



Avaliação do Risco de Subscrição de Prêmio Utilizando Inferência Bayesiana

Eduardo Fraga

Doutorando e Mestre em Finanças pelo Instituto COPPEAD de Administração/UFRJ e Graduado Magna Cum Laude em Ciências Atuariais pela UFRJ
eduardoflm@susep.gov.br

Resumo

O objetivo deste artigo é apresentar modelos para mensurar o risco de subscrição referente ao prêmio suportado pelas seguradoras utilizando inferência bayesiana e métodos de simulação estocástica via cadeias de Markov – MCMC. Este risco representa a possibilidade do montante dos sinistros a serem pagos pela sociedade seguradora em um período futuro ser maior que o montante de prêmios a ser recebido. Devido aos avanços na área de softwares matemáticos e estatísticos, os modelos propostos são fáceis de serem implementados.

Utilizando conceitos de teoria do risco em uma abordagem Bayesiana, foram propostos modelos que ajustam distribuições aos números de sinistros, valores de sinistros, população segurada e prêmio. Foram testadas as convergências dos algoritmos utilizados e feitas comparações entre os modelos a fim de se chegar ao mais adequado. Como resultado do trabalho, apropriando-se de conceitos de valor em risco – VaR, foram obtidos montantes de capital necessários para suportar o risco de subscrição referente ao prêmio em função de probabilidades críticas aos níveis de 2.5% e 5% para uma determinada seguradora com dados reais para os ramos de automóveis (acidentes pessoais para passageiros, casco e responsabilidade civil facultativa).

Palavras-Chave

Inferência bayesiana; modelos dinâmicos; MCMC; risco de subscrição; distribuição preditiva; valor em risco (VaR).

Sumário

1. Introdução; 2. Modelos de requerimento de capital; 3. Dados utilizados; 4. Modelos propostos; 5. Resultados e análise de convergência; 6. Considerações finais; 7. Referências bibliográficas.

Abstract

Risk assessment of premium subscription through bayesian inference

Eduardo Fraga, PhD and MSc in Finance from the School of Management, COPPEAD/UFRJ (The Federal University of Rio de Janeiro) and Magna Cum Laude BSc in Actuary Sciences from UFRJ
eduardoflm@susep.gov.br

The objective of this article is to present models to measure the underwriting risk regarding the premium supported by the insurance companies using bayesian inference and methods of stochastic simulation through Markov Chains – MCMC. This risk represents the possibility of the amount of claims to be paid by the insurance company in a future period be larger than the amount of premiums to be received. Due to the progresses in the area of mathematical and statistical softwares, the proposed models are easy to be implemented.

Using concepts of risk theory in a bayesian approach, models that adjust distributions to the numbers of claims, values of these claims, insured population and premiums were proposed. The convergences of the used algorithms were tested and made comparisons among the models in order to get the most appropriate one. With concepts of value at risk (VaR) and as a result of the work, capital requirements to support the underwriting risk regarding the premium were obtained according to critical probabilities at the levels of 2.5% and 5% for a certain insurance company with real data for the automobile business class.

Key Words

Bayesian Inference; dynamic models; MCMC; subscription risk; predicted distribution; value at risk (VaR).

Table of Contents

1. Introduction; 2. Capital requirement models; 3. Data used; 4. Proposed models; 5. Results and convergence analysis; 6. Final considerations; 7. References.

1. Introdução

Devido ao comportamento cada vez mais dinâmico dos fatores de risco inerentes às operações de seguradoras, é cada vez mais necessária a utilização de ferramentas de análise, gestão e mitigação destes riscos. Como formadoras de poupança popular, estas instituições mais do que quaisquer outras devem manter sólidos sistemas de controle de riscos. Assim, o sucesso desta implementação é de vital importância para a manutenção das companhias e do próprio sistema securitário, que se baseia fortemente em uma relação de confiança.

Dentre os vários riscos a que estão expostas estas instituições podemos elencar os de mercado, crédito, subscrição (atuarial), legal e operacionais. Neste contexto, serão apresentados modelos visando a avaliação do risco de subscrição referente ao prêmio suportado por uma seguradora para os ramos de automóveis (ramos de acidentes pessoais de passageiros, casco e responsabilidade civil facultativa – danos materiais e pessoais). Tais modelos possibilitarão a obtenção de capitais que cubram este risco. Ter um montante de capital próprio apropriado é importante para as empresas de seguro, para que a imagem de confiança que os segurados possuem com relação às seguradoras não seja abalada. Estudar os riscos associados ao processo de negócios das seguradoras ajuda a compreender como as margens de solvência podem ser afetadas pelas variabilidades e incertezas presentes no ambiente do negócio.

O risco de subscrição referente ao prêmio suportado pelas seguradoras representa a possibilidade do montante dos sinistros a serem pagos pela sociedade ser maior que o prêmio a ser recebido. De forma a simplificar, usualmente assume-se uma aproximação normal para o montante de sinistros com base no TCL (Teorema Central do Limite – Gerber, 1979), o que pode ser questionável se o número de sinistros (n) não for suficientemente grande. Além disso, a teoria do risco clássica baseia-se em estimativas pontuais ou intervalos de confiança, enquanto que, pelo método Bayesiano, pode-se analisar a distribuição preditiva do resultado de subscrição e, a partir desta distribuição, pode-se avaliar o capital baseado no risco, utilizando conceitos de valor em risco (VaR) com muito mais informações estatísticas do que apenas a média e a variância.

Ryan et al. (2001) analisou a importância relativa da cada fator de risco baseado em um artigo da agência de rating A.M. Best em 1999 onde foram analisados 683 casos de insolvência nos Estados Unidos entre os anos de 1969 e 1998. Neste artigo, foram classificadas as razões para insolvência por riscos de subscrição, riscos de ativos e outros riscos. Dos casos analisados 42% das insolvências foram devidas a riscos de subscrição com a seguinte separação:

- prêmios e provisões insuficientes: 22%
- crescimento acelerado: 14%
- perdas catastróficas: 6%

Riscos de ativos são responsáveis por 9% dos casos (6% – superavaliação dos ativos e 3% – falha no resseguro). Outros riscos que contribuíram para a falência das companhias foram suas subsidiárias, mudança significativa no "core business", fraude e outros riscos não identificáveis.

Este artigo foi desenvolvido com bases teóricas de Processos Estocásticos aplicados à Teoria do Risco, Modelos Hierárquicos, captando as variações inerentes à forma de agregação dos dados, Inferência Bayesiana e simulação MCMC (Scollinik, 2000). Vários modelos hierárquicos Bayesianos foram propostos para serem usados como técnicas de avaliação do capital necessário para suportar o risco de subscrição referente ao prêmio. Estes modelos foram implementados por meio do software WinBUGS (Bayesian Inference Using Gibbs Sampling – Spiegelhalter et al., 2003). O método de Simulação Estocástica via Cadeias de Markov, em particular o Amostrador de Gibbs (Gelfand & Smith, 1990; Gelfand et al., 1990), foi o método iterativo utilizado para obtenção de uma amostra das posterioris e da distribuição preditiva de interesse, já que a mesma não poderia ser obtida analiticamente.

