

**UNIVERSIDADE FEDERAL DE VIÇOSA**

**Duração da proteção de cultivares: uma abordagem da análise de sobrevivência com riscos competitivos**

Verônica Manhães Saint'Clair  
*Magister Scientiae*

**VIÇOSA - MINAS GERAIS  
2025**

**VERÔNICA MANHÃES SAINT'CLAIR**

**Duração da proteção de cultivares: uma abordagem da análise de sobrevivência com riscos competitivos**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

Orientador: Sebastiao Martins Filho

Coorientadora: Flávia Sílvia Corrêa Tomaz

**VIÇOSA - MINAS GERAIS  
2025**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade  
Federal de Viçosa - Campus Viçosa**

T

S155d Saint'Clair, Verônica Manhães, 1998-  
2025 Duração da proteção de cultivares: uma abordagem da  
análise de sobrevivência com riscos competitivos / Verônica  
Manhães Saint'Clair. – Viçosa, MG, 2025.  
1 dissertação eletrônica (83 f.): il. (algumas color.).

Inclui apêndices.

Orientador: Sebastião Martins Filho.

Dissertação (mestrado) - Universidade Federal de Viçosa,  
Departamento de Estatística, 2025.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2025.174>

Modo de acesso: World Wide Web.

1. Análise de sobrevivência (Biometria). I. Martins Filho,  
Sebastião, 1961-. II. Universidade Federal de Viçosa.  
Departamento de Estatística. Programa de Pós-Graduação em  
Estatística Aplicada e Biometria. III. Título.

CDD 22. ed. 519.546

**VERÔNICA MANHÃES SAINT'CLAIR**

**Duração da proteção de cultivares: uma abordagem da análise de sobrevivência com riscos competitivos**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

APROVADA: 18 de fevereiro de 2025.

Assentimento:

---

Verônica Manhães Saint'Clair  
Autora

---

Sebastiao Martins Filho  
Orientador

Essa dissertação foi assinada digitalmente pela autora em 03/04/2025 às 16:05:17 e pelo orientador em 03/04/2025 às 16:22:11. As assinaturas têm validade legal, conforme o disposto na Medida Provisória 2.200-2/2001 e na Resolução nº 37/2012 do CONARQ. Para conferir a autenticidade, acesse <https://siadoc.ufv.br/validar-documento>. No campo 'Código de registro', informe o código **FHQC.B4CL.GNX1** e clique no botão 'Validar documento'.

## AGRADECIMENTOS

A Deus, acima de tudo.

À minha mãe, Sandra, por sua dedicação à minha educação e pelo incentivo constante.

Ao meu pai, José (*in memoriam*), cujos ensinamentos preservo em minha conduta.

Aos meus irmãos, Ivan e Marina, pelo apoio em cada etapa.

Aos queridos amigos que fiz durante graduação, por me incentivarem ao longo da caminhada.

Ao professor Paulo César Beggio, que despertou meu interesse pela Estatística durante a graduação na Universidade Estadual do Norte Fluminense Darcy Ribeiro e me encorajou a seguir além da graduação.

Aos colegas de curso, pela acolhida, parceria nos estudos e momentos de descontração. Em especial aos amigos Beatriz, Lucas, Maurício e Renata, por cada palavra de motivação.

Ao professor Sebastião Martins Filho, meu orientador, expressei minha gratidão por sua orientação, incentivo e presença constante. Sua organização, dedicação e ensinamentos foram fundamentais para a realização deste trabalho.

Ao doutorando Marciel, pelas recomendações e disposição sempre que necessário.

Ao Departamento de Estatística, ao secretário Junior e aos professores que contribuíram para minha formação. Em especial, ao professor Paulo Roberto Cecon, pelo incentivo para a conclusão desta dissertação.

À minha coorientadora, Flávia Sílvia Corrêa Tomaz, e ao professor Vinícius Silva dos Santos, membros da banca examinadora, pelo tempo dedicado e pelas contribuições valiosas.

À Universidade Federal de Viçosa, pela oportunidade de realizar a pós-graduação.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

A todos que, de alguma forma, contribuíram para a conclusão desta etapa.

*“Cada um de nós compõe a sua história  
E cada ser em si carrega o dom de ser capaz e ser feliz”.*  
(Almir Sater e Renato Teixeira)

## RESUMO

SAINT'CLAIR, Verônica Manhães, M.Sc., Universidade Federal de Viçosa, fevereiro de 2025. **Duração da proteção de cultivares: uma abordagem da análise de sobrevivência com riscos competitivos.** Orientador: Sebastiao Martins Filho. Coorientadora: Flávia Sílvia Corrêa Tomaz.

Em diversas pesquisas, a investigação incide sobre o tempo até a ocorrência de um evento de interesse, motivando o uso da técnica estatística da análise de sobrevivência. Nos dados de sobrevivência clássica, presume-se que os indivíduos experimentem apenas um tipo de evento durante o acompanhamento, sendo empregados métodos como Kaplan-Meier e os modelos paramétricos e semiparamétricos, quando há o interesse em avaliar os efeitos de covariáveis. No entanto, em certas situações, os indivíduos podem vivenciar outros eventos que competem com o evento de interesse, capazes de impedir ou alterar a probabilidade de sua ocorrência, conhecidos como eventos competitivos. A abordagem clássica, nessas circunstâncias, trata esses eventos competitivos como censura, resultando em estimativas viesadas. Surge então, a necessidade de uma abordagem com riscos competitivos, que introduz a função de incidência acumulada, a função de risco causa-específica e a função de risco da subdistribuição. O objetivo deste trabalho foi apresentar os conceitos e aplicações relacionados à análise de sobrevivência em cenários que envolvem riscos competitivos nos contextos de tempo contínuo e discreto. Para isso, foram analisados dados de certificados de proteção de cultivares obtidos da plataforma CultivarWeb do Ministério da Agricultura, Pecuária e Abastecimento, no período de 1997 a 2024. Os eventos de expiração por prazo, renúncia, cancelamento e anulação foram analisados utilizando uma abordagem de tempo contínuo por meio do modelo de Fine-Gray e com uma abordagem de tempo discreto com o modelo de regressão binária ponderada devido à presença de empates. As covariáveis utilizadas incluem informações sobre a presença de transgenia, a característica do titular, o tipo de cultura e o ciclo de vida. Os resultados indicaram que o modelo de Fine-Gray é o mais adequado para a modelagem da duração dos certificados de proteção de cultivares.

Palavras-chave: Função de incidência acumulada; Subdistribuição; Causa-específica; Modelo de Fine-Gray; Modelo de regressão binária ponderada; Dados censurados

## ABSTRACT

SAINT'CLAIR, Verônica Manhães, M.Sc., Universidade Federal de Viçosa, February, 2025. **Duration of Cultivar Protection: An Approach Using Survival Analysis with Competing Risks**. Adviser: Sebastiao Martins Filho. Co-adviser: Flávia Sílvia Corrêa Tomaz.

In several studies, the investigation focuses on the time until the occurrence of an event of interest, which is why the statistical technique of survival analysis is used. In classic survival data, it is assumed that individuals experience only one type of event during follow-up, and methods such as Kaplan-Meier and parametric and semi-parametric models are used when there is an interest in assessing the effects of covariates. However, in certain situations, individuals may experience other events that compete with the event of interest, capable of preventing or altering the probability of its occurrence, known as competitive events. The classical approach, in these circumstances, treats these competing events as censorship, resulting in biased estimates. The need then arises for an approach with competing risks, which introduces the cumulative incidence function, the cause-specific risk function and the subdistribution risk function. The aim of this work was to present the concepts and applications related to survival analysis in scenarios involving competing risks in the contexts of continuous and discrete time. To this end, we analyzed data from cultivars protection certificates obtained from the CultivarWeb platform of the Ministério da Agricultura, Pecuária e Abastecimento, from 1997 to 2024. The events of expiration by term, resignation, cancellation and annulment were analyzed using a continuous time approach by means of the Fine-Gray model and with a discrete time approach with the weighted binary regression model due to the presence of ties. The covariates used include information on the presence of transgenics, the holder's characteristic, the type of crop and the life cycle. The results indicate that the Fine-Gray model is the most suitable for modeling the duration of plant variety protection certificates.

Keywords: Cumulative incidence function; Subdistribution; Cause-specific; Fine-Gray model; Weighted binary regression model; Censored data

## LISTA DE ILUSTRAÇÕES

**Figura 1** - Avaliação da proporcionalidade do modelo de Fine-Gray para o evento de expiração por prazo para cultivares de ciclo anual e perene: anual e com transgenia (A), anual e setor privado (B), anual e parceria (C), anual e forrageira (D), anual e olerícola (E), perene e florestal (F), perene e forrageira (G), perene e frutífera (H), perene e ornamental (I). ..... 40

**Figura 2** - Avaliação da proporcionalidade do modelo de Fine-Gray para o evento de renúncia para cultivares de ciclo anual e perene: anual e com transgenia (A), anual e setor privado (B), anual e parceria (C), anual e forrageira (D), anual e olerícola (E), anual e ornamental (F), perene e setor privado (G), perene e parceria (H), perene e florestal (I), perene e forrageira (J), perene e frutífera (K), perene e ornamental (L). ..... 41

**Figura 3** - Avaliação da proporcionalidade do modelo de Fine-Gray para os eventos de cancelamento e anulação para cultivares de ciclo anual e perene: anual e com transgenia (A), anual e setor privado (B), anual e parceria (C), perene e florestal (D), perene e forrageira (E), perene e frutífera (F), perene e olerícola (G), perene e ornamental (H). ..... 42

**Figura 4** - Curvas de incidência acumulada estimadas pelo método não paramétrico e pelo modelo de Fine-Gray para o evento de expiração por prazo: anual, instituição pública, sem transgenia e agrícola (A); anual, com transgenia, instituição pública e agrícola (B); anual, sem transgenia, instituição pública e olerícola (C); anual, sem transgenia, setor privado e agrícola (D); anual, com transgenia, setor privado e agrícola (E); anual, sem transgenia, setor privado e forrageira (F); anual, sem transgenia, setor privado e olerícola (G); perene e agrícola (H); perene e florestal (I); perene e forrageira (J); perene e frutífera (K) ; perene e ornamental (L). ..... 49

**Figura 5** - Curvas de incidência acumulada estimadas pelo método não paramétrico e pelo modelo de Fine-Gray para o evento de renúncia no ciclo anual: sem transgenia, instituição pública e agrícola (A); com transgenia, instituição pública e agrícola (B); sem transgenia, instituição pública e olerícola (C); sem transgenia, setor privado e agrícola (D); com transgenia, setor privado e agrícola (E); sem transgenia, setor privado e olerícola (F); sem transgenia, setor privado e ornamental (G); sem transgenia, parceria e agrícola (H); sem transgenia, parceria e olerícola (I). ..... 50

**Figura 6** - Curvas de incidência acumulada estimadas pelo método não paramétrico e pelo modelo de Fine-Gray para o evento de renúncia no ciclo perene: instituição pública e agrícola

(A), instituição pública e forrageira (B), instituição pública e frutífera (C), instituição pública e ornamental (D), setor privado e agrícola (E), setor privado e florestal (F), setor privado e forrageira (G), setor privado e frutífera (H), setor privado e ornamental (I), parceria e frutífera (J), parceria e ornamental (K)..... 51

**Figura 7** - Curvas de incidência acumulada estimadas pelo método não paramétrico e pelo modelo de Fine-Gray para os eventos de cancelamento e anulação nos ciclos anual e perene: anual, instituição pública e sem transgenia (A); anual, com transgenia e instituição pública (B); anual, sem transgenia e setor privado (C); anual, com transgenia e setor privado (D); anual, sem transgenia e parceria (E); perene e agrícola (F); perene e florestal (G); perene e forrageira (H); perene e frutífera (I); perene e olerícola (J); perene e ornamental (K)..... 52

**Figura 8** - Curvas ROC mostrando a área abaixo da curva (AUC), estimadas pelo modelo de regressão binária ponderada com ligação log-log complementar para os eventos de: expiração por prazo no ciclo anual (A); renúncia no ciclo anual (B); cancelamento e anulação no ciclo anual (C); expiração por prazo no ciclo perene (D); renúncia no ciclo perene (E); cancelamento e anulação no ciclo perene (F)..... 53

**Figura 9** - Curvas de incidência acumulada no ciclo anual associadas à covariável transgenia estimadas a partir do método não paramétrico para os eventos: expiração por prazo (A), renúncia (B), cancelamento e anulação (C)..... 60

**Figura 10** - Curvas de incidência acumulada no associadas à covariável titular estimadas a partir do método não paramétrico para: expiração por prazo no ciclo anual (A), renúncia no ciclo anual (B), cancelamento e anulação no ciclo anual (C), expiração por prazo no ciclo perene (D), renúncia no ciclo perene (E), cancelamento e anulação no ciclo perene (F). ..... 61

**Figura 11** - Curvas de incidência acumulada no associadas à covariável tipo de cultura estimadas a partir do método não paramétrico para: expiração por prazo no ciclo anual (A), renúncia no ciclo anual (B), cancelamento e anulação no ciclo anual (C), expiração por prazo no ciclo perene (D), renúncia no ciclo perene (E), cancelamento e anulação no ciclo perene (F). ..... 62

## LISTA DE TABELAS

<b>Tabela 1</b> - Resumo descritivo dos dados, conforme o tempo das proteções, definitiva, expiração por prazo, renunciadas e canceladas e/ou anuladas, considerando os ciclos de vida anual e perene para os eventos em estudo.....	39
<b>Tabela 2</b> - Valores estimados pelo modelo de Fine-Gray para todos os eventos associados a cultivares de ciclo anual, utilizando ausência de transgenia, instituição pública e tipo de cultura agrícola, como categorias de referência..	43
<b>Tabela 3</b> - Valores estimados pelo modelo de regressão binária ponderada com ligação log-log complementar para todos os eventos associados a cultivares de ciclo anual, utilizando ausência de transgenia, instituição pública e tipo de cultura agrícola, como categorias de referência...	44
<b>Tabela 4</b> - Valores estimados pelo modelo de Fine-Gray para todos os eventos associados a cultivares de ciclo perene, utilizando ausência de transgenia, instituição pública e tipo de cultura agrícola, como categorias de referência..	45
<b>Tabela 5</b> - Valores estimados pelo modelo de regressão binária ponderada com ligação log-log complementar para todos os eventos associados a cultivares de ciclo perene, utilizando ausência de transgenia, instituição pública e tipo de cultura agrícola, como categorias de referência...	46
<b>Tabela 6</b> - Proteções renunciadas no período de 2018-2022.....	47
<b>Tabela 7</b> - Critério de informação de Akaike (AIC) dos modelos de regressão avaliados para os dados referentes à duração das proteções de cultivares..	54
<b>Tabela 8</b> - Teste de Gray para a comparação das curvas de incidência acumulada no ciclo anual estimadas pela técnica não paramétrica.....	62
<b>Tabela 9</b> - Teste de Gray para a comparação das curvas de incidência acumulada no ciclo perene estimadas pela técnica não paramétrica..	63

## **LISTA DE SIGLAS E ABREVIATURAS**

- AIC Critério de Informação de Akaike.
- AUC Área sob a curva.
- BIC Critério de Informação Bayesiano.
- CAPES Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.
- COVID Coronavírus Disease 2019.
- FIA Função de incidência acumulada.
- MAPA Ministério da Agricultura, Pecuária e Abastecimento.
- ROC Receiver Operating Characteristic.
- SNPC Serviço Nacional de Proteção de Cultivares.

## SUMÁRIO

INTRODUÇÃO.....	12
CAPÍTULO 1: REFERENCIAL TEÓRICO.....	15
1.1. Proteção de cultivares.....	15
1.2. Análise de sobrevivência.....	16
1.2.1. Técnica não paramétrica.....	17
1.2.2. Técnica semiparamétrica.....	18
1.2.3. Técnica paramétrica.....	19
1.3. Riscos competitivos.....	20
1.4. Riscos competitivos quando o tempo é contínuo.....	21
1.4.1. Técnica não paramétrica.....	22
1.4.2. Técnicas semiparamétricas.....	24
1.5. Riscos competitivos quando o tempo é discreto.....	26
1.5.1. Técnica paramétrica.....	27
CAPÍTULO 2: ANÁLISE DE SOBREVIVÊNCIA COM RISCOS COMPETITIVOS APLICADA NA DURAÇÃO DA PROTEÇÃO DE CULTIVARES.....	34
Resumo.....	34
Introdução.....	34
Material e Métodos.....	36
Resultados e Discussão.....	38
Conclusões.....	54
Agradecimentos.....	54
Referências.....	55
APÊNDICES.....	58

## INTRODUÇÃO

Em diversas áreas de pesquisa, o tempo até a ocorrência de um determinado evento é uma variável de grande interesse a ser avaliada. No entanto, quando este tempo até o evento não pode ser observado para todos os indivíduos do estudo, isto é, quando ocorre a presença de censura, a análise desses dados exige o uso da técnica da análise de sobrevivência (Colosimo e Giolo, 2024).

A análise de sobrevivência já tem sido aplicada em vários estudos nas Ciências Agrárias. Romano e Stevanato (2020), por exemplo, utilizaram essa abordagem para estudar a germinação de sementes, considerando como censura aquelas que não germinaram até o final do período de observação. Genna et al. (2015) investigaram não apenas a germinação, mas também a infecção por patógenos, que impediam a germinação. Nesse caso, a censura foi considerada as sementes que não germinaram e nem foram infectadas ao final do estudo.

Tradicionalmente, a abordagem clássica de sobrevivência pressupõe que cada indivíduo esteja sujeito a apenas um tipo de evento ao longo do período de acompanhamento, como a morte, a remissão de uma doença ou a quebra de um equipamento. Entretanto, em muitos contextos, as unidades experimentais podem estar sujeitas a diferentes tipos de eventos que competem entre si, alterando ou impedindo a probabilidade de ocorrência do evento de interesse (Pintilie, 2006). Essa situação é conhecida como riscos competitivos.

Nas Ciências Agrárias, a análise com riscos competitivos tem sido utilizada para compreender como a ocorrência de um evento pode inviabilizar a ocorrência de outro. Por exemplo, Towe (2011) estudou a preservação de áreas agrícolas como evento de interesse e a conversão de terras para moradias como evento competitivo. Da mesma forma, Genna et al. (2015) analisaram a germinação como evento de interesse e a infecção por patógenos como evento competitivo. Enquanto Meng et al. (2020) examinaram a concessão de certificados de proteção de cultivares como evento de interesse e a renúncia a esses certificados como evento competitivo.

Uma distinção fundamental nesse tipo de análise é entre eventos com tempos contínuos (ex.: dias, meses) e discretos (ex.: ciclos, sessões). Embora, a maioria das referências bibliográficas concentra-se na análise de tempo contínuo, que costuma exigir suposições como a de riscos proporcionais, métodos de tempo discreto são mais adequados quando há muitos empates ou quando o tempo é naturalmente discreto (Nakano e Carrasco, 2006; Schmid e Berger, 2021).

Nesse contexto, o objetivo deste trabalho foi estudar e aplicar técnicas de análise de riscos competitivos em contextos de tempo contínuo e discreto. Para isso, foram analisados dados sobre certificados de proteção de cultivares obtidos da plataforma *CultivarWeb* do Ministério da Agricultura, Pecuária e Abastecimento, no período de 1997 a 2024. Os dados incluíram informações sobre 4966 registros de proteções, considerando o tempo de início e término (em meses), indicação de modificação genética, características do titular (instituição pública, setor privado e parceria), tipo de cultura (agrícola, forrageira, florestal, frutífera, oléícola, ornamental, medicinal e aromática) e ciclo de vida (anual e perene).

A expiração por término de prazo foi considerada o evento de interesse, enquanto a renúncia, o cancelamento e a anulação foram tratados como eventos competitivos, uma vez que impedem a ocorrência da expiração. Esses eventos foram analisados por meio de uma abordagem de tempo contínuo, utilizando o modelo de Fine-Gray (Fine e Gray, 1999), e de uma abordagem de tempo discreto, com o modelo de regressão binária ponderada (Berger et al., 2020), devido à presença de empates.

Este trabalho está dividido em dois capítulos, sendo o primeiro referente a uma revisão sobre análise de sobrevivência, destacando a abordagem de riscos competitivos em tempo contínuo e discreto. No segundo capítulo é apresentada a aplicação da análise de sobrevivência com riscos competitivos para investigar a duração da proteção de cultivares.

## Referências

BERGER, M.; SCHMID, M.; WELCHOWSKI, T.; SCHMITZ-VALCKENBERG, S.; BEYERSMANN, J. Subdistribution hazard models for competing risks in discrete time. **Biostatistics**, v.21, p.449–466, 2020. <https://doi.org/10.1093/biostatistics/kxy069>

COLOSIMO, E. A.; GIOLO, S. R. **Análise de sobrevivência aplicada**. 2. ed. São Paulo: Editora Edgard Blucher, 2024.

FINE, J. P.; GRAY, R. J. A proportional hazard model for subdistribution of a competing risk. **Journal of the American Statistical Association**, v. 94, n. 446, p. 496–509, 1999. <https://doi.org/10.1080/01621459.1999.10474144>

GENNA, N. G.; KANE, M. E.; PÉREZ, H. E. Simultaneous assessment of germination and infection dose-responses in fungicide-treated seeds with non- and semiparametric statistical methods. **Seed Science and Technology**, v. 43, n. 2, p. 168–186, 2015. <https://doi.org/10.1525/sst.2015.43.2.13>

MENG, T.; CAREW, R.; FLORKOWSKI, W. J. Determinants of the grant lag and the surrender lag of horticultural crop plant breeders' rights applications: Survival analysis with

competing risks. **Canadian Journal of Agricultural Economics/Revue canadienne d'agroéconomie**, v. 68, n. 4, p. 489–512, 2020. <https://doi.org/10.1111/cjag.12257>

NAKANO, E. Y.; CARRASCO, C. G. Uma Avaliação do Uso de um Modelo Contínuo na Análise de Dados Discretos de Sobrevivência. **Tendências em Matemática Aplicada e Computacional**, v. 7, n. 1, p. 91–100, 2006. <https://doi.org/10.5540/tema.2006.07.01.0091>

PINTILIE, M. **Competing risks: a practical perspective**. Chichester; Hoboken: John Wiley & Sons, 2006.

ROMANO, A.; STEVANATO, P. Germination Data Analysis by Time-to-Event Approaches. **Plants**, v.9, n. 5, p. 617, 2020. <https://doi.org/10.3390/plants9050617>

SCHMID, M.; BERGER, M. Competing risks analysis for discrete time-to-event data. **WIREs Computational Statistics**, v. 13, n. 5, p. 1529, 2021. <https://doi.org/10.1002/wics.1529>

TOWE, C. A. **A Competing Risks Model of Land Use Change**. Em: Agricultural and Applied Economics Association. Pensilvânia, 2011. Disponível em: <<https://ageconsearch.umn.edu/record/103725>>. Acesso em: 25 jul. 2024.

## **CAPÍTULO 1: REFERENCIAL TEÓRICO**

Neste capítulo, será feita uma contextualização sobre o Certificado de Proteção de Cultivar, destacando sua importância e funcionamento. Também será apresentado o embasamento teórico sobre a análise de sobrevivência, com ênfase na abordagem de riscos competitivos em tempo contínuo e discreto.

### **1.1. Proteção de cultivares**

Um cultivar é uma variedade vegetal que foi desenvolvida por meio de técnicas de melhoramento genético, visando apresentar características produtivas, decorativas ou outras que o tornem vantajoso para o cultivo (Carmo et al., 2019).

Segundo Ziomkowski et al. (2022), o melhoramento genético contribui para o avanço da agricultura, proporcionando o aumento da produção, a resistência às pragas, a adaptação a diferentes condições climáticas e de solo, além de viabilizar a produção em períodos fora da safra tradicional. Carmo et al. (2019), destacam ainda que o melhoramento representa uma atividade econômica relevante, especialmente no Brasil, em que os produtos agrícolas são os principais produtos da economia nacional.

O Certificado de Proteção de Cultivares é um instrumento para assegurar o reconhecimento da propriedade intelectual sobre novas variedades vegetais. Por meio dele, o obtentor adquire direitos exclusivos sobre a exploração comercial da variedade e o cultivo só pode ser realizado mediante sua autorização (Castro et al., 2021).

No Brasil, a regulamentação desse certificado foi estabelecida pela Lei nº 9.456/1997, que exige o cumprimento de cinco critérios para sua concessão: novidade, denominação própria, distinguibilidade, estabilidade e homogeneidade (Brasil, 1997). A análise e a decisão sobre os pedidos de proteção são de responsabilidade do Serviço Nacional de Proteção de Cultivares (SNPC), órgão vinculado ao Ministério da Agricultura, Pecuária e Abastecimento (MAPA).

O período de proteção previsto é de 18 anos para culturas perenes e de 15 anos para as culturas anuais (Brasil, 1997). No entanto, esse prazo pode ser interrompido em casos como expiração, renúncia, cancelamento ou anulação.

