

Modelos Série de Potência com Excesso de Zeros Observáveis e Latentes

Katherine Elizabeth Coaguila Zavaleta

Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Departamento de Estatística

Modelos Série de Potência com Excesso de Zeros Observáveis e Latentes

Katherine Elizabeth Coaguila Zavaleta

Orientador: Prof. Doutor Vicente Garibay Cancho

Coorientador: Prof. Doutor Adriano Kamimura Suzuki

Tese apresentada ao Departamento de Estatística da Universidade Federal de São Carlos - DEs/UFSCar como parte dos requisitos para obtenção do título de Doutor em Estatística.

UFSCar - São Carlos - SP

Setembro de 2016



Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado da candidata Katherine Elizabeth Coaguila Zavaleta, realizada em 28/09/2016:

Prof. Dr. Vicente Garibay Cancho
USP

Prof. Dr. Adriano Kamimura Suzuki
USP

Profa. Dra. Gleici da Silva Castro Perdoná
USP

Prof. Dr. Josemar Rodrigues
USP

Profa. Dra. Juliana Cobre
USP

Profa. Dra. Vera Lúcia Damasceno Tomazella
UFSCar

Agradecimentos

Agradeço à Deus por ser meu parceiro na luta, pela força para superar as dificuldades da vida e agora pela vitória desta tese de doutorado.

Ao meu orientador de doutorado, Vicente Garibay Cancho, pelas suas valiosas orientações que foram essenciais para realização deste trabalho, disponibilidade de tempo, paciência e amizade.

Ao meu co-orientador de doutorado, Adriano Kamimura Suzuki, pelo acompanhamento de supervisão e estímulo no desenvolvimento do trabalho.

Aos professores membros da banca Gleici da Silva Castro Perdoná, Juliana Cobre, Josemar Rodrigues e Vera Lucia Damasceno Tomazella pelas sugestões que ajudaram no melhoramento desta tese.

Aos professores e funcionários do Departamento de Estatística da Universidade Federal de São Carlos que me propiciaram condições para realização deste trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro concedido durante o período de doutorado.

A meus amigos e colegas que fez em São Carlos por ter me acompanhado nos momentos de alegria, conselho e estudo.

E com muito amor,

Aos meus pais Carolin e José Luis, minha irmã Karla.

Resumo

O presente trabalho teve como objetivo principal, estudar a significância de zeros numa análise de dados observáveis e latentes. Nos conjuntos de dados observáveis que ocorrem excessos de zeros, é comum a existência de sobredispersão. Neste sentido os modelos Zero-Inflacionados Série de Potência (ZISP) foram propostos para acomodar o excesso de zeros. Especificamente para a análise de dados observáveis com excesso de zeros desenvolvemos um estudo da estatística gradiente, proposta por [Terrell \(2002\)](#), para testar as hipóteses em relação ao parâmetro de inflação do modelo ZISP, baseado na avaliação da performance da estatística gradiente em comparação com as estatísticas clássicas da razão de verossimilhança ([Wilks, 1938](#)), escore ([Rao, 1948](#)) e Wald ([Wald, 1943](#)). Por outro lado, recentemente a fragilidade é modelada por distribuições discretas sob os inteiros não negativos e permite fragilidade zero, isto é, indivíduos que não apresentam o evento de interesse (fração de risco zero). Para este tipo dados de latentes, propusemos um novo modelo de sobrevivência induzida por fragilidade discreta com distribuição ZISP. Essa proposta traz uma descrição mais real dos indivíduos sem risco, pois inclui indivíduos curados devido aos fatores genéticos (imunes) modelados como a fração de risco zero determinístico, enquanto que, os indivíduos curados por tratamento são modelados pela fração de risco zero aleatório. Neste contexto desenvolvemos também a estatística gradiente para verificar a significância do parâmetro de risco zero para dados modelados pela fração de risco zero determinístico. E para completar o desenvolvimento das propostas, apresentamos os resultados de estudos de simulação e exemplos de aplicação com uso de dados reais.

Palavras-chave: Teste gradiente, Modelos zero-inflacionados Série de Potência, Modelo Série de Potência, Análise de sobrevivência, Fragilidade, Risco zero.

Abstract

The present work's main objective is to study the significance of zeros in an observable and latent data. In observable data set that occur excess of zeros, its common to have sobredispersion. In this sense, the models zero-inflated power series (ZISP) were proposed to accommodate these excesses. Specifically for the analysis of observed data, it was made a study of gradient statistic, proposed by Terrell (2002), to test the hypotheses in relation to inflation parameter ZISP models. This test is based on evaluation of the performance of gradient statistic compared with the classical likelihood ratio (Wilks, 1938), score (Rao, 1948) and Wald (Wald, 1943) statistics. In addition, recently, fragility has being modeled by discrete distributions using non-negative integers values that allows zero fragility, which means, individuals who do not present the event of interest (fraction of zero risk). For this type of latent data, we have proposed a new survival model induced by discrete frailty with ZISP distribution. This proposal brings a real description of individuals without risk, because individuals cured due to genetic factors (immune) are modeled by fraction of deterministic zero risk, while the cured by treatment are modeled by fraction of random zero risk. In this context, we also developed the gradient statistic to verify parameter significance of zero risk for data modeled by fraction of deterministic zero risk. To show our proposals, we present the results of simulation studies and applications using real data.

Keywords: Gradient statistic, Zero-Inflated Power Series Model, Power Series Model, Survival Analysis, Fragility, zero risk.

Sumário

Lista de Tabelas	vi
1 Introdução	1
1.1 Estatística gradiente	4
1.2 Objetivos	5
1.3 Organização da tese	5
2 Teste Gradiente no Modelo Zero-Inflacionado Série de Potência	7
2.1 Os modelos Zero-Inflacionado Série de Potência	7
2.1.1 Casos particulares	10
2.2 Teste gradiente no modelo de regressão zero-inflacionado Série de Potência	15
2.2.1 Casos particulares	19
2.3 Estudo de simulação	23
2.3.1 Simulações sem covariáveis no modelo ZI-Poisson	23
2.3.2 Simulações sem covariáveis no modelo ZI-Geométrica	31
2.3.3 Simulação com covariáveis no modelo ZI-Poisson	37
2.3.4 Simulações com covariáveis no modelo ZI-Geométrica	39
2.4 Aplicação	42
2.4.1 Aplicação 1: <i>Lesões em empregados da limpeza.</i>	43
2.4.2 Aplicação 2: <i>Cultivos de maçãs</i>	44

2.5	Conclusões	46
3	Um Modelo de Sobrevivência Induzida por Fragilidade Discreta: Enfoque Clássico	47
3.1	Introdução	47
3.2	O modelo	50
3.2.1	Alguns casos particulares	52
3.2.2	Inferência	55
3.2.3	Método Bootstrap	58
3.3	Estudo de simulação	59
3.4	Aplicação	61
3.5	Conclusões	66
4	Modelo de sobrevivência induzida por Fragilidade Discreta: Enfoque bayesiano	67
4.1	Inferência Bayesiana no modelo ZISP-FC	67
4.1.1	Análise de diagnóstico caso a caso	70
4.2	Estudo de simulação	72
4.3	Aplicação	74
4.4	Conclusões	78
5	Conclusões e Pesquisas Futuras	79
A		81
A.1	Obtenção da estatística Gradiente	81
A.1.1	Normalidade assintótica das estatísticas	81
B		83

B.1	Propriedades do modelo SP	83
B.2	Propriedades do modelo ZISP	84
B.2.1	Cálculo da função score e da matriz de IF	86
C		90
C.1	Método do gráfico TTT	90
C.2	Matriz de informação observada do modelo ZISP-FC	90
C.3	Vetor escore e matriz de informação observada do modelo ZIP- FC	92
Referências Bibliográficas		96

Lista de Tabelas

1.1	Derivadas da função $A(\theta)$	3
2.1	Taxas de rejeição da hipótese $H_o : \phi = 0$, no modelo ZI-Poisson, para diferentes valores de θ	24
2.2	Percentis estimados com amostra de tamanho $n = 50$ sob $H_0 : \phi = 0$	25
2.3	Percentis estimados com amostra de tamanho $n = 200$ sob $H_0 : \phi = 0$	25
2.4	Poder do teste de ϕ quando $\theta = 2$	27
2.5	Taxas de rejeição da hipótese para testar $H_o : \phi = 0$, no modelo ZIG, para diferentes valores de θ	31
2.6	Percentis estimados com amostra de tamanho $n = 100$ sob $H_0 : \phi = 0$	32
2.7	Percentis estimados com amostra de tamanho $n = 400$ sob $H_0 : \phi = 0$	32
2.8	Taxas de rejeição da hipótese para testar $H_0 : \phi = 0$ no modelo de regressão ZI-Poisson.	37
2.9	Percentis estimados sob $H_0 : \phi = 0$	38
2.10	Taxas de rejeição da hipótese $H_0 : \beta_2 = \beta_3 = 0$ para S_G, S_{LR}, S_R e S_W para diferentes tamanhos de amostra no modelo de regressão ZI-Poisson	39
2.11	Taxas de rejeição da hipótese $H_0 : \phi = 0$ no modelo de regressão ZI-Geométrica.	40
2.12	Percentis estimados sob $H_0 : \phi = 0$	41
2.13	Taxas de rejeição nula de $H_0 : \beta_2 = \beta_3 = 0$ para S_G, S_{LR}, S_R e S_W para diferentes tamanhos de amostra no modelo de regressão ZI-Geométrica	42
2.14	EMV para os parâmetros com ajuste dos modelos ZI-Poisson e ZI-Geométrica para os dados de lesões em empregados da limpeza.	43
2.15	Estimativas das estatísticas e nível descritivo do teste de hipóteses $H_0 : \omega = 0$	43