O artigo está organizado da seguinte forma: na seção 2 serão apresentados conceitos de modelos de requerimento de capital, na seção 3 é apresentada uma análise dos dados utilizados. Na seção 4 são propostos modelos dinâmicos para mensuração do capital baseado no risco de subscrição, na seção 5 é feita a análise de convergência e apresentado os resultados. Finalmente, na seção 6, são feitas as considerações finais e sugestões de extensão do trabalho.

2. Modelos de requerimento de capital

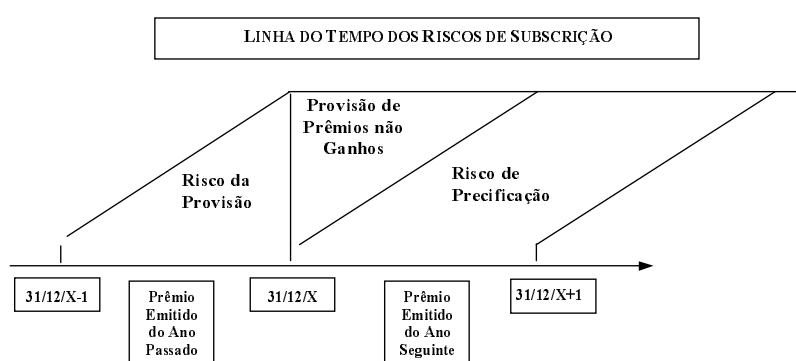
Nesta seção serão abordados conceitos sobre sistemas de classificação de risco. Um ponto comum que serve de balizador para as regras de capital é a necessidade de descrever os tipos de riscos a que as seguradoras estão expostas (Feldblum, 1996). Como exemplo é citado o Risk Based Capital, sistema de solvência americano vigente desde o início dos anos 90 desenvolvido pela NAIC (National Association of Insurance Commissioners –, para maiores detalhes consultar www.naic.org). Os vários requisitos relacionados à fórmula do RBC são inicialmente combinados em seis categorias, nomeadas de R0 a R5, como segue:

- R0: Investimento em seguradoras afiliadas e "off-balance".
- R1: Caixa, Títulos, Fator de Ajustamento do valor dos títulos, empréstimos.
- R2: Ações ordinárias, ações preferenciais, imóveis, outros ativos investidos.
- R3: Risco de crédito: recuperação de resseguro e outros recebíveis.

Riscos de Subscrição:

- R4: Risco da Provisão de Sinistros.
- R5: Risco do Prêmio Emitido.

Como pode ser notado, o risco de subscrição é dividido em risco de provisão de sinistros e risco do prêmio (precificação). Estes requisitos são calculados aplicando-se coeficientes ao do volume de provisão e de prêmios referentes a exercícios passados. De forma a facilitar o entendimento destes riscos nos negócios da entidade, é apresentado o esquema a seguir:



Data da Avaliação: 31/12/X

Risco da Provisão: Cobertura emitida e ganha, porém ainda não totalmente liquidada (desenvolvimento adverso dos sinistros).

Provisão de Prêmio não Ganho: Cobertura emitida, porém não ganha.

Risco de Precificação: Cobertura a ser emitida em períodos futuros.

Pode-se observar na figura acima que a situação representada de forma estática ocorre na realidade de forma recorrente. Para uma seguradora posicionada em 31/12/X, há o risco de um desenvolvimento adverso dos sinistros (risco de provisão), risco da provisão de prêmios não ganhos ser insuficiente, medida que foi considerada pela NAIC não significativa e o risco de precificação. Outro conceito importante que é utilizado na avaliação do capital necessário para suportar o risco de subscrição referente ao prêmio é o processo de ruína (Bowers et al. 1997):

Sejam:

R_t – ruína da seguradora no momento t
 μ – capital próprio baseado no risco
 P_t – prêmio retido no momento t
 S_t – sinistro retido no momento t

Temos que a ruína da seguradora no momento t é definida por:

$$R_t = \mu + P_t - S_t$$

O que se deseja é que $R_t = \mu + P_t - S_t > 0, \forall t$

Ou seja,

$$P_r(\mu + P_t - S_t > 0) = 1 - \alpha \Leftrightarrow P_r(P_t - S_t < -\mu) = \alpha$$

onde: α é a probabilidade crítica (ex.: 1%, 2.5% ou 5%).

Pode-se notar que o processo de ruína como definido acima não cobre todos os riscos a que estão expostas as seguradoras, entretanto o foco deste artigo é o risco de subscrição referente aos prêmios. Assim, definimos o resultado de subscrição como:

Re_{s_t} – resultado de subscrição de uma seguradora no momento t

$$Re_{s_t} = P_t - S_t$$

Neste artigo é proposta metodologia de avaliação de capital das seguradoras baseado no risco de subscrição referente ao prêmio a que elas estão expostas. Para tanto, faremos uso de métodos da inferência bayesiana para estimar a distribuição do resultado de subscrição. O grande agente catalisador das discussões sobre modelos de regulação baseado em risco foi o setor bancário. O valor que tem servido de parâmetro para a maioria dos bancos centrais tomarem alguma ação (intervenção, direção fiscal, liquidação, etc.) contra um banco é conhecido por “Value at Risk” (VaR). O VaR é uma tentativa de fornecer um número simples que resume o total de risco de um portfólio financeiro.

Segundo Hull (2002), quando se pensa na medida de VaR, o seguinte enunciado é posto: “Estamos c por cento certos que não perderemos mais do que V reais nos próximos N dias”. Seja VaR uma função de dois parâmetros: N , o horizonte de tempo; e α , o nível de confiança, ou seja, o VaR representa o nível de perda através dos N dias que estamos $100 \times \alpha\%$ certos de que não excederemos. Em geral, o VaR corresponde ao $(100 - 100 \times \alpha)$ -ésimo percentil da distribuição de mudança do valor da carteira através dos próximos N dias.

