

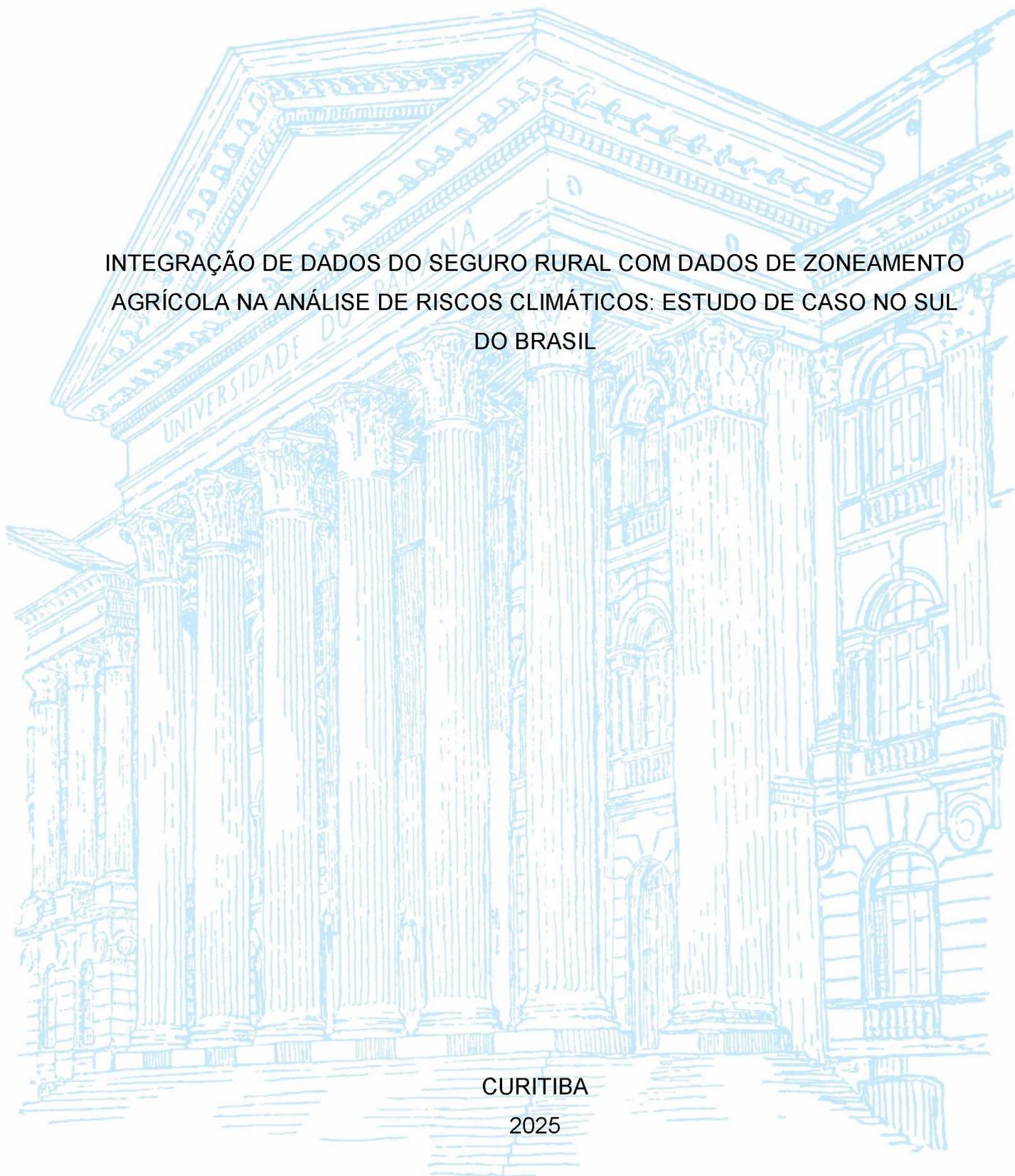
UNIVERSIDADE FEDERAL DO PARANÁ

LUCIANA JUSTINA DA SILVA

INTEGRAÇÃO DE DADOS DO SEGURO RURAL COM DADOS DE ZONEAMENTO
AGRÍCOLA NA ANÁLISE DE RISCOS CLIMÁTICOS: ESTUDO DE CASO NO SUL
DO BRASIL

CURITIBA

2025



LUCIANA JUSTINA DA SILVA

INTEGRAÇÃO DE DADOS DO SEGURO RURAL COM DADOS DE ZONEAMENTO
AGRÍCOLA NA ANÁLISE DE RISCOS CLIMÁTICOS: ESTUDO DE CASO NO SUL
DO BRASIL

Dissertação apresentada ao curso de Pós-Graduação em Métodos Numéricos em Engenharia, Setores de Tecnologia e Ciências Exatas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Métodos Numéricos em Engenharia, Área de Concentração em Programação Matemática e Linha de Pesquisa em Métodos Estatísticos.

Orientador: Prof. Dr. Anselmo Chaves Neto

Coorientador: Prof. Dr. Gilson Martins

CURITIBA

2025

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

Silva, Luciana Justina da

Integração de dados do seguro rural com dados de zoneamento agrícola na análise de riscos climáticos: estudo de caso no sul do Brasil / Luciana Justina da Silva. – Curitiba, 2025.

1 recurso on-line : PDF.

Dissertação (Mestrado) - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Métodos Numéricos em Engenharia.

Orientador: Anselmo Chaves Neto

Coorientador: Gilson Martins

1. Zoneamento agrícola. 2. Climatologia agrícola. 3. Administração de risco financeiro. 4. Teoria bayesiana de decisão estatística. I. Universidade Federal do Paraná. II. Programa de Pós-Graduação em Métodos Numéricos em Engenharia. III. Chaves Neto, Anselmo. IV. Martins, Gilson. V. Título.

Bibliotecário: Elias Barbosa da Silva CRB-9/1894



TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação MÉTODOS NUMÉRICOS EM ENGENHARIA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **LUCIANA JUSTINA DA SILVA**, intitulada: **INTEGRAÇÃO DE DADOS DO SEGURO RURAL COM DADOS DE ZONEAMENTO AGRÍCOLA NA ANÁLISE RISCOS CLIMÁTICOS: ESTUDO DE CASO NO SUL DO BRASIL**, sob orientação do Prof. Dr. ANSELMO CHAVES NETO, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestra está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 04 de Abril de 2025.

Assinatura Eletrônica

07/04/2025 09:31:32.0

ANSELMO CHAVES NETO

Presidente da Banca Examinadora

Assinatura Eletrônica

04/04/2025 15:10:24.0

ROBERTO ROCHADELLI

Avaliador Externo (UNIVERSIDADE FEDERAL DO PARANÁ - SETOR
PALOTINA)

Assinatura Eletrônica

04/04/2025 15:54:50.0

CASSIUS TADEU SCARPIN

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

07/04/2025 10:25:46.0

SOLANGE REGINA DOS SANTOS

Avaliador Externo (UNIVERSIDADE ESTADUAL DO PARANÁ)

Assinatura Eletrônica

04/04/2025 15:27:49.0

MARIANA KLEINA

Avaliador Externo (UNIVERSIDADE FEDERAL DO PARANÁ -UFPR)

Ao meu filho, Davi.

AGRADECIMENTOS

Escrever esta dissertação foi um desafio que exigiu não apenas esforço acadêmico, mas também resiliência para equilibrar os papéis de mãe, profissional e estudante. Conciliar a vida familiar, os compromissos profissionais e a dedicação ao curso foi uma jornada de aprendizado, marcada por momentos de superação e crescimento pessoal.

Agradeço primeiramente à minha família, em especial ao meu marido, Guilherme, que sempre esteve ao meu lado, oferecendo apoio incondicional, compreensão e amor. Ao meu filho Davi, cuja luz e energia me inspiraram nos dias mais desafiadores, a Clara, que compartilhou comigo essa caminhada com paciência e carinho.

Aos amigos que me acompanharam ao longo dessa trajetória, meu profundo reconhecimento. Cada conversa, palavra de incentivo e momento compartilhado ajudaram a tornar este processo mais leve e significativo.

Ao corpo docente da Universidade Federal do Paraná, minha gratidão por sua dedicação e excelência, que me proporcionaram uma formação enriquecedora. Em especial, agradeço aos meus orientadores, que com generosidade, rigor acadêmico e incentivo, guiaram minha pesquisa, tornando possível a realização deste trabalho.

A todos vocês, que foram parte fundamental desta conquista, meus mais sinceros agradecimentos.

"Liberdade é pouco. O que eu desejo ainda não tem nome."

(Clarice Lispector, 1998)

RESUMO

O aprimoramento da análise de seguro rural com dados de zoneamento agrícola é uma abordagem inovadora que busca tornar a gestão de riscos no setor agrícola mais eficiente e sustentável. Nesse contexto, o seguro agrícola surge como instrumento econômico indispensável para mitigar riscos, proporcionando estabilidade econômica aos agricultores. Com base nisto, este trabalho integra dados de zoneamento agrícola em análise de risco para seguro rural, com foco na região Sul do Brasil. A metodologia utiliza informações do Zoneamento Agrícola de Risco Climático (ZARC) e dados de seguros disponibilizados pelo Ministério da Agricultura, convertendo-os em variáveis distribucionais para um modelo probabilístico Bayesiano. Isso permite avaliação detalhada de risco, considerando cenários otimistas e pessimistas baseados em dados de solo do ZARC. Esses cenários são combinados com informações de seguros para gerar distribuições de risco mais precisas. O método possibilita a comparação de riscos entre estados, municípios e culturas agrícolas, de soja e trigo, contribuindo para uma classificação de risco robusta na região Sul. A abordagem proposta pode aprimorar significativamente a gestão de riscos no setor agrícola, beneficiando seguradoras, resseguradoras e órgãos governamentais e privados.

Palavras-chave: Gestão de riscos nas culturas agrícolas. Análise Bayesiana. Classificação de risco agrícola. Seleção adversa. Distribuição beta.

ABSTRACT

The enhancement of rural insurance analysis with agricultural zoning data is an innovative approach aimed at making risk management in the agricultural sector more efficient and sustainable. In this context, agricultural insurance emerges as an indispensable tool to mitigate risks, providing economic stability for farmers. Based on this, the present study integrates agricultural zoning data into risk analyses for rural insurance, focusing on the southern region of Brazil. The methodology utilizes information from the Agricultural Zoning for Climate Risk (ZARC) and insurance data provided by the Ministry of Agriculture, converting them into distributional variables for a Bayesian model. This approach enables detailed risk assessments by considering optimistic and pessimistic scenarios based on ZARC soil data. These scenarios are combined with insurance information to generate more accurate risk distributions. The method allows for risk comparisons across states, municipalities, and agricultural crops such as soybeans and wheat, contributing to a robust risk classification in southern Brazil. The proposed approach has the potential to significantly improve risk management in the agricultural sector, benefiting insurers, reinsurers, and both governmental and private institutions.

Keywords: Risk management in crops. Bayesian analysis. Agricultural risk classification. Adverse selection. Beta distribution.