2.16	Valores de máximo da log-verossimilhança e os critérios AIC e SBC com ajuste dos modelos ZI-Poisson e ZI-Geométrica para dados de cultivos de maçã. . .	45
2.17	EMV para os parâmetros do modelo ZI-Poisson para os dados de cultivos de maçã.	45
2.18	Estimativas das estatísticas de teste e nível descritivo para o teste $H_0 : \omega = 0$ do modelo ZI-Poisson.	45
2.19	Estimativas das estatísticas e nível descritivo para o teste $H_0 : \beta_1 = \beta_2 = 0$ do modelo ZI-Poisson.	46
2.20	Testando a significância das hipóteses, $H_0 : \beta_1 = 0$ e $H_0 : \beta_2 = 0$	46
3.1	Média das EMV, viés e raiz quadrada do erro quadrático médio (REQM) dos parâmetros do modelo de sobrevivência ZIP-FC.	60
3.2	EMVs com ajuste dos modelos de sobrevivência ZIP-FC Weibull e ZIG-FC Weibull para os dados de melanoma.	63
3.3	EMVs assintóticas e <i>bootstrap</i> com ajuste do modelo de sobrevivência ZIG-FC Weibull simplificado.	64
3.4	Estimativas das estatísticas e nível descritivo para o teste de hipótese $H_0 : \omega = 0$ nos dados de melanoma.	65
3.5	Estimativa de máxima verossimilhança para fração de curados estratificada por categoria de nódulo e idade dos pacientes.	66
4.1	As médias de Monte Carlos da média <i>a posteriori</i> , desvio padrão das médias <i>a posteriori</i> (DP), raiz quadrada do erro quadrático médio (REQM) das médias, viés e a probabilidade de cobertura (PC) do intervalo HPD de 95% de confiança para cada parâmetro do modelo ZISP-FC.	73
4.2	Porcentagens de vezes que $H_0 : \phi = 0$ foi rejeitada em 1000 amostras.	74
4.3	Média, mediana e desvio padrão (DP) <i>a posteriori</i> e o intervalo HPD de 95% de confiança dos parâmetros dos Modelos ZIP-FC e ZIG-FC, para os dados de melanoma.	74
4.4	Estimativas de de Monte Carlo das medidas ψ -divergência para os modelos ZIP-FC e ZIG-FC.	75
4.5	Estimativas de Monte Carlo do DIC, EAIC e EBIC para o ajuste dos dados de melanoma.	76
4.6	Sumário <i>a posteriori</i> , média, mediana, desvio padrão e intervalo HPD de 90% de confiança, para o modelo simplificado.	76

4.7	Estimativas Bayesianas DIC, EAIC e EBIC para o ajuste dos dados de melanoma.	77
4.8	Estimativa Bayesianas para fração de curados estratificada por categoria de nódulo e idade dos pacientes.	78

Capítulo 1

Introdução

Dados de contagem com excesso de zeros (ou zero-inflacionados) são comumente encontrados em muitas áreas, incluindo a medicina (Böhning *et al.*, 1999), saúde pública (Zhou & Tu, 2000), ciência ambiental (Agarwal *et al.*, 2002), agricultura (Hall, 2000) e aplicações industriais (Lambert, 1992). Excesso de zeros é uma indicação de sobredispersão nos dados, significando que a incidência de zero na contagem é geralmente maior do que o esperado. A incidência de zeros na contagem tem status especial do ponto de vista prático. Por exemplo, Ridout *et al.* (2001) apontaram que, na contagem de lesões das doenças em plantas, uma planta pode não ter lesões porque é resistente à doença, ou simplesmente porque não há esporos que pousaram sobre elas que causam a doença.

A literatura sobre os modelos zero-inflacionados é extensa, dentre os quais citamos os trabalhos Consul & Jain (1973), Gupta *et al.* (1995), Van den Broek (1995), Yang *et al.* (2007), del Castillo & Pérez-Casany (2005), Patil & Shirke (2007), Choo-Wosoba *et al.* (2015), Barriga & Louzada (2014), Samani *et al.* (2012) e Deng & Zhang (2015). Gupta *et al.* (1995) estudaram a distribuição zero-inflacionado Série de Potência. Van den Broek (1995) desenvolveu o teste escore para testar o excesso de zero no modelo zero-inflacionado Poisson. Já Gupta *et al.* (1996) desenvolveram o teste da razão de verossimilhança para testar a inflação de zeros no modelo zero-inflacionado Poisson generalizado. Ridout *et al.* (2001) propuseram o uso do teste escore para testar os modelos zero-inflacionado Poisson e zero-inflacionado Binomial Negativa, enquanto que Gupta *et al.* (2005) desenvolveram o teste escore para testar o parâmetro de inflação de zeros no modelo zero-inflacionado Poisson Generalizado. Xiang *et al.* (2006) desenvolveram o teste escore para o parâmetro de inflação de zeros em dados de contagem correlacionados com excesso de zeros. Deng & Zhang (2015) propuseram o teste escore para testar a incidência de zeros no modelo zero-inflacionado Binomial. Todem *et al.* (2012) estudaram a eficiência do teste escore para verificar a homogeneidade em modelos generalizados com excesso de zeros. Hsu *et al.* (2014) desenvolveram o teste de Wald para inflação ou deflação de zeros em dados de contagem. No entanto não temos conhecimento na literatura do desenvolvimento do *teste gradiente* (Terrell, 2002) em modelos para

dados de contagem com excesso de zeros. Neste trabalho desenvolvemos o *teste gradiente* para testar a inflação de zeros nos modelos zero-inflacionados Série de Potência.

A ideia básica da derivação de modelos para dados de contagem com excessos de zeros é misturar uma distribuição degenerada em zero com distribuições de probabilidade discretas (modelos de base) como Poisson, Geométrica, Binomial, Binomial Negativa, entre outras.

Gupta *et al.* (1995) introduziram o modelo Série de Potência inflacionado de zeros, ao misturar a distribuição Série de Potência com a distribuição degenerada em zero. O modelo tem a seguinte função de probabilidade

$$P(Z = z; \omega, \theta) = \begin{cases} \omega + (1 - \omega) \frac{a_0}{A(\theta)}, & z = 0 \\ (1 - \omega) \frac{a_z \theta^z}{A(\theta)}, & z = 1, 2, 3, \dots \end{cases} \quad (1.1)$$

em que $a_z > 0$ depende somente de z , $\theta \in (0, s)$ (s pode ser ∞) é o parâmetro de potência da série, $A(\theta) = \sum_{z \in \mathcal{K}} a_z \theta^z$ é a função série que atua como constante normalizadora, sendo finita e duas vezes diferenciável ($\mathcal{K}$ é um conjunto não vazio enumerável de inteiros não negativos) e ω é o parâmetro de inflação ou deflação de zeros, que pode assumir valores negativos.

Alguns modelos zero-inflacionados (ZI) importantes são as distribuições ZI-Binomial, ZI-Poisson, ZI-Binomial Negativa e ZI-Logarítmica. Por exemplo, se k é um inteiro positivo, $a_z = \binom{k}{n}$ e $A(\theta) = (1 + \theta)^k$, então (1.1) define a distribuição ZI-Binomial. Alguns membros desta família de distribuições são usados neste trabalho e são enumerados a seguir

$a_z = 1/z!$ e $A(\theta) = e^\theta$, $\theta > 0$: ZI-Poisson ,

$a_z = 1$ e $A(\theta) = (1 - \theta)^{-1}$, $0 < \theta < 10$: ZI-Geométrica ,

$a_z = \binom{k+z-1}{k-1}$ e $A(\theta) = (1 - \theta)^{-k}$, $0 < \theta < 1$: ZI-Binomial Negativa ,

$a_z = 1/(z+1)$ e $A(\theta) = -\log(1 - \theta)/\theta$, $0 < \theta < 1$: ZI-Logarítmica .

Note que quando $\omega = 0$ a distribuição em (1.1) se reduz à distribuição Série de Potência. A média e a variância para a classe de distribuição zero-inflacionados Série de Potência (ZISP) são dadas, respectivamente, por

$$E[Z] = \frac{(1 - \omega)\theta A'(\theta)}{A(\theta)} \text{ e } Var[Z] = \frac{(1 - \omega)\theta}{A(\theta)^2} (A(\theta)A'(\theta)A''(\theta) - (1 - \omega)\theta(A'(\theta))^2),$$

em que $A'(\theta) = dA(\theta)/d\theta$ e $A''(\theta) = d^2A(\theta)/d\theta^2$. As expressões da primeira, segunda e terceira derivadas da função $A(\theta)$ são escritas na Tabela 1.1 para facilitar seu uso no desenvolvimento deste trabalho.

A função geradora de probabilidades do modelo ZISP é dada por

$$G_Z(\xi) = \omega + (1 - \omega) \frac{A(\theta\xi)}{A(\theta)}, \quad |\xi| \leq 1.$$

Os modelos zero-inflacionados têm duas representações (Todem *et al.*, 2012): a representação hierárquica e a marginal. O modelo em (1.1) sob a representação hierárquica, ω representa o peso da mistura e é uma probabilidade, e $0 \leq \omega \leq 1$. Enquanto que na representação marginal, a distribuição de probabilidade, $0 \leq P(Z = z; \omega, \theta) < 1$, se $\omega + (1 - \omega)a_0/A(\theta) \geq 0$ o que implica, $\frac{-a_0}{A(\theta)-a_0} \leq \omega < 1$. Nessa representação ω é o parâmetro de inflação de zeros, quando $\omega > 0$, e é o parâmetro de deflação de zeros, quando $\omega < 0$. Neste estudo consideramos o modelo em (1.1) na representação marginal.

Tabela 1.1: Derivadas da função $A(\theta)$

Modelo	$A(\theta)$	$A'(\theta)$	$A''(\theta)$	$A'''(\theta)$
Poisson	e^θ	e^θ	e^θ	e^θ
Geométrico	$(1 - \theta)^{-1}$	$(1 - \theta)^{-2}$	$-2(1 - \theta)^{-3}$	$6(1 - \theta)^{-4}$
Logarítmica	$\frac{-\log(1-\theta)}{\theta}$	$\frac{\frac{\theta}{1-\theta} + \log(1-\theta)}{\theta^2}$	$\frac{\theta(3\theta-2) - 2(\theta-1)^2 \log(1-\theta)}{(1-\theta)^2 \theta^3}$	$\frac{(-11\theta^2 + 15\theta - 6)}{(\theta-1)^3 \theta^3} + \frac{6 \log(1-\theta)}{\theta^4}$

Os modelos de fragilidade são caracterizados pela introdução de um efeito aleatório, representado por uma variável aleatória contínua não observável no modelo. Este efeito aleatório, denominado fragilidade são utilizados para a modelagem de heterogeneidade na análise de dados de sobrevivência. Na análise desses dados a distribuição da fragilidade em general é assumida contínua (Hougaard, 1986, Vaupel *et al.*, 1979). Em alguns casos pode ser apropriado considerar distribuições discretas para a distribuição da fragilidade (Caroni *et al.*, 2010). Dados de sobrevivência contendo unidades experimentais em que o evento de interesse não aconteceu mesmo após um período longo de observação (dados de sobrevivência com fração de cura (Ibrahim *et al.*, 2001, Maller & Zhou, 1996). Nessa situação essas unidades experimentais têm fragilidade igual a zero. A fragilidade neste caso pode representar, por exemplo, a presença de um número desconhecido de fatores que levam à ocorrência do evento de interesse. Fragilidade zero corresponde a um modelo que contém uma proporção de indivíduos livre do evento de interesse (risco zero).

