

MINISTÉRIO DA EDUCAÇÃO
Universidade Federal de Alfenas - UNIFAL-MG
Instituto de Ciências Sociais Aplicadas - ICSA

WALEF MACHADO DE MENDONÇA

MODELOS *FUZZY* E PROBABILÍSTICO NO ESTUDO DA
PROBABILIDADE DE RUINA

VARGINHA/MG
2020

WALEF MACHADO DE MENDONÇA

MODELOS *FUZZY* E PROBABILÍSTICO NO ESTUDO DA
PROBABILIDADE DE RUINA

Trabalho de conclusão de curso apresentado ao Instituto de Ciências Sociais Aplicadas da Universidade Federal de Alfenas, como requisito parcial à conclusão do curso de Especialização em Modelagem em Ciência e Tecnologia

Orientador:

Leandro Ferreira
UNIFAL-MG
Campus Varginha

Co-orientador:

Lázaro Eduardo da Silva
CEFET-MG
Campus Varginha

VARGINHA/MG
2020

RESUMO

O risco de uma seguradora ficar com reserva insuficiente para pagar suas indenizações é comumente denominado como probabilidade de ruína. Além de indispensável, a probabilidade de ruína se apresenta como um grande desafio para os profissionais do mercado segurador. Isso se deve ao fato de que sua obtenção pode envolver processos complexos, e não existe uma expressão que possa ser aceita como ideal na comunidade atuarial. Nesse contexto, o objetivo deste trabalho é apresentar três abordagens para a modelagem do limite superior da probabilidade de ruína. A primeira abordagem (modelagem clássica) constitui-se da utilização do modelo probabilístico de Cramér-Lundberg para determinar o limite superior da probabilidade de ruína de uma seguradora através de dados simulados. A segunda abordagem (modelagem *fuzzy* 1) busca construir um sistema baseado em regras *fuzzy* (SBRF) integrado ao modelo de Cramér-Lundberg, sendo que a variável de entrada do SBRF é a variável aleatória “valor da indenização”, gerada através de simulação de Monte Carlo, e a variável de saída o “limite superior da probabilidade de ruína”. Para tanto, foram definidos cenários em que o valor da indenização tem distribuição exponencial e distribuição gama. Por fim, a terceira abordagem (modelagem *fuzzy* 2) busca estender o modelo probabilístico de Cramér-Lundberg sob a perspectiva da teoria dos conjuntos *fuzzy*, via princípio de extensão de Zadeh. Para isso, foi considerado que o valor das indenizações segue uma distribuição exponencial com parâmetro incerto, representado por um número *fuzzy*. Para a construção dos SBRFs, foram consideradas funções de pertinência gaussianas (5, 10 e 15 conjuntos *fuzzy*), método de inferência de Mamdani e método defuzzificação centro de gravidade, através do pacote *frbs* do *software* R. Para validação dos modelos, os dados simulados foram divididos em dois conjuntos, um para treinamento e outro para teste e foram calculados o coeficiente de determinação e o erro quadrático médio. Em todos os cenários, foram obtidos coeficientes de determinação acima de 90%, tanto para o conjunto de treinamento quanto para o conjunto de teste nos modelos com 5, 10 e 15 conjuntos *fuzzy*, isso indica que os SBRFs foram bem ajustados e têm boa capacidade para a estimação da probabilidade da ruína. Na modelagem *fuzzy* 2, observou-se que quanto menor é a incerteza referente ao parâmetro da distribuição exponencial, menor a incerteza resultante em relação ao valor do limite superior da probabilidade de ruína e ao valor médio da indenização.

Palavras-chave: Probabilidade de ruína, sistemas baseados em regras *fuzzy*, princípio de extensão de Zadeh.

ABSTRACT

The risk of an insurer being left with insufficient reserves to pay its claims is commonly referred to as the probability of ruin. In addition to being indispensable, a probability of ruin presents itself as a major challenge for professionals in the insurance market. This is due to the fact that it can include complex criteria, and there is no expression that can be accepted as the ideal in the actuarial community. In this context, this work aims to present three approaches for modeling the upper limit of the probability of ruin. The first approach (classical modeling) consists of using the Cramér-Lundberg probabilistic model to determine the upper limit of the probability of an insurer's ruin through simulated data. A second approach (fuzzy modeling 1) seeks to build a fuzzy rule-based system (FRBS) integrated with the Cramér-Lundberg model, with the FRBS input variable being "value of the claims", generated through Monte Carlo Method, and the output variable the 'upper limit of the probability of ruin'. For that, scenarios were defined in which the indemnity value has an exponential distribution and a gamma distribution. Finally, the third approach (fuzzy modeling 2) seeks to extend the Cramér-Lundberg probabilistic model from a perspective of the fuzzy set theory, via Zadeh's extension principle. For this, it was considered that the value of the indemnities follows an exponential distribution with an uncertain parameter, represented by a fuzzy number. For the construction of the FRBS were considered Gaussian pertinence functions (5, 10 and 15 fuzzy sets), Mamdani inference method and center of gravity defuzzification method. Were used the frbs package of software R. To validate the models, the simulated data were divided into two sets, one for training and the other for testing, and the determination coefficient and the mean square error were calculated. In all scenarios, determination coefficients above 90% were obtained, both for the training set and for the test set in the models with 5, 10 and 15 fuzzy sets, this indicates that the FRBS were well adjusted and have good capacity for estimating the probability of ruin. In the fuzzy modeling 2, it was observed that the lower the uncertainty regarding the exponential distribution parameter, the lower the resulting uncertainty in relation to the value of the upper limit of the probability of ruin and the average value of the indemnity.

Key-words: Ruin probability, fuzzy rules-based system, Zadeh extension principle.

SUMÁRIO

1	Introdução	4
2	Revisão de literatura	5
2.1	Teoria da ruína	5
2.1.1	Modelo clássico de Cramér-Lundberg	5
2.1.2	Probabilidade de ruína	7
2.1.3	Coefficiente de ajustamento e desigualdade de Lundberg	7
2.2	Conjuntos <i>fuzzy</i>	10
2.2.1	Subconjuntos <i>fuzzy</i> e α -nível	11
2.2.2	Princípio de extensão de Zadeh	12
2.2.3	Números <i>fuzzy</i>	12
2.2.4	Sistemas baseados em regras <i>fuzzy</i>	14
2.2.5	O método de inferência de Mandani	14
2.2.6	Métodos de defuzzificação	15
3	Metodologia	16
3.1	Recursos computacionais	16
3.1.1	O pacote <i>frbs</i>	16
3.2	Modelagem clássica	17
3.2.1	Indenizações com distribuição exponencial	18
3.2.2	Indenizações com distribuição gama	18
3.3	Modelagem <i>fuzzy</i> 1	19
3.4	Modelagem <i>fuzzy</i> 2	20
4	Resultados	21
4.1	Modelagem clássica	21
4.1.1	Indenizações com distribuição exponencial	21
4.1.2	Indenizações com distribuição gama	22
4.2	Modelagem <i>fuzzy</i> 1	24
4.3	Modelagem <i>fuzzy</i> 2	29
5	Considerações finais	30
6	Referências	31

1 Introdução

Para atuar no mercado brasileiro, uma seguradora deve obedecer alguns requisitos. Entre eles destacam-se um capital mínimo e uma margem de solvência. A margem de solvência expressa a relação entre as apólices de seguro vendidas e a capacidade da seguradora de honrar com suas obrigações (LEMOS, 2008). Os estudos de solvência envolvem a análise de vários componentes de risco, tais como; o risco legal, operacional, de crédito, das operações financeiras e o risco de subscrição. Adicionalmente, há o risco de uma seguradora ficar com reserva insuficiente para pagar suas indenizações, comumente denominado como probabilidade de ruína. Além de indispensável, determinar esse risco se apresenta como um grande desafio para os profissionais do mercado segurador, uma vez que sua obtenção pode envolver processos complexos, e não existe, entretanto, nenhuma expressão que possa ser aceita como ideal pela comunidade atuarial (FERREIRA, 2005).

Nesse contexto, este trabalho tem como objetivo, apresentar três abordagens para a modelagem da determinação do limite superior da probabilidade de ruína, baseado no modelo desenvolvido por Crámer-Lundberg. Na primeira abordagem (modelagem clássica), foram apresentados o cálculo do coeficiente de ajustamento e do limite superior da probabilidade de ruína para os casos em que o valor das indenizações segue a distribuição exponencial e a distribuição gama. Na segunda abordagem (modelagem *fuzzy* 1), foram apresentados os SBRFs utilizados para modelar o limite superior da probabilidade de ruína e o pacote *frbs* (*software* R), utilizado para a construção dos SBRFs. Os valores das indenizações foram gerados por meio de simulação de Monte Carlo. Na terceira abordagem (modelagem *fuzzy* 2), foi apresentada a extensão do modelo clássico do cálculo do limite superior da probabilidade de ruína para o contexto *fuzzy*, via princípio de extensão de Zadeh, considerando que o valor da indenização segue uma distribuição exponencial com parâmetro incerto.

O presente trabalho foi estruturado em seis seções, incluindo essa introdução. Na próxima seção, é apresentada uma revisão de literatura sobre teoria da ruína e lógica *fuzzy*. Em seguida, é apresentada a metodologia utilizada, incluindo o cálculo do coeficiente de ajustamento para os casos em que o valor das indenizações tem distribuições exponencial e gama, são apresentados também os recursos computacionais, o sistema baseado em regras *fuzzy* e a extensão do modelo probabilístico de Cramér-Lundberg para a teoria dos conjuntos *fuzzy*. Na quarta seção, são apresentados os resultados. A quinta seção traz as considerações finais, onde são discutidos os resultados alcançados. A última seção apresenta as referências bibliográficas.

2 Revisão de literatura

2.1 Teoria da ruína

A teoria da ruína trata do estudo do nível da reserva de uma seguradora para uma carteira de apólices de seguro ao longo do tempo, levando em conta os momentos em que as indenizações ocorrem, bem como seus valores (DICKSON, 2016). Uma seguradora entra em processo de ruína quando a quantidade do seu capital, na forma de reserva, não é capaz de arcar com os sinistros dos seus segurados. É importante ressaltar que, entrar em processo de ruína não significa entrar em falência (LEMOS, 2008). Várias estratégias podem ser utilizadas para reduzir o risco de entrar em processo de ruína. Entre elas se destaca o resseguro, que tem como objetivo diminuir os riscos assumidos pela seguradora direta e, conseqüentemente, a sua probabilidade de ruína. Além disso, intervenções no processo, como o pagamento de dividendos ou o aumento do prêmio por riscos com incidência de sinistros desfavoráveis, não são considerados na definição da probabilidade de ruína. Para mais, os efeitos da inflação e do retorno sobre o capital não são considerados no cálculo, e a probabilidade de ruína só responde pelo risco de seguro, não por uma possível má administração (KAAS et al 2008).