LISTA DE FIGURAS

FIGURA 1 – Mapa da Crescente Fértil.....	22
FIGURA 2 – Número de apólices de seguro agrícola contratadas por ano no Brasil para todas as culturas agrícolas.	26
FIGURA 3 – Número de apólices de seguro agrícola contratadas por ano no Brasil para a cultura da soja.....	27
FIGURA 4 – Número de apólices de seguro agrícola contratadas por ano no Brasil para a cultura da trigo.	27
FIGURA 5 – Representação da Inferência Bayesiana	37
FIGURA 6 – Curvas da distribuição Beta	44
FIGURA 7 – Critérios de classificação com base na posição das distribuições posterior, melhor cenário e pior cenário.....	52
FIGURA 8 – Períodos de Plantio de Trigo e Soja nos Estados do Sul do Brasil ao Longo do Ano.....	53
FIGURA 9 – Comparação de risco para os estados do Paraná, Rio Grande do Sul e Santa Catarina usando a distribuição posterior versus as distribuições do melhor cenário e do pior cenário.....	58
FIGURA 10 – Mapa de Risco dos Municípios – Posterioris (2017/2018 a 2021/2022)	61
FIGURA 11 – Mapa de classes por desvio padrão das Prioris (2017/2018 a 2021/2022).....	63
FIGURA 12 – Mapa das razões entre Posteriori e Priori (2017/2018 a 2021/2022)..	64

LISTA DE TABELAS

TABELA 1 – Número total de apólices, apólices com perdas e percentual de perdas para as culturas de interesse por período de análise.....	54
TABELA 2 – Contagem de municípios por categoria de risco.....	59

LISTA DE ABREVIATURAS OU SIGLAS

CNSA	- Companhia Nacional de Seguro Agrícola
COSESP	- Companhia de Seguros do Estado de São Paulo
DP	- Desvio Padrão
EMBRAPA	- Empresa Brasileira de Pesquisa Agropecuária
FESR	- Fundo de Estabilidade do Seguro Rural
ISNA	- Índice de Satisfação da Necessidade de Água
MAPA	- Ministério da Agricultura e Pecuária
PROAGRO	- Programa de Garantia da Atividade Agropecuária
PRONAF	- Programa Nacional de Fortalecimento da Agricultura Familiar
PSR	- Programa de Subvenção ao prêmio do Seguro Rural
SNSP	- Sistema Nacional de Seguros Privados
SUSEP	- Superintendência de Seguros Privados
ZARC	- Zoneamento Agrícola de Risco Climático

LISTA DE SÍMBOLOS

- Σ - somatório de números
- Π - produtório de números
- α - parâmetro alfa
- β - parâmetro beta
- $\Gamma(\cdot)$ - função Gamma;
- θ - variável theta
- Θ - espaço paramétrico

SUMÁRIO

1 INTRODUÇÃO	16
1.1 JUSTIFICATIVA	18
1.2 OBJETIVOS	19
1.2.1 Objetivos específicos.....	19
1.3 ORGANIZAÇÃO DO TRABALHO	19
2 REVISÃO DE LITERATURA	21
2.1 A CULTURA DA SOJA.....	21
2.2 A CULTURA DO TRIGO	21
2.3 GESTÃO DE RISCOS RURAIS	23
2.3.1 Seguro Agrícola no Mundo.....	23
2.3.2 Seguro Agrícola no Brasil.....	24
2.3.3 PROAGRO - Programa de Garantia da Atividade Agropecuária.....	28
2.3.4 PSR – Programa de Subvenção ao Prêmio do Seguro Rural	29
2.3.4.1 Zoneamento Agrícola de Risco Climático (ZARC).....	29
2.4 SELEÇÃO ADVERSA	31
3 REFERENCIAL TEÓRICO	34
3.1 INFERÊNCIA BAYESIANA	34
3.1.1 Abordagem Clássica ou Frequentista	34
3.1.2 Abordagem Bayesiana	35
3.1.2.1 Regra de Bayes.....	37
3.1.2.2 Escolha da Priori.....	38
3.1.2.2.1 Prioris subjetivas.....	40
3.1.2.2.2 Prioris objetivas (não informativas).....	41
3.1.2.2.3 Misturas de prioris	42
3.2 DISTRIBUIÇÃO BETA.....	43
4 MATERIAL E MÉTODOS	45
4.1 MATERIAL	45
4.2 MÉTODOS	45
4.2.1 Cálculo para a média estadual	48
4.2.2 Cenários de Análise	48
4.2.3 Dados de seguros e transformação	49
4.2.4 Aplicação do método de atualização Bayesiana e a distribuição posterior	50

4.2.5 Aplicação empírica e validação do modelo	51
5 APRESENTAÇÃO DOS RESULTADOS	53
6 CONCLUSÃO	66
6.1 RECOMENDAÇÕES PARA TRABALHOS FUTUROS	67
REFERÊNCIAS.....	68
ANEXO 1 – DETALHAMENTO BASE DE DADOS DE SEGUROS.....	72
APÊNDICE 1 – SCRIPT DESENVOLVIDO NO SOFTWARE R PARA CURVAS DA DISTRIBUIÇÃO BETA.....	73
APÊNDICE 2 – SCRIPT DESENVOLVIDO NO SOFTWARE R PARA REPRESENTAÇÃO VISUAL DA INFERÊNCIA BAYESIANA.....	74

1 INTRODUÇÃO

O aperfeiçoamento da análise de seguro rural com a introdução de dados de zoneamento agrícola envolve considerar o cenário atual da agricultura e do seguro agrícola, bem como as tendências e desafios enfrentados nesses setores.

A agricultura desempenha um papel fundamental na segurança alimentar, na economia e no sustento das pessoas no mundo. No entanto, os agricultores estão sujeitos a vários riscos, como condições climáticas adversas, pragas, doenças de plantas e flutuações de preços, que podem afetar sua produtividade e renda.

O seguro agrícola é um instrumento importante para mitigar os riscos enfrentados pelos agricultores, oferecendo proteção financeira em caso de perdas relacionadas à produção agrícola. É essencial para promover a estabilidade econômica no setor agrícola e garantir a sustentabilidade das operações agrícolas.

Inicialmente, tem-se a criação do seguro agrário no ano de 1954, operado por meio da Companhia Nacional de Seguro Agrícola – CNSA. Este foi o início das atuações mais expressivas nas políticas de gerenciamento de risco no Brasil. Entre 1964 e 1966, o governo federal impôs a obrigatoriedade do seguro agrícola nas atividades financiadas por bancos ou instituições sob o controle da União. Esta prática foi estendida a todas as instituições nacionais que financiavam o crédito rural, e então instituiu-se o Fundo de Estabilidade do Seguro Rural – FESR (RAMOS, 2009).

O programa de seguro rural atualmente vigente, foi estabelecido por meio da Lei 10.823/2003 que definiu as regras para o seguro rural como um instrumento privado para a gestão de riscos. Para ter direito à proteção de riscos, o contratante paga um prêmio ao segurador (uma empresa privada de seguros) que assume o ônus financeiro no caso de perdas e danos na atividade segurada. Um fato relevante é que tanto o seguro rural como o Proagro (Programa de Garantia da Atividade Agropecuária), podem ser utilizados como garantias para a contratação de crédito rural. Ainda que o agricultor receba, geralmente, subvenção ao prêmio do seguro do governo federal, toda a operação do seguro rural é feita por agentes privados, desde a operacionalização até as vistorias e pagamentos de indenizações no caso de sinistros (MARTINS, 2023).

A subvenção ao seguro rural foi regulamentada em 2004 com a criação do Programa de Subvenção do Seguro Rural – PSR, pela LC 5121/2004 com a finalidade de aumentar a participação dos produtores no seguro rural. Oferece ao agricultor a

oportunidade de segurar sua produção com custo reduzido, por meio de auxílio financeiro do governo federal. O PSR destina-se a subvenção do prêmio do seguro rural, ou seja, fornece um subsídio para contratação de apólices de seguro rural, de modo que parte do prêmio que seria pago pelo produtor é realizada pelo governo que faz o pagamento diretamente a seguradora (MAPA, 2024).

Para ter acesso à subvenção o produtor deve seguir o Zoneamento Agrícola de Risco Climático (ZARC) e o calendário de plantio definido pelo Ministério da Agricultura e Pecuária - MAPA. O Zoneamento Agrícola de Risco Climático é um instrumento desenvolvido pela parceria de instituições federais e estaduais, junto com a Empresa Brasileira de Pesquisa Agropecuária – EMBRAPA. O objetivo é indicar períodos de plantio por município, considerando o tipo de solo e ciclo da cultura, a fim de minimizar as chances de que eventos climáticos coincidam com o período mais vulnerável do plantio (MITIDIERI; MEDEIROS, 2008). Reduzindo assim, a ocorrência de sinistro em diversas regiões, beneficiando produtores e seguradoras e reduzindo as perdas, tornando os programas mencionados anteriormente cada vez mais eficazes.

A análise de risco no seguro agrícola tradicionalmente se baseia em dados históricos de produtividade e perdas. No entanto, o zoneamento é um instrumento de informação sobre riscos que, de maneira geral, serve como um roteiro para acessar as políticas públicas e, até mesmo, para delimitar áreas de atuação das seguradoras e Instituições Financeiras. No entanto, não há uma prática consolidada de utilizar as informações de probabilidade do ZARC como indicador de risco, em conjunto com dados relativos ao seguro.

Os avanços tecnológicos e o aumento da disponibilidade de dados trouxeram novas oportunidades para aprimorar a análise de seguro agrícola. Os dados de zoneamento agrícola, que consideram as características específicas das áreas agrícolas, como clima, solo, topografia e práticas de cultivo, podem enriquecer a análise de risco e a precificação do seguro agrícola.

Ao integrar dados de zoneamento agrícola na análise de seguro, é possível melhorar a precisão na avaliação de riscos, personalizar as políticas de seguro de acordo com as características das áreas agrícolas e promover uma gestão mais eficaz dos riscos, beneficiando tanto os agricultores quanto as seguradoras.

Ao considerar esses elementos, destaca-se a necessidade e o potencial de aprimorar a análise de seguro agrícola por meio da incorporação de dados do ZARC, visando a uma gestão mais eficiente e sustentável dos riscos neste setor.

1.1 JUSTIFICATIVA

A agricultura é uma parte vital da economia mundial. Com as necessidades de adaptações climáticas e outros desafios, a existência do seguro agrícola se tornou ainda mais crucial para garantir a segurança financeira dos agricultores e a estabilidade do setor.

Ao aprimorar a análise de seguro agrícola com dados de zoneamento agrícola, é possível identificar e avaliar melhor os riscos associados a diferentes áreas e culturas, permitindo seguros mais personalizados e precisos.

Os dados do Zoneamento sobre janelas de plantio correspondem a probabilidades de frequência de eventos desfavoráveis. Esses dados são calculados de forma científica, com base em dados históricos. De outro lado, o seguro gera dados de sinistros que refletem o perfil de risco assumido pelas seguradoras, influenciado pelo apetite de risco dos agentes econômicos, o que pode levar à ocorrência de seleção adversa. Dessa forma, a análise de riscos rurais pode ser melhorada com a análise em um modelo que possa juntar as diferentes fontes de dados em um único modelo, com objetivo de oferecer uma análise de risco adaptada à realidade do mercado.

Este trabalho aplica um modelo para integração de dados de zoneamento agrícola de risco climático (ZARC) na análise de risco de seguros. Consiste na combinação de informações públicas oficiais do programa de zoneamento brasileiro (ZARC) com dados de seguros abertos fornecidos pelo Ministério da Agricultura, Pecuária e Abastecimento do Brasil (MAPA), conforme modelo proposto por Martins e Signorini (2025).

São utilizadas neste trabalho informações sobre tipos de solo do ZARC para desenvolver cenários básicos de melhor e pior caso e calcular distribuições posteriores usando dados de seguros. A estrutura resultante permitirá a comparação de estados e municípios, tipos de culturas e classificação geral de risco. Especificamente, o estudo tem como objetivo analisar o risco das safras de soja e trigo na região sul do Brasil, resultando em classificações de risco consistentes nessas

culturas e região. A estrutura proposta tem o potencial de melhorar a análise de gestão de riscos agrícolas para resseguradores, seguradoras, agências governamentais e empresas privadas.

Enfim, este trabalho apresenta um procedimento que tem o potencial de melhorar a análise de risco e a tomada de decisão no setor agrícola.

Os dados de zoneamento agrícola podem ajudar na tomada de decisão informada sobre políticas agrícolas, investimentos em infraestrutura rural e desenvolvimento sustentável e, desta forma, contribuindo para o crescimento e a resiliência do setor.

1.2 OBJETIVOS

Este trabalho tem como objetivo principal apresentar um procedimento que pode melhorar a análise de riscos agrícolas, por meio da aplicação de modelo estatístico para análise conjunta de dados de zoneamento agrícola (ZARC) com dados de Seguro Rural.

1.2.1 Objetivos específicos

a) Validação de um modelo proposto por Martins e Signorini (2025) de análise de integração de dados do ZARC, com dados do Seguro Agrícola;

b) Aplicação empírica do modelo para a análise de riscos da soja e do trigo nos três estados do Sul do Brasil.

1.3 ORGANIZAÇÃO DO TRABALHO

O Capítulo 2, é realizada uma revisão da literatura, abordando tópicos relevantes como as culturas de soja e trigo e as estratégias de gestão de riscos no setor rural.

O Capítulo 3 é dedicado ao referencial teórico, explorando conceitos importantes como a Inferência Bayesiana e suas características particulares, além de discutir a distribuição Beta e o fenômeno da seleção adversa no contexto de seguros agrícolas.

