PREVISÃO MULTICRITÉRIO DA PROBABILIDADE DE INADIMPLÊNCIA PARA PORTFÓLIOS DE BAIXO *DEFAULT*

Ayrton Benedito Gaia do Couto
Analista de Sistemas, BNDES
Av. República do Chile, 100, Centro, Rio de Janeiro, CEP 20031-917, RJ, Brazil, Phone: +55 21 3747-7658
ayrtoncouto@gmail.com

Luiz Flavio Autran Monteiro Gomes
Professor, Ibmec/RJ
Av. Presidente Wilson, 118, Centro, Rio de Janeiro, CEP 20030-020, RJ, Brazil, Phone: +55 21 4503-4053
luiz.gomes@ibmec.edu.br

RESUMO

Este estudo analisa, por meio de um enfoque multicritério (regras de decisão "*if...,then...*") e com a aplicação simultânea e complementar das teorias dos conjuntos nebulosos (*Fuzzy sets*) e aproximativos (*Rough sets*), a possibilidade de previsão da probabilidade de inadimplência (ou calote) para portfólios de baixo *default* (ou seja, inadimplência). Na prática, o cálculo dessa probabilidade está sujeito a julgamentos de especialistas, os quais implicam em subjetividade, imprecisão e incerteza dos resultados. Assim, vis-à-vis à imprecisão e incerteza inerentes ao processo de cálculo e tomada de decisão, neste estudo as regras inferidas sobre uma tabela de decisão para uma simulação com 50 empresas permitiram a geração de uma base de conhecimento para a previsão de classes de probabilidades de inadimplência. Fez-se uso do *package* "RoughSets ®", em "R", como suporte à análise e previsão.

PALAVRAS CHAVE. Probabilidade de inadimplência, Análise Multicritério, Teoria dos Conjuntos Aproximativos e Nebulosos.

Tópicos: ADM - Apoio à Decisão Multicritério

ABSTRACT

This study analyzes by means of a multicriteria approach ("*if...,then...*" decision rules), with the simultaneous and complementary application of fuzzy and rough sets theories, the possibility of forecasting the default probability for low-default portfolios. In practice, the calculation of this probability is subject to judgments of experts, which imply subjectivity, imprecision and uncertainty of the results. In this study, the inferred rules about a decision table for a simulation with 50 companies allowed the generation of a knowledge base for the forecasting of classes of default probabilities. The "RoughSets ®" package, in "R", was used as support for analysis and forecasting.

KEYWORDS. Default probability, Multi-criteria Analysis, Rough and Fuzzy Sets Theory.

Paper topics: ADM - Multicriteria Decision Support

https://proceedings.science/p/85316?lang=pt-br

1. Introdução

O cálculo da probabilidade de *default* (PD) ("inadimplência ou calote de determinada obrigação") para portfólios de baixa ocorrência de *default*, é um árduo processo de apuração das inúmeras variáveis qualitativas e quantitativas que podem influenciar esse cálculo. A escolha de um modelo (estatístico), a definição de parâmetros, a identificação dos dados relevantes, etc., requer um julgamento e que está sujeito a críticas e justificativas. No caso de instituições financeiras, essas justificativas representam uma importante parte do processo de validação do modelo estatístico proposto. Essa análise estatística pode ser aprimorada quando se acrescenta o conhecimento ("julgamento") de um especialista na definição dos valores dos parâmetros ou mesmo na escolha do modelo, por exemplo. Mas, como todo resultado de julgamento, implica em possuir um certo grau de subjetividade, incerteza e imprecisão.

Para tratar essa incerteza e imprecisão, há duas teorias complementares: Conjuntos Nebulosos (*Fuzzy*) e Aproximativos (*Rough Sets*). Essas teorias foram estabelecidas no início dos anos 1960 e 1980, respectivamente, e visam tratar a imprecisão e incerteza de dados, cada uma sob uma ótica própria [Pawlak 1982], [Zadeh 1965].

O objetivo deste estudo é realizar a predição da "classe de PD" (espaço ou região fuzzy), concomitantemente, com a indução de regras de decisão multicritério ("*if...,then...*") para um conjunto de 50 empresas, por meio das teorias dos Conjuntos *Fuzzy* e Aproximativos, e com o uso do *package* "RoughSets", em "R". Ademais, a indução de regras de decisão multicritério é de fundamental importância para o campo do aprendizado ("machine learning") e inteligência artificial, além de permitir a identificação de padrões "escondidos" em dados [Pawlak 1991], [Pawlak et al. 1995]. O presente estudo visa também, apresentar uma alternativa (enfoque multicritério) ao estudo de um caso prático de cálculo de PD para portfólio de baixo *default* frente a estudos correntes [BIS 2004], [Flores et al. 2010], [Jacobs e Kiefer 2010], [Jacobs e Kiefer 2011], [Kiefer 2006], [Kiefer 2007], [Kiefer 2008a], [Kiefer 2008b], [Kiefer 2009], [Pluto e Tasche 2005].