Nesse trabalho, probabilidade de ruína significa a probabilidade da reserva de uma seguradora ficar negativa em algum instante do tempo, dado um capital inicial, ou seja, é a possibilidade de uma companhia de seguro não ter reserva suficiente para pagar as indenizações resultantes de algum sinistro.

Alguns trabalhos sobre probabilidade de ruína são: Eade (1983), Beekman (1985), Gerber (1987), De Vylder (1988), Hipp (1989), Seah (1990), Dickson e Waters (1992) e Asmussen (2010).

2.1.1 Modelo clássico de Cramér-Lundberg

O modelo de Cramér-Lundberg, ou modelo clássico de risco coletivo, foi inicialmente apresentado em 1903 por Filip Oskar Lundberg. Porém, devido à sua terminologia de difícil entendimento e ao baixo desenvolvimento da teoria de processos estocásticos na época, o trabalho foi pouco utilizado pelos atuários. Mais tarde, em 1930, a teoria de Lundberg foi retomada por Harald Cramér que a desenvolve e a torna mais acessível (LEMOS, 2008).

A partir daí, iniciam-se os primeiros estudos acerca de probabilidade de ruína e seu comportamento no tempo infinito, ou no tempo finito, que constituem o campo de estudo da Teoria da Ruína.

O modelo de Cramér-Lundberg é um processo estocástico e pode ser apresentado da seguinte forma

$$U(t) = u + ct - S(t), \quad t \geq 0, \quad (1)$$

em que $U(t)$ representa a reserva de uma seguradora até o instante de tempo $t \geq 0$. As indenizações agregadas relativas ao intervalo $(0, t]$ são representadas por $S(t) = \sum_{i=1}^{N(t)} X_i$, $N(t)$ é o número de indenizações ocorridas no intervalo de tempo $(0, t]$ e ocorre segundo um processo de Poisson(λt), com $\lambda > 0$. Assume-se que X_i são variáveis aleatórias, identicamente distribuídas e não negativas e representam os valores das indenizações individuais (LEMOS, 2008). Para esse modelo, assume-se que a reserva inicial $U(0) = u$ e os prêmios são recebidos pela seguradora a uma taxa constante $c > 0$. Também assume-se que os prêmios recebidos são superiores ao valor esperado das indenizações agregadas, ou seja, $c > \lambda p_1$, em que p_1 é a esperança de X_i .

É importante destacar que o modelo de Cramér-Lundberg não leva em consideração despesas administrativas com os contratos de seguro, rendimentos de investimentos ou taxas de juros.

A Figura 1 apresenta a trajetória da reserva de uma seguradora segundo o modelo de Cramér-Lundberg, em que a reserva inicial é igual a 10 ($U(0) = 10$), a taxa de recebimento dos prêmios é igual a 2 ($c = 2$) e o parâmetro da distribuição do número de indenizações é 1 ($\lambda = 1$).

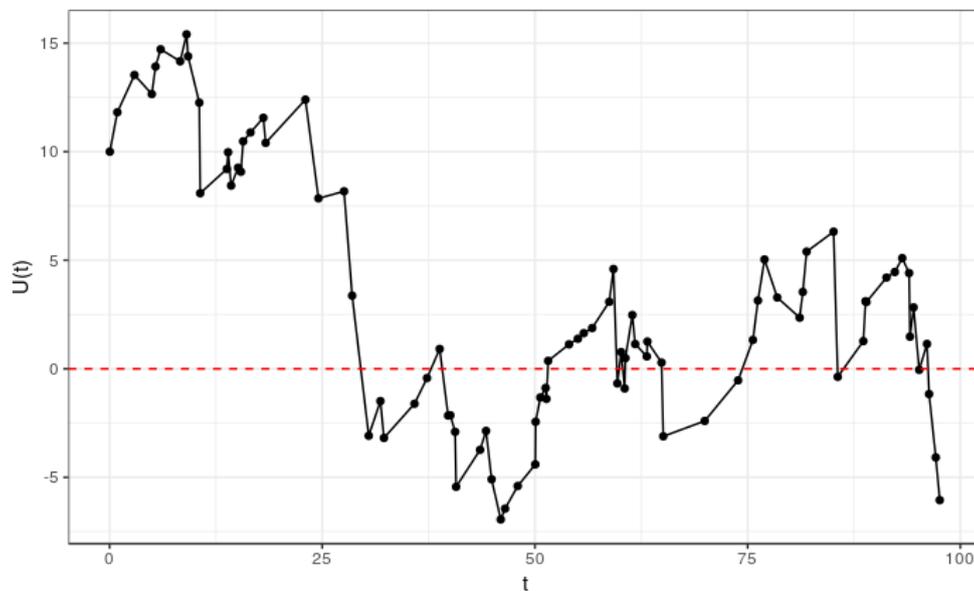


Figura 1: Trajetória do processo de reserva de uma seguradora.
Fonte: Elaboração própria.

Os valores da esperança e variância da variável $U(t)$ são dados respectivamente por

$$\begin{aligned}
E(U(t)) &= u + ct - E(S(t)) \\
&= u + ct - \lambda tp_1 \\
&= u + t(c - \lambda p_1) \text{ e}
\end{aligned}$$

$$Var(U(t)) = \lambda tp_2.$$

2.1.2 Probabilidade de ruína

A probabilidade de ruína, denotada por $\psi(u)$, decorrente de uma reserva inicial u em um horizonte temporal infinito, é dada por:

$$\begin{aligned}
\psi(u) &= P(T < \infty | U(0) = u) = P(U(t) < 0 \text{ para um determinado valor fixo } x > 0 | U(0) = u) \\
&= P(u + ct - S(t) < 0 | U(0) = u) = P(S(t) > u + ct | U(0) = u) \\
&= P\left(\sum_{i=1}^{N(t)} X_i > u + ct | U(0) = u\right) = \int_{u+ct}^{\infty} p_s(x) dx,
\end{aligned}$$

em que $p_s(x)$ é a função de densidade das indenizações agregadas e a variável aleatória $T = \inf\{t > 0 \text{ e } U(t) < 0\}$ é o momento de ocorrência da ruína para uma dada reserva inicial $U(0) = u$ (LEMOS, 2008).

É importante destacar que, como u é a reserva da seguradora, a probabilidade de ruína é inversamente proporcional a u . Ou seja, a probabilidade de ruína é alta para valores baixos de u , o que pode ser visualizado na Figura 2, considerando que $X \sim Exponencial(0, 5)$.

2.1.3 Coeficiente de ajustamento e desigualdade de Lundberg

Nessa seção serão apresentadas algumas definições e alguns resultados necessários para descrever o coeficiente de ajustamento e a desigualdade de Lundberg.

Ramos (2014) define o coeficiente de segurança, associado ao modelo clássico de risco, como o número real $\theta > 0$ tal que o prêmio é $c = (1 + \theta)\lambda p_1$, onde θ é um carregamento de segurança. Portanto, o coeficiente de ajustamento é definido como a única raiz positiva $r = R$ da equação

$$1 + (1 + \theta)p_1 r = M_X(r), \quad (2)$$

em que $M_X(r)$ denota a função geradora de momentos da distribuição de X .

Utilizando a definição de Ramos (2014), a igualdade apresentada na Equação 2 pode ser reescrita da seguinte forma $rc = \lambda[M_X(r) - 1]$, isto é, $e^{rc} = M_S(r)$, em

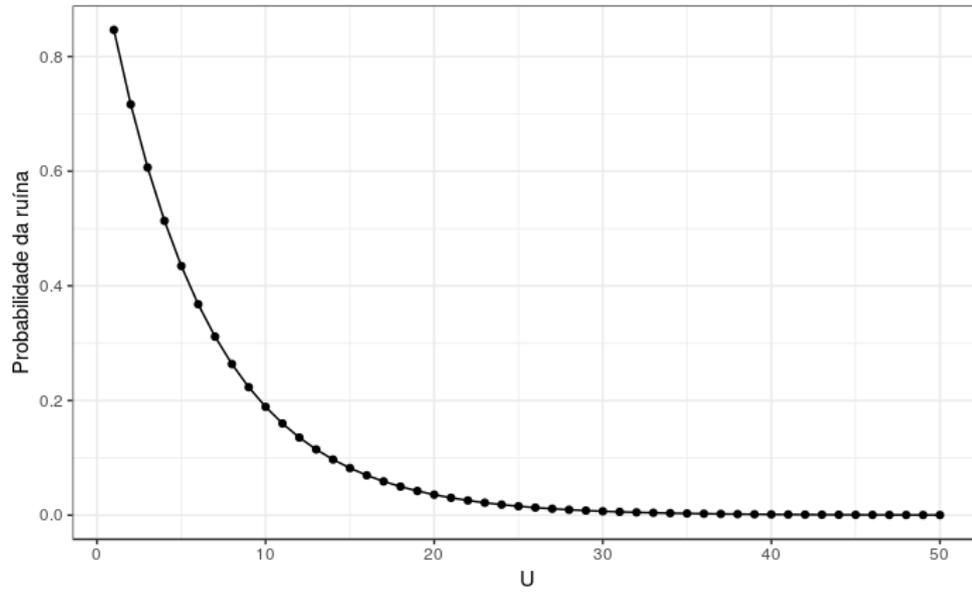


Figura 2: Probabilidade de ruína em função da reserva inicial u da seguradora.
Fonte: Elaboração própria.

que $S = \sum_{i=1}^N X_i$, com $N \sim \text{Poisson}(\lambda)$. Sendo assim, o coeficiente de ajustamento, comumente denotado como R , representa o equilíbrio ou ajustamento entre os valores que a seguradora recebe como prêmio dos seguros (rc) e os valores que ela gasta com as indenizações ($M_S(r)$) (RAMOS, 2014).

A Equação 2 tem apenas uma solução positiva

$$M_X(r) = 1 + \frac{c}{\lambda}r$$

A desigualdade de Lundberg pode ser demonstrada através do seguinte teorema apresentado por Lemos (2008).

Teorema 2.1 *Admitindo a existência do coeficiente de ajustamento R , a desigualdade de Lundberg é dada por*

$$\psi(u) \leq e^{-Ru}. \quad (3)$$

Demonstração. Considerando que a ruína ocorre na primeira indenização ($n = 1$), por definição, o tempo decorrente até a primeira indenização tem função de densidade $k(t)$. Então,

$$\begin{aligned}\psi_1(u) &= P([0 < T < \infty] \cap [U(t) < 0]) = P([0 < T < \infty] \cap [u + ct - S(t) < 0]) \\ &= P([0 < T < \infty] \cap [X_1 > u + ct]) = \int_0^\infty k(t) \int_{u+ct}^\infty p(x) dx dt.\end{aligned}$$

Já que por definição R é positivo e $x > u + ct$, logo, $u + ct - x < 0$, e isso implica que $e^{-R(u+ct-x)}$ seja uma valor maior que 1. Portanto,

$$\psi_1(u) \leq \int_0^\infty k(t) \int_{u+ct}^\infty e^{-R(u+ct-x)} p(x) dx dt.$$

Aumentando o limite de integração, obtém-se

$$\begin{aligned}\psi_1(u) &\leq \int_0^\infty k(t) \int_0^\infty e^{-R(u+ct-x)} p(x) dx dt \\ &= e^{-Ru} \int_0^\infty k(t) e^{-Rct} \int_0^\infty p(x) dx dt \\ &= e^{-Ru} E(e^{-cRT_i}) E(e^{RX_i}) = e^{-Ru};\end{aligned}$$

esse resultado se deve ao fato de que o produto destes dois valores esperados é igual a 1.