No Capítulo 4, são apresentados os materiais e os métodos empregados na pesquisa, detalhando a abordagem utilizada para a análise dos dados.

A apresentação dos resultados ocorre no Capítulo 5, onde são discutidos os principais achados da pesquisa. Por fim, a conclusão é apresentada no capítulo 6 e junto sugestões para futuros estudos.

2 REVISÃO DE LITERATURA

2.1 A CULTURA DA SOJA

A soja (*Glycine max (L.) Merrill*) é uma das culturas mais importantes tanto no cenário mundial quanto nacional. Seu uso como alimento remonta a mais de 5.000 anos, tendo a China como centro de origem, onde era cultivada pelo imperador Shennung como uma fonte alternativa de nutrição. Atualmente, o uso de seus grãos cresce anualmente, servindo como fonte de proteína para a alimentação humana e animal devido ao seu alto teor proteico (40%) (SEDIYAMA, 2009), além de ser utilizada na agroindústria, indústria química e, mais recentemente, como matéria-prima para a produção de biodiesel (COSTA NETO e ROSSI, 2000).

No contexto global, a soja desempenha um papel significativo no agronegócio internacional, sendo o quarto grão mais produzido e consumido no mundo, ficando atrás do arroz, trigo e milho. Na safra de 2021/2022, foram produzidas aproximadamente 758 milhões de toneladas (FAO, 2023). O Brasil é o principal produtor dessa cultura, seguido pelos Estados Unidos e Argentina, correspondendo a 33,7%, 31,4% e 11,9% da produção mundial, respectivamente (FAO, 2023).

No Brasil, a soja é a principal cultura em termos de área e volume de produção, com o estado do Mato Grosso liderando como principal produtor, representando cerca de 29% da produção total do país. Em 2023, a produção nacional alcançou aproximadamente 152 milhões de toneladas de grãos, segundo levantamento do IBGE (IBGE, 2024). Isso valoriza ainda mais o setor agrícola brasileiro, que desempenha um papel crucial no desenvolvimento do país como provedor de alimentos e matérias-primas, além de contribuir para a geração de empregos, demandando mão de obra e fomentando o mercado interno e externo com o aumento nos níveis de exportação (BUAINAIN e GARCIA, 2015; MARANHÃO e VIEIRA FILHO, 2016).

2.2 A CULTURA DO TRIGO

O trigo (*Triticum spp.*) é considerado uma das culturas mais importantes tanto no contexto mundial quanto nacional. Sua origem data de cerca de 10.000 anos atrás e ocorreu na região conhecida como Crescente Fértil, figura 1 adiante.

Geograficamente, a região do Crescente Fértil fica entre os rios Tigre, Eufrates, Jordão e Nilo, e tem um formato que se assemelha a uma lua crescente. A figura 1 mostra a situação dessa região.

FIGURA 1 – Mapa da Crescente Fértil.



FONTE: Wikipédia, 2021.

A cultura do trigo tem desempenhado um papel crucial na alimentação humana desde os primórdios da civilização. Este cereal foi uma das primeiras plantas a serem cultivadas de forma sistemática, permitindo o desenvolvimento de sociedades agrárias estáveis (TUDGE, 2000).

Atualmente, o trigo continua a ser uma cultura vital, não apenas como um alimento básico para bilhões de pessoas, mas também como uma *commodity* crucial para a economia global. O teor médio de proteína do trigo está em torno de 12-14%. Este cereal é amplamente utilizado na produção de pão, massas, bolos e uma ampla gama de produtos alimentícios (SHEWRY, 2009).

No contexto global, o trigo é um dos cereais mais cultivados, ocupando o segundo lugar em termos de área plantada e produção total, logo após o milho, e seguido pelo arroz (FAO, 2022). Em 2022, a produção mundial de trigo foi estimada

em aproximadamente 778 milhões de toneladas, com a União Europeia, China, e Índia liderando como os maiores produtores (FAO, 2023).

No Brasil, o trigo é uma cultura de grande importância, especialmente nas regiões Sul e Sudeste. Em 2022, a produção brasileira de trigo atingiu cerca de 9,5 milhões de toneladas, com o estado do Paraná sendo o maior produtor, seguido pelo Rio Grande do Sul (CONAB, 2022). O cultivo de trigo no Brasil enfrenta desafios climáticos, mas avanços na pesquisa agrícola e no desenvolvimento de cultivares adaptados têm contribuído para aumentos constantes na produtividade (EMBRAPA, 2023).

Além de sua relevância alimentar e econômica, o trigo também tem um impacto significativo na cultura e na sociedade. Desde os rituais religiosos da antiguidade até os pratos modernos da culinária global, o trigo continua a ser um símbolo de sustento e prosperidade.

Tanto a soja quanto o trigo exigem um solo com pH alcalino, geralmente variando entre 6,0 e 7,5. Fora dessa faixa, é necessário realizar a calagem, que consiste em incorporar calcário ao solo para corrigir a acidez e adequar o pH às necessidades das culturas. Na década de 70 foi criado o programa PROCAL de financiamento para os agricultores corrigirem a acidez dos solos das suas propriedades agrícolas. Esse investimento possibilitou o aumento da produtividade agrícola da Região Sul do Brasil, bem como a introdução massiva de novas culturas.

2.3 GESTÃO DE RISCOS RURAIS

Os diversos riscos naturais que ameaçam as atividades agrícolas exigem a busca por alternativas que possam minimizar os impactos dos eventos adversos. Assim, o seguro agrícola desponta como um recurso essencial para o desenvolvimento e adoção de tecnologias, além de garantir estabilidade financeira no agronegócio, reduzindo o êxodo rural e as dívidas dos empréstimos dos produtores (BRISOLARA, 2013).

2.3.1 Seguro Agrícola no Mundo

Ao contrário do que possa parecer, a demanda por seguros agrícolas pelos agricultores é relativamente baixa em quase todo o mundo, com exceção dos Estados

Unidos, onde essa modalidade alcançou grande sucesso (OZAKI, 2005). Contudo, esse processo exigiu um prazo muito longo, de algumas décadas, sendo muito dispendioso e precisando, de forma determinante, da participação efetiva do Estado para reduzir a desconfiança dos produtores e das empresas agrícolas (GUIMARÃES & NOGUEIRA, 2009).

Desde a década de 1930, ou seja, com a implantação do *New Deal* pelo presidente Franklin Delano Roosevelt no vale do rio Tennessee, os Estados Unidos vêm desenvolvendo medidas para apoio ao seguro agrícola. Entretanto, somente a partir da década de 90, houve um aumento expressivo das áreas rurais seguradas, quando então o seguro foi definitivamente incorporado à política agrícola norte-americana (OZAKI, 2005). Atualmente, o seguro agrícola é praticamente uma obrigação para os agricultores americanos, onde mais de 90% das lavouras possuem seguro (Sistema Nacional de Agricultura (SNA), AGROANALYSIS, 2007).

Sendo assim, os países que visam utilizar esse instrumento em sua política agrícola, como o Brasil, precisam avaliar os requisitos para o seu sucesso e estar atentos para possíveis problemas gerados por conta de sua massificação. Na maioria dos países onde o seguro agrícola está mais desenvolvido, é comum a intervenção estatal por meio de subsídios, tanto para os prêmios de seguro quanto para as despesas operacionais das seguradoras. Esse apoio visa aumentar a viabilidade econômica do seguro agrícola, tornando-o mais atrativo tanto para os agricultores quanto para as seguradoras (GUIMARÃES & NOGUEIRA, 2009). No entanto, esse tipo de programa demanda grandes montantes de recursos fiscais, que estão disponíveis em quantidade suficiente geralmente apenas em países desenvolvidos, a exemplo do caso americano (GUIMARÃES & NOGUEIRA, 2009).

2.3.2 Seguro Agrícola no Brasil

A primeira experiência brasileira com seguros rurais aconteceu em 1939, quando o Estado de São Paulo instituiu o seguro obrigatório (Lei nº 10.554, de outubro de 1939) para danos causados por granizo nas lavouras de algodão. Outros programas similares surgiram no estado, como a Carteira de Seguro Agrícola contra Granizo para a Viticultura em 1948 (Lei nº 11, de 19 de julho de 1948) e a Carteira de

Seguro Agrícola contra Geada para Horticultura em 1964 (Lei nº 8.375, de 28 de outubro de 1964) (SILVA; TEIXEIRA; SANTOS, 2014).

No âmbito regional, destaca-se a Associação dos Fumicultores do Brasil (Afubra), criada em 1955, que utiliza recursos dos próprios associados para indenizar produtores de fumo do Paraná, Santa Catarina e Rio Grande do Sul por perdas causadas por granizo e ventos fortes. Na esfera federal, o Instituto de Resseguros do Brasil foi criado em 1948 para garantir maior estabilidade aos produtores de rebanhos e reduzir perdas inesperadas, posteriormente incluindo culturas como trigo, café, videira, arroz e algodão (SILVA; TEIXEIRA; SANTOS, 2014).

Em 1954, foi criada a Companhia Nacional de Seguros Agrícolas (CNSA) pela Lei nº 2.168, mas encerrou suas atividades em 1966. No mesmo período, o Sistema Nacional de Seguros Privados (SNSP) foi instituído, regulando as atividades de seguros no Brasil com a criação de decretos em 1966 e 1967. Em 1967, o seguro obrigatório para cobrir danos agropecuários e o Fundo de Estabilidade do Seguro Rural (FESR) foram introduzidos. Em 1969, foi criada a Companhia de Seguros do Estado de São Paulo (COSESP), que expandiu em 1997, mas sofreu grandes prejuízos. Em 1970, a Resolução nº 5 do Conselho Nacional de Seguros Privados determinou as modalidades de seguros agrários, sendo eles:

- Seguro agrícola: destinado a cobertura de perdas decorrentes de fenômenos meteorológicos, doenças e pragas;
- Seguro pecuário: destinado a indenizações em relação a mortes de animais causadas por acidentes ou doenças;
- Seguro de crédito: destinado a comercialização de produtos agropecuários.

Em 1973, o Programa de Garantia da Atividade Agropecuária (PROAGRO) foi criado pela Lei nº 5.969, vinculando o seguro às operações de crédito agropecuário (SILVA; TEIXEIRA; SANTOS, 2014).

Através da Companhia Nacional do Seguro Agrícola, o seguro agrícola brasileiro passou a ganhar escala expressiva. Nos meados dos anos 90, outras seguradoras tentaram, sem sucesso, ingressar no setor agropecuário brasileiro. Um dos principais problemas enfrentados eram os preços elevados negociados, inibindo sua difusão no mercado (OZAKI & CAMPOS, 2017).

Uma década depois, em 2003, o governo federal, através da Lei nº 10.823 (BRASIL, 2003), criou o Programa de Subvenção ao Prêmio do Seguro Rural (PSR)

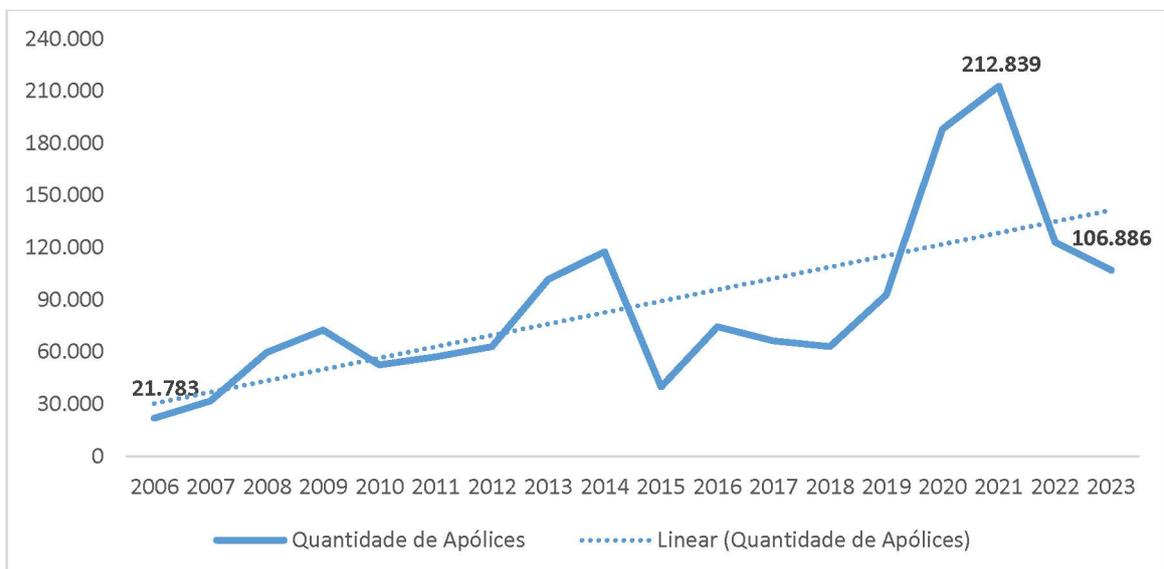
com o objetivo de reduzir o prêmio do seguro pago pelos produtores rurais e incentivar a demanda no mercado de seguro agrícola privado.

