, pode-se afirmar que, por VP-DRSA, o conjunto $POS_p^{0,75}(CI^{\geq 1}) = \{3,4,8\}$.

Para o cálculo de $POS_p^{0,75}(CI^{\leq 0})$, a tabela 7 facilita o uso do algoritmo de classificação em VP-DRSA:

X	$D_p^+(x) \cap CI^{\leq 0}$	$D_p^-(x) \cap CI^{\geq 1}$
1	{5,6,7,8}	{1}
2	{8}	{1,2}
3	{8}	{1,2,3}
4	{8}	{1,2,4}
5	{5,7,8}	{1}
6	{6,7,8}	{1}
7	{7,8}	{1}
8	{8}	{1,2,3,4}

Tabela 7 - Interseção de Dominância (Dominado) com as Classes

Fonte: autor.

A equação 23 leva aos seguintes cálculos:

$$POS_p^{0,75}(CI^{\leq 0}) = \left\{ 1 \in U : \frac{|D_p^+(1) \cap CI^{\leq 0}|}{|D_p^+(1) \cap CI^{\leq 0}| + |D_p^-(1) \cap CI^{\geq 1}|} = \frac{4}{(4+1)} \geq 0,75 \right\} (verdadeiro)$$

$$POS_p^{0,75}(CI^{\leq 0}) = \left\{ 2 \in U : \frac{|D_p^+(2) \cap CI^{\leq 0}|}{|D_p^+(2) \cap CI^{\leq 0}| + |D_p^-(2) \cap CI^{\geq 1}|} = \frac{1}{(1+2)} \geq 0,75 \right\} (falso)$$

$$\begin{aligned}
POS_p^{0,75}(CI^{\leq 0}) &= \left\{ 3 \in U : \frac{|D_p^+(3) \cap CI^{\leq 0}|}{|D_p^+(3) \cap CI^{\leq 0}| + |D_p^-(3) \cap CI^{\geq 1}|} = \frac{1}{(1+3)} \geq 0,75 \right\} (falso) \\
POS_p^{0,75}(CI^{\leq 0}) &= \left\{ 4 \in U : \frac{|D_p^+(4) \cap CI^{\leq 0}|}{|D_p^+(4) \cap CI^{\leq 0}| + |D_p^-(4) \cap CI^{\geq 1}|} = \frac{1}{(1+3)} \geq 0,75 \right\} (falso) \\
POS_p^{0,75}(CI^{\leq 0}) &= \left\{ 5 \in U : \frac{|D_p^+(5) \cap CI^{\leq 0}|}{|D_p^+(5) \cap CI^{\leq 0}| + |D_p^-(5) \cap CI^{\geq 1}|} = \frac{3}{(3+1)} \geq 0,75 \right\} (verdadeiro) \\
POS_p^{0,75}(CI^{\leq 0}) &= \left\{ 6 \in U : \frac{|D_p^+(6) \cap CI^{\leq 0}|}{|D_p^+(6) \cap CI^{\leq 0}| + |D_p^-(6) \cap CI^{\geq 1}|} = \frac{3}{(3+1)} \geq 0,75 \right\} (verdadeiro) \\
POS_p^{0,75}(CI^{\leq 0}) &= \left\{ 7 \in U : \frac{|D_p^+(7) \cap CI^{\leq 0}|}{|D_p^+(7) \cap CI^{\leq 0}| + |D_p^-(7) \cap CI^{\geq 1}|} = \frac{2}{(2+1)} \geq 0,75 \right\} (falso) \\
POS_p^{0,75}(CI^{\leq 0}) &= \left\{ 8 \in U : \frac{|D_p^+(8) \cap CI^{\leq 0}|}{|D_p^+(8) \cap CI^{\leq 0}| + |D_p^-(8) \cap CI^{\geq 1}|} = \frac{1}{(1+4)} \geq 0,75 \right\} (falso)
\end{aligned}$$

Portanto, pode-se afirmar que, por VP-DRSA, o conjunto $POS_p^{0,75}(CI^{\leq 0}) = \{1,5,6\}$.

Pela definição mostrada anteriormente, tem-se que $NNG_p^l(CI_t^{\geq}) = U - POS_p^l(CI_{t-1}^{\leq})$ e $NNG_p^l(CI_t^{\leq}) = U - POS_p^l(CI_{t+1}^{\geq})$. Portanto, a equação 24 leva aos seguintes valores:

$$NNG_p^{0,75}(CI^{\geq 1}) = U - POS_p^l(CI^{\leq 0}) = \{1,2,3,4,5,6,7,8\} - \{1,5,6\} = \{2,3,4,7,8\} \text{ e}$$

$$NNG_p^{0,75}(CI^{\leq 0}) = U - POS_p^{0,75}(CI^{\geq 1}) = \{1,2,3,4,5,6,7,8\} - \{3,4,8\} = \{1,2,5,6,7\}.$$

É possível notar que as duas regiões não-negativas criadas contêm as duas regiões P-positivas.

A figura 9 detalha os conjuntos formados pelas regiões P-positivas de acordo com VP-DRSA:

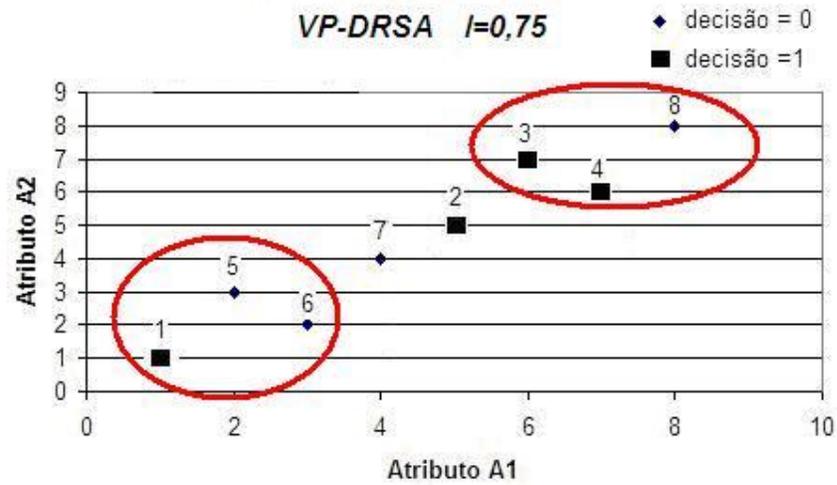


Figura 9 - Regiões P-positivas (VP-DRSA)
Fonte: autor.

Nota-se que as duas regiões P-positivas contêm objetos que não pertencem estritamente a sua classe, ou seja:

$$POS_p^l(Cl_i^{\geq}) \not\subset Cl_i^{\geq} \text{ e } POS_p^l(Cl_i^{\leq}) \not\subset Cl_i^{\leq}.$$

Para o uso do VP-DRSA tem-se a seguinte qualidade da aproximação (equação 16):

$$\gamma_B(X) = \sum_i \frac{|P(X_i)|}{|U|} = \frac{3+3}{8} = \frac{6}{8}.$$

Ou seja, 75% dos objetos são capazes de gerar regras ao nível mínimo de consistência $l = 0,75$.

Deng, Wang e Hu (2011) propuseram uma modificação com um ajuste mínimo no critério de entrada, onde somente seriam candidatos a pertencer à aproximação inferior, membros da própria classe Cl_i^{\geq} ou Cl_i^{\leq} . Portanto, por Deng, Wang e Hu (2011), as aproximações inferiores (antes chamada de região P-positiva por Inuiguchi, Yoshioka e Kusunoki (2009)) são definidas como:

$$\begin{aligned} \underline{P}_{VP}^l(Cl_t^{\geq}) &= \left\{ x \in Cl_t^{\geq} : \frac{|D_p^-(x) \cap Cl_t^{\geq}|}{|D_p^-(x) \cap Cl_t^{\geq}| + |D_p^+(x) \cap Cl_{t-1}^{\leq}|} \geq l \right\} e \\ \underline{P}_P^l(Cl_t^{\leq}) &= \left\{ x \in Cl_t^{\leq} : \frac{|D_p^+(x) \cap Cl_t^{\leq}|}{|D_p^+(x) \cap Cl_t^{\leq}| + |D_p^-(x) \cap Cl_{t+1}^{\geq}|} \geq l \right\} \end{aligned}$$

Equação 25 - Aproximação Inferior em VP-DRSA Alterada

Usando o conceito de dualidade das aproximações inferiores, as aproximações superiores podem ser definidas por:

$$\begin{aligned} \overline{P}_{VP}^l(Cl_t^{\geq}) &= \left\{ x \in U : \frac{|D_p^-(x) \cap Cl_t^{\geq}|}{|D_p^-(x) \cap Cl_t^{\geq}| + |D_p^+(x) \cap Cl_{t-1}^{\leq}|} \geq 1-l \right\} e \\ \overline{P}_P^l(Cl_t^{\leq}) &= \left\{ x \in U : \frac{|D_p^+(x) \cap Cl_t^{\leq}|}{|D_p^+(x) \cap Cl_t^{\leq}| + |D_p^-(x) \cap Cl_{t+1}^{\geq}|} \geq 1-l \right\} \end{aligned}$$

Equação 26 - Aproximação Superior em VP-DRSA Alterada

Deng, Wang e Hu (2011) notaram que a entrada de objetos nas aproximações inferiores ainda estava associada à cardinalidade de $D_p^-(x) \cap Cl_t^{\geq}$ ou de $D_p^+(x) \cap Cl_t^{\leq}$ o que leva a problemas semelhantes de inclusão em aproximações inferiores aos estudos iniciais de VC-DRSA. Sendo assim, os autores propuseram um novo critério denominado ISVP-DRSA (*inclusion degree and supported degree for DRSA*).

Supondo que $x \in Cl_t^{\geq}$ e que o subconjunto de critérios P está contido no conjunto de todos os critérios condicionantes C ($P \subseteq C$), um coeficiente β pode ser gerado por:

$$\beta(x) = \max \left(\frac{|D_p^+(x) \cap Cl_t^{\geq}|}{|D_p^+(x)|}, \frac{|D_p^-(x) \cap Cl_t^{\geq}|}{|D_p^-(x) \cap Cl_t^{\geq}| + |D_p^+(x) \cap Cl_{t-1}^{\leq}|} \right)$$

Equação 27 - Coeficiente Beta em ISVP-DRSA

Ou seja, a primeira parte da função de maximização $\left(\frac{|D_P^+(x) \cap Cl_t^{\geq}|}{|D_P^+(x)|} \right)$ é exatamente a mesma função de entrada de objetos na aproximação inferior de VC-DRSA. A segunda parte da função de maximização $\left(\frac{|D_P^-(x) \cap Cl_t^{\geq}|}{|D_P^-(x) \cap Cl_t^{\geq}| + |D_P^+(x) \cap Cl_{t-1}^{\leq}|} \right)$ é a mesma proposta de entrada de objetos na região P-positiva de Cl_t^{\geq} .

De uma maneira análoga, pode-se encontrar o coeficiente β para um objeto $x \in Cl_t^{\leq}$:

Supondo que $x \in Cl_t^{\leq}$ e que o subconjunto de critérios P está contido no conjunto de todos os critérios condicionantes C ($P \subseteq C$), um coeficiente β pode ser gerado por:

$$\beta(x) = \max \left(\frac{|D_P^-(x) \cap Cl_t^{\leq}|}{|D_P^-(x)|}, \frac{|D_P^+(x) \cap Cl_t^{\leq}|}{|D_P^+(x) \cap Cl_t^{\leq}| + |D_P^-(x) \cap Cl_{t+1}^{\geq}|} \right).$$

Desta forma, a aproximação inferior em ISVP-DRSA será o conjunto de objetos que satisfaça:

$$\underline{P}_{ISVP}^l(Cl_t^{\geq}) = \{x \in Cl_t^{\geq} : \beta(x) \geq l\}, \underline{P}_{ISVP}^l(Cl_t^{\leq}) = \{x \in Cl_t^{\leq} : \beta(x) \geq l\}$$

Equação 28 - Aproximação Inferior em ISVP-DRSA

Assim como em VC-DRSA, as aproximações superiores podem ser geradas por:

$$\overline{P}_{ISVP}^l(Cl_t^{\geq}) = U - \underline{P}_{ISVP}^l(Cl_{t-1}^{\leq}), \overline{P}_{ISVP}^l(Cl_t^{\leq}) = U - \underline{P}_{ISVP}^l(Cl_{t+1}^{\geq})$$

Equação 29 - Aproximação Superior em ISVP-DRSA

No exemplo estudado com 8 objetos os índices β gerados (equação 26) para cada objeto pode ser resumido em:

$$\text{Objetos pertencentes a } Cl^{\geq 1} = \{1,2,3,4\}$$

Objeto	$\frac{ D_p^+(x) \cap Cl^{\geq 1} }{ D_p^+(x) }$	$\frac{ D_p^-(x) \cap Cl^{\geq 1} }{ D_p^-(x) \cap Cl^{\geq 1} + D_p^+(x) \cap Cl^{\leq 0} }$	Max
			β
1	4/8	1/5	4/8
2	2/4	2/3	3/4
3	1/2	3/4	3/4
4	1/2	3/4	3/4

Tabela 8 - Índice para Aproximação Inferior (ISVP-DRSA)

Fonte: autor.

Utilizando o mesmo valor de $l = 0,75$. Tem-se a seguinte formação de aproximação inferior (utilizando a equação 28):

$$P_{ISVP}^{0,75}(Cl^{\geq 1}) = \{x \in Cl^{\geq 1} : \beta(x) \geq 0,75\} = \{2,3,4\}$$

Objetos pertencentes a $Cl^{\leq 0} = \{5,6,7,8\}$

Objeto	$\frac{ D_p^-(x) \cap Cl_r^{\leq 0} }{ D_p^-(x) }$	$\frac{ D_p^+(x) \cap Cl^{\leq 0} }{ D_p^+(x) \cap Cl^{\leq 0} + D_p^-(x) \cap Cl^{\geq 1} }$	Max
			β
5	1/2	3/4	3/4
6	1/2	3/4	3/4
7	3/4	2/3	3/4
8	4/8	1/5	4/8

Tabela 9 - Índice para Aproximação Inferior (ISVP-DRSA)

Fonte: autor.

Utilizando o mesmo valor de $l = 0,75$. Tem-se a seguinte formação de aproximação inferior (utilizando a equação 28):

$$P_{ISVP}^{0,75}(Cl^{\leq 0}) = \{x \in Cl^{\leq 0} : \beta(x) \geq 0,75\} = \{5,6,7\}.$$

Desta maneira, todos os membros de uma aproximação inferior de uma classe pertencem a esta própria classe. As aproximações superiores são estabelecidas da seguinte forma (equação 28):

$$\bar{P}_{ISVP}^l(CI^{\geq 1}) = U - \underline{P}_{ISVP}^l(CI^{\leq 0}) = \{1,2,3,4,5,6,7,8\} - \{5,6,7\} = \{1,2,3,4,8\}$$

$$\bar{P}_{ISVP}^{0,75}(CI^{\leq 0}) = U - \underline{P}_{ISVP}^{0,75}(CI^{\geq 1}) = \{1,2,3,4,5,6,7,8\} - \{2,3,4\} = \{1,5,6,7,8\}$$

A figura 10 detalha os conjuntos formados pelas aproximações inferiores de acordo com ISVP-DRSA:

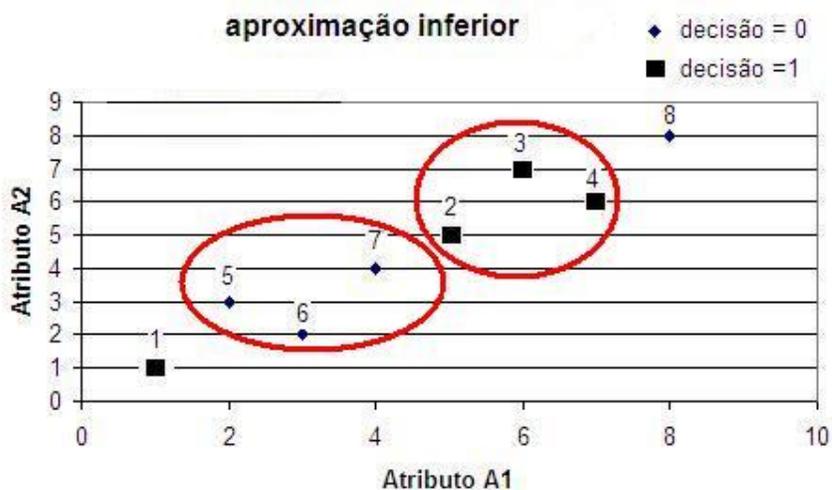


Figura 10 - Regiões P-positivas (ISVP-DRSA)

Fonte: autor.

Para o uso do ISVP-DRSA tem-se a seguinte qualidade da aproximação (equação 16):

$$\gamma_B(X) = \sum_i \frac{|P(X_i)|}{|U|} = \frac{3+3}{8} = \frac{6}{8}$$

Ou seja, 75% dos objetos são capazes de gerar regras ao nível mínimo de consistência $l = 0,75$. Pode-se observar que é o mesmo valor encontrado para VP-DRSA, porém os objetos que pertencem à aproximação inferior de ambas as classes não são os mesmos.

5 REDUÇÃO DE VALORES PARA O ATRIBUTO DE DECISÃO

5.1 INTRODUÇÃO

Conforme apresentado no capítulo anterior, se um determinado conjunto de valores possui uma inconsistência na sua classificação o resultado da qualidade da aproximação será impactado. A influência deste impacto no índice de qualidade da aproximação dependerá do índice l utilizado e de qual a técnica utilizada. As inconsistências surgem porque os objetos do sistema de informação original possuem dois estados: “presentes” (fazendo parte da aproximação inferior) ou “ausentes” (não fazendo parte da aproximação inferior).

Blaszczynski *et al.* (2006) apresentam possíveis violações que levam a uma inconsistência que viola o princípio de dominância. Supondo que todos os atributos possuem dominância, dois objetos do sistema de informação são apresentados:

Objeto	Matemática	Física	Literatura	Avaliação
1	Bom	Médio	Ruim	Ruim
2	Médio	Médio	Ruim	Médio

Tabela 10 - Inconsistência entre dois atributos

A tabela 10 permite observar que “1” não tem avaliações piores do que “2” nos três atributos condicionantes. Porém, para o critério de decisão, a avaliação de “1” é pior do que a de “2”. Estes dois objetos são inconsistentes e prejudicarão a qualidade da aproximação.

Com o objetivo de estudar possíveis impactos na qualidade da aproximação, SLOWINSKI (1992) define indiscernibilidade estrita (*strict indiscernibility*) e indiscernibilidade fraca (*weak indiscernibility*) para atributos quantitativos, com domínio de

valores em uma região contínua, baseada na região de fronteira gerada pelas aproximações superior e inferior. Esse autor, baseado em variáveis quantitativas, com domínio bem definido em um intervalo de números reais (por exemplo: pressão sanguínea ou temperatura corporal – que tem infinitos valores da sua região de domínio), assinala que a alocação dos objetos em classes, definidas arbitrariamente segundo normas pré-estabelecidas (por exemplo: baixo, médio e alto) pode influenciar o resultado da análise de conjuntos aproximativos.

Ao invés de pré-estabelecer as normas, Slowinski (1992) propõe iniciar a divisão do intervalo do domínio da variável quantitativa em questão em m classes. Naturalmente, o limite superior da classe q é o limite inferior da classe $q+1$. Também é definido um “valor de fronteira” (*threshold*), para a classe q , definido como $t_q(b_q)$, no qual a classe q tem os limites inferior (i) e superior (s) expandidos para:

$$\rho^c(x, q) = \left\{ \rho(x, q) \in \left[b_q^I - t_q(b_q), b_q^S + t_q(b_q) \right] \right\}$$

Equação 30 - Divisão de Intervalos - Variável Quantitativa

Essa expansão de intervalos é definida pelo autor como o código do intervalo q . A cardinalidade total, número máximo de objetos, de $\rho^c(x, q)$ é, no máximo, dois. Se dois objetos pertencerem a uma mesma classe, e a cardinalidade das classes é igual a 1, os objetos possuem indiscernibilidade estrita, ou seja:

$$\rho^c(x, q) = \rho^c(y, q) \quad e \quad \text{card} [\rho^c(x, q)] = \text{card} [\rho^c(y, q)] = 1 \quad .$$

Se $\rho^c(x, q) \cap \rho^c(y, q) \neq \emptyset$ os objetos possuem indiscernibilidade fraca, pois um objeto pertence, ao mesmo tempo, a duas classes distintas com limites expandidos.

Ou seja, na prática, os limites para a classificação de determinado objeto x em uma classe são expandidos para mais ou menos um valor limite pré-estabelecido. Se algum outro objeto y (com seus limites expandidos) encontrar a região do objeto x , e somente y com seus limites expandidos encontrar, é definida uma indiscernibilidade forte. Se houver outros objetos na área expandida, há uma indiscernibilidade fraca. Isto tem certa relação com o nível de consistência l pré-estabelecido em VC-DRSA. Quando se coloca um valor de fronteira muito baixo as chances de indiscernibilidades fracas e fortes diminuem. O mesmo se pode

pensar do índice de consistência l : se for muito próximo de 1, poucos objetos entrarão na aproximação inferior além dos encontrados em DRSA.

5.2 UNIÃO DE ATRIBUTOS NO CRITÉRIO DE DECISÃO

Valores de atributos de decisão poucos frequentes podem representar uma classe não muito importante ou não corretamente definida em meio às outras possibilidades de valores dos atributos. Se estas forem adicionadas aos seus vizinhos imediatos, criando uma nova classificação, algumas inconsistências poderão ser superadas.

Com este objetivo, desenvolve-se aqui uma metodologia baseada na identificação de uma distribuição de probabilidade contínua que se adeque ao histograma original dos valores do atributo de decisão. Com o auxílio do *software Input Analyzer*[®], a melhor distribuição é encontrada.

Se as classes tiverem a mesma participação, uma distribuição uniforme será a mais adequada (pois o histograma desta distribuição é uma reta paralela ao eixo x). Porém, outras distribuições podem representar melhor situações em que se observem concentrações de valores. Nessas distribuições com concentração de valores, a importância das classes será considerada.