Este estudo está dividido nas seguintes seções e subseções: 1-Introdução; 2-Fundamentação teórica; 2.1- Abordagem Bayesiana; 2.2-Teoria dos Conjuntos Aproximativos; 2.3-Teoria dos Conjuntos *Fuzzy*; 2.4-Conjuntos *Fuzzy* e Aproximativos; 3-Extensão da análise do cálculo de PD a uma simulação; e, 4-Conclusões e estudos futuros.

2. Fundamentação teórica

No sentido de permitir a utilização das Teorias dos Conjuntos Aproximativos, Conjuntos *Fuzzy* e *Fuzzy*-Aproximativos para tratar a incerteza e imprecisão no cálculo da probabilidade de *default*, esta seção foi dividida nas subseções seguintes.

2.1. Abordagem Bayesiana

Inicialmente, no contexto do cálculo da probabilidade de *default*, a abordagem Bayesiana permite a incorporação formal de informação ("julgamento") por meio de regras da probabilidade [Jacobs e Kiefer 2010]. Como não se sabe quando ocorrerá (ou não) um *default*, utiliza-se da distribuição de probabilidade para modelar este evento incerto, e que reflete um conhecimento parcial do mecanismo de *default*. À distribuição de probabilidade sobre dados históricos, informações fornecidas ("julgamentos") por especialistas são acrescidas ao modelo, tendo em vista que, cálculos de PD somente com base em dados históricos não são suficientes, pois provavelmente, somente algumas inadimplências seriam observadas em um portfólio de baixo *default* [Kiefer 2006], [Kiefer 2008a], [Kiefer 2008b], [Kiefer 2009].

A informação fornecida pelo especialista constitui-se em um mix de análise subjetiva – informações não modeladas e experiência com bases de dados. Cabe ao especialista decidir sobre quais segmentos de ativos formarão a base de risco, o período de amostragem dos dados e qual modelo estatístico será utilizado. O uso da probabilidade para descrever e modelar incertezas sobre *default* com modelos estatísticos é questão central para o gerenciamento de risco e também para atender os requerimentos de "Basileia II" [Kiefer 2007]. É usual na modelagem estatística escolher um modelo que gere todas as probabilidades relevantes como função de um pequeno

conjunto de parâmetros, sendo o mais comum o de Bernoulli, onde os *defaults* são independentes dos ativos e períodos de tempo, e ocorrem com uma probabilidade "θ". Considerando-se D = {d_i, i = 1, ..., n} um conjunto de dados, onde d_i representa um *default* (d_i = 1) ou não (d_i = 0), r = r(D) = Σ_i d_i representa a quantidade de *defaults*, tem-se [Kiefer 2007]:

$$P(D|\theta) = \prod \theta^{d_i} (1-\theta)^{1-d_i} = \theta^r (1-\theta)^{n-r} \quad (1)$$

Assim, a distribuição da probabilidade de "r" *defaults*, cada um com probabilidade "θ" (constante) em "n" tentativas (fixas), é dada pela distribuição Binomial:

$$p(r|\theta) = \binom{n}{r} \theta^r (1-\theta)^{n-r} \quad (2)$$

Mas, há limitações: se as probabilidades de *default* (θ) forem heterogêneas devido a mudanças macroeconômicas por exemplo, os *defaults* irão se aglutinar, o que tornará essa informação difícil de ser capturada por esse modelo. A distribuição Binomial representa melhor situações quando se trata de condições homogêneas ("θ" constante), e que pode não ser factível para longos períodos de tempo ou quando se trata de portfólios (ativos) distintos. Na equação (2), como "θ" permanece desconhecido, uma forma conveniente de distribuição de probabilidade a "priori" pode ser então utilizada [Kiefer 2007]:

$$p(\theta) = B(\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1} \quad (3)$$

onde $B(\alpha, \beta)$ representa a função Beta; Γ(n) = (n - 1)!, representa a função Gama e,

$$\bar{x} = \alpha/(\alpha+\beta) \quad (4)$$
$$S^2 = \alpha\beta/(\alpha+\beta)^2(\alpha+\beta+1) \quad (5)$$

onde $\bar{x}$ e S^2 são respectivamente, a média e variância. A partir de (3), pode-se obter a distribuição de probabilidade conjunta a "posteriori" com "r", por meio do teorema de Bayes (p(θ|r) = p(θ,r) / p(r)). Portanto,

$$p(\theta|r) = \frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+r)\Gamma(\beta+n-r)} \theta^{\alpha+r-1}(1-\theta)^{\beta+n-r-1} \quad (6)$$