O Programa de Subvenção do Prêmio do Seguro Rural (PSR) foi criado como uma fonte de apoio aos produtores rurais na proteção de suas lavouras contra eventos meteorológicos adversos (OZAKI, 2010). Por meio da subvenção econômica ao prêmio do seguro rural, o Governo Federal apoia financeiramente produtores que contratarem essa modalidade de garantia, arcando com parcela dos custos de aquisição do seguro. O percentual de subvenção pago pelo Governo Federal varia de 30% a 35%, de acordo com as prioridades da política agrícola formulada pelo Ministério da Agricultura, Pecuária e Abastecimento (MAPA). As modalidades de seguro rural amparadas pelo PSR são: agrícola, pecuária, florestal e aquícola (MAPA, 2019).

De acordo com o Ministério da Agricultura, Pecuária e Abastecimento (MAPA, 2019), através de seu Programa de Subvenção ao Prêmio do Seguro Rural (PSR), o número de apólices de seguro agrícola contratadas por produtores rurais passou de 21 mil para aproximadamente 106 mil entre os anos de 2006 e 2023 (Figura 2).

Analisando esses números e destacando as apólices contratadas apenas para a soja, neste mesmo período, tem-se que iniciando do patamar de 18 mil contratos atingiu-se aproximadamente 42 mil no período representando um crescimento de aproximadamente 230%, conforme demonstrado na Figura 3.

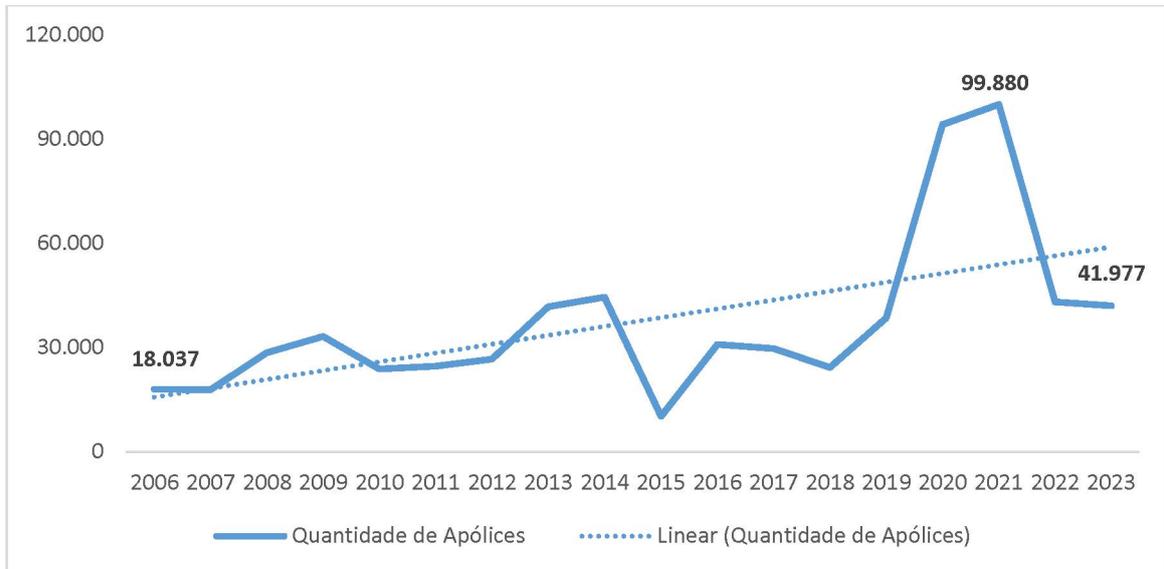
FIGURA 2 – Número de apólices de seguro agrícola contratadas por ano no Brasil para todas as culturas agrícolas.



FONTE: Ministério da Agricultura, Pecuária e Abastecimento – Programa de Subvenção ao Prêmio do

Seguro Rural (PSR). Elaboração: A autora (2025)

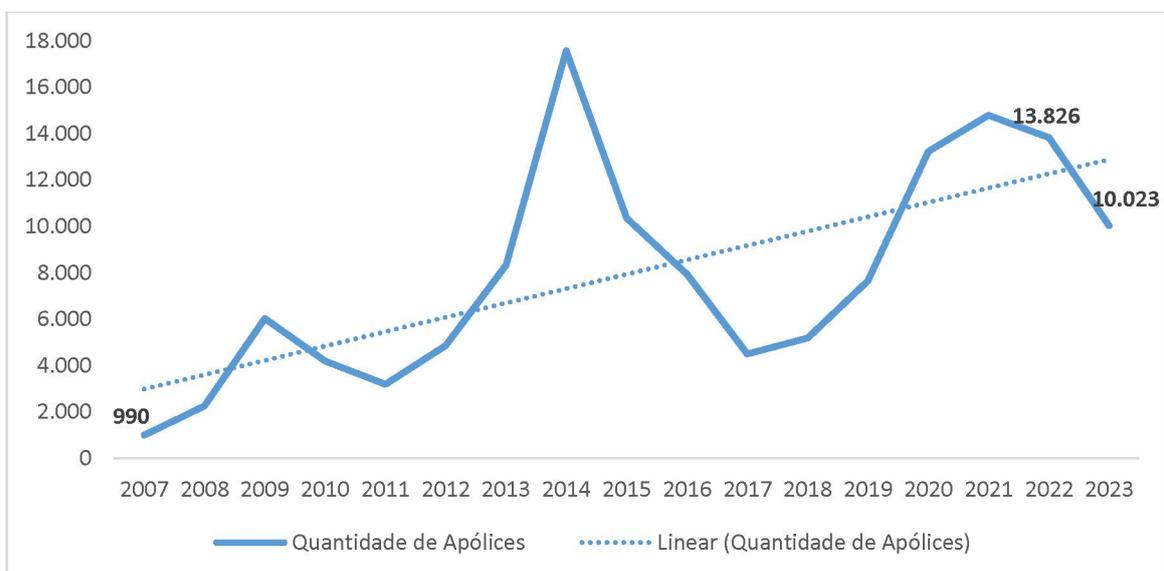
FIGURA 3 – Número de apólices de seguro agrícola contratadas por ano no Brasil para a cultura da soja.



FONTE: Ministério da Agricultura, Pecuária e Abastecimento – Programa de Subvenção ao Prêmio do Seguro Rural (PSR). Elaboração: A autora (2025)

Em relação ao trigo, conforme Figura 4, apesar das oscilações ao longo dos anos, os dados evidenciam uma tendência de crescimento na quantidade total de apólices de seguro entre 2007 e 2023. Partindo de 990 contratos em 2007 e alcançando 10.023 em 2023, representando um crescimento de mais de 1.000%.

FIGURA 4 – Número de apólices de seguro agrícola contratadas por ano no Brasil para a cultura da trigo.



FONTE: Ministério da Agricultura, Pecuária e Abastecimento – Programa de Subvenção ao Prêmio do Seguro Rural (PSR). Elaboração: A autora (2025)

De acordo com Agroanalysis (2007), a falta de difusão do seguro agrícola no Brasil, quando comparado a outros países agrícolas do mundo é evidente. O mercado brasileiro de seguros apresenta um grande potencial a ser explorado. No entanto, para garantir o sucesso dessa modalidade de seguro, é necessário um esforço conjunto dos setores governamental, privado e acadêmico. Isso inclui a implementação de políticas, subvenções e resseguro por parte do governo; a participação ativa de seguradoras, resseguradores e agentes financeiros; e a realização de estudos acadêmicos para definir o modelo ideal de proteção para as lavouras.

2.3.3 PROAGRO - Programa de Garantia da Atividade Agropecuária

Este programa foi criado por meio da Lei nº 5.969 de 1973, criação do Programa de Garantia da Atividade Agropecuária (Proagro), representou uma nova etapa do seguro rural ao vincular o seguro às operações de crédito agropecuário.

O principal objetivo do Proagro era isentar o produtor das obrigações financeiras do crédito rural em caso de perda de produção devido a eventos naturais. Inicialmente, cobria apenas o valor do crédito de custeio contratado. Em 1991, o programa foi reformulado devido a problemas financeiros, fraudes e má gestão, mas o aumento dos prêmios inviabilizou sua operação (BUAINAIN; VIEIRA, 2011).

A partir de 1995, os prêmios e coberturas passaram a ser calculados com base no zoneamento agrícola, permitindo regionalizar datas de plantio e diferenciar prêmios por local e cultura. Houve uma redução no valor dos prêmios e na taxa de sinistralidade, mas também no número de contratações, com a concentração no cultivo de milho, soja, arroz, feijão, algodão, trigo e maçã, e a exclusão de sinistros por déficit hídrico (BUAINAIN; VIEIRA, 2011).

A contratação da cobertura do Proagro ou do Proagro Mais, exige que o produtor pague um valor, denominado Adicional, que corresponde à aplicação da alíquota definida pelo Banco Central sobre o valor protegido. Após análise realizada pelo Banco Central foi identificado, nos últimos anos um déficit expressivo no Programa e, isso gerou a necessidade de recorrer aos recursos do Tesouro Nacional para o pagamento das indenizações.

Até 2023, era comum que uma determinada cultura tivesse uma alíquota única estabelecida para várias regiões ou até mesmo para todo o país. Portanto, nessas condições a alíquota não representava o risco do local/município da área

segurada. Visando melhorar o equilíbrio do Programa, o Banco Central definiu novas alíquotas para as operações contratadas a partir de 1º de julho/2024.

O Proagro é similar ao seguro, mas com importantes diferenças. Primeiramente, ele é administrado diretamente pelo Bacen, fora do alcance regulatório da Susep (Superintendência de Seguros Privados). Além disso, no Proagro, não há seguradoras ou resseguradoras envolvidas, apenas instituições financeiras. Isso significa que o Governo atua como ressegurador do programa, embora existam limites orçamentários a serem seguidos.

2.3.4 PSR – Programa de Subvenção ao Prêmio do Seguro Rural

O Programa de Subvenção ao Prêmio do Seguro Rural (PSR) foi instituído pela Lei nº 10.823, de 19 de dezembro de 2003, e pelo Decreto nº 5.121, de 2004. Este programa tornou-se a principal política do Governo Federal para incentivar o seguro rural no Brasil. O PSR tem como objetivo fornecer subvenção econômica ao valor do prêmio do seguro rural contratado com seguradoras autorizadas (BRASIL, 2018; MAIA; ROITMAN; DE CONTI, 2011; SILVA; TEIXEIRA; SANTOS, 2014).

O PSR (Programa de Subvenção ao Prêmio do Seguro Rural) é direcionado a produtores que não participam dos programas Proagro e Proagro Mais. Este programa oferece subsídios para a contratação de apólices de seguro rural, cobrindo parte do prêmio que seria pago pelo produtor. O governo assume essa parte do custo e realiza o pagamento diretamente às seguradoras (MAPA, 2020).

2.3.4.1 Zoneamento Agrícola de Risco Climático (ZARC)

O Zoneamento Agrícola de Risco Climático (ZARC) fundamenta-se na integração de modelos de simulação do crescimento e desenvolvimento de culturas, bancos de dados climáticos e de solo, técnicas de análise de decisão e instrumentos de geoprocessamento (CUNHA; ASSAD, 2001). Utilizando dados climáticos históricos, é realizada uma simulação do balanço hídrico para avaliar o risco de déficit hídrico nas culturas. Esses balanços hídricos são calculados em períodos de 10 dias, chamados decêndios, e são aplicados a mais de 40 culturas, desde as mais tradicionais e amplamente cultivadas, como milho, soja e café, até as mais exóticas (MONTEIRO, 2009).