De acordo com as Informações aos Usuários do SNPC (Brasil, 2023), a expiração ocorre ao fim do prazo legal, tornando a cultivar de domínio público. A renúncia acontece quando o titular abdica dos direitos. O cancelamento pode ser motivado pela falta de pagamento da

anuidade, perda de homogeneidade ou estabilidade, ausência de um procurador no Brasil ou não apresentação de uma amostra viva. Já a anulação ocorre quando não são cumpridos os requisitos de novidade e distinguibilidade, há violação de direitos de terceiros, quando o título não corresponde ao objeto ou há omissão de exigências legais.

De acordo com Bulsing et al. (2010), a proteção de cultivares estimula investimentos em melhoramento genético, proporcionando variedades mais adaptadas às necessidades e com maior competitividade. Além disso, a proteção contribui para o fortalecimento do setor de sementes e possibilita a incorporação de materiais estrangeiros para aprimorar os programas nacionais de melhoramento (Sá e Saes, 2015).

Diante dessa perspectiva, destaca-se a relevância do estudo dos fatores que influenciam o tempo de duração da proteção de cultivares. A análise de sobrevivência se apresenta como um conjunto de técnicas estatísticas adequadas, pois possibilita investigar o tempo de duração considerando os diferentes modos de encerramento da proteção, os quais são: expiração por prazo, renúncia, cancelamento e anulação.

## 1.2. Análise de sobrevivência

A análise de sobrevivência é uma abordagem utilizada em estudos em que a variável resposta é o tempo até a ocorrência de um evento de interesse. Devido à natureza dos dados é necessário empregar técnicas estatísticas adequadas, considerando a presença de falhas e censuras. As falhas referem-se à ocorrência do evento de interesse, enquanto as censuras ocorrem quando o tempo até o evento não pode ser observado para todos os indivíduos do estudo.

De acordo com Colosimo e Giolo (2024), as censuras podem ocorrer por diversas causas, como a perda de acompanhamento dos indivíduos por um motivo não relacionado ao evento em estudo ou a conclusão do estudo sem que o evento tenha ocorrido.

Desta forma, a variável resposta nos estudos de análise de sobrevivência, além de ser composta pelo tempo até a falha também inclui uma variável indicadora de censura, a saber

$$C = \begin{cases} 1, & \text{se é um tempo de falha.} \\ 0, & \text{se é um tempo de censura.} \end{cases}$$

A variável aleatória não negativa  $T$ , que representa o tempo até a ocorrência da falha, é comumente especificada pela função de sobrevivência ou pela função de risco (Chiesa et al.,

2020). Além disso, Colosimo e Giolo (2024) afirmam que o procedimento inicial é encontrar uma estimativa para essas funções e, a partir delas, calcular as estatísticas de interesse como o tempo médio, mediana, percentis, etc.

A função de sobrevivência  $S(t)$ , corresponde a probabilidade de um indivíduo sobreviver ao tempo  $t$ . Portanto, é definida por

$$S(t) = P(T \geq t) = 1 - P(T < t) = 1 - F(t). \quad (1)$$

Nesta equação,  $F(t)$  representa a função de distribuição acumulada, que expressa a probabilidade de um indivíduo não sobreviver até o tempo  $t$ , ou seja, falhar antes deste tempo.

A função de risco  $\lambda(t)$ , fornece o risco instantâneo de um indivíduo experimentar o evento de interesse em um intervalo de tempo entre  $t$  e  $t + \Delta t$ , considerando que tenha sobrevivido até o tempo  $t$ . Esta função aborda a falha de indivíduos que estão em risco, ou seja, aqueles que ainda não experimentaram o evento de interesse e não foram censurados.

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \right\}. \quad (2)$$

A função de risco acumulado representa o risco de ocorrência do evento até um determinado tempo. Colosimo e Giolo (2024) destacam que sua interpretação está relacionada à avaliação da inclinação da função de risco  $\lambda(t)$  em cada tempo  $t$ . Essa função é definida por

$$\Lambda(t) = \int_0^t \lambda(u) du = -\ln[S(t)]. \quad (3)$$

Os métodos que podem ser utilizados para o estudo dessas funções são classificados em não paramétricos, semiparamétricos e paramétricos.

### 1.2.1. Técnica não paramétrica

Segundo Saccaro et al. (2019), o método não paramétrico de Kaplan-Meier (Kaplan e Meier, 1958) é uma das ferramentas mais utilizadas para estimar a função de sobrevivência. Por ser um método de estimação não paramétrico, ele não exige suposições sobre distribuições de probabilidade subjacentes aos dados (Romano e Stevanato, 2020).

Considerando que  $t_1, t_2, \dots, t_m$  sejam os únicos tempos ordenados de falha, então, o estimador de Kaplan-Meier da função de sobrevivência é definido como

$$\hat{S}(t) = \prod_{j: t_j \leq t} \left(1 - \frac{d_j}{n_j}\right). \quad (4)$$

em que  $d_j$  é o número de falhas que ocorrem no tempo  $t_j$  e  $n_j$  é o número de indivíduos sob risco de experimentar o evento imediatamente anterior a  $t_j$ .

O teste de significância mais utilizado para verificar a igualdade de duas funções de sobrevivência é o teste não paramétrico *log-rank* (Mantel, 1966). Esse teste é adequado sob a suposição de que a razão das funções de risco dos grupos a serem comparados é constante, ou seja, assume a proporcionalidade dos riscos.

Uma alternativa ao teste *log-rank* é o teste de Wilcoxon (Gehan, 1965), que ajusta o peso conforme o número de indivíduos sob risco, atribuindo maior peso aos dados no início da curva de sobrevivência, quando há mais indivíduos em risco (Colosimo e Giolo, 2024). Isso assegura que as falhas iniciais tenham mais peso em comparação com as falhas que ocorrem posteriormente.

Apesar de sua simplicidade e facilidade de aplicação, o método não paramétrico apresenta a limitação de não considerar o efeito das covariáveis na análise, o que frequentemente é o interesse na maioria dos estudos (Carvalho et al., 2011).

Diante disso, para estimar o efeito das covariáveis no tempo de sobrevivência, a proposta é utilizar as técnicas semiparamétrica e paramétrica.

### 1.2.2. Técnica semiparamétrica

O modelo de Cox (Cox, 1972) é uma técnica semiparamétrica utilizada para estimar o efeito das covariáveis sobre a função de risco. Sua principal vantagem é a flexibilidade, pois não exige uma suposição prévia sobre a distribuição dos tempos até o evento, ao contrário dos modelos paramétricos (Rizopoulos, 2011). Dessa forma, o modelo de Cox é expresso por

$$\lambda(t|\mathbf{x}) = \lambda_0(t) \exp\{\mathbf{x}^T \boldsymbol{\beta}\}. \quad (5)$$

em que  $\lambda_0(t)$  é a função de risco basal,  $\boldsymbol{\beta}$  é o vetor de parâmetros associados às covariáveis e  $\mathbf{x}$  é o vetor de covariáveis.

O modelo de Cox é considerado um método semiparamétrico porque combina um componente paramétrico e um não paramétrico. A parte não paramétrica é representada pela função de risco basal, que não assume uma forma específica. A paramétrica é representada pela suposição de que as covariáveis agem multiplicativamente sobre o risco (Carvalho et al., 2011).

Para a utilização desse modelo, a razão entre os riscos deve ser proporcional (Brito et al., 2020). Isso significa que a razão entre os riscos de dois indivíduos deve permanecer constante ao longo do tempo

$$\frac{\lambda(t|\mathbf{x}_1)}{\lambda(t|\mathbf{x}_2)} = \exp\{\boldsymbol{\beta}^T(\mathbf{x}_1 - \mathbf{x}_2)\}. \quad (6)$$

Essa suposição pode ser avaliada por teste estatístico e método gráfico baseado nos resíduos padronizados de Schoenfeld (Colosimo e Giolo, 2024). A ausência de inclinação no gráfico dos coeficientes de regressão estimados *versus* o tempo fornece evidências a favor da proporcionalidade.

Além disso, adequação do modelo pode ser verificada por meio da análise gráfica dos resíduos de Cox-Snell (Cox e Snell, 1968). Para que o modelo seja considerado apropriado, o gráfico dos resíduos  $e_i$  definidos para o modelo selecionado em relação à função de risco acumulado  $\hat{\Lambda}(e_i)$  deve se aproximar de uma reta que passa pela origem com inclinação igual a um (Colosimo e Giolo, 2024).

### 1.2.3. Técnica paramétrica

De acordo com o modelo de Cox, a suposição de proporcionalidade das taxas de falha deve ser atendida. Quando essa condição não é válida, uma alternativa é utilizar a abordagem paramétrica, que assume uma distribuição de probabilidade para os tempos de sobrevivência (Resende et al., 2014).

Entre os principais modelos paramétricos utilizados na análise de sobrevivência estão os modelos exponencial, Weibull, log-normal e log-logístico.

A seleção do melhor modelo pode ser realizada por meio do Critério de Informação de Akaike (AIC) (Akaike, 1974) e do Critério de Informação Bayesiano (BIC) (Schwarz, 1978).

A adequação do modelo escolhido pode ser avaliada com o uso dos resíduos de Cox-Snell, de forma semelhante à abordagem adotada para o modelo de Cox (Colosimo e Giolo, 2024).

### 1.3. Riscos competitivos

Nos dados de sobrevivência clássica discutidos na seção anterior, presume-se que os indivíduos experimentem apenas um tipo de evento durante o acompanhamento. No entanto, em determinadas situações, os indivíduos podem experimentar outros tipos de eventos, os quais podem impedir ou alterar a probabilidade de ocorrência do evento de interesse (Gooley et al., 1999). Tais eventos são denominados eventos competitivos.

Para cada indivíduo, os dados são também representados como um par  $(T, C)$ , mas que

$$C = \begin{cases} k, & \text{se é um tempo de falha.} \\ 0, & \text{se é um tempo de censura.} \end{cases}$$

com  $k \in \{1, 2, \dots, c\}$  representando o tipo da primeira falha observada.

O tempo de falha  $T$  pode ser contínuo ou discreto. De acordo com Ndlovu et al. (2022), dados discretos de tempo até o evento ocorrem quando a unidade de tempo é representada de forma intrinsecamente discreta ou devido a presença de empates.

Colosimo e Giolo (2024) destacam que, ao considerar a ordenação crescente dos tempos de falha na construção da função de verossimilhança, a estimativa do vetor de parâmetros do modelo se torna problemática na presença de empates. Isso ocorre porque, nesses casos, não é possível estabelecer uma ordenação única, sendo necessário considerar todas as permutações possíveis das observações empatadas.

Na presença de riscos competitivos, a interpretação da função de sobrevivência perde sua validade, uma vez que a ocorrência de um evento pode impedir a ocorrência dos demais (Colosimo e Giolo, 2024). Alternativamente, utiliza-se a função de incidência acumulada (FIA), também chamada de função de subdistribuição (Kalbfleisch e Prentice, 2002).

A FIA representa a probabilidade de um indivíduo experimentar um dos  $k$  tipos de eventos até o tempo  $t$  e é definida por

$$F_k(t) = P(T \leq t, C = k). \quad (7)$$

No entanto, a probabilidade de um indivíduo experimentar um evento de qualquer tipo até o tempo  $t$  é dada pela função de distribuição global, definida por

$$F(t) = P(T \leq t) = \sum_{k=1}^c F_k(t). \quad (8)$$

em que  $F_k(t)$  é a função de incidência acumulada para o evento do tipo  $k$  e  $c$  é o número total de tipos de eventos observados.

A expressão (8) implica a suposição de que os eventos devem ser independentes.

Para a análise considerando as covariáveis, pode ser utilizada duas funções de risco distintas: a função de risco causa-específica e a função de risco da subdistribuição. No contexto de tempo contínuo e discreto, essas funções de risco possuem diferenças em sua interpretação, que serão abordadas nas subseções seguintes.

#### 1.4. Riscos competitivos quando o tempo é contínuo

A função de risco causa-específica, denotada por  $\lambda_k(t)$ , representa a taxa instantânea da ocorrência do  $k$ -ésimo evento em indivíduos que não experimentaram nenhum dos eventos até o tempo  $t$ . Essa função é utilizada para estimar o efeito das covariáveis sobre um evento específico e é definida por

$$\lambda_k(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t < T \leq t + \Delta t, C = k | T > t)}{\Delta t} \right\}. \quad (9)$$

Segundo Austin e Fine (2017), é incorreto inferir que uma determinada covariável está associada a um aumento ou diminuição da incidência do evento de interesse ao utilizar a função de risco causa-específica. Portanto, essa abordagem é insuficiente quando o objetivo principal é analisar a FIA, o que justifica o uso da função de risco da subdistribuição.

A função de risco da subdistribuição, denotada por  $\tilde{\lambda}_k(t)$ , pode ser interpretada como a taxa instantânea de ocorrência do evento do tipo  $k$ , considerando que o evento de interesse ainda não ocorreu ou que um evento de risco competitivo já foi observado. Assim como a função de risco causa-específica, ela também é utilizada para estimar o efeito das covariáveis sobre o evento e é definida por

$$\tilde{\lambda}_k(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t < T \leq t + \Delta t, C = k | T > t \text{ ou } (T < t, C \neq k))}{\Delta t} \right\}. \quad (10)$$

Essa função apresenta uma relação direta com a FIA, dada por

$$\tilde{\lambda}_k(t) = -\frac{d \log(1 - F_k(t))}{dt}. \quad (11)$$

De acordo com Colosimo e Giolo (2024), a distinção entre as duas funções de risco é destacada pelo grupo de risco considerado. A função de risco da subdistribuição inclui no grupo de risco tanto os indivíduos que ainda não experimentaram nenhum evento quanto aqueles que já experimentaram um evento diferente do evento  $k$ . Em contraste, a função de risco causa-específica considera apenas os indivíduos que ainda não experimentaram nenhum evento.

#### 1.4.1. Técnica não paramétrica

Na presença de riscos competitivos, o uso do estimador de Kaplan-Meier pode gerar estimativas viesadas, pois o indivíduo é removido do conjunto de risco ao experimentar um evento competitivo (Zhang, 2017). Como alternativa, emprega-se o estimador não paramétrico da função de incidência acumulada (FIA).

De acordo com Colosimo e Giolo (2024), o estimador não paramétrico da FIA considera pontos de massa nos tempos em que ocorreram eventos do tipo  $k$ , dependendo do total de indivíduos que experimentam o evento e do total de indivíduos sob risco.

$$\hat{F}_k(t) = \sum_{t_j \leq t} \hat{\lambda}_k(t_j) \hat{S}(t_{j-1}). \quad (12)$$

em que  $\hat{\lambda}_k(t_j)$  é a função de risco causa-específica da ocorrência do evento do tipo  $k$  e  $\hat{S}(t_{j-1})$  é a estimativa de sobrevivência global em  $t_{j-1}$  obtida pelo estimador de Kaplan-Meier.

A função de risco causa-específica da ocorrência do evento do tipo  $k$ , utilizada na expressão (12), pode ser estimada como

$$\hat{\lambda}_k(t_j) = \frac{d_{kj}}{n_j}. \quad (13)$$

em que  $d_{kj}$  é o número de eventos do tipo  $k$  que ocorre no tempo  $t_j$  e  $n_j$  é o número de indivíduos sob risco em  $t_j$ .

Assim como na análise de sobrevivência clássica, o efeito de uma covariável pode ser investigado por meio da estimativa não paramétrica das curvas de incidência acumulada e da comparação entre elas (Puter, 2007), sendo o teste de Gray o mais recomendado.

O teste de Gray baseia-se na comparação de médias ponderadas dos riscos da subdistribuição entre os grupos para o evento de interesse (Gray, 1988). Desta forma, considerando  $i$  grupos, a estatística de teste é definida por um *score* dado por

$$z_i(\tau) = \int_0^{\tau} W_i(t) \{ \tilde{\lambda}_i(t) - \tilde{\lambda}_0(t) \} dt. \quad (14)$$

em que  $\tau$  representa o tempo máximo de acompanhamento,  $W_i(t)$  é uma função peso,  $\tilde{\lambda}_i(t)$  é a função de risco da subdistribuição para o  $i$ -ésimo grupo e  $\tilde{\lambda}_0(t)$  é a função de risco da subdistribuição considerando todos os grupos.

A função peso  $W_i(t)$  é dada por  $W_i(t) = L(t) R_i(t)$ , em que  $L(t)$  é geralmente utilizada na forma  $L(t) = 1$ , equivalente ao teste *log-rank*, e  $R_i(t)$  representa o número ajustado de indivíduos sob risco no tempo  $t$ , determinado por

$$R_i(t) = n_i(t) \frac{1 - \hat{F}_i(t-)}{S_i(t-)} \quad (15)$$

em que  $n_i(t)$  é o número de indivíduos sob risco no tempo  $t$  no  $i$ -ésimo grupo,  $\hat{F}_i(t-)$  é o valor da FIA imediatamente anterior a  $t$  para o evento de interesse e  $S_i(t-)$  é o valor da probabilidade de estar livre de eventos até o momento anterior a  $t$ .

A estatística do teste consiste em uma forma quadrática que envolve  $Z = \sum_{i=1}^v z_i$  e sua respectiva matriz de variância e covariância. Sob a hipótese nula  $H_0: F_1(t) = \dots = F_v(t)$ , a estatística segue uma distribuição qui-quadrado com  $v - 1$  graus de liberdade, em que  $v$  é o número de grupos que se deseja testar.

A função *cuminc* do pacote *cmprsk* (Gray, 2024) do *software* R (R Core Team, 2025) pode ser utilizada para obter o estimador não paramétrico da função de incidência acumulada e o teste de Gray.

### 1.4.2. Técnicas semiparamétricas

Para a estimativa das funções de risco, dois modelos semiparamétricos são frequentemente utilizados: o modelo de taxas de falha causa-específica e o modelo de Fine-Gray. A escolha entre esses modelos depende da questão de interesse, conforme destacado por Pintilie (2006).

Van Der Pas et al. (2018) sugerem que o modelo de taxas de falha causa-específica é mais apropriado para investigar questões etiológicas ou biológicas, enquanto o modelo de Fine-Gray é mais indicado para predições sobre a incidência acumulada dos eventos.

O modelo de taxas de falha causa-específica é uma extensão do modelo de Cox de taxas de falhas proporcionais. Ele permite estimar o efeito de covariáveis sobre um evento específico, tratando os demais eventos como censurados (Carvalho et al., 2011). Para a causa  $k$ , o modelo de taxas de falha causa-específica é dado por

$$\lambda_k(t|\mathbf{x}) = \lambda_{0k}(t) \exp\{\mathbf{x}^T \boldsymbol{\beta}\}. \quad (16)$$

em que para a  $k$ -ésima causa,  $\lambda_{0k}(t)$  é a função de risco basal que não assume uma forma específica,  $\boldsymbol{\beta}$  é o vetor de parâmetros associados às covariáveis e  $\mathbf{x}$  é o vetor de covariáveis.

Para estimar a função de risco causa-específica utilizando o modelo de taxas de falha causa-específica, o vetor de coeficientes de regressão é estimado maximizando a verossimilhança parcial dada por

$$L(\boldsymbol{\beta}) = \prod_{j=1}^m \frac{\exp(\mathbf{x}_j^T \boldsymbol{\beta})}{\sum_{l \in R(t_j)} \exp(\mathbf{x}_l^T \boldsymbol{\beta})}. \quad (17)$$

em que  $R(t_j)$  representa o número de indivíduos sob risco no tempo  $t_j$  e  $\mathbf{x}_j$  o vetor de covariáveis do indivíduo que falhou em  $t_j$ .

A função *coxph* do pacote *survival* (Therneau et al., 2024) do *software* R (R Core Team, 2025) pode ser utilizada para essa análise.

Putter et al. (2007) destacam que, na presença de riscos competitivos, o efeito das mesmas covariáveis para diferentes causas de falha não se estende para as funções de incidência acumulada. Austin e Fine (2017) afirmam que, por essa razão, é incorreto inferir que uma

determinada covariável está associada ao aumento ou diminuição da incidência do evento de interesse.

O modelo de Fine-Gray (Fine e Gray, 1999) é baseado na função de risco da subdistribuição e é o mais adequado para a análise de incidência em riscos competitivos. Assim, se uma variável aumenta a função de risco da subdistribuição, ela também aumentará a incidência do evento (Austin e Fine, 2017). Para a causa  $k$ , o modelo de Fine-Gray é dado por

$$\tilde{\lambda}_k(t|\mathbf{x}) = \tilde{\lambda}_{0k}(t) \exp\{\mathbf{x}^T \boldsymbol{\beta}\}. \quad (18)$$

em que para a  $k$ -ésima causa,  $\tilde{\lambda}_{0k}(t)$  representa a função de risco basal da subdistribuição,  $\boldsymbol{\beta}$  é o vetor de parâmetros associados às covariáveis e  $\mathbf{x}$  é o vetor de covariáveis.

Para estimar a função de risco da subdistribuição utilizando o modelo de Fine-Gray, o vetor de coeficientes de regressão é estimado maximizando a verossimilhança parcial dada por

$$L(\boldsymbol{\beta}) = \prod_{j=1}^m \left[ \frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta})}{\sum_{l \in R(t_j)} w_{jl} \exp(\mathbf{x}_l^T \boldsymbol{\beta})} \right]^{C=k}. \quad (19)$$

em que  $R(t_j)$  representa o número de indivíduos sob risco para a causa  $k$  no tempo  $t_j$  e  $w_{jl}$  é uma função peso.

Para evitar viés nas estimativas, indivíduos que experimentaram eventos diferentes de  $k$  continuam no grupo de risco (Assane et al., 2021) e sua contribuição é determinada por

$$w_{jl} = \begin{cases} 1, & \text{se o indivíduo não experimentou evento ou censura em } t_j. \\ \frac{\hat{G}(t_j)}{\hat{G}(t_l)}, & \text{se o indivíduo experimentou o evento competitivo em } t_l < t_j. \\ 0, & \text{se o indivíduo foi censurado.} \end{cases} \quad (20)$$

sendo  $\hat{G}$  a estimativa de Kaplan-Meier da função da distribuição da censura, a qual é definida pelo par  $(T_i, \delta_i)$ , em que  $T_i$  é o tempo até a ocorrência primeiro evento, e  $\delta_i$  é uma variável indicadora que assume o valor 1 se nenhum evento foi observado e 0 caso qualquer tipo de evento tenha sido observado (Pintilie, 2006).

A função *crr* do pacote *cmprsk* (Gray, 2024) do *software* R (R Core Team, 2025) pode ser utilizada para essa análise.

Para que ambos os modelos ajustados sejam considerados adequados, é necessário avaliar a suposição de proporcionalidade das taxas de falha e a qualidade do ajuste do modelo.

No modelo de taxas de falha causa-específica, a suposição de proporcionalidade pode ser verificada por meio dos resíduos padronizados de Schoenfeld, enquanto a qualidade do ajuste pode ser avaliada utilizando os resíduos de Cox-Snell, de forma semelhante à abordagem adotada para o modelo de Cox (Pintilie, 2006).

No modelo de Fine-Gray, a suposição de proporcionalidade pode ser verificada por meio dos resíduos padronizados de Schoenfeld, também de forma similar ao modelo de Cox. A qualidade do ajuste é analisada comparando a proximidade das funções de incidência acumulada obtidas pelo estimador não paramétrico com as previstas pelos modelos (Colosimo e Giolo, 2024).

Contudo, distribuições paramétricas como exponencial, gama e Weibull também podem ser empregadas na modelagem de dados com riscos competitivos, conforme sugerido por Flehinger et al. (2002) e Sarhan et al. (2010).

### 1.5. Riscos competitivos quando o tempo é discreto

No contexto de tempo discreto, as funções de risco diferem em sua interpretação, como mencionado anteriormente. Nesse cenário, as funções de risco representam uma probabilidade, ao invés de uma taxa, como ocorre no caso contínuo. Segundo Schmid e Berger (2021), essa interpretação é geralmente mais intuitiva e acessível para os pesquisadores em comparação com as funções de risco em tempo contínuo.

Supondo que  $T$  assuma valores discretos, com sua função de probabilidade dada por  $p(t) = P(T = t)$  e  $\mathbf{x}$  como o vetor de covariáveis, a função de risco causa-específica é definida como a probabilidade de falha associada à causa  $k$  dado que nenhum outro evento ocorreu entre os indivíduos sob risco até o tempo  $t$  (Meir e Gorfine, 2023). Essa função é expressa como

$$\lambda_k(t|\mathbf{x}) = P(T = t, C = k | T \geq t). \quad (21)$$

Além disso, é possível obter outras funções relevantes no contexto de tempo discreto, incluindo a função de risco geral, a função de sobrevivência e a probabilidade do evento específico para cada causa, expressas respectivamente por

$$\lambda(t|\mathbf{x}) = \sum_{k=1}^c \lambda_k(t|\mathbf{x}) = P(T = t|T \geq t), \quad (22)$$

$$S(t|\mathbf{x}) = P(T > t|\mathbf{x}) = \prod_{j=1}^m (1 - \lambda(j|\mathbf{x})), \quad (23)$$

$$P(T = t, C = k|\mathbf{x}) = \lambda_k(t|\mathbf{x}) S(t-1). \quad (24)$$

Em contraste com a abordagem de tempo contínuo, a função de risco da subdistribuição em tempo discreto é definida em termos da probabilidade de se experimentar o evento de interesse no tempo  $t$ , dado que nenhum evento ocorreu até então, ou que um evento concorrente já tenha ocorrido antes de  $t$ . Segundo Berger et al. (2020), a função de risco da subdistribuição discreta é dada por

$$\tilde{\lambda}_k(t|\mathbf{x}) = P(T = t, C = k | (T \geq t) \text{ ou } (T \leq t, C \neq k)). \quad (25)$$

Assim como na abordagem contínua, essa função apresenta uma relação direta com a função de incidência acumulada, que no caso discreto é expressa como

$$F_k(t|\mathbf{x}) = 1 - \prod_{j=1}^m (1 - \tilde{\lambda}_k(j|\mathbf{x})) = 1 - S_k(t|\mathbf{x}). \quad (26)$$

em que  $S_k(t|\mathbf{x})$  é a função de sobrevivência discreta para um evento do tipo  $k$ .