Pode-se associar o VaR a uma medida de probabilidade crítica, postulando que o conceito de risco em aplicações financeiras está associado à variabilidade da distribuição de perdas ou aos ganhos monetários. Como caso especial, o Value at Risk é definido por:

$$P[y \leq \text{VaR}] = 1 - \alpha = \int_{-\infty}^{\text{VaR}} p(y) dy$$

Neste artigo estamos interessados em determinar um dado VaR, que a partir de agora chamaremos de requerimento de capital (RC) dada uma probabilidade crítica e um horizonte de tempo. Assim como os bancos, as seguradoras possuem riscos que devem ser gerenciados. Contudo, a sua operação possui características diferenciadas daqueles. Portanto, é necessário que seja desenvolvido modelo adequado a sua operação que forneça a mesma medida que o VaR utilizado pelos bancos. Sendo assim, serão apropriados os conceitos de probabilidade crítica para avaliar o capital baseado no risco de subscrição de prêmio (precificação). Serão utilizados modelos bayesianos para avaliação dos fluxos de caixa das perdas futuras da seguradora.

3. Dados utilizados

Os dados utilizados nesta aplicação foram os sinistros e exposições de uma seguradora de automóveis do mercado brasileiro no período de análise de Janeiro/2003 a Dezembro/2003 (1 ano). A chave primária do arquivo era região, código tarifário, tipo de contrato e trimestre. As variáveis são valor do sinistro, frequência, população exposta ao risco e prêmio ganho. Cada uma destas variáveis está discriminada por tipo: casco, responsabilidade civil facultativa – danos materiais (RCF-DM), responsabilidade civil facultativa – danos pessoais (RCF-DP) e acidentes pessoais por passageiro (APP). A tabela 1 apresenta os fatores de risco para avaliação do risco de subscrição de prêmio:

Tabela 1 – Fatores de risco ou “quebra” do arquivo de sinistros, exposição (população) e prêmios ganhos

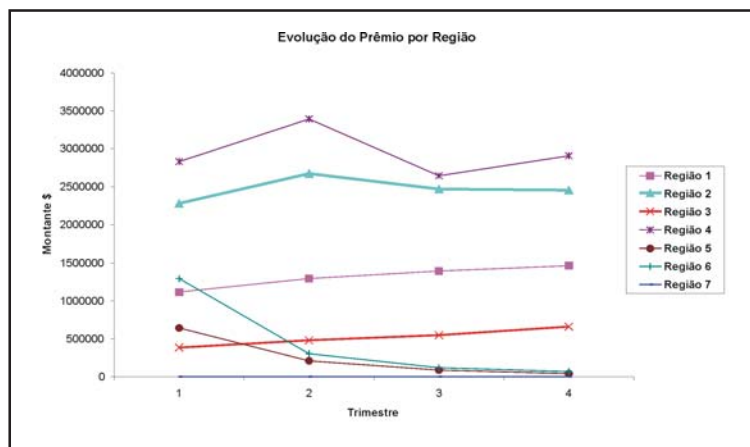
Região	1. RS, SC, PR 2. SP 3. MG 4. RJ 5. ES, BA 6. SE, PE, PB, RN, AL, CE, PI, MA, PA, AM, AP, RO, RR, AC, MT, MS, DF, GO, TO 7. Outras*	
Código tarifário	1. Passeio nacional 2. Passeio importado 3. Pick-up (nacional e importado) 4. Veículo de Carga (nac. e imp.)	5. Motocicleta (nac. e imp.) 6. Ônibus (nac. e imp.) 7. Utilitários (nac. e imp.) 8. Outros
Tipo de contrato	1. Pessoa Física 2. Pessoa Jurídica	
Trimestre	1º, 2º, 3º ou 4º	

* Código de identificação de regiões não classificadas de acordo com o *lay-out* definido pela Circular nº 197/2002 (www.susep.gov.br).

A seguir são apresentados gráficos da evolução temporal do prêmio ganho, população exposta, montante de sinistros, número de sinistros e sinistro médio (montante de sinistros/número de sinistros) e frequência de sinistros para algumas das “quebras” expostas na tabela acima.

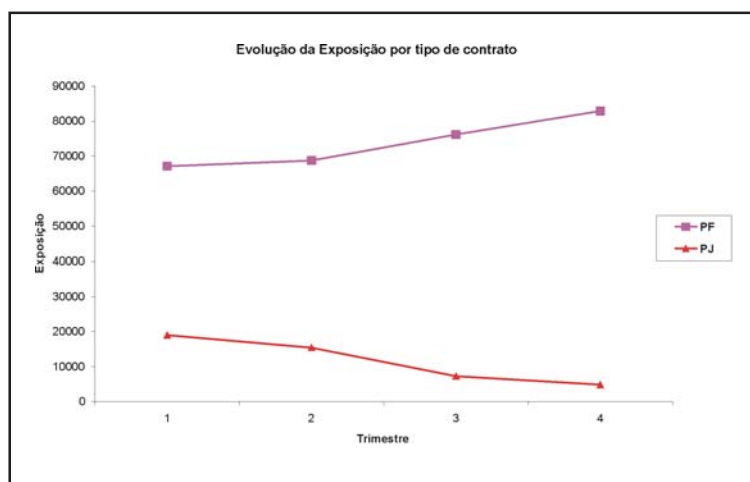


Gráfico 1 – Evolução do prêmio ganho por região para os ramos de automóveis por trimestre no ano de 2003



Neste gráfico, pode-se notar que as diferentes taxas de crescimento variam de acordo com a região. Não há um crescimento ou decréscimo uniforme do prêmio ganho pelas regiões.

Gráfico 2 – Evolução da exposição (população) por tipo de contrato (pessoa física ou jurídica) para os ramos de automóveis por trimestre no ano de 2003





Neste gráfico, fica clara a diferença do comportamento do número de expostos por trimestre em 2003. Alguns dos modelos a serem apresentados conterão parâmetros capazes de captar tal diferença.

Gráfico 3 – Evolução do montante de sinistros por código tarifário para os ramos de automóveis por trimestre no ano de 2003

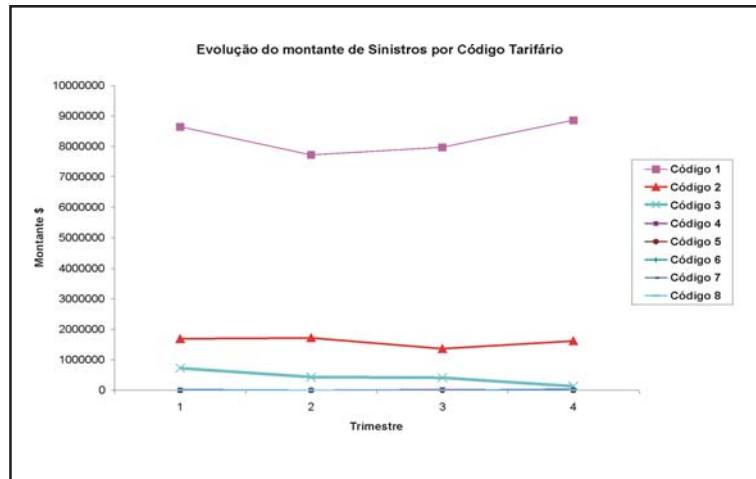


Gráfico 4 – Evolução do número de sinistros por região para os ramos de automóveis por trimestre no ano de 2003