Para o caso de $n = 2$, usando raciocínio análogo, onde a ruína pode acontecer na primeira ou na segunda indenização, temos:

$$\psi_2(u) = P([0 < T < \infty] \cap [X_1 > u+ct]) + P([0 < T < \infty] \cap [X_1 > u+ct] \cap (X_2 \text{ ocorre ruína}))$$

Conseqüentemente, se a ruína ocorreu na primeira indenização, a segunda indenização ocorrerá com o sistema em ruína com probabilidade igual a 1.

$$\psi_2(u) = \int_0^\infty k(t) \int_{u+ct}^\infty p(x) dx dt + \int_0^\infty k(t) \int_0^{u+ct} p(x) \psi_1(u + ct - x) dx dt$$

$$\begin{aligned}
\psi_2(u) &\leq \int_0^\infty \int_{u+ct}^\infty e^{-R(u+ct-x)} p(x) dx dt + \int_0^\infty \int_0^{u+ct} e^{-R(u+ct-x)} p(x) dx dt \\
&= \int_0^\infty k(t) \left[\int_{u+ct}^\infty e^{-R(u+ct-x)} p(x) dx dt + \int_0^{u+ct} e^{-R(u+ct-x)} p(x) dx dt \right] \\
&= \int_0^\infty k(t) \int_0^\infty e^{-R(u+ct-x)} p(x) dx dt = e^{-Ru} \int_0^\infty e^{-Rct} k(t) dt \int_0^\infty e^{rx} p(x) dx \\
&= e^{-Ru} E(e^{-cRT_i}) E(e^{RX_i}) = e^{-Ru},
\end{aligned}$$

assim como observado em $\psi_1(u)$. Admitindo válido para n , é necessário provar para $n + 1$. De forma análoga e, considerando que a ruína acontece da primeira ou nas n indenizações seguintes, segue

$$\begin{aligned}
\psi_{n+1}(u) &= \int_0^\infty k(t) \int_{u+ct}^\infty p(x) dx dt + \int_0^\infty k(t) \int_0^{u+ct} p(x) \psi_n(u + ct - x) dx dt \\
&\leq \int_0^\infty k(t) \int_{u+ct}^\infty e^{-R(u+ct-x)} p(x) dx dt + \int_0^\infty k(t) \int_0^{u+ct} e^{-R(u+ct-x)} p(x) dx dt \\
&= \int_0^\infty k(t) \left[\int_{u+ct}^\infty e^{-R(u+ct-x)} p(x) dx dt + \int_0^{u+ct} e^{-R(u+ct-x)} p(x) dx dt \right] \\
&= \int_0^\infty k(t) \int_0^\infty e^{-R(u+ct-x)} p(x) dx dt = e^{-Ru} \int_0^\infty k(t) e^{-Rct} dt \int_0^\infty e^{Rx} p(x) dx \\
&= e^{-Ru} E(e^{-cRT_i}) E(e^{RX_i}) = e^{-Ru}.
\end{aligned}$$

Consequentemente, $\psi_n(u) \leq e^{-Ru}$. □

2.2 Conjuntos *fuzzy*

A teoria dos conjuntos *fuzzy*, introduzida por Zadeh (1965), é uma extensão da teoria dos conjuntos clássicos, e seu principal objetivo é dar tratamento matemático para conjuntos cujos elementos possuem graus de associação ou níveis de pertinência (BERGMEIR, BEN, 2015). Se o nível de pertinência de um elemento é igual a 1, então isso significa que o elemento é membro do conjunto; se o valor for 0 o elemento não pertence ao conjunto; e valores nesse intervalo mostram um grau parcial de pertencimento ao conjunto. O nível de pertinência de um dado elemento é definido pela função de pertinência, que caracteriza o conjunto *fuzzy* (BERGMEIR, BEN, 2015).

A teoria proposta por Zadeh (1965) apresenta este novo conceito de conjunto, que generaliza o conceito clássico, e consequentemente, exige adequações nas definições para as operações com conjuntos, ou seja, união e intersecção, complementares e assim por diante. Isso, por sua vez, levou à extensão de muitos outros conceitos, como número,

intervalo, equação etc (RIZA et al., 2015).

A teoria dos conjuntos *fuzzy* fornece as ferramentas para dar tratamento matemático a conceitos, variáveis e regras imprecisos, provenientes de conceitos da linguagem humana, que é inerentemente vaga, tornando-se um modelo natural para representar o conhecimento humano (BARROS, BASSANEZI, 2006).

Um importante conceito é a variável linguística, definida como uma variável cujos valores são termos linguísticos, com uma semântica descrita por um conjunto *fuzzy* (ZADEH, 1975). Um valor de uma variável linguística representa conhecimento e tem significado determinado pelo seu grau da função de pertinência (BERGMEIR, BEN, 2015). Por exemplo, na Figura 3, a variável “valor de indenização”, definida no intervalo [1.000,15.000], pode ser representada pelos conjuntos *fuzzy* baixo, médio e alto, nos intervalos [1.000, 8.000], [6.000, 12.000] e [10000,15000], respectivamente. No caso, um valor de indenização igual a 7.500 (linha tracejada) pertencerá a cada conjunto *fuzzy* com diferentes graus de pertinência.

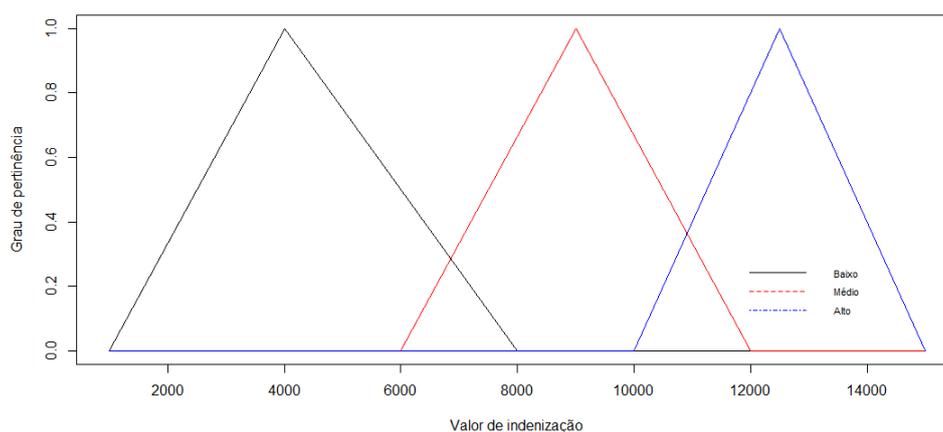


Figura 3: Exemplo de conjuntos *fuzzy* para a variável “valor de indenização”.
Fonte: Elaboração própria.

Nos últimos anos, a pesquisa nessa área tem crescido constantemente e a literatura é vasta. Para mais aplicações de lógica *fuzzy* no contexto atuarial ver Shapiro (2004).

2.2.1 Subconjuntos *fuzzy* e α -nível

Seja U um conjunto clássico e A um subconjunto *fuzzy* de U . A função $\varphi_A : U \rightarrow [0, 1]$ é definida como função de pertinência do subconjunto *fuzzy* A , sendo que $\varphi_A(x) \in [0, 1]$ indica o grau de pertinência do elemento x de U no conjunto *fuzzy* A . O α -nível de A é o subconjunto clássico de U definido por

$$[A]^\alpha = \{x \in U : \varphi_A(x) \geq \alpha\}, \text{ para } 0 < \alpha \leq 1.$$

Além disso, define-se o nível zero de um subconjunto *fuzzy*, $[A]^0$, como o fecho do suporte de A , que é indicado por $\overline{\text{supp}A}$.

2.2.2 Princípio de extensão de Zadeh

O princípio de extensão de Zadeh nos dá uma ferramenta para estender conceitos matemáticos da teoria dos conjuntos clássicos para a teoria dos conjuntos *fuzzy* (BARROS; BASSANEZI, 2010).

Definição: Seja f uma função clássica tal que $f : X \rightarrow Z$ e A um subconjunto *fuzzy* de X . A extensão de Zadeh de f é a função \hat{f} que, aplicada a A , fornece o subconjunto *fuzzy* $\hat{f}(A)$ de Z , cuja função de pertinência é dada por

$$\varphi_{\hat{f}(A)}(z) = \begin{cases} \sup_{f^{-1}(z)} \varphi_A(x) & \text{se } f^{-1}(z) \neq \emptyset, \\ 0 & \text{se } f^{-1}(z) = \emptyset. \end{cases} \quad (4)$$

$f^{-1}(z) = \{x; f(x) = z\}$ denomina-se a pré-imagem de z .

Teorema 2.2 *Sejam $f : X \rightarrow Z$ uma função contínua e A um subconjunto fuzzy de X . Então, para todo $\alpha \in [0, 1]$ vale*

$$[\hat{f}(A)]^\alpha = f([A]^\alpha).$$

O Teorema 2.2 indica que os α -níveis do conjunto *fuzzy*, obtido pelo princípio de extensão de Zadeh, coincidem com as imagens dos α -níveis da função clássica (BARROS; BASSANEZI, 2010).

2.2.3 Números fuzzy

Um número *fuzzy* pode ser definido como um subconjunto *fuzzy* A quando o conjunto universo no qual φ_A está definida é o conjunto dos números reais \mathbb{R} e satisfaz às seguintes condições:

1. todos os α -níveis de A são não vazios, com $0 \leq \alpha \leq 1$;
2. todos os α -níveis de A são intervalos fechados de \mathbb{R} ;
3. $\text{supp } A = \{x \in \mathbb{R} : \varphi_A(x) > 0\}$ é limitado.

Os α -níveis de um número *fuzzy* A podem ser denotados como

$$[A]^\alpha = [a_1^\alpha, a_2^\alpha], \text{ para todo } \alpha \in [0, 1].$$

Os números *fuzzy* mais comuns são os triangulares, trapezoidais e os gaussianos. Um número *fuzzy* A é da forma triangular se sua função de pertinência é da forma

$$\varphi_A(x) = \begin{cases} 0, & \text{se } x \leq a; \\ \frac{x-a}{u-a}, & \text{se } a < x \leq u; \\ \frac{b-x}{b-u}, & \text{se } u < x \leq b; \\ 0, & \text{se } x \geq b. \end{cases} \quad (5)$$

Um número *fuzzy* triangular pode ser representado graficamente por um triângulo com a base sendo o intervalo $[a, b]$ e com vértice no ponto $(u, 1)$. Assim, os números reais a , u e b definem um número *fuzzy* triangular A que pode ser denotado pela terna $(a; u; b)$.