### 1.5.1. Técnica paramétrica

O modelo de regressão multinomial tem sido amplamente utilizado como o padrão para estimar a função de risco causa-específica em tempo discreto (Ndlovu et al., 2022). Baseado em uma transformação de regressão, esse modelo pode ser descrito pela equação

$$h(\lambda_k(t|\mathbf{x})) = \alpha_k(t) + \mathbf{x}^T \boldsymbol{\beta}. \quad (27)$$

em que  $h$  é uma função de ligação e  $\alpha_k(t)$ , quando não especificada, atua de forma semelhante a uma função de risco basal no modelo de riscos proporcionais de Cox (Meier e Gorfine, 2023).

Ao adotar a função de ligação logit, a função de risco causa-específica é expressa como

$$\lambda_k(t|\mathbf{x}) = \frac{\exp\{\alpha_k(t) + \mathbf{x}^T\boldsymbol{\beta}\}}{1 + \exp\{\alpha_k(t) + \mathbf{x}^T\boldsymbol{\beta}\}}. \quad (28)$$

De acordo com Beyersmann et al. (2012), os coeficientes de regressão obtidos com esse modelo devem ser próximos aos estimados pelo modelo de Cox para os riscos específicos da causa.

A estimação dos parâmetros pode ser realizada por meio de uma abordagem baseada em modelos lineares generalizados. Tutz e Schmid (2016) recomendam a função *dataLongCompRisk* do pacote *discSurv* (Welchowski et al., 2022) para converter dados de sobrevivência com eventos concorrentes em um quadro de dados com resposta multinomial. Já Beyersmann et al. (2012) sugerem o ajuste do modelo multinomial com função de ligação logit, utilizando a função *multinom* do pacote *nnet* (Ripley e Venables, 2025). Ambos os pacotes do *software R* (R Core Team, 2025).

Berger et al. (2020) apresentam uma extensão da abordagem de modelagem para tempos de eventos contínuos de Fine e Gray para dados de tempos discretos, aplicada à função de risco da subdistribuição. Eles recomendam o uso do modelo de regressão binária ponderada para modelar essa função, o que pode ser expresso por

$$h\left(\tilde{\lambda}_k(t|\mathbf{x})\right) = \beta_0(t) + \mathbf{x}^T\boldsymbol{\beta}. \quad (29)$$

em que  $h$  é uma função de ligação e  $\beta_0(t)$  refere-se aos coeficientes de linha de base.

O método pode ser implementado utilizando *software* de regressão binária, em que os dados são transformados para conter uma matriz aumentada e um vetor de pesos. A matriz de dados aumentada é construída de forma que cada indivíduo contribua com várias linhas de dados, uma para cada intervalo de tempo em que está sob risco. Os pesos são ajustados para refletir a probabilidade de o indivíduo permanecer sob risco em cada intervalo.

Os pesos são aplicados para ajustar a contribuição de cada observação na função de verossimilhança e são definidos de forma análoga aos pesos utilizados nos modelos de risco da subdistribuição para tempo contínuo.

Dessa forma, ao adotar o modelo de regressão binária ponderada com função de ligação log-log complementar (c log-log), conforme indicado por Schmid e Berger (2018), a função de risco da subdistribuição é modelada por

$$\tilde{\lambda}_k(t|\mathbf{x}) = 1 - \exp(-\exp(\beta_0(t) + \mathbf{x}^T \boldsymbol{\beta})). \quad (30)$$

em  $\beta_0(t)$  refere-se aos coeficientes de linha de base,  $\boldsymbol{\beta}$  é o vetor de parâmetros associados às covariáveis e  $\mathbf{x}$  é o vetor de covariáveis

A estimação dos parâmetros pode ser realizada por meio de uma abordagem baseada em modelos lineares generalizados. O procedimento pode ser implementado utilizando a função *dataLongSubDist* do pacote *discSurv* (Welchowski et al., 2022), em conjunto com a função *glm* do *software R* (R Core Team, 2025).

O desempenho dos modelos pode ser analisado por meio da curva ROC (*Receiver Operating Characteristic*), sendo a área sob a curva (AUC) classificada segundo os critérios: inferior a 0,7 como suficiente, entre 0,7 e 0,8 como aceitável e superior a 0,8 como excelente (Hosmer et al., 2013).

## Referências

AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v. 19, n. 6, p. 716–723, 1974. <https://doi.org/10.1109/TAC.1974.1100705>

ASSANE, C. C.; PEREIRA, B. D. B.; MUAUALO, M. A. M. Modelos de riscos competitivos para a análise de sobrevivência de pacientes com hipertensão arterial em tratamento no hospital Universitário, HUCFF/UFRJ. **Latin American Journal of Development**, v.2, p.488–504, 2021. <https://doi.org/10.46814/lajdv2n6-013>

AUSTIN, P. C.; FINE, J. P. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. **Statistics in Medicine**, v. 36, n. 27, p. 4391–4400, 2017. <https://doi.org/10.1002/sim.7501>

BERGER, M.; SCHMID, M.; WELCHOWSKI, T.; SCHMITZ-VALCKENBERG, S.; BEYERSMANN, J. Subdistribution hazard models for competing risks in discrete time. **Biostatistics**, v. 21, n. 3, p. 449–466, 2020. <https://doi.org/10.1093/biostatistics/kxy069>

BEYERSMANN, J.; ALLIGNOL, A.; SCHUMACHER, M. **Competing Risks and Multistate Models with R**. New York, NY: Springer New York, 2012.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Informações aos Usuários do SNPC**. 2023. Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/insumos->

agropecuarios/insumos-agricolas/protecao-de-cultivar/informacoes-publicacoes/informacoes-aos-usuarios-do-snpc-fevereiro-de-2020>. Acesso em: 22 set. 2024.

BRASIL. Lei nº 9.456, de 25 de abril de 1997. Institui a Lei de Proteção de Cultivares e dá outras providências. Diário Oficial da União, Brasília, DF, 28 abr. 1997. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/leis/19456.htm](https://www.planalto.gov.br/ccivil_03/leis/19456.htm)>. Acesso em: 22 set. 2024.

BRITO, A. D. L.; XAVIER JUNIOR, S. F.; MENDONÇA, E. B. de; XAVIER, E. F. M.; SANTOS, T. T. de M.; OLIVEIRA, T. A. de. Ajuste de modelos de fragilidade e riscos proporcionais aplicados a dados de retinopatia diabética. **Research, Society and Development**, v. 9, n. 8, p. e478985691, 2020. <https://doi.org/10.33448/rsd-v9i8.5691>

BULSING, A. C.; AVIANI, D. S.; PACHECO, L. G. A.; MACHADO, R. Z. Proteção de cultivares. In Brasil. Ministério da Agricultura Pecuária e Abastecimento (ed.) **Curso de propriedade intelectual & inovação no agronegócio**. MAPA/EaD/UFSC, Brasília/Florianópolis, 2010.

CARMO, F. L. do; HANNA, S. A.; UCHÔA, S. B. B.; ANGELI, R.; BRUCH, K. L. Cultivares – O que são, como se Apropriar, como Consultar. In: SANTOS, W. P. C. dos (org.). **Conceitos e aplicações de propriedade intelectual**. Salvador (BA): IFBA, p. 349-399, 2019.

CARVALHO, M. S.; ANDREOZZI, V. L.; CODEÇO, C. T.; CAMPOS, D. P.; BARBOSA, M. T. S.; SHIMAKURA, S. E. **Análise de sobrevivência: teoria e aplicações em saúde**. Fiocruz Editora, 2011.

CASTRO, E. D. C.; PACHECO, A.; SOUZA, C. B.; WANDER, A. Estruturas de governança de transações das organizações detentoras de direitos de proteção em cultivares de feijão-comum no Brasil. **Revista Razão Contábil e Finanças**, v. 12, n. 2, p. 1-22, 2021.

CHALITA, L. V. A. S.; COLOSIMO, E. A.; DEMÉTRIO, C. G. B. Likelihood approximations and discrete models for tied survival data. **Communications in Statistics - Theory and Methods**, v. 31, n. 7, p. 1215–1229, 2002. <https://doi.org/10.1081/STA-120004920>

CHIESA, E. M. F.; MORAES, A. B.; RIGÃO, M. H. Análise de sobrevivência através de um método não-paramétrico. **Ciência e Natura**, v. 42, p. e19, 2020. <https://doi.org/10.5902/2179460X40479>

COLOSIMO, E. A.; GIOLO, S. R. **Análise de sobrevivência aplicada**. 2. ed. São Paulo, SP: Editora Edgard Blucher, 2024.

COX, D. R. Regression Models and Life-Tables. **Journal of the Royal Statistical Society Series B: Statistical Methodology**, v. 34, n. 2, p. 187–202, 1972. <https://doi.org/10.1111/j.2517-6161.1972.tb00899.x>

COX, D. R.; SNELL, E. J. A General Definition of Residuals. **Journal of the Royal Statistical Society Series B: Statistical Methodology**, v. 30, n. 2, p. 248–265, 1968. <https://doi.org/10.1111/j.2517-6161.1968.tb00724.x>

FINE, J. P.; GRAY, R. J. A proportional hazard model for subdistribution of a competing risk. **Journal of the American Statistical Association**, v. 94, n. 446, p. 496–509, 1999. <https://doi.org/10.1080/01621459.1999.10474144>

FLEHINGER, B. J.; REISER, B.; YASHCHIN, E. Parametric modeling for survival with competing risks and masked failure causes. **Lifetime Data Analysis**, v. 8, n. 2, p. 177–203, 2002. <https://doi.org/10.1023/A:1014891707936>

GEHAN, E. A. A Generalized Wilcoxon Test for Comparing Arbitrarily Singly-Censored Samples. **Biometrika**, v. 52, n. 1/2, p. 203, 1965. <https://doi.org/10.2307/2333825>

GOOLEY, T. A.; LEISENRING, W.; CROWLEY, J.; STORER, B. E. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. **Statistics in Medicine**, v. 18, n. 6, p. 695–706, 1999. [https://doi.org/10.1002/\(SICI\)1097-0258\(19990330\)18:6<695::AID-SIM60>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<695::AID-SIM60>3.0.CO;2-O)

GRAY, B. **cmprsk: Subdistribution Analysis of Competing Risks**. 2024. Disponível em: <<https://CRAN.R-project.org/package=cmprsk>>. Acesso em: 24 set. 2024.

GRAY, R. J. A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. **The Annals of Statistics**, v. 16, n. 3, 1988. <https://doi.org/10.1214/aos/1176350951>

HOSMER J. R., D. W.; LEMESHOW, S.; STURDIVANT, R. X. **Applied Logistic Regression**. 3. ed. Hoboken, John Wiley & Sons, 2013.

KALBFLEISCH, J. D.; PRENTICE, R. L. **The statistical analysis of failure time data**. 2. ed. Hoboken, N.J: Wiley-Interscience, 2002.

KAPLAN, E. L.; MEIER, P. Nonparametric Estimation from Incomplete Observations. **Journal of the American Statistical Association**, v. 53, n. 282, p. 457–481, 1958. <https://doi.org/10.1080/01621459.1958.10501452>

MANTEL, N. Evaluation of pacote data and two new rank order statistics arising in its consideration. **Cancer Chemotherapy Reports**, v. 50, n. 3, p. 163–170, 1966.

MEIR, T.; GORFINE, M. Discrete-time Competing-Risks Regression with or without Penalization. **arXiv**, 2023. <https://doi.org/10.48550/arXiv.2303.01186>

NDLOVU, B.; MELESSE, S.; ZEWOTIR, T. A regression analysis of discrete time competing risks data using a vertical model approach. **South African Statistical Journal**, v. 56, n. 1, p. 21–36, 2022. <https://doi.org/10.37920/sasj.2022.56.1.2>

PINTILIE, M. **Competing risks: a practical perspective**. Chichester, England; Hoboken, NJ: John Wiley & Sons, 2006.

PUTTER, H.; FIOCCO, M.; GESKUS, R. B. Tutorial in biostatistics: competing risks and multi-state models. **Statistics in Medicine**, v. 26, n. 11, p. 2389–2430, 2007. <https://doi.org/10.1002/sim.2712>

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing. 2025. Disponível em: <<https://www.r-project.org/>>.

RESENDE, M. D. V. de; SILVA, F. F. E; AZEVEDO, C. F. **Estatística matemática, biométrica e computacional**. Viçosa: Suprema, 2014.

RIPLEY, B.; VENABLES, W. **nnet: Feed-Forward Neural Networks and Multinomial Log-Linear Models**. 2025. Disponível em: <<https://CRAN.R-project.org/package=nnet>>. Acesso em: 24 jan. 2025.

RIZOPOULOS, D. Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. **Biometrics**, v. 67, n. 3, p. 819–829, 2011. <https://doi.org/10.1111/j.1541-0420.2010.01546.x>

ROMANO, A.; STEVANATO, P. Germination Data Analysis by Time-to-Event Approaches. **Plants**, v.9, n. 5, p. 617, 2020. <https://doi.org/10.3390/plants9050617>

SÁ, C. D. de; SAES, M. S. M. Propriedade intelectual na cadeia de flores e plantas ornamentais: uma análise da legislação brasileira de proteção de cultivares. **Revista Brasileira de Inovação**, v.14, p.49–76, 2015. <https://doi.org/10.20396/rbi.v14i1.8649089>

SACCARO, A.; FRANÇA, M. T. A.; JACINTO, P. D. A. Fatores Associados à Evasão no Ensino Superior Brasileiro: um estudo de análise de sobrevivência para os cursos das áreas de Ciência, Matemática e Computação e de Engenharia, Produção e Construção em instituições públicas e privadas. **Estudos Econômicos (São Paulo)**, v. 49, n. 2, p. 337–373, 2019. <https://doi.org/10.1590/0101-41614925amp>

SARHAN, A. M.; HAMILTON, D. C.; SMITH, B. Statistical analysis of competing risks models. **Reliability Engineering & System Safety**, v.95, n. 9, p. 953–962, 2010. <https://doi.org/10.1016/j.ress.2010.04.006>

SCHMID, M.; BERGER, M. Competing risks analysis for discrete time-to-event data. **WIREs Computational Statistics**, v.13, n. 5, p. e1529, 2021. <https://doi.org/10.1002/wics.1529>

SCHWARZ, G. Estimating the Dimension of a Model. **Annals of Statistics**, v.6, p. 461-464, 1978.

THERNEAU, T. M.; LUMLEY, T.; ELIZABETH, A.; CYNTHIA, C. **survival: Survival Analysis**. 2024. Disponível em: <<https://CRAN.R-project.org/package=survival>>. Acesso em: 22 dez. 2024.

TUTZ, G.; SCHMID, M. **Modeling Discrete Time-to-Event Data**. Cham: Springer International Publishing, 2016.

VAN DER PAS, S.; NELISSEN, R.; FIOCCO, M. Different competing risks models for different questions may give similar results in arthroplasty registers in the presence of few events. **Acta Orthopaedica**, v. 89, n. 2, p. 145–151, 2018. <https://doi.org/10.1080/17453674.2018.1427314>

WELCHOWSKI, T.; BERGER, M.; KOEHLER, D.; SCHMID, M. **discSurv: Discrete Time Survival Analysis**. 2022. Disponível em: <<https://CRAN.R-project.org/package=discSurv>>. Acesso em: 24 set. 2024.

ZHANG, Z. Survival analysis in the presence of concurrent risks. **Annals of Translational Medicine**, v. 5, n. 3, p. 47, 2017. <https://doi.org/10.21037/atm.2016.08.62>

ZIOMKOWSKI, P.; BRUCH, K. L.; VOLPATO, G. Propriedade Intelectual, Registro e Transferência de Tecnologia: Estudo de Caso de cultivares de aveia da Universidade Federal do Rio Grande do Sul (UFRGS). **Cadernos de Prospecção**, v. 15, n. 1, p. 18–35, 2022.

## **CAPÍTULO 2: ANÁLISE DE SOBREVIVÊNCIA COM RISCOS COMPETITIVOS APLICADA NA DURAÇÃO DA PROTEÇÃO DE CULTIVARES**

### **Resumo**

A proteção de cultivares assegura direitos de propriedade intelectual sobre novas variedades vegetais por meio do Certificado de Proteção de Cultivar. No entanto, eventos como renúncia, cancelamento e anulação podem encerrar essa proteção antes do prazo legal, configurando um cenário de riscos competitivos. Esta pesquisa investiga como as características do titular: instituição pública, setor privado e parcerias e do cultivar: presença de transgenia, tipo de cultura e ciclo de vida, influenciam a duração do certificado. Foram analisadas 4966 proteções obtidas da plataforma *CultivarWeb* do Ministério da Agricultura, Pecuária e Abastecimento, entre 1997 e 2024. O evento de interesse (expiração por prazo) e os eventos competitivos (renúncia, cancelamento e anulação) foram analisados utilizando uma abordagem de tempo contínuo por meio do modelo de Fine-Gray e com uma abordagem de tempo discreto com o modelo de regressão binária ponderada devido à presença de empates. Os resultados indicaram que cultivares transgênicas apresentam menor incidência tanto de expiração quanto de eventos competitivos. Proteções do setor privado têm menor incidência de expiração, mas com maior propensão à renúncia, cancelamento e anulação. Cultivares ornamentais registraram as maiores incidências de renúncia, independentemente do ciclo de vida. Dessa forma, conclui-se que características do titular e da cultivar impactam a duração da proteção.

**Termos para indexação:** dados censurados, função de subdistribuição, modelo de Fine-Gray, modelo de regressão binária ponderada, propriedade intelectual.

### **Introdução**

No Brasil a concessão do Certificado de Proteção de Cultivar foi instaurada em 1997 e exige o cumprimento de cinco critérios: novidade, denominação própria, distinguibilidade, estabilidade e homogeneidade (Brasil, 1997).

O período do Certificado de Proteção de Cultivares é de 18 anos para culturas perenes e de 15 anos para culturas anuais. No entanto, a proteção pode ser interrompida por: expiração do prazo, que ocorre ao término do período legal, quando a cultivar entra em domínio público; renúncia, solicitada pelo titular; cancelamento, devido a perda de homogeneidade ou estabilidade da cultivar, falta de pagamento da anuidade, ausência de um procurador qualificado

ou não apresentação de amostra viva; e anulação, quando não são atendidos critérios como novidade e distinguibilidade, ou em caso de desrespeito a direitos de terceiros ou omissão de requisitos legais (Brasil, 2023).

A proteção de cultivares incentiva investimentos em melhoramento genético e promove o desenvolvimento de variedades mais produtivas, resistentes a pragas e doenças, e adaptadas a diferentes necessidades (Bulsing et al., 2010; Glenn et al., 2017). Além disso, beneficia instituições públicas e privadas de pesquisa, que podem obter recursos financeiros por meio dos direitos sobre as cultivares desenvolvidas, sustentando programas de melhoramento e o lançamento de novas variedades de interesse para o agronegócio (Carvalho et al., 2009; Castro et al., 2021).

No entanto, o desenvolvimento e a proteção de novas cultivares demandam recursos financeiros, tecnologia, infraestrutura e material genético (ISF & ABRASEM, 2021; Yu & Chung, 2021). Segundo Santos et al. (2012), o tempo necessário para desenvolver e proteger uma nova variedade varia entre 6 e 30 anos, dependendo da espécie. Além disso, ressaltam ainda que esse período pode ser ainda maior devido à exigência de testes para validação, como os de Distinguibilidade, Homogeneidade e Estabilidade (DHE), que aumentam tanto a duração quanto os custos do processo.

Diante de tais esforços, compreender os motivos que levam à expiração da proteção de cultivares torna-se de notória relevância. A análise de sobrevivência, um conjunto de técnicas que investiga o tempo até a ocorrência de um evento, emprega conceitos como falha e censura (Colosimo & Giolo, 2024) e pode contribuir para esse entendimento. No contexto da proteção de cultivares, quando o certificado se encerra ao final do prazo estabelecido por lei, ocorre uma falha, pois o evento de interesse (expiração por prazo) aconteceu. No entanto, as proteções que continuam vigentes até o final do período de observação são tratadas como censuras, uma vez que o término da proteção não ocorreu. Além disso, a proteção pode ser encerrada por renúncia, cancelamento ou anulação. Nessas situações, o término da proteção por qualquer um desses motivos impede que ela atinja o prazo máximo previsto em lei, ou seja, o evento de expiração por prazo não ocorre.

Quando, em um estudo, outros tipos de eventos acontecem, podendo impedir ou alterar a probabilidade de ocorrência do evento de interesse, a abordagem recomendada para utilizar a análise de sobrevivência é a de riscos competitivos (Austin et al., 2016). Diversos estudos nas ciências agrárias têm explorado essa abordagem, como os de Towe (2011) em que foi analisada a conversão de terras agrícolas para uso residencial em contraposição à preservação de áreas agrícolas. Genna et al. (2015) investigaram a germinação de sementes em contraste com

infecções por patógenos. Meng et al. (2020) examinaram a concessão de certificados de proteção de cultivares frente à renúncia desses certificados. Wei et al. (2024), por sua vez, examinaram dois mecanismos distintos de mortalidade em mudas de *Chamaecyparis obtusa* var. *formosana*: a morte induzida por fatores ambientais e a provocada por herbivoria.

Desta forma, o objetivo desta pesquisa foi investigar como as características do titular: instituição pública, setor privado e parcerias e do cultivar: presença de transgenia, tipo de cultura e ciclo de vida, influenciam a duração do certificado.

## Material e Métodos

Os dados utilizados neste trabalho foram obtidos na plataforma *CultivarWeb* (MAPA, 2024), abrangendo o período de 1997 a 2024. Incluem informações sobre 4966 proteções, considerando o tempo de início e término da proteção (em meses), indicação de modificação genética, características do titular (instituição pública, setor privado e parceria), tipo de cultura (agrícola, forrageira, florestal, frutífera, olerícola, ornamental, medicinal e aromática) e ciclo de vida (anual e perene).

O evento de interesse considerado foi a expiração da proteção devido ao término do prazo, enquanto os eventos competitivos analisados foram a renúncia, o cancelamento e a anulação da proteção. As observações censuradas referem-se as cultivares cuja proteção permaneceu vigente durante o período estudado.

Os métodos estatísticos empregados para análise dos eventos foram as abordagens de riscos competitivos baseados em tempo contínuo e tempo discreto, devido à presença de empates. Como as cultivares de ciclo perene possuem o tempo de proteção menor em relação as cultivares de ciclo anual, as análises foram realizadas separadamente para os dois ciclos.

A função de risco da subdistribuição permite interpretar os efeitos das covariáveis, fornecendo estimativas relacionadas à incidência acumulada dos eventos, considerando a presença de todos os demais eventos na análise (Austin & Fine, 2017). Nesse contexto, o grupo de risco inclui tanto os indivíduos que experimentaram o evento competitivo quanto aqueles que ainda não vivenciaram nenhum dos eventos (Schuster et al., 2020).

Segundo Carvalho et al. (2011), a contribuição desses indivíduos para o grupo de risco é ponderada por uma função peso que decresce ao longo do tempo, definida por

$$w_{ji} = \begin{cases} 1, & \text{se o indivíduo não experimentou evento ou censura em } t_j. \\ \frac{\hat{G}(t_j)}{\hat{G}(t_1)}, & \text{se o indivíduo experimentou o evento competitivo em } t_1 < t_j. \\ 0, & \text{se o indivíduo foi censurado.} \end{cases}$$

sendo  $\hat{G}$  a estimativa de Kaplan-Meier da função da distribuição da censura, a qual é definida pelo par  $(T_i, \delta_i)$ , em que  $T_i$  é o tempo até a ocorrência primeiro evento, e  $\delta_i$  é uma variável indicadora que assume o valor 1 se nenhum evento foi observado e 0 caso qualquer tipo de evento tenha sido observado (Pintilie, 2006).