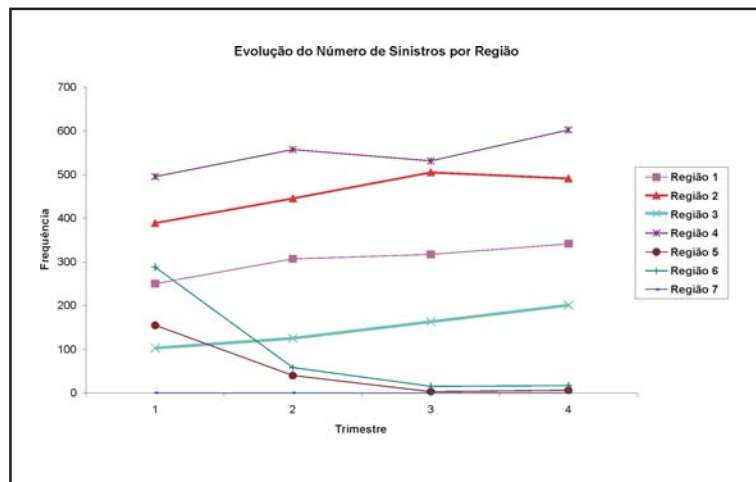
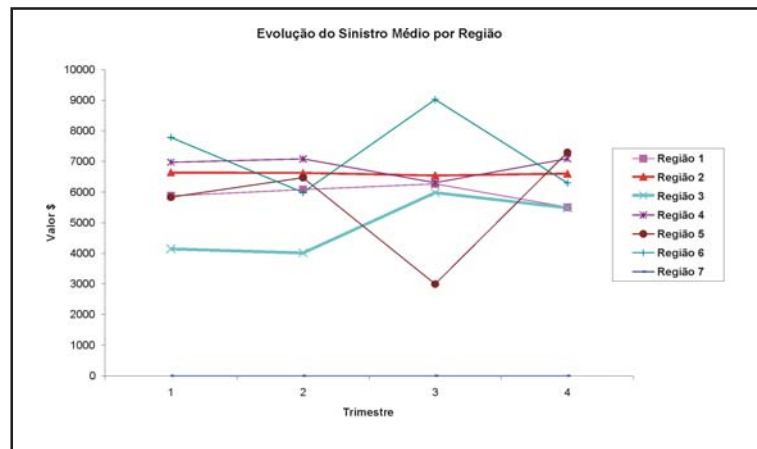




Gráfico 5 – Evolução do sinistro médio por região para os ramos de automóveis por trimestre no ano de 2003



Observando-se os gráficos acima, facilmente nota-se a influência dos fatores de risco (código tarifário, região e tipo de contrato) no montante de sinistros, no número de sinistros, no sinistro médio e na evolução deles durante o ano. Os códigos tarifários 4 a 8 quase não apresentaram sinistros. O mesmo ocorre quando analisamos a frequência de sinistros: diferenças significativas na evolução temporal de acordo com cada região (Gráfico 4). Por isso, como base nos gráficos, parece ser razoável que seja levado em consideração os diferentes códigos tarifários, regiões e tipos de contrato ao modelarmos as distribuições de número, valores de sinistros, prêmio e população para se chegar no requerimento de capital para o risco de subscrição de prêmio.

4. Modelos propostos

Para facilitar a compreensão dos modelos que são introduzidos neste artigo abordaremos brevemente a teoria do risco coletivo (Bowers et al. 1997), que é uma das metodologias mais utilizadas para a tarifação de seguros. Nesta teoria a soma dos valores dos sinistros (sinistro agregado) é denotada por:

$$S_{col} = \sum_{i=1}^N X_i$$

onde S_{col} é a variável aleatória montante de sinistros; X_i é a variável aleatória valor de um sinistro e N é a variável aleatória número de sinistros. Assumindo a distribuição de N como Poisson (λ), obtém-se a distribuição de S_{col} através de uma aproximação normal, quando a verdadeira distribuição de S_{col} é a Poisson Composta. A média e variância de S_{col} pode ser facilmente obtida através das equações:

$$E[S_{col}] = E[N] * E[X] = \lambda * E[X]$$

$$V[S_{col}] = E[N] * V[X] + E[X] * V[N] = \lambda * V[X] + E[X]^2 * \lambda$$

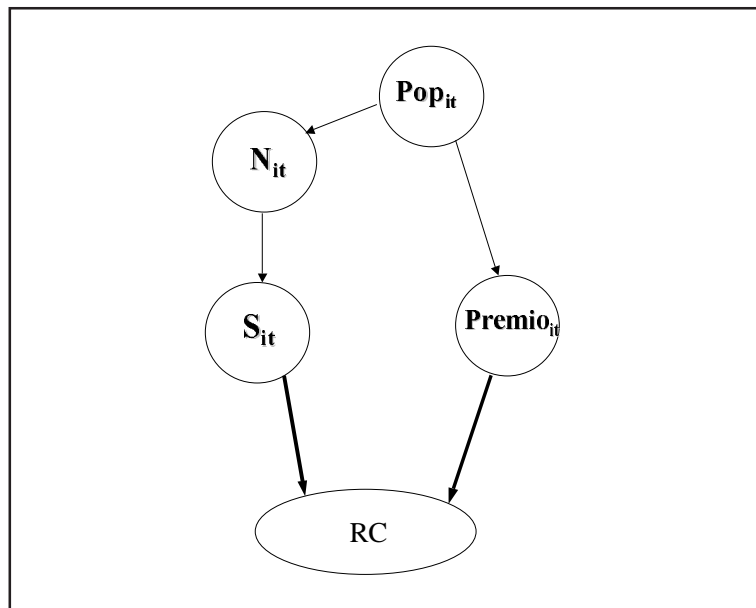
Em uma abordagem clássica ou frequentista, o valor de λ é estimado pelo número de sinistros observados no período de análise (n) e $E[X]$ e $V[X]$ pela média e variância amostrais respectivamente. Uma das fórmulas de cálculo de estimação do montante de sinistros futuros dado um nível de confiança de $(1-\alpha)$ consiste em aplicar um carregamento de segurança sobre o valor esperado de S_{col} .

$$M_{T;\alpha} = E[S_{col}] + Z_{1-\alpha} * \sqrt{V[S_{col}]}$$

onde α é tal que $P(S_{col} < P_T) = 1-\alpha$ e $Z_{1-\alpha}$ é o percentil $1-\alpha$ da distribuição Normal padrão. Entretanto, a abordagem utilizada neste trabalho será simular as distribuições das variáveis de interesse e dos parâmetros em um contexto bayesiano.