Para cada decêndio, é atribuído um nível de risco (20%, 30% e 40%) conforme os cálculos realizados pela Embrapa, estabelecendo as chamadas janelas de plantio. O tipo de solo é uma das variáveis levada em consideração no cálculo do risco, pois a disponibilidade de água para as plantas é significativamente influenciada pela textura do solo. Regiões com solos mais arenosos (Tipo 1) têm uma capacidade muito baixa de retenção de água, resultando em maior risco e janelas de plantio mais curtas (menos decêndios) no ZARC, em comparação com solos do Tipo 3. O zoneamento considera os seguintes tipos de solo:

- Tipo 1 (arenoso): teor de argila maior que 10% e menor ou igual a 15%;
- Tipo 2 (textura média): teor de argila maior que 15% e menor que 35%, com a diferença entre os percentuais de areia e argila sendo menor que 50%;
- Tipo 3 (argiloso): teor de argila maior que 35%.

Um novo sistema de classificação da água disponível no solo foi disponibilizado pela EMBRAPA em 2023, o novo sistema de classificação de solos que deixa de utilizar o tradicional sistema de solos “Tipo 1”, “Tipo 2” e “Tipo 3” e passa a adotar uma nova metodologia que define seis classes de água disponível (AD1, AD2, AD3, AD4, AD5, AD6). Nesse sistema, a classe de AD do solo é determinada com base na sua composição textural completa, ou seja, com os teores de areia, silte e argila. No sistema utilizado nesse trabalho, os tipos do solo eram estabelecidos, basicamente, pelo teor de argila. Para o presente trabalho, foi utilizada a classificação de solos tradicional, com o intuito de seguir a linha de pesquisa utilizada por Martins e Signorini (2024).

Outra variável é o ciclo da cultura em questão, ou seja, se a variedade a ser plantada é de ciclo curto (Grupo I), médio (Grupo II) ou longo (Grupo III). Normalmente, variedades de ciclo mais longo apresentam maior risco e mais restrições quanto à janela de plantio. Isso ocorre porque as plantas permanecem mais tempo no campo, aumentando a probabilidade de serem afetadas por eventos climáticos adversos. Evidentemente, a duração dos ciclos depende do tipo de cultura (MARTINS, 2023). Isso permite aos produtores identificarem o melhor local e período para o cultivo de cada cultura nas diferentes regiões brasileiras, além de relacionar os tipos de cultivares mais adequados e as fases mais sensíveis de cada cultura, minimizando perdas agrícolas (ASSAD et al., 2007).

A metodologia do ZARC considera elementos que influenciam diretamente o desenvolvimento das plantas e, conseqüentemente, a produtividade das culturas, tais

como temperatura, precipitação, umidade relativa do ar, água disponível no solo e demanda hídrica das culturas. No entanto, a principal variável utilizada no ZARC é o índice de satisfação da necessidade de água (ISNA), que expressa a evapotranspiração relativa. Isso é, a razão entre a evapotranspiração real da cultura (ET_r), determinada pela disponibilidade de água no solo, e a evapotranspiração máxima da cultura (ET_c) durante a sua fase crítica (MONTEIRO, 2009). Conforme MONTEIRO (2009), os níveis de ISNA toleráveis pelas culturas variam de 0,55 a 0,65, significando que as culturas podem suportar déficits hídricos relativos $(1 - ET_r/ET_c)$ de 0,35 a 0,45 sem comprometer a produtividade. Para que uma época de semeadura seja considerada de baixo risco, pelo menos 80% dos anos da série histórica analisada devem apresentar um ISNA igual ou superior ao limiar da cultura (0,55 a 0,65). Se essa probabilidade estiver entre 70 e 80%, o risco é classificado como moderado-baixo; valores entre 60 e 70% indicam risco moderado-alto; e abaixo de 60%, o risco é considerado alto, não sendo recomendados os plantios nessas condições.

Seguir as recomendações do ZARC permite aos produtores acessarem os benefícios do Programa de Subvenção ao Prêmio do Seguro Rural (PSR) e obter facilidades na concessão de financiamentos do crédito rural junto às instituições financeiras. Estas, por sua vez, têm a garantia de que os riscos climáticos associados às culturas seguradas e/ou financiadas serão minimizados. Este sistema de zoneamento de risco foi pioneiro a nível mundial e hoje serve de modelo para outros países.

2.4 SELEÇÃO ADVERSA

A seleção adversa é um conceito desenvolvido por George Akerlof em seu artigo de 1970, "*The Market for Lemons*", que trata da assimetria de informações em mercados. A ideia central é que, quando uma das partes em uma transação (geralmente o vendedor) tem mais informações sobre a qualidade de um produto do que a outra parte (geralmente o comprador), isso pode levar a resultados indesejáveis, como a redução da qualidade média dos produtos disponíveis no mercado. Akerlof exemplificou esse fenômeno no mercado de carros usados, onde vendedores de carros de baixa qualidade ("*lemons*") podem explorar a falta de conhecimento dos compradores.

Esse problema é particularmente relevante no setor de seguros. Em mercados de seguros, a seleção adversa ocorre quando indivíduos com maior risco de sinistros têm maior probabilidade de buscar seguros, enquanto aqueles de menor risco podem optar por não participar, principalmente se considerarem os prêmios muito altos para o seu perfil de risco.

Akerlof sugeriu que a introdução de mecanismos para reduzir a assimetria de informações poderia ajudar a mitigar o impacto da seleção adversa. Com o aumento no valor dos sinistros, as companhias de seguros, para manter sua viabilidade financeira, tendem a elevar os valores dos prêmios cobrados. Contudo, esse reajuste pode resultar em uma carteira de segurados composta predominantemente por clientes de maior risco, reduzindo a diversificação e expondo a seguradora a uma probabilidade ainda maior de elevação nos sinistros pagos (BOYER e PORRINI, 2008).

A seleção adversa pode desencadear uma espiral econômica negativa, caracterizada por sucessivas perdas na carteira de seguros. Esse processo ocorre devido ao crescente desequilíbrio entre os valores dos sinistros e os prêmios arrecadados, além da redução da diversificação dos riscos, já que clientes de maior risco passam a predominar na base de segurados.

No contexto das mudanças climáticas e do aumento na frequência e intensidade de eventos climáticos extremos, destaca-se o risco de eventos generalizados como o principal catalisador para o surgimento da seleção adversa no mercado de seguro agrícola. O aumento da exposição a esses eventos eleva os riscos associados ao seguro agrícola, forçando as seguradoras a aumentar os prêmios para assegurar sua sustentabilidade financeira. Isso, por sua vez, pode levar à retenção na carteira apenas dos segurados de maior risco, que possuem maior necessidade de proteger suas produções agrícolas, agravando ainda mais o ciclo de seleção adversa.

O aumento na frequência e na intensidade dos eventos climáticos extremos gera uma expectativa de elevação na probabilidade de sinistros no setor agrícola. Diferentemente de outras modalidades de seguro, onde os eventos de sinistro são geralmente independentes entre os segurados, o seguro agrícola apresenta a particularidade do fenômeno de catástrofe. Esse fenômeno expõe simultaneamente um grande número de segurados em uma mesma região ao mesmo risco, característica que o torna único (BUAINAIN e VIEIRA, 2011). Dessa forma, esses

eventos podem ocasionar valores de sinistros superiores aos montantes arrecadados com os prêmios, comprometendo a sustentabilidade financeira das seguradoras.

Esse cenário é suscetível à ocorrência de seleção adversa. O aumento no risco de sinistros leva as seguradoras a elevar os valores dos prêmios, o que, por sua vez, pode resultar em uma carteira de clientes predominantemente composta por segurados de maior risco (BOYER e PORRINI, 2008). Portanto, identificar, mapear e analisar adequadamente a seleção adversa é essencial para propor soluções eficazes. Entre as possíveis estratégias, destaca-se o incentivo a investimentos em medidas de adaptação aos impactos das mudanças climáticas. Tais ações poderiam mitigar os riscos de sinistro e romper o ciclo econômico negativo provocado pela seleção adversa no mercado de seguros agrícolas.

Além disso, a seleção adversa também pode ser minimizada por meio de regulamentações governamentais, como a exigência de cobertura universal, que obriga tanto indivíduos de baixo risco quanto de alto risco a se inscreverem, diluindo o risco entre uma população maior e tornando os prêmios mais acessíveis para todos.

Portanto, a teoria de Akerlof sobre a seleção adversa oferece uma base sólida para entender como as assimetrias de informações podem impactar o mercado de seguros e por que soluções regulatórias e de mercado são frequentemente necessárias para mitigar esses efeitos.

3 REFERENCIAL TEÓRICO

3.1 INFERÊNCIA BAYESIANA

Atualmente, duas abordagens estatísticas são empregadas na análise de dados. A primeira é a abordagem clássica ou frequentista, que avalia os procedimentos com base em amostras repetidas, partindo de um modelo específico (a verossimilhança), no qual a distribuição condicional de probabilidade é definida com parâmetros desconhecidos. Nesse contexto, as propriedades do procedimento são avaliadas dentro de um cenário de repetição, lidando com parâmetros fixos, mas desconhecidos.

Por outro lado, a abordagem Bayesiana requer tanto um modelo para as amostras quanto uma distribuição a priori para todas as quantidades desconhecidas do modelo, incluindo parâmetros e dados ausentes. A combinação da distribuição a priori com a verossimilhança resulta na obtenção da distribuição a posteriori, que é utilizada para inferência e tomada de decisão.

3.1.1 Abordagem Clássica ou Frequentista

A abordagem clássica ou frequentista, cujos principais pesquisadores expoentes incluem Karl Pearson, Ronald A. Fisher e Jerzy Neyman, tem como objetivo central fazer generalizações sobre uma população com base em uma amostra representativa dessa população. A amostra é definida como um conjunto de observações ou dados estatísticos obtidos a partir de experimentos repetidos em condições semelhantes (também chamados de *ensembles*). Já a população refere-se ao conjunto de todas as observações possíveis de serem observadas sob as mesmas condições. Definindo $\underline{x} = (x_1, x_2, x_3, \dots, x_n)$ como o vetor que representa uma amostra específica ou conjunto de dados e onde n é o tamanho da amostra é possível estabelecer um conjunto \underline{X} que contém todas as amostras possíveis, o que nos permite defini-lo como o espaço amostral (PAULINO et al., 2018).

É importante destacar que a inferência clássica envolve o reconhecimento da variabilidade de uma amostra para outra. Ao estabelecer inferências, deve-se lembrar que os dados observados representam apenas um dos muitos conjuntos de dados que poderiam ter sido obtidos sob as mesmas condições. A interpretação desses dados depende não só do conjunto observado, mas também das hipóteses adotadas

em relação aos possíveis conjuntos alternativos. Seja F_0 uma função de distribuição que descreve a variabilidade ou incerteza na observação da característica populacional X , definida como uma variável aleatória e representada por observações amostrais. Quando se analisa um fenômeno a fim de modelá-lo, usa-se o conhecimento prévio sobre o fenômeno ou processo aleatório para propor uma família de distribuições \mathcal{F} , à qual F_0 pertence, estabelecendo assim um modelo estatístico para a variável aleatória X (PAULINO et al., 2018).

A família \mathcal{F} muitas vezes pode ser parametrizável, ou seja, seus elementos podem ser indexados por um parâmetro. Considerando, então, uma família de distribuições \mathcal{F} indexada pelo parâmetro θ , que está definido no espaço paramétrico Θ . Assim, é possível definir a família como:

$$\mathcal{F} = \{F_\theta: \theta \in \Theta\} \quad (1)$$

Na perspectiva da inferência clássica, o parâmetro θ é desconhecido, mas é tratado como fixo. Uma amostra aleatória de X é extraída de uma população cuja característica de interesse é modelada por F_θ , e com base nos valores observados são feitas inferências sobre o parâmetro θ . No entanto, em muitas situações, essa suposição de fixidez para o parâmetro pode não ser a mais apropriada, ou seja, o parâmetro θ é considerado uma variável aleatória e não tem um valor fixo. Por isso, é útil recorrer à Inferência Bayesiana para fazer estimativas sobre θ , considerando a incerteza de forma mais explícita.