A função de risco da subdistribuição apresenta diferenças, tanto na interpretação quanto na modelagem utilizada para ajuste, quando o tempo é considerado discreto ou contínuo. No tempo contínuo, essa função é interpretada como uma taxa instantânea, enquanto no tempo discreto, ela é entendida como a probabilidade de ocorrência do evento de interesse no tempo  $t$ , dado que o evento ainda não aconteceu ou que um evento de risco competitivo já foi observado (Schmid & Berger, 2021). Para ambas as situações, as expressões são definidas, respectivamente, por

$$\tilde{\lambda}_k(t|\mathbf{x}) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t < T \leq t + \Delta t, C = k | (T > t) \text{ ou } (T < t, C \neq k))}{\Delta t} \right\},$$

$$\tilde{\lambda}_k(t|\mathbf{x}) = P(T = t, C = k | (T \geq t) \text{ ou } (T \leq t, C \neq k)).$$

Na abordagem de tempo contínuo, o modelo desenvolvido por Fine e Gray é um dos mais utilizados para o ajuste (Assane et al., 2021). Assim como o modelo de Cox, ele estabelece a suposição de proporcionalidade para os riscos da subdistribuição, e a expressão para a  $k$ -ésima causa de falha pode ser escrita como

$$\tilde{\lambda}_k(t|\mathbf{x}) = \tilde{\lambda}_{0k}(t) \exp\{\mathbf{x}^T \boldsymbol{\beta}\}.$$

em que  $\tilde{\lambda}_{0k}(t)$  representa a função de risco de base da subdistribuição,  $\boldsymbol{\beta}$  é o vetor de parâmetros associados às covariáveis e  $\mathbf{x}$  é o vetor de covariáveis.

Para o ajuste deste modelo, foi utilizada a função *crr* do pacote *cmprsk* (Gray, 2024) do *software* R (R Core Team, 2025), adotando para as variáveis explicativas as seguintes categorias de referência: ausência de transgenia, instituição pública e tipo de cultura agrícola.

A significância dos parâmetros foi avaliada pelo teste de Wald, com nível de significância de 5%. A suposição de proporcionalidade foi verificada por meio dos gráficos dos resíduos Schoenfeld. A qualidade do ajuste foi analisada pela comparação entre as curvas de incidências acumuladas obtidas pelo estimador não paramétrico e as previstas pelo modelo, considerando subconjuntos formados pelas combinações das categorias das covariáveis.

Para o ajuste na abordagem de tempo discreto, Berger et al. (2020) recomendam o uso do modelo de regressão binária ponderada com a função de ligação log-log complementar, considerando uma matriz de dados aumentada e um vetor de pesos. A matriz de dados aumentada é construída replicando os tempos de risco para cada indivíduo, associando a cada linha uma variável indicadora do evento e um peso correspondente, sendo este peso baseado na expressão  $w_{ji}$  descrita anteriormente.

Dessa forma, a função de risco da subdistribuição é modelada por

$$\tilde{\lambda}_k(t|\mathbf{x}) = 1 - \exp(-\exp(\boldsymbol{\beta}_0(t) + \mathbf{x}^T \boldsymbol{\beta})).$$

em  $\boldsymbol{\beta}_0(t)$  refere-se aos coeficientes de linha de base,  $\boldsymbol{\beta}$  é o vetor de parâmetros associados às covariáveis e  $\mathbf{x}$  é o vetor de covariáveis.

A estimação do modelo de regressão binária ponderada na abordagem de tempo discreto foi realizada por meio de uma abordagem baseada em modelos lineares generalizados, utilizando as funções *dataLongSubDist* do pacote *discsurv* (Welchowski et al., 2022) e *glm* do *software R* (R Core Team, 2025). Também se adotou para as variáveis explicativas as categorias de referência: ausência de transgenia, instituição pública e tipo de cultura agrícola. A significância dos parâmetros foi avaliada pelo teste de Wald, com nível de significância de 5%.

O desempenho do modelo de regressão binária ponderada foi analisado por meio da curva ROC (*Receiver Operating Characteristic*), sendo a área sob a curva (AUC) classificada segundo os critérios: inferior a 0,7 como insuficiente, entre 0,7 e 0,8 como aceitável e superior a 0,8 como excelente (Hosmer et. al, 2013).

## Resultados e Discussão

Entre os 4966 registros de proteção analisados, 2511 permaneceram em vigor, 744 expiraram pelo término de prazo, 1465 foram renunciados e 246 foram cancelados ou anulados.

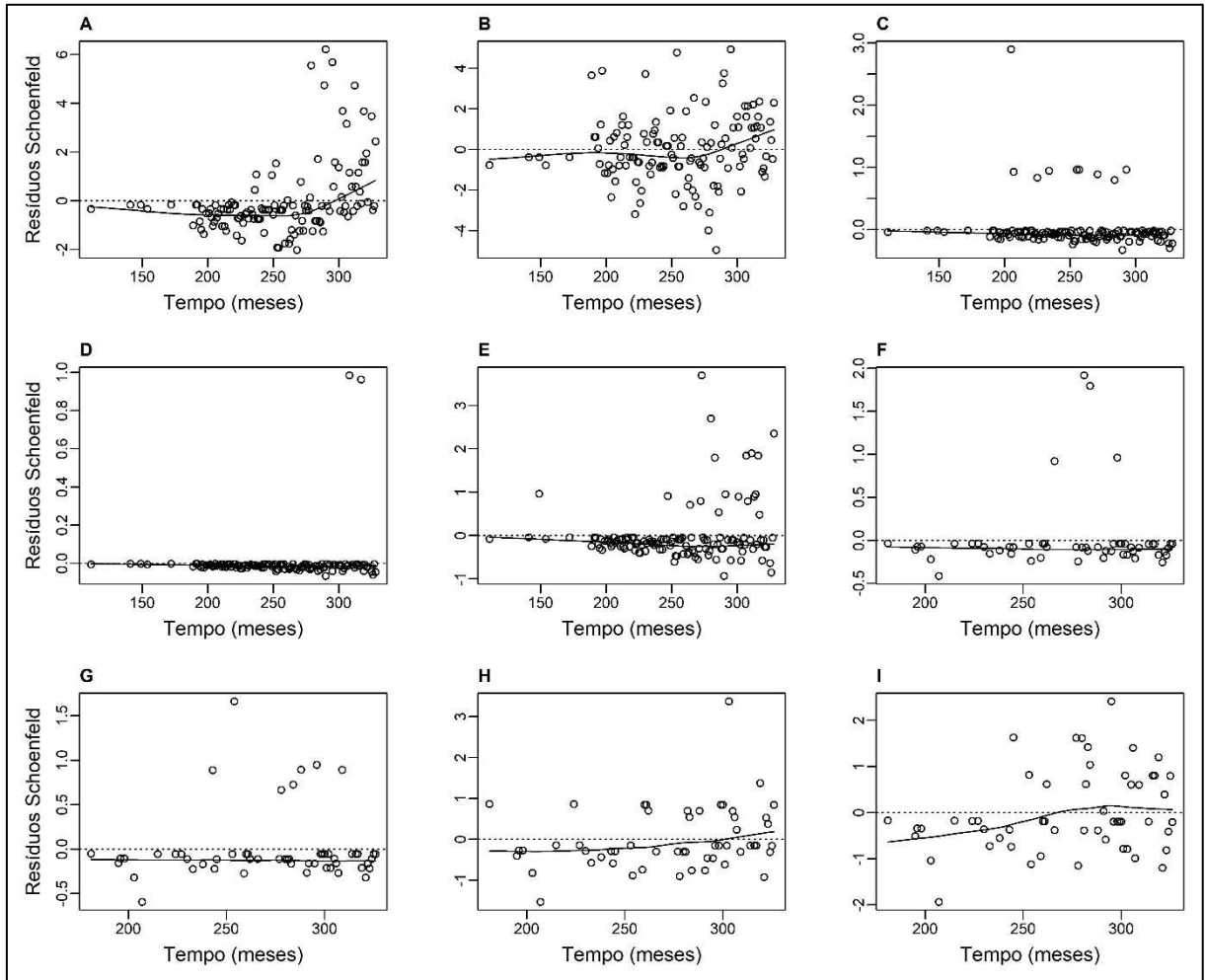
A Tabela 1 apresenta um resumo descritivo das covariáveis considerando os ciclos de vida anual e perene para os eventos em estudo.

**Tabela 1** - Resumo descritivo dos dados, conforme o tempo das proteções, definitiva, expiração por prazo, renunciadas e canceladas e/ou anuladas, considerando os ciclos de vida anual e perene para os eventos em estudo.

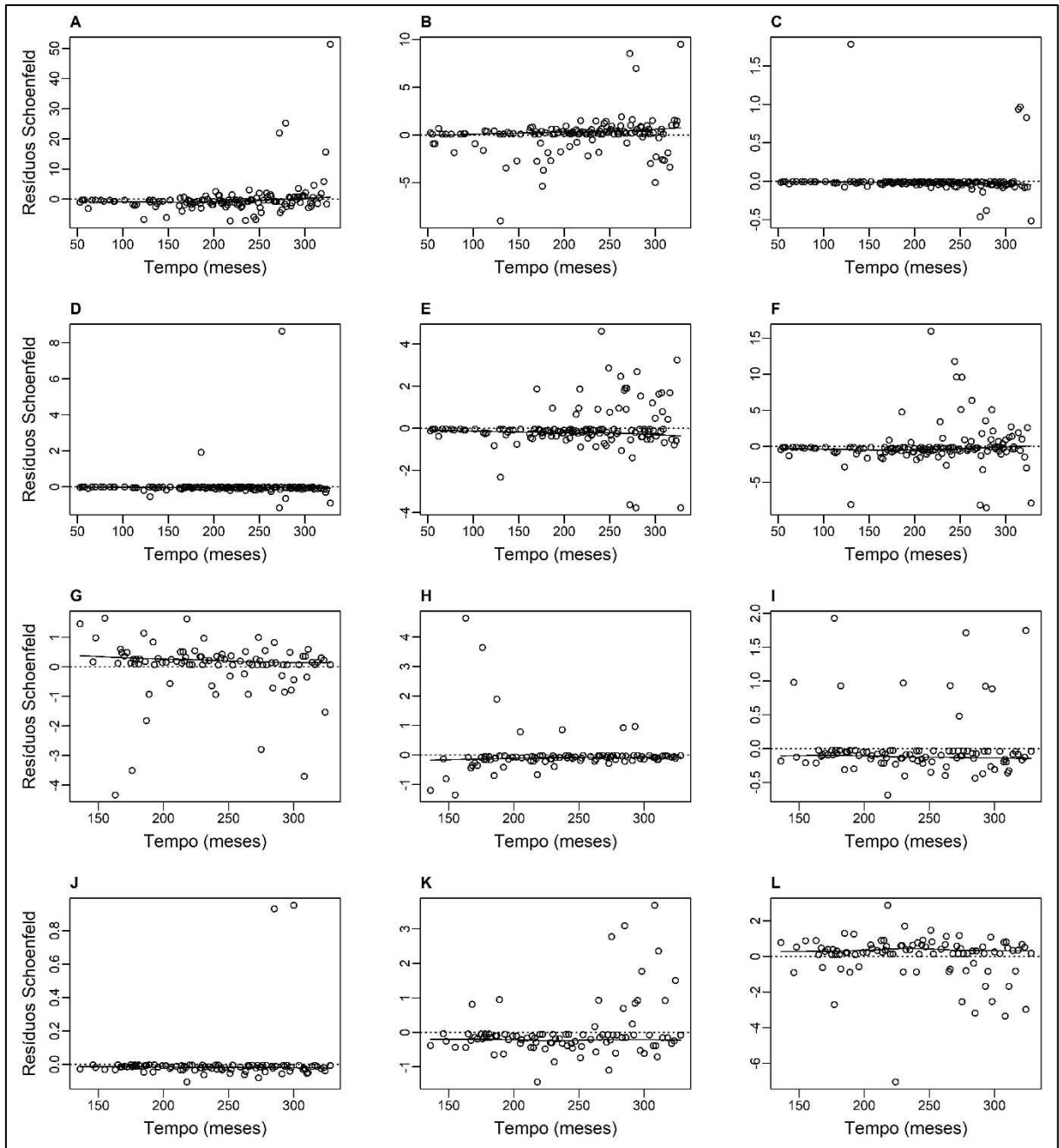
Covariáveis	Definitiva	Expiração	Renúncia	Cancelamento/ Anulação
Ciclo de vida anual				
Transgenia				
Sem transgenia	849	480	628	141
Com transgenia	765	117	411	30
Titular				
Instituição pública	472	328	87	39
Setor privado	1103	258	947	130
Parceria	39	11	5	2
Tipo de cultura				
Agrícola	1322	566	843	135
Forageira	19	2	11	6
Olerícola	177	29	51	30
Ornamental	96	0	134	0
Ciclo de vida perene				
Transgenia				
Sem transgenia	891	147	426	75
Com transgenia	6	0	0	0
Titular				
Instituição pública	282	56	18	11
Setor privado	607	90	392	64
Parceria	8	1	16	0
Tipo de cultura				
Agrícola	182	83	12	5
Florestal	164	6	13	9
Forageira	36	8	2	3
Frutífera	234	22	27	18
Olerícola	16	0	0	1
Ornamental	247	28	351	38
Medicinal e aromática	16	0	0	1

A suposição de proporcionalidade das taxas de falha, exigida pelo modelo de Fine-Gray foi atendida em ambos os ciclos para os eventos de expiração por prazo, renúncia, cancelamento

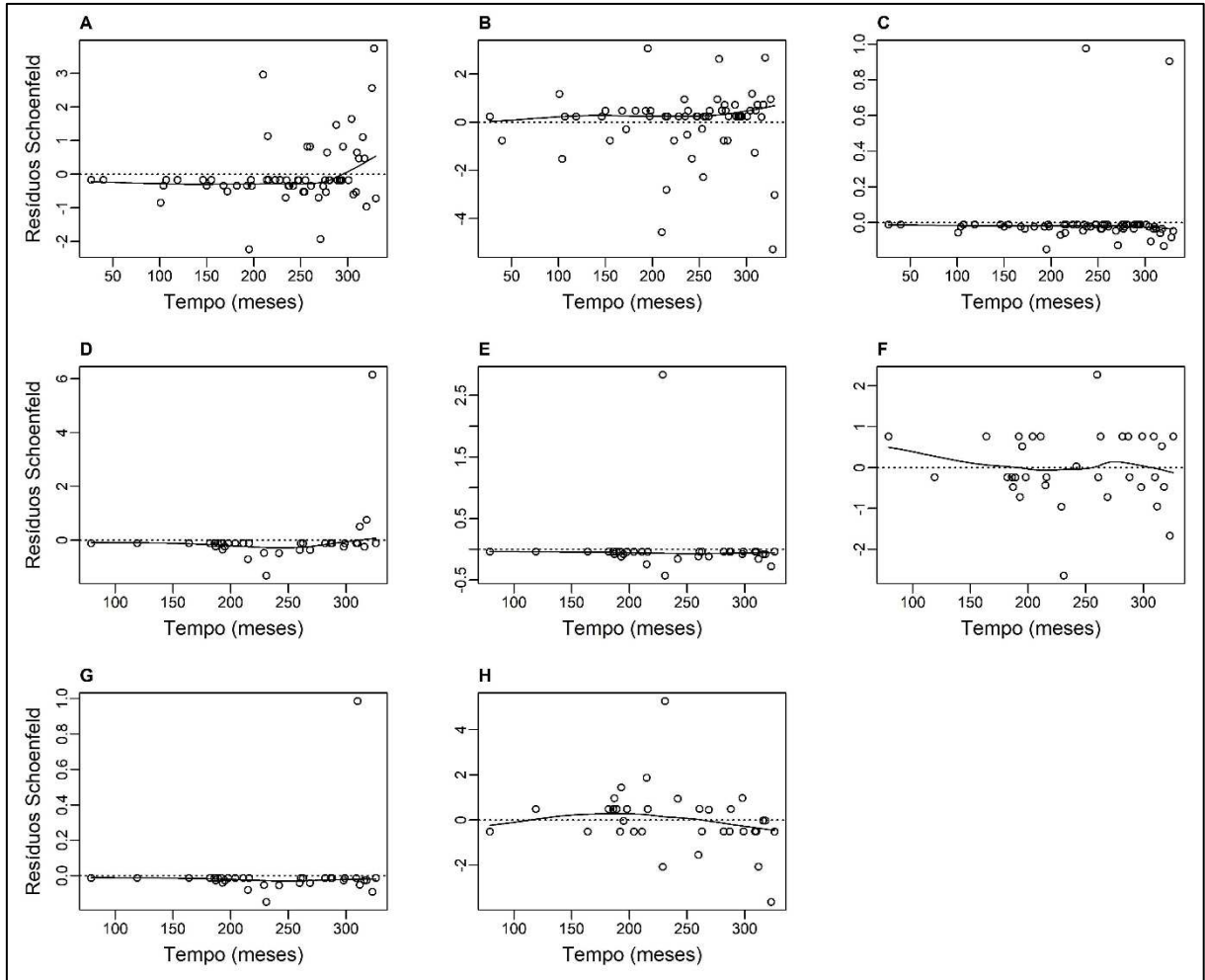
e anulação, conforme evidenciado pelos gráficos de resíduos Schoenfeld (Figuras 1, 2 e 3). As curvas suavizadas oscilaram em torno de zero, indicando que não há evidências de violação da hipótese de proporcionalidade.



**Figura 1** - Avaliação da proporcionalidade do modelo de Fine-Gray para o evento de expiração por prazo para cultivares de ciclo anual e perene: anual e com transgenia (A), anual e setor privado (B), anual e parceria (C), anual e forrageira (D), anual e olerícola (E), perene e florestal (F), perene e forrageira (G), perene e frutífera (H), perene e ornamental (I).



**Figura 2** - Avaliação da proporcionalidade do modelo de Fine-Gray para o evento de renúncia para cultivares de ciclo anual e perene: anual e com transgenia (A), anual e setor privado (B), anual e parceria (C), anual e forrageira (D), anual e olerícola (E), anual e ornamental (F), perene e setor privado (G), perene e parceria (H), perene e florestal (I), perene e forrageira (J), perene e frutífera (K), perene e ornamental (L).



**Figura 3** - Avaliação da proporcionalidade do modelo de Fine-Gray para os eventos de cancelamento e anulação para cultivares de ciclo anual e perene: anual e com transgenia (A), anual e setor privado (B), anual e parceria (C), perene e florestal (D), perene e forrageira (E), perene e frutífera (F), perene e olerícola (G), perene e ornamental (H).

Com o ajuste do modelo de Fine-Gray e do modelo de regressão binária ponderada com ligação log-log complementar para os eventos em estudo, observou-se que, no ciclo anual, a presença de transgenia, em comparação com sua ausência, está associada a uma redução na incidência desses eventos, conforme indicado pelos coeficientes negativos e pelas razões de taxas de falha (RTFs) menores que um (Tabelas 2 e 3). No ciclo perene, apenas cultivares não transgênicas apresentaram proteções que não foram expiradas, renunciadas, canceladas ou anuladas. Diante disso, a análise foi conduzida sem incluir essa covariável.

**Tabela 2** - Valores estimados pelo modelo de Fine-Gray para todos os eventos associados a cultivares de ciclo anual, utilizando ausência de transgenia, instituição pública e tipo de cultura agrícola, como categorias de referência.

Covariável	Estimativa	Erro-padrão	Valor- $p^1$	RTFs <sup>2</sup>	IC 95%
Evento: expiração por prazo					
Transgenia: sim	-1,0147	0,1057	< <b>0,0001</b>	0,3625	(0,2947; 0,4459)
Titular: setor privado	-0,9328	0,0872	< <b>0,0001</b>	0,9395	(0,3317; 0,4668)
Titular: parceria	-0,2820	0,3310	0,3900	0,7543	(0,3942; 1,4431)
Cultura: forrageira	-1,7218	0,7080	<b>0,0150</b>	0,1788	(0,0446; 0,7160)
Cultura: olerícola	-1,0726	0,2049	< <b>0,0001</b>	0,3421	(0,2289; 0,5112)
Evento: renúncia					
Transgenia: sim	-0,4865	0,0713	< <b>0,0001</b>	0,6148	(0,5346; 0,7071)
Titular: setor privado	1,7261	0,1196	< <b>0,0001</b>	5,6189	(4,4443; 7,1040)
Titular: parceria	0,2128	0,4681	0,6500	1,2371	(0,4942; 3,0964)
Cultura: forrageira	-0,3291	0,2951	0,2600	0,7196	(0,4035; 1,2833)
Cultura: olerícola	-0,8590	0,1479	< <b>0,0001</b>	0,4236	(0,3169; 0,5661)
Cultura: ornamental	0,3092	0,1017	<b>0,0020</b>	1,3623	(1,1161; 1,6629)
Eventos: cancelamento e anulação (inicial)					
Transgenia: sim	-1,3347	0,2291	< <b>0,0001</b>	0,2580	(0,1647; 0,4043)
Titular: setor privado	0,6941	0,2057	<b>0,0007</b>	2,0018	(1,3376; 2,9960)
Titular: parceria	-0,4795	0,7565	0,5300	0,6191	(0,1405; 2,7272)
Cultura: forrageira	0,7056	0,4137	0,0880	2,0251	(0,9002; 4,5558)
Cultura: olerícola	0,3518	0,2178	0,1100	1,4217	(0,9277; 2,1788)
Eventos: cancelamento e anulação (final)					
Transgenia: sim	-1,4539	0,2176	< <b>0,0001</b>	0,2337	(0,1525; 0,3580)
Titular: setor privado	0,7519	0,1984	<b>0,0002</b>	2,1210	(1,4376; 3,1290)
Titular: parceria	-0,2943	0,7208	0,6800	0,7451	(0,1814; 3,0600)

<sup>1</sup> Valor- $p$  do teste de Wald; <sup>2</sup> Razão de taxas de falha

No ciclo anual, as cultivares protegidas pelo setor privado apresentaram um aumento na incidência dos eventos de renúncia, cancelamento e anulação com o modelo de Fine-Gray (Tabela 2), bem como com o modelo de regressão binária ponderada com ligação log-log complementar (Tabela 3). Por outro lado, as cultivares também protegidas pelo setor privado demonstraram a maior redução na incidência de expiração por prazo pelos dois modelos.

**Tabela 3** - Valores estimados pelo modelo de regressão binária ponderada com ligação log-log complementar para todos os eventos associados a cultivares de ciclo anual, utilizando ausência de transgenia, instituição pública e tipo de cultura agrícola, como categorias de referência.

Covariável	Estimativa	Erro-padrão	z	Valor- $p^1$
Evento: expiração por prazo				
Intercepto	-6,5878	0,0565	-116,6480	< <b>0,0001</b>
Transgenia: sim	-0,9253	0,1105	-8,3780	< <b>0,0001</b>
Titular: setor privado	-0,8460	0,0886	-9,5450	< <b>0,0001</b>
Titular: parceria	-0,3064	0,3132	-0,9880	0,3232
Cultura: forrageira	-1,5985	0,7089	-2,2550	<b>0,0241</b>
Cultura: olerícola	-0,9553	0,1963	-4,8670	< <b>0,0001</b>
Evento: renúncia				
Intercepto	-8,0361	0,1078	-74,5430	< <b>0,0001</b>
Transgenia: sim	-0,3795	0,0709	-5,3550	< <b>0,0001</b>
Titular: setor privado	1,5926	0,1159	13,7360	< <b>0,0001</b>
Titular: parceria	0,1958	0,4622	0,4240	0,6718
Cultura: forrageira	-0,2808	0,3055	-0,9190	0,3580
Cultura: olerícola	-0,7573	0,1492	-5,0770	< <b>0,0001</b>
Cultura: ornamental	0,2459	0,1008	2,4390	<b>0,0147</b>
Evento: cancelamento e anulação (inicial)				
Intercepto	-8,8783	0,1626	-54,5690	< <b>0,0001</b>
Transgenia: sim	-1,3296	0,2123	-6,2630	< <b>0,0001</b>
Titular: setor privado	0,6753	0,1884	3,5850	< <b>0,0001</b>
Titular: parceria	-0,4737	0,7335	-0,6460	0,5184
Cultura: forrageira	0,7075	0,4211	1,6800	0,0929
Cultura: olerícola	0,3506	0,2111	1,6610	0,0967
Evento: cancelamento e anulação (final)				
Intercepto	-8,8276	0,1604	-55,0390	< <b>0,0001</b>
Transgenia: sim	-1,4281	0,2050	-6,9650	< <b>0,0001</b>
Titular: setor privado	0,7313	0,1860	3,9320	0,0001
Titular: parceria	-0,2920	0,7250	-0,4030	0,6870

<sup>1</sup> Valor- $p$  do teste de Wald

No ciclo perene, os resultados dos modelos indicam que as proteções associadas a parcerias estão relacionadas a um aumento na incidência de renúncia (Tabelas 4 e 5). No entanto, para os eventos de expiração por prazo, cancelamento e anulação, os titulares não apresentaram efeito significativo, o que levou à exclusão dessa covariável da análise.

**Tabela 4** - Valores estimados pelo modelo de Fine-Gray para todos os eventos associados a cultivares de ciclo perene, utilizando ausência de transgenia, instituição pública e tipo de cultura agrícola, como categorias de referência.