De forma simplificada podemos apresentar o modelo de avaliação do requerimento de capital baseado no risco de subscrição de prêmio (precificação) com diagramas de influência. Abaixo, encontra-se uma versão simplificada do diagrama. Os fatores de risco considerados neste trabalho foram a região, o código tarifário e o tipo de contrato (pessoa física ou jurídica).

Figura 1 – Diagrama de influência do requerimento de capital baseado no risco de subscrição de prêmio (precificação)



Onde:

Pop_t – população exposta nos fatores de risco i no período t .

N_{it} – número de sinistros para os fatores de risco i no período t .

S_{it} – montante de sinistros para os fatores de risco i no período t .

$Premio_{it}$ – montante de prêmios para os fatores de risco i no período t .

RC – requerimento de capital.

- Modelos dinâmicos propostos

Foram propostos 5 modelos distintos para os dados com diferentes distribuições assumidas para os montantes de sinistros, para os números de sinistros, população segurada e prêmio.

- I. Modelo hierárquico dinâmico que assume Gama como distribuição dos montantes de sinistros (S) e prêmios (P), Poisson como distribuição do número de sinistros (N), modela a distribuição para população segurada (π) com um fator macroeconômico (taxa SELIC) e considera todos os fatores de risco no tempo t.
- II. Modelo hierárquico que assume Gama como distribuição dos montantes de sinistros (S) e prêmios (P), Poisson como distribuição do número de sinistros (N), modela a distribuição para população segurada (π) com um fator macroeconômico (taxa SELIC) e não considera a evolução dos fatores de risco no tempo t.
- III. Modelo hierárquico que assume Gama como distribuição dos montantes de sinistros (S) e prêmios (P), Poisson como distribuição do número de sinistros (N), modela a distribuição para população segurada (π) com um fator macroeconômico (taxa SELIC), não considera a evolução dos fatores de risco no tempo t e não considera o fator região.
- IV. Modelo hierárquico que assume Gama como distribuição dos montantes de sinistros (S) e prêmios (P), Poisson como distribuição do número de sinistros (N), modela a distribuição para população segurada (π) com um fator macroeconômico (taxa SELIC), não considera a evolução dos fatores de risco no tempo t e não considera os fatores região e tipo de contrato.
- V. Modelo hierárquico que assume Gama como distribuição dos montantes de sinistros (S) e prêmios (P), Poisson como distribuição do número de sinistros (N), modela a distribuição para população segurada (π) com um fator macroeconômico (taxa SELIC), não considera a evolução dos fatores de risco no tempo t e não considera os fatores região, sexo e código tarifário.

A taxa SELIC foi escolhida como preditor da população por apresentar um alto índice de correlação com esta variável. Além disso, é economicamente aceitável que, em um ambiente de taxas de juros mais elevadas, haja uma contração no nível de consumo de contratos de seguros e vice-versa em um ambiente de taxas em declínio. Não é foco do artigo discutir a utilização deste preditor, mas pode ser testada a sua contribuição ao modelo, assim como podem ser testadas outras variáveis, como nível de renda, desemprego, agregados monetários, índices de consumo e etc.

As equações a seguir descrevem o Modelo I. Os outros modelos podem ser escritos de maneira análoga:



- Modelando a população segurada π , onde π_{sft} é a população segurada para a cobertura s (casco, RCF-DM, RCF-DP e APP) dos fatores de risco f (região, código tarifário e tipo de contrato), no trimestre t :

Supondo que π_{sft} tenha um crescimento exponencial e modelando o seu logaritmo, temos que este valor depende de uma parcela relativa à região, código tarifário e tipo de contrato, outra também relativa a estes fatores de risco porém crescente no tempo e uma terceira considerando os fatores de risco com a variável macroeconômica “taxa SELIC”, que apresentou alta correlação negativa com a variável exposição (-0.74). $\beta_0, \beta_1, \beta_2, \tau_0, \tau_1, \tau_2$ e τ_δ são chamados hiperparâmetros e foram modelados com distribuições a priori não informativas.

$$\begin{aligned} \log(\pi_{sft}) | \beta_{0sf}, \beta_{1sf}, \beta_{2sf}, \tau_\delta &\sim N(\beta_{0sf} + t * \beta_{1sf} + Selic_t * \beta_{2sf}, 1/\tau_\delta) & (1) \\ \beta_{0sf} | \beta_0, \tau_0 &\sim N(\beta_0, 1/\tau_0) & \beta_{1sf} | \beta_1, \tau_1 &\sim N(\beta_1, 1/\tau_1) \\ \beta_{2sf} | \beta_2, \tau_2 &\sim N(\beta_2, 1/\tau_2) \\ \beta_0 &\sim N(a_{\beta_0}, 1/b_{\beta_0}) & \beta_1 &\sim N(a_{\beta_1}, 1/b_{\beta_1}) & \beta_2 &\sim N(a_{\beta_2}, 1/b_{\beta_2}) \\ \tau_0 &\sim Gamma(a_0, b_0) & \tau_1 &\sim Gamma(a_1, b_1) & \tau_2 &\sim Gamma(a_2, b_2) \\ \tau_\delta &\sim Gamma(a_\delta, b_\delta) \end{aligned}$$

- Modelando o número de sinistros N , onde N_{sft} é o número de sinistros da cobertura s , nos fatores de risco f , no trimestre t :

Supondo que o número de sinistros tem uma distribuição de Poisson, o parâmetro desta distribuição (sua média e variância) deve depender do tamanho da população segurada (π_{sft}) e de uma taxa de ocorrência de sinistro (λ_{fs}) para a cobertura (s) e fatores de risco (f). Os hiperparâmetros a e b foram modelados com distribuições a priori não informativas.

$$\begin{aligned} N_{fst} | \pi_{fst}, \lambda_{fs} &\sim Poisson(\pi_{fst} * \lambda_{fs}) \\ \lambda_{fs} | a, b &\sim Beta(a, b) \end{aligned}$$

- Modelando o montante de sinistro S , onde S_{sft} é o montante de sinistros da cobertura (s), fatores de risco (f), no trimestre t :

Supondo que o montante de sinistros tem uma distribuição Gamma, os parâmetros desta distribuição devem depender do número de sinistros (N_{ft}) e do valor esperado de cada sinistro (θ_{fs}^{-1}) no trimestre (t), fator de risco (f) e tipo de cobertura (s). Os hiperparâmetros a_w e b_w foram modelados com distribuições a priori não informativas.

$$\begin{aligned} S_{sft} | N_{sft}, \theta_{sf} &\sim Gamma(N_{sft}, \theta_{sf}) \\ \theta_{sf} &\sim Gamma(a_w, b_w) \end{aligned}$$

- Modelando o montante de prêmios P , onde P_{sft} é o montante de prêmios da cobertura (s), fatores de risco (f), no trimestre t :