Neste trabalho, também propomos uma classe de modelos de sobrevivência induzidos por fragilidade discreta com distribuição dada em (1.1) para a modelagem de dados de sobrevivência com fração de cura (Maller & Zhou, 1996, Rodrigues *et al.*, 2009a). Uma análise estatística do modelo de sobrevivência induzido por fragilidade discreta é obtido considerando as abordagens clássica e Bayesiana. Além disso, o teste *gradiente* foi desenvolvido para testar as hipóteses a respeito da proporção de indivíduos com risco zero.

1.1 Estatística gradiente

Grande parte da literatura trata os problemas de teste de hipóteses em modelos paramétricos, tais como os modelos com excesso de zeros, os modelos de sobrevivência, entre outros. As estatísticas de teste usuais para grandes amostras são as estatísticas da razão de verossimilhança (Wilks, 1938), escore (Rao, 1948) e Wald (Wald, 1943), no entanto, recentemente uma nova estatística foi proposta por Terrell (2002) para resolver problemas de testes de hipóteses em alguns modelos paramétricos (Lemonte, 2016, Lemonte & Ferrari, 2011, 2012). A formulação das estatísticas para o caso de uma hipótese nula simples são apresentadas a seguir.

Considere que $y_1, \dots, y_n$ denotam n realizações de alguma característica de uma população com função densidade ou probabilidade, $f(y, \boldsymbol{\vartheta})$, $\boldsymbol{\vartheta} \in \Theta \subset R^p$. Quando não há parâmetros de perturbação, o interesse é testar a hipóteses nula simples $H_0 : \boldsymbol{\theta} = \boldsymbol{\vartheta}_0$ versus $H_1 : \boldsymbol{\vartheta} \neq \boldsymbol{\vartheta}_0$, em que $\boldsymbol{\vartheta}_0$ é um vetor especificado para $\boldsymbol{\vartheta} \in \Theta$. Sejam $\ell(\boldsymbol{\vartheta})$, $\mathbf{U}(\boldsymbol{\vartheta})$ e $\mathbf{K}(\boldsymbol{\vartheta})$ a log-verossimilhança, o vetor escore e a matriz de informação de Fisher relativas ao vetor $\boldsymbol{\vartheta}$, respectivamente. As estatísticas comumente usadas para testar H_0 versus H_1 são: razão de verossimilhança (S_{LR}), escore de Rao (S_R) e Wald (S_W), que são dadas, respectivamente, por

$$S_{LR} = 2[\ell(\hat{\boldsymbol{\vartheta}}) - \ell(\boldsymbol{\vartheta}_0)], \quad (1.2)$$

$$S_R = \mathbf{U}(\boldsymbol{\vartheta}_0)^\top \mathbf{K}(\boldsymbol{\vartheta}_0) \mathbf{U}(\boldsymbol{\vartheta}_0), \quad (1.3)$$

e

$$S_W = (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^\top \mathbf{K}^{-1}(\hat{\boldsymbol{\vartheta}}) (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0), \quad (1.4)$$

em que $\hat{\boldsymbol{\vartheta}}$ é o estimador de máxima verossimilhança de $\boldsymbol{\theta}$.

Terrell (2002) propôs uma estatística denominada gradiente. Esta estatística é obtida das estatísticas de escore e de Wald, e é dada por

$$S_G = \mathbf{U}(\boldsymbol{\vartheta}_0)^\top (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0). \quad (1.5)$$

Note que a estatística gradiente é muito mais simples de ser calculada, já que não envolve a matriz de informação observada ou esperada, tampouco a inversa da matriz. Lemonte & Ferrari (2011) e Lemonte *et al.* (2016) compararam por meio de simulações, os testes da razão de verossimilhança, escore, Wald e gradiente em modelos de Birbaun-Saunders com regressão linear e não linear, respectivamente. Lemonte & Ferrari (2012) obtiveram os poderes locais dos testes baseados com as estatísticas de razão de verossimilhança, escore, Wald e gradiente

para os parâmetros da família Exponencial uniparamétrica e concluíram que nenhuma das estatísticas é uniformemente superior às outras. O livro de [Lemonte \(2016\)](#) apresenta resultados analíticos e evidências numéricas do teste gradiente. A obtenção do teste gradiente é mostrado no Apêndice [A.1](#).

1.2 Objetivos

No contexto apresentado acima, o objetivo desta tese é propor procedimentos alternativos aos testes clássicos, para os testes de hipóteses dos parâmetros de uma família de distribuição discreta zero-inflacionado (distribuição ZISP) que unifica vários modelos. Além disso, propor um novo modelo de sobrevivência induzido por fragilidade discreta com distribuição ZISP, que inclui como caso particular o modelo de tempo de promoção ([Yakovlev *et al.*, 1993](#)) e o modelo proposto por [Cancho *et al.* \(2013\)](#), entre outros. Dessa forma, podemos relacionar os seguintes objetivos específicos que foram atingidos no decorrer desta tese:

- (i) desenvolver o *teste gradiente* para testar a inflação (ou deflação) no modelo ZISP dada em (1.1) e avaliar por meio de um estudo de simulação o teste para pequenas, moderadas e grandes tamanhos amostrais;
- (ii) fazer um estudo comparativo do poder do teste proposto com os testes escore, Wald e razão de verossimilhança para o procedimento descrito em (i);
- (iii) estudar as propriedades estruturais do modelo de sobrevivência induzida por fragilidade discreta com distribuição dada em (1.1) e desenvolver os procedimentos inferenciais por meio da perspectiva clássica e Bayesiana;
- (iv) desenvolver o teste gradiente para o novo modelo de sobrevivência com fração de cura.

1.3 Organização da tese

A tese se desenvolve como segue. No Capítulo 2, apresentamos o modelo ZISP na presença e ausência de covariáveis, obtemos a função escore e a matriz de informação de Fisher, desenvolvemos os testes, gradiente, escore, Wald e da razão de verossimilhança. Além disso, comparamos o poder dos testes das estatísticas de teste gradiente, escore, Wald e razão de verossimilhança por meio de um estudo de simulação e realizamos uma aplicação a um conjunto de dados reais. No Capítulo 3 apresentamos um novo modelo de sobrevivência induzido por fragilidade discreta com distribuição ZISP. Uma abordagem clássica utilizando o método de máxima verossimilhança é considerada para fazer a inferência sobre os parâmetros do modelo. Além disso, desenvolvemos o teste gradiente para avaliar a adequabilidade

do modelo de tempo de promoção. Também, realizamos um estudo de simulação com o objetivo de avaliar o desempenho do método de estimação e do teste proposto e, uma aplicação do modelo para um conjunto de dados reais da literatura. No Capítulo 4 desenvolvemos métodos Bayesianos de estimação para o modelo ZISP-FC usando o método Monte Carlo Cadeias de Markov (MCMC). Finalmente, as principais conclusões desta tese com base nos resultados obtidos, bem como algumas perspectivas futuras de pesquisa são discutidas no Capítulo 5.

Capítulo 2

Teste Gradiente no Modelo Zero-Inflacionado Série de Potência

Neste Capítulo, mostramos a obtenção da estatística *gradiente* para testar a inflação (ou deflação) de zeros no modelo zero-inflacionado Série de Potência dada em (1.1), ou seja, testar a hipótese nula composta, $H_0 : \omega = \omega_0$ contra $H_1 : \omega \neq \omega_0$, em que ω_0 é uma quantidade especificada para ω . Além disso, comparamos a performance da estatística *gradiente* com as estatísticas da razão de verossimilhança, escore e Wald para pequenas, moderadas e grandes amostras.

2.1 Os modelos Zero-Inflacionado Série de Potência

A distribuição do modelo zero-inflacionado Série de Potência (ZISP) dada em (1.1), reparametrizado em $\omega = \phi/(1 + \phi)$ (veja, [Van den Broek, 1995](#)), tem a função de probabilidade dada por

$$P(Z; \theta, \phi) = \begin{cases} \frac{\phi + a_0 A(\theta)^{-1}}{1 + \phi} & , z = 0 \\ \frac{a_z \theta^z A(\theta)^{-1}}{1 + \phi} & , z = 1, 2, 3, \dots \end{cases} \quad (2.1)$$

em que $\theta \in (0, s)$ (s pode ser ∞) e $-a_0/A(\theta) \leq \phi < \infty$. Como observado por [Van den Broek \(1995\)](#), com essa parametrização as estimativas de máxima verossimilhança dos parâmetros do modelo são muito mais estáveis numericamente. A média e a variância do ZIPS nesta parametrização são dadas, respectivamente, por

$$E[Z] = \frac{\theta A'(\theta)}{(1 + \phi)A(\theta)} \text{ e } Var[Z] = \frac{\theta^2 A''(\theta)}{(1 + \phi)A(\theta)} - \left[\frac{\theta A'(\theta)}{(1 + \phi)A(\theta)} \right]^2.$$

Note que o modelo ZISP se reduz ao modelo Série de Potência quando $\phi = 0$ e $\phi > 0$ indica sobredispersão. Também pode-se demonstrar que o modelo (2.1) pertence à família Exponencial bi-paramétrica (veja a Proposição 1 no Apêndice B.2).