Assim, (6) representa uma distribuição a "posteriori" (incorpora julgamento(s) de um especialista). Um modelo para o cálculo de PD também é proposto por "Basileia" (Comitê Modelo de Capital) [Kiefer 2007]:

$$\bar{\theta} = (\alpha + r) / (\alpha + \beta + n) \quad (7)$$

2.2. Teoria dos Conjuntos Aproximativos

A Teoria dos Conjuntos Aproximativos (*Rough Sets Theory*, RST) teve origem com Zdzislaw Pawlak no início dos anos 1980, como ferramenta matemática para o tratamento da imprecisão e incerteza de dados [Pawlak 1982], [Pawlak et al. 1995]. Em Pawlak [1991], encontra-se o conceito de Sistema de Informação: K = (U, A), onde U é um conjunto finito e não-vazio de objetos e A é um conjunto finito e não-vazio de atributos, tal que, a: U → V_a para todo a ϵ A, onde V_a é o conjunto de valores que pode ser atribuído ao atributo a. Inicialmente, há o conceito sobre "relação de indiscernibilidade" ("indiscernibility relation"): dado um sistema de informação (U,A) e para qualquer B ⊆ A, uma "relação de equivalência" (ou "classificação", indistintamente) R_B é definida por [Riza et al. 2014]:

$$R_B(x,y) = \{(x,y) \epsilon U^2 | \forall a \epsilon B, a(x) = a(y)\} \quad (8)$$

https://proceedings.science/p/85316?lang=pt-br

Se (x,y) ϵ R_B(x,y), então x e y tem exatamente os mesmos valores para os atributos em B. De acordo com Pawlak [1991], nem sempre é possível expressar exatamente algum conjunto de objetos com o conhecimento disponível. Assim, torna-se possível expressar um conjunto de objetos por outros dois subconjuntos: aproximações "inferior" ("lower", $\underline{R}$) e "superior" ("upper", $\overline{R}$), dado $X \subseteq U$ e uma relação de equivalência R:

$$\underline{R}X = \cup \{Y \epsilon U|R: Y \subseteq X\}; \overline{R}X = \cup \{Y \epsilon U|R: Y \cap X \neq \emptyset\} \quad (9)$$

$\underline{R}X$ é o subconjunto de elementos de U, que "certamente" podem ser classificados como elementos de X; $\overline{R}X$ é o subconjunto de elementos de U, que "possivelmente" podem ser classificados como elementos de X. Assim, um conjunto é "aproximativo" ("rough") com relação a R, se e somente se, $\underline{R}X \neq \overline{R}X$. Do conceito de Sistema de Informação (K = (U, A)), pode-se obter o conceito de "tabela de decisão": quando A é formado por dois subconjuntos: C e D, atributos de condição e de decisão, respectivamente (C, D $\subset$ A); ou, T = (U, A, C, D). E, classes de equivalência das relações R_B e R_C são conhecidas como classes de condição e decisão, respectivamente. Para cada x ϵ U, associa-se uma função d_x: A $\rightarrow$ V, tal que, $d_x(a) = a(x)$, para todo $a \epsilon$ C $\cup$ D; a função d_x é conhecida como "regra de decisão" (em T) [Pawlak 1991]. De acordo com Pawlak [1991] e Riza et al. [2014], define-se região "positiva-B", o conjunto de objetos em U que certamente pode ser classificado usando os atributos condicionais B:

$$POS_B = \bigcup_{x \epsilon U} \underline{R}X \quad (10)$$

E, pode-se computar o grau de dependência do atributo de decisão em relação ao conjunto de atributos de condição B:

$$\gamma_B = |POS_B| / |U| \quad (11)$$

Assim, uma tabela de decisão A é considerada consistente se $\gamma_A = 1$. De acordo com Riza et al. [2014], um "superreduto" ("superreduct") é um conjunto de atributos B $\subset$ A tal que $R_B = R_A$, onde R_B e R_A são relações de indiscernibilidade definidas por B e A, respectivamente. Se essa relação for também mínima, trata-se de um "reduto" ("reduct"). A interseção de todos os redutos é um "núcleo" ("core"), de um sistema de informação. Há também, o conceito de Matriz de Discernibilidade (*M*(A)): seja (U, A) um sistema de informação; *M*(A) é uma matriz simétrica "n x n" cujos elementos c_{ij} são definidos como: $c_{ij} = \{ a \in A : a(x_i) \neq a(x_j) \}$, para i, j = 1, ..., n. Ou seja, c_{ij} contém aqueles atributos para diferentes objetos x_i e x_j. De acordo com Slowinski et al. [2012], há o princípio de Dominância: "objetos que possuem uma melhor avaliação ou que possuem no mínimo a mesma avaliação ("classe de decisão"), não podem estar associados a uma classe pior de decisão, considerados todos os critérios de decisão". Consequentemente, as relações de indiscernibilidade são substituídas por "relações de dominância" nas aproximações das classes de decisão. Em Slowinski et al. [2012], apresenta-se o conceito para "decisão multicritério": aquela que envolve problemas relativos a um conjunto finito de objetos (alternativas, ações, soluções, etc.) que podem ser avaliadas por um conjunto finito de critérios (atributos, características, variáveis, etc.).