Supondo que o montante de prêmios tem uma distribuição Gamma, os parâmetros desta distribuição devem depender da exposição (população) (π_{sft}) e do valor esperado de cada prêmio (v_{fs}^{-1}) no trimestre (t), fator de risco (f) e tipo de cobertura (s). Os hiperparâmetros a_e e b_e foram modelados com distribuições a priori não informativas.

$$P_{sft} | \pi_{fst}, v_{sf} \sim \text{Gamma}(\pi_{fst}, v_{sf})$$

$$v_{sf} \sim \text{Gamma}(a_e, b_e)$$

Nossa intenção é estudar a distribuição preditiva do resultado de subscrição da seguradora para o ramo de automóveis:

$$Res = \sum P_{sft} - \sum S_{sft}$$

Visto que a distribuição de interesse (distribuição do resultado de subscrição da seguradora) não pode ser obtida de forma analítica, é necessário o uso de métodos numéricos, como, por exemplo, a simulação estocástica via Cadeias de Markov utilizada neste trabalho por meio do algoritmo Amostrador de Gibbs.

5. Resultados e análise de convergência

A metodologia proposta utiliza Modelos Hierárquicos com diferentes distribuições assumidas para os montantes de sinistros, para os números de sinistros, população segurada e prêmio. Estes modelos foram implementados no BUGS (Bayesian Inference Using Gibbs Sampling) – Software especializado em Inferência Bayesiana que utiliza os métodos de simulação MCMC. Por meio deste programa foi obtida a distribuição preditiva do resultado de subscrição da seguradora. Foram utilizados como probabilidades críticas os níveis de 2,5% e 5%.

A inovação deste artigo é calcular o capital baseado no risco de subscrição referente ao prêmio através da análise da distribuição preditiva do resultado de subscrição da seguradora, ou seja, obter resultados baseados em mais informações estatísticas da seguradora do que estimadores pontuais e intervalos de confiança dos parâmetros de interesse. O algoritmo Amostrador de Gibbs (utilizado no software BUGS) é descrito da seguinte maneira: seja a distribuição a posteriori de interesse: $p_i(\theta_i | D)$ onde $\theta = (\theta_1 \dots \theta_p)$ são os parâmetros desconhecidos do modelo e suponha que sejam conhecidas as distribuições condicionais completas $p_i(\theta_i) = p(\theta_i | \theta_{-i}, D)$, $i = 1, \dots, p$, onde θ_{-i} representa o vetor das componentes de θ , exceto a i -ésima.

O algoritmo consiste em fornecer valores iniciais aos parâmetros $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_p^{(0)})$ e iterativamente obter uma atualização destes valores, por meio da geração das amostras das distribuições condicionais completas.

$$\theta_1^{(i)} \sim p(\theta_1 | \theta_2^{(i-1)}, \dots, \theta_p^{(i-1)}, D)$$

$$\theta_2^{(i)} \sim p(\theta_2 | \theta_1^{(i)}, \theta_3^{(i-1)}, \dots, \theta_p^{(i-1)}, D)$$

$$\vdots$$

$$\theta_p^{(i)} \sim p(\theta_p | \theta_1^{(i)}, \dots, \theta_{p-1}^{(i)}, D)$$

Quando o algoritmo converge, obtemos uma amostra da distribuição de interesse. Existem diferentes métodos para estudar a convergência deste algoritmo. Pode-se dizer que uma simulação por MCMC convergiu quando os valores gerados (saída) deste ponto em diante podem ser considerados provenientes da distribuição estacionária da cadeia de Markov.

– Verificação da Convergência:

A partir da habilidade do software WinBUGS de simular múltiplas cadeias (independentes) em paralelo e fornecer a medida de diagnóstico de Gelman-Rubin (Gelman & Rubin, 1992) baseada nestas cadeias, simulamos, para cada algoritmo, três cadeias independentes com valores iniciais distintos com o objetivo de evitar que as amostras geradas concentrem-se em regiões em torno de uma moda local, no caso de multimodalidade da posteriori (Gamerman, 1996), e verificamos a convergência destes algoritmos.

Segundo Brooks & Gelman (1998), operacionalmente, a convergência efetiva da simulação da cadeia de Markov é alcançada quando inferências das quantidades de interesse não dependem do ponto inicial da simulação. Podemos calcular estatísticas sobre os parâmetros baseados nas saídas das três cadeias juntas ou nas três separadamente. No primeiro caso teremos uma única estatística e no segundo três estatísticas diferentes. Sejam:

R_p – tamanho do intervalo de confiança (IC com 80% de confiança) baseado nas três cadeias juntas *em verde*.

R_w – média dos tamanhos dos intervalos de confiança (IC com 80% de confiança) baseado nas três cadeias individualmente *em azul*.

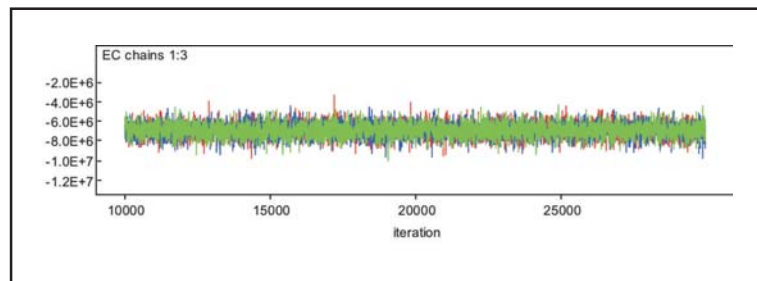
$$R = \frac{R_p}{R_w} \text{ em vermelho}$$

A medida que a simulação por MCMC converge, os valores de R_p e R_w devem convergir para um valor estável e a razão R deve convergir para 1 (um). Se isso não ocorrer e R continuar sendo bem maior do que 1, deve-se continuar a simulação por mais alguns passos ou assumir que há considerações no modelo que levam à lenta convergência.

Para facilitar a verificação da convergência, o software WinBUGS disponibiliza um gráfico chamado GRdiag que contém os três valores com as cores descritas acima. Note que os valores de R_p e R_w foram normalizados para terem o valor máximo de 1 (um). Através destes gráficos podemos também determinar aproximadamente a partir de qual iteração a simulação parece ter convergido e então descartar (como 'aquecimento') as primeiras iterações utilizando apenas as iterações após a convergência para estudar os resultados do modelo.