Covariável	Estimativa	Erro-padrão	Valor- $p^1$	RTFs <sup>2</sup>	IC 95%
Evento: expiração por prazo (inicial)					
Titular: setor privado	0,3526	0,1902	0,0640	1,4228	(0,9799; 2,0657)
Titular: parceria	-0,2239	1,0487	0,8300	0,7994	(0,1024; 6,2430)
Cultura: florestal	-2,6114	0,4354	< <b>0,0001</b>	0,0734	(0,0313; 0,1724)
Cultura: forrageira	-0,7091	0,3701	0,0550	0,4921	(0,2383; 1,0163)
Cultura: frutífera	-1,5644	0,2393	< <b>0,0001</b>	0,2092	(0,1309; 0,3344)
Cultura: ornamental	-2,3551	0,2408	< <b>0,0001</b>	0,0949	(0,0592; 0,1521)
Evento: expiração por prazo (final)					
Cultura: florestal	-2,4173	0,4222	< <b>0,0001</b>	0,0892	(0,0390; 0,2040)
Cultura: forrageira	-0,6928	0,3695	0,0610	0,5002	(0,2425; 1,0319)
Cultura: frutífera	-1,5563	0,2371	< <b>0,0001</b>	0,2109	(0,1325; 0,3357)
Cultura: ornamental	-2,1702	0,2169	< <b>0,0001</b>	0,1142	(0,0746; 0,1746)
Evento: renúncia					
Titular: setor privado	0,4655	0,2930	0,1100	1,5929	(0,8969; 2,8290)
Titular: parceria	2,4833	0,3365	< <b>0,0001</b>	11,9801	(6,1951; 23,167)
Cultura: florestal	0,2261	0,4475	0,6100	1,2537	(0,5215; 3,0140)
Cultura: forrageira	-0,0390	0,7581	0,9600	0,9617	(0,2177; 4,2500)
Cultura: frutífera	0,5448	0,3628	0,1300	1,7242	(0,8467; 3,5110)
Cultura: ornamental	2,6372	0,3541	< <b>0,0001</b>	13,9746	(6,9809; 27,975)
Eventos: cancelamento e anulação (inicial)					
Titular: setor privado	0,6584	0,3915	0,0930	1,9320	(0,8968; 4,1610)
Cultura: florestal	0,6547	0,6115	0,2800	1,9250	(0,5805; 6,3810)
Cultura: forrageira	1,2513	0,7418	0,0920	3,4950	(0,8165; 14,957)
Cultura: frutífera	1,2678	0,5026	<b>0,0120</b>	3,5530	(1,3269; 9,5150)
Cultura: olerícola	1,3105	1,0619	0,2200	3,7080	(0,4626; 29,721)
Cultura: ornamental	0,8784	0,5400	0,1000	2,4070	(0,8353; 6,9370)
Eventos: cancelamento e anulação (final)					
Cultura: florestal	0,9706	0,5520	0,0790	2,6400	(0,8947; 7,7870)
Cultura: forrageira	1,2701	0,7321	0,0830	3,5610	(0,8480; 14,954)
Cultura: frutífera	1,2691	0,5036	<b>0,0120</b>	3,5580	(1,3259; 9,5460)
Cultura: olerícola	1,2034	1,0786	0,2600	3,3320	(0,4023; 27,588)
Cultura: ornamental	1,1900	0,4731	<b>0,0120</b>	3,2870	(1,3005; 8,3080)

<sup>1</sup> Valor- $p$  do teste de Wald; <sup>2</sup> Razão de taxas de falha

**Tabela 5** - Valores estimados pelo modelo de regressão binária ponderada com ligação log-log complementar para todos os eventos associados a cultivares de ciclo perene, utilizando ausência de transgenia, instituição pública e tipo de cultura agrícola, como categorias de referência.

Covariável	Estimativa	Erro-padrão	z	Valor- <i>p</i> <sup>1</sup>
Evento: expiração por prazo (inicial)				
Intercepto	-7,0846	0,1444	-49,0650	< <b>0,0001</b>
Titular: setor privado	0,2807	0,1879	1,4940	0,1353
Titular: parceria	-0,2586	1,0147	-0,2550	0,7988
Cultura: florestal	-2,4571	0,4331	-5,6740	< <b>0,0001</b>
Cultura: forrageira	-0,6374	0,3702	-1,7220	0,0851
Cultura: frutífera	-1,4498	0,2405	-6,0290	< <b>0,0001</b>
Cultura: ornamental	-2,2016	0,2374	-9,2750	< <b>0,0001</b>
Evento: expiração por prazo (final)				
Intercepto	-6,9557	0,1098	-63,3700	< <b>0,0001</b>
Cultura: florestal	-2,3075	0,4227	-5,4580	< <b>0,0001</b>
Cultura: forrageira	-0,6309	0,3702	-1,7040	0,0884
Cultura: frutífera	-1,4532	0,2398	-6,0600	< <b>0,0001</b>
Cultura: ornamental	-2,0604	0,2185	-9,4280	< <b>0,0001</b>
Evento: renúncia				
Intercepto	-9,2382	0,3347	-27,6040	< <b>0,0001</b>
Titular: setor privado	0,5611	0,2928	1,9170	0,0553
Titular: parceria	1,5853	0,3726	4,2550	< <b>0,0001</b>
Cultura: florestal	0,1899	0,4187	0,4530	0,6502
Cultura: forrageira	-0,0485	0,7635	-0,0640	0,9494
Cultura: frutífera	0,6729	0,3478	1,9350	0,0530
Cultura: ornamental	2,3945	0,3172	7,5480	< <b>0,0001</b>
Evento: cancelamento e anulação (inicial)				
Intercepto	-10,1698	0,5045	-20,1580	< <b>0,0001</b>
Titular: setor privado	0,6544	0,3881	1,6860	0,0918
Cultura: florestal	0,6596	0,5787	1,1400	0,2544
Cultura: forrageira	1,2409	0,7303	1,6990	0,0893
Cultura: frutífera	1,2553	0,5055	2,4830	<b>0,0130</b>
Cultura: olerícola	1,2998	1,0972	1,1850	0,2361
Cultura: ornamental	0,8681	0,4989	1,7400	0,0819
Evento: cancelamento e anulação (inicial)				
Intercepto	-9,8286	0,4472	-21,9800	< <b>0,0001</b>
Cultura: florestal	0,9729	0,5577	1,7440	0,0811
Cultura: forrageira	1,2524	0,7303	1,7150	0,0863
Cultura: frutífera	1,2569	0,5055	2,4860	<b>0,0129</b>
Cultura: olerícola	1,1971	1,0954	1,0930	0,2745
Cultura: ornamental	1,1771	0,4750	2,4780	<b>0,0132</b>

<sup>1</sup> Valor-*p* do teste de Wald

Os resultados obtidos estão alinhados com o estudo de Meng et al. (2020), que analisou como a entrega, a anulação ou a revogação do certificado de proteção são influenciados pelas características do requerente e do tipo de cultura hortícola, utilizando dados da *Canadian Food Inspection Agency*.

De acordo com Meng et al. (2020), instituições públicas geralmente mantêm a proteção de cultivares até o término do prazo estabelecido, uma vez que as receitas geradas por *royalties* são reinvestidas em pesquisa e desenvolvimento de novas variedades. Esse aspecto também é destacado por Degani et al. (2021), que reforçam a importância dos *royalties* obtidos com as cultivares protegidas.

Por outro lado, empresas privadas, orientadas por fins lucrativos, enfrentam maior pressão econômica para desenvolver novas variedades. A decisão de renunciar à proteção de uma cultivar no setor privado está associada à necessidade de se manterem competitivas, já que precisam constantemente reduzir custos e diferenciar-se no mercado (Freitas & Albrecht, 2015).

Esses fatores, como a perda de interesse comercial e a substituição por variedades mais rentáveis, explicam o resultado de que as cultivares protegidas pelo setor privado apresentam menor incidência de expiração por prazo, mas maior incidência de renúncia, cancelamento e anulação da proteção.

A Tabela 6 apresenta a quantidade de proteções de cultivares renunciadas no período de 2018 a 2022. Observa-se que, em 2020, houve um aumento expressivo, com 253 renúncias, ano marcado pela pandemia de COVID-19.

**Tabela 6** - Proteções renunciadas no período de 2018-2022.

Ano	2018	2019	2020	2021	2022
Proteções renunciadas	55	72	257	91	74

O pico observado em 2020 pode estar associado aos impactos negativos da pandemia no agronegócio, como destacado por Schneider et al. (2020), Silva et al. (2021) e Pinheiro et al. (2022). Entre esses impactos, destacam-se o aumento nos preços de insumos e equipamentos, decorrente do fechamento de portos e da valorização do dólar, além das dificuldades relacionadas à mão de obra no campo.

Em relação aos tipos de cultura, no ciclo anual, as cultivares olerícolas apresentaram a maior redução na incidência de renúncia, enquanto as ornamentais registraram o maior aumento, em comparação com os demais tipos (Tabelas 2 e 3). As cultivares forrageiras, por

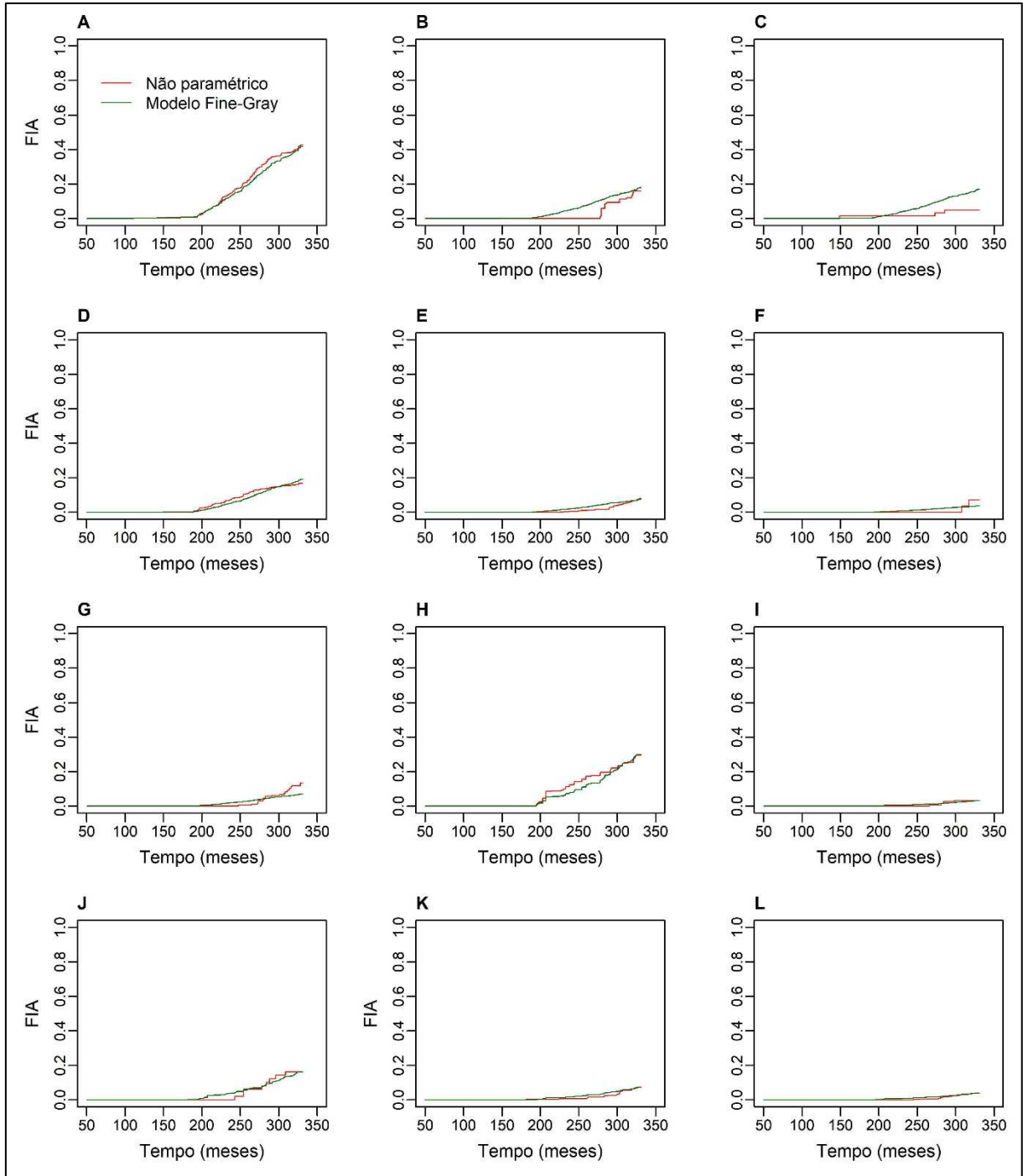
sua vez, estão associadas à maior redução na incidência de expiração por prazo, em comparação com as cultivares agrícolas e olerícolas. Entretanto, para os eventos de cancelamento e anulação, não foi observado efeito significativo entre os tipos de cultura, resultando na exclusão dessa covariável do modelo.

No ciclo perene, as cultivares forrageiras apresentaram a menor redução na incidência de expiração da proteção por término de prazo e um aumento na incidência dos eventos de cancelamento e anulação (Tabelas 5 e 6). Já as cultivares ornamentais tiveram o maior aumento na incidência de renúncia, enquanto as forrageiras registraram a maior redução.

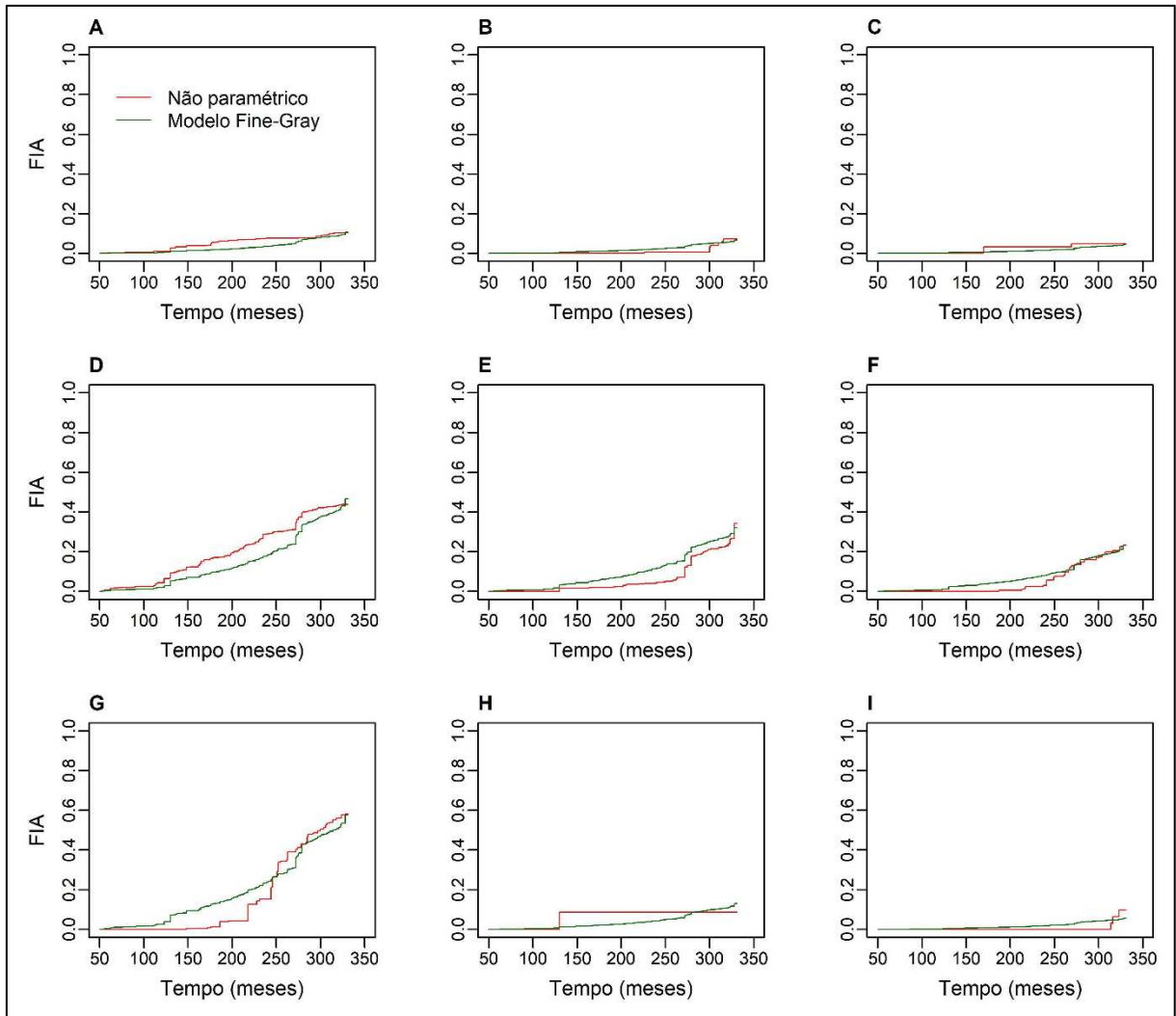
De acordo com Vilela et al. (2016), a intensificação da produção animal em sistemas de pastagem, impulsionada pela necessidade de maior produtividade e rentabilidade na pecuária leiteira, exige a adoção de forrageiras mais produtivas e adaptadas às condições específicas de cada região. Essa dinâmica do setor pode explicar, em parte, o aumento no cancelamento da proteção de cultivares forrageiras, bem como a redução no número de cultivares que atingem o término do prazo de vigência da proteção.

O estudo de Meng et al. (2020) destaca que as cultivares ornamentais apresentam a menor incidência de manutenção da proteção, um fato atribuído às rápidas inovações tecnológicas no setor. Esse resultado está alinhado com os deste trabalho, que também indicam uma maior incidência de renúncia à proteção entre cultivares ornamentais de ciclo anual e perene. Nesse contexto, Neitzke et al. (2016) enfatizam que um dos principais desafios para os melhoristas de plantas ornamentais é desenvolver cultivares que atendam às exigências do mercado de floricultura em termos de qualidade, durabilidade e frescor (Junqueira & Peetz, 2008).

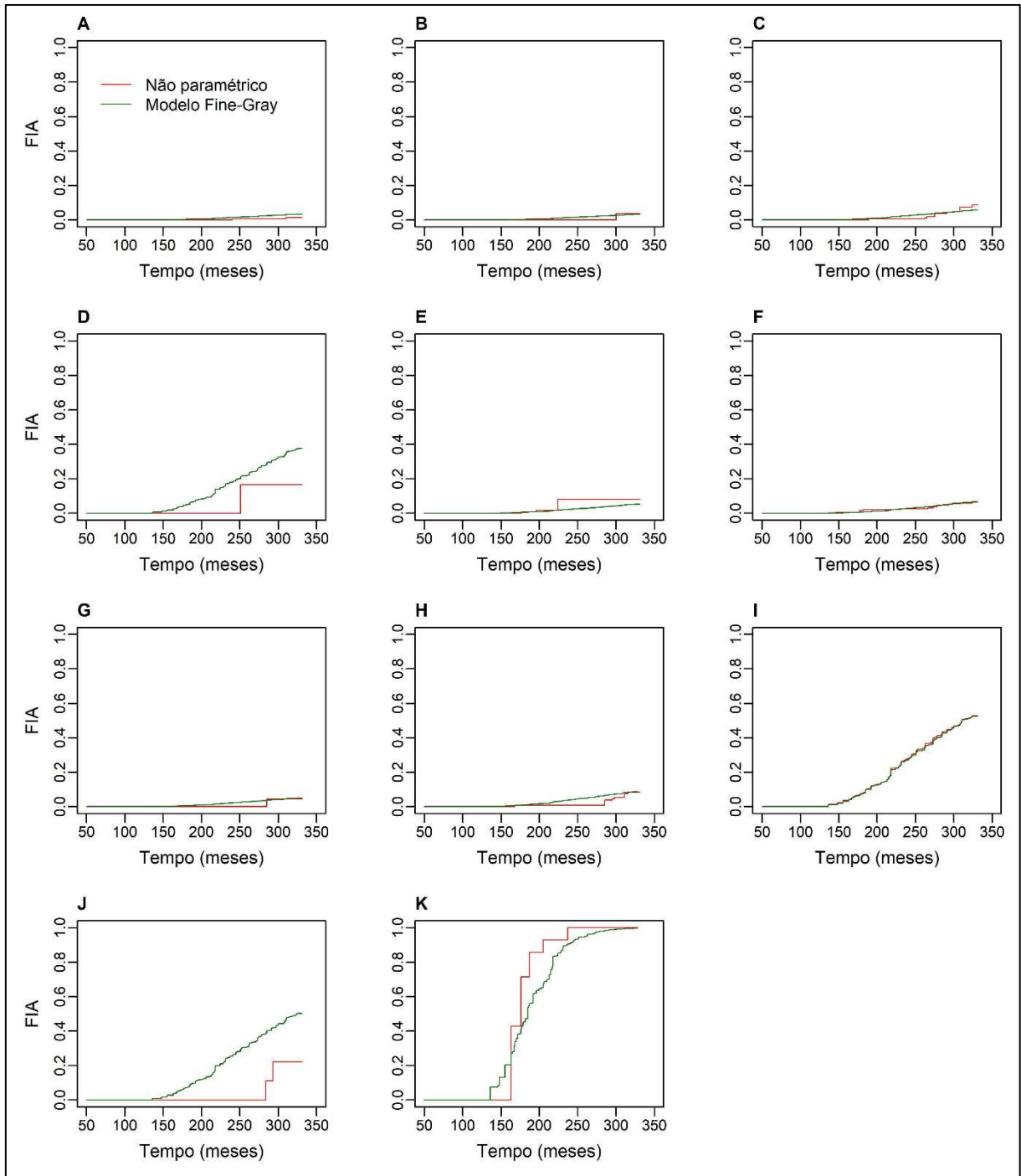
A proximidade entre as FIAs obtidas pelo estimador não paramétrico e as preditas pelo modelo, considerando subconjuntos formados pelas combinações das categorias das covariáveis (Figuras 4, 5, 6 e 7), indicou um ajuste satisfatório do modelo de Fine-Gray para os eventos. Nos casos em que as curvas apresentaram divergências, foi devido ao fato de que as combinações das categorias das covariáveis possuíam um número reduzido de observações.



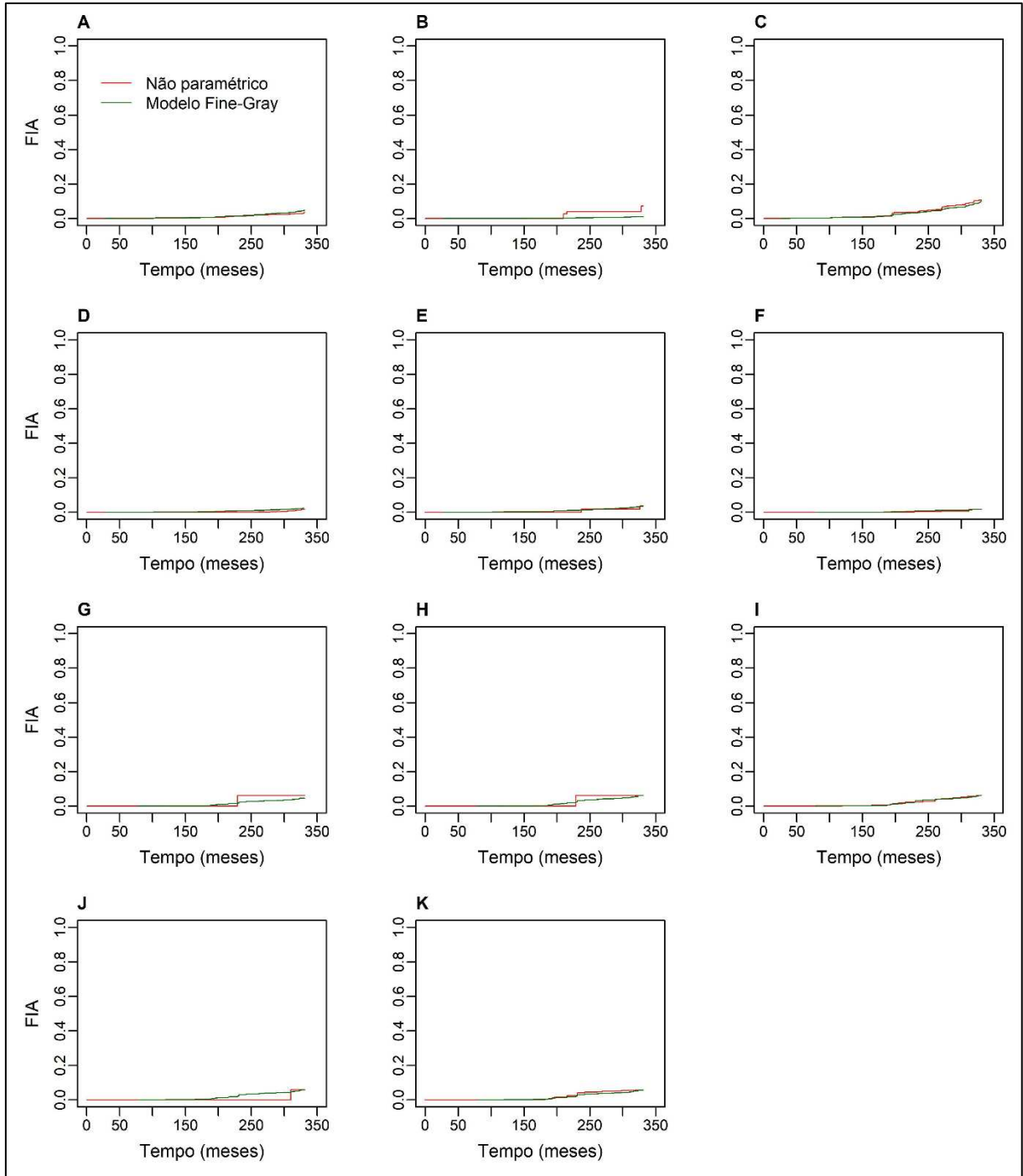
**Figura 4** - Curvas de incidência acumulada estimadas pelo método não paramétrico e pelo modelo de Fine-Gray para o evento de expiração por prazo: anual, instituição pública, sem transgenia e agrícola (A); anual, com transgenia, instituição pública e agrícola (B); anual, sem transgenia, instituição pública e olerícola (C); anual, sem transgenia, setor privado e agrícola (D); anual, com transgenia, setor privado e agrícola (E); anual, sem transgenia, setor privado e forrageira (F); anual, sem transgenia, setor privado e olerícola (G); perene e agrícola (H); perene e florestal (I); perene e forrageira (J); perene e frutífera (K) ; perene e ornamental (L).