Se o sistema possui n objetos e t classes de valores para o atributo de decisão, uma distribuição uniforme sugeriria uma frequência de n/t valores em cada classe. Havendo concentração de valores em determinadas classes, algumas classes ficarão abaixo da frequência n/t . Essas classes serão candidatas a unirem-se às classes vizinhas para formar uma classe com maior representação em termos de frequência.

As distribuições normal e triangular, por exemplo, sugerem que as classes iniciais e finais têm pouca participação na frequência, a primeira com simetria e a segunda podendo representar decréscimo mais suave de um lado (dependendo dos parâmetros da distribuição triangular). Uma distribuição exponencial pode corresponder a classes finais de um dos lados com muito pouca participação.

O critério para que uma classe se una à classe vizinha (imediate superior ou inferior) será o impacto adicional desta classe na função acumulada de probabilidade. Se for uma contribuição pequena, a classe é candidata a esta união.

5.3 IMPACTOS DA UNIÃO DE VALORES NO ATRIBUTO DE DECISÃO

Um exemplo pode mostrar diretamente como um valor *outlier* impacta na qualidade da aproximação em DRSA, mesmo após o uso de técnicas de VC-DRSA.

Suponha um sistema de informações simples com dois atributos condicionantes (a_1 e a_2) e um critério de decisão (dec). Suponha ainda que todos os atributos têm relação de dominância, ou seja, há uma ordenação natural nos valores dos atributos condicionantes e nos do atributo de decisão.

O sistema de informação é resumido na tabela 11:

objeto	Atributo a_1	Atributo a_2	Atributo de Decisão
1	1	1	A
2	1	2	A
3	2	2	A
4	2	3	A
5	2	4	A
6	3	2	A
7	3	3	A
8	4	5	B
9	6	4	B
10	7	7	C
11	8	8	C
12	8	9	C
13	8	8	C
14	8	9	B
15	10	8	C
16	10	9	C

Tabela 11 - Sistema de Informação Original
Fonte: autor.

A tabela 12, de conjuntos P-dominantes (D_{p+}) e P-dominados (D_{p-}), é resumida (sendo P o conjunto formado pelos dois atributos condicionantes A_1 e A_2 – e utilizando as equações 8 e 9):

Obj.	D_{p+} (conj. dominante)	Card	D_{p-} (conj. dominado)	Card
1	{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16}	16	{1}	1
2	{2,3,4,5,6,7,8,9,10,11,12,13,14,15,16}	15	{1,2}	2
3	{3,4,5,6,7,8,9,10,11,12,13,14,15,16}	14	{1,2,3}	3
4	{4,5,7,8,9,10,11,12,13,14,15,16}	12	{1,2,3,4}	4
5	{5,8,9,10,11,12,13,14,15,16}	10	{1,2,3,4,5}	5
6	{6,7,8,9,10,11,12,13,14,15,16}	11	{1,2,3,6}	4
7	{7,8,9,10,11,12,13,14,15,16}	10	{1,2,3,4,6,7}	6
8	{8,10,11,12,13,14,15,16}	8	{1,2,3,4,5,6,7,8}	8
9	{9,10,11,12,13,14,15,16}	8	{1,2,3,4,5,6,7,9}	8
10	{10,11,12,13,14,15,16}	7	{1,2,3,4,5,6,7,8,9,10}	10
11	{11,12,13,14,15,16}	6	{1,2,3,4,5,6,7,8,9,10,11}	11
12	{12,14,16}	3	{1,2,3,4,5,6,7,8,9,10,11,12}	12
13	{13,14,15,16}	4	{1,2,3,4,5,6,7,8,9,10,11,13}	12
14	{14,16}	2	{1,2,3,4,5,6,7,8,9,10,11,12,13,14}	14
15	{15,16}	2	{1,2,3,4,5,6,7,8,9,10,11,13,15}	13
16	{16}	1	{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16}	16

Tabela 12 - Conjuntos Dominantes e Dominados

Fonte: autor.

As classes obtidas a partir do sistema de informações original são:

$$\begin{aligned}
 CI_{\leq a} &= \{1,2,3,4,5,6,7\} \\
 CI_{\leq b} &= \{8,9,10,11,12,13,14,15,16\} \\
 CI_{\geq c} &= \{10,11,12,13,15,16\}
 \end{aligned}$$

Com essas três classes é possível definir todas as aproximações e a qualidade da aproximação. Fazendo o primeiro estudo sem relaxar a entrada de objetos na aproximação inferior, tem-se a seguinte formação de conjuntos (equações 14 e 15):

$$\begin{aligned}
 \underline{P}(CI^{\leq a}) &\Rightarrow \{x \in U : D_p^-(x) \subseteq CI^{\leq a}\} = \{1,2,3,4,5,6,7\} \\
 \overline{P}(CI^{\leq a}) &= \{x \in U : D_p^+(x) \cap CI^{\leq a} \neq \emptyset\} = \{1,2,3,4,5,6,7\} \\
 \underline{P}(CI^{\geq b}) &\Rightarrow \{x \in U : D_p^+(x) \subseteq CI^{\geq b}\} = \{8,9,10,11,12,13,14,15,16\} \\
 \overline{P}(CI^{\geq b}) &= \{x \in U : D_p^-(x) \cap CI^{\geq b} \neq \emptyset\} = \{8,9,10,11,12,13,14,15,16\} \\
 \underline{P}(CI^{\geq c}) &\Rightarrow \{x \in U : D_p^+(x) \subseteq CI^{\geq c}\} = \{15,16\} \\
 \overline{P}(CI^{\geq c}) &= \{x \in U : D_p^-(x) \cap CI^{\geq c} \neq \emptyset\} = \{10,11,12,13,14,15,16\}
 \end{aligned}$$

As regiões de fronteiras (Rf) são, então, estabelecidas (equação 13):

$$Rf(CI^{\leq a}) = \overline{P}(CI^{\leq a}) - \underline{P}(CI^{\leq a}) = \emptyset$$

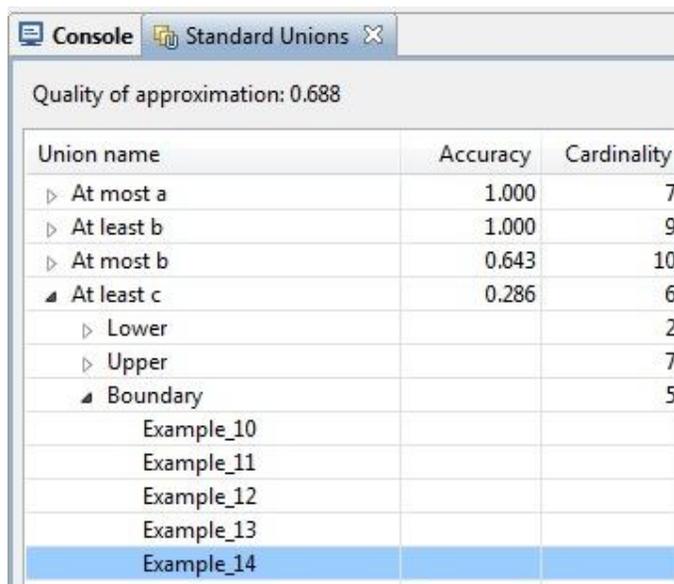
$$Rf(CI^{\geq b}) = \overline{P}(CI^{\geq b}) - \underline{P}(CI^{\geq b}) = \emptyset$$

$$Rf(CI^{\geq c}) = \overline{P}(CI^{\geq c}) - \underline{P}(CI^{\geq c}) = \{10,11,12,13,14\}$$

Como definido por Blaszczynski, Greco e Slowinski (2006), a qualidade da aproximação (equação 16) é dada por:

$$\gamma_P(CI) = \frac{|U - (\cup Rf(CI_i^{\geq}))|}{|U|} = \frac{16-5}{16} = \frac{11}{16} = 0,6875$$

Algumas informações calculadas acima podem ser obtidas no *software* jMAF[®] (o código fonte dos dados para leitura no *software* jMAF[®] está no anexo 3. A figura 11 apresenta as informações calculadas:



The screenshot shows a window titled 'Standard Unions' with a 'Console' tab. It displays the 'Quality of approximation: 0.688'. Below this is a table with three columns: 'Union name', 'Accuracy', and 'Cardinality'. The table lists several unions, with 'Example_14' highlighted in blue.

Union name	Accuracy	Cardinality
▷ At most a	1.000	7
▷ At least b	1.000	9
▷ At most b	0.643	10
▲ At least c	0.286	6
▷ Lower		2
▷ Upper		7
▲ Boundary		5
Example_10		
Example_11		
Example_12		
Example_13		
Example_14		

Figura 11 - Região de fronteira
Fonte: autor.

O índice de entrada de cada um dos objetos pertencentes à região de fronteira pode ser calculado por:

$$\begin{aligned} \text{objeto 10: } & \frac{|D_p^+(10) \cap Cl^{\geq c}|}{|D_p^+(10)|} = \frac{6}{7}; \\ \text{objeto 11: } & \frac{|D_p^+(11) \cap Cl^{\geq c}|}{|D_p^+(11)|} = \frac{5}{6}; \\ \text{objeto 12: } & \frac{|D_p^+(12) \cap Cl^{\geq c}|}{|D_p^+(12)|} = \frac{2}{3}; \\ \text{objeto 13: } & \frac{|D_p^+(13) \cap Cl^{\geq c}|}{|D_p^+(13)|} = \frac{3}{4}; \\ \text{objeto 14: } & \frac{|D_p^+(14) \cap Cl^{\geq c}|}{|D_p^+(14)|} = \frac{1}{2}. \end{aligned}$$

É possível perceber que o objeto 14 não permite um aumento na qualidade da aproximação. Se a entrada for relaxada, é possível aumentar a entrada de objetos em aproximações inferiores. O valor do índice que permite a entrada de cada objeto pertencente à região de fronteira pode ser calculado de acordo com Blaszczynski, Greco e Slowinski (2006) por (equação 19):

$$\begin{aligned} \varepsilon_p^{X_i^{\geq}}(x) &= \max_{y \in D_p^+(x) \cap X_i^{\geq}} \frac{|D_p^+(y) \cap X_i^{\geq}|}{|D_p^+(y)|}; \\ \text{objeto 10: } & \max_{10} \left[\frac{6}{7} \right] = \frac{6}{7}; \\ \text{objeto 11: } & \max_{10,11} \left[\frac{6}{7}; \frac{5}{6} \right] = \frac{6}{7}; \\ \text{objeto 12: } & \max_{10,11,12} \left[\frac{6}{7}; \frac{5}{6}; \frac{2}{3} \right] = \frac{6}{7}; \\ \text{objeto 13: } & \max_{10,11,13} \left[\frac{6}{7}; \frac{5}{6}; \frac{3}{4} \right] = \frac{6}{7}; \\ \text{objeto 14: } & \max_{14} \left[\frac{1}{2} \right] = \frac{1}{2}. \end{aligned}$$

Portanto, se o índice de consistência pré-determinado for menor ou igual a 6/7 (85,71%) os objetos 10, 11, 12 e 13 farão parte das aproximações inferiores, ficando apenas o objeto 14 em uma região de fronteira. O novo índice de qualidade da aproximação será igual a

15/16 (15 objetos bem definidos no sistema original de objetos em um total de 16), ou seja, 93,75%.

A figura 12 representa a situação escrita no parágrafo anterior:

Quality of approximation: 0.938		
Union name	Accuracy	Cardinality
▸ At most a	1.000	7
▸ At least b	1.000	9
▸ At most b	0.900	10
▾ At least c	0.857	6
▸ Lower		6
▸ Upper		7
▾ Boundary		1
Example_14		

Figura 12 - Região de fronteira (VC-DRSA)

Fonte: autor.

Qual seria o impacto de uma possível união de atributos no critério de decisão? Quais motivos poderiam justificar essa união de atributos? É evidente que se a classe pertencente ao objeto 14 for a mesma dos objetos vizinhos a este objeto, o resultado da qualidade da aproximação será diferente.

A figura 13 resume a aplicação de DRSA, ou seja, sem a aplicação do índice de consistência l nos dados do sistema anterior com a substituição dos objetos classificados como classe b para objetos classificados como classe c.

Quality of approximation: 1.000		
Union name	Accuracy	Cardinality
▾ At most a	1.000	7
▸ Lower		7
▸ Upper		7
Boundary		0
▾ At least c	1.000	9
▸ Lower		9
▸ Upper		9
Boundary		0

Figura 13 - Região de fronteira - classes unidas

Fonte: autor.

Portanto, a substituição de uma classe possivelmente mal estabelecida ou pouco frequente pode alterar substancialmente a qualidade da aproximação da aplicação da TCA com relações de dominância. Na sequencia do trabalho, será apresentado um algoritmo para perceber possíveis classes que podem ter seus atributos unidos a classes vizinhas.

5.4 DISTRIBUIÇÃO CONTÍNUA A SER UTILIZADA NO CRITÉRIO DE DECISÃO

5.4.1 Distribuição exponencial

O *software Input Analyzer*[®] permite que se encontre a melhor distribuição contínua que se adequa ao histograma com os dados do critério de decisão. Automaticamente, o *software* gera o p-valor de um teste Qui-quadrado onde a hipótese nula é a de que o ajuste dos dados discretos para uma função de distribuição contínua é satisfatória contra a hipótese alternativa de que o ajuste não é adequado, ou seja:

$$\begin{cases} H_0 = o \text{ ajuste é satisfatório} \\ H_1 = o \text{ ajuste não é satisfatório} \end{cases}$$

Ou seja, a hipótese nula H_0 faz referência sobre a possibilidade das frequências observadas (dados originais) serem iguais às frequências esperadas (função densidade encontrada). Navidi (2006) observa que o teste proposto por Karl Pearson é baseado nas diferenças quadráticas entre as frequências esperadas e observadas dividido pelo numero de observações.

Para aceitação da hipótese nula, o valor das diferenças quadráticas médias deve ser o menor possível. A comparação do valor encontrado é feita de acordo com uma tabela. Se o valor calculado for menor do que o encontrado na tabela, aceita-se a hipótese nula H_0 . O teste qui-quadrado será detalhado nas próximas seções para testar se uma função contínua se adequa à frequência de dados discretos.

Como o p-valor informa o valor mínimo da probabilidade definida a priori para que a hipótese nula seja rejeitada, uma função bem ajustada aos dados discretos gera um valor alto para o p-valor. O objetivo é usar a função sugerida pelo *software*; portanto, espera-se a

aceitação da hipótese nula e, conseqüentemente, um alto valor para o p-valor (nas saídas do *software* “p-value”).

Caso o p-valor seja baixo, não será possível utilizar a função apresentada pelo *software*. No caso de uso da função, o importante será identificar a função densidade de probabilidade e os parâmetros sugeridos pelo *software*. A sugestão do *software* é que o p-valor mínimo seja de 0,10.

Como exemplo do uso da distribuição contínua gerada pelo *Input Analyzer*[®], a figura 14 faz referência ao histograma gerado pela frequência dos valores do atributo de decisão dos objetos apresentados no anexo 5.

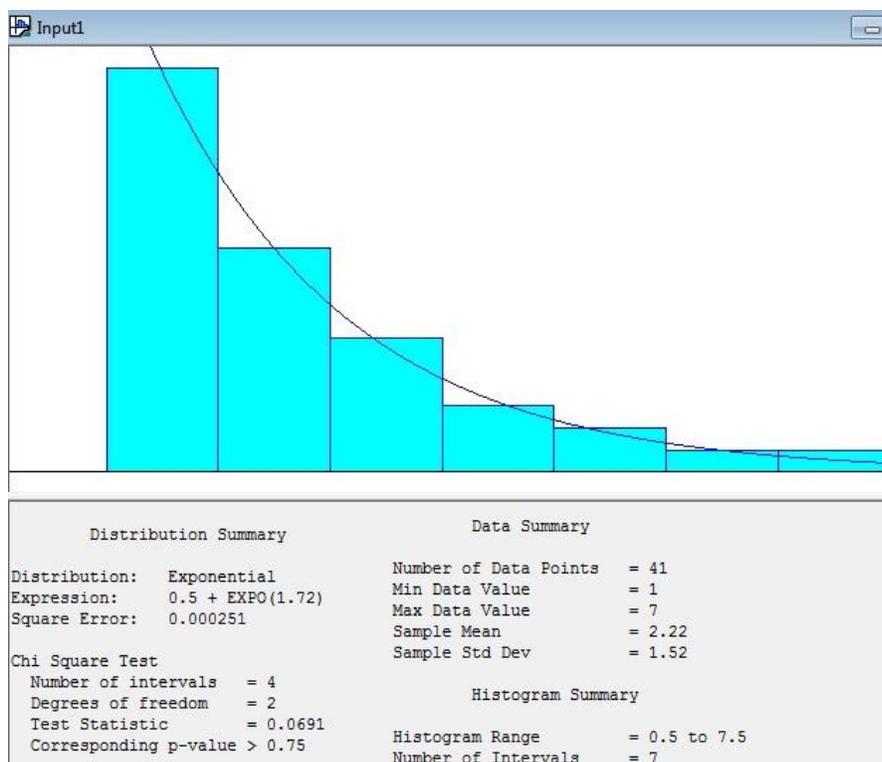


Figura 14 - Aproximação para uma distribuição exponencial
Fonte: autor.

Percebe-se uma grande concentração de dados à esquerda, em consequência de baixas avaliações dos objetos no critério de decisão. Isso sugere uma distribuição com concentração de valores baixos, como é o caso, por exemplo, da exponencial.

Os valores do atributo de decisão variam de 1 até 7. Há um total de 41 observações, a média e o desvio-padrão amostrais são, respectivamente, 2,22 e 1,52. O *software* usa um conceito de aproximação de dados discretos para contínuo semelhante ao apresentado por

Bussab e Morettin (2003) no qual um valor discreto tem a sua área alterada para uma função contínua para uma faixa de valores diminuída e acrescida de 0,5. Desta maneira, há 7 classes no histograma variando, no limite inferior, de 0,5 (menor valor diminuído de 0,5) para, no limite superior, 7,5 (maior valor acrescido de 0,5). Essas observações podem ser visualizadas em “*Data Summary*” e “*Histogram Summary*”.

A distribuição que melhor se adequa aos dados é a exponencial. A expressão sugerida pelo *software* é: $0,5 + \text{Exp}(1,72)$. “*Square Error*” na figura 14 representa a soma dos quadrados dos erros da aproximação, ou seja, das diferenças ao quadrado entre o valor percentual da classe (no seu ponto médio – o próprio valor do atributo de decisão nos dados originais) e o valor da função densidade no ponto médio da classe adequada (a classe adequada será posteriormente definida para esta função). A função apresentada pelo *software* é a que gera o menor somatório de erros quadráticos entre todas as distribuições disponíveis na biblioteca do *Input Analyzer*[®]. O teste qui-quadrado apresentou um p-valor superior a 10% sugerindo que a aproximação das classes discretas pela distribuição contínua exponencial é adequada.

O valor da constante 0,5 na expressão sugere que todos os valores das classes devem ser deslocados para a esquerda, ou seja, diminuídos de 0,5. Assim, na prática, o valor mínimo das classes (que era de 0,5) passa a ser zero. O ponto inicial da primeira classe na distribuição exponencial sempre será zero. Como o valor do menor atributo de decisão é 1 e na mudança para contínua representa o intervalo de 0.5 até 1.5, uma constante equivalente a 0.5 surge na função sugerida pelo *Input Analyzer*. Isso faz com que todas as classes sejam adequadas, como mostra a tabela 13, esta tabela facilita o entendimento das classes dos atributos de decisão:

Classes	Valor Original	Aprox. Cont	Adequação
1	1	0,5 a 1,5	0,0 a 1,0
2	2	1,5 a 2,5	1,0 a 2,0
3	3	2,5 a 3,5	2,0 a 3,0
4	4	3,5 a 4,5	3,0 a 4,0
5	5	4,5 a 5,5	4,0 a 5,0
6	6	5,5 a 6,5	5,0 a 6,0
7	7	6,5 a 7,5	6,0 a 7,0

Tabela 13 - Aproximação de dados discretos para contínuos - Distr. Exponencial
Fonte: autor.

A soma quadrática dos erros, como anteriormente definida, pode ser obtida pela soma dos quadrados das diferenças entre a frequência percentual da classe apresentada no anexo 5 subtraída da função densidade da distribuição contínua, ou seja:

$$\begin{aligned} \sum \text{erros}^2 = & \left[\frac{18}{41} - f(0,5) \right]^2 + \left[\frac{10}{41} - f(1,5) \right]^2 + \left[\frac{6}{41} - f(2,5) \right]^2 + \left[\frac{3}{41} - f(3,5) \right]^2 + \\ & + \left[\frac{2}{41} - f(4,5) \right]^2 + \left[\frac{1}{41} - f(5,5) \right]^2 + \left[\frac{18}{41} - f(6,5) \right]^2 \approx 0,000251 \end{aligned}$$

O parâmetro apresentado pelo software para a exponencial foi 1,72. A distribuição exponencial (NAVIDI, 2006) tem a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}, \text{ a média sugerida pelo software é de 1,72. Na distribuição}$$

exponencial o parâmetro λ é igual a:

$$E(x) = \int_0^{+\infty} x f(x) dx = \int_0^{+\infty} x \lambda e^{-\lambda x} dx = \frac{1}{\lambda} = 1,72 \Rightarrow \lambda = \frac{1}{1,72}.$$

Assim, os erros quadráticos, apresentados anteriormente, podem ser calculados (em valores aproximados) por:

$$\begin{aligned} \sum \text{erros}^2 = & [0,439 - 0,441]^2 + [0,244 - 0,247]^2 + [0,146 - 0,138]^2 + [0,073 - 0,077]^2 + \\ & + [0,049 - 0,043]^2 + [0,024 - 0,024]^2 + [0,0244 - 0,0135]^2 \approx 0,000251 \end{aligned}$$

A função acumulada da distribuição exponencial pode ser definida por:

$$F(x) = P(X \leq x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x} = \begin{cases} 1 - e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Como a exponencial depende somente do parâmetro λ (já conhecido) é possível calcular todas as probabilidades das classes aproximadas pela função exponencial usando a função densidade de probabilidade ou a função acumulada.