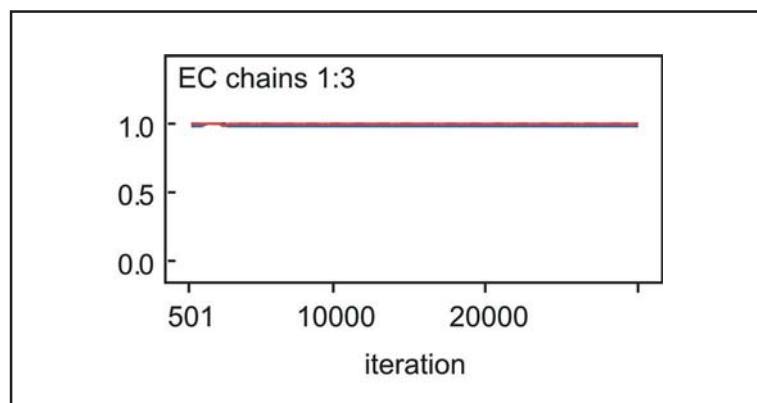
Uma outra disponibilidade do software WinBUGS de verificação da convergência é a análise dos chamados Trace Plots para as múltiplas cadeias. Os Trace Plots são gráficos que acompanham os valores estimados dos parâmetros no decorrer das iterações. Para uma simulação que tenha convergido, as cadeias devem se misturar neste gráfico (e por terem cores distintas isto é facilmente identificado graficamente). Ou seja, não deve haver nenhum comportamento diferente para um mesmo parâmetro entre cadeias distintas após a convergência. Apesar de não ser um método quantitativo de verificação de convergência, ele pode identificar aquelas simulações que realmente não convergiram.

Gráfico 6 – Trace Plot do requerimento de capital (RC) referente ao Modelo V



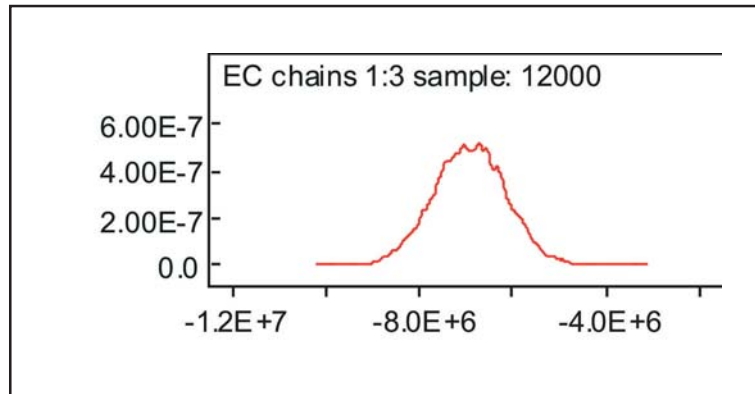
Como resultado deste teste de convergência, os 5 modelos propostos convergiram com um. Abaixo está o gráfico do diagnóstico de Gelman-Rubin para a variável “Requerimento de Capital” do modelo V.

Gráfico 7 – Diagnóstico de Gelman-Rubin referente ao Modelo V



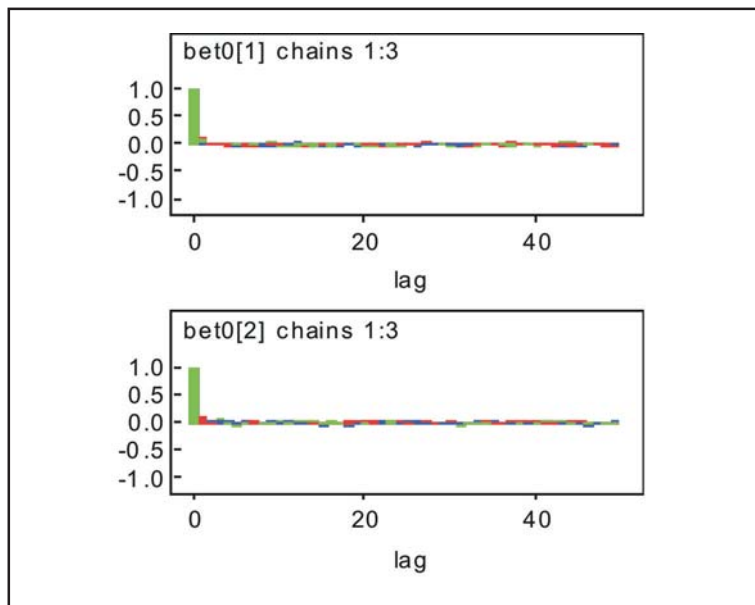
O gráfico demonstra convergência para o algoritmo. Apresentamos as densidades estimadas pelo WinBUGS para o Requerimento de Capital da seguradora para um ano (incluindo todos os fatores de risco). São gráficos de uma estimativa da densidade Kernel aplainada (smoothed kernel density estimate) da variável, se esta for contínua, ou um histograma se for discreta.

Gráfico 8 – Densidade Kernel para a variável “Requerimento de Capital”



Os gráficos de autocorrelação (a função de autocorrelação da variável com lags de 1 a 50) do parâmetro “ β_0 ” se mostraram como os gráficos de autocorrelação de todos os outros parâmetros: valores bem próximos de zero indicando que as estimativas cada vez menos dependem da iteração anterior e que a amostra poderia ser considerada independente.

Gráfico 9 – Gráficos de autocorrelação do parâmetro β_0 para o modelo V



- Comparação de Modelos:

Para comparar os modelos, foi utilizado um critério exposto no artigo de Gelfand and Ghosh (1998). Este critério calcula uma medida de comparação (D) que é a soma de dois termos:

$$D(m) = G(m) + P(m)$$

$$G(m) = \sum_{i=1}^n [\mu_i(m) - S_{i,obs}]^2$$

$$\text{onde } \mu_i(m) = E[S_{i,rep} | S_{i,obs}]$$

$$P(m) = \sum_{i=1}^n \sigma_i^2(m)$$

$$\text{onde } \sigma_i^2(m) = V[S_{i,rep} | S_{i,obs}, m]$$

onde:

- G(m) é uma soma dos quadrados dos resíduos (ou seja, se o modelo estiver muito distante das verdadeiras distribuições, este termo deve ser grande).
- P(m) é termo de penalidade, ou seja, se colocarmos parâmetros a mais sem necessidade, este termo deverá ser grande.

Portanto, quanto menor o valor de D(m), melhor é o ajuste do modelo. Os valores obtidos para os 5 modelos implementados foram:

Tabela 2 – Resultados do critério de comparação de modelos de Gelfand & Ghosh (1998)

	G(M)	P(M)	D(M)
Modelo I	2.47E+07	6.19E+11	6.19E+11
Modelo II	2.86E+07	6.15E+11	6.15E+11
Modelo III	1.77E+07	6.05E+11	6.05E+11
Modelo IV	6.21E+07	5.88E+11	5.88E+11
Modelo V	3.57E+07	5.85E+11	5.85E+11

Pode-se notar que o que determinou este resultado foi o termo P(M). O termo que leva em consideração somente a soma dos resíduos ao quadrado indica que o Modelo III seria o mais adequado. Porém, o critério exposto no artigo de Gelfand & Ghosh (1998) indica o modelo V como o mais adequado dos 5 modelos propostos. Cabe ressaltar que apesar do modelo V se apresentar como mais adequado, ele não permite que seja feita uma análise detalhada de quais fatores de risco estão sensibilizando mais o requerimento de capital, uma vez que, para este modelo, não são consideradas as influências dos fatores código tarifário, região, tipo de contrato e nem trimestre. Abaixo, constam os requerimentos de capital (RC) calculados para cada um dos 5 modelos em função das probabilidades críticas de 2,5% e 5%. Pode ser notado que os valores não diferem muito um do outro entre os 5 modelos.