Considere $Z_1, \dots, Z_n$ uma amostra aleatória de tamanho n de uma população com distribuição dada em (2.1). O logaritmo da função verossimilhança de $\boldsymbol{\vartheta} = (\theta, \phi)$, dada a amostra observada $\mathbf{z} = (z_1, \dots, z_n)$, é dada por

$$\begin{aligned} \ell(\boldsymbol{\vartheta}) = \sum_i \{ & -\log(1 + \phi) + I_{(z_i=0)} \log(\phi + a_0 A(\theta)^{-1}) + I_{(z_i>0)} [\log(a_{z_i}) \\ & + z_i \log(\theta) - \log(A(\theta))] \}. \end{aligned} \quad (2.2)$$

em que $I_{\{Z_i=0\}}$ é a função indicadora sendo igual a 1 se $Z_i = 0$, e 0 em caso contrário. Os estimadores de máxima verossimilhança (EMV) dos parâmetros do modelo ZISP, denotados por $\hat{\theta}$ e $\hat{\phi}$, são os valores de θ e ϕ , respectivamente, que maximizam a função log-verossimilhança.

Com a primeira derivada da função de log-verossimilhança dada em (2.2) em relação a cada parâmetro do modelo, obtemos a função (ou vetor) escore $\mathbf{U}(\boldsymbol{\vartheta}) = (U_1(\boldsymbol{\vartheta}), U_2(\boldsymbol{\vartheta}))^\top$, cujos elementos são dados por

$$\begin{aligned} U_1(\boldsymbol{\vartheta}) &= \frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \theta} = \frac{-a_0 n_0}{\phi + a_0 A(\theta)^{-1}} \frac{A'(\theta)}{A(\theta)^2} + \left((n - n_0) \frac{\bar{z}}{\theta} - \frac{A'(\theta)}{A(\theta)} \right), \\ U_2(\boldsymbol{\vartheta}) &= \frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \phi} = \frac{-n}{1 + \phi} + \frac{n_0}{\phi + a_0 A(\theta)^{-1}}, \end{aligned} \quad (2.3)$$

em que $n_0 = \sum_i I_{(z_i=0)}$ é a contagem de todas as observações iguais a zero.

Os EMV de θ e ϕ são obtidos resolvendo o sistema de equações não lineares dadas em (2.3), fazendo $\mathbf{U}(\boldsymbol{\vartheta}) = \mathbf{0}$. Dessas equações obtemos a seguinte relação entre os EMV dos parâmetros do modelo ZISP

$$\begin{aligned} \hat{\phi} &= \frac{n_0 - n a_0 A(\hat{\theta})^{-1}}{n - n_0}, \\ \bar{z} &= \frac{\hat{\theta} A'(\hat{\theta})}{A(\hat{\theta}) - a_0}, \end{aligned} \quad (2.4)$$

em que $\bar{z} = \frac{\sum_i z_i I_{(z_i>0)}}{n - n_0}$ é a média das observações positivas. As equações em (2.4) para os modelos ZI-Poisson e ZI-Geométrica têm soluções explícitas. Inferências sobre os parâmetros do modelo podem ser baseadas na normalidade assintótica que, sob certas condições de regularidade (Cox & Hinkley, 1979), têm distribuição assintoticamente Normal com vetor de médias $\boldsymbol{\vartheta}$ e matriz de variância-covariância, a inversa da matriz de informação de Fisher avaliada no EMV de $\boldsymbol{\vartheta}$.

A matriz de informação de Fisher para $\boldsymbol{\vartheta}$ é dada por

$$K(\boldsymbol{\vartheta}) = -E \left[\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}^\top \partial \boldsymbol{\vartheta}} \right] = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}, \quad (2.5)$$

em que $K_{11} = -E \left[\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \theta^2} \right]$, $K_{22} = -E \left[\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \phi^2} \right]$ e $K_{12} = K_{21} = -E \left[\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \theta \partial \phi} \right]$, sendo que cada componente da matriz de informação de Fisher para $\boldsymbol{\vartheta}$ é dado por

$$\begin{aligned} K_{11} &= \frac{na_0}{1+\phi} \left[\frac{A(\theta)A''(\theta)(a_0 + \phi A(\theta)) - A'(\theta)^2(a_0 + 2\phi A(\theta))}{A(\theta)^4(\phi + a_0 A(\theta)^{-1})} \right] \\ &\quad + \frac{n}{1+\phi} \left[\frac{\theta A'(\theta)}{A(\theta)} + \frac{A(\theta)A''(\theta) - (A'(\theta))^2}{A(\theta)^2} (1 - a_0 A(\theta)^{-1}) \right], \\ K_{12} = K_{21} &= \frac{na_0}{(1+\phi)(\phi + a_0 A(\theta)^{-1})} \frac{A'(\theta)}{A(\theta)^2} \text{ e } K_{22} = \frac{-n(1 - a_0 A(\theta)^{-1})}{(1+\phi)^2(\phi + a_0 A(\theta)^{-1})}. \end{aligned}$$

A inversa de matriz de informação de Fisher é dada por

$$K^{-1}(\boldsymbol{\vartheta}) = \begin{bmatrix} K^{11} & K^{12} \\ K^{21} & K^{22} \end{bmatrix} \quad (2.6)$$

em que $K^{11} = (K_{11} - K_{12}K_{22}^{-1}K_{21})^{-1}$, $K^{12} = K^{21} = -K_{11}^{-1}K_{12}(K_{22} - K_{21}K_{11}^{-1}K_{12})^{-1}$ e $K^{22} = (K_{22} - K_{21}K_{11}^{-1}K_{12})^{-1}$. As componentes K^{jj} , $j = 1, 2$ é a variância do EMV $\boldsymbol{\vartheta}_j$ e K^{jk} , $j \neq k = 1, 2$, é a covariância entre os EMV de $\boldsymbol{\vartheta}_j$ e $\boldsymbol{\vartheta}_k$.

A hipótese de interesse é testar se o parâmetro de inflação (ou deflação) de zeros no modelo ZISP, pode assumir um valor pré-especificado, ou seja, tem-se interesse em testar as hipóteses

$$H_0 : \phi = \phi_0, \text{ contra } H_1 : \phi \neq \phi_0, \quad (2.7)$$

em que ϕ_0 é valor especificado de ϕ . A estatística gradiente para testar as hipóteses em (2.7) é dada por

$$S_G = \left(\frac{-n}{1+\phi_0} + \frac{n_0}{\phi_0 + a_0 A(\tilde{\theta})^{-1}} \right) (\hat{\phi} - \phi_0), \quad (2.8)$$

em que $\tilde{\theta}$ é o EMV de θ sob a hipóteses nula e $\hat{\phi}$ é o EMV no espaço paramétrico.

As estatísticas tradicionais usadas para testar as hipóteses em (2.7) são: a estatística

razão de verossimilhança dada por

$$S_{\text{LR}} = 2 \left[-n \log \left(\frac{1 + \hat{\phi}}{1 + \phi_0} \right) + n_0 \log \left(\frac{\hat{\phi} + a_0 A(\hat{\theta})^{-1}}{\phi_0 + a_0 A(\tilde{\theta})^{-1}} \right) \right. \\ \left. + (n - n_0) \left(\bar{z} \log \left(\frac{\hat{\theta}}{\tilde{\theta}} \right) - \log \left(\frac{A(\hat{\theta})}{A(\tilde{\theta})} \right) \right) \right], \quad (2.9)$$

em que $\tilde{\theta}$ é o EMV de θ sob H_0 e as estatísticas escore e de Wald dadas, respectivamente, por

$$S_{\text{R}} = \left(\frac{-n}{1 + \phi_0} + \frac{n_0}{\phi_0 + a_0 A(\tilde{\theta})^{-1}} \right)^2 K^{22}(\tilde{\boldsymbol{\vartheta}}) \quad (2.10)$$

e

$$S_{\text{W}} = K^{22}(\tilde{\boldsymbol{\vartheta}})^{-1} (\hat{\phi} - \phi_0)^2, \quad (2.11)$$

em que

$$K^{22}(\tilde{\boldsymbol{\vartheta}}) = \left[\left(\frac{-n(1 - a_0 A(\tilde{\theta})^{-1})}{(1 + \phi_0)^2 (\phi_0 + a_0 A(\tilde{\theta})^{-1})} \right) - \left(\frac{na_0}{(1 + \phi_0)(\phi_0 + a_0 A(\tilde{\theta})^{-1})} \frac{A'(\tilde{\theta})}{A(\tilde{\theta})^2} \right)^2 K_{11}^{-1} \right]^{-1}$$

com

$$K_{11}(\tilde{\boldsymbol{\vartheta}}) = \frac{na_0}{1 + \phi_0} \left[\frac{A(\tilde{\theta})A''(\tilde{\theta})(a_0 + \phi_0 A(\tilde{\theta})) - A'(\tilde{\theta})^2(a_0 + 2\phi_0 A(\tilde{\theta}))}{A(\tilde{\theta})^4(\phi_0 + a_0 A(\tilde{\theta})^{-1})} \right] \\ + \frac{n}{1 + \phi_0} \left[\frac{\tilde{\theta} A'(\tilde{\theta})}{A(\tilde{\theta})} + \frac{A(\tilde{\theta})A''(\tilde{\theta}) - (A'(\tilde{\theta}))^2}{A(\tilde{\theta})^2} (1 - a_0 A(\tilde{\theta})^{-1}) \right].$$

As estatísticas de teste em (2.8)-(2.11), sob H_0 , têm distribuição assintoticamente Qui-quadrado com *um* grau de liberdade.

2.1.1 Casos particulares

Nesta subseção apresentamos alguns casos particulares do modelo ZISP, conjuntamente com resultados obtidos, tais como o vetor escore, a matriz de informação de Fisher e as estatísticas de teste dadas em (2.8)-(2.11) para testar a hipótese H_0 dada em (2.7). Quando $A(\theta) = e^\theta$, $\theta > 0$ e $a_z = 1/z!$ no modelo em (2.1), obtém-se o modelo ZI-Poisson

(ZIP) cuja função de probabilidade é dada por

$$P(Z_i; \phi, \theta) = \begin{cases} \frac{\phi + e^{-\theta}}{1 + \phi} & , z_i = 0 \\ \frac{e^{-\theta} \theta^{z_i}}{(1 + \phi) z_i!} & , z_i > 0 \end{cases} \quad (2.12)$$

em que $-e^{-\theta} \leq \phi < \infty$ é o parâmetro de inflação ou deflação de zeros.