Os cálculos, para cada uma das sete classes do atributo de decisão, são apresentados abaixo:

$$\begin{aligned} \text{Classe 1} &\Rightarrow \text{atributo 1} \Rightarrow \text{de 0,5 até 1,5} \xrightarrow{\text{aproximado}} \text{de 0,0 até 1,0} \\ &\Rightarrow F(1,0) - F(0) \Rightarrow 0,441 - 0,000 = 0,441 \end{aligned}$$

$$\begin{aligned} \text{Classe 2} &\Rightarrow \text{atributo 2} \Rightarrow \text{de 1,5 até 2,5} \xrightarrow{\text{aproximado}} \text{de 1,0 até 2,0} \\ &\Rightarrow F(2,0) - F(1,0) \Rightarrow 0,687 - 0,441 = 0,246 \end{aligned}$$

$$\begin{aligned} \text{Classe 3} &\Rightarrow \text{atributo 3} \Rightarrow \text{de 3,5 até 2,5} \xrightarrow{\text{aproximado}} \text{de 2,0 até 3,0} \\ &\Rightarrow F(3,0) - F(2,0) \Rightarrow 0,825 - 0,687 = 0,138 \end{aligned}$$

$$\begin{aligned} \text{Classe 4} &\Rightarrow \text{atributo 4} \Rightarrow \text{de 3,5 até 4,5} \xrightarrow{\text{aproximado}} \text{de 3,0 até 4,0} \\ &\Rightarrow F(4,0) - F(3,0) \Rightarrow 0,902 - 0,825 = 0,077 \end{aligned}$$

$$\begin{aligned} \text{Classe 5} &\Rightarrow \text{atributo 5} \Rightarrow \text{de 5,5 até 4,5} \xrightarrow{\text{aproximado}} \text{de 4,0 até 5,0} \\ &\Rightarrow F(5,0) - F(4,0) \Rightarrow 0,945 - 0,902 = 0,043 \end{aligned}$$

$$\begin{aligned} \text{Classe 6} &\Rightarrow \text{atributo 6} \Rightarrow \text{de 6,5 até 5,5} \xrightarrow{\text{aproximado}} \text{de 5,0 até 6,0} \\ &\Rightarrow F(6,0) - F(5,0) \Rightarrow 0,969 - 0,945 = 0,024 \end{aligned}$$

$$\begin{aligned} \text{Classe 7} &\Rightarrow \text{atributo 7} \Rightarrow \text{de 6,5 até 7,5} \xrightarrow{\text{aproximado}} \text{de 6,0 até 7,0} \\ &\Rightarrow F(7,0) - F(6,0) \Rightarrow 0,983 - 0,969 = 0,014 \end{aligned}$$

Com os dados calculados anteriormente, tem-se o seguinte resumo dos dados na tabela 14:

Utilizando os dados da distribuição exponencial			
Classe	Atributo	Prob. da Classe	Prob. Acum.
1	1	0,441	0,441
2	2	0,246	0,687
3	3	0,138	0,825
4	4	0,077	0,902
5	5	0,043	0,945
6	6	0,024	0,969
7	7	0,014	0,983

Tabela 14 - Probabilidades das classes geradas pela distribuição exponencial
Fonte: autor.

Pela análise da frequência das classes do atributo de decisão, tem-se o seguinte resumo dos dados na tabela 15:

Utilizando os dados discretos				
Classe	Atributo	Freq. Da Classe	Prob. da Classe	Prob. Acum.
1	1	18	18/41=0,439	0,439
2	2	10	10/41=0,244	0,683
3	3	6	6/41=0,146	0,829
4	4	3	3/41=0,0732	0,902
5	5	2	2/41=0,0488	0,951
6	6	1	1/41=0,0244	0,976
7	7	1	1/41=0,0244	1

Tabela 15 - Frequência dos dados originais (discretos) – Distr. Exponencial
Fonte: autor.

Esses valores podem ser observados diretamente pelo software *Input Analyzer*[®] de acordo com a figura 15:

Int. No.	No. of Data Pts.	x	Probability Density		Cumulative Distribution	
			Data	Function	Data	Function
0	18	1.00	0.439	0.441	0.439	0.441
1	10	2.00	0.244	0.247	0.683	0.687
2	6	3.00	0.146	0.138	0.829	0.825
3	3	4.00	0.0732	0.0770	0.902	0.902
4	2	5.00	0.0488	0.0431	0.951	0.945
5	1	6.00	0.0244	0.0241	0.976	0.969
6	1	7.00	0.0244	0.0135	1.00	0.983

Figura 15 - Classes dos dados e da distribuição exponencial
Fonte: autor.

Nota-se uma grande aderência entre os cálculos com dados originais e com valores aproximados pela distribuição exponencial. Essa aderência é refletida na decisão de não rejeitar a hipótese nula do teste de hipótese, ou seja, a hipótese de que a distribuição exponencial é adequada ao histograma original.

5.4.2 Distribuição normal

A figura 16 faz referência ao histograma gerado pela frequência dos valores do atributo de decisão dos objetos apresentados no anexo 1:

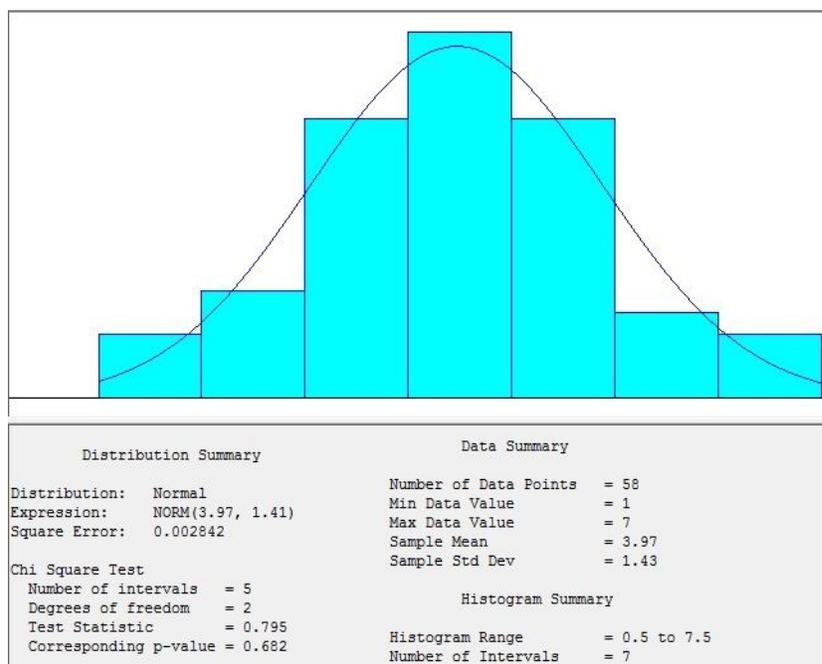


Figura 16 - Aproximação para uma distribuição normal
Fonte: autor.

Pela figura 16, percebe-se uma grande concentração de dados no centro da distribuição. Isso sugere uma distribuição com concentração de valores em torno da média e com simetria, como é o caso, por exemplo, da distribuição normal.

Os valores do atributo de decisão variam de 1 até 7. Há um total de 58 observações, a média e o desvio-padrão amostrais são, respectivamente, 3,97 e 1,43. Um valor discreto, ao ser transformado para contínuo, tem a sua área alterada para uma faixa de valores diminuída e acrescida de 0,5. Desta maneira, há 7 classes no histograma variando, no limite inferior, de 0,5 (menor valor diminuído de 0,5) para, no limite superior, 7,5 (maior valor acrescido de 0,5). Essas observações podem ser visualizadas em “*Data Summary*” e “*Histogram Summary*”.

A distribuição que melhor se adequa aos dados é a normal. A expressão sugerida pelo *software* é: Norm (3,97; 1,41²). “*Square Error*” na figura 16 representa a soma dos quadrados dos erros da aproximação, ou seja, das diferenças ao quadrado entre o valor percentual da classe (no seu ponto médio – o próprio valor do atributo de decisão nos dados originais) e o valor da função densidade no ponto médio da classe adequada (será o próprio valor discreto porque o ponto médio da classe aproximada pela contínua é o próprio ponto). A função apresentada pelo *software* é a que gera o menor somatório de erros quadráticos entre todas as distribuições disponíveis na biblioteca do *Input Analyzer*[®]). O teste qui-quadrado apresentou

um p-valor superior a 10% sugerindo que a aproximação das classes discretas pela distribuição contínua normal é adequada.

Dessa maneira, a tabela 16 facilita o entendimento das classes dos atributos de decisão:

Classes	Valor Original	Aprox. Cont
1	1	0,5 a 1,5
2	2	1,5 a 2,5
3	3	2,5 a 3,5
4	4	3,5 a 4,5
5	5	4,5 a 5,5
6	6	5,5 a 6,5
7	7	6,5 a 7,5

Tabela 16 - Aproximação de dados discretos para contínuos - Distr. Normal
Fonte: autor.

A soma quadrática dos erros, como anteriormente, apresentada, pode ser obtida pela soma das seguintes parcelas quadráticas:

$$\begin{aligned} \sum \text{erros}^2 = & \left[\frac{3}{58} - f(1,0) \right]^2 + \left[\frac{5}{58} - f(2,0) \right]^2 + \left[\frac{13}{58} - f(3,0) \right]^2 + \left[\frac{17}{58} - f(4,0) \right]^2 + \\ & + \left[\frac{13}{58} - f(5,0) \right]^2 + \left[\frac{4}{58} - f(6,0) \right]^2 + \left[\frac{3}{58} - f(7,0) \right]^2 \approx 0,002842 \end{aligned}$$

Os parâmetros apresentados pelo software para a distribuição normal foram 3,97 e 1,43 (para a média e desvio-padrão, respectivamente). A distribuição normal (NAVIDI, 2006) tem a seguinte função densidade de probabilidade:

$$f(x) = \left\{ \frac{1}{\sqrt{2\pi}\delta} \exp\left\{ \frac{-1}{2} \left(\frac{x-\mu}{\delta} \right)^2 \right\} \right\}, \text{ onde } \mu \text{ significa a média e } \delta \text{ o desvio-padrão:}$$

Assim, os erros quadráticos, como apresentados anteriormente, podem ser calculados (em valores aproximados) por:

$$\begin{aligned} \sum \text{erros}^2 = & [0,052 - 0,034]^2 + [0,086 - 0,109]^2 + [0,224 - 0,221]^2 + [0,293 - 0,276]^2 + \\ & + [0,224 - 0,214]^2 + [0,069 - 0,102]^2 + [0,052 - 0,030]^2 \approx 0,002842 \end{aligned}$$

A função acumulada da distribuição normal pode ser definida por:

$$F(x) = P(X \leq x) = \int_0^x f(t) dt$$

Os cálculos, para cada uma das sete classes do atributo de decisão, são apresentados abaixo:

$$\text{Classe 1} \Rightarrow \text{atributo 1} \Rightarrow \text{de 0,5 até 1,5} \xrightarrow{\text{aproximado}} F(1,5) - F(0,5) \Rightarrow 0,040 - 0,007 = 0,033$$

$$\text{Classe 2} \Rightarrow \text{atributo 2} \Rightarrow \text{de 1,5 até 2,5} \xrightarrow{\text{aproximado}} F(2,5) - F(1,5) \Rightarrow 0,149 - 0,040 = 0,109$$

$$\text{Classe 3} \Rightarrow \text{atributo 3} \Rightarrow \text{de 2,5 até 3,5} \xrightarrow{\text{aproximado}} F(3,5) - F(2,5) \Rightarrow 0,369 - 0,149 = 0,220$$

$$\text{Classe 4} \Rightarrow \text{atributo 4} \Rightarrow \text{de 3,5 até 4,5} \xrightarrow{\text{aproximado}} F(4,5) - F(3,5) \Rightarrow 0,646 - 0,369 = 0,277$$

$$\text{Classe 5} \Rightarrow \text{atributo 5} \Rightarrow \text{de 4,5 até 5,5} \xrightarrow{\text{aproximado}} F(5,5) - F(4,5) \Rightarrow 0,861 - 0,646 = 0,215$$

$$\text{Classe 6} \Rightarrow \text{atributo 6} \Rightarrow \text{de 5,5 até 6,5} \xrightarrow{\text{aproximado}} F(6,5) - F(5,5) \Rightarrow 0,964 - 0,861 = 0,103$$

$$\text{Classe 7} \Rightarrow \text{atributo 7} \Rightarrow \text{de 6,5 até 7,5} \xrightarrow{\text{aproximado}} F(7,5) - F(6,5) \Rightarrow 0,994 - 0,964 = 0,030$$

Com os dados calculados anteriormente, tem-se o seguinte resumo dos dados na tabela

17:

Utilizando os dados da distribuição normal			
Classe	Atributo	Prob. da Classe	Prob. Acum.
1	1	0,033	0,040
2	2	0,109	0,149
3	3	0,220	0,369
4	4	0,277	0,646
5	5	0,215	0,861
6	6	0,103	0,964
7	7	0,030	0,994

Tabela 17 - Probabilidades das classes geradas pela distribuição normal

Fonte: autor.

Na tabela 17, o valor acumulado da primeira classe não é o acumulado da própria classe porque é preciso somar a parcela da normal que acumula de $-\infty$ até 0,5. Ou seja, a probabilidade acumulada da primeira classe, na verdade, é o valor que pode ser representado

por $F(1,5)$. Pela análise da frequência das classes do atributo de decisão, tem-se, na tabela 18, o seguinte resumo dos dados:

Utilizando os dados discretos				
Classe	Atributo	Freq. Da Classe	Prob. da Classe	Prob. Acum.
1	1	3	3/58=0,0517	0,0517
2	2	5	5/58=0,0862	0,138
3	3	13	13/58=0,224	0,362
4	4	17	17/58=0,293	0,655
5	5	13	13/58=0,224	0,879
6	6	4	4/58=0,069	0,948
7	7	3	3/58=0,0517	1

Tabela 18 - Frequência dos dados originais (discretos). Distr. Normal
Fonte: autor.

Os valores da tabela 18, podem ser observados diretamente pelo *software Input Analyzer*[®] de acordo com a figura 17:

Int. No.	No. of Data Pts.	x	Probability Density		Cumulative Distribution	
			Data	Function	Data	Function
0	3	1.00	0.0517	0.0335	0.0517	0.0406
1	5	2.00	0.0862	0.109	0.138	0.150
2	13	3.00	0.224	0.221	0.362	0.371
3	17	4.00	0.293	0.276	0.655	0.647
4	13	5.00	0.224	0.214	0.879	0.861
5	4	6.00	0.0690	0.102	0.948	0.963
6	3	7.00	0.0517	0.0303	1.00	0.994

Figura 17 - Classes dos dados e da distribuição normal
Fonte: autor.

Nota-se uma grande aderência entre os cálculos com dados originais e com valores aproximados pela distribuição normal. Essa aderência é refletida na decisão de não rejeitar a hipótese nula do teste de hipótese, ou seja, a hipótese de que a distribuição normal é adequada ao histograma original.

5.4.3 Distribuição beta

A figura 18 faz referência ao histograma gerado pela frequência dos valores do atributo de decisão dos objetos apresentados no anexo 7:

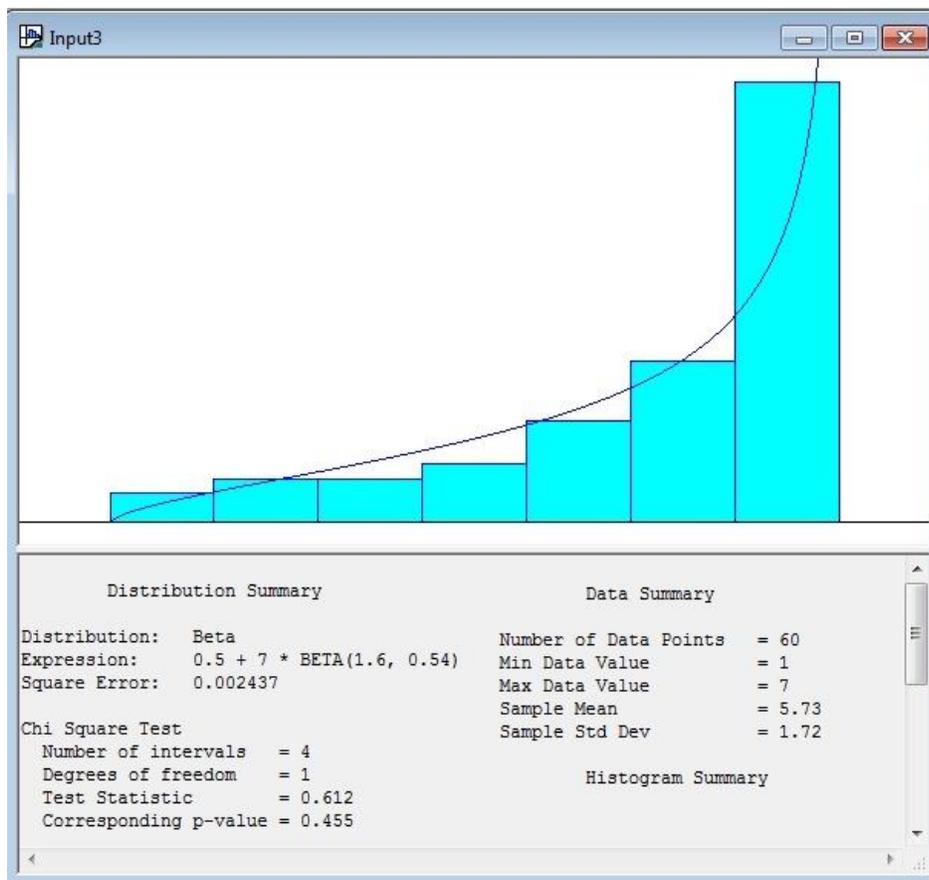


Figura 18 - Aproximação para uma distribuição beta
Fonte: autor.

Na figura 18, percebe-se uma grande concentração de dados à direita da distribuição. Isso sugere uma distribuição com concentração de valores à direita, como é o caso, por exemplo, da distribuição beta com parâmetros alpha e beta adequados.

Os valores do atributo de decisão variam de 1 até 7. Há um total de 60 observações os parâmetros sugeridos pelo *Input Analyzer*[®] para a distribuição beta são: alpha igual a 1,60 e beta igual a 0,54. Um valor discreto, ao ser transformado para contínuo, tem a sua área alterada para uma faixa de valores diminuída e acrescida de 0,5. Desta maneira, há 7 classes no histograma variando, no limite inferior, de 0,5 (menor valor diminuído de 0,5) para, no limite superior, 7,5 (maior valor acrescido de 0,5). Essas observações podem ser visualizadas em “*Data Summary*” e “*Histogram Summary*”.

A distribuição que melhor se adequa aos dados é a beta. A expressão sugerida pelo *software* é: $0,5 + \text{Beta}(1,6 ; 0,54)$. “*Square Error*” na figura 18 representa a soma dos quadrados dos erros da aproximação, ou seja, das diferenças ao quadrado entre o valor percentual da classe (no seu ponto médio – o próprio valor do atributo de decisão nos dados

originais) e o valor da função densidade no ponto médio da classe adequada (a classe adequada será posteriormente definida para esta função). A função apresentada pelo software é a que gera o menor somatório de erros quadráticos entre todas as distribuições disponíveis na biblioteca do *Input Analyzer*[®]. O teste qui-quadrado apresentou um p-valor superior a 10% sugerindo que a aproximação das classes discretas pela distribuição contínua exponencial é adequada.

O valor da constante 0,5 na expressão sugere que todos os valores das classes devem ser deslocados para a esquerda, ou seja, diminuídos de 0,5. Assim, na prática, o valor mínimo das classes (que era de 0,5) passa a ser zero. O ponto inicial da primeira classe na distribuição beta sempre será zero. Como o valor do menor atributo de decisão é 1 e, na mudança para contínua representa o intervalo de 0.5 até 1.5, uma constante equivalente a 0.5 surge na função sugerida pelo *Input Analyzer*[®]. Isso faz com que todas as classes sejam adequadas, como mostra a tabela 19.

A tabela 19 facilita o entendimento das classes dos atributos de decisão:

Classes	Valor Original	Aprox. Cont	Adequação
1	1	0,5 a 1,5	0,0 a 1,0
2	2	1,5 a 2,5	1,0 a 2,0
3	3	2,5 a 3,5	2,0 a 3,0
4	4	3,5 a 4,5	3,0 a 4,0
5	5	4,5 a 5,5	4,0 a 5,0
6	6	5,5 a 6,5	5,0 a 6,0
7	7	6,5 a 7,5	6,0 a 7,0

Tabela 19 - Aproximação de dados discretos para contínuos - Distr. Beta
Fonte: autor.

A soma quadrática dos erros, previamente apresentada, pode ser obtida pela soma das seguintes parcelas quadráticas obtidas da frequência percentual da classe apresentada no anexo 7 subtraído da função densidade da distribuição contínua, ou seja:

$$\begin{aligned} \sum \text{erros}^2 = & \left[\frac{2}{60} - f(0,5) \right]^2 + \left[\frac{3}{60} - f(1,5) \right]^2 + \left[\frac{3}{60} - f(2,5) \right]^2 + \left[\frac{4}{60} - f(3,5) \right]^2 + \\ & + \left[\frac{7}{60} - f(4,5) \right]^2 + \left[\frac{11}{60} - f(5,5) \right]^2 + \left[\frac{30}{60} - f(6,5) \right]^2 \approx 0,002437 \end{aligned}$$

Os parâmetros apresentados pelo *software* para a distribuição beta foram 0,54 e 1,60. A distribuição beta (LARSON, 1982) tem a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, & 0 < x < 1 \\ 0 & , c.c. \end{cases}, \text{ os parâmetros a serem utilizados são:}$$

$\alpha=1,60$ e $\beta=0,54$. A função gama descrita na densidade de probabilidade anterior pode ser definida como:

$$\Gamma(\lambda) = \int_0^{+\infty} e^{-x} x^{\lambda-1} dx, \quad \lambda > 0$$

Assim, os erros quadráticos apresentados anteriormente podem ser calculados (em valores aproximados) por:

$$\begin{aligned} \sum \text{erros}^2 &= [0,033 - 0,021]^2 + [0,05 - 0,047]^2 + [0,050 - 0,070]^2 + [0,067 - 0,096]^2 + \\ &+ [0,117 - 0,131]^2 + [0,183 - 0,188]^2 + [0,500 - 0,528]^2 \approx 0,002437 \end{aligned}$$

Como a distribuição gama depende somente dos parâmetros já estabelecidos é possível calcular todas as probabilidades das classes aproximadas pela a função densidade de probabilidade e usando a definição de função acumula $F(x)$.