A Inferência Bayesiana, por outro lado, trata da avaliação de hipóteses usando a maior probabilidade da amostra ocorrer, ou seja, o Princípio da Máxima Verossimilhança. Essa ideia decorre da fórmula de Bayes, e é fundamental para métodos computacionais relacionados à inteligência, tais como mineração de dados, entre outros.

3.1.2 Abordagem Bayesiana

A Inferência Bayesiana tem suas origens nos estudos de Thomas Bayes (1701-1761), Harold Jeffreys (1891-1989) e Bruno de Finetti (1906-1985), cujas contribuições foram fundamentais para o desenvolvimento dessa abordagem. Ao contrário da Inferência Clássica, a Inferência Bayesiana interpreta a probabilidade como uma medida de incerteza, expressando o grau de crença em uma hipótese com

base nas evidências disponíveis. Isso permite que se atualizem as crenças sobre um parâmetro à medida que novas informações são adquiridas, proporcionando uma interpretação mais flexível e dinâmica da incerteza.

Na Inferência Bayesiana, o parâmetro θ pode ser interpretado como uma realização de uma variável aleatória θ , cuja variabilidade é descrita por uma distribuição de probabilidade $\pi(\theta)$. Essa distribuição é chamada de distribuição a priori e reflete a crença ou as informações disponíveis sobre o parâmetro antes da realização do experimento. Dessa forma, é possível incorporar todas as informações prévias sobre o parâmetro, sejam elas provenientes de especialistas ou de experimentos anteriores, para abordar o problema atual.

Em alguns casos, como será discutido mais adiante, $\pi(\theta)$ não é uma distribuição de probabilidade. Quando isso ocorre, e $\pi(\theta)$ não é interpretada como uma distribuição probabilística, ela é chamada simplesmente de priori.

Após coletar uma amostra de uma população indexada por θ , as informações sobre o parâmetro são atualizadas com os dados dessa amostra por meio do teorema de Bayes. Seja $f(x|\theta)$ a função de verossimilhança associada à variável aleatória X condicionada a θ . Assim, define-se como distribuição a posteriori a expressão:

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{f(x)} \quad (2)$$

onde $f(x)$ é a função de probabilidade marginal dos dados e é definida como:

- Para um espaço paramétrico discreto de θ :

$$f(x) = \sum_{\theta} f(x|\theta)\pi(\theta) \quad (3)$$

- Para um espaço paramétrico contínuo de θ :

$$f(x) = \int f(x|\theta)\pi(\theta)d\theta \quad (4)$$

A distribuição a posteriori, $\pi(\theta|x)$, combina a informação inicial sobre o parâmetro, representada pela distribuição a priori $\pi(\theta)$, com a informação obtida dos dados, expressa pela função de verossimilhança $f(x|\theta)$. Na figura 5, é possível visualizar essa representação.

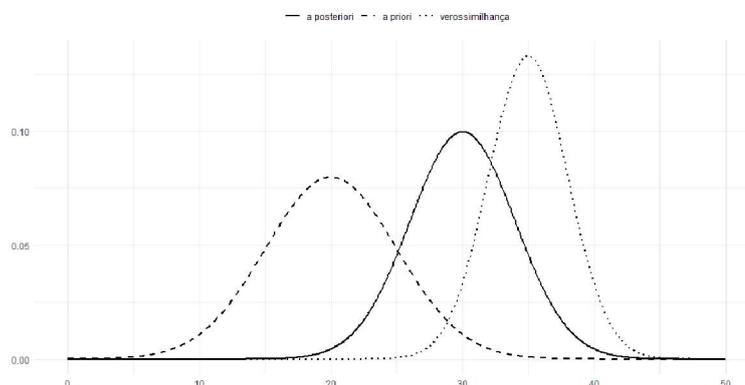
Antes de realizar o experimento, todas as informações disponíveis sobre o parâmetro são extraídas da priori. Após a coleta dos dados, essas informações são atualizadas com os resultados da amostra, resultando na distribuição a posteriori, que reflete nosso conhecimento atualizado sobre θ .

A distribuição a posteriori fornece todas as informações necessárias para realizar inferências sobre o parâmetro θ . No entanto, por ser uma fórmula matemática, é conveniente utilizar medidas resumo para sintetizar e comunicar essas informações de maneira clara e acessível. Entre as principais medidas resumo estão: média, mediana, moda, além de medidas de dispersão e assimetria. Essas medidas permitem captar as características essenciais da distribuição a posteriori. As inferências realizadas com base nessas medidas estão detalhadas em O'Hagan (1994).

Além disso, assim como na inferência clássica, a abordagem bayesiana permite realizar testes de hipóteses e selecionar os modelos mais adequados para resolver o problema em questão, oferecendo flexibilidade e precisão na análise dos dados.

O Teorema de Bayes é a base teórica da Inferência Bayesiana, fornecendo o fundamento para a regra de atualização que quantifica o aumento de informação. Esse teorema permite que, partindo de uma distribuição de probabilidade inicial, chamada de priori $\pi(\theta)$, seja possível obter a distribuição a posteriori $p(\theta|x)$, que reflete o grau de crença no parâmetro θ após observar os dados x . A função de verossimilhança desempenha um papel crucial nessa fórmula, pois é através dela que os dados observados são incorporados à distribuição a priori, atualizando-a e resultando na distribuição a posteriori. Isso faz com que a Inferência Bayesiana se adapte às novas informações de forma contínua e precisa.

FIGURA 5 – Representação da Inferência Bayesiana



FONTE: A autora (2025)

3.1.2.1 Regra de Bayes

Ao invés de considerar θ como um parâmetro fixo, pode-se interpretá-lo como uma variável aleatória. Nesse contexto, a distribuição de probabilidade atribuída a θ

resume todas as informações prévias disponíveis sobre ele, independentemente dos dados y . Essa distribuição é denominada distribuição a priori.

Assim como a verossimilhança (*likelihood*) depende dos parâmetros θ , a distribuição a priori também é parametrizada frequentemente por hiperparâmetros η . Esses hiperparâmetros são usualmente assumidos como conhecidos, e a distribuição a priori pode ser escrita como $\pi(\theta) \equiv \pi(\theta|\eta)$. A inferência sobre θ é então baseada na distribuição a posteriori, definida como:

$$p(\theta|y) = \frac{p(y|\theta)}{p(y)} = \frac{p(y|\theta)}{\int p(y, \theta) d\theta} = \frac{p(y|\theta)\pi(\theta)}{\int p(y, \theta)\pi(\theta) d\theta} \quad (5)$$

Essa equação, conhecida como o Teorema de Bayes, mostra que a distribuição a posteriori é o produto da verossimilhança $p(y|\theta)$ e da distribuição a priori $\pi(\theta)$, devidamente normalizado para garantir que a posteriori seja uma distribuição de probabilidade válida. A equação acima apresenta a versão contínua do Teorema de Bayes, enquanto a versão discreta, mais familiar em estatística, é expressa como:

$$P(B_j|A) = \frac{P(A|B_j)}{P(A)} = \frac{P(A|B_j)}{\sum_j P(A|B_j)} = \frac{P(A|B_j)P(B_j)}{\sum_j P(A|B_j)P(B_j)} \quad (6)$$

Um dos principais desafios na obtenção da distribuição a posteriori é o cálculo da integral presente no denominador. Em certos casos, essa integral pode ser resolvida de forma analítica, resultando em uma distribuição a posteriori com solução fechada (CARLIN; LOUIS, 2008). Entretanto, quando isso não é possível, métodos numéricos ou aproximações são frequentemente necessários.

3.1.2.2 Escolha da Priori

A informação a priori que se deseja incorporar na análise refere-se ao conhecimento prévio disponível sobre o tema, frequentemente oriundo de especialistas na área, e pode conter elementos subjetivos. Esses elementos podem também ser baseados em fontes objetivas, como dados históricos ou problemas semelhantes. No entanto, em diversas situações, o conhecimento de especialistas pode não estar disponível, o que leva à utilização de outros tipos de distribuições a priori. Um método alternativo é o uso de priori não informativas, nas quais a informação prévia tem menor relevância para a análise da distribuição. Esse tipo de priori assume que todos os possíveis valores do parâmetro em questão são igualmente prováveis, ou seja, têm distribuição Uniforme.

A distribuição a priori desempenha um papel crucial na Inferência Bayesiana, conforme destacado por DeGroot e Schervish (2012). Sem a definição de uma distribuição a priori adequada, não é possível calcular a distribuição a posteriori, o que compromete a realização da análise bayesiana. A priori serve como ponto de partida para a atualização das crenças com base nos dados observados, sendo essencial para o processo de inferência.

Ao analisar uma população, o pesquisador pode possuir algum nível de conhecimento prévio sobre o parâmetro de interesse θ . A Inferência Bayesiana permite que essa informação seja incorporada de forma sistemática na análise estatística. Esse conhecimento inicial é representado por meio de uma função ou densidade de probabilidade $\pi(\theta)$, conhecida como distribuição a priori.

Na abordagem Bayesiana, a estimativa dos parâmetros é baseada em dois elementos principais:

- a) A distribuição amostral $f(y|\theta)$, que descreve o comportamento dos dados θ , sendo utilizada para definir a função de verossimilhança $L(\theta|y)=f(y_1|\theta)\cdot\dots\cdot f(y_n|\theta)$.
- b) A distribuição a priori $\pi(\theta)$, que reflete as informações prévias sobre o parâmetro antes da observação dos dados.

A partir desses dois componentes, a informação sobre θ é atualizada após a observação dos dados y , resultando na distribuição a posteriori, $\pi(\theta|y)$, expressa como:

$$\pi(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\theta)\pi(\theta)d(\theta)} \quad (7)$$

Nessa fórmula, o denominador corresponde a uma constante em relação a θ e assegura que a posteriori seja uma distribuição de probabilidade válida.

A distribuição a posteriori também pode ser reescrita em termos da verossimilhança como:

$$\pi(\theta|y) = \frac{L(\theta|y)\pi(\theta)}{h(y)} \quad (8)$$

$h(y)$ é denominada constante de normalização e representa a densidade marginal dos dados y . Essa constante é calculada da seguinte maneira:

$$h(y) = \int \pi(y, \theta)d\theta = \int L(\theta|y)\pi(\theta)d\theta. \quad (9)$$

Essa formulação evidencia como a Inferência Bayesiana utiliza tanto os dados observados quanto o conhecimento prévio para atualizar as informações sobre o parâmetro, resultando em uma distribuição posteriori que reflete o estado mais atualizado do conhecimento.

A escolha da distribuição a priori é subjetiva e foi realizada ao longo do desenvolvimento desse projeto, buscando otimizar o desempenho da análise e alcançar os melhores resultados possíveis a partir dos dados disponíveis.

3.1.2.2.1 Prioris subjetivas

Existem situações em que informações sobre os parâmetros do modelo permitem definir uma forma específica para a priori. No entanto, conforme destacado por Paulino, Turkman e Murteira (2003), há uma dificuldade recorrente no campo científico: a obtenção e quantificação dessas informações de maneira que possam ser efetivamente utilizadas na inferência bayesiana. Um recurso valioso para auxiliar na construção da priori são as medidas resumo. Por exemplo, consultar especialistas para obter estimativas sobre a média da distribuição a priori ou outras características pode ser crucial para determinar uma distribuição adequada. O'Hagan (1994) ilustra um caso em que as informações disponíveis incluem a média e o desvio-padrão de θ .

Nesse cenário, uma escolha comum seria a distribuição Normal, quando θ pode assumir qualquer valor, uma distribuição Gama se θ for restrito a valores positivos, ou ainda uma distribuição Uniforme quando não se tem conhecimento nenhum.

a) Método do histograma

Esse método consiste em dividir o espaço paramétrico Θ em k intervalos $\Theta = \cup_{j=1}^k \Theta_j$ e solicitar ao especialista que estime a probabilidade de θ estar em cada intervalo. Com essas informações, é possível construir um histograma que serve como base para estimar uma densidade de probabilidade apropriada. Essa abordagem fornece uma aproximação da distribuição que pode indicar uma possível escolha para a priori (Paulino, Turkman e Murteira, 2003).

b) Método preditivo de eliciação

Em geral, especialistas possuem um entendimento detalhado sobre o processo em análise, mas podem ter dificuldade em fornecer informações diretas sobre o

parâmetro θ . Por isso, é mais intuitivo para o especialista refletir sobre os resultados do processo ou sobre medidas resumo das observações do que diretamente sobre θ . Esse procedimento é conhecido como método preditivo de eliciação.