De (2.3), os elementos do vetor escore são dados por

$$\begin{aligned} U_1(\boldsymbol{\vartheta}) &= \frac{-n_0 e^{-\theta}}{\phi + e^{-\theta}} - (n - n_0) \left(1 - \frac{\bar{z}}{\theta}\right) \\ U_2(\boldsymbol{\vartheta}) &= \frac{-n}{1 + \phi} + \frac{n_0}{\phi + e^{-\theta}}. \end{aligned} \quad (2.13)$$

O EMV de ϕ é dado por

$$\hat{\phi} = \frac{n_0 - n e^{-\hat{\theta}}}{n - n_0}, \quad (2.14)$$

em que $\hat{\theta}$ é o EMV de θ , que resulta da solução de $\hat{\theta}$ na equação

$$\bar{z} = \frac{\hat{\theta}}{1 - e^{-\hat{\theta}}}.$$

Considerando a função de Lambert W (Weisstein, 2002), o EMV do parâmetro θ é dado por

$$\hat{\theta} = \bar{z} + LW(-\bar{z} e^{-\bar{z}}). \quad (2.15)$$

De (2.5) temos que a matriz de informação de Fisher dos parâmetros do modelo ZIP resulta em

$$K(\theta, \phi) = \begin{bmatrix} -\frac{n}{(1+\phi)} \left(\frac{\phi e^{-\theta}}{(\phi + e^{-\theta})} - \frac{1}{\theta} \right) & -\frac{n e^{-\theta}}{(\phi + e^{-\theta})(1+\phi)} \\ -\frac{n e^{-\theta}}{(\phi + e^{-\theta})(1+\phi)} & n \left(\frac{1 - e^{-\theta}}{(1+\phi)^2 (\phi + e^{-\theta})} \right) \end{bmatrix}. \quad (2.16)$$

A estatística gradiente dada em (2.8) para testar a hipótese H_0 em (2.7) no modelo ZIP resulta em

$$S_G = \left(\frac{-n}{1 + \phi_0} + \frac{n_0}{\phi_0 + e^{-\hat{\theta}}} \right) (\hat{\phi} - \phi_0),$$

em que $\hat{\theta}$ é o EMV de θ sob H_0 , que nesse caso é a média amostral dos dados de contagem.

As estatísticas de teste da razão de verossimilhança, escore e Wald dadas em (2.9)-(2.11) para testar a hipótese H_0 em (2.7), correspondente ao modelo ZIP são dadas, respectivamente por

$$S_{LR} = 2 \left[-n \log \left(\frac{1 + \hat{\phi}}{1 + \phi_0} \right) + n_0 \log \left(\frac{\hat{\phi} + e^{-\hat{\theta}}}{\phi_0 + e^{-\tilde{\theta}}} \right) + (n - n_0) \left(\bar{z} \log \left(\frac{\hat{\theta}}{\tilde{\theta}} \right) - (\hat{\theta} - \tilde{\theta}) \right) \right],$$

$$S_R = \left(\frac{-n}{1 + \phi_0} + \frac{n_0}{\phi_0 + e^{-\tilde{\theta}}} \right)^2 K^{22}(\tilde{\theta}, \phi_0)$$

e

$$S_W = (\hat{\phi} - \phi_0)^2 K^{22}(\hat{\theta}, \hat{\phi})^{-1},$$

sendo

$$K^{22}(\tilde{\theta}, \phi_0) = \left\{ n \left(\frac{1 - e^{-\tilde{\theta}}}{(1 + \phi_0)^2 (\phi_0 + e^{-\tilde{\theta}})} \right) - \left(-\frac{ne^{-\tilde{\theta}}}{(\phi_0 + e^{-\tilde{\theta}})(1 + \phi_0)} \right)^2 \left(-\frac{n}{(1 + \phi_0)} \left(\frac{\phi_0 e^{-\tilde{\theta}}}{(\phi_0 + e^{-\tilde{\theta}})} - \frac{1}{\tilde{\theta}} \right) \right)^{-1} \right\}^{-1},$$

e

$$K^{22}(\hat{\theta}, \hat{\phi}) = \left\{ n \left(\frac{1 - e^{-\hat{\theta}}}{(1 + \hat{\phi})^2 (\hat{\phi} + e^{-\hat{\theta}})} \right) - \left(-\frac{ne^{-\hat{\theta}}}{(\hat{\phi} + e^{-\hat{\theta}})(1 + \hat{\phi})} \right)^2 \left(-\frac{n}{(1 + \hat{\phi})} \left(\frac{\hat{\phi} e^{-\hat{\theta}}}{(\hat{\phi} + e^{-\hat{\theta}})} - \frac{1}{\hat{\theta}} \right) \right)^{-1} \right\}^{-1}.$$

Agora, no caso em que $A(\theta) = (1 - \theta)^{-1}$, $\theta \in (0, 1)$ e $a_{z_i} = 1$ no modelo ZISP em (2.1), obtém-se o modelo ZI-Geométrica (ZIG), com função de probabilidade dada por

$$P(Z_i; \phi, \theta) = \begin{cases} \frac{\phi + (1 - \theta)}{1 + \phi} & , z_i = 0 \\ \frac{\theta^{z_i}(1 - \theta)}{(1 + \phi)} & , z_i > 0 \end{cases} \quad (2.17)$$

em que $-(1 - \theta) \leq \phi < \infty$ é o parâmetro de inflação ou deflação de zeros.

De (2.3), os elementos do vetor escore são dados por

$$\begin{aligned} U_1(\boldsymbol{\vartheta}) &= \frac{-n_0}{\phi + (1 - \theta)} + (n - n_0) \left(\frac{\bar{z}}{\theta} - \frac{1}{(1 - \theta)} \right), \\ U_2(\boldsymbol{\vartheta}) &= \frac{-n}{1 + \phi} + \frac{n_0}{\phi + (1 - \theta)}. \end{aligned} \quad (2.18)$$

A partir de (2.4), obtém-se os EMVs dos parâmetros do modelo ZIG que resultam em

$$\hat{\phi} = \frac{n_0 - n(1 - \hat{\theta})}{n - n_0} \quad (2.19)$$

e

$$\hat{\theta} = 1 - 1/\bar{z}. \quad (2.20)$$

De (2.5), obtemos a matriz de informação dos parâmetros do modelo ZIG dada por

$$K(\theta, \phi) = \begin{bmatrix} \frac{n}{1+\phi} \left(\frac{1}{\phi+(1-\theta)} - \frac{\theta}{(1-\theta)^2} \right) & \frac{-n}{(1+\phi)(\phi+(1-\theta))} \\ \frac{-n}{(1+\phi)(\phi+(1-\theta))} & \frac{n(1-(1-\theta))}{(1+\phi)^2(\phi+(1-\theta))} \end{bmatrix}.$$

A estatística gradiente (2.8) para testar a hipótese H_0 em (2.7), para o modelo ZIG, é dada por

$$S_G = \left(\frac{-n}{1+\phi_0} + \frac{n_0}{\phi_0 + (1-\tilde{\theta})} \right) (\hat{\phi} - \phi_0),$$

em que $\tilde{\theta}$ é o EMV de θ sob a hipótese nula.

As estatísticas do teste da razão de verossimilhança, escore e Wald dadas em (2.9)-(2.11) para testar a hipótese H_0 em (2.7), correspondente ao modelo ZIG são dadas, respectivamente, por

$$S_{LR} = 2 \left[-n \log \left(\frac{1+\hat{\phi}}{1+\phi_0} \right) + n_0 \log \left(\frac{\hat{\phi} + (1-\hat{\theta})}{\phi_0 + (1-\tilde{\theta})} \right) + (n - n_0) \left(\bar{z} \log \left(\frac{\hat{\theta}}{\tilde{\theta}} \right) + \log \left(\frac{1-\hat{\theta}}{1-\tilde{\theta}} \right) \right) \right],$$

$$S_R = \left(\frac{-n}{1+\phi_0} + \frac{n_0}{\phi_0 + (1-\tilde{\theta})} \right)^2 K^{22}(\tilde{\theta}, \phi_0),$$

$$S_W(\hat{\phi} - \phi_0)^2 K^{22}(\hat{\theta}, \hat{\phi})^{-1},$$

em que

$$K^{22}(\tilde{\theta}, \phi_0) = \left(\frac{n(1-(1-\tilde{\theta}))}{(1+\phi_0)^2(\phi_0 + (1-\tilde{\theta}))} - \left(\frac{n}{(1+\phi_0)(\phi_0 + (1-\tilde{\theta}))} \right)^2 \left[\frac{n}{1+\phi_0} \left(\frac{1}{\phi_0 + (1-\tilde{\theta})} + \frac{\tilde{\theta}}{1-\tilde{\theta}} \right) \right]^{-1} \right)^{-1},$$

$$K^{22}(\hat{\theta}, \hat{\phi}) = \left(\frac{n(1-(1-\hat{\theta}))}{(1+\hat{\phi})^2(\hat{\phi} + (1-\hat{\theta}))} - \left(\frac{n}{(1+\hat{\phi})(\hat{\phi} + (1-\hat{\theta}))} \right)^2 \left[\frac{n}{1+\hat{\phi}} \left(\frac{1}{\hat{\phi} + (1-\hat{\theta})} + \frac{\hat{\theta}}{1-\hat{\theta}} \right) \right]^{-1} \right)^{-1}.$$

Finalmente, quando $A(\theta) = -\frac{\log(1-\theta)}{\theta}$, $\theta \in (0, 1)$ e $a_{z_i} = \frac{1}{z_i+1}$ em (2.1), o modelo ZISP se reduz ao modelo ZI-Logaritmica (ZIL) com função de probabilidade dada por

$$P(Z_i; \phi, \theta) = \begin{cases} \frac{1}{1+\phi} \left(\phi - \frac{\theta}{\log(1-\theta)} \right) & , z_i = 0 \\ \frac{-1}{1+\phi} \frac{\theta^{z_i+1}}{(z_i+1) \log(1-\theta)} & , z_i > 0 \end{cases}$$

em que $\frac{\theta}{\log(1-\theta)} < \phi < \infty$ é o parâmetro de inflação ou deflação de zeros.