2.3. Teoria dos Conjuntos *Fuzzy* (*Fuzzy Sets*, FS)

De acordo com Zadeh [1965], seja X um conjunto de objetos e "x" um elemento qualquer desse conjunto (X = {x}). Um conjunto *fuzzy* (nebuloso, difuso) A em X é caracterizado por uma função de "pertinência" ou "associação" ("membership") $f_A(x)$ (ou "μ"), que associa à cada objeto em X um número real no intervalo [0, 1]. Assim, $f_A(x)$ representa o grau de "pertinência ou associação" de x em A. Como exemplo, seja X um conjunto de números reais e A um conjunto *fuzzy* de números muito maiores que 1. Então, valores representativos para a função poderiam ser definidos: $f_A(0) = 0$; $f_A(1) = 0$; $f_A(5) = 0{,}01$; $f_A(10) = 0{,}2$; $f_A(100) = 0{,}95$; $f_A(500) =$

1. Embora a função de associação tenha alguma semelhança com função de probabilidade quando X é discreto (ou função de densidade de probabilidade quando X é contínuo), a natureza de um conjunto *fuzzy* é completamente "não estatística". Assim, de acordo com Zadeh [2008], na lógica *fuzzy* tudo é ou pode ser granulado – um grânulo é um grupo de valores de atributos unidos pela indistinguibilidade, similaridade, proximidade ou funcionalidade. Essa granulação de forma gradual, é inspirada pela maneira como os humanos lidam com a complexidade e a imprecisão. Como exemplo (Figura 1), a idade de uma pessoa pode ser granulada e, consequentemente, estar associada a um ou mais conjuntos *fuzzy* (jovem, meia-idade, idoso), conforme determinada gradação – processo conhecido como "fuzzificação". Neste exemplo, "jovem", "meia-idade" e "idoso" são variáveis "linguísticas". O processo inverso da "fuzzificação" é a "desfuzzificação": de variáveis linguísticas para valores esperados.

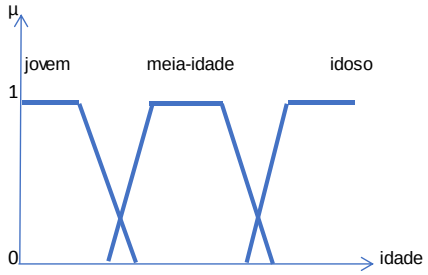


Figura 1 – Conjuntos *fuzzy* "jovem", "meia-idade" e "idoso"; fonte: Zadeh 2008

2.4. Conjuntos *Fuzzy* e Aproximativos

De acordo com Radzikowska e Kerre [2002] e Jensen e Cornelis [2010], os estudos sobre conjuntos *Fuzzy*-Aproximativos começaram com D. Dubois e H. Prade, no início dos anos 1990. Predominantemente, o foco foi na "fuzzificação" das aproximações inferior e superior encontradas nos conjuntos Aproximativos (equações em 9). Consequentemente, dois princípios foram amplamente adotados: a) um conjunto A pode ser generalizado para um conjunto *fuzzy* em X, permitindo que os objetos possam pertencer a um dado conceito (possui certas características) segundo uma "gradação"; b) e, não somente identificar objetos indiscerníveis mas, medir a qualidade da aproximação por meio de uma relação *fuzzy* R. Como resultado, objetos podem ser categorizados em classes (ou grânulos) com limites tênues, com base na sua similaridade. Assim, transições abruptas entre classes são substituídas por gradações, as quais permitem que um objeto possa pertencer a mais de uma classe, indicando-se o seu grau. Há autores ainda que adicionam mais uma condição de τ −transitividade, isto é, uma *t-norm* τ,

$$\tau\big(\mathrm{R}(x,y),\mathrm{R}(y,z)\big) \leq R(x,z) \qquad (12)$$

para qualquer x, y e z em X; R é então chamada uma relação de equivalência- τ *fuzzy*. τ representa uma *t-norm* - uma função [0, 1] x [0, 1] → [0, 1], que satisfaz as seguintes propriedades: a) comutatividade: τ (a, b) = τ (b, a); b) monotonicidade: τ (a, b) ≤ τ (c, d) se a ≤ c e b ≤ d; c) associatividade: τ (a, τ (b, c)) = τ (τ (a, b), c); e) tem o "1" como elemento identidade: τ (a, 1) = a. De acordo com Jensen e Cornelis [2010], para um atributo qualitativo (nominal) *a*, a maneira clássica de diferenciar objetos, $\mathrm{R}_a(x,y) = 1$ se $a(x) = a(y)$ e $\mathrm{R}a(x,y) = 0$, caso contrário, define-se que, para qualquer subconjunto B de A, a relação de indiscernibilidade fuzzy B por:

$$\mathrm{R}_B(x,y) = \tau(\mathrm{R}a(x,y)) \quad (13)$$

Assim, as relações em (9), para as aproximações "inferior" e "superior", para um conjunto *fuzzy* ficam, respectivamente [Jensen e Cornelis 2010]:

$$(\mathrm{R}{\downarrow}\, A)(y) = inf_{x\epsilon X}\, \zeta(R(x,y), A(x)) \qquad (14)$$
$$(\mathrm{R}{\uparrow}\, A)(y) = sup_{x\epsilon X}\, \tau(\mathrm{R}(x,y), A(x)) \qquad (15)$$

para qualquer y em X; ζ representa um "implicador *fuzzy*". Em ambas as teorias [Pawlak et al. 1995], [Zadeh 2008], a indução de regras de decisão multicritério torna-se uma poderosa linguagem de sumarização dos dados através do uso da granulação. A combinação dos conceitos propostos na RST e FS conforme Greco et al. [2008], permite a aplicação do princípio de Dominância a um conjunto *fuzzy*. Em Greco et al. [2006] e Du e Hu [2017], encontram-se exemplos dessa combinação de conceitos para a indução de regras.