Nesse método, se propõe um modelo probabilístico ($f(x|\theta)$) para o processo e se formula perguntas ao especialista sobre as observações ou sobre estatísticas resumo dessas observações. A distribuição usada como referência nas questões é a marginal em relação ao parâmetro θ , definida como:

$$f(x) = \int_{\Theta} f(x|\theta)\pi(\theta)d\theta \quad (10)$$

$\pi(\theta)$ é a priori que se deseja determinar, é necessário empregar métodos que ajustem essa distribuição de maneira que a integral reflita corretamente o conhecimento do especialista sobre $f(x)$. Paulino, Turkman e Murteira (2003) sugerem diversos métodos para lidar com essa questão, incluindo o uso de prioris pertencentes à família conjugada natural. Nesse caso, o desafio é reduzido à estimação dos hiperparâmetros da priori, que são os parâmetros que caracterizam a distribuição a priori.

c) Família conjugada

O cálculo da posteriori pode ser complexo em muitos casos, especialmente quando a priori escolhida resulta em distribuições que não possuem soluções conhecidas. Para esses cenários, é possível recorrer a métodos computacionais. Outra alternativa prática é o uso de prioris conjugadas, que têm a propriedade de pertencerem à mesma classe de distribuições que a posteriori. Assim, a atualização do conhecimento ocorre apenas pela modificação dos hiperparâmetros (O'Hagan, 1994).

Essa característica facilita significativamente o cálculo da posteriori e, conseqüentemente, a implementação da inferência bayesiana.

3.1.2.2.2 Prioris objetivas (não informativas)

Quando há pouca ou nenhuma informação a priori sobre um parâmetro, podem ser usadas distribuições a priori não informativas, que minimizam vieses e não favorecem valores específicos de θ (BERGER, 1985). Essas prioris são úteis para situações onde o conhecimento subjetivo é insuficiente, para comparar inferências

bayesianas com a teoria clássica ou avaliar o impacto de prioris subjetivas. Entre elas, destacam-se:

- a) Priori de Jeffreys: Baseada na matriz de informação de Fisher, a priori é proporcional à raiz quadrada do determinante da matriz de Fisher. Em alguns casos, como no exemplo de distribuições Binomial-Poisson, a priori resultante não é própria (não integrável).
- b) Priori de Máxima Entropia: Maximiza a incerteza (ou entropia) dentro das restrições conhecidas. Quando o espaço paramétrico é finito e discreto, a distribuição uniforme maximiza a entropia e é considerada não informativa.
- c) Priori de Referência (Berger e Bernardo): Define padrões de referência para estatísticos e busca minimizar a influência de subjetividade nas inferências. Ela maximiza a informação mútua entre os dados observados e os parâmetros desconhecidos. Em casos unidimensionais, coincide com a priori de Jeffreys, que é, assim, um caso particular da priori de referência.

Essas técnicas são ferramentas essenciais em inferência bayesiana, especialmente em cenários onde informações objetivas são limitadas.

3.1.2.2.3 Misturas de prioris

Quando os dados são organizados em grupos com comportamentos distintos, as prioris convencionais podem não ser adequadas. Nesse contexto, as misturas de prioris permitem modelar essas diferenças. A priori é definida como uma combinação ponderada de distribuições individuais:

$$\pi(\theta) = \sum_{i=1}^k w_i \pi_i(\theta) \quad (11)$$

onde $k \geq 1$, $w_i > 0$, e $\sum_{i=1}^k w_i = 1$. As distribuições individuais $\pi_i(\theta)$ são densidades válidas ($\int \pi_i(\theta) d\theta = 1$). Quando os pesos w_i são desconhecidos, eles são atribuídos a uma distribuição a priori, como a Distribuição de Dirichlet com parâmetros $(\alpha_1, \alpha_2, \dots, \alpha_k)$, que assegura uma modelagem flexível das proporções entre os grupos.

Neste trabalho as prioris não são “não informativas” ou subjetivas. De fato, são informações bem estruturadas, que podem ser atualizadas para prever especificamente o grau de risco das operações de seguro.

3.2 DISTRIBUIÇÃO BETA

A distribuição Beta é uma função densidade de probabilidade contínua, e que pode ser definida no intervalo $(0,1)$. De um modo geral é caracterizada por dois parâmetros positivos de forma, comumente representados por a e b .

A forma mais comum dessa distribuição é conhecida como a distribuição Beta do primeiro tipo. No entanto, também há uma versão mais abrangente chamada de distribuição Beta do segundo tipo, que constitui uma generalização da distribuição Beta e que é definida com 4 parâmetros. Sua função densidade de probabilidade é apresentada por Johnson, Kotz e Balakrishnan (1995), conforme descrito na equação de probabilidade.

$$f(x; a, b, c, d) = \frac{1}{B(a, b)} \frac{(x - c)^{(a-1)}(d - x)^{(b-1)}}{(d - c)^{a+b+1}} I_{X(c,d)} \quad (12)$$

Sendo $a, b, c, d > 0$ e $B(a, b) = \int_0^1 x^{a-1} (1 - x)^{b-1} dx$ é a função Beta, que pode ser reescrita pela relação com a distribuição Gama, ou seja, $B(a, b) = \Gamma(a + b)/(\Gamma(a)\Gamma(b))$ onde $\Gamma(a)$ é a conhecida função Gama definida por: $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$ $a > 0$.

Se for feita a transformação $Y = (X - c)/(d - c)$, será obtida a formulação usual da distribuição Beta:

$$f(y; a, b) = \frac{1}{\beta(a, b)} y^{(a-1)}(1 - y)^{(b-1)} I(y)_{(0,1)} \quad (13)$$

Johnson, Kotz e Balakrishnan (1995) discutem a notável flexibilidade da distribuição Beta, atribuindo essa característica à sua capacidade de assumir diversas formas mediante diferentes configurações dos parâmetros a e b . Em virtude dessa adaptabilidade, a distribuição Beta é amplamente recomendada para modelar dados contínuos no intervalo $(0,1)$.

Conforme observado por Oliveira (2011), a distribuição Beta pode facilmente se transformar em distribuições simétricas, assimétricas ou constantes, apenas ajustando os valores de seus parâmetros.

É comum introduzir a distribuição Beta por meio da relação entre a função beta e a função gama, onde $\beta(a, b) = \Gamma(a)\Gamma(b)/(\Gamma(a + b))$. Para este trabalho,

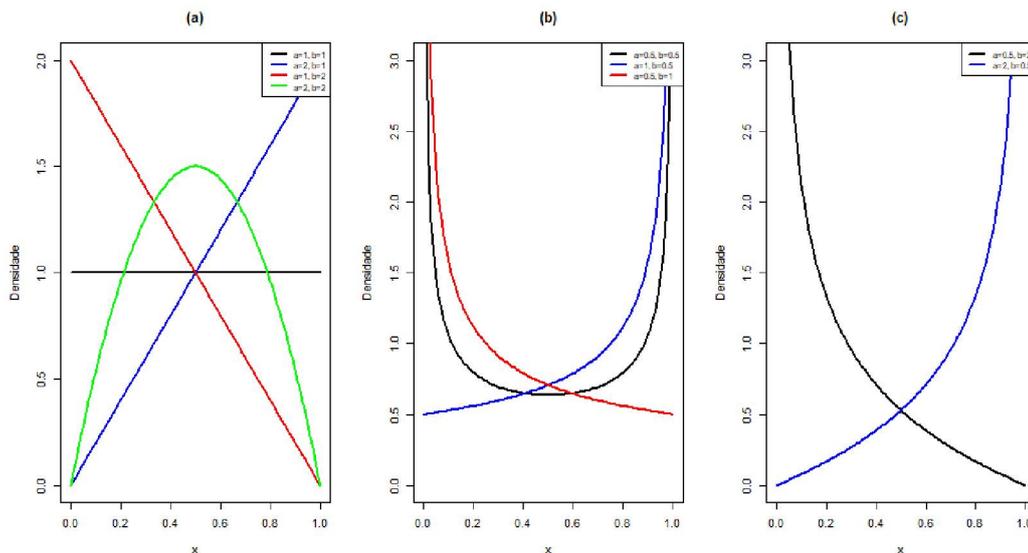
denotaremos a e b como α e β . Dessa forma, a densidade da distribuição Beta fica expressa pela seguinte equação:

$$f(\theta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1} \text{ para } 0 \leq \theta \leq 1 \quad (14)$$

em que α e β são maiores que zero e $\Gamma(\cdot)$ denota a função gama. Neste estudo, sempre que a distribuição Beta for referenciada, será conforme a formulação mencionada na equação 14.

Na Figura 6, são apresentadas curvas da distribuição Beta com diferentes configurações de parâmetros. Na Figura (a), observa-se que, quando $a = 1$ e $b = 1$, a curva corresponde a uma linha reta horizontal; para $a = 2$ e $b = 1$, a curva é crescente e decrescente; para $a = 1$ e $b = 2$, e quando $a = 2$ e $b = 2$, forma-se uma parábola com ponto máximo em $y = 1/2$. Já na Figura (b), quando $a < 1$ e $b < 1$, a curva adquire a forma de "U"; quando $a \geq 1$ e $b < 1$, assume a forma de "J"; e quando $a < 1$ e $b \geq 1$, tem o formato de "J" invertido. Além disso, na Figura (c), verifica-se que, se $a < b$, a curva é assimétrica à esquerda (positiva), enquanto que, se $b < a$, ela se torna assimétrica à direita (negativa).

FIGURA 6 – Curvas da distribuição Beta



FONTE: A autora (2025)

4 MATERIAL E MÉTODOS

4.1 MATERIAL

A avaliação do risco do seguro agrícola foi construída usando as frequências de perdas previstas pelo ZARC e as recomendações de semeadura para as culturas de soja e trigo, ao longo de um período de cinco safras que abrangem de 2017/2018 a 2021/2022.

Para garantir uma análise robusta, analisou-se um período de cinco anos de dados de seguros, disponíveis publicamente através da plataforma "dados abertos" mantida pelo governo federal do Brasil, disponibilizados no site oficial do Ministério da Agricultura, Pecuária e Abastecimento. Essa abordagem está alinhada com o período adotado pelas seguradoras em suas projeções de mercado, permitindo-nos avaliar de forma abrangente o desempenho da apólice de seguro. O detalhamento da base de dados utilizada para o seguro agrícola está disponível no Anexo 1 deste trabalho.

Esses dados forneceram informações cruciais sobre o perfil do solo de cada estado, bem como sobre as informações de seguro utilizadas, o que permitiria tomar decisões fundamentadas em dados reais.

No entanto, como os dados exatos sobre o tipo do solo não estavam disponíveis em sua totalidade, houve a necessidade de se basear em um limiar estimado. Para isso, foi utilizado um modelo baseado no método bayesiano, no qual uma apreciação subjetiva foi incorporada ao processo de análise. Diante dessa limitação, usou-se grupos de maturação, representados por k e tipos de solo, representados por j , com o pressuposto de que a distribuição exata dos solos não era totalmente conhecida. Um critério razoável foi adotar uma média que considerasse os extremos de risco: uma para o pior cenário e outra para o melhor.

Neste trabalho utilizou-se o Programa Estatístico R e Planilha de Cálculo EXCEL para se trabalhar com os dados.

4.2 METÓDOS

O ZARC (zoneamento agrícola de risco climático) estima probabilidades (z) como uma soma de eventos desfavoráveis u dividida pela duração do ciclo da cultura (codificada como janelas de 10 dias) e média ao longo da série histórica de 30 anos.

Esses cálculos baseados em simulações são padrões para todos os sistemas de cultivo e municípios incluídos no ZARC e podem ser resumidos da seguinte forma:

$$z_{ij} = p(u) \quad (15)$$

onde z_{ij} são probabilidades simuladas para três texturas de solo j e variando o período de 10 dias i em que as culturas são semeadas. As texturas do solo são consideradas representações dos tipos de solo e são definidas da seguinte forma:

- $j = 1$ quando o teor de argila é maior que 10% e menor que 15%.
- $j = 2$ quando o teor de argila está entre 15% e 35% e o teor de areia é inferior a 70%.
- $j = 3$ quando o teor de argila é superior a 35%.