De (2.3), temos que os elementos do vetor escore são dados por

$$U_1(\boldsymbol{\vartheta}) = \frac{-n_0}{\phi \log(1-\theta) - \theta} \frac{\frac{\theta}{1-\theta} + \log(1-\theta)}{\log(1-\theta)} + (n - n_0) \left(\frac{\bar{z}}{\theta} + \frac{1}{(1-\theta) \log(1-\theta)} - \frac{1}{\theta} \right),$$

$$U_2(\boldsymbol{\vartheta}) = \frac{-n}{1+\phi} + \frac{n_0 \log(1-\theta)}{\phi \log(1-\theta) - \theta}.$$
(2.21)

De (2.4) os EMVs dos parâmetros do modelo ZIL resultam da solução em $\hat{\theta}$ e $\hat{\phi}$ nas equações

$$\hat{\phi} = \frac{n_0 \log(1-\hat{\theta}) + n\hat{\theta}}{(n - n_0) \log(1-\hat{\theta})} \quad e \quad \bar{z} = \frac{\frac{\hat{\theta}}{1-\hat{\theta}} + \log(1-\hat{\theta})}{-\log(1-\hat{\theta}) - \hat{\theta}}.$$
(2.22)

De (2.5) obtemos a matriz de informação dos parâmetros do modelo ZIL dada por

$$K(\theta, \phi) = \begin{bmatrix} -n \frac{\phi - \frac{\theta}{\log(1-\theta)}}{1+\phi} \mathbf{A} + \frac{n(\theta+1 - \frac{\theta}{\log(1-\theta)})}{\theta^2(1+\phi)1 - \frac{\theta}{\log(1-\theta)}} - n \frac{1 - \frac{\theta}{\log(1-\theta)}}{1+\phi} \mathbf{B} & -n \frac{\phi - \frac{\theta}{\log(1-\theta)}}{1+\phi} \frac{(\theta-1) \log(1-\theta) - \theta}{(\theta-1)(\phi \log(1-\theta) - \theta)^2} \\ -n \frac{\phi - \frac{\theta}{\log(1-\theta)}}{1+\phi} \frac{(\theta-1) \log(1-\theta) - \theta}{(\theta-1)(\phi \log(1-\theta) - \theta)^2} & \frac{-n}{(1+\phi)^2} + n \frac{\phi - \frac{\theta}{\log(1-\theta)}}{1+\phi} \frac{\log^2(1-\theta)}{(\phi \log(1-\theta) - \theta)^2} \end{bmatrix}$$

em que

$$\mathbf{A} = \frac{-2\theta - (2-\theta) \log(1-\theta)}{(\phi + A(\theta)^{-1})(1-\theta)^2 \log^3(1-\theta)} + \frac{\theta + (1-\theta \log(1-\theta))}{(\phi + A(\theta)^{-1})^2(1-\theta) \log(1-\theta)}$$

$$\mathbf{B} = \frac{-\theta^2 + \theta^2 \log(1-\theta) - (1-\theta)^2 \log^2(1-\theta)}{(1-\theta)^2 \theta^2 \log^2(1-\theta)}$$

A estatística gradiente (2.8) para testar a hipótese H_0 em (2.7) para o modelo ZIL, é dada por

$$S_G = \left(\frac{-n}{1+\phi_0} + \frac{n_0 \log(1-\tilde{\theta})}{\phi_0 \log(1-\tilde{\theta}) - \tilde{\theta}} \right) (\hat{\phi} - \phi_0),$$

em que $\tilde{\theta}$ é o EMV de θ sob a hipótese nula.

As estatísticas do teste da razão de verossimilhança, escore e Wald dadas em (2.9)-(2.11) para testar a hipótese H_0 em (2.7), correspondente ao modelo ZIL, são dadas, respectivamente, por

$$S_{LR} = 2 \left[-n \log \left(\frac{1 + \hat{\phi}}{1 + \phi_0} \right) + n_0 \log \left(\frac{\hat{\phi} - \frac{\tilde{\theta}}{\log(1-\tilde{\theta})}}{\phi_0 - \frac{\tilde{\theta}}{\log(1-\tilde{\theta})}} \right) + (n - n_0) \left(\bar{z} \log \left(\frac{\hat{\theta}}{\tilde{\theta}} \right) - \frac{\tilde{\theta}}{\tilde{\theta}} \log \left(\frac{1 - \hat{\theta}}{1 - \tilde{\theta}} \right) \right) \right],$$

$$S_R = \left(\frac{-n}{1 + \phi_0} + \frac{n_0 \log(1 - \tilde{\theta})}{\phi_0 \log(1 - \tilde{\theta}) - \tilde{\theta}} \right)^2 K^{22}(\tilde{\theta}, \phi_0)$$

e

$$S_W = (\hat{\phi} - \phi_0)^2 (K^{22}(\hat{\theta}, \hat{\phi}))^{-1},$$

em que

$$K^{22}(\tilde{\theta}, \phi_0) = \left(\frac{-n}{(1 + \phi_0)^2} + n \frac{\phi_0 - \frac{\tilde{\theta}}{\log(1-\tilde{\theta})}}{1 + \phi_0} \frac{\log^2(1 - \tilde{\theta})}{(\phi_0 \log(1 - \tilde{\theta}) - \tilde{\theta})^2} - \left(\frac{n}{(1 + \phi_0)(\phi_0 - \frac{\tilde{\theta}}{\log(1-\tilde{\theta})})} \frac{\frac{\tilde{\theta}}{1-\tilde{\theta}} + \log(1 - \tilde{\theta})}{\log^2(1 - \tilde{\theta})} \right)^2 \tilde{K}_{11}^{-1} \right)^{-1}$$

e

$$\begin{aligned} \tilde{K}_{11} = & -n \frac{\phi_0 - \frac{\tilde{\theta}}{\log(1-\tilde{\theta})}}{1 + \phi_0} \tilde{\mathbf{A}} + \frac{n(\tilde{\theta} + 1 - \frac{\tilde{\theta}}{\log(1-\tilde{\theta})})}{\tilde{\theta}^2(1 + \phi_0)1 - \frac{\tilde{\theta}}{\log(1-\tilde{\theta})}} - n \frac{1 - \frac{\tilde{\theta}}{\log(1-\tilde{\theta})}}{1 + \phi_0} \tilde{\mathbf{B}} \\ & - n \frac{\phi_0 - \frac{\tilde{\theta}}{\log(1-\tilde{\theta})}}{1 + \phi_0} \frac{(\tilde{\theta} - 1) \log(1 - \tilde{\theta}) - \tilde{\theta}}{(\tilde{\theta} - 1)(\phi_0 \log(1 - \tilde{\theta}) - \tilde{\theta})^2}. \end{aligned}$$

2.2 Teste gradiente no modelo de regressão zero-inflacionado Série de Potência

Em muitas aplicações práticas é comum assumir que os parâmetros do modelo dependam de covariáveis. Nesta Seção, consideramos covariáveis no parâmetro de potência θ no modelo (2.1) e desenvolvemos a estatística gradiente para o modelo de regressão zero-inflacionado Série de Potência. Para a i -ésima observação, o parâmetro θ_i relaciona-se com as covariáveis $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})$ por meio de uma função de ligação da seguinte forma

$$g(\theta_i) = \eta_i = \eta(\mathbf{x}_i, \boldsymbol{\beta}), \quad i = 1, \dots, n. \quad (2.23)$$

em que $g(\cdot)$, representa uma função de ligação contínua monótona e duas vezes diferenciável com domínio R^+ , $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ ($p < n$) é o vetor de parâmetros de regressão ($\boldsymbol{\beta} \in R^p$) associado ao vetor de covariáveis $\mathbf{x}_i$ e $\eta(\cdot; \cdot)$ é uma função possivelmente linear ou não-linear, contínua e diferenciável com respeito aos componentes de $\boldsymbol{\beta}$. Seja $X = X(\boldsymbol{\beta}) = \partial\eta/\partial\boldsymbol{\beta}^\top$ a matriz de derivadas de $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)$ com respeito a $\boldsymbol{\beta}^\top$ e assume-se que o

$posto(X) = p$.

Para $Z_1, \dots, Z_n$ uma amostra aleatória de tamanho n em que cada Z_i segue o modelo de regressão ZISP ($Z_i \sim \text{ZISP}(\theta_i, \phi)$) e com a expressão (2.23), temos que o logaritmo da função de verossimilhança de $\boldsymbol{\vartheta}$ dada a amostra observada $\mathbf{z} = (z_1, \dots, z_n)$ pode ser escrita como

$$\begin{aligned} \ell(\boldsymbol{\vartheta}) = \sum_{i=1}^n \{ & -\log(1 + \phi) + I_{(z_i=0)} \log(\phi + a_0 A(\theta_i)^{-1}) + I_{(z_i>0)} [\log(a_{z_i}) \\ & + z_i \log(\theta_i) - \log(A(\theta_i))] \}. \end{aligned} \quad (2.24)$$

Diferenciando (2.24) em relação aos parâmetros do modelo, $\boldsymbol{\beta}$ e ϕ , obtemos o vetor escore $\mathbf{U}^\top(\boldsymbol{\vartheta}) = (U_1^\top(\boldsymbol{\vartheta}), U_2(\boldsymbol{\vartheta}))$, cujos elementos são

$$U_1(\boldsymbol{\vartheta}) = \frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^n \left\{ I_{(z_i=0)} \frac{-a_0 A'(\theta_i) A(\theta_i)^{-2}}{\phi + a_0 A(\theta_i)^{-1}} + I_{(z_i>0)} \left[\frac{z_i}{\theta_i} - \frac{A'(\theta_i)}{A(\theta_i)} \right] \right\} \frac{\partial \theta_i}{\partial \boldsymbol{\beta}}, \quad (2.25)$$

$$U_2(\boldsymbol{\vartheta}) = \frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \phi} = \frac{-n}{1 + \phi} + \sum_{i=1}^n \left\{ I_{(z_i=0)} \frac{1}{\phi + a_0 A(\theta_i)^{-1}} \right\},$$

em que $\frac{\partial \theta_i}{\partial \boldsymbol{\beta}} = \frac{\partial g^{-1}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\beta}}$.