Os α -níveis de um número *fuzzy* triangular podem ser apresentados de forma simplificada como

$$[a_1^\alpha, a_2^\alpha] = [(u - a)\alpha + a, -(b - u)\alpha + b],$$

para todo $\alpha \in [0, 1]$.

Como exemplo, a Figura 4 apresenta o número *fuzzy* triangular $(4; 6; 8)$ para a variável x .

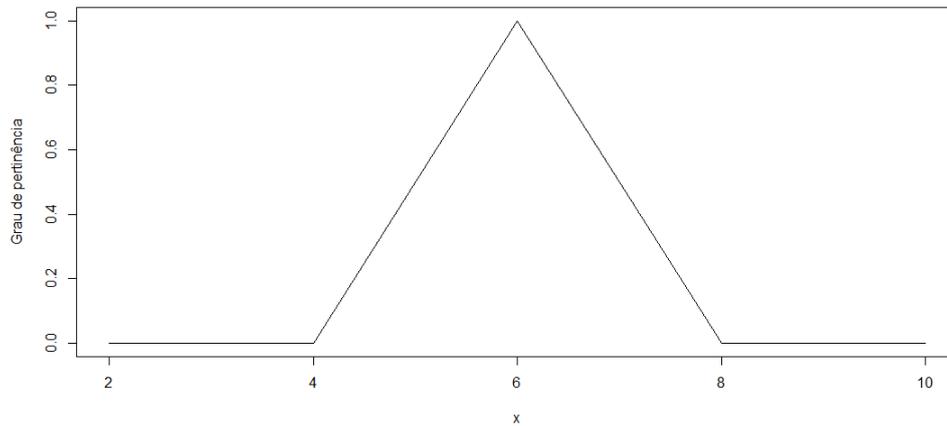


Figura 4: Exemplo de número *fuzzy* triangular.
Fonte: Elaboração própria.

2.2.4 Sistemas baseados em regras *fuzzy*

Sistemas baseados em regras *fuzzy* (SBRF) (em inglês *Fuzzy Rule-Based Systems - FRBS*) são uma extensão de sistemas baseados em regras clássicas e são basicamente regras expressas na forma “SE A ENTÃO B”. Nos SBRFs *fuzzy* A e B são conjuntos *fuzzy*, sendo A e B chamados de parte antecedente e parte consequentes da regra, respectivamente (RIZA et al., 2015).

Os sistemas baseados em regras *fuzzy* envolvem em sua aplicação uma base de conhecimento e um processador. A base de conhecimento fornece funções de pertinência e regras *fuzzy* necessárias para o processamento das variáveis de entrada (SHAPIRO, 2004).

Na fase de processamento, utilizando uma base de dados, as variáveis de entrada do sistema passam pelo módulo de fuzzificação, onde são transformados em conjuntos *fuzzy*. No módulo de inferência, as variáveis de entrada e saída são conectadas, criando-se a base de regras. Os conjuntos *fuzzy* resultantes (saída *fuzzy*), do módulo de inferência, são então transformados, pelo módulo de defuzzificação, em valores numéricos (saída crisp) e se tornam a saída do sistema. A Figura 5 ilustra o processo de inferência nos sistemas baseados em regras *fuzzy* (SHAPIRO, 2004).

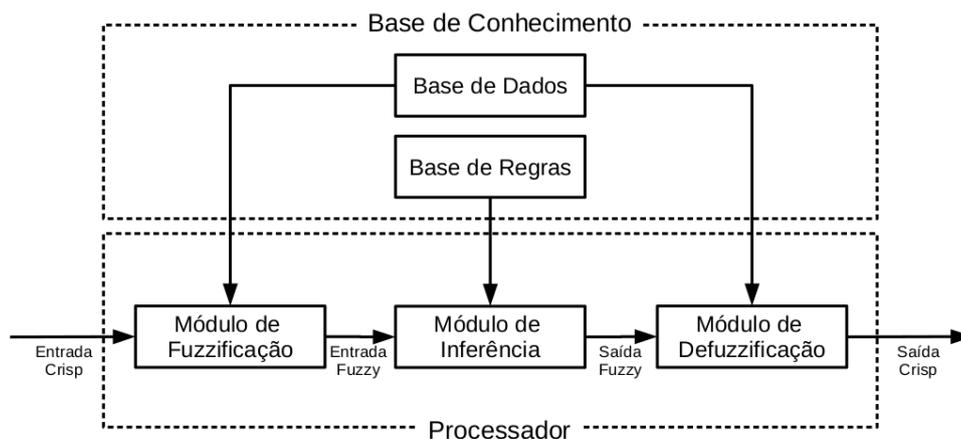


Figura 5: Sistema baseado em regras *fuzzy*.

Fonte: adaptado de Shapiro (2004).

2.2.5 O método de inferência de Mandani

Introduzido por Mamdani (1974) e Mamdani e Assilian (1975) este tipo de modelo é constituído por conjuntos *fuzzy* em ambas as partes, antecedentes e consequentes, do

sistema de regras. Além disso, é constituído por várias variáveis de entradas e uma única variável de saída. As regras *fuzzy* são da seguinte forma:

$$\text{SE } X_1 \text{ é } A_1 \text{ e } \dots \text{ e } X_n \text{ é } A_n \text{ ENTÃO } Y \text{ é } B, \quad (6)$$

em que X_i e Y são variáveis de entrada e saída, respectivamente, e A_i e B são conjuntos *fuzzy*.

O modelo Mamdani conta com quatro componentes: fuzzificação, base de conhecimento, mecanismo de inferência e defuzzificador.

O processo de fuzzificação transforma as variáveis de entradas *crisp* em valores linguísticos representados por conjuntos *fuzzy*. A base de conhecimento é constituída de um banco de informações, que contém as definições dos conjuntos *fuzzy* e os parâmetros das funções de pertinência, e uma base de regras do tipo “SE ENTÃO”. O mecanismo de inferência executa as operações do conjunto de regras nos dados de entrada. Por sua vez, o defuzzificador produz valores *crisp* dos conjuntos *fuzzy* resultantes do processo de inferência. Uma vantagem fundamental do modelo Mamdani é sua maior interpretabilidade e flexibilidade para formular a base de conhecimento (RIZA et al., 2015).

2.2.6 Métodos de defuzzificação

Nos sistemas baseados em regras *fuzzy*, o módulo de inferência recebe uma variável de entrada e produz uma variável de saída, ambas *fuzzy*. Entretanto, se a variável de entrada do sistema for um número real, é esperado que a saída correspondente também seja um número real. Para que isso ocorra é necessário fazer uso de um método de defuzzificação, que transforma um conjunto *fuzzy* em um número real (BARROS, BASSANEZI, 2006).

Centro de gravidade

O centro de gravidade, também conhecido como centroide, é um dos métodos mais utilizados na literatura. O cálculo desse método se assemelha com o cálculo da esperança de uma variável aleatória e fornece a média da área dos gráficos que representam os graus de pertinência de um conjunto *fuzzy* (BARROS, BASSANEZI, 2006). A Equação 7 apresenta a forma de cálculo do centro de gravidade no caso discreto:

$$G(B) = \frac{\sum_{i=0}^n x_i \varphi_B(x_i)}{\sum_{i=0}^n \varphi_B(x_i)}, \quad (7)$$

em que $\varphi_B(x_i)$ representa os pesos que indicam o grau de pertencimento do valor x_i à variável modelada pelo conjunto *fuzzy* B . A Equação 8 apresenta a forma de cálculo do centro de gravidade no caso contínuo:

$$G(B) = \frac{\int_{\mathbb{R}} x\varphi_B(x)dx}{\int_{\mathbb{R}} \varphi_B(x)dx}, \quad (8)$$

em que \mathbb{R} é a região de integração.

3 Metodologia

Nesta seção, serão apresentadas as três abordagens propostas para modelagem do limite superior da probabilidade de ruína. Na primeira abordagem (modelagem clássica), serão apresentados: o cálculo do coeficiente de ajustamento (R) e do limite superior da probabilidade de ruína para os casos em que o valor das indenizações (X) segue a distribuição exponencial ($X \sim \text{Exponencial}(\beta)$) e a distribuição gama ($X \sim \text{Gama}(\alpha, \eta)$).

Na segunda abordagem (modelagem *fuzzy* 1), serão apresentados os SBRFs utilizados para modelar o limite superior da probabilidade de ruína e o pacote *frbs*, utilizado para a construção dos SBRFs. Na terceira abordagem (modelagem *fuzzy* 2), será apresentada a extensão do modelo clássico do cálculo do limite superior da probabilidade de ruína para o contexto *fuzzy*, via princípio de extensão de Zadeh, considerando que o valor da indenização segue uma distribuição exponencial com parâmetro incerto.

3.1 Recursos computacionais

Esse estudo foi feito utilizando a linguagem de programação R (R DEVELOPMENT CORE TEAM, 2020) através do programa gratuito e de código aberto RStudio (RSTUDIO, 2020), que é um ambiente de desenvolvimento integrado e tem por objetivo tornar o uso do R mais produtivo e intuitivo.

Foram utilizados três pacotes, sendo eles: *ggplot2*, que é destinado à visualização de dados; *dplyr*, que fornece um conjunto consistente de funções destinadas à manipulação de dados, e *frbs*, que implementa os sistemas baseados em regras *fuzzy* mais utilizados, tais como, Mamdani e Takagi Sugeno Kang, além de algumas variantes (WICKHAM, 2009; WICKHAM; FRANCOIS, 2015; RIZA et al., 2015).

3.1.1 O pacote *frbs*

O pacote *frbs* visa fornecer os modelos de SBRFs mais comuns e também implementar os procedimentos de aprendizagem mais amplamente utilizados (RIZA et al.,

2014). Esse pacote traz a vantagem da possibilidade de criação das regras *fuzzy* através de métodos de aprendizado e, além disso, possibilita a criação manual a partir de conhecimento especializado (RIZA et al., 2015).

O pacote *frbs* permite que sejam construídos modelos como o Mamdani e o Takagi Sugeno Kang, entre outros. Em relação às abordagens de aprendizagem para construir um SBRF, o pacote apresenta cinco grupos principais: *frbs* baseado na partição espacial, que usa a estratégia de dividir o espaço da variável e, a partir dessas divisões, obter os parâmetros das funções de pertinência; *frbs* baseado no gradiente descendente, que faz uso da abordagem de gradiente descendente para otimizar parâmetros nas partes antecedentes e consequentes das regras; *frbs* baseado em algoritmos genéticos, que se refere aos sistemas genéticos *fuzzy*; que é uma combinação do SBRF com algoritmos genéticos, onde os algoritmos genéticos são usados para otimizar os parâmetros das funções de pertinência; também há os *frbs* baseado em redes neurais e *frbs* baseado em *clustering*.