O ZARC rotula as classes de risco usando as probabilidades máximas dos limiares. As estimativas relatadas \hat{z}_{ij} são de 20% para $\hat{z}_{ij} \leq 0.2$, 30% para $0.2 < \hat{z}_{ij} \leq 0.3$, 40% para $0.3 < \hat{z}_{ij} \leq 0.4$, e "-" quando $\hat{z}_{ij} > 0.4$.

O ZARC deixa explícito para os produtores que a semeadura não é recomendada nos períodos de 10 dias marcados com um sinal de "-". Semeadura de culturas em todos os outros períodos de 10 dias sugere que as perdas de produção ocorrem com frequências máximas de 20%, 30% ou 40%.

Reformulou-se as estimativas fornecidas pelo ZARC (\hat{z}_{ij}) para desenvolver uma distribuição a priori das perdas de produção. O processo de transformação dos dados do ZARC pode ser descrito em três etapas:

1. Convertendo as frequências máximas \hat{z}_{ij} em frequências esperadas de perda. Uma simples transformação nos leva a \bar{z}_{ij} com valores contidos no conjunto $\bar{\mathbb{Z}}$, onde $\bar{\mathbb{Z}} = \{0, 1; 0,25; 0,35, "-"\}$.
2. Realizou-se o cálculo da média simples das frequências de perda de produção transformadas \bar{z}_{ij} ao longo de todas as janelas de semeadura possíveis. Esse cálculo visa fornecer uma estimativa direta da frequência média esperada de perda, considerando as diversas janelas de semeadura sem aplicar pesos baseados na área semeada. A fórmula final resultante será uma média aritmética de \bar{z}_{ij} , conforme o ZARC. Em resumo:

$$E(\theta) = \frac{1}{n} \sum_{i=1}^n \bar{z}_{ij} \quad (16)$$

3. Ao se multiplicar o termo resultante $E(\theta)$ pelo número de períodos de 10 dias (n) com recomendações positivas de semeadura, chegou-se ao número estimado de eventos com perda de produção ao longo de todas as janelas de semeadura possíveis, representado por α_{kjm} . Em notação matemática, o número esperado de tentativas de produção agrícola com perda α_{kjm} e o número esperado de tentativas de produção bem-sucedidas (sem perda) tornam-se:

$$\alpha_{kjm} = E(\theta).n \quad (17)$$

$$\beta_{kjm} = n - \alpha_{kjm} \quad (18)$$

As estimativas de α e β foram normalizadas em uma base percentual para garantir que a distribuição siga as mesmas proporções utilizadas no ZARC. As estimativas, bem como o comportamento naturalmente limitado de Y , com $0 \leq Y \leq 1$, tornam-se parâmetros convenientes para estudar a frequência de perdas de produção sob a perspectiva da literatura de finanças e seguros. Utilizou-se α e β como parâmetros para ajustar uma distribuição beta, amplamente usada para modelar perdas de seguros. A distribuição beta tem a vantagem de ser flexível e assumir várias formas à medida que α e β variam.

Em nossa aplicação, se $Y \sim \text{Beta}(\alpha, \beta)$, então a probabilidade de uma tentativa de produção Y sofrer uma perda segue a função de densidade de probabilidade, onde $f(Y)$ se refere à função de densidade de probabilidade das frequências de perda derivadas do ZARC. Em nossa aplicação, por exemplo, $f(Y > 0,1) = 0,5$ significa que metade das tentativas de produção provavelmente experimentará uma perda a cada dez anos de cultivo.

Dependendo dos objetivos analíticos ou da granularidade de interesse, analistas que conduzem pesquisas para companhias de seguros ou formuladores de políticas podem aplicar os passos 1 a 3 e agregar ou desagregar os resultados conforme necessário. Os parâmetros α e β podem ser calculados para a média estadual somando as estimativas de todos os municípios, representando por m e dividindo pelo número de municípios, representando por M . As equações do item 4.2.1 resumem o procedimento de agregação. Para os propósitos desta aplicação, agregou-se α e β sobre grupos de maturação e tipos de solo, resultando em A_k e B_k . Características específicas, como grupos de maturação e tipos de solo, também estão

ausentes nos contratos de seguros, então a agregação desses parâmetros não compromete o esforço de combinar projeções de perdas do zoneamento agrícola com os dados de sinistros de seguros.

A distribuição Beta é bastante flexível, assumindo várias formas dependendo dos valores dos parâmetros α e β . Ela é adequada para modelar probabilidades, já que seu suporte está restrito ao intervalo $[0,1]$, o que é útil para modelar frequências de perda de produção.

A equação 19 descreve a função de densidade de probabilidade (fdp) da distribuição Beta, que é utilizada para modelar o risco de perda de produção no contexto do ZARC. A equação geral da distribuição Beta é:

$$f(\theta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1 - \theta)^{\beta-1} \text{ para } 0 \leq \theta \leq 1 \quad (19)$$

- α e β são parâmetros da distribuição que foram normalizados para seguir as proporções do ZARC, foram multiplicados por 100;
- $\Gamma(\cdot)$ é a função Gamma;
- θ é a variável aleatória que representa a frequência de perda de produção.

4.2.1 Cálculo para a média estadual

Os parâmetros α e β podem ser calculados para a média de um estado somando as estimativas de todos os municípios, representado por m e dividindo pelo número de municípios, representado por M .

$$A_k = \frac{1}{M} \sum_{m=1}^M \alpha_{km} \quad (20)$$

$$\beta_k = \frac{1}{M} \sum_{m=1}^M \beta_{km} \quad (21)$$

4.2.2 Cenários de Análise

a) Melhor cenário (Solo Argiloso)

Para o melhor cenário, que considera o solo argiloso, as fórmulas utilizadas foram:

$$A_{k, \text{argiloso}} = \frac{1}{M} \sum_{m=1}^M \alpha_{k, \text{argiloso}, m} \quad (22)$$

$$\beta_{k, \text{argiloso}} = \frac{1}{M} \sum_{m=1}^M \beta_{k, \text{argiloso}, m} \quad (23)$$

b) Pior cenário (Solo Arenoso)

No pior cenário, que considera o solo arenoso, as fórmulas utilizadas foram:

$$A_{k, \text{arenoso}} = \frac{1}{M} \sum_{m=1}^M \alpha_{k, \text{arenoso}, m} \quad (24)$$

$$\beta_{k, \text{arenoso}} = \frac{1}{M} \sum_{m=1}^M \beta_{k, \text{arenoso}, m} \quad (25)$$

Essas fórmulas representam os parâmetros agregados A_k e B_k para os três cenários: (1) geral, considerando todos os tipos de solo, (2) o melhor cenário, com solo argiloso, e (3) o pior cenário, com solo arenoso. Esses parâmetros foram utilizados para plotar as distribuições e fornecer uma visão mais detalhada do risco de perda de produção em diferentes condições de solo por Estado.

4.2.3 Dados de seguros e transformação

Os dados sobre apólices individuais estão disponíveis desde 2006. Neste trabalho, recuperou-se dados para contratos de seguro individuais emitidos para soja e trigo cultivados em todos os municípios do Paraná, Rio Grande do Sul e Santa Catarina durante as safras de 2017/2018 a 2021/2022. O conjunto de dados consiste em 21 variáveis para todos os 27 estados do Brasil. Focou-se em um subconjunto de quatro variáveis: município, cultura, valor da indenização e ano da apólice. O valor total das indenizações alcançou a expressiva quantia de R\$ 3.374.656.082,47, distribuída em 46.891 pagamentos, efetuados com base em 275.479 apólices contratadas nos três estados. A variável que relata os valores das indenizações é de interesse central. Quando o pagamento da indenização é maior que zero para um determinado contrato, atribuímos o valor de 1 e 0, caso contrário. Seja y o número de ocorrências de pagamentos de indenização em todos os contratos de seguro c emitidos para a cultura de interesse no município m , tal que $Y \in \{0, 1, 2, \dots, c\}$.

Considerando a origem do conjunto de dados e os mecanismos institucionais em vigor para monitorar e verificar a perda de produção agrícola, pode-se inferir plausivelmente que a ocorrência de pagamentos de indenização é condicional à ocorrência de perda. Portanto, a probabilidade de y condicional à ocorrência de perda de produção, $P(y|\theta)$, segue uma distribuição binomial com parâmetros c e θ . Em notação estatística:

$$P(y|\theta) = (cy)\theta^y(1 - \theta)^{c-y} \quad y = 0, 1, 2, \dots, c \quad (26)$$

Embora a probabilidade condicional de pagamentos de indenização seja válida para todas as culturas de interesse em todos os municípios, representados por m , o argumento de agregação apresentado acima também é aplicável. Analistas e formuladores de políticas podem considerar útil examinar a probabilidade condicional de indenizações a nível estadual, por exemplo.

4.2.4 Aplicação do método de atualização Bayesiana e a distribuição posterior

Sabendo que a distribuição a priori das frequências de perdas de produção, derivadas do ZARC, segue uma distribuição Beta (α, β) e que os dados transformados sobre a ocorrência de pagamentos de indenização Y dão origem a uma probabilidade condicional que segue uma distribuição binomial (n, θ) podemos aplicar o método de atualização bayesiana para obter a distribuição preditiva a posteriori das perdas de produção. O cálculo da distribuição a posteriori segue o procedimento demonstrado por Hoff. (2009, pp. 35-38):

$$\begin{aligned} P(y) &= \frac{P(\theta)P(y|\theta)}{P(y)} \\ &= \frac{1}{P(y)} \cdot \frac{(\alpha+\beta)}{(\alpha)(\beta)} \theta^{\alpha-1}(1 - \theta)^{\beta-1} x \binom{c}{y} \theta^y(1 - \theta)^{c-y} \\ &= k(c, y, \alpha, \beta) \cdot \theta^{\alpha+y-1}(1 - \theta)^{\beta+c-y-1} \\ &= \text{beta}(\alpha + y, \beta + c - y) \end{aligned} \quad (27)$$

A distribuição Beta a posteriori torna-se uma combinação da distribuição a priori e dos dados transformados de sinistros de seguros, resultando em momentos brutos facilmente reconhecíveis. Nas equações 28 e 29, é possível identificar as essas expressões da esperança e variância da distribuição Beta a posteriori.

$$E(\theta|y) = \frac{\alpha + y}{\alpha + \beta + c} \quad (28)$$

$$Var(\theta|y) = \frac{(\theta|y)E(1 - \theta|y)}{\alpha + \beta + c + 1} \quad (29)$$

4.2.5 Aplicação empírica e validação do modelo

Cálculos empíricos foram realizados no software estatístico R. Foi construído um sistema de equações, conforme 4.2.2, para simular cenários alternativos para diferentes culturas e municípios. Foi utilizada a classificação desenvolvida por Martins e Signorini (2024) para validar a metodologia e comparar o poder de previsibilidade da distribuição posterior contra as frequências de perda de produção derivadas do sistema de zoneamento agrícola ZARC.

Embora a distribuição posterior resultante tenha utilizado uma priori beta com parâmetros agregados sobre características específicas de culturas e tipos de solo, nossa avaliação comparativa empírica requer a preparação de duas distribuições beta adicionais. A primeira seguiu os passos procedimentais 1 a 3 listados em 3.2 acima, mas manteve solos arenosos, para o tipo de solo em vez de unificar o parâmetro. A segunda distribuição beta utilizou solos argilosos, para o tipo de solo. Assumimos que a distribuição beta de solos arenosos representa o pior cenário de produção para grãos, enquanto a distribuição beta de solos argilosos corresponde ao melhor cenário, conforme 3.2.3 e 3.2.2. As razões subjacentes para essa suposição são duas. Primeiro, solos argilosos e arenosos estão nos dois extremos de um espectro contínuo de capacidade de retenção de água, onde solos argilosos retêm mais água e atendem às necessidades hídricas das culturas por mais tempo do que solos arenosos. Assim, pode-se argumentar plausivelmente que solos argilosos tendem a reduzir a exposição das culturas ao estresse hídrico em comparação com solos arenosos. Também estende as janelas de plantio em comparação com solos arenosos. Essas são tendências comuns para todos os municípios ou culturas consideradas neste trabalho. Na verdade, as razões fornecidas para a determinação dos limites de risco em função dos tipos de solo convergem à medida que os cálculos do ISNA levam às classes de risco do ZARC z_{ij} e se baseiam em características técnicas de solos e culturas.

Os critérios de classificação compreendem três categorias amplas e sete subcategorias com base na comparação direta de valores esperados e distribuições