Os EMV de $\boldsymbol{\vartheta}$ denotado por $\hat{\boldsymbol{\vartheta}}$ podem ser obtidos ao maximizar a função log-verossimilhança (2.24), o qual é equivalente a resolver o sistema de equações $U_1(\hat{\boldsymbol{\vartheta}}) = \mathbf{0}$ e $U_2(\hat{\boldsymbol{\vartheta}}) = 0$. Estas equações são não lineares e não possuem solução explícita, então consideramos o algoritmo do Newton-Rapson, implementado no pacote *pscl* no *software* R (R Development Core Team, 2009) para estimar as soluções.

Sob usuais condições de regularidade, de (2.5) temos que a matriz de informação de Fisher dos parâmetros $\boldsymbol{\beta}$ e ϕ no modelo de regressão ZISP tem como componentes

$$\begin{aligned} K_{\boldsymbol{\beta}\boldsymbol{\beta}} = -E \left(\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right) &= \frac{a_0}{1 + \phi} \sum_{i=1}^n \left[\frac{A''(\theta_i)}{A(\theta_i)^2} - \frac{(1 + 2A(\theta_i))}{\phi + a_0 A(\theta_i)^{-1}} \frac{(A'(\theta_i))^2}{A(\theta_i)^4} \right] \frac{\partial \theta_i}{\partial \boldsymbol{\beta}^\top} \frac{\partial \theta_i}{\partial \boldsymbol{\beta}} \\ &\quad - \sum_{i=1}^n \left[\frac{\mu_i}{\theta_i^2} + \frac{1 - a_0 A(\theta_i)^{-1}}{1 + \phi} \left(\frac{A''(\theta_i)}{A(\theta_i)} - \frac{(A'(\theta_i))^2}{A(\theta_i)^2} \right) \right] \frac{\partial \theta_i}{\partial \boldsymbol{\beta}^\top} \frac{\partial \theta_i}{\partial \boldsymbol{\beta}} \\ &\quad - \frac{a_0}{1 + \phi} \sum_{i=1}^n \frac{A'(\theta_i)}{A(\theta_i)^2} \frac{\partial^2 \theta_i}{\partial \boldsymbol{\beta}^\top \partial \boldsymbol{\beta}} + \sum_{i=1}^n \left[\frac{\mu_i}{\theta_i} - \frac{1 - a_0 A(\theta_i)^{-1}}{1 + \phi} \frac{A'(\theta_i)}{A(\theta_i)} \right] \frac{\partial^2 \theta_i}{\partial \boldsymbol{\beta}^\top \partial \boldsymbol{\beta}}, \end{aligned} \quad (2.26)$$

$$K_{\beta\phi} = -E \left(\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \beta \partial \phi} \right) = \frac{a_0}{1+\phi} \sum_{i=1}^n \frac{A'(\theta_i)}{(\phi + a_0 A(\theta_i)^{-1}) A(\theta_i)^2} \frac{\partial \theta_i}{\partial \beta} \quad (2.27)$$

e

$$K_{\phi\phi} = -E \left(\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \phi^2} \right) = -\frac{1}{(1+\phi)^3} \sum_{i=1}^n \phi + a_0 A(\theta_i)^{-1} + \frac{1}{1+\phi} \sum_{i=1}^n \frac{1 - a_0 A(\theta_i)^{-1}}{(\phi + a_0 A(\theta_i)^{-1})^2}, \quad (2.28)$$

em que $\frac{\partial^2 \theta_i}{\partial \beta^\top \partial \beta} = \frac{\partial g^{-1}(\eta)}{\partial \eta} \frac{\partial^2 \eta}{\partial \beta^\top \partial \beta} + \frac{\partial^2 g^{-1}(\eta)}{\partial \eta^2} \left(\frac{\partial \eta}{\partial \beta} \right)^2$ e $E[z_i I(z_i > 0)] = \mu_i = \frac{\theta_i A'(\theta_i)}{(1+\phi)A(\theta_i)}$.

Note que $K_{\beta\beta}$ representa uma submatriz de dimensão $(p \times p)$, $K_{\beta\phi} = K_{\phi\beta}$ com dimensão $p \times 1$ e a dimensão de $K_{\phi\phi}$ é 1×1 correspondente a um valor escalar.

O interesse é formular a estatística gradiente sob os testes de hipóteses em relação aos parâmetros do modelo de regressão ZISP. Sejam $\boldsymbol{\vartheta}_1$ e $\boldsymbol{\vartheta}_2$ subconjuntos disjuntos do vetor de parâmetros $\boldsymbol{\vartheta}$ e geram uma partição do vetor escore e a matriz de informação de Fisher, dadas por

$$U(\boldsymbol{\vartheta}) = \begin{bmatrix} U_{\boldsymbol{\vartheta}_1} \\ U_{\boldsymbol{\vartheta}_2} \end{bmatrix}, \quad K(\boldsymbol{\vartheta}) = \begin{bmatrix} K_{\boldsymbol{\vartheta}_1 \boldsymbol{\vartheta}_1} & K_{\boldsymbol{\vartheta}_1 \boldsymbol{\vartheta}_2} \\ K_{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_1}^\top & K_{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_2} \end{bmatrix}. \quad (2.29)$$

em que $K_{\boldsymbol{\vartheta}_1 \boldsymbol{\vartheta}_1} = -E \left[\frac{\partial^2 \ell}{\partial \boldsymbol{\vartheta}_1^\top \partial \boldsymbol{\vartheta}_1} \right]$, $K_{\boldsymbol{\vartheta}_1 \boldsymbol{\vartheta}_2} = K_{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_1}^\top = -E \left[\frac{\partial^2 \ell}{\partial \boldsymbol{\vartheta}_1^\top \partial \boldsymbol{\vartheta}_2} \right]$ e $K_{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_2} = -E \left[\frac{\partial^2 \ell}{\partial \boldsymbol{\vartheta}_2^\top \partial \boldsymbol{\vartheta}_2} \right]$, e uma das componentes da matriz inversa de informação de Fisher relacionada ao vetor de parâmetros que deseja-se testar é dada por

$$K^{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_2} = (K_{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_2} - K_{\boldsymbol{\vartheta}_1 \boldsymbol{\vartheta}_2}^\top K_{\boldsymbol{\vartheta}_1 \boldsymbol{\vartheta}_1}^{-1} K_{\boldsymbol{\vartheta}_1 \boldsymbol{\vartheta}_2})^{-1}. \quad (2.30)$$

A hipótese para testar é:

$$H_0 : \boldsymbol{\vartheta}_2 = \boldsymbol{\vartheta}_{20} \text{ contra } H_1 : \boldsymbol{\vartheta}_2 \neq \boldsymbol{\vartheta}_{20}$$

em que $\boldsymbol{\vartheta}_{20}$ é um vetor pré-especificado.

Para formular as estatísticas de teste, primeiramente considere $\hat{\boldsymbol{\vartheta}} = (\hat{\boldsymbol{\vartheta}}_1^\top, \hat{\boldsymbol{\vartheta}}_2^\top)^\top$ e $\tilde{\boldsymbol{\vartheta}} = (\tilde{\boldsymbol{\vartheta}}_1^\top, \tilde{\boldsymbol{\vartheta}}_{20}^\top)^\top$, que são as estimativas de máxima verossimilhança de $\boldsymbol{\vartheta}$ sob o espaço paramétrico completo e sob a hipótese nula H_0 , respectivamente. A estatística gradiente para testar a hipótese H_0 é dada por

$$S_G = \tilde{U}_{\boldsymbol{\vartheta}_2}^\top (\hat{\boldsymbol{\vartheta}}_2 - \boldsymbol{\vartheta}_{20}), \quad (2.31)$$

em que $\tilde{U}_{\boldsymbol{\vartheta}_2} = U_{\boldsymbol{\vartheta}_2}(\boldsymbol{\vartheta})|_{\boldsymbol{\vartheta}=\tilde{\boldsymbol{\vartheta}}}$.

A formulação das estatísticas usuais como o teste da razão de verossimilhança, escore

e Wald são dadas, respectivamente, por

$$S_{LR} = 2 \left(\ell(\hat{\boldsymbol{\vartheta}}) - \ell(\tilde{\boldsymbol{\vartheta}}) \right), \quad (2.32)$$

$$S_R = \tilde{U}_{\boldsymbol{\vartheta}_2}^\top \tilde{K}^{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_2} \tilde{U}_{\boldsymbol{\vartheta}_2} \quad (2.33)$$

e

$$S_W = (\hat{\boldsymbol{\vartheta}}_2 - \boldsymbol{\vartheta}_{20})^\top (\hat{K}^{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_2})^{-1} (\hat{\boldsymbol{\vartheta}}_2 - \boldsymbol{\vartheta}_{20}), \quad (2.34)$$

em que $\tilde{K}^{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_2} = K^{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_2}(\boldsymbol{\vartheta})|_{\boldsymbol{\vartheta}=\tilde{\boldsymbol{\vartheta}}}$, e $(\hat{K}^{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_2})^{-1} = (K^{\boldsymbol{\vartheta}_2 \boldsymbol{\vartheta}_2}(\boldsymbol{\vartheta}))^{-1}|_{\boldsymbol{\vartheta}=\hat{\boldsymbol{\vartheta}}}$ são obtidos de (2.30).

As quatro estatísticas sob H_0 e algumas condições de regularidade, convergem em distribuição para uma distribuição Qui-quadrado com $\dim(\boldsymbol{\vartheta}_2)$ graus de liberdade.