Os cálculos, para cada uma das sete classes do atributo de decisão, são apresentados a seguir:

Classe 1 \Rightarrow atributo 1 \Rightarrow de 0,5 até 1,5 $\overset{\text{aproximado}}{\Rightarrow}$ de 0,0 até 1,0

$$\Rightarrow F(1,0) - F(0) \Rightarrow 0,0211 - 0,000 = 0,0211$$

Classe 2 \Rightarrow atributo 2 \Rightarrow de 1,5 até 2,5 $\overset{\text{aproximado}}{\Rightarrow}$ de 1,0 até 2,0

$$\Rightarrow F(2,0) - F(1,0) \Rightarrow 0,0670 - 0,0211 = 0,0460$$

Classe 3 \Rightarrow atributo 3 \Rightarrow de 3,5 até 2,5 $\overset{\text{aproximado}}{\Rightarrow}$ de 2,0 até 3,0

$$\Rightarrow F(3,0) - F(2,0) \Rightarrow 0,1357 - 0,0670 = 0,0687$$

Classe 4 \Rightarrow atributo 4 \Rightarrow de 3,5 até 4,5 $\overset{\text{aproximado}}{\Rightarrow}$ de 3,0 até 4,0

$$\Rightarrow F(4,0) - F(3,0) \Rightarrow 0,2302 - 0,1357 = 0,0945$$

Classe 5 \Rightarrow atributo 5 \Rightarrow de 5,5 até 4,5 $\overset{\text{aproximado}}{\Rightarrow}$ de 4,0 até 5,0

$$\Rightarrow F(5,0) - F(4,0) \Rightarrow 0,3588 - 0,2302 = 0,1286$$

Classe 6 \Rightarrow atributo 6 \Rightarrow de 6,5 até 5,5 $\overset{\text{aproximado}}{\Rightarrow}$ de 5,0 até 6,0

$$\Rightarrow F(6,0) - F(5,0) \Rightarrow 0,5440 - 0,3588 = 0,1852$$

Classe 7 \Rightarrow atributo 7 \Rightarrow de 6,5 até 7,5 $\overset{\text{aproximado}}{\Rightarrow}$ de 6,0 até 7,0

$$\Rightarrow F(7,0) - F(6,0) \Rightarrow 1,00 - 0,544 = 0,456$$

Com os dados calculados anteriormente, tem-se o seguinte resumo dos dados na tabela 20:

Utilizando os dados da distribuição beta			
Classe	Atributo	Prob. da Classe	Prob. Acum.
1	1	0,0211	0,0211
2	2	0,0460	0,0670
3	3	0,0687	0,1357
4	4	0,0945	0,2302
5	5	0,1286	0,3588
6	6	0,1852	0,5440
7	7	0,456	1,00

Tabela 20 - Probabilidades das classes geradas pela distribuição beta
Fonte: autor

Pela análise da frequência das classes do atributo de decisão, tem-se o seguinte resumo dos dados na tabela 21:

Utilizando os dados discretos				
Classe	Atributo	Freq. Da Classe	Prob. da Classe	Prob. Acum.
1	1	2	$2/60=0,033$	0,033
2	2	3	$3/60=0,050$	0,0833
3	3	3	$3/60=0,050$	0,1333
4	4	4	$4/60=0,067$	0,200
5	5	7	$7/60=0,117$	0,317
6	6	11	$11/60=0,188$	0,500
7	7	30	$30/60=0,500$	1,00

Tabela 21 - Frequência dos dados originais (discretos). Distr. Beta
Fonte: autor.

Os valores da tabela 21 podem ser observados diretamente pelo *software Input Analyzer*[®] de acordo com a figura 19:

Int. No.	No. of Data Pts.	x	Probability Density		Cumulative Distribution	
			Data	Function	Data	Function
0	2	1.00	0.0333	0.0214	0.0333	0.0214
1	3	2.00	0.0500	0.0468	0.0833	0.0682
2	3	3.00	0.0500	0.0699	0.133	0.138
3	4	4.00	0.0667	0.0961	0.200	0.234
4	7	5.00	0.117	0.131	0.317	0.365
5	11	6.00	0.183	0.188	0.500	0.553
6	30	7.00	0.500	0.528	1.00	1.08

Figura 19 - Classes dos dados e da distribuição beta
Fonte: autor.

Os dados gerados da função acumulada da distribuição gama podem ser gerados por diversos *softwares* como excel ou matlab. Por exemplo, a figura 20 mostra a obtenção do valor acumulado da classe relativa ao atributo de decisão 4 utilizando o *software* matlab:

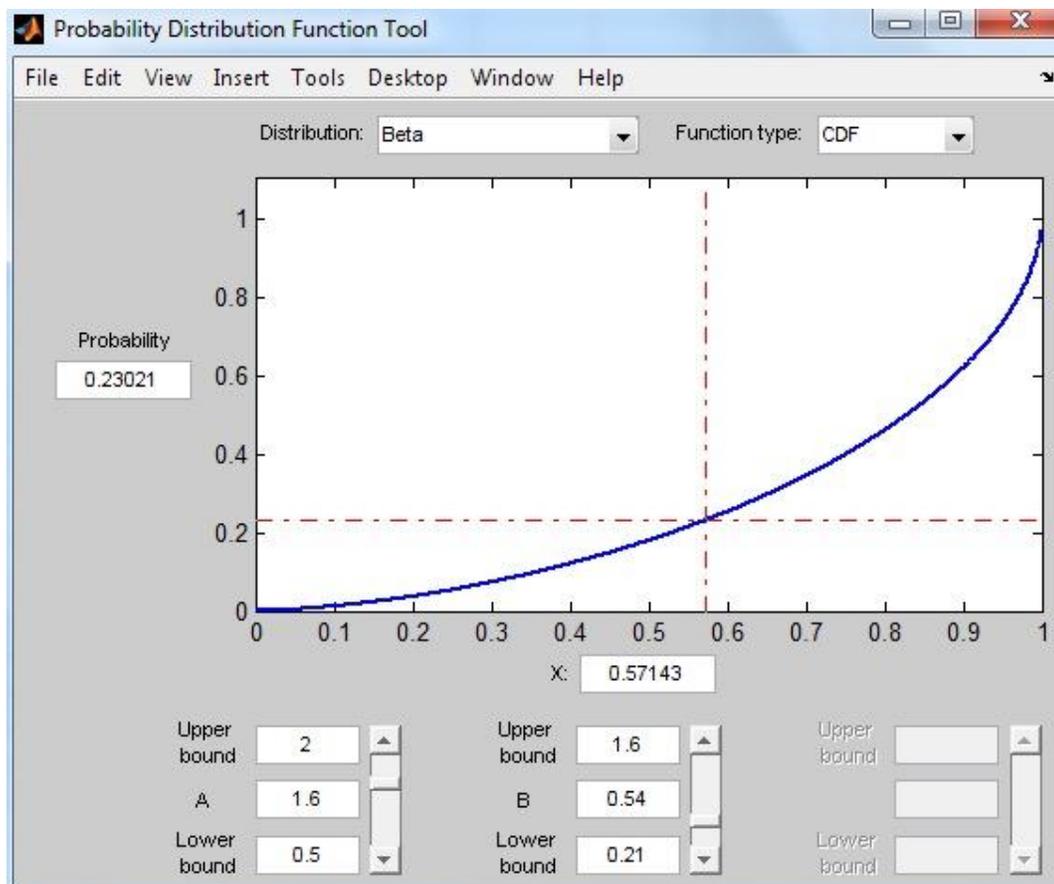


Figura 20 - Função acumulada beta utilizando Matlab
Fonte: autor.

O algoritmo para a obtenção dos valores da função em particular deve ser ampliado.

O domínio da função no *Matlab*[®] é de 0 até 1. O domínio da função estabelecida pelos atributos é de 0 até 7. Portanto, uma transformada simples deve ser feita: Como 4/7 representa 57,14% do domínio, o ponto $x=0$, na função cumulativa apresentada na figura 20, equivale ao acumulado até o ponto 4 em um domínio de 0 até 7. Os demais pontos podem ser obtidos da mesma forma.

Nota-se uma grande aderência entre os cálculos com dados originais e com valores aproximados pela distribuição beta. Essa aderência é refletida na decisão de não rejeitar a hipótese nula do teste de hipótese, ou seja, a hipótese de que a distribuição beta é adequada ao histograma original.

5.4.4 Outras distribuições

Os exemplos apresentados na seção anterior podem ser estendidos para qualquer distribuição contínua. O algoritmo para o cálculo deve ser semelhante aos três exemplos apresentados. Uma distribuição lognormal, por exemplo, pode ser usada para representar uma distribuição com concentração de dados tanto à esquerda como à direita, dependendo do histograma original e dos parâmetros sugeridos pelo *Input Analyzer*[®].

Outra possibilidade a seguir explora a ideia de que os valores observados são de variáveis aleatórias. Isto pode conduzir a substituir, como proposto em Sant'Anna (2002), cada valor do atributo de decisão pela probabilidade de o exemplo respectivo vir a apresentar o maior (ou o menor) valor em uma replicação. A probabilidade de atingir um valor extremo, sendo pequena para todos os valores afastados desse extremo, pode ser considerada igual a zero, de modo que as classes de valores no extremo considerado menos importante são candidatas a união.

6 CONJUNTOS APROXIMATIVOS COM REDUÇÃO DE CLASSES NO ATRIBUTO DE DECISÃO

6.1 PROPOSTA 1 – UNIÃO POR FUNÇÃO DE DENSIDADES DE PROBABILIDADE

A ideia central para a redução de classes do atributo de decisão é a possibilidade de diminuir as inconsistências (mais precisamente, as regiões de fronteiras que surgem pela diferença entre a aproximação superior e a aproximação inferior de cada classe) que prejudicam a qualidade da aproximação para o sistema de informações original. A proposta, apresentada abaixo, baseia-se na união de classes pouco representativas (ou pouco frequentes).

Como apresentado no capítulo 5, o uso de uma função densidade de probabilidades pode representar o histograma dos valores do atributo de decisão, a proposta é a união de classes nas caudas das funções densidade de probabilidade. Por exemplo, uma distribuição que tenha concentração de atributos de dominância à direita, possui a cauda da função densidade de probabilidade à esquerda (ou para valores baixos do atributo de decisão). Uma distribuição que tenha concentração de valores de atributos de dominância à esquerda, possui a cauda da função densidade de probabilidade à direita (ou para valores altos do atributo de decisão).

Se a distribuição tiver uma concentração central, como a distribuição normal, considera-se que a função em questão é bi-caudal e, portanto, as classes sujeitas à união estão tanto à esquerda como à direita.

Da união destas classes caudais surge uma única classe (ou duas se a distribuição tiver concentração central). O pesquisador deve procurar a característica desta nova classe, uma

vez que, dado que o atributo de decisão tem relação de dominância, as classes com as piores (ou melhores) classificações irão ficar unidas.

Desta união de classes, espera-se que as regiões de fronteiras tenham suas cardinalidades diminuídas e, naturalmente, que a qualidade da aproximação aumente. Assim como em VC-DRSA (quando o índice l para a entrada de objetos na aproximação inferior não é fixo) não haverá um único valor fixo para a união de classes. O índice a ser apresentado, nomeado de l^* , poderá ser alterado de acordo com a conveniência do pesquisador.

Da mesma maneira que em VC-DRSA um valor muito baixo de l representa um relaxamento muito grande para a entrada de valores nas aproximações inferiores, nesta proposta um valor alto de l^* possibilita a união de muitas classes. Portanto, o valor de l^* não será fixado, mas um alto valor deste índice pode permitir a união de classes que não necessariamente estejam na cauda da função densidade de probabilidade.

A proposta para a redução de classes no atributo de decisão é apresentada da seguinte forma para os três tipos de concentração:

- Se a concentração de valores for à direita: as frequências baixas ficam à esquerda (ou para valores baixos do atributo de decisão). As classes candidatas à união são as que incluem os valores abaixo do limite $s = s(l^*)$ tal que

$$\int_{-\infty}^s f(x)dx = l^*$$

Equação 31 - Limite para União de Classes Abaixo de s

- Se a concentração de valores for à esquerda: as frequências baixas ficam à direita (ou para valores altos do atributo de decisão). As classes candidatas à união são as que incluem os valores acima do limite $s = s(l^*)$ tal que

$$\int_s^{+\infty} f(x)dx = l^*$$

Equação 32 - Limites para União de Classes Acima de s

- Se a concentração de valores for central: as frequências baixas ficam à esquerda (para valores baixos do atributo de decisão) e à direita (para valores altos do atributo de decisão). As classes candidatas à união são as que incluem os valores

abaixo de um limite $s_1 = s_1(l^*)$ e as que incluem os valores acima de outro limite $s_2 = s_2(l^*)$.

$$\int_{-\infty}^{s_1} f(x)dx = \frac{l^*}{2} \quad e \quad \int_{s_2}^{+\infty} f(x)dx = \frac{l^*}{2}$$

Equação 33 - Limites para União de Classes Abaixo de s_1 e Acima de s_2

O valor para o limite “s” na proposta acima representa, no caso de concentração à direita, o valor limite para que as classes devam ser reunidas. Todos os limites superiores das classes inferiores ao valor de s deverão ter as classes reunidas.

No caso de concentração à esquerda, o valor s representa o valor mínimo que as classes devem ter para que sejam candidatas à união de classes.

Quando a concentração é central, o valor de l^* deve ser dividido em duas partes iguais e o mesmo algoritmo de cálculo deve ser executado para que possa haver união tanto em classes inferiores como em classes superiores.

A redução de classes proposta acima satisfaz as 3 condições de Blaszczynsk *et al.* (2006), apresentadas inicialmente na seção 3.2 deste trabalho.

- $\underline{P}(X) \subseteq X \subseteq \overline{P}(X)$: a aproximação inferior continuará sendo um subconjunto da própria classe de objetos pertencentes a classe de no máximo $(X = Cl_t^{\leq})$ ou a classe de pelo menos $(X = Cl_t^{\geq})$. A aproximação superior continuará a conter a própria classe. Os critérios de entrada nas aproximações não se alteram. Um objeto P-domina e é P-dominado por ele mesmo, portanto o próprio objeto sempre fará parte da aproximação superior. Desta maneira, a classe $(X = Cl_t^{\leq})$ ou a classe $(X = Cl_t^{\geq})$ sempre estará presente na aproximação superior. Como o critério de entrada na aproximação inferior é mais restritivo que na aproximação superior, continua não sendo possível um objeto que não pertence a $(X = Cl_t^{\leq})$ ou a $(X = Cl_t^{\geq})$ fazer parte do conjunto da aproximação inferior, por isso a

aproximação inferior estará sempre contida na classe objetos pertencentes à classe $(X = Cl_t^{\geq})$ ou à classe $(X = Cl_t^{\leq})$.

- $\forall_{X \subseteq U} \underline{P}(X) = U - \overline{P}(U - X)$: a ideia de complementaridade continua a valer mesmo após a união de classes no atributo de decisão. O que irá se alterar é que haverá menos classes e, portanto, menos relações de complementaridade. Porém, todas as complementaridades continuarão a existir, ou seja, a aproximação inferior de um conjunto X é a diferença entre todos os objetos do sistema e a aproximação superior do complemento de X em relação ao universo de objetos – *complementarity*.
- $\forall_{X \subseteq U} \forall_{P, R \subseteq C, P \subseteq R} \underline{P}(X) \subseteq \underline{R}(X) \text{ e } \overline{R}(X) \subseteq \overline{P}(X)$: a monotonicidade da acurácia da aproximação continuará a valer porque não há mudanças de valores nos atributos condicionantes, portanto, os subconjuntos criados P e R (através do conjunto de critérios C) continuarão a ser os mesmos após a mudança do atributo de decisão. As regras para as aproximações inferior e superior continuam as mesmas e, por consequência, as relações de suas aproximações inferior e superior.

6.2 EXEMPLO COM USO DA DISTRIBUIÇÃO NORMAL – PROPOSTA 1

Nesta seção, consideramos um sistema de informações apresentado da seguinte maneira:

- Atributos condicionantes: acabamento, compatibilidade e durabilidade (todos com dominância com valores de 1 a 5);
- Atributo de decisão: avaliação (com dominância e valores de 1 a 7).

A aplicação da TCA foi feita com o auxílio do *software* jMAF[®]. O código fonte para criação do sistema de informações está no Anexo 1 deste artigo. O total de 58 objetos resultou em uma qualidade de aproximação de 0,172 (10 dos 58 objetos são consistentes pelas regras de decisão). A figura 21 mostra a saída do *software* jMAF[®].

Union name	Accuracy	Cardinality
At most 1	0.000	3
At least 2	0.862	55
At most 2	1.000	8
At least 3	1.000	50
At most 3	0.429	21
At least 4	0.652	37
At most 4	0.608	38
At least 5	0.259	20
At most 5	1.000	51
At least 6	1.000	7
At most 6	0.897	55
At least 7	0.000	3

Figura 21 - Aproximação inicial. Distr. Normal
Fonte: autor.

A acurácia (*accuracy*) refere-se à divisão da quantidade de objetos na aproximação inferior pela quantidade de objetos na aproximação superior e foi definida no capítulo 2 como aproximação de imprecisão. A cardinalidade representa o número de objetos contidos em cada união (*union name*).

A figura 21 representa o cálculo das aproximações feitas com índice de consistência l de 1,00. Ou seja, nenhum relaxamento é permitido na entrada de objetos na aproximação inferior. A regra rígida de entrada em qualquer aproximação inferior, estabelecida na equação 11, leva a:

$$\underline{P}(CI_t^{\geq}) = \{x \in U : Dp^+(x) \subseteq CI_t^{\geq}\} \text{ e } \underline{P}(CI_t^{\leq}) = \{x \in U : Dp^-(x) \subseteq CI_t^{\leq}\}$$

Ou seja, quando todos os objetos (dominantes ou dominados) estão contidos na classe ao menos t ou pelo menos t .

Se for utilizado o algoritmo de relaxamento de entrada de objetos na aproximação inferior através da utilização do índice $l=0,85$, a qualidade da aproximação irá aumentar, pois o critério de entrada na aproximação inferior foi relaxado (o algoritmo de cálculo para a entrada de objetos na aproximação inferior foi apresentado no capítulo 4). O novo valor de aproximação encontrado está na figura 22:

Quality of approximation: 0.621		
Union name	Accuracy	Cardinality
▷ At most 1	0.000	3
▷ At least 2	0.948	55
▷ At most 2	1.000	8
▷ At least 3	1.000	50
▷ At most 3	0.818	21
▷ At least 4	0.900	37
▷ At most 4	0.745	38
▷ At least 5	0.478	20
▷ At most 5	1.000	51
▷ At least 6	1.000	7
▷ At most 6	0.948	55
▷ At least 7	0.000	3

Figura 22 - Aproximação com índice de consistência. Distr. Normal
Fonte: autor.

Agora, 36 objetos (ou 62,1% do total de 58 objetos) são capazes de gerar regras consistentes com o índice pré-estabelecido de 0,85.

A análise gráfica dos valores do atributo de decisão pode permitir a união de algumas classes para aumentar a qualidade da aproximação. O histograma e a distribuição mais adequada ao histograma foram apresentados no capítulo anterior. A distribuição sugerida foi normal com média 3,97 e desvio-padrão 1,41.

Como a distribuição tem concentração de valores centrais (em torno da média) a suposição é a de que as classes nos extremos são candidatas a se unirem. Um nível pré-determinado de l^* deve ser estabelecido. Serão usados os seguintes parâmetros: nível de consistência $l=0,85$ (com 15 % de “relaxamento” para entrada de objetos na aproximação inferior) e l^* igual a 0,30 (um relaxamento de 30% na união de classes).

Por se tratar de uma distribuição normal, os limites para união de classes são determinados por (utilizando a equação 33):

$$\int_{-\infty}^{s_1} f(x)dx = \frac{0,30}{2} = 0,15 \quad e$$

$$\int_{s_2}^{+\infty} f(x)dx = \frac{0,30}{2} = 0,15$$

Utilizando informações da distribuição normal e dos parâmetros estabelecidos de média e desvio-padrão, têm-se os seguintes valores como limites de união de classes:

$$\int_0^{2,51} f(x)dx = \frac{0,30}{2} = 0,15 \quad e$$

$$\int_{5,43}^{+\infty} f(x)dx = \frac{0,30}{2} = 0,15$$

Desta maneira, as classes com limites superiores inferiores a 2,51 devem ser unidas e com limites inferiores maiores que 5,43 também devem unir-se.

Portanto, as classes 1 e 2 (limite superior equivalente a 2,5) e as classe 6 (limite inferior equivalente a 5,5) e 7 devem unir-se. Sendo assim, as novas classes a serem estudadas pelo pesquisador são:

- a: representando as classes 1 e 2 originais (baixa avaliação geral);
- b: representando a classe 3 original;
- c: representado a classes 4 original;
- d: representando a classe 5 original;
- e: representando as classes 6 e 7 originais (alta avaliação geral).

O novo código fonte está representado no Anexo 2. A qualidade da aproximação, já com o índice de consistência de 0,85, passou de 0,621 (36 em 58) para 0,724 (42 em 58) conforme a figura 23:

Quality of approximation: 0.724		
Union name	Accuracy	Cardinality
▸ At most a	1.000	8
▸ At least b	1.000	50
▸ At most b	0.818	21
▸ At least c	0.900	37
▸ At most c	0.745	38
▸ At least d	0.478	20
▸ At most d	1.000	51
▸ At least e	1.000	7

Figura 23 - Qualidade da aproximação com dados alterados. Distr. Normal
Fonte: autor.

A cardinalidade 8 na figura 23 representa o conjunto das classes 1 e 2 no sistema de informação anterior. O mesmo raciocínio pode ser estendido à classe e (com cardinalidade igual a 7 que representa a união das classes 6 e 7 do sistema de informações anterior).