3. Extensão da análise do cálculo de PD a uma simulação

Para o presente estudo, utilizou-se de uma simulação para o cálculo de PD para 50 empresas (Tabela 2), considerando-se as seguintes variáveis ou critérios (Tabela 1) [Cardoso et al. 2013]:

Variável (critério)	Limite inferior	Limite superior
dívida líquida/ebitda	-8,068	103,094
cobertura de juros	-24,457	1788,988
ROA (retorno sobre ativos)	-0,947	-0,947
"utility"	0	1
passivo/ativo	0	24,017
tamanho (total ativo)	12,284	27,266

Tabela 1 – Variáveis ou critérios utilizados no estudo; fonte: Cardoso et al. 2013

Subentende-se que, "ebitda" é "lucro antes de juros, impostos, depreciação e amortização"; "utility": "1" ou "0", se é integrante ou não do setor de serviços públicos, respectivamente. Para este estudo, o valor de PD para portfólio de baixo *default* foi simulado no intervalo [0,0001; 0,0100] (ou [0,01%; 1,00%]). Ademais, o cálculo de PD poderia ser realizado em função de outras variáveis ou indicadores, como exemplo, lucros retidos sobre o ativo e saldo de tesouraria sobre vendas, conforme Brito e Assaf Neto [2008]; taxa anual de juros e limite de crédito, conforme Flores et al. [2010]. Os modelos para cálculo de PD foram aperfeiçoados ao longo do tempo: Altman [1968] já apontava modelos para previsão de falências de empresas. Assim, dadas a incerteza e imprecisão concernentes ao processo de cálculo do valor de PD para um portfólio de baixo *default* [Kiefer 2006], foram utilizados os modelos matemáticos RST e FS conjuntamente, para tratamento desses resultados. Inicialmente, definiu-se um conjunto *fuzzy* "classe baixa de PD", com valores simulados de PD no intervalo [0; 0,01]; definiu-se também que, a "classe baixa de PD" seria então, dividida em outros três (sub)conjuntos *fuzzy*: "superior", "médio" e "inferior". Como há uma tênue delimitação entre esses conjuntos, conforme demonstra a Figura 2, definiu-se que os limites (internos) estão em torno dos valores 0,0030 e 0,0070 de PD. Assim, os conjuntos *fuzzy* superior, médio e inferior estão delimitados pelos seguintes intervalos de PD: [0; 0,0030], [0,0030; 0,0070] e [0,0070; 0,0100], respectivamente. Assim, ao conjunto *fuzzy* "superior" da "classe baixa de PD", por exemplo, associou-se o menor intervalo de valores de PD ([0; 0,0030]). Ou seja, quanto menor o valor de PD, mais se destaca esse valor, classificando-o como um valor "superior" dentro da "classe baixa de PD".



Figura 2 – Conjuntos *fuzzy* "superior", "médio" e "inferior", inclusos na "classe baixa de PD"