**Figura 5** - Curvas de incidência acumulada estimadas pelo método não paramétrico e pelo modelo de Fine-Gray para o evento de renúncia no ciclo anual: sem transgenia, instituição pública e agrícola (A); com transgenia, instituição pública e agrícola (B); sem transgenia, instituição pública e olerícola (C); sem transgenia, setor privado e agrícola (D); com transgenia, setor privado e agrícola (E); sem transgenia, setor privado e olerícola (F); sem transgenia, setor privado e ornamental (G); sem transgenia, parceria e agrícola (H); sem transgenia, parceria e olerícola (I).

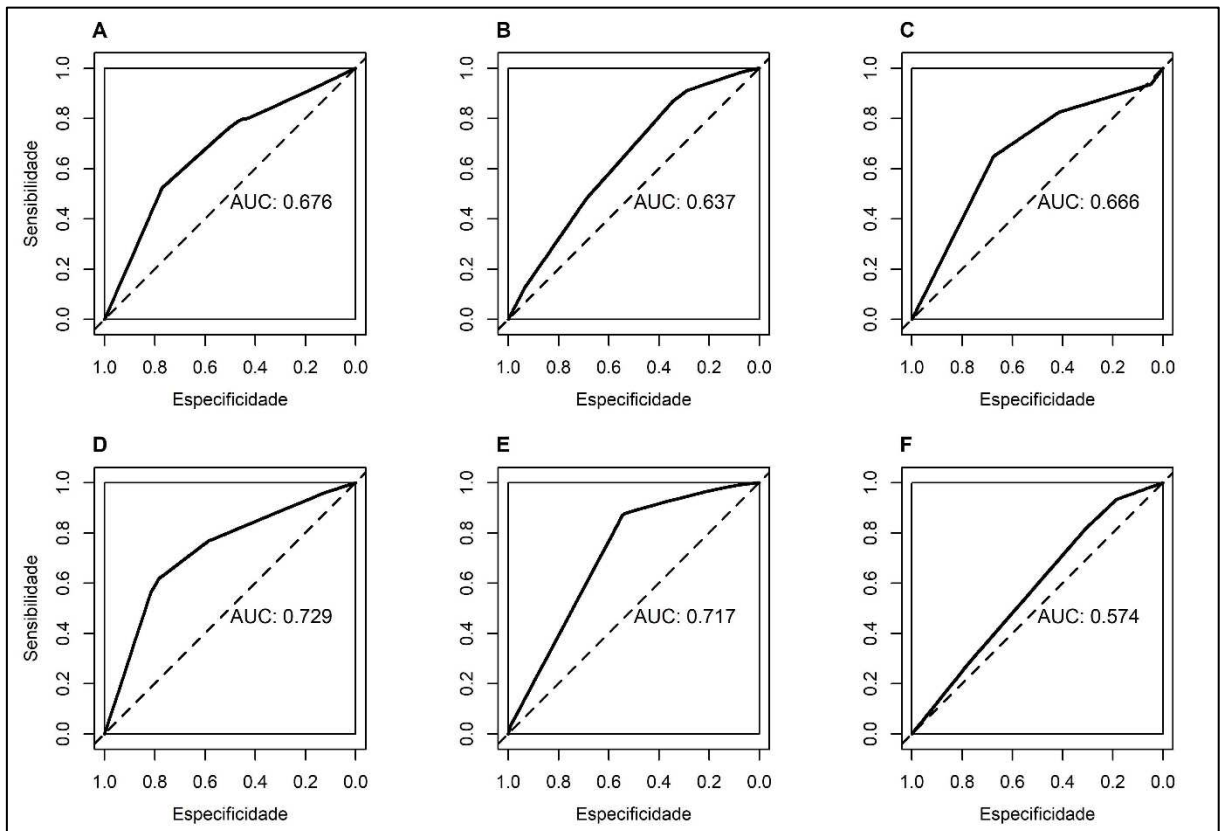


**Figura 6** - Curvas de incidência acumulada estimadas pelo método não paramétrico e pelo modelo de Fine-Gray para o evento de renúncia no ciclo perene: instituição pública e agrícola (A), instituição pública e forrageira (B), instituição pública e frutífera (C), instituição pública e ornamental (D), setor privado e agrícola (E), setor privado e florestal (F), setor privado e forrageira (G), setor privado e frutífera (H), setor privado e ornamental (I), parceria e frutífera (J), parceria e ornamental (K).



**Figura 7** - Curvas de incidência acumulada estimadas pelo método não paramétrico e pelo modelo de Fine-Gray para os eventos de cancelamento e anulação nos ciclos anual e perene: anual, instituição pública e sem transgenia (A); anual, com transgenia e instituição pública (B); anual, sem transgenia e setor privado (C); anual, com transgenia e setor privado (D); anual, sem transgenia e parceria (E); perene e agrícola (F); perene e florestal (G); perene e forrageira (H); perene e frutífera (I); perene e olerícola (J); perene e ornamental (K).

Na Figura 8, são apresentadas as curvas ROC com o objetivo de avaliar o desempenho do modelo de regressão binária ponderada com ligação log-log complementar. Embora as estimativas dos parâmetros sejam relativamente próximas às do modelo de Fine-Gray, a área sob a curva (AUC) para os eventos de expiração por prazo e renúncia no ciclo perene foram de 0,729 e 0,717 respectivamente, indicando um desempenho aceitável do modelo. No entanto, nas demais curvas, os valores de AUC são inferiores a 0,7, o que corresponde a um desempenho insuficiente.



**Figura 8** - Curvas ROC mostrando a área abaixo da curva (AUC), estimadas pelo modelo de regressão binária ponderada com ligação log-log complementar para os eventos de: expiração por prazo no ciclo anual (A); renúncia no ciclo anual (B); cancelamento e anulação no ciclo anual (C); expiração por prazo no ciclo perene (D); renúncia no ciclo perene (E); cancelamento e anulação no ciclo perene (F).

Outra medida que demonstra a melhor adequação do modelo de Fine-Gray para analisar a duração do certificado de proteção de cultivares é o Critério de Informação de Akaike (AIC) (Akaike, 1974), sendo desejáveis valores menores de AIC. O modelo de Fine-Gray apresentou

os menores valores de AIC em todas as análises, exceto na modelagem do evento de renúncia no ciclo anual (Tabela 7).

**Tabela 7** - Critério de informação de Akaike (AIC) dos modelos de regressão avaliados para os dados referentes à duração das proteções de cultivares.

Eventos	Modelo de regressão de Fine-Gray		Modelo de regressão binária ponderada	
	anual	perene	anual	perene
Expiração por prazo	9173,9813	2009,6739	9793,4704	2566,0961
Renúncia	16154,9509	5634,3149	16102,1229	6423,7433
Cancelamento e anulação	2688,7357	1095,1981	3265,7416	1471,8116

De forma geral, os resultados pelo modelo de Fine-Gray indicaram que a presença de transgenia nas cultivares anuais reduz a incidência de expiração da proteção por término de prazo, bem como de renúncia, cancelamento e anulação. As cultivares protegidas pelo setor privado apresentam menor incidência de expiração por prazo, mas, em contrapartida, possuem maior incidência de renúncia, cancelamento e anulação. As cultivares forrageiras estão associados a uma menor redução incidência de expiração da proteção no ciclo perene e a uma maior redução dessa expiração no ciclo anual. Ainda no ciclo perene, cultivares forrageiras apresentam uma redução na incidência de renúncia e um aumento de cancelamento e anulação. Tanto no ciclo anual quanto no perene, cultivares ornamentais apresentam a maior taxa de renúncia.

### Conclusões

1. O modelo de riscos competitivos em tempo contínuo mostra-se promissor para modelar a duração do certificado de proteção de cultivares.
2. A análise indica que a transgenia, o tipo de titularidade e de cultura influenciam significativamente a dinâmica de manutenção da proteção de cultivares.

### Agradecimentos

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES - 001) pelo apoio financeiro deste estudo.

## Referências

- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v.19, p.716–723, 1974. DOI: 10.1109/TAC.1974.1100705
- ASSANE, C. C.; PEREIRA, B. D. B.; MUAUALO, M. A. M. Modelos de riscos competitivos para a análise de sobrevivência de pacientes com hipertensão arterial em tratamento no hospital Universitário, HUCFF/UFRJ. **Latin American Journal of Development**, v.2, p.488–504, 2021. DOI: 10.46814/lajdv2n6-013
- AUSTIN, P. C.; FINE, J. P. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. **Statistics in Medicine**, v.36, p.4391–4400, 2017. DOI: 10.1002/sim.7501
- AUSTIN, P. C.; LEE, D. S.; FINE, J. P. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. **Circulation**, v.133, p. 01–609, 2016. DOI: 10.1161/CIRCULATIONAHA.115.017719
- BERGER, M.; SCHMID, M.; WELCHOWSKI, T.; SCHMITZ-VALCKENBERG, S.; BEYERSMANN, J. Subdistribution hazard models for competing risks in discrete time. **Biostatistics**, v.21, p.449–466, 2020. DOI: 10.1093/biostatistics/kxy069
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Informações aos Usuários do SNPC**. 2023. Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/insumos-agropecuarios/insumos-agricolas/protecao-de-cultivar/informacoes-publicacoes/informacoes-aos-usuarios-do-snpc-fevereiro-de-2020>>. Acesso em: 22 set. 2024.
- BRASIL. Lei nº 9.456, de 25 de abril de 1997. Institui a Lei de Proteção de Cultivares e dá outras providências. Diário Oficial da União, Brasília, DF, 28 abr. 1997. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/leis/l9456.htm](https://www.planalto.gov.br/ccivil_03/leis/l9456.htm)>. Acesso em: 22 set. 2024.
- BULSING, A. C.; AVIANI, D. S.; PACHECO, L. G. A.; MACHADO, R. Z. Proteção de cultivares. In Brasil. Ministério da Agricultura Pecuária e Abastecimento (ed.) **Curso de propriedade intelectual & inovação no agronegócio**. MAPA/EaD/UFSC, Brasília e Florianópolis, 2010.
- CARVALHO, S. I. D.; BIANCHETTI, L. D. B.; REIFSCHNEIDER, F. J. Registro e proteção de cultivares pelo setor público: a experiência do programa de melhoramento de Capsicum da Embrapa Hortaliças. **Horticultura Brasileira**, v.27, p.135–138, 2009. DOI: 10.1590/S0102-05362009000200002
- CARVALHO, M. S.; ANDREOZZI, V. L.; CODEÇO, C. T.; CAMPOS, D. P.; BARBOSA, M. T. S.; SHIMAKURA, S. E. Análise de sobrevivência: teoria e aplicações em saúde. Fiocruz Editora, 2011.
- CASTRO, E. D. C.; PACHECO, A.; SOUZA, C. B.; WANDER, A. Estruturas de governança de transações das organizações detentoras de direitos de proteção em cultivares de feijão-comum no Brasil. **Revista Razão Contábil e Finanças**, v.12, n.2, p.1-22, 2021.
- COLOSIMO, E. A.; GIOLO, S. R. **Análise de sobrevivência aplicada**. 2. ed. São Paulo: Editora Edgard Blucher, 2024. 380p.

DEGANI, P. C.; LEITNER, C. P. S.; BAGGENSTOSS, S.; TORKOMIAN, A. L. V.; ALVES FILHO, A. G. Inovação em cultivares: panorama das proteções e registros da soja no Brasil. **Navus - Revista de Gestão e Tecnologia**, v.11, p.01–16, 2021. DOI: 10.22279/navus.2021.v11.p01-16.1524

FREITAS, F. D. S.; ALBRECHT, A. J. P. Estudo do Mercado das Mantenedoras das Cultivares de Soja no Estado do Mato Grosso. **Revista IPecege**, v.1, p.7–25, 2015. DOI: 10.22167/r.ipecege.2015.3-4.7

GENNA, N. G.; KANE, M. E.; PÉREZ, H. E. Simultaneous assessment of germination and infection dose-responses in fungicide-treated seeds with non- and semiparametric statistical methods. **Seed Science and Technology**, v.43, p.168–186, 2015. DOI: 10.15258/sst.2015.43.2.13

GLENN, K. C.; ALSOP, B.; BELL, E.; GOLEY, M.; JENKINSON, J.; LIU, B.; MARTIN, C.; PARROTT, W.; SOUDER, C.; SPARKS, O.; URQUHART, W.; WARD, J. M.; VICINI, J. L. Bringing new plant varieties to market: Plant breeding and selection practices advance beneficial characteristics while minimizing unintended changes. **Crop Science Society of America**, v.57, p.2906–2921, 2017. DOI: 10.2135/cropsci2017.03.0199

GRAY, B. **cmprsk: Subdistribution Analysis of Competing Risks**. 2024. Disponível em: <<https://CRAN.R-project.org/package=cmprsk>>. Acesso em: 24 set. 2024.

HOSMER J. R., D. W.; LEMESHOW, S.; STURDIVANT, R. X. **Applied Logistic Regression**. 3. ed. Hoboken, John Wiley & Sons, 2013.

INTERNATIONAL SEED FEDERATION (ISF); ASSOCIAÇÃO BRASILEIRA DE SEMENTES E MUDAS (ABRASEM). **Guia de Propriedade Intelectual**. 2021. Disponível em: <<https://worldseed.org/document/isf-ip-guide-2021/guia-de-propriedade-intelectual/>>. Acesso em: 09 fev. 2025.

JUNQUEIRA, A. H.; PEETZ, M. S. Mercado interno para os produtos da floricultura brasileira: características, tendências e importância sócio-econômica recente. **Revista Brasileira de Horticultura Ornamental**, v.14, n.1, 2008. DOI: 10.14295/rbho.v14i1.230

MENG, T.; CAREW, R.; FLORKOWSKI, W. J. Determinants of the grant lag and the surrender lag of horticultural crop plant breeders' rights applications: Survival analysis with competing risks. **Canadian Journal of Agricultural Economics/Revue canadienne d'agroéconomie**, v.68, p.489–512, 2020. DOI: 10.1111/cjag.12257

MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO. **CultivarWeb**. 2024. Disponível em: <<https://sistemas.agricultura.gov.br/snpc/cultivarweb/>>. Acesso em: 25 set. 2024.

NEITZKE, R. S.; FISCHER, S. Z.; VASCONCELOS, C. S.; BARBIERI, R. L.; TREPTOW, R. O. Pimentas ornamentais: aceitação e preferências do público consumidor. **Horticultura Brasileira**, v.34, p.102–109, 2016. DOI: 10.1590/S0102-053620160000100015

PINHEIRO, Y. A.; KONDA, S. T.; BONINI, L. M. D. M. Impactos da pandemia COVID-19 na importação de fertilizantes para o agronegócio brasileiro. In: CARVALHO, A. C.; CASTRO, A. C. (Ed.). **Implicações Socioeconômicas da COVID-19 no Brasil e no Mundo**. Editora Científica Digital, 2022. p.148-156. DOI: 10.37885/211006353

PINTILIE, M. **Competing risks: a practical perspective**. Chichester, England; Hoboken, NJ: John Wiley & Sons, 2006. 224p.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, 2025. Disponível em: <<https://www.r-project.org/>>.

SANTOS, F. S.; AVIANI, D. M.; HIDALGO, J. A. F.; MACHADO, R. Z.; ARAÚJO, S. P. Evolution, importance and evaluation of cultivar protection in Brazil: the work of the SNP. **Crop Breeding and Applied Biotechnology**, v.12, p.99–110, 2012. DOI: 10.1590/S1984-70332012000500011

SCHNEIDER, S.; CASSOL, A.; LEONARDI, A.; MARINHO, M. M. Os efeitos da pandemia da Covid-19 sobre o agronegócio e a alimentação. **Estudos Avançados**, v.34, p.167–188, 2020. DOI: 10.1590/s0103-4014.2020.34100.011

SCHMID, M.; BERGER, M. Competing risks analysis for discrete time-to-event data. **WIREs Computational Statistics**, v.13, p.e1529, 2021. DOI: 10.1002/wics.1529

SCHUSTER, N. A.; HOOGENDIJK, E. O.; KOK, A. A. L.; TWISK, J. W. R.; HEYMANS, M. W. Ignoring competing events in the analysis of survival data may lead to biased results: a nonmathematical illustration of competing risk analysis. **Journal of Clinical Epidemiology**, v.122, p.42–48, 2020. DOI: 10.1016/j.jclinepi.2020.03.004

SILVA, M. C. M.; RODRIGUES, J. M. A.; YAMASHITA, O. M. Impacto da pandemia de COVID-19 no agronegócio brasileiro. **Colloquium Socialis**, v. 5, p. 63–70, 2021. DOI: 10.5747/cs.2021.v5.n1.s131

TOWE, C. A. **A Competing Risks Model of Land Use Change**. Em: Agricultural and Applied Economics Association. Pensilvânia, 2011. Disponível em: <<https://ageconsearch.umn.edu/record/103725>>. Acesso em: 25 jul. 2024.

VILELA, D.; FERREIRA, R. P.; FERNANDES, E. N.; JUNTOLLI, F. V. **Pecuária de leite no Brasil: cenários e avanços tecnológicos**. 1. ed. Brasília, DF: Embrapa, 2016. 435p.

WEI, S.; TSENG, Y. P.; ZELENÝ, D. The ecological characteristics of the safe sites for early-stage establishment of *Chamaecyparis obtusa* var. *formosana* seedlings in Taiwan. **Forest Ecology and Management**, 555, 121714, 2024. DOI: 10.1016/j.foreco.2024.121714

WELCHOWSKI, T.; BERGER, M.; KOEHLER, D.; SCHIMID, M. **discSurv: Discrete Time Survival Analysis**. 2022. Disponível em: <<https://CRAN.R-project.org/package=discSurv>>. Acesso em: 25 jul. 2024.

YU, J. K.; CHUNG, Y.-S. Plant Variety Protection: Current Practices and Insights. **Genes**, v.12, p.1127, 2021. DOI: 10.3390/genes12081127

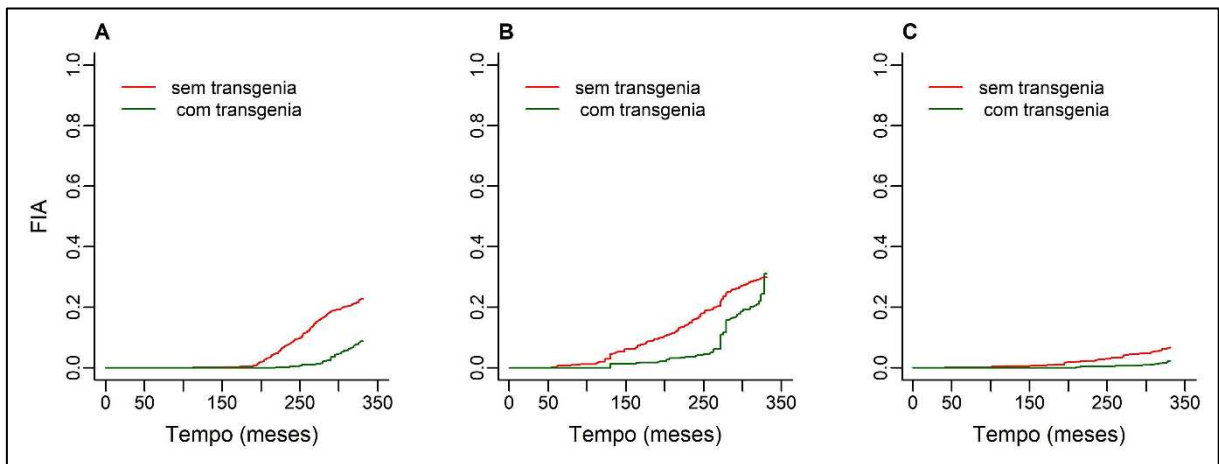
**APÊNDICES**

APÊNDICE A – Estimador não paramétrico da função de incidência acumulada.....	59
APÊNDICE B – Rotina do modelo de Fine-Gray para o evento de expiração por prazo.....	64
APÊNDICE C – Rotina do modelo de regressão binária ponderada para os eventos de expiração por prazo, renúncia, cancelamento e anulação.....	71
APÊNDICE D – Rotina do estimador não paramétrico da função de incidência acumulada para os eventos de expiração por prazo, renúncia, cancelamento e anulação.....	75

## APÊNDICE A – Estimador não paramétrico da função de incidência acumulada

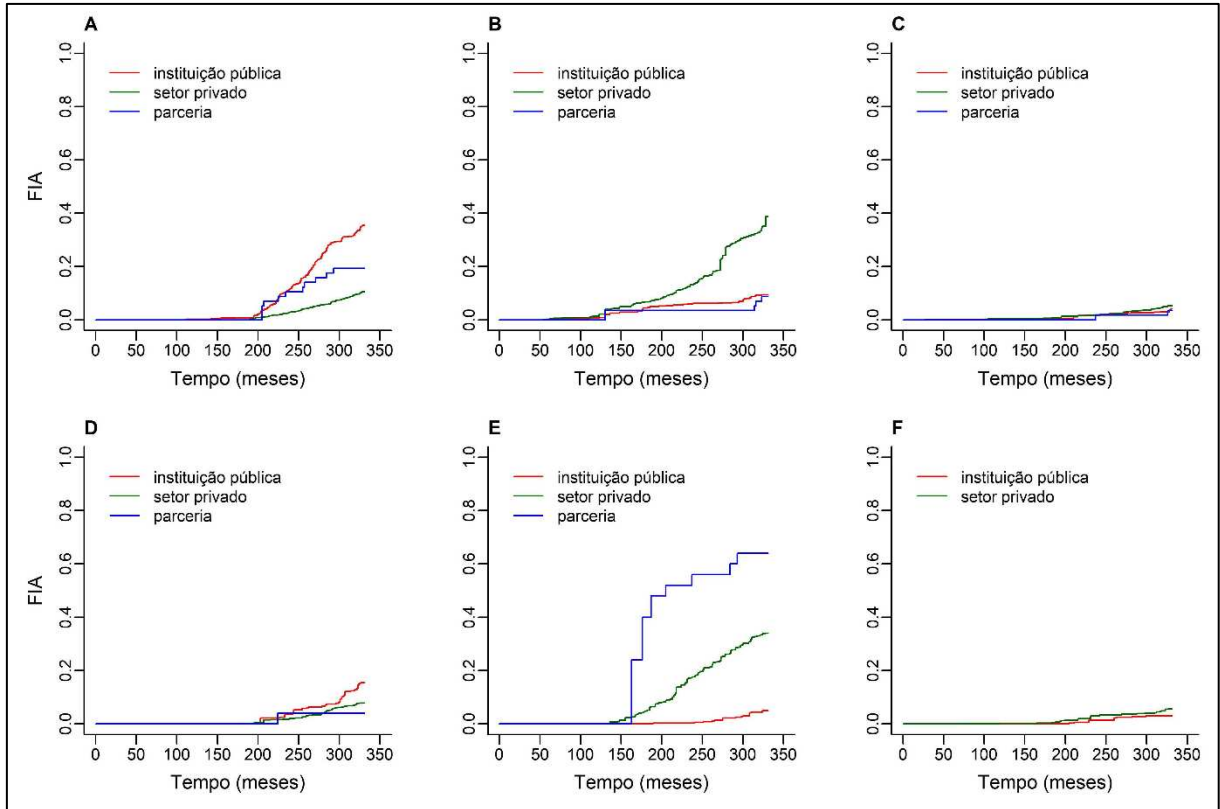
Neste apêndice, são apresentados os procedimentos referentes ao método não paramétrico descrito na subseção 1.4.1. do Capítulo 1. Dessa forma, são apresentadas as curvas de incidência acumulada associadas às covariáveis (transgenia, titular e tipo de cultura) para cada causa de falha (expiração por prazo, renúncia, cancelamento e anulação) nos ciclos de vida anual e perene.

A Figura 9 apresenta as curvas de incidência acumulada para cultivares com e sem transgenia no ciclo de vida anual. Observa-se que cultivares não transgênicas possuem maior incidência de expiração por prazo, cancelamento e anulação. Para o evento de renúncia, não houve rejeição da hipótese de igualdade das FIAs (Tabela 8).



**Figura 9** - Curvas de incidência acumulada no ciclo anual associadas à covariável transgenia estimadas a partir do método não paramétrico para os eventos: expiração por prazo (A), renúncia (B), cancelamento e anulação (C).

A Figura 10 apresenta as curvas de incidência acumulada para a covariável titular. O teste de Gray confirma diferenças significativas entre as curvas, com exceção dos eventos de cancelamento e anulação (Tabelas 8 e 9). No ciclo anual, instituições públicas apresentam a maior incidência de proteção até o término do prazo. No entanto, cultivares protegidas pelo setor privado exibem a maior incidência de renúncia à proteção.