Este trabalho fez uso do método de aprendizagem baseados em algoritmos genéticos, mais especificamente, o método de ajustamento genético lateral (método GFS.LT.RS) (RIZA et al., 2015). O ajustamento consiste em aprimorar a definição da base de dados, uma vez que a base de regras foi obtida. Ele se dá através de uma variação na forma das funções de pertinência o que melhora a interação global das funções e têm o objetivo de induzir uma maior cooperação entre as regras. O objetivo do ajustamento é encontrar a melhor configuração global das funções de pertinência, e não apenas encontrar funções de pertinência independentes (ALCALÁ, ALCALÁ-FDEZ, HERRERA, 2007).

O ajustamento lateral, por sua vez, permite com que as funções de pertinência sejam deslocadas lateralmente, considerando apenas um parâmetro, o que simplifica a otimização do modelo (RIZA et al., 2015).

3.2 Modelagem clássica

A modelagem clássica é constituída por simulações de Monte Carlo para que geram os valores de indenização. A partir desses valores, através dos coeficientes de ajustamento, é calculado o limite superior da probabilidade de ruína. Os valores de indenização simulados também são utilizados como variável de entrada dos SBRFs (modelagem *fuzzy* 1).

A Tabela 1 apresenta as simulações de Monte Carlo. Em todas as simulações, a variável aleatória “número de sinistros ocorridos (N)” segue uma distribuição de Poisson, com parâmetro λ igual a 1 ($N \sim \text{Poisson}(1)$).

Tabela 1: Simulações para os valores das indenizações.

Simulação	Distribuição	Parâmetros
1	$X \sim \text{Exponencial}(\beta)$	$u = 1, c = 3, \lambda = 1$ e $\beta \in [0, 4; 0, 8]$
2	$X \sim \text{Gama}(\alpha, \eta)$	$u = 1, c = 3, \lambda = 1, \alpha = 2, 5$ e $\eta \in [1, 8; 2, 2]$
3	$X \sim \text{Gama}(\alpha, \eta)$	$u = 1, c = 3, \lambda = 1, \eta = 2$ e $\alpha \in [2, 0; 3, 0]$

3.2.1 Indenizações com distribuição exponencial

Para os casos em que o valor das indenizações tem distribuição exponencial, ou seja, $X \sim \text{Exponencial}(\beta)$, o valor do coeficiente de ajustamento R pode ser obtido através de uma solução fechada. Partindo da definição apresentada na Equação 2, tem-se que $e^{rc} = M_S(r)$, com $M_S(r) = M_N[\ln M_X(r)]$.

Substituindo a função geradora de momentos da distribuição exponencial, tem-se que:

$$rc = \lambda[M_X(r) - 1] = \lambda \left[\frac{\beta}{\beta - r} - 1 \right] = \lambda \left[\frac{\beta - \beta + r}{\beta - r} \right] = \frac{\lambda r}{\beta - r}.$$

Isolando R , segue que:

$$R = \frac{-\lambda + c\beta}{c}.$$

que é o coeficiente de ajustamento quando $X \sim \text{Exponencial}(\beta)$ e $N \sim \text{Poisson}(\lambda)$.

3.2.2 Indenizações com distribuição gama

Para os casos em que o valor das indenizações tem distribuição gama, ou seja, $X \sim \text{gama}(\alpha, \eta)$, tem-se que

$$e^{rc} = M_S(r) = M_N[\ln M_X(r)] = M_N \left[\ln \left(\frac{\eta}{\eta - r} \right)^\alpha \right] = e^{\lambda \left[e^{\ln \left(\frac{\eta}{\eta - r} \right)^\alpha} - 1 \right]},$$

$$rc = \lambda \left[\left(\frac{\eta}{\eta - r} \right)^\alpha - 1 \right]. \quad (9)$$

Como essa a Equação 9 não possui solução exata, é necessário recorrer a uma rotina numérica para encontrar suas raízes (KAAS et al., 2008). Nesse trabalho foi utilizada a função *uniroot* da linguagem de programação R, considerando que

$$rc = \lambda \left[\left(\frac{\eta}{\eta - r} \right)^\alpha - 1 \right],$$

$$rc = \lambda(M_X(r) - 1),$$

$$\lambda M_X(r) - \lambda - rc = 0.$$

3.3 Modelagem *fuzzy* 1

A Tabela 2 apresenta os três modelos propostos baseados em SBRFs. O primeiro utiliza como variável de entrada valores da indenização advindos da Simulação 1 (Tabela 1), em que $X \sim Exponencial(\beta)$. Os modelos 2 e 3 utilizam como variável de entrada valores de indenização advindos das simulações 2 e 3 (Tabela 1), sendo que ($X \sim Gama(\alpha, \eta)$). No modelo 2, o parâmetro α é considerado fixo e no modelo 3, o parâmetro η é fixo. Os SBRFs foram criados utilizando a função de pertinência gaussiana, método de inferência de Mamdani e método defuzzificação centro de gravidade.

Tabela 2: Variáveis de entrada e saída utilizadas nos SBRFs - modelagem *fuzzy* 1.

Modelo	Variável de entrada	Variável de saída
1	$X \sim Exponencial(\beta)$	ψ
2	$X \sim Gama(\alpha, \eta)$, α fixo	ψ
3	$X \sim Gama(\alpha, \eta)$, η fixo	ψ

A Figura 6 ilustra o SBRF criado para encontrar o limite superior da probabilidade de ruína utilizando como variável de entrada o valor da indenização (X). Nesse sistema, a base de regras é construída por meio de um método de aprendizagem baseado em algoritmos genéticos por meio do pacote *frbs*, no *software* R.

Para a implementação dos SBRFs via pacote *frbs*, foram consideradas 10000 simulações de Monte Carlo de cada distribuição de X , sendo 60% das simulações utilizadas como dados de treinamento e 40% como dados de teste. As funções de pertinência gaussianas foram geradas automaticamente pelo pacote *frbs*.

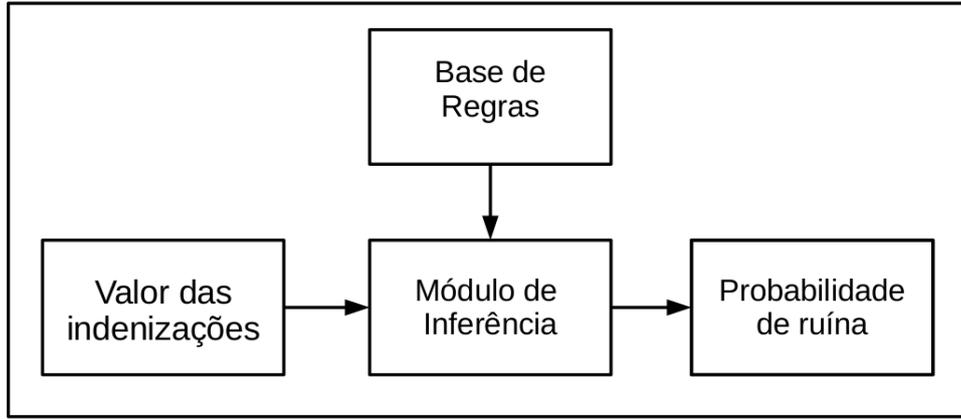


Figura 6: SBRF para a probabilidade de ruína - modelagem *fuzzy* 1.
Fonte: Elaboração própria.

3.4 Modelagem *fuzzy* 2

A modelagem *fuzzy* 2 buscou estender o modelo probabilístico de Cramér-Lundberg para a teoria dos conjuntos *fuzzy*, via princípio de extensão de Zadeh. Para tanto, foi considerado que o valor das indenizações segue uma distribuição exponencial com parâmetro β incerto, representado por $\tilde{\beta}$. Dessa maneira, $X \sim \text{Exponencial}(\tilde{\beta})$, sendo $\tilde{\beta}$ um número *fuzzy* triangular representado pela terna $(a; u; b)$, com função de pertinência dada por

$$\varphi_{\tilde{\beta}}(\beta) = \begin{cases} 0, & \text{se } \beta \leq a; \\ \frac{\beta-a}{u-a}, & \text{se } a \leq \beta \leq u; \\ \frac{b-\beta}{b-u}, & \text{se } u < \beta \leq b; \\ 0, & \text{se } \beta \geq b. \end{cases} \quad (10)$$

Os α -níveis de $\tilde{\beta}$ são os intervalos

$$\begin{aligned} [\tilde{\beta}]^\alpha &= [\tilde{\beta}_{\text{inf}}^\alpha, \tilde{\beta}_{\text{sup}}^\alpha] \\ &= [(u-a)\alpha + a, -(b-u)\alpha + b], \end{aligned}$$

para todo $\alpha \in [0, 1]$.

Utilizando o Princípio de Extensão de Zadeh e o Teorema 2.2, tem-se que os α -níveis do limite superior da probabilidade de ruína em função de $\tilde{\beta}$ são dados pelos intervalos:

$$\begin{aligned}\psi([\tilde{\beta}]^\alpha) &= [\psi(\tilde{\beta})]^\alpha \\ &= \left[\exp \left\{ - \left(\frac{-\lambda + c\tilde{\beta}_{sup}^\alpha}{c} \right) \right\}, \exp \left\{ - \left(\frac{-\lambda + c\tilde{\beta}_{inf}^\alpha}{c} \right) \right\} \right].\end{aligned}$$

4 Resultados

4.1 Modelagem clássica

4.1.1 Indenizações com distribuição exponencial

A Tabela 3 apresenta os resultados obtidos para o limite superior da probabilidade de ruína pelo modelo de Cramér-Lundberg, considerando que as indenizações são variáveis aleatórias com distribuição exponencial ($X \sim \text{Exponencial}(\beta)$). Os resultados da Tabela 3 foram alcançados considerando a reserva inicial igual a 1, a taxa de recebimento dos prêmios da seguradora constante e igual a 3 e o parâmetro $\lambda = 1$ ($N \sim \text{Poisson}(1)$). Como esperado, observa-se que o limite superior da probabilidade de ruína diminui quando o valor de β aumenta.

A Figura 7 ilustra a relação entre o limite superior da probabilidade de ruína e os valores de β utilizados para as simulações. A Figura 7 ilustra os resultados apresentados na Tabela 3 e observa-se que a probabilidade de ruína diminui à medida que o valor do β é maior.

Tabela 3: Estimativas do limite superior da probabilidade de ruína para indenizações com distribuição exponencial para $u = 1$, $c = 3$, $\lambda = 1$ e diferentes valores de β .