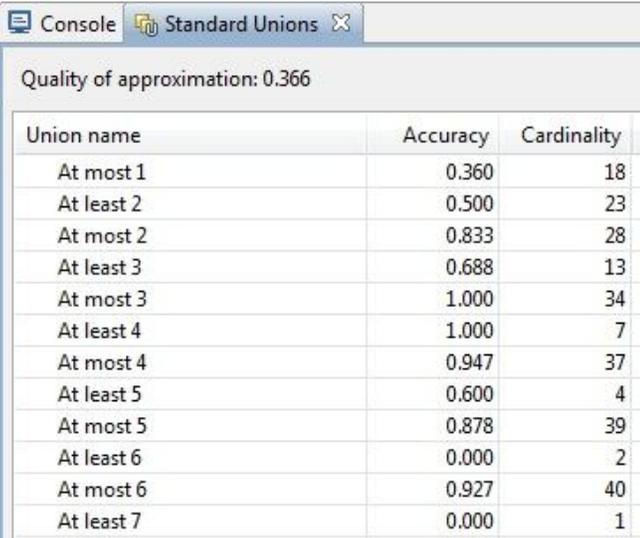
Assim como no sistema anterior, o core é formado pelos três atributos, portanto, a exclusão de qualquer atributo impactará na qualidade da aproximação. O core e possíveis reduções não são alterados pela união de atributos no critério de decisão porque não há alteração de valores nos atributos condicionantes.

6.3 EXEMPLO COM USO DA DISTRIBUIÇÃO EXPONENCIAL – PROPOSTA 1

Nesta seção, considera-se um sistema de informações apresentado da seguinte maneira:

- Atributos condicionantes: acabamento, compatibilidade e durabilidade (todos com dominância com valores de 1 a 5);
- Atributo de decisão: avaliação (com dominância e valores de 1 a 7).

A aplicação da TCA foi feita com o auxílio do *software* jMAF[®]. O código fonte para criação do sistema de informações está no anexo 5 deste artigo. O total de 41 objetos resultou em uma qualidade de aproximação de 0,366 (15 dos 41 objetos são consistentes pelas regras de decisão). A figura 24 mostra a saída do *software* jMAF[®].



Quality of approximation: 0.366		
Union name	Accuracy	Cardinality
At most 1	0.360	18
At least 2	0.500	23
At most 2	0.833	28
At least 3	0.688	13
At most 3	1.000	34
At least 4	1.000	7
At most 4	0.947	37
At least 5	0.600	4
At most 5	0.878	39
At least 6	0.000	2
At most 6	0.927	40
At least 7	0.000	1

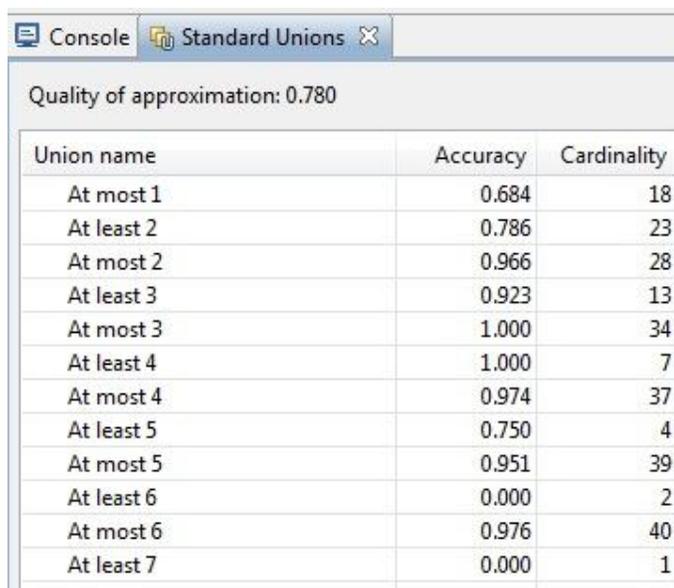
Figura 24 - Qualidade da aproximação inicial. Distr. Exponencial
Fonte: autor.

A acurácia (*accuracy*) refere-se à divisão da quantidade de objetos na aproximação inferior pela quantidade de objetos na aproximação superior e foi definida no capítulo 2 como aproximação de imprecisão. A cardinalidade representa o número de objetos contidos em cada união (*union name*).

A figura 24 representa o cálculo das aproximações feitas com índice de consistência l de 1,00. Ou seja, nenhum relaxamento é permitido na entrada de objetos na aproximação inferior. A regra rígida de entrada em qualquer aproximação inferior, estabelecida na equação 11, leva a:

$\underline{P}(CI_t^{\geq}) = \{x \in U : Dp^+(x) \subseteq CI_t^{\geq}\}$ e $\underline{P}(CI_t^{\leq}) = \{x \in U : Dp^-(x) \subseteq CI_t^{\leq}\}$, ou seja, quando todos os objetos (dominantes ou dominados) estão contidos na classe ao menos t ou pelo menos t .

Se for utilizado o algoritmo de relaxamento de entrada de objetos na aproximação inferior através da utilização do índice $l=0,85$. A qualidade da aproximação irá aumentar, pois o critério de entrada na aproximação inferior foi relaxado (o algoritmo de cálculo para a entrada de objetos na aproximação inferior foi apresentado no capítulo 4). O novo valor de aproximação encontrado está na figura 25:



Union name	Accuracy	Cardinality
At most 1	0.684	18
At least 2	0.786	23
At most 2	0.966	28
At least 3	0.923	13
At most 3	1.000	34
At least 4	1.000	7
At most 4	0.974	37
At least 5	0.750	4
At most 5	0.951	39
At least 6	0.000	2
At most 6	0.976	40
At least 7	0.000	1

Figura 25 - Aproximação com índice de consistência. Distr. Exponencial
Fonte: autor.

Agora, 32 objetos (ou 78,0% do total de 41 objetos) são capazes de gerar regras consistentes com o índice pré-estabelecido de 0,85.

A análise gráfica dos valores do atributo de decisão pode permitir a união de algumas classes para aumentar a qualidade da aproximação. O histograma e a distribuição mais adequada ao histograma foram apresentados no capítulo anterior. A distribuição sugerida foi exponencial com média 1,72.

Como a distribuição tem concentração de valores à esquerda (maior frequência para valores baixos para o atributo de decisão) a suposição é a de que as classes no extremo superior são candidatas a se unirem. Um nível pré-determinado de l^* deve ser estabelecido. Serão usados os seguintes parâmetros: nível de consistência $l=0,85$ (com 15 % de “relaxamento” para entrada de objetos na aproximação inferior) e l^* igual a 0,31 (um relaxamento de 31% na união de classes).

Por se tratar de uma distribuição exponencial, os limites para união de classes são determinados por (utilizando a equação 32):

$$\int_s^{+\infty} f(x)dx = l^*$$

Utilizando informações da distribuição exponencial e do parâmetro estabelecido de média, tem-se o seguinte valor como limite de união de classes:

$$\int_{2,00}^{+\infty} f(x)dx = 0,31;$$

Desta maneira, as classes superiores com limites inferiores maiores do que 2,00 devem ser unidas.

Portanto, as classes 3 (classe com limite inferior igual a 2,0), 4 (classe superior com limite inferior equivalente a 3,0), a classe 5 (classe superior com limite inferior equivalente a 4,0), a classe 6 (classe superior com limite inferior equivalente a 5,0) e a classe 7 (classe superior com limite inferior equivalente a 6,0) devem unir-se. Sendo assim, as novas classes a serem estudadas pelo pesquisador são:

- a: representando a classe 1 original;
- b: representando a classe 2 original;
- c: representado as classes 3,4,5,6 e 7 originais (alta avaliação geral).

O novo código fonte está representado no Anexo 6. A qualidade da aproximação, já com o índice de consistência de 0,85, passou de 0,78 (32 em 41) para 0,829 (34 em 41) conforme a figura 26:

Quality of approximation: 0.829		
Union name	Accuracy	Cardinality
▷ At most a	0.684	18
▷ At least b	0.786	23
▷ At most b	0.966	28
▷ At least c	0.923	13

Figura 26 - Qualidade da aproximação com dados alterados. Distr. Exponencial
Fonte: autor.

A cardinalidade 13 na figura 26 representa o conjunto das classes 3,4,5,6 e 7 no sistema de informação anterior. Assim como no sistema anterior, o core é formado pelos atributos acabamento e durabilidade. Portanto, a exclusão de qualquer um destes dois atributos impactará na qualidade da aproximação. O core (e possíveis reduções) não são alterados pela união de atributos no critério de decisão porque não há alteração de valores nos atributos condicionantes. Neste exemplo, o atributo condicionante - compatibilidade pode ser retirado do sistema que não impactará na qualidade de aproximação.

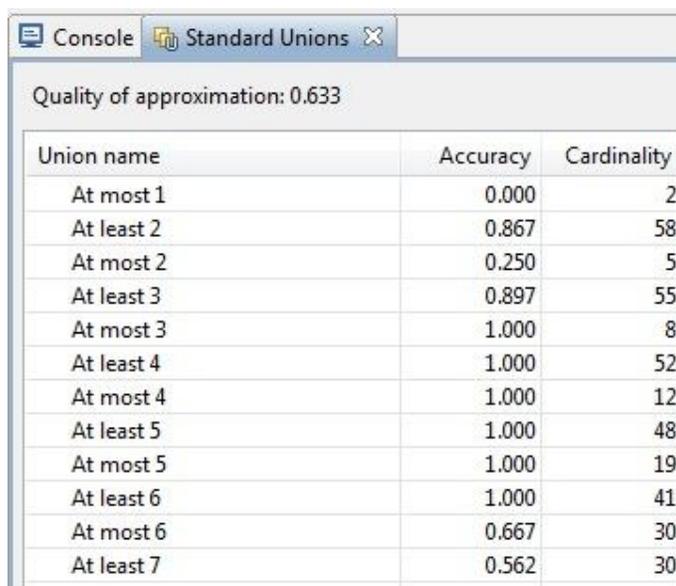
A qualidade da aproximação aumentou de 0,789 para 0,829 com o relaxamento na entrada de objetos na aproximação inferior e a união das classes {3,4,5,6,7}. Se o nível de consistência l diminuísse ainda mais, a qualidade da aproximação aumentaria, porém os critérios para a entrada de objetos na aproximação inferior perderiam força. O mesmo pode ser pensado para o nível l^* de união de classes para o atributo de decisão.

6.4 EXEMPLO COM USO DA DISTRIBUIÇÃO BETA – PROPOSTA 1

Nesta seção, considera-se um sistema de informações apresentado da seguinte maneira:

- Atributos condicionantes: acabamento, compatibilidade e durabilidade (todos com dominância com valores de 1 a 5);
- Atributo de decisão: avaliação (com dominância e valores de 1 a 7).

A aplicação da TCA foi feita com o auxílio do *software* jMAF[®]. O código fonte para criação do sistema de informações está no anexo 7 deste artigo. O total de 60 objetos resultou em uma qualidade de aproximação de 0,633 (38 dos 60 objetos são consistentes pelas regras de decisão). A figura 27 mostra a saída do *software* jMAF[®]:



Quality of approximation: 0.633		
Union name	Accuracy	Cardinality
At most 1	0.000	2
At least 2	0.867	58
At most 2	0.250	5
At least 3	0.897	55
At most 3	1.000	8
At least 4	1.000	52
At most 4	1.000	12
At least 5	1.000	48
At most 5	1.000	19
At least 6	1.000	41
At most 6	0.667	30
At least 7	0.562	30

Figura 27 - Qualidade da aproximação inicial. Distr. Beta
Fonte: autor.

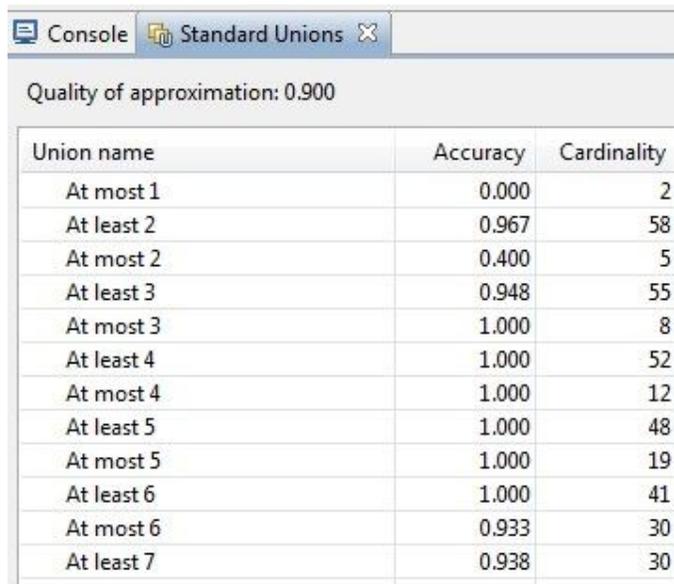
A acurácia (*accuracy*) refere-se à divisão da quantidade de objetos na aproximação inferior pela quantidade de objetos na aproximação superior e foi definida no capítulo 2 como aproximação de imprecisão. A cardinalidade representa o número de objetos contidos em cada união (*union name*).

A figura anterior representa o cálculo das aproximações feitas com índice de consistência l de 1,00. Ou seja, nenhum relaxamento é permitido na entrada de objetos na aproximação inferior. A regra rígida de entrada em qualquer aproximação inferior, estabelecida na equação 11, leva a:

$$\underline{P}(CI_t^{\geq}) = \{x \in U : Dp^+(x) \subseteq CI_t^{\geq}\}$$
 e
$$\underline{P}(CI_t^{\leq}) = \{x \in U : Dp^-(x) \subseteq CI_t^{\leq}\}$$
, ou seja, quando todos os objetos (dominantes ou dominados) estão contidos na classe ao menos t ou pelo menos t .

Se for utilizado o algoritmo de relaxamento de entrada de objetos na aproximação inferior através da utilização do índice $l=0,85$. A qualidade da aproximação irá aumentar, pois

o critério de entrada na aproximação inferior foi relaxado (o algoritmo de cálculo para a entrada de objetos na aproximação inferior foi apresentado no capítulo 4). O novo valor de aproximação encontrado está na figura 28:



Union name	Accuracy	Cardinality
At most 1	0.000	2
At least 2	0.967	58
At most 2	0.400	5
At least 3	0.948	55
At most 3	1.000	8
At least 4	1.000	52
At most 4	1.000	12
At least 5	1.000	48
At most 5	1.000	19
At least 6	1.000	41
At most 6	0.933	30
At least 7	0.938	30

Figura 28 - Aproximação com índice de consistência. Distr. Beta
Fonte: autor.

Agora, 54 objetos (ou 90,0% do total de 60 objetos) são capazes de gerar regras consistentes com o índice pré-estabelecido de 0,85.

A análise gráfica dos valores do atributo de decisão pode permitir a união de algumas classes para aumentar a qualidade da aproximação. O histograma e a distribuição mais adequada ao histograma foram apresentados no capítulo anterior. A distribuição sugerida foi beta com parâmetros alpha igual a 1,60 e beta igual a 0,54.

Como a distribuição tem concentração de valores à direita (maior frequência para valores altos para o atributo de decisão) a suposição é a de que as classes no extremo inferior são candidatas a se unirem. Um nível pré-determinado de l^* deve ser estabelecido. Serão usados os seguintes parâmetros: nível de consistência $l=0,85$ (com 15 % de “relaxamento” para entrada de objetos na aproximação inferior) e l^* igual a 0,15 (um relaxamento de 15% na união de classes).

Por se tratar de uma distribuição beta (com concentração de valores à direita), os limites para união de classes são determinados por (utilizando a equação 31):

$$\int_0^s f(x)dx = l^*$$

Utilizando informações da distribuição beta e dos parâmetros estabelecidos para alpha e beta, tem-se o seguinte valor como limite de união de classes:

$$\int_0^{3,172} f(x)dx = 0,15 ;$$

Desta maneira, as classes com limites superiores menores do que 3,172 devem ser unidas.

Portanto, as classes 1 (limite superior equivalente a 1,0), a classe 2 (limite superior equivalente a 2,0) e a classe 3 (limite superior equivalente a 3,0) devem unir-se. Sendo assim, as novas classes a serem estudadas pelo pesquisador são:

- a: representando as classes 1,2,3 originais;
- b: representando a classe 4 original;
- c: representando a classe 5 original;
- d: representando 6 original;
- e: representando a classe 7 original.

O novo código fonte está representado no Anexo 8. A qualidade da aproximação, já com o índice de consistência de 0,85, passou de 0,90 (54 em 60) para 0,967 (58 em 60) conforme a figura 29:

Union name	Accuracy	Cardinality
At most a	1.000	8
At least b	1.000	52
At most b	1.000	12
At least c	1.000	48
At most c	1.000	19
At least d	1.000	41
At most d	0.933	30
At least e	0.938	30

Figura 29 - Qualidade da aproximação com dados alterados. Distr. Beta
Fonte: autor.

A cardinalidade 8 na figura representa o conjunto das classes $\{1,2,3\}$ no sistema de informação anterior. Assim como no sistema anterior, o core é formado pelos três atributos acabamento, compatibilidade e durabilidade, portanto, a exclusão de qualquer um destes três atributos impactará na qualidade da aproximação. O core (e possíveis reduções) não são alterados pela união de atributos no critério de decisão porque não há alteração de valores nos atributos condicionantes.

A qualidade da aproximação aumentou de 0,633 para 0,90 com o relaxamento na entrada de objetos na aproximação inferior. Se o nível de consistência l diminuísse ainda mais, a qualidade da aproximação aumentaria, porém os critérios para a entrada de objetos na aproximação inferior perderiam força.

O mesmo raciocínio pode ser estendido para o critério de união de classes que se encontram nas caudas da distribuição sugerida, quanto maior o índice usado, maiores as chances de união das classes e, conseqüentemente, maior será a qualidade da aproximação. Utilizando um l^* de 0,31 as classes $\{1,2,3\}$ uniram-se, aumentando a qualidade da aproximação de 0,90 para 0,967. Deve-se buscar um equilíbrio entre o valor do índice usado e as classes candidatas à união. Em um limite, se l^* for igual a 1, todas as classes se unirão, o que não auxiliará o pesquisador, pois com somente um valor no critério de decisão perde o sentido a aplicação da teoria dos conjuntos aproximativos.

6.5 DIFERENÇAS ENCONTRADAS NAS CLASSES UNIDAS – PROPOSTA 1

Será agora verificado o impacto da união das classes em cada um dos exemplos apresentados anteriormente. A hipótese de unir algumas classes pouco representativas no atributo de decisão permite um aumento na qualidade da aproximação devido ao menor número de valores no próprio atributo de decisão, o que tende a causar menos inconsistências entre todas as classificações. Em todos os exemplos que serão apresentados, a análise será feita em cima dos sistemas de informação (original e transformado com redução de atributos do atributo de decisão) com a utilização de um índice de consistência igual a 0,85.

O motivo para a escolha de analisar apenas utilizando VC-DRSA é a quantidade menor de objetos nas regiões de fronteira que impactam diretamente a qualidade da aproximação. Porém, a extensão da análise com um índice de consistência igual a 1 segue os mesmos princípios.

6.5.1 Exemplo usando distribuição normal

No exemplo da distribuição normal com o sistema de dados original a qualidade de aproximação obtida foi de 62,1%, ou seja, 36 objetos definem bem as regras e o restante, $58-36=22$ objetos, fazem parte da região de fronteira e não são bem definidos no sistema de informações. Como visto anteriormente, a qualidade da aproximação em técnicas utilizando o conceito de dominância é diretamente relacionada aos objetos que formam as regiões de fronteira. A figura 30 ilustra as regiões de fronteira das classes CI^{\geq} (o mesmo valor da qualidade de aproximação seria encontrado se fossem utilizada as regiões de fronteira das classes CI^{\leq}).

Para a distribuição normal, com o sistema de informações original, as regiões de fronteiras geradas pelo *software* jMAF[®] são apresentadas:

Quality of approximation: 0.621		
Union name	Accuracy	Cardinality
▸ At most 1	0.000	3
▲ At least 2	0.948	55
▸ Lower		55
▸ Upper		58
▲ Boundary		3
Example_1		
Example_2		
Example_3		
▸ At most 2	1.000	8
▲ At least 3	1.000	50
▸ Lower		50
▸ Upper		50
Boundary		0
▸ At most 3	0.818	21
▲ At least 4	0.900	37
▸ Lower		36
▸ Upper		40
▲ Boundary		4
Example_11		
Example_13		
Example_16		
Example_23		
▸ At most 4	0.745	38
▲ At least 5	0.478	20
▸ Lower		11
▸ Upper		23
▲ Boundary		12
Example_25		
Example_28		
Example_35		
Example_39		
Example_41		
Example_42		
Example_44		
Example_45		
Example_47		
Example_49		
Example_50		
Example_51		
▸ At most 5	1.000	51
▲ At least 6	1.000	7
▸ Lower		7
▸ Upper		7
Boundary		0
▸ At most 6	0.948	55
▲ At least 7	0.000	3
Lower		0
▸ Upper		3
▲ Boundary		3
Example_56		
Example_57		
Example_58		

Figura 30 - Regiões de Fronteira - Distr. Normal - Sistema de Informações Original
Fonte: autor.