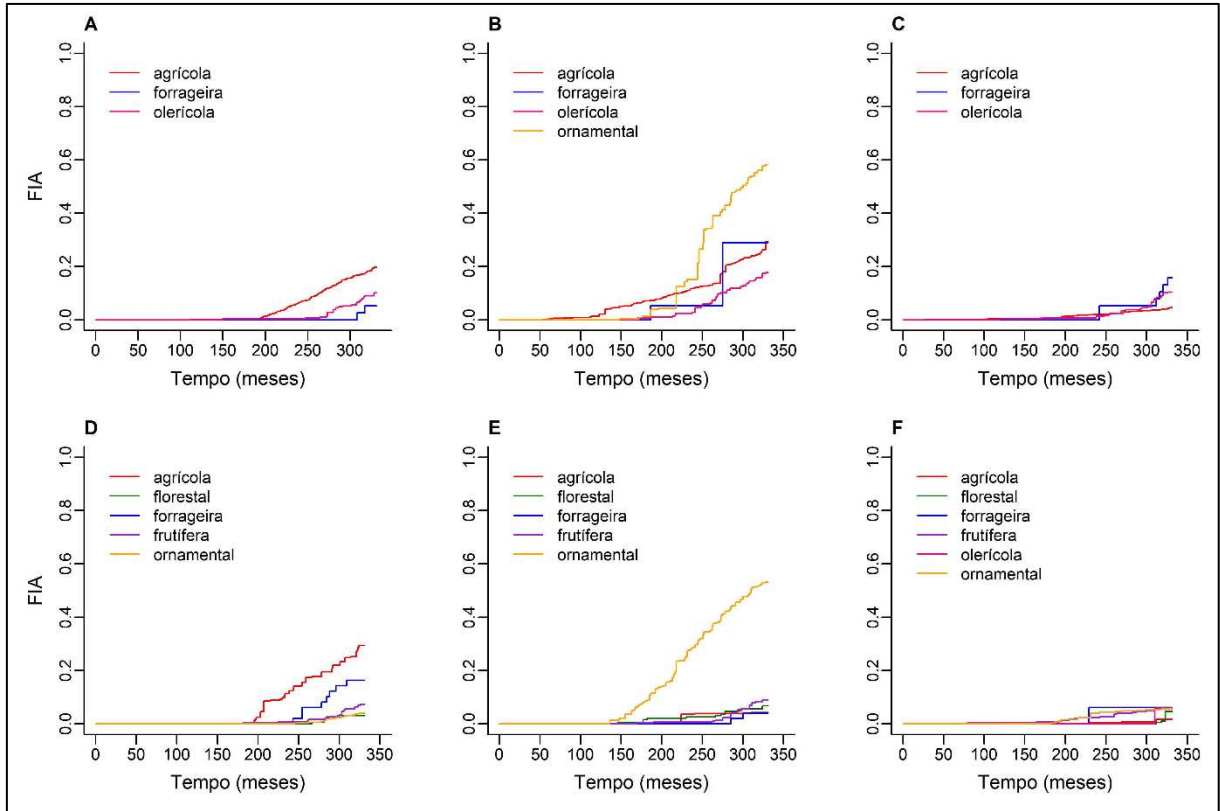


**Figura 10** - Curvas de incidência acumulada no associadas à covariável titular estimadas a partir do método não paramétrico para: expiração por prazo no ciclo anual (A), renúncia no ciclo anual (B), cancelamento e anulação no ciclo anual (C), expiração por prazo no ciclo perene (D), renúncia no ciclo perene (E), cancelamento e anulação no ciclo perene (F).

No ciclo perene, cultivares protegidas pelo setor privado registram a maior incidência de renúncia à proteção em cerca dos primeiros 160 meses. Após esse período, cultivares provenientes de parcerias entre o setor privado e instituições públicas passam a apresentar a maior incidência, enquanto instituições públicas mantêm a menor incidência de renúncia.

A Figura 11 exhibe as curvas de incidência acumulada associadas à covariável tipo de cultura. No ciclo anual, cultivares ornamentais apresentaram a maior incidência de renúncia, enquanto cultivares agrícolas destacaram-se por apresentar maior incidência de expiração por prazo e menor incidência de cancelamento e anulação da proteção (Tabela 8).

De forma semelhante ao ciclo anual, cultivares perenes destinadas a fins ornamentais também apresentaram a maior incidência de renúncia. Por outro lado, cultivares dos tipos florestal, frutífera e ornamental registraram a menor incidência de expiração por prazo. Os eventos de cancelamento e anulação apresentaram baixa probabilidade de ocorrência, sem diferenças significativas entre os tipos analisados (Tabela 9).



**Figura 11** - Curvas de incidência acumulada no associadas à covariável tipo de cultura estimadas a partir do método não paramétrico para: expiração por prazo no ciclo anual (A), renúncia no ciclo anual (B), cancelamento e anulação no ciclo anual (C), expiração por prazo no ciclo perene (D, renúncia no ciclo perene (E), cancelamento e anulação no ciclo perene (F).

**Tabela 8** - Teste de Gray para a comparação das curvas de incidência acumulada no ciclo anual estimadas pela técnica não paramétrica.

Comparações	Expiração		Renúncia		Cancelamento/ Anulação	
	estatística	valor- <i>p</i>	estatística	valor- <i>p</i>	estatística	valor- <i>p</i>
Transgenia						
sem vs com	120,0300	<b>&lt;0,0001</b>	1,1390	0,2858	34,3470	<b>&lt;0,0001</b>
Titular						
todas são iguais	297,2320	<b>&lt;0,0001</b>	270,3160	<b>&lt;0,0001</b>	2,0760	0,3542
pública vs privado	297,8570	<b>&lt;0,0001</b>	255,4370	<b>&lt;0,0001</b>	1,7880	0,1812
pública vs parceria	5,3068	<b>0,0212</b>	0,0361	0,8493	0,0671	0,7957
privado vs parceria	4,8829	<b>0,0271</b>	19,7999	<b>&lt;0,0001</b>	0,3811	0,5370
Tipo de cultura						
todas são iguais	74,2700	<b>&lt;0,0001</b>	113,8900	<b>&lt;0,0001</b>	39,2200	<b>&lt;0,0001</b>
agrícola vs forrageira	5,1524	<b>0,0232</b>	0,0067	0,9345	9,6972	<b>0,0018</b>
agrícola vs olerícola	17,0800	<b>0,0001</b>	17,7500	<b>&lt;0,0001</b>	16,9100	<b>&lt;0,0001</b>
agrícola vs ornamental	54,4300	<b>&lt;0,0001</b>	87,0700	<b>&lt;0,0001</b>	-	-
forrageira vs olerícola	0,9539	0,3287	2,6980	0,1005	0,8926	0,3448
forrageira vs ornamental	-	-	14,3158	<b>0,0002</b>	-	-
olerícola vs ornamental	-	-	86,8084	<b>&lt;0,0001</b>	-	-

**Tabela 9** - Teste de Gray para a comparação das curvas de incidência acumulada no ciclo perene estimadas pela técnica não paramétrica.

Comparações	Expiração		Renúncia		Cancelamento/ Anulação	
	estatística	valor- <i>p</i>	estatística	valor- <i>p</i>	estatística	valor- <i>p</i>
Titular						
todas são iguais	18,7930	<b>0,0001</b>	142,607	<b>&lt;0,0001</b>	5,1020	0,0785
pública vs privado	17,7920	<b>&lt;0,0001</b>	118,790	<b>&lt;0,0001</b>	3,753	0,0527
pública vs parceria	2,2301	0,1353	113,8136	<b>&lt;0,0001</b>	-	-
privado vs parceria	0,4683	0,4938	17,7813	<b>&lt;0,0001</b>	-	-
Tipo de cultura						
todas são iguais	174,5450	<b>&lt;0,0001</b>	431,6980	<b>&lt;0,0001</b>	7,7540	0,1703
agrícola vs florestal	51,8470	<b>&lt;0,0001</b>	1,4190	0,2335	3,32640	0,0708
agrícola vs forrageira	3,7526	0,0527	0,0049	0,9445	3,3937	0,0654
agrícola vs frutífera	50,0060	<b>&lt;0,0001</b>	4,9000	<b>0,0269</b>	6,8760	0,0874
agrícola vs olerícola	-	-	-	-	1,3865	0,2390
agrícola vs ornamental	132,7960	<b>&lt;0,0001</b>	188,3370	<b>&lt;0,0001</b>	6,8710	0,0876
florestal vs forrageira	12,4807	<b>0,0004</b>	0,5062	0,4768	0,2133	0,6442
florestal vs frutífera	3,7292	0,0535	0,6776	0,4104	0,4673	0,4942
florestal vs olerícola	-	-	-	-	0,0667	0,7962
florestal vs ornamental	0,3025	0,5823	122,8902	<b>&lt;0,0001</b>	0,3050	0,5808
frutífera vs olerícola	-	-	-	-	0,0015	0,9687
frutífera vs ornamental	4,8079	<b>0,0283</b>	172,1709	<b>&lt;0,0001</b>	0,0572	0,8109
olerícola vs ornamental	-	-	-	-	0,0009	0,9748

## APÊNDICE B – Rotina do modelo de Fine-Gray para o evento de expiração por prazo

```

rm(list=ls())

# Bibliotecas
library(survival)
library(cmprsk)
library(readxl)
library(dplyr)

# Dados
data <- read_excel("dados.xlsx")

data$transgenia <- as.factor(data$transgenia)
data$ciclo <- ifelse(data$ciclo == "anual", 0, 1)
data$titular <- as.factor(data$titular)
data$grupo <- as.factor(data$grupo)
data$ciclo <- as.factor(data$ciclo)

##### ANALISE NO CICLO: ANUAL #####
data_anual <- data %>% filter(ciclo == 0)
data_anual <- data_anual %>% filter(!grupo %in% c(1, 3, 5, 6))
data_anual$grupo <- factor(data_anual$grupo)

# Modelo de Fine-Gray
cens1 <- ifelse(data_anual$status == 1, 1, 0)
mod1 <- coxph(Surv(tempo, cens1) ~ transgenia + relevel(titular,
  ref = "1") + grupo, data=data_anual, x=TRUE)
xmod1 <- model.matrix(mod1)
fit1 <- with(data_anual, crr(tempo, status, xmod1, failcode=1))
print(summary(fit1), digits=4)

k1 <- length(fit1$coef)
aic1 <- -2*fit1$loglik + 2*k1
print(aic1, digits=8)

##### ANALISE NO CICLO: PERENE #####
data_perene <- data %>% filter(ciclo == 1)
data_perene <- data_perene %>% filter(!grupo %in% c(4, 6))
data_perene$grupo <- factor(data_perene$grupo)

# Modelo de Fine-Gray
cens2 <- ifelse(data_perene$status == 1, 1, 0)
mod.2 <- coxph(Surv(tempo, cens2) ~ relevel(titular, ref = "1")+
  grupo, data=data_perene, x=TRUE)
xmod.2 <- model.matrix(mod.2)
fit.2 <- with(data_perene, crr(tempo, status, xmod.2, failcode=1))
print(summary(fit.2), digits=4)

```

```

mod2 <- coxph(Surv(tempo, cens2) ~ grupo, data = data_perene,
              x=TRUE)
xmod2 <- model.matrix(mod2)
fit2 <- with(data_perene, crr(tempo, status, xmod2, failcode=1))
print(summary(fit2), digits=4)

k2 <- length(fit2$coef)
aic2 <- -2*fit2$loglik + 2*k2
print(aic2, digits=9)

##### PROPORCIONALIDADE #####
tiff("prop-exp.tiff",width=220,height=180,units="mm",res=600)
par(mfrow=c(3,3),mar=c(4,4,2,1),oma=c(0,0,0,0),mgp=c(2,0.5,0))

# Ciclo: anual
names(fit1$coef) = c("A","B","C","D","E")
for (j in 1:ncol(fit1$res)) {
  if (j == 1 || j == 4) {ylab_text = 'Resíduos Schoenfeld'}
  else {ylab_text = ''}
  scatter.smooth(fit1$uft, fit1$res[, j],
                 cex.lab = 1.3, cex.main = 1.4,
                 lpars = list(col = 'black', lwd = 1),
                 xlab = 'Tempo (meses)',
                 ylab = ylab_text,xaxt="n")
  abline(h=0, lty=3)
  axis(1, at=c(0,50,100,150,200,250,300,350))
  title(main=paste(names(fit1$coef)[j],sep=""),adj=0)}

# Ciclo: perene
names(fit2$coef) = c("F","G","H","I")
for (j in 1:ncol(fit2$res)) {
  if (j == 2) {ylab_text = 'Resíduos Schoenfeld'}
  else {ylab_text = ''}
  scatter.smooth(fit2$uft, fit2$res[, j],
                 cex.lab = 1.3, cex.main = 1.4,
                 lpars = list(col = 'black', lwd = 1),
                 xlab = 'Tempo (meses)',
                 ylab = ylab_text,xaxt="n")
  abline(h=0, lty=3)
  axis(1, at=c(0,50,100,150,200,250,300,350))
  title(main=paste(names(fit2$coef)[j],sep=""),adj=0)}

dev.off()

##### QUALIDADE: CICLO ANUAL #####
tiff("qual-exp.tiff",width=190,height=220,units="mm",res=600)
par(mfrow=c(4,3),mar=c(4,4,2,1),oma=c(0,0,0,0),mgp=c(2,0.5,0))

```

```

## Titular: Pública (1)
# sem transgenia e grupo agrícola (0)
comb4 <- subset(data_anual, transgenia==0 & titular==1 & grupo==0)
fia4 <- with(comb4, cuminc(tempo, status, rho=0, cencode=0))
fia.14 <- cbind(fia4$`1 1`$ time, fia4$`1 1`$ est)
fia.14[,1][fia.14[,1] == 0] <- 50
fia14 <- cbind(c(fia.14[,1], 331), c(fia.14[,2], fia.14[198,2]))

fp.4 <- predict(fit1, rbind(c(0,0,0,0,0))) #seleciona com xmod1
fp4 <- cbind(c(50, fp.4[,1], 331), c(0, fp.4[,2], fp.4[125,2]))

plot(fia14, type="s", xlab="Tempo (meses)", ylab="FIA",
     cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
     xlim=c(50,350), lty=1, lwd=0.75, col="red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp4, type="s", lty=1, col="darkgreen", lwd=0.75)
title(main = "A", adj = 0)

legend(50, 0.9, lty=c(1,1), col=c("red","darkgreen"),
      lwd=c(0.75,0.75), title(main = "A", adj = 0, bty="n",
      c("Não paramétrico", "Modelo Fine-Gray"), cex=1.1))

# com transgenia e grupo agrícola (0)
comb5 <- subset(data_anual, transgenia==1 & titular==1 & grupo==0)
fia5 <- with(comb5, cuminc(tempo, status, rho=0, cencode = 0))
fia.15 <- cbind(fia5$`1 1`$ time, fia5$`1 1`$ est)
fia.15[,1][fia.15[,1] == 0] <- 50
fia15 <- cbind(c(fia.15[,1], 331), c(fia.15[,2], fia.15[20,2]))

fp.5 <- predict(fit1, rbind(c(1,0,0,0,0))) #seleciona com xmod1
fp5 <- cbind(c(50, fp.5[,1], 331), c(0, fp.5[,2], fp.5[125,2]))

plot(fia15, type="s", xlab="Tempo (meses)", ylab="",
     cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
     xlim=c(50,350), lty=1, lwd = 0.75, col = "red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp5, type="s", lty=1, col = "darkgreen", lwd = 0.75)
title(main = "B", adj = 0)

# sem transgenia e grupo olericola (4)
comb6 <- subset(data_anual, transgenia==0 & titular==1 & grupo==4)
fia6 <- with(comb6, cuminc(tempo, status, rho=0, cencode = 0))
fia.16 <- cbind(fia6$`1 1`$ time, fia6$`1 1`$ est)
fia.16[,1][fia.16[,1] == 0] <- 50
fia16 <- cbind(c(fia.16[,1], 331), c(fia.16[,2], fia.16[8,2]))

fp.6 <- predict(fit1, rbind(c(0,0,0,0,1))) #seleciona com xmod1

```

```

fp6 <- cbind(c(50,fp.6[,1],331),c(0,fp.6[,2],fp.6[125,2]))

plot(fia16, type="s", xlab="Tempo (meses)", ylab="",
     cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
     xlim=c(50,350), lty=1, lwd = 0.75, col = "red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp6, type="s", lty=1, col = "darkgreen", lwd = 0.75)
title(main = "C", adj = 0)

## Titular: Privada (2)
# sem transgenia e grupo agrícola (0)
comb7 <-subset(data_anual,transgenia==0 & titular==2 & grupo==0)
fia7 <- with(comb7, cuminc(tempo, status, rho=0, cencode = 0))
fia.17 <- cbind(fia7$`1 1`$ time,fia7$`1 1`$ est)
fia.17[,1][fia.17[,1] == 0] <- 50
fia17 <- cbind(c(fia.17[,1],331),c(fia.17[,2],fia.17[132,2]))

fp.7 <- predict(fit1, rbind(c(0,1,0,0,0))) #selecciona com xmod1
fp7 <- cbind(c(50,fp.7[,1],331),c(0,fp.7[,2],fp.7[125,2]))

plot(fia17, type="s", xlab="Tempo (meses)", ylab="FIA",
     cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
     xlim=c(50,350), lty=1, lwd = 0.75, col = "red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp7, type="s", lty=1, col = "darkgreen", lwd = 0.75)
title(main = "D", adj = 0)

# com transgenia e grupo agrícola (0)
comb8 <-subset(data_anual,transgenia==1 & titular==2 & grupo==0)
fia8 <- with(comb8, cuminc(tempo, status, rho=0, cencode = 0))
fia.18 <- cbind(fia8$`1 1`$ time,fia8$`1 1`$ est)
fia.18[,1][fia.18[,1] == 0] <- 50
fia18 <- cbind(c(fia.18[,1],331),c(fia.18[,2],fia.18[76,2]))

fp.8 <- predict(fit1, rbind(c(1,1,0,0,0))) #selecciona com xmod1
fp8 <- cbind(c(50,fp.8[,1],331),c(0,fp.8[,2],fp.8[125,2]))

plot(fia18, type="s", xlab="Tempo (meses)", ylab="",
     cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
     xlim=c(50,350), lty=1, lwd = 0.75, col = "red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp8, type="s", lty=1, col = "darkgreen", lwd = 0.75)
title(main = "E", adj = 0)

```

```

# sem transgenia e grupo forrageira (2)
comb9 <- subset(data_anual,transgenia==0 & titular==2 & grupo==2)
fia9 <- with(comb9, cuminc(tempo, status, rho=0, cencode = 0))
fia.19 <- cbind(fia9$`1 1`$ time,fia9$`1 1`$ est)
fia.19[,1][fia.19[,1] == 0] <- 50
fia19 <- cbind(c(fia.19[,1],331),c(fia.19[,2],fia.19[6,2]))

fp.9 <- predict(fit1, rbind(c(0,1,0,1,0))) #seleciona com xmod1
fp9 <- cbind(c(50,fp.9[,1],331),c(0,fp.9[,2],fp.9[125,2]))

plot(fia19, type="s", xlab="Tempo (meses)", ylab=" ",
      cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
      xlim=c(50,350), lty=1, lwd = 0.75, col = "red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp9, type="s", lty=1, col = "darkgreen", lwd = 0.75)
title(main = "F", adj = 0)

# sem transgenia e grupo olericola (4)
comb10 <- subset(data_anual,transgenia==0&titular==2&grupo==4)
fia10 <- with(comb10, cuminc(tempo, status, rho=0, cencode = 0))
fia.110 <- cbind(fia10$`1 1`$ time,fia10$`1 1`$ est)
fia.110[,1][fia.110[,1] == 0] <- 50
fia110 <-cbind(c(fia.110[,1],331),c(fia.110[,2],fia.110[34,2]))

fp.10 <- predict(fit1, rbind(c(0,1,0,0,1))) #seleciona com xmod1
fp10 <- cbind(c(50,fp.10[,1],331),c(0,fp.10[,2],fp.10[125,2]))

plot(fia110, type="s", xlab="Tempo (meses)", ylab="FIA",
      cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
      xlim=c(50,350), lty=1, lwd = 0.75, col = "red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp10, type="s", lty=1, col = "darkgreen", lwd = 0.75)
title(main = "G", adj = 0)

##### QUALIDADE: CICLO PERENE #####
# grupo agricola (0)
comb16 <-subset(data_perene, grupo==0)
fia16 <- with(comb16, cuminc(tempo, status, rho=0, cencode = 0))
fia.116 <- cbind(fia16$`1 1`$ time,fia16$`1 1`$ est)
fia.116[,1][fia.116[,1] == 0] <- 50
fia116 <-cbind(c(fia.116[,1],331),c(fia.116[,2],fia.116[52,2]))

fp.16 <- predict(fit2, rbind(c(0,0,0,0))) #seleciona com xmod2
fp16 <- cbind(c(50,fp.16[,1],331),c(0,fp.16[,2],fp.16[54,2]))

plot(fia116, type="s", xlab="Tempo (meses)", ylab=" ",

```

```

cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
xlim=c(50,350), lty=1, lwd = 0.75, col = "red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp16, type="s", lty=1, col = "darkgreen", lwd = 0.75)
title(main = "H", adj = 0)

# grupo florestal (1)
comb17 <- subset(data_perene, grupo==1)
fia17 <- with(comb17, cuminc(tempo, status, rho=0, cencode = 0))
fia.117 <- cbind(fia17$`1 1`$ time, fia17$`1 1`$ est)
fia.117[,1][fia.117[,1] == 0] <- 50
fia117 <- cbind(c(fia.117[,1], 331), c(fia.117[,2], fia.117[10,2]))

fp.17 <- predict(fit2, rbind(c(1,0,0,0))) #seleciona com xmod2
fp17 <- cbind(c(50, fp.17[,1], 331), c(0, fp.17[,2], fp.17[54,2]))

plot(fia117, type="s", xlab="Tempo (meses)", ylab="",
      cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
      xlim=c(50,350), lty=1, lwd = 0.75, col = "red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp17, type="s", lty=1, col = "darkgreen", lwd = 0.75)
title(main = "I", adj = 0)

# grupo forrageira (2)
comb18 <- subset(data_perene, grupo==2)
fia18 <- with(comb18, cuminc(tempo, status, rho=0, cencode = 0))
fia.118 <- cbind(fia18$`1 1`$ time, fia18$`1 1`$ est)
fia.118[,1][fia.118[,1] == 0] <- 50
fia118 <- cbind(c(fia.118[,1], 331), c(fia.118[,2], fia.118[16,2]))

fp.18 <- predict(fit2, rbind(c(0,1,0,0))) # seleciona com xmod2
fp18 <- cbind(c(50, fp.18[,1], 331), c(0, fp.18[,2], fp.18[54,2]))

plot(fia118, type="s", xlab="Tempo (meses)", ylab="FIA",
      cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
      xlim=c(50,350), lty=1, lwd = 0.75, col = "red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp18, type="s", lty=1, col = "darkgreen", lwd = 0.75)
title(main = "J", adj = 0)

# grupo frutifera (3)
comb23 <- subset(data_perene, grupo==3)
fia23 <- with(comb23, cuminc(tempo, status, rho=0, cencode = 0))
fia.123 <- cbind(fia23$`1 1`$ time, fia23$`1 1`$ est)
fia.123[,1][fia.123[,1] == 0] <- 50

```

```

fia123 <-cbind(c(fia.123[,1],331),c(fia.123[,2],fia.123[38,2]))

fp.23 <- predict(fit2, rbind(c(0,0,1,0))) #seleciona com xmod2
fp23 <- cbind(c(50,fp.23[,1],328),c(0,fp.23[,2],fp.23[54,2]))

plot(fia123, type="s", xlab="Tempo (meses)", ylab="",
     cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
     xlim=c(50,350), lty=1, lwd = 0.75, col = "red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp23, type="s", lty=1, col = "darkgreen", lwd = 0.75)
title(main = "K", adj = 0)

# grupo ornamental (5)
comb19 <-subset(data_perene, grupo==5)
fia19 <- with(comb19, cuminc(tempo, status, rho=0, cencode = 0))
fia.119 <- cbind(fia19$`1 1`$ time,fia19$`1 1`$ est)
fia.119[,1][fia.119[,1] == 0] <- 50
fia119 <-cbind(c(fia.119[,1],331),c(fia.119[,2],fia.119[40,2]))

fp.19 <- predict(fit2, rbind(c(0,0,0,1))) #seleciona com xmod2
fp19 <- cbind(c(50,fp.19[,1],331),c(0,fp.19[,2],fp.19[54,2]))

plot(fia119, type="s", xlab="Tempo (meses)", ylab="",
     cex.lab=1.2, xaxt="n", yaxt="n", ylim=c(0,1),
     xlim=c(50,350), lty=1, lwd = 0.75, col = "red", main="")

axis(1, at=c(50,100,150,200,250,300,350))
axis(2, at=seq(0,1,by=0.2))
lines(fp19, type="s", lty=1, col = "darkgreen", lwd = 0.75)
title(main = "L", adj = 0)

dev.off()