β	\bar{x}	$\psi(u)$
0,4000000	2,471234	0,8751733
0,4444444	2,208988	0,8007374
0,4888889	2,067006	0,7326325
0,5333333	1,872354	0,6703200
0,5777778	1,709096	0,6133075
0,6222222	1,611897	0,5611440
0,6666667	1,508938	0,5134171
0,7111111	1,413843	0,4697496
0,7555556	1,332563	0,4297961
0,8000000	1,258115	0,3932407

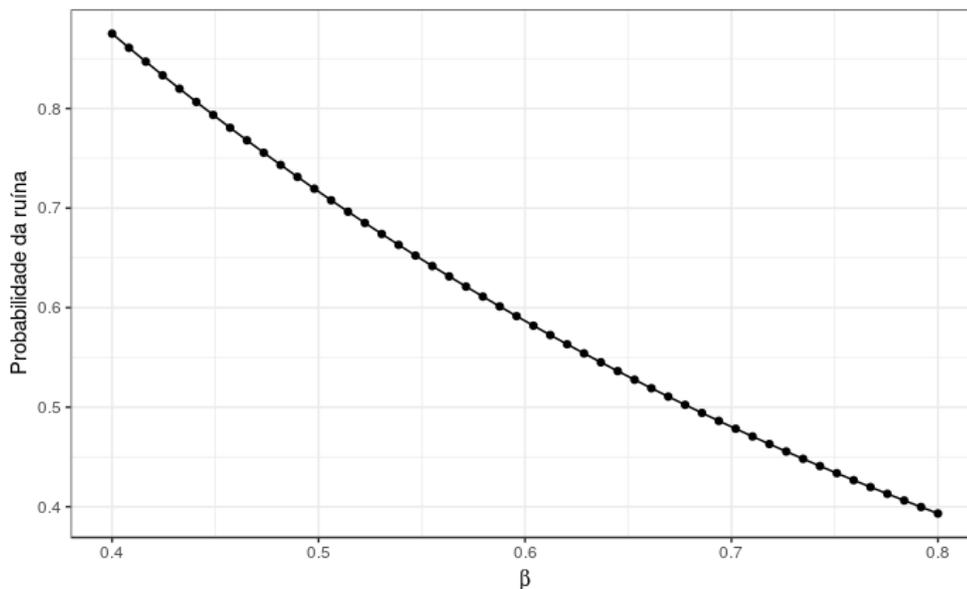


Figura 7: Relação entre o limite superior da probabilidade de ruína e valores de β , para $u = 1$, $c = 3$, $\lambda = 1$.

Fonte: Elaboração própria.

4.1.2 Indenizações com distribuição gama

A Tabela 4 apresenta os resultados obtidos para a probabilidade de ruína modelada por Cramér-Lundberg, considerando que as indenizações são variáveis aleatórias com distribuição gama com parâmetro $\alpha = 2,5$ ($X \sim \text{Gama}(\alpha, \eta)$), variando os valores de η . A Tabela 5 apresenta os resultados obtidos para a probabilidade de ruína, variando os valores de α , considerando o parâmetro η igual a 2. Os resultados foram obtidos considerando a reserva inicial igual a 1, a taxa de recebimento dos prêmios da seguradora constante e igual a 3 e o parâmetro $\lambda = 1$ ($N \sim \text{Poisson}(1)$). Como esperado, observa-se que a probabilidade de ruína diminui quando o valor de η e α aumentam, como pode ser observado nas Figuras 8 e 9, respectivamente.

Tabela 4: Estimativas do limite superior da probabilidade de ruína para indenizações com distribuição gama, para $u = 1$, $c = 3$, $\lambda = 1$, $\alpha = 2, 5$ e diferentes valores de η .

η	\bar{x}	$\psi(u)$
1,800000	1,383844	0,5404103
1,844444	1,354722	0,5242143
1,888889	1,340873	0,5084404
1,933333	1,289743	0,4930753
1,977778	1,264809	0,4781077
2,022222	1,231806	0,4635544
2,066667	1,221883	0,4493689
2,111111	1,197245	0,4355754
2,155556	1,152902	0,4221615
2,200000	1,140412	0,4091165

Tabela 5: Estimativas do limite superior da probabilidade de ruína para indenizações com distribuição gama, para $u = 1$, $c = 3$, $\lambda = 1$, $\eta = 2$ e diferentes valores de α .

α	\bar{x}	$\psi(u)$
2,000000	1,005679	0,3678794
2,111111	1,052645	0,3906372
2,222222	1,107169	0,4135475
2,333333	1,168451	0,4364708
2,444444	1,220689	0,4593668
2,555556	1,278819	0,4821477
2,666667	1,342679	0,5047595
2,777778	1,386829	0,5271369
2,888889	1,444785	0,5492426
3,000000	1,504786	0,5710304

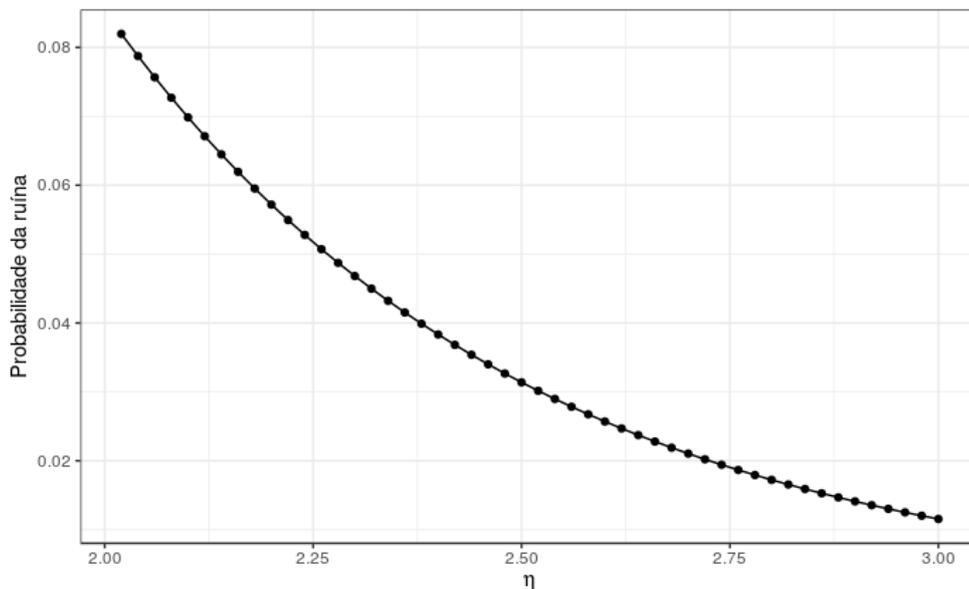


Figura 8: Relação entre o limite superior da probabilidade de ruína e valores de η , para $\alpha = 2,5$, $u = 1$, $c = 3$ e $\lambda = 1$.

Fonte: Elaboração própria.

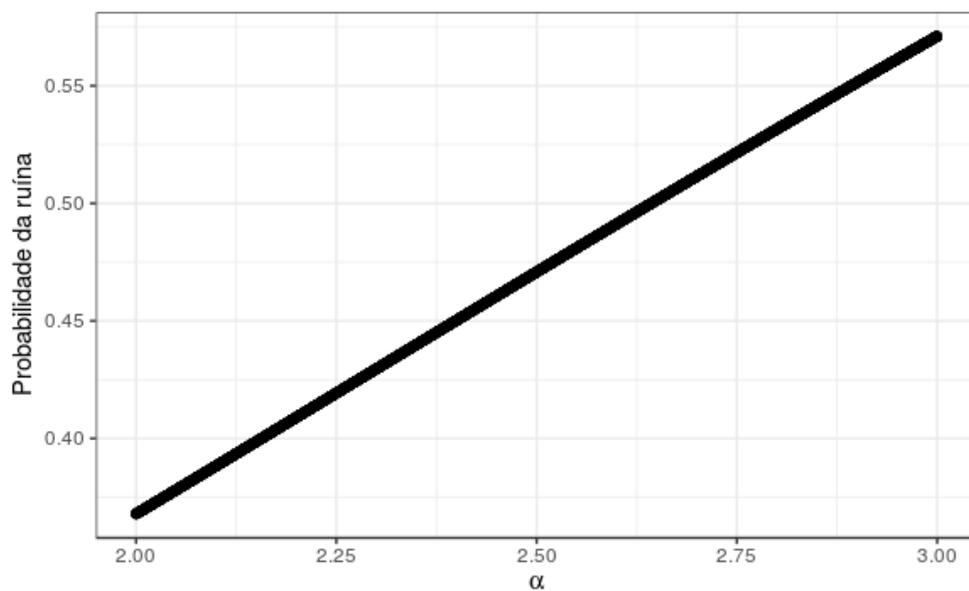


Figura 9: Relação entre o limite superior da probabilidade de ruína e valores de α , para $\eta = 2$, $u = 1$, $c = 3$ e $\lambda = 1$.

Fonte: Elaboração própria.

4.2 Modelagem *fuzzy* 1

A Tabela 6 traz os resultados dos SBRFs criados com 5, 10 e 15 conjuntos *fuzzy* para os casos em que a variável de entrada é o valor das indenizações com distribuição

exponencial e distribuição gama. É possível observar que o erro quadrado médio, tanto dos dados de treinamento quanto de teste alcançam valores menores à medida que mais conjuntos *fuzzy* são utilizados. Além disso, os coeficientes de determinação (R^2) em ambos os SBRFs tendem a aumentar quando mais conjuntos *fuzzy* são adotados e, em todos os casos, apresentam valores acima de 95%, o que indica que os SBRFs têm boa capacidade de previsão do limite superior da probabilidade de ruína.

Tabela 6: Estatísticas dos SBRFs para diferentes distribuições do valor das indenizações.

Modelo 1 - Indenizações com distribuição exponencial ($u = 1, c = 3, \lambda = 1$ e diferentes valores de β)				
Nº de conjuntos <i>fuzzy</i>	R^2 treinamento	Erro treinamento	R^2 teste	Erro teste
5	0,9874869	0,0028093	0,9889606	0,0027673
10	0,9741854	0,0007739	0,9909758	0,0007576
15	0,9953996	0,0003897	0,9953120	0,0004043
Modelo 2 - Indenizações com distribuição gama ($u = 1, c = 3, \lambda = 1, \alpha = 2, 5$ e diferentes valores de η)				
Nº de conjuntos <i>fuzzy</i>	R^2 treinamento	Erro treinamento	R^2 teste	Erro teste
5	0,9398887	0,0070875	0,9406672	0,0072997
10	0,9786645	0,0019749	0,9769295	0,0019738
15	0,9846479	0,0013135	0,9856412	0,0012655
Modelo 3 - Indenizações com distribuição gama ($u = 1, c = 3, \lambda = 1, \eta = 2$ e diferentes valores de α)				
Nº de conjuntos <i>fuzzy</i>	R^2 treinamento	Erro treinamento	R^2 teste	Erro teste
5	0,9903475	0,0040186	0,9899505	0,0038129
10	0,9946615	0,0008708	0,9945196	0,0008789
15	0,9943193	0,0005918	0,9943019	0,0005952

As Figuras 10 a 12 apresentam os resultados alcançados pelos SBRFs com 15 conjuntos *fuzzy* referentes a cada modelo proposto, pois com esse número de conjuntos *fuzzy* foram obtidos os melhores resultados, conforme Tabela 6. A Figura 10 apresenta os resultados do SBRF para os dados de treinamento e teste para o caso em que o valor das indenizações tem distribuição exponencial, considerando $u = 1$, $c = 3$, $\lambda = 1$ e diversos valores de β . Na Figura 11, são apresentados os resultados do SBRF para os dados de treinamento e teste para o cenário em que o valor das indenizações tem distribuição gama, com $u = 1$, $c = 3$, $\lambda = 1$ e $\alpha = 2, 5$, com diferentes valores de η . Os resultados do SBRF para os dados de treinamento e teste para o cenário em que o valor das indenizações tem distribuição gama, com $u = 1$, $c = 3$, $\lambda = 1$ e $\eta = 2$, com diversos valores de α , são apresentados na Figura 12.