empresa	dívida líquida/ ebitda	cobertura juros	ROA	"utility"	passivo /ativo	tamanho (total ativo)	valor PD	classe baixa PD
001	77,000	1344,000	-0,200	0	9,000	14,000	0,0054	medio
002	86,000	1201,000	0,100	0	7,000	22,000	0,0083	inferior
003	40,000	977,000	-0,600	1	17,000	26,000	0,0002	superior
004	-1,000	886,000	0,000	0	14,000	21,000	0,0032	medio
005	28,000	175,000	0,000	1	22,000	17,000	0,0023	superior
006	89,000	910,000	0,000	0	20,000	17,000	0,0014	superior
007	51,000	1277,000	0,100	1	15,000	14,000	0,0022	superior
008	46,000	770,000	-0,300	1	12,000	23,000	0,0025	superior
009	-3,000	204,000	-0,900	0	23,000	18,000	0,0043	medio
010	21,000	1705,000	-0,800	1	7,000	27,000	0,0067	medio
011	2,000	1218,000	-0,900	0	4,000	19,000	0,0067	medio
012	102,000	952,000	0,400	1	0,000	27,000	0,0042	medio
013	86,000	824,000	0,200	1	10,000	18,000	0,0020	superior
014	40,000	802,000	0,400	0	14,000	23,000	0,0010	superior
015	15,000	461,000	-0,100	1	14,000	19,000	0,0064	medio
016	51,000	731,000	0,400	1	21,000	17,000	0,0063	medio
017	66,000	688,000	-0,700	0	8,000	21,000	0,0026	superior
018	87,000	384,000	0,100	0	13,000	19,000	0,0021	superior
019	3,000	637,000	-0,200	1	7,000	23,000	0,0026	superior
020	41,000	1750,000	0,300	1	0,000	26,000	0,0034	medio
021	98,000	107,000	-0,100	1	3,000	20,000	0,0043	medio
022	11,000	1227,000	-0,600	0	5,000	20,000	0,0074	inferior
023	46,000	777,000	-0,400	1	16,000	21,000	0,0012	superior
024	100,000	1562,000	-0,800	0	11,000	15,000	0,0013	superior
025	73,000	973,000	0,400	0	22,000	20,000	0,0058	medio
026	35,000	714,000	-0,400	0	14,000	18,000	0,0041	medio
027	64,000	758,000	0,600	1	10,000	16,000	0,0096	inferior
028	39,000	988,000	-0,900	0	12,000	16,000	0,0015	superior
029	70,000	112,000	-0,200	1	13,000	26,000	0,0072	inferior
030	54,000	202,000	-0,800	0	23,000	19,000	0,0048	medio
031	78,000	8,000	0,200	0	22,000	21,000	0,0007	superior
032	35,000	930,000	0,500	0	14,000	19,000	0,0025	superior
033	65,000	1259,000	-0,200	1	23,000	19,000	0,0000	superior
034	66,000	159,000	-0,300	0	0,000	13,000	0,0034	medio
035	57,000	816,000	-0,800	0	22,000	23,000	0,0031	medio
036	65,000	391,000	0,000	0	9,000	15,000	0,0074	inferior
037	102,000	1084,000	-0,300	1	20,000	16,000	0,0019	superior
038	86,000	1454,000	-0,200	0	20,000	15,000	0,0071	inferior
039	96,000	740,000	0,300	0	20,000	21,000	0,0055	medio
040	24,000	460,000	0,500	0	11,000	26,000	0,0064	medio
041	27,000	1182,000	-0,900	0	16,000	18,000	0,0004	superior
042	46,000	842,000	0,400	0	12,000	25,000	0,0016	superior
043	21,000	1617,000	0,600	0	18,000	16,000	0,0080	inferior
044	70,000	916,000	-0,200	0	17,000	21,000	0,0025	superior
045	93,000	94,000	-0,300	1	23,000	15,000	0,0098	inferior
046	73,000	1479,000	-0,400	0	24,000	25,000	0,0091	inferior
047	-7,000	283,000	0,600	1	11,000	18,000	0,0051	medio
048	70,000	910,000	0,400	1	6,000	22,000	0,0089	inferior
049	37,000	1548,000	-0,700	0	16,000	16,000	0,0037	medio
050	93,000	1630,000	-0,400	0	3,000	22,000	0,0032	medio

Tabela 2 – Simulação do cálculo de PD; fonte: Cardoso et al. 2013

No contexto deste estudo, visto que se trata de uma simulação, não há necessariamente, relação dos valores dos indicadores com os respectivos valores e classes de PD. Assim, considerando-se a Tabela 2 como uma "tabela de decisão": "classe baixa PD", como critério de decisão, e demais critérios como de condição, exceto "empresa" e "valor PD", submeteu-se esse conjunto de dados à análise pelo *package* "RoughSets" (versão 1.3-0, de 03/09/2015), do *software* "R" [Riza et al. 2014], [Riza et al. 2015]. Os resultados obtidos dessa análise foram os seguintes:

a) Regras de decisão: a indução de regras de decisão (multicritério) pode ser realizada pela identificação dos critérios de condição mais relevantes - por "seleção de características" ("feature selection") [Jensen et al. 2009], com o uso da função "RI.hybridFS.FRST"; ou sob forma mais geral [Zhao et al. 2010], pelo uso da função "RI.GFRS.FRST".

a.1) Exemplos de indução de regras de decisão por critérios de condição mais relevantes

The rules :
[[1]]
[1] "IF div_liquida_ebitda is around -1 and passivo_ativo is around 14 THEN classe_baixa_pd is medio"

[[2]]
[1] "IF div_liquida_ebitda is around -3 and passivo_ativo is around 23 THEN classe_baixa_pd is medio"

[[3]]
[1] "IF div_liquida_ebitda is around 102 and passivo_ativo is around 0 THEN classe_baixa_pd is medio"

a.2) Exemplos de indução de regras de decisão multicritério sob o método geral

The rules :
[[1]]
[1] "IF passivo_ativo is around 14 and div_liquida_ebitda is around 40 and tamanho is around 23 THEN classe_baixa_pd is superior"

[[2]]
[1] "IF cobertura_juros is around 159 and passivo_ativo is around 0 THEN classe_baixa_pd is medio"

[[3]]
[1] "IF cobertura_juros is around 202 and roa is around -0.8 THEN classe_baixa_pd is medio"

b) Previsão para a "classe baixa de PD" considerando-se a mesma tabela de decisão com as "50 empresas" (Tabela 2), com o uso das funções RI.hybridFS.FRST (método híbrido) e RI.GFRS.FRST (método geral). Jensen e Cornelis [2011] propõem um algoritmo (sequência finita de instruções) "classe vizinha mais próxima" ("k-nearest neighbor algorithm") para predições com base em aproximações inferior e superior utilizando-se das teorias RST e FS. Em Keller et al. [1985], há uma proposta desse algoritmo para conjuntos *Fuzzy*.