```

## APÊNDICE C – Rotina do modelo de regressão binária ponderada para os eventos de expiração por prazo, renúncia, cancelamento e anulação

```

rm(list=ls())

# Bibliotecas
library(survival)
library(cmprsk)
library(pROC)
library(discSurv)
library(readxl)
library(dplyr)

# Dados
data <- as.data.frame(read_excel("dados.xlsx"))
data$transgenia <- as.factor(data$transgenia)
data$titular <- as.factor(data$titular)
data$grupo <- as.factor(data$grupo)
class(data)

##### ANALISE NO CICLO: ANUAL #####
data_anual <- data %>% filter(ciclo == "anual")
data_anual1 <- data_anual %>% filter(!grupo %in% c(1, 3, 5, 6))
data_anual1$grupo <- factor(data_anual1$grupo)
data_anual2 <- data_anual %>% filter(!grupo %in% c(1, 3, 6))
data_anual2$grupo <- factor(data_anual2$grupo)

# Evento: expiração por prazo
dados1 <- dataLongSubDist(dataShort=data_anual1,
                        timeColumn = "tempo",
                        eventColumns=c("exp", "ren", "canc"),
                        eventFocus="exp")

fit1 <- glm(y ~ transgenia + relevel(titular, ref = "1") + grupo,
           family=binomial(link = "cloglog"), data = dados1,
           weights = dados1$subDistWeights)
print(summary(fit1), digits=4)
print(AIC(fit1), digits=9)

# Evento: renuncia
dados2 <- dataLongSubDist(dataShort=data_anual2,
                        timeColumn = "tempo",
                        eventColumns=c("exp", "ren", "canc"),
                        eventFocus="ren")

fit2 <- glm(y ~ transgenia + relevel(titular, ref = "1") + grupo,
           family=binomial(link = "cloglog"), data = dados2,
           weights = dados2$subDistWeights)

```

```

print(summary(fit2), digits=4)
print(AIC(fit2), digits=9)

# Evento: cancelamento e anulacao
dados3 <- dataLongSubDist(dataShort=data_anual1,
                          timeColumn = "tempo",
                          eventColumns=c("exp", "ren", "canc"),
                          eventFocus="canc")

fit3 <- glm(y ~ transgenia + relevel(titular, ref = "1") + grupo,
            family=binomial(link = "cloglog"), data = dados3,
            weights = dados3$subDistWeights)
print(summary(fit3), digits=4)

fit4 <- glm(y ~ transgenia + relevel(titular, ref = "1"),
            family=binomial(link = "cloglog"), data = dados3,
            weights = dados3$subDistWeights)
print(summary(fit4), digits=4)
print(AIC(fit4), digits=9)

##### ANALISE NO CICLO: PERENE #####
data_perene <- data %>% filter(ciclo == "perene")
data_perene1 <- data_perene %>% filter(!grupo %in% c(4, 6))
data_perene1$grupo <- factor(data_perene1$grupo)
data_perene2 <- data_perene %>% filter(!titular %in% 3) %>%
  %>% filter(!grupo %in% 6)
data_perene2$titular <- factor(data_perene2$titular)
data_perene2$grupo <- factor(data_perene2$grupo)

# Evento: expiração por prazo
dados4 <- dataLongSubDist(dataShort=data_perene1,
                          timeColumn = "tempo",
                          eventColumns=c("exp", "ren", "canc"),
                          eventFocus="exp")

fit5 <- glm(y ~ relevel(titular, ref = "1") + grupo,
            family=binomial(link = "cloglog"), data = dados4,
            weights = dados4$subDistWeights)
print(summary(fit5), digits=4)

fit6 <- glm(y ~ grupo, family=binomial(link = "cloglog"),
            data = dados4, weights = dados4$subDistWeights)
print(summary(fit6), digits=4)
print(AIC(fit6), digits=9)

# Evento: renuncia
dados5 <- dataLongSubDist(dataShort=data_perene1,
                          timeColumn = "tempo",
                          eventColumns=c("exp", "ren", "canc"),
                          eventFocus="ren")

```

```

fit7 <- glm(y ~ relevel(titular, ref = "1") + grupo,
            family=binomial(link = "cloglog"), data = dados5,
            weights = dados5$subDistWeights)
print(summary(fit7), digits=4)
print(AIC(fit7), digits=9)

# Evento: cancelamento e anulacao
dados6 <- dataLongSubDist(dataShort=data_perene2,
                          timeColumn = "tempo",
                          eventColumns=c("exp","ren","canc"),
                          eventFocus="canc")

fit8 <- glm(y ~ relevel(titular, ref = "1") + grupo,
            family=binomial(link = "cloglog"), data = dados6,
            weights = dados6$subDistWeights)
print(summary(fit8), digits=4)

fit9 <- glm(y ~ grupo, family=binomial(link = "cloglog"),
            data = dados6, weights = dados6$subDistWeights)
print(summary(fit9), digits=4)
print(AIC(fit9), digits=9)

##### CURVAS ROC #####
roc1 <- plot.roc(dados1$y, fitted(fit1))
roc2 <- plot.roc(dados2$y, fitted(fit2))
roc3 <- plot.roc(dados3$y, fitted(fit4))
roc4 <- plot.roc(dados4$y, fitted(fit6))
roc5 <- plot.roc(dados5$y, fitted(fit7))
roc6 <- plot.roc(dados6$y, fitted(fit9))

tiff("curvasroc.tiff",width=190,height=130,units="mm",res=600)
par(mfrow=c(2,3),mar=c(4,4,2,1),oma=c(0,0,0,0),mgp=c(2,0.5,0))

plot(smooth(roc1, method="density"), xlab="Especificidade",
     ylab="Sensibilidade", print.auc=TRUE,
     auc.polygon=TRUE, grud=c(0.1,0.2),
     max.auc.polygon=TRUE, auc.polygon.col=FALSE,
     grid.col=FALSE, max.auc.polygon.col=FALSE,
     identity.lty=2, identity.lwd=1.5,
     identity.col="black", print.thres=FALSE,
     print.auc.cex=1.1)
title(main = "A", adj = 0)

plot(smooth(roc2, method="density"), xlab="Especificidade",
     ylab=" ", print.auc=TRUE,
     auc.polygon=TRUE, grud=c(0.1,0.2),
     max.auc.polygon=TRUE, auc.polygon.col=FALSE,
     grid.col=FALSE, max.auc.polygon.col=FALSE,
     identity.lty=2, identity.lwd=1.5,
     identity.col="black", print.thres=FALSE,

```

```
        print.auc.cex=1.1)
title(main = "B", adj = 0)

plot(smooth(roc3, method="density"), xlab="Especificidade",
      ylab="", print.auc=TRUE,
      auc.polygon=TRUE, grud=c(0.1,0.2),
      max.auc.polygon=TRUE, auc.polygon.col=FALSE,
      grid.col=FALSE, max.auc.polygon.col=FALSE,
      identity.lty=2, identity.lwd=1.5,
      identity.col="black", print.thres=FALSE,
      print.auc.cex=1.1)
title(main = "C", adj = 0)

plot(smooth(roc4, method="density"), xlab="Especificidade",
      ylab="Sensibilidade", print.auc=TRUE,
      auc.polygon=TRUE, grud=c(0.1,0.2),
      max.auc.polygon=TRUE, auc.polygon.col=FALSE,
      grid.col=FALSE, max.auc.polygon.col=FALSE,
      identity.lty=2, identity.lwd=1.5,
      identity.col="black", print.thres=FALSE,
      print.auc.cex=1.1)
title(main = "D", adj = 0)

plot(smooth(roc5, method="density"), xlab="Especificidade",
      ylab="", print.auc=TRUE,
      auc.polygon=TRUE, grud=c(0.1,0.2),
      max.auc.polygon=TRUE, auc.polygon.col=FALSE,
      grid.col=FALSE, max.auc.polygon.col=FALSE,
      identity.lty=2, identity.lwd=1.5,
      identity.col="black", print.thres=FALSE,
      print.auc.cex=1.1)
title(main = "E", adj = 0)

plot(smooth(roc6, method="density"), xlab="Especificidade",
      ylab="", print.auc=TRUE,
      auc.polygon=TRUE, grud=c(0.1,0.2),
      max.auc.polygon=TRUE, auc.polygon.col=FALSE,
      grid.col=FALSE, max.auc.polygon.col=FALSE,
      identity.lty=2, identity.lwd=1.5,
      identity.col="black", print.thres=FALSE,
      print.auc.cex=1.1)
title(main = "F", adj = 0)

dev.off()
```

## APÊNDICE D – Rotina do estimador não paramétrico da função de incidência acumulada para os eventos de expiração por prazo, renúncia, cancelamento e anulação

```

rm(list=ls())

# Bibliotecas
library(survival)
library(cmprsk)
library(readxl)
library(dplyr)

# Dados
data <- read_excel("dados.xlsx")
data$transgenia <- as.factor(data$transgenia)
data$titular <- as.factor(data$titular)
data$grupo <- as.factor(data$grupo)

##### ANALISE NO CICLO: ANUAL #####
data_anual <- data %>% filter(ciclo == "anual")
data_anual <- data_anual %>% filter(!grupo %in% c(1, 3, 6))
data_anual$grupo <- factor(data_anual$grupo)

##### ANALISE NO CICLO: PERENE #####
data_perene <- data %>% filter(ciclo == "perene")
data_perene <- data_perene %>% filter(!grupo %in% 6)
data_perene$grupo <- factor(data_perene$grupo)

##### GRAFICOS #####
# Ciclo: anual
tiff("eventos.tiff",width=200,height=120,units="mm",res=600)
par(mfrow = c(1,2))

f1 <- with(data_anual, cuminc(tempo, status, cencode=0))

plot(f1, ylim = c(0, 0.5), xlim = c(0, 350),
     xlab = "Tempo (meses)", ylab = "FIA", lty=c(1,1,1),
     col=c("red","darkgreen","blue"), lwd=c(1,1,1),
     curvlab = c("Expiração por prazo", "Renúncia",
                 "Cancelamento e Anulação"), main = "", xatx="n")
axis(1,at=c(0,50,100,150,200,250,300,350))
title(main = "A", adj = 0)

# Ciclo: perene
f2 <- with(data_perene, cuminc(tempo, status, cencode=0))

plot(f2, ylim = c(0, 0.5), xlim = c(0, 350),
     xlab = "Tempo (meses)", ylab = "FIA", lty=c(1,1,1),

```

```

        col=c("red","darkgreen","blue"), lwd=c(1,1,1),
        curvlab = c("Expiração por prazo", "Renúncia",
        "Cancelamento e Anulação"), main = "", xatx="n")
axis(1,at=c(0,50,100,150,200,250,300,350))
title(main = "B", adj = 0)

dev.off()

### FIAs associadas a covariavel: transgenia

# Ciclo: anual
fia2 <- with(data_anual, cuminc(tempo, status, transgenia,
                               cencode="0"))
fia2$Tests

tiff("anu-trans.tiff",width=185,height=70,units="mm",res=600)
par(mfrow=c(1,3),mar=c(4,4,2,1),oma=c(0,0,0,0),mgp=c(2,0.5,0))

prazol <- list(list(fia2$`0 1`$time, fia2$`0 1`$est),
               list(c(fia2$`1 1`$time,331),
                    c(fia2$`1 1`$est,fia2$`1 1`$est[86])))

plot.cuminc(prazol, ylim = c(0, 1), xlim = c(0, 350),
            cex.lab=1.2, main="",xatx="n",
            curvlab = c('sem transgenia', 'com transgenia'),
            col=c("red","darkgreen"), lty = c(1, 1),
            xlab = 'Tempo (meses)', ylab = 'FIA')
axis(1,at=c(0,50,100,150,200,250,300,350))
title(main = "A", adj = 0)

ren1 <- list(list(fia2$`0 2`$time, fia2$`0 2`$est),
             list(c(fia2$`1 2`$time,331),
                  c(fia2$`1 2`$est,fia2$`1 2`$est[94])))

plot.cuminc(ren1, ylim = c(0, 1), xlim = c(0, 350),
            cex.lab=1.2, main="",xatx="n",
            curvlab = c('sem transgenia', 'com transgenia'),
            col=c("red","darkgreen"), lty = c(1, 1),
            xlab = 'Tempo (meses)', ylab = 'FIA')
axis(1,at=c(0,50,100,150,200,250,300,350))
title(main = "B", adj = 0)

cancl <- list(list(fia2$`0 3`$time, fia2$`0 3`$est),
              list(c(fia2$`1 3`$time,331),
                   c(fia2$`1 3`$est,fia2$`1 3`$est[34])))

plot.cuminc(cancl, ylim = c(0, 1), xlim = c(0, 350),
            cex.lab=1.2, main="",xatx="n",
            curvlab = c('sem transgenia', 'com transgenia'),
            col=c("red","darkgreen"), lty = c(1, 1),
            xlab = 'Tempo (meses)', ylab = 'FIA')

```

```

axis(1,at=c(0,50,100,150,200,250,300,350))
title(main = "C", adj = 0)

dev.off()

### FIAs associadas a covariavel: titular

# Ciclo: anual
fia3<-with(data_anual, cuminc(tempo,status, titular,
                             cencode = "0"))
fia3$Tests

tiff("anpe-tit.tiff",width=210,height=140,units="mm",res=600)
par(mfrow=c(2,3),mar=c(4,4,2,1),oma=c(0,0,0,0),mgp=c(2,0.5,0))

prazo2 <- list(list(fia3$`1 1`$time,fia3$`1 1`$est),
                list(fia3$`2 1`$time, fia3$`2 1`$est),
                list(fia3$`3 1`$time,fia3$`3 1`$est))

plot.cuminc(prazo2, ylim = c(0, 1), xlim = c(0, 350),
             cex.lab=1.2, curvlab = c('instituição pública',
                                       'setor privado', 'parceria'), lty=c(1,1,1),
             col=c("red","darkgreen","blue"), main="", xatx="n",
             xlab = 'Tempo (meses)', ylab = 'FIA')
axis(1, at=c(0,50,100,150,200,250,300,350))
title(main = "A", adj = 0)

ren2 <- list(list(fia3$`1 2`$time,fia3$`1 2`$est),
              list(fia3$`2 2`$time, fia3$`2 2`$est),
              list(fia3$`3 2`$time,fia3$`3 2`$est))

plot.cuminc(ren2, ylim = c(0, 1), xlim = c(0, 350),
             cex.lab=1.2, curvlab = c('instituição pública',
                                       'setor privado', 'parceria'), lty=c(1,1,1),
             col=c("red","darkgreen","blue"), main="", xatx="n",
             xlab = 'Tempo (meses)', ylab = 'FIA')
axis(1, at=c(0,50,100,150,200,250,300,350))
title(main = "B", adj = 0)

canc2 <- list(list(fia3$`1 3`$time,fia3$`1 3`$est),
               list(fia3$`2 3`$time, fia3$`2 3`$est),
               list(fia3$`3 3`$time,fia3$`3 3`$est))

plot.cuminc(canc2, ylim = c(0, 1), xlim = c(0, 350),
             cex.lab=1.2, curvlab = c('instituição pública',
                                       'setor privado', 'parceria'), lty=c(1,1,1),
             col=c("red","darkgreen","blue"), main="", xatx="n",
             xlab = 'Tempo (meses)', ylab = 'FIA')
axis(1,at=c(0,50,100,150,200,250,300,350))
title(main = "C", adj = 0)

```



```

        col=c("red","darkgreen"), main="", xatx="n",
        xlab = 'Tempo (meses)', ylab = 'FIA')
axis(1,at=c(0,50,100,150,200,250,300,350))
title(main = "F", adj = 0)

dev.off()

# Testes
dados_pt12 <- data_perene[data_perene$titular %in% c(1,2), ]
fia_pt12 <- with(dados_pt12, cuminc(tempo, status, titular,
                                cencode = "0"))
fia_pt12$Tests

dados_pt13 <- data_perene[data_perene$titular %in% c(1,3), ]
fia_pt13 <- with(dados_pt13, cuminc(tempo, status, titular,
                                cencode = "0"))
fia_pt13$Tests

dados_pt23 <- data_perene[data_perene$titular %in% c(2,3), ]
fia_pt23 <- with(dados_pt23, cuminc(tempo, status, titular,
                                cencode = "0"))
fia_pt23$Tests

### FIAs associadas a covariavel: grupo

# Ciclo: anual
fia5<-with(data_anual, cuminc(tempo,status, grupo,
                             cencode = "0"))
fia5$Tests

tiff("anpe-grup.tiff",width=210,height=140,units="mm",res=600)
par(mfrow=c(2,3),mar=c(4,4,2,1),oma=c(0,0,0,0),mgp=c(2,0.5,0))

prazo4 <- list(list(fia5$`0 1`$time, fia5$`0 1`$est),
               list(fia5$`2 1`$time, fia5$`2 1`$est),
               list(fia5$`4 1`$time, fia5$`4 1`$est))

plot.cuminc(prazo4, ylim= c(0, 1), xlim = c(0, 335),
            cex.lab=1.2, curvlab=c('agrícola','forrageira',
            'olerícola'),col=c("red","blue", "#FF007F"),
            lty=c(1,1,1), xlab = 'Tempo (meses)', ylab = 'FIA',
            main="",xatx="n")
axis(1,at=c(0,50,100,150,200,250,300,350))
title(main = "A", adj = 0)

ren4 <- list(list(fia5$`0 2`$time, fia5$`0 2`$est),
             list(fia5$`2 2`$time, fia5$`2 2`$est),
             list(fia5$`4 2`$time, fia5$`4 2`$est),
             list(fia5$`5 2`$time, fia5$`5 2`$est))

plot.cuminc(ren4, ylim = c(0, 1), xlim = c(0, 350), cex.lab=1.2,

```

```

        curvlab=c('agrícola', 'forrageira', 'olerícola',
        'ornamental'), col=c("red", "blue", "#FF007F",
        "orange"), lty=c(1,1,1,1), xlab = 'Tempo (meses)',
        ylab = '', main="", xatx="n")
axis(1, at=c(0, 50, 100, 150, 200, 250, 300, 350))
title(main = "B", adj = 0)

canc4 <- list(list(fia5$`0 3`$time, fia5$`0 3`$est),
        list(fia5$`2 3`$time, fia5$`2 3`$est),
        list(fia5$`4 3`$time, fia5$`4 3`$est))

plot.cuminc(canc4, ylim = c(0, 1), xlim = c(0, 350), cex.lab=1.2,
        curvlab=c('agrícola', 'forrageira', 'olerícola'),
        col=c("red", "blue", "#FF007F"), lty=c(1,1,1),
        xlab = 'Tempo (meses)', ylab = '', main="", xatx="n")
axis(1, at=c(0, 50, 100, 150, 200, 250, 300, 350))
title(main = "C", adj = 0)

# Testes
dados_ag02 <- data_anual[data_anual$grupo %in% c(0, 2), ]
fia_ag02 <- with(dados_ag02, cuminc(tempo, status, grupo,
        cencode = "0"))
fia_ag02$Tests

dados_ag04 <- data_anual[data_anual$grupo %in% c(0, 4), ]
fia_ag04 <- with(dados_ag04, cuminc(tempo, status, grupo,
        cencode = "0"))
fia_ag04$Tests

dados_ag05 <- data_anual[data_anual$grupo %in% c(0, 5), ]
fia_ag05 <- with(dados_ag05, cuminc(tempo, status, grupo,
        cencode = "0"))
fia_ag05$Tests

dados_ag24 <- data_anual[data_anual$grupo %in% c(2, 4), ]
fia_ag24 <- with(dados_ag24, cuminc(tempo, status, grupo,
        cencode = "0"))
fia_ag24$Tests

dados_ag25 <- data_anual[data_anual$grupo %in% c(2, 5), ]
fia_ag25 <- with(dados_ag25, cuminc(tempo, status, grupo,
        cencode = "0"))
fia_ag25$Tests

dados_ag45 <- data_anual[data_anual$grupo %in% c(4, 5), ]
fia_ag45 <- with(dados_ag45, cuminc(tempo, status, grupo,
        cencode = "0"))
fia_ag45$Tests

# Ciclo: perene
fia9<-with(data_perene, cuminc(tempo, status, grupo,

```

```

        cencode = "0"))
fia9$Tests

prazo7 <- list(list(fia9$`0 1`$time, fia9$`0 1`$est),
               list(fia9$`1 1`$time, fia9$`1 1`$est),
               list(fia9$`2 1`$time, fia9$`2 1`$est),
               list(fia9$`3 1`$time, fia9$`3 1`$est),
               list(fia9$`5 1`$time, fia9$`5 1`$est))

plot.cuminc(prazo7, ylim = c(0, 1), xlim = c(0, 350),
            cex.lab=1.2, curvlab = c('agrícola', 'florestal',
            'forrageira', 'frutífera', 'ornamental'), col=
            c("red", "darkgreen", "blue", "purple", "orange"),
            lty=c(1,1,1,1,1), xlab = 'Tempo (meses)',
            ylab = 'FIA', main="", xatx="n")
axis(1, at=c(0,50,100,150,200,250,300,350))
title(main = "D", adj = 0)

ren7 <- list(list(fia9$`0 2`$time, fia9$`0 2`$est),
              list(fia9$`1 2`$time, fia9$`1 2`$est),
              list(fia9$`2 2`$time, fia9$`2 2`$est),
              list(fia9$`3 2`$time, fia9$`3 2`$est),
              list(fia9$`5 2`$time, fia9$`5 2`$est))

plot.cuminc(ren7, ylim = c(0, 1), xlim = c(0, 350), cex.lab=1.2,
            curvlab = c('agrícola', 'florestal', 'forrageira',
            'frutífera', 'ornamental'), col=c("red", "darkgreen",
            "blue", "purple", "orange"), lty=c(1,1,1,1,1), xlab=
            'Tempo (meses)', ylab='', main="")
axis(1, at=c(0,50,100,150,200,250,300,350))
title(main = "E", adj = 0)

canc7 <- list(list(fia9$`0 3`$time, fia9$`0 3`$est),
              list(fia9$`1 3`$time, fia9$`1 3`$est),
              list(fia9$`2 3`$time, fia9$`2 3`$est),
              list(fia9$`3 3`$time, fia9$`3 3`$est),
              list(fia9$`4 3`$time, fia9$`4 3`$est),
              list(fia9$`5 3`$time, fia9$`5 3`$est))

plot.cuminc(canc7, ylim = c(0, 1), xlim = c(0, 350), cex.lab=1.2,
            curvlab = c('agrícola', 'florestal', 'forrageira',
            'frutífera', 'olerícola', 'ornamental'), col=c("red",
            "darkgreen", "blue", "purple", "#FF007F", "orange"),
            lty=c(1,1,1,1,1,1), xlab='Tempo (meses)', ylab = '',
            main="", xatx="n")
axis(1, at=c(0,50,100,150,200,250,300,350))
title(main = "F", adj = 0)

dev.off()
# Testes
dados_pg01 <- data_perene[data_perene$grupo %in% c(0, 1), ]

```

```
fia_pg01 <- with(dados_pg01, cuminc(tempo, status, grupo,
                                cencode = "0"))
fia_pg01$Tests

dados_pg02 <- data_perene[data_perene$grupo %in% c(0, 2), ]
fia_pg02 <- with(dados_pg02, cuminc(tempo, status, grupo,
                                cencode = "0"))
fia_pg02$Tests

dados_pg03 <- data_perene[data_perene$grupo %in% c(0, 3), ]
fia_pg03 <- with(dados_pg03, cuminc(tempo, status, grupo,
                                cencode = "0"))
fia_pg03$Tests

dados_pg04 <- data_perene[data_perene$grupo %in% c(0, 4), ]
fia_pg04 <- with(dados_pg04, cuminc(tempo, status, grupo,
                                cencode = "0"))
fia_pg04$Tests

dados_pg05 <- data_perene[data_perene$grupo %in% c(0, 5), ]
fia_pg05 <- with(dados_pg05, cuminc(tempo, status, grupo,
                                cencode = "0"))
fia_pg05$Tests

dados_pg12 <- data_perene[data_perene$grupo %in% c(1, 2), ]
fia_pg12 <- with(dados_pg12, cuminc(tempo, status, grupo,
                                cencode = "0"))
fia_pg12$Tests

dados_pg13 <- data_perene[data_perene$grupo %in% c(1, 3), ]
fia_pg13 <- with(dados_pg13, cuminc(tempo, status, grupo,
                                cencode = "0"))
fia_pg13$Tests

dados_pg14 <- data_perene[data_perene$grupo %in% c(1, 4), ]
fia_pg14 <- with(dados_pg14, cuminc(tempo, status, grupo,
                                cencode = "0"))
fia_pg14$Tests

dados_pg15 <- data_perene[data_perene$grupo %in% c(1, 5), ]
fia_pg15 <- with(dados_pg15, cuminc(tempo, status, grupo,
                                cencode = "0"))
fia_pg15$Tests

dados_pg34 <- data_perene[data_perene$grupo %in% c(3, 4), ]
fia_pg34 <- with(dados_pg34, cuminc(tempo, status, grupo,
                                cencode = "0"))
fia_pg34$Tests

dados_pg35 <- data_perene[data_perene$grupo %in% c(3, 5), ]
fia_pg35 <- with(dados_pg35, cuminc(tempo, status, grupo,
```