As classes destacadas na figura 30 são:

$$CI^{\geq 2} = \{4,5,6,7,8,9,\dots,58\}$$

$$CI^{\geq 3} = \{9,10,11,12,\dots,58\}$$

$$CI^{\geq 4} = \{22,23,24,25,\dots,58\}$$

$$CI^{\geq 5} = \{39,40,41,42,\dots,58\}$$

$$CI^{\geq 6} = \{52,53,54,55,56,57,58\}$$

$$CI^{\geq 7} = \{56,57,58\}$$

A união de todas as fronteiras destacadas e a sua cardinalidade são (equação 13):

$$\begin{aligned} \cup Rf(CI_i^{\geq}) &= Rf(CI^{\geq 2}) + Rf(CI^{\geq 3}) + Rf(CI^{\geq 4}) + Rf(CI^{\geq 5}) + Rf(CI^{\geq 6}) + Rf(CI^{\geq 7}) = \\ &= \{1,2,3\} + \{ \} + \{11,13,16,23\} + \{25,28,35,39,41,42,44,45,47,49,50,51\} + \{ \} + \{56,57,58\} = \\ &= \{1,2,3,11,13,16,23,25,28,35,39,41,42,44,45,47,49,50,51,56,57,58\} \\ |\cup Rf(CI_i^{\geq})| &= 22 \end{aligned}$$

Com a cardinalidade da união das regiões de fronteiras formada pelas classes CI_i^{\geq} a qualidade da aproximação é dada por (equação 16):

$$\gamma(CI) = \frac{|U - (\cup Rf(CI_i^{\geq}))|}{|U|} = \frac{58 - 22}{58} = 0,621$$

O resultado anterior já era conhecido. A informação adicional que pode ser verificada aqui é quais objetos pertencem às regiões de fronteira causando impacto na qualidade da aproximação. As classes originais foram unidas no sistema de informações alterado (por se tratar de uma distribuição normal, a proposta foi a união das duas classes à direita e das duas classes à esquerda do histograma original).

$$CI^a = CI^1 \cup CI^2$$

$$CI^b = CI^3$$

$$CI^c = CI^4$$

$$CI^d = CI^5$$

$$CI^e = CI^6 \cup CI^7$$

$$CI^{\geq b} = \{9,10,11,12,\dots,58\}$$

$$CI^{\geq c} = \{22,23,24,25,\dots,58\}$$

$$CI^{\geq d} = \{39,40,41,42,\dots,58\}$$

$$CI^{\geq e} = \{52,53,54,55,56,57,58\}$$

Para a distribuição normal com o sistema de informações alterado as regiões de fronteira geradas pelo *software* jMAF[®] são:

Quality of approximation: 0.724		
Union name	Accuracy	Cardinality
▷ At most a	1.000	8
▲ At least b	1.000	50
▷ Lower		50
▷ Upper		50
Boundary		0
▷ At most b	0.818	21
▲ At least c	0.900	37
▷ Lower		36
▷ Upper		40
▲ Boundary		4
Example_11		
Example_13		
Example_16		
Example_23		
▷ At most c	0.745	38
▲ At least d	0.478	20
▷ Lower		11
▷ Upper		23
▲ Boundary		12
Example_25		
Example_28		
Example_35		
Example_39		
Example_41		
Example_42		
Example_44		
Example_45		
Example_47		
Example_49		
Example_50		
Example_51		
▷ At most d	1.000	51
▲ At least e	1.000	7
▷ Lower		7
▷ Upper		7
Boundary		0

Figura 31 - Regiões de Fronteira - Distr. Normal - Sistema de Informações Transformado
Fonte: autor.

As classes destacadas na figura 31 são:

$$CI^{\geq b} = \{9,10,11,12,13,\dots,58\}$$

$$CI^{\geq c} = \{22,23,24,25,26,\dots,58\}$$

$$CI^{\geq d} = \{39,40,41,42,\dots,58\}$$

$$CI^{\geq e} = \{52,53,54,55,56,57,58\}$$

A união de todas as fronteiras destacadas e a sua cardinalidade são (equação 12):

$$\begin{aligned} \cup Rf(CI_i^{\geq}) &= Rf(CI^{\geq b}) + Rf(CI^{\geq c}) + Rf(CI^{\geq d}) + Rf(CI^{\geq e}) = \\ &= \{ \} + \{11,13,16,23\} + \{25,28,35,39,41,42,44,45,47,49,50,51\} + \{ \} = \\ &= \{11,13,16,23,25,28,35,39,41,42,44,45,47,49,50,51\} \\ |\cup Rf(CI_i^{\geq})| &= 16 \end{aligned}$$

Com a cardinalidade da união das regiões de fronteiras formada pelas classes CI_i^{\geq} a qualidade da aproximação é dada por (equação 16):

$$\gamma(CI) = \frac{|U - (\cup Rf(CI_i^{\geq}))|}{|U|} = \frac{58 - 16}{58} = 0,724$$

O resultado anterior já era conhecido. A informação adicional que pode ser verificada aqui é quais objetos que pertencem às regiões de fronteira causando impacto na qualidade da aproximação. A diferença entre os dois conjuntos de região de fronteira é dada por:

$$\begin{aligned} \cup Rf(CI_{original}^{\geq}) - \cup Rf(CI_{transformado}^{\geq}) &= \{1,2,3,56,57,58\} \\ |\cup Rf(CI_{original}^{\geq}) - \cup Rf(CI_{transformado}^{\geq})| &= 6 \end{aligned}$$

Todos os objetos que afetaram a qualidade da aproximação fazem parte das classes que foram unidas. Portanto, a melhora da qualidade da aproximação está diretamente relacionada à ausência de inconsistências entre as classes 1 e 2 e as classes 6 e 7.

6.5.2 Exemplo usando distribuição exponencial

No exemplo da distribuição exponencial com o sistema de dados original a qualidade de aproximação obtida foi de 78,0%, ou seja, 32 objetos definem bem as regras e o restante, 41-32=09 objetos, fazem parte da região de fronteira e não são bem definidos no sistema de informações. Como visto anteriormente, a qualidade da aproximação em técnicas utilizando o conceito de dominância é diretamente relacionada aos objetos que formam as regiões de fronteira. A figura 32 ilustra as regiões de fronteira das classes CI^{\geq} (o mesmo valor da

qualidade de aproximação seria encontrado se fossem utilizada as regiões de fronteira das classes CI^s).

Para a distribuição exponencial, com o sistema de informações original, as regiões de fronteiras geradas pelo *software* jMAF[®] são apresentadas:

Quality of approximation: 0.780		
Union name	Accuracy	Cardinality
▷ At most 1	0.684	18
▲ At least 2	0.786	23
▷ Lower		22
▷ Upper		28
▲ Boundary		6
Example_4		
Example_7		
Example_9		
Example_11		
Example_16		
Example_19		
▷ At most 2	0.966	28
▲ At least 3	0.923	13
▷ Lower		12
▷ Upper		13
▲ Boundary		1
Example_31		
▷ At most 3	1.000	34
▲ At least 4	1.000	7
▷ Lower		7
▷ Upper		7
Boundary		0
▷ At most 4	0.974	37
▲ At least 5	0.750	4
▷ Lower		3
▷ Upper		4
▲ Boundary		1
Example_41		
▷ At most 5	0.951	39
▲ At least 6	0.000	2
Lower		0
▷ Upper		2
▲ Boundary		2
Example_40		
Example_41		
▷ At most 6	0.976	40
▲ At least 7	0.000	1
Lower		0
▷ Upper		1
▲ Boundary		1
Example_41		

Figura 32 - Regiões de Fronteira - Distr. Exponencial - Sistema de Informações Original
Fonte: autor.

As classes destacadas na figura 32 são:

$$CI^{\geq 2} = \{19,21,21,22,23,24,\dots,41\}$$

$$CI^{\geq 3} = \{29,30,31,32,33,34,\dots,41\}$$

$$CI^{\geq 4} = \{35,36,37,38,39,40,41\}$$

$$CI^{\geq 5} = \{38,39,40,41\}$$

$$CI^{\geq 6} = \{40,41\}$$

$$CI^{\geq 7} = \{41\}$$

A união de todas as fronteiras destacadas e a sua cardinalidade são (equação 13):

$$\begin{aligned} \cup Rf(CI_i^{\geq}) &= Rf(CI^{\geq 2}) + Rf(CI^{\geq 3}) + Rf(CI^{\geq 4}) + Rf(CI^{\geq 5}) + Rf(CI^{\geq 6}) + Rf(CI^{\geq 7}) = \\ &= \{4,7,9,11,16,19\} + \{31\} + \{ \} + \{41\} + \{40,41\} + \{41\} = \\ &= \{4,7,9,11,16,19,31,40,41\} \\ |\cup Rf(CI_i^{\geq})| &= 9 \end{aligned}$$

Com a cardinalidade da união das regiões de fronteiras formada pelas classes CI_i^{\geq} a qualidade da aproximação é dada por (equação 16):

$$\gamma(CI) = \frac{|U - (\cup Rf(CI_i^{\geq}))|}{|U|} = \frac{41-9}{41} = 0,780$$

O resultado anterior já era conhecido. A informação adicional que pode ser verificada aqui é quais objetos que pertencem às regiões de fronteira causando impacto na qualidade da aproximação. As classes originais foram unidas no sistema de informações alterado (por se tratar de uma distribuição exponencial – com concentração à esquerda) a proposta foi a união das classes à direita do histograma original).

$$CI^a = CI^1$$

$$CI^b = CI^2$$

$$CI^c = CI^3 \cup CI^4 \cup CI^5 \cup CI^6 \cup CI^7$$

$$CI^{\geq b} = \{19,20,21,22,23,\dots,41\}$$

$$CI^{\geq c} = \{29,30,31,32,33,\dots,41\}$$

Para a distribuição exponencial com o sistema de informações alterado as regiões de fronteiras geradas pelo *software* jMAF[®] são apresentadas:

Quality of approximation: 0.829		
Union name	Accuracy	Cardinality
▷ At most a	0.684	18
▲ At least b	0.786	23
▷ Lower		22
▷ Upper		28
▲ Boundary		6
Example_4		
Example_7		
Example_9		
Example_11		
Example_16		
Example_19		
▷ At most b	0.966	28
▲ At least c	0.923	13
▷ Lower		12
▷ Upper		13
▲ Boundary		1
Example_31		

Figura 33 - Regiões de Fronteira - Distr. Exponencial - Sistema de Informações Transformado
Fonte: autor.

As classes destacadas na figura 33 são:

$$CI^{\geq b} = \{19,20,21,22,23,\dots,41\}$$

$$CI^{\geq c} = \{29,30,31,32,33,\dots,41\}$$

A união de todas as fronteiras destacadas e a sua cardinalidade são (equação 13):

$$\begin{aligned} \cup Rf(CI_t^{\geq}) &= Rf(CI^{\geq b}) + Rf(CI^{\geq c}) = \\ &= \{4,7,9,11,16,19\} + \{31\} = \{4,7,9,11,16,19,31\} \\ |\cup Rf(CI_t^{\geq})| &= 7 \end{aligned}$$

Com a cardinalidade da união das regiões de fronteiras formada pelas classes CI_t^{\geq} a qualidade da aproximação é dada por (equação 16):

$$\gamma(CI) = \frac{|U - (\cup Rf(CI_t^{\geq}))|}{|U|} = \frac{41 - 7}{41} = 0,829$$

O resultado anterior já era conhecido. A informação adicional que pode ser verificada aqui é quais objetos que pertencem às regiões de fronteira causando impacto na qualidade da aproximação. A diferença entre os dois conjuntos de região de fronteira é dada por:

$$\cup Rf(CI_{original}^{\geq}) - \cup Rf(CI_{transformado}^{\geq}) = \{40,41\}$$

$$|\cup Rf(CI_{original}^{\geq}) - \cup Rf(CI_{transformado}^{\geq})| = 2$$

Todos os objetos que afetaram a qualidade da aproximação fazem parte das classes que foram unidas. Portanto, a melhora da qualidade da aproximação está diretamente relacionada à ausência de inconsistências entre as classes 3,4,5,6 e 7. O objeto 31 ainda permanece na região de fronteira desta nova classe pela avaliação do critério “acabamento”, como esta avaliação é muito baixa, o critério de união utilizando um nível de 15% para união de classes não foi o suficiente para que ele saísse da zona de inconsistência, se a união da nova classe transformada englobasse, também, a classe 2, não haveria esta inconsistência. Por outro lado, o “*trade-off*” seria um critério muito pouco rigoroso para união de classes.

6.5.3 Exemplo usando distribuição beta

No exemplo da distribuição beta com o sistema de dados original a qualidade de aproximação obtida foi de 90,0%, ou seja, 54 objetos definem bem as regras e o restante, 60-54=06 objetos, fazem parte da região de fronteira e não são bem definidos no sistema de informações. Como visto anteriormente, a qualidade da aproximação em técnicas utilizando o conceito de dominância é diretamente relacionada aos objetos que formam as regiões de fronteira. A figura 34 ilustra as regiões de fronteira das classes CI^{\geq} (o mesmo valor da qualidade de aproximação seria encontrado se fossem utilizadas as regiões de fronteira das classes CI^{\leq}).

Para a distribuição beta, com o sistema de informações original, as regiões de fronteiras geradas pelo *software* jMAF[®] são apresentadas na figura 34:

Quality of approximation: 0.900		
Union name	Accuracy	Cardinality
▷ At most 1	0.000	2
▲ At least 2	0.967	58
▷ Lower		58
▷ Upper		60
▲ Boundary		2
Example_1		
Example_2		
▷ At most 2	0.400	5
▲ At least 3	0.948	55
▷ Lower		55
▷ Upper		58
▲ Boundary		3
Example_2		
Example_4		
Example_5		
▷ At most 3	1.000	8
▲ At least 4	1.000	52
▷ Lower		52
▷ Upper		52
Boundary		0
▷ At most 4	1.000	12
▲ At least 5	1.000	48
▷ Lower		48
▷ Upper		48
Boundary		0
▷ At most 5	1.000	19
▲ At least 6	1.000	41
▷ Lower		41
▷ Upper		41
Boundary		0
▷ At most 6	0.933	30
▲ At least 7	0.938	30
▷ Lower		30
▷ Upper		32
▲ Boundary		2
Example_21		
Example_22		

Figura 34 - Regiões de Fronteira - Distr. Beta - Sistema de Informações Original
Fonte: autor.

As classes destacadas na figura 34 são:

$$CI^{\geq 2} = \{3,4,5,6,7,8,9,10,11,\dots,60\}$$

$$CI^{\geq 3} = \{6,7,8,9,10,11,12,13,\dots,60\}$$

$$CI^{\geq 4} = \{9,10,11,12,13,14,15,\dots,60\}$$

$$CI^{\geq 5} = \{13,14,15,16,17,18,19,\dots,60\}$$

$$CI^{\geq 6} = \{20,21,22,23,24,25,26,\dots,60\}$$

$$CI^{\geq 7} = \{31,32,33,34,35,36,37,\dots,60\}$$

A união de todas as fronteiras destacadas e a sua cardinalidade são (equação 13):

$$\begin{aligned} \cup Rf(CI_i^{\geq}) &= Rf(CI^{\geq 2}) + Rf(CI^{\geq 3}) + Rf(CI^{\geq 4}) + Rf(CI^{\geq 5}) + Rf(CI^{\geq 6}) + Rf(CI^{\geq 7}) = \\ &= \{1,2\} + \{2,4,5\} + \{ \} + \{ \} + \{ \} + \{21,22\} = \\ &= \{1,2,4,5,21,22\} \\ |\cup Rf(CI_i^{\geq})| &= 6 \end{aligned}$$

Com a cardinalidade da união das regiões de fronteiras formada pelas classes CI_i^{\geq} a qualidade da aproximação é dada por (equação 16):

$$\gamma(CI) = \frac{|U - (\cup Rf(CI_i^{\geq}))|}{|U|} = \frac{60 - 6}{60} = 0,90$$

O resultado anterior já era conhecido. A informação adicional que pode ser verificada aqui é quais objetos que pertencem às regiões de fronteira causando impacto na qualidade da aproximação. As classes originais foram unidas no sistema de informações alterado (por se tratar de uma distribuição beta – com concentração à direita) a proposta foi a união das classes à esquerda do histograma original.

$$\begin{aligned} CI^a &= CI^1 \cup CI^2 \cup CI^3 \\ CI^b &= CI^4 \\ CI^c &= CI^5 \\ CI^d &= CI^6 \\ CI^e &= CI^7 \\ CI^{\geq b} &= \{9,10,11,12,13,14,\dots,60\} \\ CI^{\geq c} &= \{13,14,15,16,17,18,\dots,60\} \\ CI^{\geq d} &= \{20,21,22,23,24,25,\dots,60\} \\ CI^{\geq e} &= \{31,32,33,34,35,36,\dots,60\} \end{aligned}$$

Para a distribuição beta, com o sistema de informações alterado, as regiões de fronteiras geradas pelo *software* jMAF[®] são apresentadas na figura 35:

Quality of approximation: 0.967		
Union name	Accuracy	Cardinality
▷ At most a	1.000	8
▲ At least b	1.000	52
▷ Lower		52
▷ Upper		52
Boundary		0
▷ At most b	1.000	12
▲ At least c	1.000	48
▷ Lower		48
▷ Upper		48
Boundary		0
▷ At most c	1.000	19
▲ At least d	1.000	41
▷ Lower		41
▷ Upper		41
Boundary		0
▷ At most d	0.933	30
▲ At least e	0.938	30
▷ Lower		30
▷ Upper		32
▲ Boundary		2
Example_21		
Example_22		

Figura 35 - Regiões de Fronteira - Distr. Beta - Sistema de Informações Transformado
Fonte: autor.

As classes destacadas na figura 35 são:

$$CI^{\geq b} = \{9,10,11,12,13,14,\dots,60\}$$

$$CI^{\geq c} = \{13,14,15,16,17,18,\dots,60\}$$

$$CI^{\geq d} = \{20,21,22,23,24,25,\dots,60\}$$

$$CI^{\geq e} = \{31,32,33,34,35,36,\dots,60\}$$

A união de todas as fronteiras destacadas e a sua cardinalidade são (equação 13):

$$\begin{aligned} \cup Rf(CI_i^{\geq}) &= Rf(CI^{\geq b}) + Rf(CI^{\geq c}) + Rf(CI^{\geq d}) + Rf(CI^{\geq e}) = \\ &= \{ \} + \{ \} + \{ \} + \{21,22\} = \{21,22\} \\ |\cup Rf(CI_i^{\geq})| &= 2 \end{aligned}$$

Com a cardinalidade da união das regiões de fronteiras formada pelas classes CI_i^{\geq} a qualidade da aproximação é dada por (equação 16):

$$\gamma(CI) = \frac{|U - (\cup Rf(CI_i^{\geq}))|}{|U|} = \frac{60 - 2}{60} = 0,967$$

O resultado anterior já era conhecido. A informação adicional que pode ser verificada aqui é quais objetos que pertencem às regiões de fronteira causando impacto na qualidade da aproximação. A diferença entre os dois conjuntos de região de fronteira é dada por:

$$\begin{aligned} \cup Rf(CI_{original}^{\geq}) - \cup Rf(CI_{transformado}^{\geq}) &= \{1,2,4,5\} \\ |\cup Rf(CI_{original}^{\geq}) - \cup Rf(CI_{transformado}^{\geq})| &= 4 \end{aligned}$$

Todos os objetos que afetaram a qualidade da aproximação fazem parte das classes que foram unidas. Portanto, a melhora da qualidade da aproximação está diretamente relacionada à ausência de inconsistências entre as classes 1,2 e 3.

Portanto, ao unir classes de valores de atributos as inconsistências entre as classes unidas, que são as regiões de fronteira do sistema de informações original, não são mais observadas. Cabe ao pesquisador verificar qual o nível l^* deve ser usado para unir classes. Como pôde ser observado no exemplo da distribuição exponencial, um aumento do nível l^* poderia causar um aumento na qualidade da informação com um critério pouco rigoroso na união de classes.

6.6 PROPOSTA 2 – UNIÃO POR TRANSFORMAÇÃO PROBABILÍSTICA

Nesta proposta, os valores do atributo de decisão são tratados como médias de distribuições normais. O uso da distribuição normal para cada classificação do atributo de decisão permite considerar que esta classificação tem uma imprecisão. A variância de cada uma dessas distribuições normais pode ser estimada ou arbitrada.

Por exemplo, pode-se assumir que todas as classificações possuem a mesma variância e usar como estimador dessa variância a variância da amostra de valores observados para o atributo. Sendo assim, quanto maior for a variação do atributo de decisão em torno de seu

valor médio, maiores são as chances de uma classificação inferior ser maior que uma outra avaliação superior (uma vez que os valores são transformados em normais com média igual a própria avaliação e com a variância igual a variância da amostra de valores).

Assim, um atributo de decisão (com relação de dominância), tem média μ e variância σ^2 , sendo:

$$\mu = \frac{\sum_{i=1}^n X_i}{n} \quad ; \quad \sigma^2 = \sqrt{\frac{(X_i - \mu)^2}{n}} ;$$

$n = \text{total de observações} ; X_i = \text{valor do atributo de decisão do objeto } i$

Equação 34 - Média e Variância do Universo

Todo atributo com classificação semelhante terá a mesma distribuição normal, uma vez que a variância estimada para todos os objetos é a mesma e a média de cada distribuição é igual a classificação do próprio atributo.

Uma vez pensados os valores absolutos como médias de distribuições independentes, pode-se calcular a probabilidade de, em uma replicação da avaliação, cada uma dessas distribuições apresentar um valor mais alto que todas as outras, se o objetivo for encontrar a probabilidade de uma determinada classificação ser a maior entre o universo de objetos.

Se o objetivo for encontrar a probabilidade de uma determinada classificação ser a menor entre o universo de objetos, deve-se calcular a probabilidade desta distribuição ter o menor valor entre todas as distribuições.

Portanto, cada valor do atributo terá uma probabilidade de maximizar (ou minimizar) o valor do universo (agora transformado em funções de probabilidade). Como um objeto sempre será o máximo (ou o mínimo) de uma replicação, a soma de todas as probabilidades do vetor criado será sempre 1.

Na prática, a probabilidade de maximizar de um objeto x, com valor igual no atributo de decisão de outro objeto y, é a mesma. Porém, nunca será a mesma que, por exemplo, um objeto z que pertence a uma classe vizinha superior. A probabilidade de um objeto com média maior que outro ser o máximo de uma replicação qualquer é maior do que a probabilidade do objeto com média menor. O raciocínio se inverte quando a transformação é baseada na probabilidade de ser o mínimo em uma replicação qualquer.

O que se pode esperar é que a probabilidade de maximizar (ou minimizar) seja tão próxima do valor zero que algumas classes podem se unir por representar o conjunto de

classes com probabilidades próximas a zero de ser o máximo (ou o mínimo) da replicação. O critério aqui adotado será a aproximação até a terceira casa decimal, ou seja, probabilidades menores que 0,0005 (que quando arredondadas para três casas decimais ficam iguais a zero) serão consideradas equivalentes a zero, portanto são objetos pertencentes a classes candidatas à união.