[1] "Definição classe de decisão por previsão, método híbrido:"
> print(resultado #1)
classe_baixa_pd
1 medio
2 inferior
3 superior
4 medio

5 superior
6 superior
7 superior
8 superior
9 medio
10 medio
11 medio
12 medio
13 superior
14 superior
15 medio
16 medio
17 superior
18 superior
19 superior
20 medio
21 medio
22 inferior
23 superior
24 superior
25 medio
26 medio
27 inferior
28 superior
29 inferior
30 medio
31 superior
32 superior
33 superior
34 medio
35 medio
36 inferior
37 superior
38 inferior
39 medio
40 medio
41 superior
42 superior
43 inferior
44 superior
45 inferior
46 inferior
47 medio
48 inferior
49 medio
50 medio

[1] "Definição classe de decisão por previsão, método geral:"
> print(resultado #2)
classe_baixa_pd
1 superior
2 superior
3 superior
4 superior
5 superior
6 superior
7 superior
8 superior

9	**medio**
10	**superior**
11	**superior**
12	**superior**
13	**superior**
14	**superior**
15	**superior**
16	**superior**
17	**superior**
18	**superior**
19	**superior**
20	**superior**
21	**medio**
22	**superior**
23	**superior**
24	**superior**
25	**superior**
26	**medio**
27	**superior**
28	**superior**
29	**superior**
30	**medio**
31	**superior**
32	**superior**
33	**superior**
34	**medio**
35	**superior**
36	**superior**
37	**superior**
38	**superior**
39	**superior**
40	**superior**
41	**superior**
42	**superior**
43	**inferior**
44	**superior**
45	**superior**
46	**inferior**
47	**medio**
48	**superior**
49	**superior**
50	**superior**

Pelos resultados anteriores, não foram encontradas discordâncias na previsão da "classe baixa de PD" com o uso do método híbrido (resultado #1) - taxa de concordância = 100%. Quando utilizado o método geral (resultado #2), houve 22 discordâncias (empresas 1, 2, 4, 10, 11, 12, 15, 16, 20, 22, 25, 27, 29, 35, 36, 38, 39, 40, 45, 48, 49 e 50); assim, a taxa de concordância = (50 – 22)/50 = 28/50 ou 56%.

Assim, por meio da simulação do cálculo de PD para 50 empresas, o estudo mostra que é possível tratar essa incerteza e imprecisão advinda (naturalmente) do processo de cálculo de PD para esse tipo de portfólio (baixo *default*), onde as ocorrências de *default* são raras em razoáveis intervalos de tempo. Ao invés de se trabalhar diretamente com pequenos valores de probabilidade (ex, 0,003 ou 0,3%), procurou-se "classificar" os valores de PD segundo um espaço ou região *fuzzy* – "superior", "médio", "inferior", dentro de uma classe de baixa PD. Consequentemente, tornou-se possível extrair regras de decisão muticritério e, consequentemente, gerar uma base de conhecimento para efetuar predições para a "classe baixa de PD" – métodos híbrido e geral, com "*fuzzy-rough*".

4. Conclusões e estudos futuros

Dadas as incertezas na forma e conteúdo do cálculo de PD para um portfólio de baixo *default*, este estudo visou ampliar a análise do cálculo de PD por meio de ferramentas matemáticas (complementares) direcionadas para o tratamento da incerteza e imprecisão de dados – Teorias dos Conjuntos Nebulosos (*Fuzzy Sets*) e Aproximativos (*Rough Sets*), sob enfoque multicritério. Assim, por meio de simulação, tornou-se possível a predição de "classes de PD" (espaços ou regiões *fuzzy*) com base na indução de regras de decisão multicritério ("*if...,then...*"), como alternativa ao cálculo direto da probabilidade de default para portfólios de baixo *default*, além de permitir ser instrumento de "aprendizagem por exemplos" (uso de critérios de condição e de decisão). E, no caso específico dessa versão do *package* "RoughSets", em "R", verificou-se que, gerar uma matriz de Discernibilidade somente é possível quando se utiliza de critérios de decisão nominais; também, não se encontrou referência ao uso do "princípio de Dominância".

Como estudo futuro, a extensão da análise dessa simulação com o uso dos *softwares* jMAF (*Dominance-Based Rough Set Data Analysis Framework*) [Blaszczynski et al. 2012] e WEKA (*Waikato Environment for Knowledge Analysis*) [Riza et al. 2014], e a "fuzzificação" dos critérios de condição pela classificação dos valores em regiões.

Agradecimentos

Pesquisa sob suporte do CNPq (projeto nº 305732/2012-9)**.**

Referências

Altman, E. J. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The Journal of Finance*, 23, 4, 589-609.

BIS (Bank for International Settlements) (2004). Basel Committee on Banking Supervision, International Convergence of Capital Measurement and Capital Standards.

Blaszczynski, J., Greco, S., Matarazzo, B., Slowinski, R. e Szeląg, M. (2012). *jMAF*. Instituto de Ciência da Computação, Universidade de Tecnologia de Poznan, Polônia. Web page: http://www.cs.put.poznan.pl/jblaszczynski/Site/jRS.html. Acessado: 2013-03-21.

Brito, G. A. S. e Assaf Neto, A. (2008). Modelo de classificação de risco de crédito de empresas, *Revista Contabilidade & Finanças* (USP), 19, 46, 18-29.