As Figuras 10 a 12 apresentam a relação entre os valores simulados por Monte Carlo e os valores simulados pelos SBRFs, permitindo avaliar a qualidade do ajuste dos modelos. Verifica-se que a dispersão dos dados ajustados pelo SBRFs se aproxima da dispersão dos dados simulados por Monte Carlo em todos os modelos. Nas Figuras 10 a 12, é possível observar que, tanto para os dados de treinamento quanto de teste, os SBRFs apresentaram boa capacidade de ajuste.

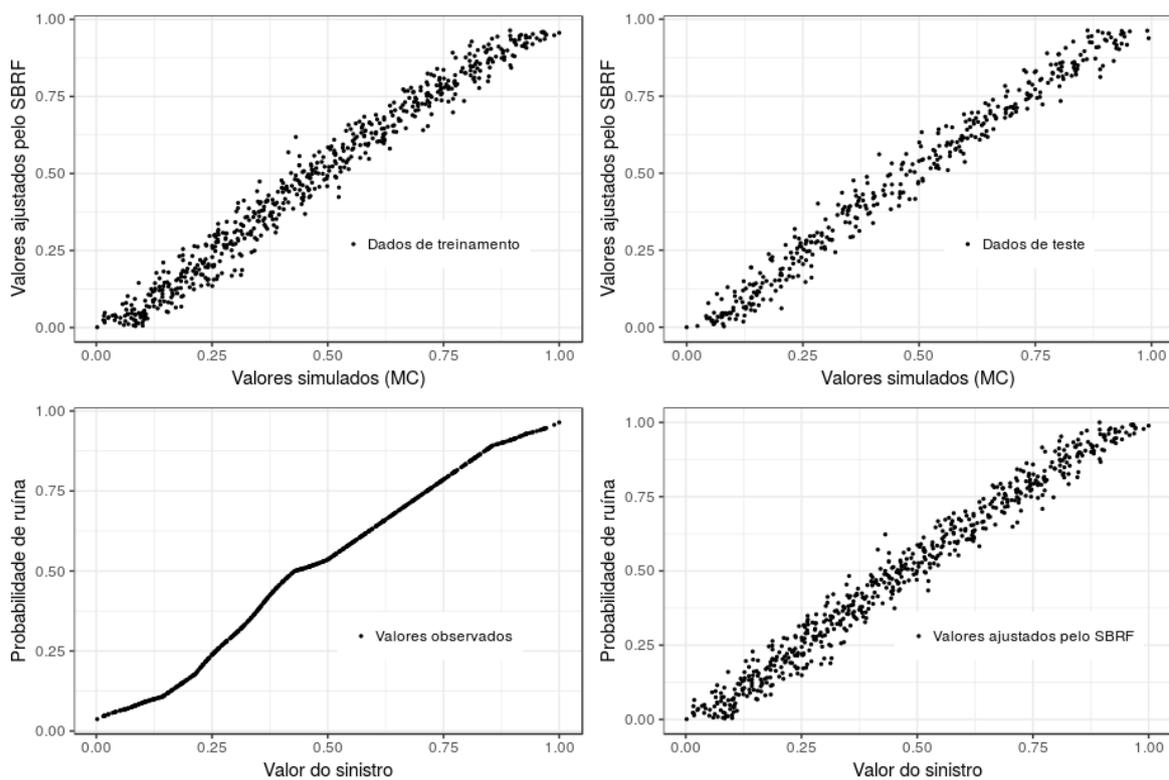


Figura 10: Valores ajustados pelo SBRF (15 conjuntos *fuzzy*) em função dos valores simulados por Monte Carlo para dados ($X \sim \text{Exponencial}(\beta)$) de treinamento e teste e estimativas do limite superior da probabilidade de ruína.

Fonte: Elaboração própria.

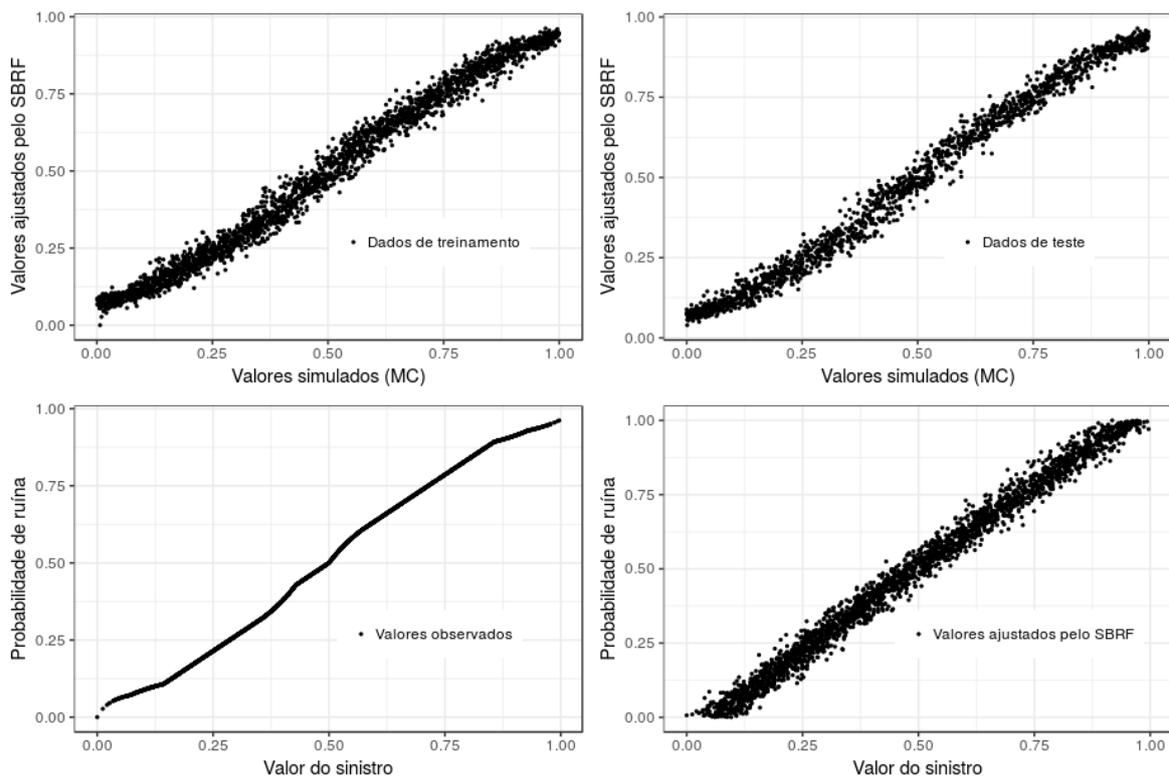


Figura 11: Valores ajustados pelo SBRF (15 conjuntos *fuzzy*) em função dos valores simulados por Monte Carlo para dados ($X \sim \text{Gama}(\alpha, \eta), \alpha \text{ fixo}$) de treinamento e teste e estimativas do limite superior da probabilidade de ruína.

Fonte: Elaboração própria.

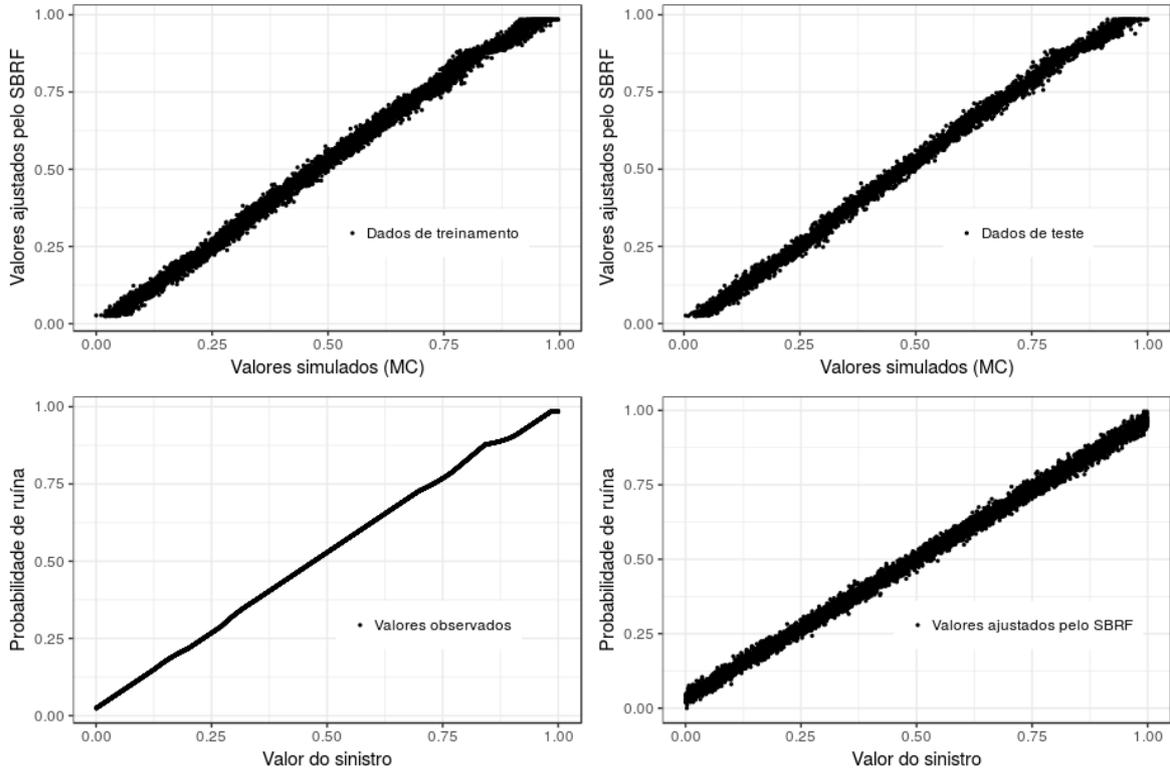


Figura 12: Valores ajustados pelo SBRF (15 conjuntos *fuzzy*) em função dos valores simulados por Monte Carlo para dados ($X \sim \text{Gama}(\alpha, \eta), \eta$ fixo) de treinamento e teste e estimativas do limite superior da probabilidade de ruína.