Tabela 3 – Resultados do requerimento de capital (valor absoluto) calculado para cada um dos 5 modelos ajustados

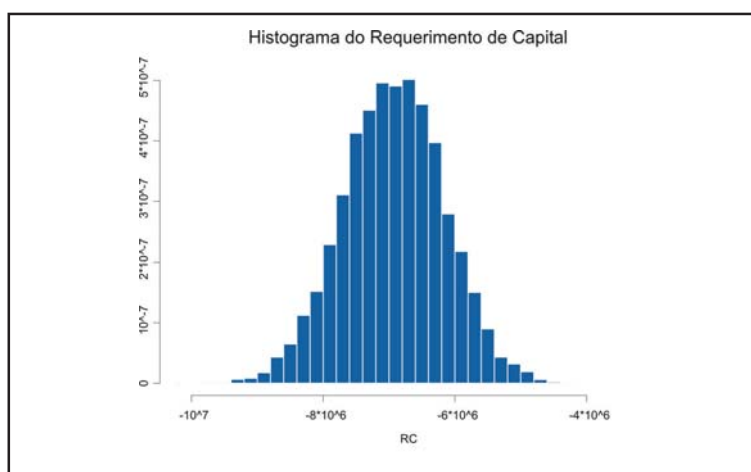
RC	Percentil	
	2.50%	5.00%
Modelo I	8.460	8.210
Modelo II	8.450	8.210
Modelo III	8.440	8.210
Modelo IV	8.440	8.200
Modelo V	8.430	8.200

*Valores em milhares.

**Estimativas obtidas com 40.000 simulações (descarte das 20.000 primeiras) e lag de tamanho 10.

Definido o modelo mais adequado, abaixo apresentamos o histograma do requerimento de capital para suportar o risco de subscrição referente ao prêmio para o horizonte de 1 ano. Pode-se observar uma certa simetria da distribuição. Chamo a atenção para os valores negativos da distribuição, pois o requerimento de capital para o risco de subscrição de prêmio representa o excesso de sinistros sobre os prêmios.

Gráfico 10 – Histograma do requerimento de capital para suportar o risco de subscrição referente ao prêmio



6. Considerações Finais

Neste artigo foram apresentados modelos dinâmicos para a avaliação do risco de subscrição referente ao prêmio para uma determinada seguradora utilizando-se dados dos ramos de automóveis. Para a obtenção da distribuição preditiva do requerimento de capital baseado neste risco foram executados simulações MCMC no pacote WinBUGS. Foram considerados nesta tarefa os fatores de risco que captaram as diferenças inerentes à região, código tarifário e tipo de contrato (pessoa física ou jurídica) para o ramo de automóveis (coberturas de casco, RCF-DM, RCF-DP e APP).

Assim, foram propostos diversos modelos levando-se em consideração a evolução temporal ocorrida nos parâmetros. Após a análise de convergência e comparação destes modelos, foi verificado que o modelo mais adequado é o *Modelo III* descrito na seção 4. Mais importante do que os valores obtidos de requerimento de capital de acordo com a probabilidade crítica, o que deve ser apreciado é a

abordagem bayesiana. Com tal ferramenta foi possível obter a distribuição do resultado de subscrição por meio de simulações, ou seja, resultados baseados em mais informações estatísticas do que médias e variâncias dos valores observados.

É importante frisar que para se modelar o risco de subscrição total da companhia faz-se necessário executar a rotina de avaliação do risco de prêmio (precificação) em conjunto com a de avaliação do risco de provisão. Tal procedimento captaria as possíveis correlações entre estes riscos. Como trabalho futuro, pode ser considerado no desconto dos fluxos de caixa, principalmente para o caso de contratos de longo prazo, a estrutura a termo de taxa de juros. Outros fatores que devem ser sensibilizados para grandes riscos são as características dos contratos de resseguro. Estes contratos produzem mudanças significativas no risco total a que as seguradoras estão expostas e sensibilizam o resultado de subscrição.

7. Referências Bibliográficas

- Bowers, N.L. et al. (1997). Actuarial Mathematics. Second Edition. SOA – Society of Actuaries.
- Brooks, S.P., and Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics* 7 (4) pp. 434-455.
- Felblum, S. "NAIC Property/Casualty Insurance Company Risk Based Capital Requirements". In: *Proceedings of the Casualty Actuaries Society*, 83, 1996, pp. 297-435.
- Gamerman, D. (1996). Simulação Estocástica via Cadeias de Markov. XII SINAPE – ABE.
- Gelfand, A.E., and Ghosh, S.K. (1998). Model choice: A minimum posterior predictive loss approach. *Biometrika*, 85, 1, pp. 1-11.
- Gelfand, A.E., and Smith, A.F.M. (1990). Sampling-based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association* 85 pp. 398-409.
- Gelfand, A.E., Hills, S.E., Racine-Poon, A., and Smith, A.F.M. (1990). Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling. *Journal of the American Statistical Association* 85 pp. 972-985.
- Gelman, A., and Rubin, D. (1992). Inference from Iterative Simulation using Multiple Sequences. *Statistical Science* 7, pp. 457-511.
- Gerber, H. U. An introduction to mathematical risk theory. Homework. Huebner Foundation Studies, 1979.
- Hull, J. C., Options, Futures and other Derivatives. 5ª ed., 2002, Prentice Hall.
- Migon, H. S. & Gamerman, D. Statistical Inference: an Integrated Approach, Arnold Publisher, 1999.
- Ryan, J. P., Archer-Lock, P. R., Czernuszewics, A. J., Gillott, N. R., Hinton, P. H., Ibeson, D., Malde, S. A., Paul, D. and Shah, N. (2001). Financial Condition Assessment. Institute of Actuaries and Faculty of Actuaries. Apresentado ao Institute of Actuaries, Londres – Março/2001.
- Scollnik, D.P.M. (2000). Actuarial Modeling with MCMC and BUGS.
- Spiegelhalter, D.J. et al (2003). WinBUGS User Manual, version 1.4. Disponível em www.mrc-bsu.cam.ac.uk/bugs.