Se o objetivo for explorar a probabilidade de uma classe apresentar o máximo em uma replicação dos valores, as probabilidades menores serão das classes à direita (ou com valores baixos para o atributo de decisão). De maneira análoga, se o objetivo for explorar a probabilidade de uma classe ser o mínimo em uma replicação dos valores, as probabilidades menores serão das classes à esquerda (ou com valores baixos para o atributo de decisão).

Assim como nas técnicas apresentadas (VC-DRSA, VP-DRSA, ISVP-DRSA), se o critério de união de classes for menos rigoroso que 0,0005, ou seja, um valor maior para permitir uniões de classes, mais classes irão se unir. O valor aqui proposto, assim como o índices l e l^* , pode ser alterado.

6.6.1 Exemplo usando a probabilidade de ser o maior valor

Para aplicar a proposta, será utilizado o exemplo apresentado primeiramente na seção 6.2. O sistema de informações tem 58 objetos e o atributo de decisão tem valores com relação de dominância indo de 1 (menor avaliação) até 7 (maior avaliação).

Ao utilizar um índice $l=0,85$ (em VC-DRSA) a qualidade da aproximação foi 62,1% (figura 22 – seção 6.2), o que significa que 36 dos 58 objetos estão bem definidos nas aproximações inferiores. A proposta aqui apresentada transforma os valores do atributo (discretos com valores inteiros de 1 a 7) em distribuições de probabilidades normais com média igual ao valor do próprio atributo e uma variância igual à variância do universo de valores do atributo de decisão.

Resumidamente, os valores do atributo de decisão deste sistema de informações têm os seus valores discretos alterados para as seguintes funções de probabilidade normais:

$$\begin{aligned} \text{valor do atributo :1} &\Rightarrow N(1; 1,4138^2); \\ \text{valor do atributo :2} &\Rightarrow N(2; 1,4138^2) \\ \text{valor do atributo :3} &\Rightarrow N(3; 1,4138^2) \\ \text{valor do atributo :4} &\Rightarrow N(4; 1,4138^2) \\ \text{valor do atributo :5} &\Rightarrow N(5; 1,4138^2) \\ \text{valor do atributo :6} &\Rightarrow N(6; 1,4138^2) \\ \text{valor do atributo :7} &\Rightarrow N(7; 1,4138^2) \end{aligned}$$

Para o vetor de valores do atributo de decisão, deve-se calcular a probabilidade de cada um dos objetos ser o maior valor em uma replicação qualquer. O código-fonte do arquivo usado no *software* Matlab[®] para o cálculo das probabilidades está apresentado no anexo 9 deste trabalho.

O valor da probabilidade de o objeto ser o máximo em uma replicação depende da sua média. Portanto, para cada média do atributo, uma probabilidade é apresentada na tabela 22:

Média da Distribuição	Probab. de um objeto pertencente à classe ser o máximo em uma replicação	Nova Classe
1	0,000	a
2	0,000	a
3	0,000	a
4	0,002	d
5	0,013	e
6	0,057	f
7	0,188	g

Tabela 22 - Classes unidas por probabilidade de maxim.
Fonte: autor.

Pela proposta de considerar valores arredondados menores até a terceira casa decimal como sendo zero, as classes inferiores com atributos 1, 2 e 3 se unem, formando a classe “a”. As classes 4, 5, 6 e 7 são respectivamente as classes “d”, “e”, “f” e “g” no novo sistema.

A qualidade da aproximação, já com o índice de consistência de $l=0,85$, passou de 0,621 (36 em 58) para 0,672 (39 em 58) conforme a figura 36:

Union name	Accuracy	Cardinality
At most a	0.818	21
At least d	0.900	37
At most d	0.745	38
At least e	0.478	20
At most e	1.000	51
At least f	1.000	7
At most f	0.948	55
At least g	0.000	3

Figura 36 - Qualidade da aproximação com dados alterados - probab. de maxim.
Fonte: autor.

A cardinalidade 21 na figura 36 representa o conjunto das classes 1,2 e 3 no sistema de informação anterior. Assim como no sistema anterior, o core é formado pelos três atributos, portanto, a exclusão de qualquer atributo impactará na qualidade da aproximação. O core (e possíveis reduções) não são alterados pela união de atributos no critério de decisão porque não há alteração de valores nos atributos condicionantes.

6.6.2 Exemplo usando a probabilidade de ser o menor valor

Para aplicar a proposta com a transformação em probabilidade de ser o menor valor em uma replicação, será utilizado o exemplo apresentado na seção 6.3. O sistema de informações tem 41 objetos e o atributo de decisão tem valores com relação de dominância indo de 1 (menor avaliação) até 7 (maior avaliação).

Ao utilizar um índice $l=0,85$ (em VC-DRSA) a qualidade da aproximação foi 78,0% (figura 25 – seção 6.3), o que significa que 32 dos 41 objetos estão bem definidos nas aproximações inferiores. A proposta aqui apresentada transforma os valores do atributo (discretos com valores inteiros de 1 a 7) em distribuições de probabilidades normais com média igual ao valor do próprio atributo e uma variância igual à variância do universo de valores do atributo de decisão.

Resumidamente, os valores do atributo de decisão deste sistema de informações têm os seus valores discretos alterados para as seguintes funções de probabilidade normais:

$$\begin{aligned} \text{valor do atributo :1} &\Rightarrow N(1; 1,50623^2); \\ \text{valor do atributo :2} &\Rightarrow N(2; 1,50623^2) \\ \text{valor do atributo :3} &\Rightarrow N(3; 1,50623^2) \\ \text{valor do atributo :4} &\Rightarrow N(4; 1,50623^2) \\ \text{valor do atributo :5} &\Rightarrow N(5; 1,50623^2) \\ \text{valor do atributo :6} &\Rightarrow N(6; 1,50623^2) \\ \text{valor do atributo :7} &\Rightarrow N(7; 1,50623^2) \end{aligned}$$

Para o vetor de valores do atributo de decisão, deve-se calcular a probabilidade de cada um dos objetos ser o menor valor em uma replicação qualquer. Para o cálculo destas probabilidades, será usado o mesmo algoritmo apresentado no anexo 9, porém os valores das médias devem ser multiplicados por -1. Desta maneira, a probabilidade de um objeto x , com média z , ser o mínimo em uma replicação qualquer é a mesma de um objeto qualquer, com média $-z$, ser o máximo em uma replicação (dado que todas as observações tiveram seus valores de média multiplicados por menos1). A inversão dos valores de média não altera o valor da variância entre os valores do atributo de decisão.

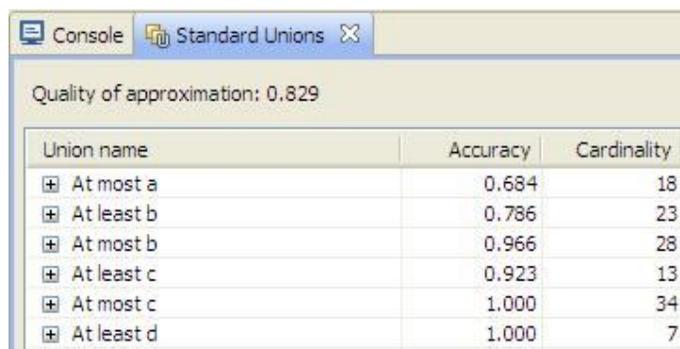
O valor da probabilidade do objeto ser o mínimo em uma replicação depende da sua média. Portanto, para cada média do atributo, uma probabilidade é apresentada na tabela 23:

Média da Distribuição	Probab. de um objeto pertencente à classe ser o mínimo em uma replicação	Nova Classe
1	0,048	a
2	0,011	b
3	0,002	c
4	0,000	d
5	0,000	d
6	0,000	d
7	0,000	d

Tabela 23 - Classes unidas por probabilidade de minim.

Pela proposta de considerar valores arredondados menores até a terceira casa decimal como sendo zero, as classes superiores com atributos 4, 5, 6 e 7 se unem, formando a classe “d”. As classes 1,2 e 3 são, respectivamente, as classes “a”, “b” e “c” no novo sistema.

A qualidade da aproximação, já com o índice de consistência de $l=0,85$, passou de 0,780 (32 em 41) para 0,829 (34 em 41) conforme a figura 37:



Union name	Accuracy	Cardinality
At most a	0.684	18
At least b	0.786	23
At most b	0.966	28
At least c	0.923	13
At most c	1.000	34
At least d	1.000	7

Figura 37 - Qualidade da aproximação com dados alterados - prob. de minim.
Fonte: autor.

A cardinalidade 7 na figura 37 representa o conjunto das classes 4,5,6 e 7 no sistema de informação anterior. Assim como no sistema anterior, o core é formado pelos três atributos, portanto, a exclusão de qualquer atributo impactará na qualidade da aproximação. O core (e possíveis reduções) não são alterados pela união de atributos no critério de decisão porque não há alteração de valores nos atributos condicionantes.

O núcleo (*core*) é formado pelos atributos acabamento e durabilidade. Portanto, a exclusão de qualquer um destes dois atributos impactará na qualidade da aproximação. O core (e possíveis reduções) não são alterados pela união de atributos no critério de decisão porque não há alteração de valores nos atributos condicionantes. Neste exemplo, o atributo condicionante - compatibilidade pode ser retirado do sistema que não impactará na qualidade de aproximação.

Assim como nas propostas previamente apresentadas, a qualidade da aproximação aumentou com o relaxamento na entrada de objetos na aproximação inferior. Se o critério de união de classes por probabilidade abaixo de 0,0005 aumentasse, ou seja, a probabilidade mínima para união de classes fosse maior do que 0,0005, a qualidade da aproximação aumentaria, porém os critérios para a entrada de objetos na aproximação inferior perderiam força.

6.7 DIFERENÇAS ENCONTRADAS NAS CLASSES UNIDAS – PROPOSTA 2

Será agora verificado o impacto da união das classes em cada um dos exemplos apresentados anteriormente na proposta 2. A hipótese de unir algumas classes que possuem uma chance mínima de ser o máximo ou o mínimo em uma replicação permite um aumento na qualidade da aproximação devido ao menor número de valores no próprio atributo de decisão o que tende a causar menos inconsistências entre todas as classificações. Em todos os exemplos que serão apresentados, a análise será feita em cima dos sistemas de informação (original e transformado com redução de atributos do atributo de decisão) com a utilização de um índice de consistência igual a $l=0,85$.

O motivo para a escolha de analisar apenas utilizando VC-DRSA é a quantidade menor de objetos nas regiões de fronteira que impactam diretamente a qualidade da aproximação. Porém, a extensão da análise com um índice de consistência igual a 1 segue os mesmos princípios.

6.7.1 Exemplo usando a probabilidade de ser o maior valor

No exemplo apresentado, com o sistema de dados original, a qualidade de aproximação obtida foi de 62,1%, ou seja, 36 objetos definem bem as regras e o restante, $58-36=22$ objetos, fazem parte da região de fronteira e não são bem definidos no sistema de informações. Como visto anteriormente, a qualidade da aproximação em técnicas utilizando o conceito de dominância é diretamente relacionada aos objetos que formam as regiões de fronteira. Na seção 6.5.1 mostrou-se as regiões de fronteiras geradas pelo sistema de informações original com índice de consistência $l = 0,85$. A união das regiões de fronteira das classes Cl^{\geq} , com cardinalidade igual a 22 foi (informações encontradas na figura 30 deste trabalho (equação 13)):

$$\begin{aligned} \cup Rf(Cl_i^{\geq}) &= Rf(Cl^{\geq 2}) + Rf(Cl^{\geq 3}) + Rf(Cl^{\geq 4}) + Rf(Cl^{\geq 5}) + Rf(Cl^{\geq 6}) + Rf(Cl^{\geq 7}) = \\ &= \{1,2,3\} + \{ \} + \{1,13,16,23\} + \{25,28,35,39,41,42,44,45,47,49,50,51\} + \{ \} + \{56,57,58\} = \\ &= \{1,2,3,11,13,16,23,25,28,35,39,41,42,44,45,47,49,50,51,56,57,58\} \\ |\cup Rf(Cl_i^{\geq})| &= 22 \end{aligned}$$

Com a cardinalidade da união das regiões de fronteiras formada pelas classes Cl_i^z a qualidade da aproximação é dada por (equação 16):

$$\gamma(Cl) = \frac{|U - (\cup Rf(Cl_i^z))|}{|U|} = \frac{58 - 22}{58} = 0,621$$

Para o exemplo dado, utilizando a união de classes apresentadas na proposta 2, as regiões de fronteiras geradas pelo *software* jMAF[®] são apresentadas (na proposta anteriormente definida pela probabilidade de ser o maior valor, as classes 1,2,3 formam a classe “a”) na figura 38:

Quality of approximation: 0.672

Union name	Accuracy	Cardinality
At most a	0.818	21
At least d	0.900	37
Lower		36
Upper		40
Boundary		4
Example_11		
Example_13		
Example_16		
Example_23		
At most d	0.745	38
At least e	0.478	20
Lower		11
Upper		23
Boundary		12
Example_25		
Example_28		
Example_35		
Example_39		
Example_41		
Example_42		
Example_44		
Example_45		
Example_47		
Example_49		
Example_50		
Example_51		
At most e	1.000	51
At least f	1.000	7
Lower		7
Upper		7
Boundary		0
At most f	0.948	55
At least g	0.000	3
Lower		0
Upper		3
Boundary		3
Example_56		
Example_57		
Example_58		

Figura 38 - Regiões de Fronteira - classes unidas pela probab. de maxim.

As classes destacadas na figura 38 são:

$$CI^{\geq d} = \{22,23,24,25,26,\dots,58\}$$

$$CI^{\geq e} = \{39,40,41,42,43,\dots,58\}$$

$$CI^{\geq f} = \{52,53,54,55,56,57,58\}$$

$$CI^{\geq g} = \{56,57,58\}$$

A união de todas as fronteiras destacadas e a sua cardinalidade são (equação 13):

$$\begin{aligned} \cup Rf(CI_i^{\geq}) &= Rf(CI^{\geq d}) + Rf(CI^{\geq e}) + Rf(CI^{\geq f}) + Rf(CI^{\geq g}) = \\ &= \{1,1,1,3,16,23\} + \{25,28,35,39,41,42,44,45,47,49,50,51\} + \{ \} + \{56,57,58\} = \\ &= \{1,1,1,3,16,23,25,28,35,39,41,42,44,45,47,49,50,51,56,57,58\} \\ |\cup Rf(CI_i^{\geq})| &= 19 \end{aligned}$$

Com a cardinalidade da união das regiões de fronteiras formada pelas classes CI_i^{\geq} a qualidade da aproximação é dada por (equação 16):

$$\gamma(CI) = \frac{|U - (\cup Rf(CI_i^{\geq}))|}{|U|} = \frac{58 - 19}{58} = 0,672$$

O resultado anterior já era conhecido. A informação adicional que pode ser verificada aqui é quais objetos que pertencem às regiões de fronteira causando impacto na qualidade da aproximação. A diferença entre os dois conjuntos de região de fronteira é dada por:

$$\begin{aligned} \cup Rf(CI_{original}^{\geq}) - \cup Rf(CI_{transformado}^{\geq}) &= \{1,2,3\} \\ |\cup Rf(CI_{original}^{\geq}) - \cup Rf(CI_{transformado}^{\geq})| &= 3 \end{aligned}$$

Todos os objetos que afetaram a qualidade da aproximação fazem parte das classes que foram unidas. Portanto, a melhora da qualidade da aproximação está diretamente relacionada à ausência de inconsistências entre as classes 1, 2 e 3.

6.7.2 Exemplo usando a probabilidade de ser o menor valor

No exemplo apresentado, com o sistema de dados original, a qualidade de aproximação obtida foi de 78,0%, ou seja, 32 objetos definem bem as regras e o restante, $41 - 32 = 9$ objetos, fazem parte da região de fronteira e não são bem definidos no sistema de informações. Como visto anteriormente, a qualidade da aproximação em técnicas utilizando o conceito de dominância é diretamente relacionada aos objetos que formam as regiões de fronteira. Na seção 6.5.2 mostrou-se as regiões de fronteiras geradas pelo sistema de informações original com índice de consistência $l = 0,85$. A união das regiões de fronteira das classes Cl_i^{\geq} , com cardinalidade igual a 9 foi (informações encontradas na figura 32 deste trabalho):

$$\begin{aligned} \cup Rf(Cl_i^{\geq}) &= Rf(Cl^{\geq 2}) + Rf(Cl^{\geq 3}) + Rf(Cl^{\geq 4}) + Rf(Cl^{\geq 5}) + Rf(Cl^{\geq 6}) + Rf(Cl^{\geq 7}) = \\ &= \{4,7,9,11,16,19\} + \{31\} + \{ \} + \{41\} + \{40,41\} + \{41\} = \\ &= \{4,7,9,11,16,19,31,40,41\} \\ |\cup Rf(Cl_i^{\geq})| &= 9 \end{aligned}$$

Com a cardinalidade da união das regiões de fronteiras formada pelas classes Cl_i^{\geq} a qualidade da aproximação é dada por (equação 16):

$$\gamma(Cl) = \frac{|U - (\cup Rf(Cl_i^{\geq}))|}{|U|} = \frac{41 - 9}{41} = 0,780$$

Para o exemplo dado, utilizando a união de classes apresentadas na proposta 2, as regiões de fronteiras geradas pelo *software* jMAF[®] são apresentadas (na proposta anteriormente definida pela probabilidade de ser o menor valor, as classes 4,5,6 e 7 formam a classe “d”) na figura 39:

Union name	Accuracy	Cardinality
At most a	0.684	18
At least b	0.786	23
Lower		22
Upper		28
Boundary		6
Example_4		
Example_7		
Example_9		
Example_11		
Example_16		
Example_19		
At most b	0.966	28
At least c	0.923	13
Lower		12
Upper		13
Boundary		1
Example_31		
At most c	1.000	34
At least d	1.000	7
Lower		7
Upper		7
Boundary		0

Figura 39 - Regiões de Fronteira - Classes unidas pela probab. de minim.
Fonte: autor.

As classes destacadas na figura 39 são:

$$CI^{\geq b} = \{19,20,21,22,23, \dots, 41\}$$

$$CI^{\geq c} = \{29,30,31,32,33, \dots, 41\}$$

$$CI^{\geq d} = \{35,36,37,38,39,40,41\}$$

A união de todas as fronteiras destacadas e a sua cardinalidade são (equação 13):

$$\begin{aligned} \cup Rf(CI_i^{\geq}) &= Rf(CI^{\geq b}) + Rf(CI^{\geq c}) + Rf(CI^{\geq d}) = \\ &= \{4,7,9,11,16,19\} + \{31\} + \{\} = \{4,7,9,11,16,19,31\} \\ |\cup Rf(CI_i^{\geq})| &= 7 \end{aligned}$$

Com a cardinalidade da união das regiões de fronteiras formada pelas classes CI_i^{\geq} a qualidade da aproximação é dada por (equação 16):

$$\gamma(Cl) = \frac{|U - (\cup Rf(Cl_i^{\geq}))|}{|U|} = \frac{41-7}{41} = 0,829$$

O resultado anterior já era conhecido. A informação adicional que pode ser verificada aqui é quais objetos que pertencem às regiões de fronteira causando impacto na qualidade da aproximação. A diferença entre os dois conjuntos de região de fronteira é dada por:

$$\begin{aligned} \cup Rf(Cl_{original}^{\geq}) - \cup Rf(Cl_{transformado}^{\geq}) &= \{40,41\} \\ |\cup Rf(Cl_{original}^{\geq}) - \cup Rf(Cl_{transformado}^{\geq})| &= 2 \end{aligned}$$

Todos os objetos que afetaram a qualidade da aproximação fazem parte das classes que foram unidas. Portanto, a melhora da qualidade da aproximação está diretamente relacionada à ausência de inconsistências entre as classes 4,5,6 e 7.

Em comparação com a união deste sistema proposta pelo uso da distribuição exponencial, percebe-se que a qualidade da aproximação foi a mesma. Na primeira proposta as classes 3,4,5,6 e 7 se uniram. Quando se usou a probabilidade de ser o mínimo valor em uma replicação, as classes 4,5,6 e 7 se uniram. Como não há inconsistências de objetos pertencentes à classe Cl^3 e objetos pertencentes a $Cl^{\geq 4}$ (na figura 32 nota-se apenas o objeto 31 pertencendo à região de fronteira da classe $Cl^{\geq 3}$, e o objeto 31 pertence a Cl^3) Portanto, a união de classes só poderá aumentar a qualidade da aproximação, nunca diminuir. E esse aumento dar-se-á pela ausência de inconsistências de objetos pertencentes às classes unidas. O que não foi o caso do objeto 31 neste exemplo.

Ao analisar a proposta baseada na distribuição normal (proposta 1), nota-se que as classes 6 e 7 se uniram. Este fato permitiu que os objetos 56, 57 e 58 saíssem da região de fronteira. Na proposta 2 (aplicando a probabilidade do valor ser o máximo em uma replicação), apenas as classes com valor baixo de atributo foram reunidas o que impacta no ganho de qualidade da aproximação. Quanto mais classes forem reunidas, maiores as chances de aumento na qualidade da aproximação, porém menos rigoroso será o critério de entrada das classes do atributo de decisão.

7 CONCLUSÃO

A importância de um sistema de informação que permita uma análise eficiente e eficaz sempre foi uma preocupação para os pesquisadores. Muitos dados estão disponíveis e o uso racional dos mesmos torna-se uma medida de vital importância.

A TCA permite usar um sistema de informações para extrair regras de decisão baseadas nos dados coletados, sem premissas de distribuição de probabilidade, independência entre objetos e conhecimento passado sobre os objetos. Isso torna a TCA uma técnica interessante com o uso e divulgação aumentando com o passar dos anos desde o desenvolvimento por Pawlak em 1982.

Com o avanço dos estudos empregando a TCA, surgiram técnicas que estudam relações de dominância das avaliações dos atributos do sistema de informação original. As propostas iniciais avaliam a inclusão ou exclusão de objetos em relação a índices e em relação aos demais objetos do sistema de informação original.