Fonte: Elaboração própria.

4.3 Modelagem *fuzzy* 2

Na modelagem *fuzzy* 2, foi considerado que $X \sim \text{Exponencial}(\tilde{\beta})$, sendo $\tilde{\beta}$ um número *fuzzy* triangular representado pela terna $(0, 4; 0, 5; 0, 6)$. Os α -níveis de $\tilde{\beta}$ são os intervalos $[\tilde{\beta}]^\alpha = [0, 1\alpha + 0, 4, -0, 1\alpha + 0, 6]$, para todo $\alpha \in [0, 1]$.

A Tabela 7 apresenta os resultados obtidos para o limite superior da probabilidade de ruína, considerando diferentes valores de α -níveis. Nota-se que quanto menor a incerteza referente a β , ou seja, quanto maior o valor de α , menor a incerteza quanto ao valor do limite superior da probabilidade de ruína e ao valor médio da indenização (\bar{x}). Para o caso clássico, ou seja, sem incertezas, em que $\alpha = 1$, tem-se que $\tilde{\beta}_{\text{inf}}^{1,0} = \tilde{\beta}_{\text{sup}}^{1,0} = 0, 5$, $\bar{x}_{\text{inf}} = \bar{x}_{\text{sup}} = 2, 0$ e $\psi(\tilde{\beta}_{\text{inf}}^{1,0}) = \psi(\tilde{\beta}_{\text{sup}}^{1,0}) = 0, 72$. Tomando como exemplo, para $\alpha = 0, 2$, tem-se o intervalo do limite superior da probabilidade de ruína $[0, 61; 0, 84]$ e o intervalo de valor médio da indenização $[1, 73; 2, 38]$. Foi possível notar que quanto maior a incerteza relacionada ao parâmetro da distribuição exponencial, maior é a amplitude dos intervalos do limite superior da probabilidade de ruína.

Tabela 7: Resultados obtidos para o limite superior da probabilidade de ruína, considerando diferentes valores de α -níveis.

α	$\tilde{\beta}_{\text{inf}}^\alpha$	$\tilde{\beta}_{\text{sup}}^\alpha$	$\psi(\tilde{\beta}_{\text{inf}}^\alpha)$	$\psi(\tilde{\beta}_{\text{sup}}^\alpha)$	\bar{x}_{inf}	\bar{x}_{sup}
0,0	0,40	0,60	0,59	0,88	1,67	2,50
0,2	0,42	0,58	0,61	0,84	1,73	2,38
0,4	0,44	0,56	0,64	0,81	1,79	2,27
0,6	0,46	0,54	0,66	0,78	1,85	2,18
0,8	0,48	0,52	0,69	0,75	1,92	2,08
1,0	0,50	0,50	0,72	0,72	2,00	2,00

5 Considerações finais

Este trabalho apresentou durante o seu desenvolvimento, os fundamentos teóricos da teoria da ruína, mais especificamente o modelo clássico de Crámer-Lundberg, incluindo o cálculo do limite superior da probabilidade de ruína, do coeficiente de ajustamento e da desigualdade de Lundberg. Além disso, foi feita uma breve introdução à teoria dos conjuntos *fuzzy*, a fim de apresentar o embasamento teórico necessário para a construção dos sistemas baseados em regras *fuzzy* (SBRFs). Também foram apresentados os principais métodos de inferência utilizados pelos SBRFs e os métodos de defuzzificação. Além disso, foi apresentado o princípio de extensão de Zadeh e a noção de números *fuzzy*, utilizados para estender o cálculo do limite superior da probabilidade de ruína para a teoria dos conjuntos *fuzzy*.

Através da desigualdade de Lundberg, e do cálculo do coeficiente de ajustamento, foram feitas simulações utilizando simulação de Monte Carlo para o limite superior da probabilidade da ruína, para o caso em que o valor das indenizações segue distribuições exponencial e gama.

Na modelagem *fuzzy* 1, os valores das indenizações, obtidos através das simulações de Monte Carlo, foram utilizados como variável de entrada nos SBRFs. Os SBRFs foram criados por meio do pacote *frbs* (*software* R), adotando a função de pertinência gaussiana, o método de inferência de Mamdani e o método de defuzzificação centro de gravidade, com o objetivo de encontrar o limite superior da probabilidade de ruína. Os resultados mostraram que, tanto para o caso em que o valor das indenizações tem distribuição exponencial quanto para o caso em que tem distribuição gama, os sistemas baseados em regras *fuzzy* foram bem ajustados e alcançaram um bom desempenho na estimação do limite superior da probabilidade de ruína. Em todas as simulações, os SBRFs alcançaram um coeficiente de determinação superior a 90%.

Na modelagem *fuzzy* 2, foi considerado que o valor das indenizações segue uma distribuição exponencial com parâmetro incerto, representado por um número *fuzzy* triangular. Observou-se que quanto menor a incerteza referente ao parâmetro, menor a

incerteza quanto ao valor do limite superior da probabilidade de ruína e ao valor médio da indenização.

Destaca-se que em trabalhos futuros, pretende-se analisar o desempenho dos modelos *fuzzy* utilizando outras aproximações para a probabilidade de ruína, como a De Vylder (DE VYLDER, 1978) e Beekman-Bowers (BEEKMAN, 1969).

6 Referências

- ALCALÁ, R; ALCALÁ-FDEZ, J; HERRERA, F. A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection. **IEEE Transactions on Fuzzy Systems**, v. 15, n. 4, p. 616-635, 2007.
- ASMUSSEN S. Ruin Probabilities (2nd Edition). 2010. Disponível em: <http://site.ebrary.com/lib/royallibrary/docDetail.action?docID=10480047>.
- BARROS, L. C.; BASSANEZI, R. C. **Tópicos de lógica fuzzy e biomatemática**. Grupo de Biomatemática, Instituto de Matemática, Estatística e Computação Científica (IMECC), Universidade Estadual de Campinas (UNICAMP), 2010.
- BEEKMAN, J. A. A Ruin Function Approximation. **Transactions of the Society of Actuaries** v. 21, p. 41-48, 1969
- BEEKMAN, J A. A series for infinite time ruin probabilities. **Insurance Mathematics and Economics**, v. 4, n. 2, p. 129-134, 1985.
- BERGMEIR, C; BEN, M. frbs : Fuzzy Rule-Based Systems for Classification. **Journal of Statistical Software**, v. 65, n. 6, p. 1-30, 2015.
- DE VYLDER, F.; GOOVAERTS, M. J. Recursive calculation of finite-time ruin probabilities. **Insurance: Mathematics and Economics**, v. 7, n. 1, p. 1-7, 1988.
- DE VYLDER, F. A Practical Solution to the Problem of Ultimate Ruin Probability. **Scandinavian Actuarial Journal**, p. 114-119, 1978
- DICKSON, D. C. M. **Insurance risk and ruin**. Cambridge University Press, 2016.
- DICKSON, D. C. M; WATERS, H. R. The probability and severity of ruin in finite and infinite time. **ASTIN Bulletin: The Journal of the IAA**, v. 22, n. 2, p. 177-190, 1992.
- EADE, J. P. The ruin problem for mixed Poisson risk processes. **Scandinavian Actuarial Journal**, v. 1983, n. 4, p. 193-210, 1983.
- EMBRECHTS, P.; SCHMIDLI, H. Ruin estimation for a general insurance risk

model. **Advances in applied probability**, v. 26, n. 2, p. 404-422, 1994.

EVERITT, B. S.; HOTHORN, T. **An introduction to applied multivariate analysis with R**. New York: Springer-Verlag, 2011.

FERREIRA, P. P. Modelos de precificação e ruína para seguros de curto prazo. Rio de Janeiro: FUNENSEG, 2005. p. 210.

GERBER, H. U.; GOOVAERTS, M. J.; KAAS, R. On the probability and severity of ruin. **The Journal of the IAA**, v. 17, n. 2, p. 151-163, 1987.

HIPP, C. Estimators and Bootstrap Confidence Intervals for Ruin Probabilities. **ASTIN Bulletin**, v. 9, n. 01, p. 57-70, 1989.

KAAS, R., GOOVAERTS, M., DHAENE, J., DENUIT, M. Modern actuarial risk theory: using R. Springer Science & Business Media, 2008.

LEMOS, S. R. R. Probabilidade de Ruína no Mercado de Seguros: Fundamentos Teóricos e alguns resultados de Simulação. 2008. Dissertação e mestrado em estatística. UFPE, 2008.

MAMDANI, E. H. Application of fuzzy algorithms for control of simple dynamic plant. In: **Proceedings of the institution of electrical engineers**. IET, 1974. p. 1585-1588.

RAMOS, P. A. F. L. Princípios de cálculo de prémios e de medidas de risco em modelos atuariais. Dissertação de mestrado. Universidade do Minho, Portugal, 2014.

R DEVELOPMENT CORE TEAM. R: A Language and Environment for Statistical Computing. 2020. Acesso em: 10 de março de 2020. Disponível em: <http://www.r-project.org>.

RSTUDIO. **RStudio**: integrated development for R. 2020. Disponível em: <http://rstudio.org/>. Acesso em: 10 de março de 2020.

RIZA, L. S. et al. Learning from data using the R package FRBS. **IEEE International Conference on Fuzzy Systems** p. 2149-2155. 2014.

RIZA, L. S. et al. Fuzzy Rule-Based Systems for Classification and Regression Tasks. R package version 3.1-0, 2015. Disponível em: <https://cran.r-project.org/package=frbs>.

SEAH, E. S. Computing the Probability of Eventual Ruin. **Transaction of Society of Actuaries**, v. 42, p. 421-446, 1990. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.124.6446>.

SHAPIRO, A. F. Fuzzy logic in insurance. **Insurance: Mathematics and Economics**, v. 35, n. 2 SPEC. ISS., p. 399-424, 2004.

SUGENO, M; TAKAGI, T. Fuzzy identification of systems and its applications to modeling and control. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 15, n. 1, p. 116-132, 1985.

WICKHAM, H. **ggplot2**: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009. Disponível em: <https://CRAN.R-project.org/package=ggplot2>. Acesso em: 4 de junho 2019

WICKHAM, H.; FRANCOIS, R. **dplyr**: A Grammar of Data Manipulation. R package version 0.4.3, 2015. Disponível em: <https://CRAN.R-project.org/package=dplyr>. Acesso em: 4 de junho 2019

ZADEH, L. A. Fuzzy Sets. **Information and control**, v. 8, p. 338-353, 1965.

ZADEH, L. A. The concept of a linguistic variable and its application to approximate reasoning. **Information Sciences**, v. 8, n. 3, p. 199-249. 1975.