Cardoso, V. S., Guimarães, A. L. S., Macedo, H. F. e Lima, J. C. C. O. (2013). Assessing corporate risk: a pd model based on credit ratings, *ACRN Journal of Finance and Risk Perspectives*, 2, 1, 51-58.

Du, W. S. e Hu, B. Q. (2017). Dominance-based rough fuzzy set approach and its application to rule induction, *European Journal of Operational Research*, 261, 2, 690-703.

Flores, J. A. E., Basualdo, T. L. e Sordo, A. R. Q. (2010). Regulatory use of system-wide estimations of PD, LGD and EAD, Financial Stability Institute, Bank for International Settlements (BIS).

Greco, S., Inuiguchi, M. e Slowinski, R. (2006). Fuzzy rough sets and multiple-premise gradual decision rules, *International Journal of Approximate Reasoning*, 41, 179–211.

Greco, S., Matarazzo, B. e Slowinski, R. (2008). Fuzzy Set Extensions of the Dominance-Based Rough Set Approach, H. Bustince et al., (eds.). Fuzzy Sets and Their Extensions: Representation, Aggregation and Models, Springer.

Jacobs, M. e Kiefer, N. M. (2010). The Bayesian Approach to default Risk: A Guide. *CAE Working Paper*, No. 10-01, Center for Analytical Economics (CAE), Cornell University.

Jacobs Jr., M. e Kiefer, N. M. (2011). The Bayesian Approach to default Risk Analysis and the Prediction of default Rates. Yale SOM, *Econometrics Summer Conference*.

Jensen, R. e Cornelis, C. (2010). Fuzzy-rough instance selection, *WCCI IEEE World Congress on Computational Intelligence*, 1776-1782.
Jensen, R. e Cornelis, C. (2011). Fuzzy rough nearest neighbour classification and prediction", *Theoretical Computer Science*, 412, 5871-5884.
Jensen, R., Cornelis, C. e Shen, Q. (2009). Hybrid fuzzy-rough induction and feature selection, *FUZZ-IEEE*, Korea, 1151-1156.
Keller, J. M., Gray, M. R. e Givens Jr., J. A. (1985). A fuzzy k-nearest neighbor algorithm, *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-15, 4, 580-585.
Kiefer, N. M. (2006). Default estimation for low-default portfolios. *OCC Economics Working Paper*, 2006-2.
Kiefer, N. M. (2007). The probability approach to default probabilities. *OCC Economics Working Paper*, 2007-1.
Kiefer, N. M. (2008a). Default estimation and expert information. *Journal of Business and Economic Statistics*, 28(2), 320–328.
Kiefer, N. M. (2008b). Default estimation, correlated defaults and expert information. *CAE Working Paper*, 08-02, Center for Analytical Economics (CAE), Cornell University.
Kiefer, N. M. (2009). Correlated defaults, temporal correlation, expert information and predictability of default rates. *CAE Working Paper*, 09-12, Center for Analytical Economics (CAE), Cornell University.
Pawlak, Z. (1982). Rough sets. *Int. J. Comput. Inf. Sci*. 11, 341-356.
Pawlak, Z. (1991). Rough Sets. Theoretical Aspects of Reasoning about Data. Kluwer, Netherlands.
Pawlak, Z., Grzymala-Busse, J., Slowinski, R. e Ziarko, W. (1995). Rough sets, *Communications of the ACM*, 38, 11, 88-95.
Pluto, K. e Tasche, D. (2005). Thinking Positively, *Risk*, August, 72-78.
Radzikowska, A. M. e Kerre, E. E. (2002). A comparative study of fuzzy rough sets, Fuzzy Sets and Systems, 126, 2, 137-155. Web page: https://dokumen.tips/documents/a-comparative-study-of-fuzzy-rough-sets.html. Acessado: 2018-02-12.
Riza, L. S., Janusz, A., Bergmeir, C., Cornelis, C., Herrera, F., Slezak, D. e Benitez, J. M. (2014). Implemeting algorithms of rough set theory and fuzzy rough set theory in the R package "roughsets", *Information Sciences*, 287, 68-89.
Riza, L. S., Janusz, A., Slezak, D., Cornelis, C., Herrera, F., Benitez, J. M., Bergmeir, C. e Stawicki, S. (2015). Data analysis using rough set and fuzzy rough set theories. Web page: https://cran.r-project.org/web/packages/RoughSets/RoughSets.pdf. Acessado: 2018-01-13.
Slowinski, R., Greco, S. e Matarazzo, B. (2012). Rough set and rule-based multicriteria decision aiding, *Pesquisa Operacional*, 32 (2), 213-269.
Zadeh, L. A. (1965). Fuzzy Sets, *Information and Control*, 8, 338-353.
Zadeh, L. A. (2008). Is there a need for fuzzy logic?, *Information Sciences*, 178, 2751-2779.
Zhao, S. Y., Tsang, E. C. C., Chen , D. G. e Wang X. Z. (2010). Building a rule-based classifier – a fuzzy-rough set approach, *IEEE Transactions on Knowledge and Data Engineering*, 22, 5, 624-638.