O critério original de entrada de objetos em uma aproximação inferior é rigoroso e inconsistências podem ocorrer. Diversos autores definiram técnicas de relaxamento desta entrada para que mais objetos possam gerar regras de apoio à decisão. Antes deste trabalho, não havia estudos sobre os impactos da mudança de classes no atributo de decisão na qualidade da aproximação. O critério de entrada na aproximação inferior era modificado, sem alterar dados originais.

Este estudo investigou o *trade-off* entre a quantidade de classes no atributo de decisão e as possibilidades de inconsistência. Diminuindo as classes de um atributo de decisão com dominância, segundo algum critério matemático, a qualidade da aproximação, relação direta entre a quantidade de objetos bem definidos e o total de objetos, melhora. Na prática, é importante avaliar o significado da união de algumas classes no atributo de decisão.

Duas novas propostas foram nesta tese apresentadas com resultados diferentes pelo critério de entrada. Ambas exploram a nova ideia de rever os valores do atributo de decisão. Apesar do mesmo objetivo final, as propostas usam técnicas diferentes: uma por meio de uma adequação do histograma dos valores do atributo de decisão a uma função de densidade de probabilidades e outra por transformadas probabilísticas com cálculo de probabilidades de maximizar ou minimizar.

Não há uma proposta melhor do que a outra. O que o pesquisador deve saber é que, quanto mais relaxados os critérios de união de classes, maiores são as chances de aumento no índice de qualidade da aproximação. O mesmo acontece com a aplicação de VC-DRSA. Particularmente, pelas propostas apresentadas neste trabalho, as inconsistências notadas entre as classes candidatas à união não serão mais inconsistências no novo sistema (por serem unidas em apenas uma classe). Isso causa um ganho direto na qualidade da aproximação.

As propostas apresentadas seguiram definições probabilísticas, empregando a aproximação por funções de densidade de probabilidade e a transformação na probabilidade, buscando critérios racionais para possíveis uniões do atributo de decisão. Orientando a busca da união por critérios numéricos, tem-se uma base natural para essa avaliação.

Outra possibilidade apresentada foi proposta por Greco, Matarazzo, Slowinski e Stefanowski (2001), e trabalhada posteriormente por diversos autores, na qual os dados do sistema de informação não são alterados, porém os critérios para a classificação de objetos em torno das aproximações superior e inferior são alterados. O uso da união de classes no atributo de decisão não impede a aplicação simultânea de técnicas com alterações nos critérios de entrada na aproximação inferior.

Qualquer alteração apresentada, seja nos dados originais ou no relaxamento das classificações, tende a aumentar a qualidade da aproximação. Cabe ao pesquisador fazer o uso correto destas possibilidades de modo a preservar o interesse original na pesquisa ao alterar os dados originais ou ao relaxar o critério de entrada na aproximação inferior (VC-DRSA).

As duas propostas de união de classes de atributos apresentadas neste trabalho não impedem o uso em conjunto de técnicas de VC-DRSA. Caberá ao pesquisador julgar a necessidade de aplicação de união e VC-DRSA ao mesmo tempo.

Ao aplicar as técnicas de união de classes no atributo de decisão, nenhuma premissa para aplicação de TCA com relação de dominância é violada. Apenas algumas classes do atributo de decisão foram unidas. As propriedades das aproximações inferior e superior continuam preservadas.

O que deve, sempre, ser ponderado pelo pesquisador é o quanto se pode efetuar alterações no sistema de informações original e não perder o sentido da pesquisa ou a obtenção de regras que continuem auxiliando o tomador de decisão, o grande objetivo de todo sistema de apoio à decisão.

Como desenvolvimentos futuros, podem ser feitos estudos de propostas de união nos atributos condicionantes com as mesmas propostas de união aqui apresentadas para o atributo de decisão. Ainda pode ser estudada a possibilidade de utilização de um polinômio não-linear que represente as frequências encontradas nos histogramas dos atributos em substituição ao uso de funções de densidade de probabilidade para verificar a possibilidade de união de classes.

8 REFERÊNCIAS

ALEXANDRE, J. W. C.; ANDRADE, D. F.; VASCONCELOS, A. P.; ARAUJO, A. M. S.; BATISTA, M. J. Análise do uso de categorias da escala likert aplicada à gestão da qualidade total através da teoria da resposta ao item. *Anais do ENEGEP-2003*.

BLASCZYNSKI, J.; GRECO, S.; SLOWINSKI, R.; SZELAG, M. *On Variable Consistency Dominance-based Rough Set Approaches*. LNAI, v. 4259, Springer-Verlag, Berlim, 2006, pp. 191-202.

BUSSAB, W. de O.; MORETTIN, P. A. *Estatística Básica*. 5. ed. São Paulo: Saraiva, 2003.

CABRAL, J. E. J. *Detecção de Fraudes em Unidades Consumidoras de Energia Elétrica Usando Rough Sets*. Dissertação (Mestrado) – Engenharia Elétrica. UFMS, 2005.

COUTO, A. B. G.; GOMES, L. F. A. M. G. A Tomada de Decisão em Recursos Humanos com Dados Replicados e Inconsistentes: Uma Aplicação da Teoria dos Conjuntos Aproximativos. *Pesquisa Operacional*, v. 30, n. 3, 2010, p. 657-686.

DEMBCZYNSKI, K.; SALVATORE, G.; SLOWINSKI, R. *Dominance-bases Rough Sets Approach: Features, Extensions and Application*. Disponível em: <<http://www.iiasa.ac.at/~marek/ftppub/Pubs/csm02/dembcz.ppt>>. (novembro/2011).

DENG, W.; WANG, G.; YANG, S.; HU, FENG. A new method for Inconsistent Multicriteria Classification. *Rough Sets and Knowledge Technology*, v. 6954, 2011.

DENG, W.; WANG, G.; YANG, S.; HU, FENG. An Improved Variable Precision Model of Dominance-Based Rough Set Approach. *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing. 13th International Conference*, Moscow, Russia, pp. 60-67, 2011.

GOMES, L. F. A. M.; GOMES, C. F. S; ALMEIDA, A. T. *Tomada de decisão gerencial: enfoque multicritério*. São Paulo: Atlas, 2006.

GRECO, S.; MATARAZZO, B.; SLOWINSKI, R.; STEFANOWSKI, J. Variable Consistency Dominance-based Rough Sets Approach. *International Conference on Rough Sets and Current Trends in Computing*. Londres – UK, p. 170-181, 2001.

GRZYMALA-BUSSE, J. W. Knowledge acquisition under uncertainty – a rough set approach. *Journal of Intelligent and Robotic Systems*, 1, p. 3-16, 1988.

HEIN, N.; KROENKE, A. Escólios sobre a Teoria dos Conjuntos Aproximativos. *Revista CIATEC – UPF*, v. 2 (1), p. 13-20, 2010.

INUIGUCHI, M.; YOSHIOKA, Y. Variable-Precision Dominance-Based Rough Set Approach. *Rough Sets and Current Trends in Computing. 5th International Conference*, Kobe-Japan, pp. 203-213, 2006.

INUIGUCHI, M.; YOSHIOKA, Y.; KUSUNOKI, Y. Variable-precision dominance-based rough set approach and attribute reduction. *International Journal of Approximate Reasoning*, v. 50, pp. 1199-1214. 2009.

LABORATORY OF INTELLIGENT DECISION SUPPORT SYSTEMS OF THE POZNAN UNIVERSITY OF TECHNOLOGY. Software jMAF[®]. Disponível em: <<http://www-idss.cs.put.poznan.pl>>. Março, 2011.

LAHDELMA, R.; HOKKANEN, J.; SALMINEN, P. SMAA – stochastic multiobjective acceptability analysis. *European Journal of Operational Research*, 106: p. 137-143, 1998.

LARSON, H. J. *Introduction to Probability Theory and Statistical Inference*. Singapore: John Wiley & Sons, 1982.

LIU, J. J. H.; TZENG, G. A dominance-based Rough Set Approach to customer behavior in the airline market. *Information Science*, v. 180, pp. 2230-2238, 2010.

LIU, J. J. H.; TANG, C.; YEH, W.; TSAI, C. A decision rules approach for improvement of airport service quality. *Expert Systems with Applications*, v. 38, n. 11, pp. 13723-13730, 2011.

MARAKAS, G. *Decision Support System in the 21st century*. New York: Prentice Hall, 1998.

MOREIRA FILHO, R. M.; SANT'ANNA, A. P. Impactos da Redução de Valores no Critério de Decisão em Aplicações de Rough Sets. *Anais do Enegep*, 2011.

NAVIDI, W. *Statistics for Engineers and Scientists*. New York: Mc Graw Hill, 2006.

PATRÍCIO, C. M.; PINTO, J. O. P.; SOUZA, C. C. Rough Sets – Técnica de Redução de Atributos e Geração de Regras para Classificação de Dados. *Anais do CNMAC*, CD volume 28, 2005.

PAWLAK, Z. Rough Sets and Decision Analysis. *Information Systems & Operational Research*, 38(3), 2000, p. 132-144.

PAWLAK, Z. Rough Sets. *International Journal of Computer and Information Science*. pp. 341-356, 1982.

PAWLAK, Z. *Rough Sets: Theoretical Aspects of Reasoning About Data*. Springer, 1991.

PAWLAK, Z.; SKOWRON, A. Rudiments of Rough Sets. *Information Sciences*, v. 177, n. 1, p.03-27, 2007.

PEREIRA, J. M.; GÓMEZ, F. D.; LÓPEZ, M. R. Aplicação da Teoria dos Rough Sets na Previsão do Fracasso Empresarial. *Revista de Estudos Politécnicos*, v. VI, n.10, p. 199-220, 2008.

PESSOA, A. S. A. *Mineração de Dados Meteorológicos pela Teoria dos Conjuntos Aproximativos na Previsão de Clima por Redes Neurais Artificiais*. Dissertação (Mestrado) – INPE – Computação Aplicada. São José dos Campos, 2009.

PILA, A. D. *Seleção de Atributos Relevantes para Aprendizado de Máquina Utilizando a Abordagem de Rough Sets*. Dissertação (Mestrado) – USP – Instituto de Ciências Matemáticas e de Computação – UFSCar, São Carlos, 2001.

POLKOWSKI, L.; KACPRZYK, J.; SKOWRON, A. *Rough Sets in knowledge discovery 2: Applications, case studies, and software systems*. Physica-Verlag, 1998.

SANT'ANNA A. P. Rough sets in the Probabilistic Composition of Preferences. In: De BAETS, B.; D CALUWE, R.; De TRÉ, G.; FODOR, J.; KACPRZYK, J.; ZADROZNY, S. *Current Issues in Data and Knowledge Engineering*, 407- 414, 2004.

SANT'ANNA, A. P. Índices para a Qualidade da Aproximação Baseados em Comparação Completa. *Relatórios de Pesquisa em Engenharia de Produção*, v. 6, n. 7, 2006.

SANT'ANNA, A. P.; SANT'ANNA, L. A. F. P. A probabilistic approach to evaluate the exploitation of the geographic situation of hydroelectric plants. *Energy Policy*, v. 36, n. 7, p. 2320-2329. 2008.

SASSI, R. J. Aplicação dos conceitos da Teoria dos Conjuntos Aproximados no tratamento da indiscernibilidade. *Exacta*, v. 8, p. 89-98, 2010.

<http://buscador.periodicos.capes.gov.br.ez25.periodicos.capes.gov.br/V/636FKNP6HM2BLCE47NKTQJTIC4BG3SF7TAAJX39CXTFDI956V9-41663?func=lateral-link&doc_number=064100188&line_number=0004>. Acesso em: março 2012.

<http://buscador.periodicos.capes.gov.br.ez25.periodicos.capes.gov.br/V/636FKNP6HM2BLCE47NKTQJTIC4BG3SF7TAAJX39CXTFDI956V9-41662?func=lateral-link&doc_number=064100188&line_number=0005>. Acesso em: março 2012.

SLOWINSKI, R. Rough sets with strict and weak indiscernibility relations. *Proceedings of IEEE International Conference of Fuzzy Systems*, San Diego, California, EUA, 1992, p. 695-702.

SLOWINSKI, R.; GRECO, S.; MATARAZZO, B. *Inducing Robust Decision Rules from Rough Approximations of a Preference Relation*. Springer-Verlag, Berlin, Heidelberg, 2004, pp. 118-132.

ZIARKO W. *Rough Sets, Fuzzy Sets and Knowledge Discovery*. New York: Springer Verlag, 1994.

ZIARKO W. Variable Precision Rough Sets Model. *Journal of Computer and Systems Sciences*, v. 46, n. I, p. 39-59, 1993.

ZHU, C.; CAI, W.; LU, X. Fuzzy Measure and Appraisalment of Influencing Factors on Real State Price Base don Dominance-based Rough Sets Approach. *Fifth International Conference on Fuzzy Systems and Knowledge Discovery*. pp. 381-385, 2008.

ZADEH, L. A. Fuzzy Sets. *Information and Control*, v. 8, p.338-353, 1965.

ANEXOS

ANEXO 1

**ATTRIBUTES

**ATTRIBUTES

+ acabam : [1,2,3,4,5]

+ compat : [1,2,3,4,5]

+ durabi : [1,2,3,4,5]

+ avalia : [1,2,3,4,5,6,7]

decision : avalia

**PREFERENCES

acabam: gain

compat: gain

durabi: gain

avalia: gain

**EXAMPLES

2 2 1 1

2 2 2 1

1 2 2 1

1 1 2 2

1 2 1 2

2 2 2 2

2 1 1 2

1 2 1 2

3 2 3 3

3 3 3 3

4 4 3 3

3 3 2 3

3 4 3 3

2333

1433

3433

3333

2323

2313

2133

2333

3324

2324

4344

4444

3434

3244

4544

4534

3434

3444

5434

4344

4354

4544

4324

3444

4334

4545

5435

4545

4445

5435

4445

4445

4355

4445

4 3 5 5

4 5 3 5

4 4 4 5

4 4 3 5

5 4 5 6

5 4 5 6

5 5 5 6

5 5 4 6

4 5 5 7

5 5 5 7

5 4 5 7

**END

ANEXO 2

****ATTRIBUTES**

+ acabam : [1,2,3,4,5]

+ compat : [1,2,3,4,5]

+ durabi : [1,2,3,4,5]

+ avalia : [a,b,c,d,e]

decision : avalia

****PREFERENCES**

acabam: gain

compat: gain

durabi: gain

avalia: gain

****EXAMPLES**

2 2 1 a

2 2 2 a

1 2 2 a

1 1 2 a

1 2 1 a

2 2 2 a

2 1 1 a

1 2 1 a

3 2 3 b

3 3 3 b

4 4 3 b

3 3 2 b

3 4 3 b

2 3 3 b

1 4 3 b

3 4 3 b

3 3 3 b

2 3 2 b

231b
213b
233b
332c
232c
434c
444c
343c
324c
454c
453c
343c
344c
543c
434c
435c
454c
432c
344c
433c
454d
543d
454d
444d
543d
444d
444d
435d
444d
435d
453d
444d
443d
545e

545 e

555 e

554 e

455 e

555 e

545 e

**END

ANEXO 3

****ATTRIBUTES**

+ a1 : [1,2,3,4,5,6,7,8,9,10]

+ a2 : [1,2,3,4,5,6,7,8,9,10]

+ dec: [a,b,c]

decision : dec

****PREFERENCES**

a1: gain

a2: gain

dec: gain

****EXAMPLES**

1 1 a

1 2 a

2 2 a

2 3 a

2 4 a

3 2 a

3 3 a

4 5 b

6 4 b

7 7 c

8 8 c

8 9 c

9 8 c

9 9 b

10 8 c

10 9 c

****END**

ANEXO 4

**ATTRIBUTES

+ G1: [1, 2, 3, 4, 5, 6, 7, 8], gain

+ G2: [1, 2, 3, 4, 5, 6, 7, 8], gain

+ PS: [0, 1], gain, decision

decision: PS

**PREFERENCES

G1: gain

G2: gain

PS: gain

**EXAMPLES

Example_1: 1, 1, 1

Example_2: 5, 5, 1

Example_3: 6, 7, 1

Example_4: 7, 6, 1

Example_5: 2, 3, 0

Example_6: 3, 2, 0

Example_7: 4, 4, 0

Example_8: 8, 8, 0

**END

ANEXDO 5

**ATTRIBUTES

+ acabam : [1,2,3,4,5]

+ compat : [1,2,3,4,5]

+ durabi : [1,2,3,4,5]

+ avalia : [1,2,3,4,5,6,7]

decision : avalia

**PREFERENCES

acabam: gain

compat: gain

durabi: gain

avalia: gain

**EXAMPLES

1 2 2 1

1 1 2 1

2 1 1 1

2 1 2 1

1 2 1 1

2 1 1 1

3 2 1 1

2 1 1 1

2 2 2 1

1 2 1 1

2 1 2 1

1 1 1 1

1 1 1 1

2 1 1 1

1 1 2 1

2 3 1 1

1 2 1 1

1 1 1 1

2 1 1 2
2 2 2 2
2 2 1 2
2 3 2 2
2 2 1 2
2 2 2 2
2 2 3 2
2 3 2 2
2 1 2 2
2 2 1 2
3 2 2 3
3 3 2 3
2 3 2 3
3 2 2 3
3 3 3 3
2 2 3 3
4 3 4 4
4 5 4 4
5 4 4 4
5 5 4 5
4 5 5 5
4 4 5 6
5 4 4 7
**END

ANEXO 6

**ATTRIBUTES

+ acabam : [1,2,3,4,5]

+ compat : [1,2,3,4,5]

+ durabi : [1,2,3,4,5]

+ avalia : [a,b,c]

decision : avalia

**PREFERENCES

acabam: gain

compat: gain

durabi: gain

avalia: gain

**EXAMPLES

1 2 2 a

1 1 2 a

2 1 1 a

2 1 2 a

1 2 1 a

2 1 1 a

3 2 1 a

2 1 1 a

2 2 2 a

1 2 1 a

2 1 2 a

1 1 1 a

1 1 1 a

2 1 1 a

1 1 2 a

2 3 1 a

1 2 1 a

1 1 1 a

2 1 1 b
2 2 2 b
2 2 1 b
2 3 2 b
2 2 1 b
2 2 2 b
2 2 3 b
2 3 2 b
2 1 2 b
2 2 1 b
3 2 2 c
3 3 2 c
2 3 2 c
3 2 2 c
3 3 3 c
2 2 3 c
4 3 4 c
4 5 4 c
5 4 4 c
5 5 4 c
4 5 5 c
4 4 5 c
5 4 4 c
**END

ANEXO 7

****ATTRIBUTES**

+ acabam : [1,2,3,4,5]
+ compat : [1,2,3,4,5]
+ durabi : [1,2,3,4,5]
+ avalia : [1,2,3,4,5,6,7]
decision : avalia

****PREFERENCES**

acabam: gain
compat: gain
durabi: gain
avalia: gain

****EXAMPLES**

1 2 2 1
2 2 2 1
1 2 2 2
2 2 2 2
2 2 2 2
2 2 2 3
2 1 1 3
2 1 2 3
3 3 3 4
4 3 3 4
4 3 3 4
4 3 3 4
4 4 3 5
4 4 3 5
3 3 4 5
4 4 3 5
3 4 4 5
3 4 4 5

3445
4446
4556
5546
4446
4456
5446
5446
4456
4456
4446
4546
5547
5557
5547
4557
5557
5557
5457
5457
5557
5557
5457
4557
5557
5557
5557
4557
5547
5547
5547
5557
4557
5547

5 5 5 7

5 5 5 7

4 5 5 7

5 4 5 7

4 5 5 7

5 5 5 7

5 5 5 7

5 5 5 7

**END

ANEXO 8

****ATTRIBUTES**

+ acabam : [1,2,3,4,5]

+ compat : [1,2,3,4,5]

+ durabi : [1,2,3,4,5]

+ avalia : [a,b,c,d,e]

decision : avalia

****PREFERENCES**

acabam: gain

compat: gain

durabi: gain

avalia: gain

****EXAMPLES**

1 2 2 a

2 2 2 a

1 2 2 a

2 2 2 a

2 2 2 a

2 2 2 a

2 1 1 a

2 1 2 a

3 3 3 b

4 3 3 b

4 3 3 b

4 3 3 b

4 4 3 c

4 4 3 c

3 3 4 c

4 4 3 c

3 4 4 c

3 4 4 c

3 4 4 c
4 4 4 d
4 5 5 d
5 5 4 d
4 4 4 d
4 4 5 d
5 4 4 d
5 4 4 d
4 4 5 d
4 4 5 d
4 4 4 d
4 5 4 d
5 5 4 e
5 5 5 e
5 5 4 e
4 5 5 e
5 5 5 e
5 5 5 e
5 4 5 e
5 4 5 e
5 5 5 e
5 5 5 e
5 4 5 e
4 5 5 e
5 5 5 e
5 5 5 e
5 5 5 e
4 5 5 e
5 5 4 e
5 5 4 e
5 5 4 e
5 5 5 e
4 5 5 e
5 5 4 e

5 5 5 e

5 5 5 e

4 5 5 e

5 4 5 e

4 5 5 e

5 5 5 e

5 5 5 e

5 5 5 e

**END

ANEXO 9

```

filename=uigetfile('*.txt', 'Selezione o arquivo');
dados = importdata(filename, ' ');
tol=1.0e-6;
escala=10^(size(dados,1)^(1/3));
reavaliados=zeros(size(dados));
    for j=1:size(dados,2)
        desvio=desvio_da_normal(dados(:,j));
        minimo=min(dados(:,j))-4*desvio;
        maximo=max(dados(:,j))+4*desvio;
        for i=1:size(dados,1)
            M=dados(:,j);
            M(i)=[];
            prov=@(x)produtocdfnormal_vet(x,M,desvio).*normpdf(x,dados(i,j),desvio)*escala;
            reavaliados(i,j)=quad(prov,minimo,maximo,tol)/escala;
        end
    end
fid = fopen(resultados,'w');
for i = 1:size(dados,1)
    fprintf(fid,'%0.4f\t', reavaliados(i,:));
    fprintf(fid,'\n');
end
fclose(fid);
function y=produtocdfnormal_vet(x,M,b)
% produto das funções acumuladas da distribuição normal, vetorizada.

y=1;
for i=1: length(M)
    y=y.*normcdf(x,M(i),b);
end

```