Com as expressões das estatísticas gradiente, da razão de verossimilhança, escore e Wald em (2.31) - (2.34), temos interesse em desenvolver procedimentos sobre o parâmetro de inflação (ou deflação) de zeros no modelo de regressão ZISP, ou seja, deseja-se testar a hipótese

$$H_0 : \phi = \phi_0, \text{ contra } H_1 : \phi \neq \phi_0, \quad (2.35)$$

em que ϕ_0 é valor especificado do ϕ . A estatística gradiente para testar a hipótese nula em (2.35) é dada por

$$S_G = \left(\frac{-n}{1 + \phi_0} + \sum_{i=1}^n I_{(z_i=0)} \frac{1}{\phi_0 + a_0 A(\tilde{\theta}_i)^{-1}} \right) (\hat{\phi} - \phi_0), \quad (2.36)$$

em que $\tilde{\theta}_i = g^{-1}(\eta(\mathbf{x}_i; \tilde{\boldsymbol{\beta}}))$ é EMV de θ_i sob hipótese nula e $\hat{\phi}$ é o EMV de ϕ no espaço paramétrico completo. As estatísticas da razão de verossimilhança, escore e Wald são dadas, respectivamente, por

$$\begin{aligned} S_{LR} = & 2 \left[-n \log \left(\frac{1 + \hat{\phi}}{1 + \phi_0} \right) + \sum_{i=1}^n I_{(z_i>0)} \log \left(\frac{\hat{\phi} + a_0 A(\hat{\theta}_i)^{-1}}{\phi_0 + a_0 A(\tilde{\theta}_i)^{-1}} \right) \right. \\ & \left. + \sum_{i=1}^n z_i I_{(z_i>0)} \log \left(\frac{\hat{\theta}_i}{\tilde{\theta}_i} \right) - \sum_{i=1}^n I_{(z_i>0)} \log \left(\frac{A(\hat{\theta}_i)}{A(\tilde{\theta}_i)} \right) \right], \end{aligned} \quad (2.37)$$

em que $\hat{\theta}_i = g(\eta(\mathbf{x}_i; \hat{\boldsymbol{\beta}}))$ é EMV de θ_i sob o espaço paramétrico completo.

$$S_R = \left(\frac{-n}{1 + \phi_0} + \sum_{i=1}^n I_{(z_i=0)} \frac{1}{\phi_0 + a_0 A(\tilde{\theta}_i)^{-1}} \right) \tilde{K}^{\phi\phi}(\boldsymbol{\vartheta}) \quad (2.38)$$

e

$$S_W = (\hat{\phi} - \phi_0)^2 \hat{K}^{\phi\phi}(\boldsymbol{\vartheta})^{-1}. \quad (2.39)$$

As estatísticas dadas em (2.36) - (2.39) para testar H_0 convergem em distribuição para uma distribuição Qui-quadrado com $\dim(\phi) = 1$ graus de liberdade. Também é de interesse testar o parâmetro de potência, θ_i , com objetivo de verificar a significância das covariáveis $\mathbf{x}_i$. Sejam β_1 e β_2 subconjuntos disjuntos do vetor de parâmetros β . Deseja-se testar

$$H_0 : \beta_2 = \beta_{20}, \text{ contra } H_1 : \beta_2 \neq \beta_{20}, \quad (2.40)$$

em que β_{20} é o vetor especificado de β_2 .

Para testar (2.40), considere as partições $U_{\beta_1}(\boldsymbol{\vartheta})$ e $U_{\beta_2}(\boldsymbol{\vartheta})$ do vetor escore $U_1(\boldsymbol{\vartheta})$ (2.25). Seja $U_{\beta_2}(\boldsymbol{\vartheta})$ um vetor com r elementos, enquanto que $U_{\beta_1}(\boldsymbol{\vartheta})$ contém $p - r$ elementos. Da submatriz $K^{\beta\beta}$ definida em (2.27) escolha-se a partição $K_{\beta_2}^{\beta\beta}$ que corresponde a submatriz de dimensão $r \times r$ formada pelas r linhas e as r colunas relacionadas aos parâmetros β_2 . Desta forma apresentamos a formulação da estatística gradiente para testar a hipótese nula, dada por

$$S_G = \tilde{U}_{\beta_2}^\top (\hat{\beta}_2 - \beta_{20}). \quad (2.41)$$

As estatísticas da razão de verossimilhança, escore e Wald são dadas, respectivamente, por

$$S_{LR} = 2(\ell(\hat{\boldsymbol{\vartheta}}) - \ell(\tilde{\boldsymbol{\vartheta}})), \quad (2.42)$$

$$S_R = \tilde{U}_{\beta_2}^\top \tilde{K}_{\beta_2}^{\beta\beta}(\boldsymbol{\vartheta}) \tilde{U}_{\beta_2} \quad (2.43)$$

e

$$S_W = (\hat{\beta}_2 - \beta_{20})^\top \hat{K}_{\beta_2}^{\beta\beta}(\boldsymbol{\vartheta})^{-1} (\hat{\beta}_2 - \beta_{20}), \quad (2.44)$$

em que $\hat{\boldsymbol{\vartheta}} = (\hat{\beta}_1^\top, \hat{\beta}_2^\top, \hat{\phi})$ e $\tilde{\boldsymbol{\vartheta}} = (\tilde{\beta}_1^\top, \tilde{\beta}_2^\top, \tilde{\phi})$ são os EMVs de $\boldsymbol{\vartheta}$ sob o espaço paramétrico completo e sob H_0 , respectivamente. $\tilde{U}_{\beta_2}$, $\tilde{K}_{\beta_2}^{\beta\beta}(\boldsymbol{\vartheta})$ e $\hat{K}_{\beta_2}^{\beta\beta}(\boldsymbol{\vartheta})^{-1}$ são obtidos por meio de (2.30).

As estatísticas dadas em (2.41)-(2.44) sob H_0 e algumas condições de regularidade, convergem em distribuição para uma distribuição Qui-quadrado com $\dim(\beta_2) = r$ graus de liberdade.

2.2.1 Casos particulares

Nesta subseção apresentamos alguns casos particulares do modelo de regressão ZISP, e considere que o parâmetro de potência relacionado às covariáveis $\mathbf{x}_i$ será por meio de uma função linear que adotará diferentes funções de ligação dependendo do caso particular do modelo ZISP. De (2.12) temos que o modelo de regressão ZI-Poisson com covariáveis no parâmetro θ_i considerando a função de ligação logarítmica e uma função linear em (2.23)

dada por

$$\log(\theta_i) = \mathbf{x}_i \boldsymbol{\beta} \quad \text{ou} \quad \theta_i = \exp(\mathbf{x}_i \boldsymbol{\beta}),$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ é o vetor de parâmetros de regressão ($\boldsymbol{\beta} \in R^p$) associado ao vetor de covariáveis $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})$ com $i = 1, \dots, n$.

De (2.25), os elementos do vetor escore com respeito a $\boldsymbol{\beta}$ e ϕ correspondente ao modelo de regressão ZI-Poisson, são escritos como

$$\begin{aligned} U_1(\boldsymbol{\vartheta}) &= \frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^n \left\{ I_{(z_i=0)} \frac{-e^{-\theta_i} \theta_i}{\phi + e^{-\theta_i}} + I_{(z_i>0)} (z_i - \theta_i) \right\} \mathbf{x}_i^\top \\ U_2(\boldsymbol{\vartheta}) &= \frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \phi} = \sum_{i=1}^n \left\{ \frac{-1}{1 + \phi} + I_{(z_i=0)} \left(\frac{1}{\phi + e^{-\theta_i}} \right) \right\}. \end{aligned} \quad (2.45)$$

Por meio de (2.5), temos que a matriz de informação de Fisher dos parâmetros $\boldsymbol{\beta}$ e ϕ no modelo de regressão ZI-Poisson tem como componentes

$$\begin{aligned} K_{\boldsymbol{\beta}\boldsymbol{\beta}} &= -E \left(\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right) = \sum_{i=1}^n \left\{ \frac{e^{-\theta_i} (1 - \phi \theta_i) + \phi}{(1 + \phi)(\phi + e^{-\theta_i})} \right\} \theta_i \mathbf{x}_i^\top \mathbf{x}_i, \\ K_{\boldsymbol{\beta}\phi} &= -E \left(\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\beta} \partial \phi} \right) = - \sum_{i=1}^n \left\{ \frac{e^{-\theta_i}}{(1 + \phi)(\phi + e^{-\theta_i})} \right\} \theta_i \mathbf{x}_i^\top, \\ K_{\phi\phi} &= -E \left(\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \phi^2} \right) = \sum_{i=1}^n \left\{ \frac{1 - e^{-\theta_i}}{(1 + \phi)^2 (\phi + e^{-\theta_i})} \right\}. \end{aligned} \quad (2.46)$$

Na forma matricial

$$K_{\boldsymbol{\beta}\boldsymbol{\beta}} = \sum_{i=1}^n \mathbf{d}_i \mathbf{x}_i^\top \mathbf{x}_i = \mathbf{X}^\top \text{diag}(\mathbf{d}) \mathbf{X} \quad \text{e} \quad K_{\boldsymbol{\beta}\phi} = - \sum_{i=1}^n \mathbf{w}_i \mathbf{x}_i^\top = -\mathbf{X}^\top \mathbf{W},$$

em que $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^\top$ e $\text{diag}(\mathbf{d})$ é uma matriz diagonal $n \times n$ com (i, i) -ésima entrada dada por

$$\mathbf{d}_i(\boldsymbol{\vartheta}) = \frac{(e^{-\theta_i} (1 - \phi \theta_i) + \phi) \theta_i}{(1 + \phi)(\phi + e^{-\theta_i})},$$

e $\mathbf{W} = \text{diag}(\mathbf{W}) \mathbf{X} \mathbf{e}_p$ em que $\mathbf{e}_p$ é um vetor que tem como primeiro valor igual a 1 e têm os outros elementos iguais a 0 e $\text{diag}(\mathbf{W})$ é uma matriz diagonal $n \times n$ com (i, i) -ésima componente dada por

$$\mathbf{w}_i(\boldsymbol{\vartheta}) = \frac{\theta_i e^{-\theta_i}}{(1 + \phi)(\phi + e^{-\theta_i})}.$$

De (2.6), tem-se

$$K^{\phi\phi}(\boldsymbol{\beta}, \phi) = \left(\sum_{i=1}^n \left\{ \frac{1 - e^{-\theta_i}}{(1 + \phi)^2(\phi + e^{-\theta_i})} \right\} - \mathbf{W}^\top \mathbf{X} [\mathbf{X}^\top \text{diag}(\mathbf{d}) \mathbf{X}]^{-1} \mathbf{X}^\top \mathbf{W} \right)^{-1}. \quad (2.47)$$

As estatísticas gradiente, da razão de verossimilhança, escore e Wald dadas em (2.36) - (2.39) para testar H_0 em (2.35), correspondentes ao modelo de regressão ZI-Poisson são dadas, respectivamente, por

$$S_G = (\hat{\phi} - \phi_0) \sum_{i=1}^n \left\{ \frac{-1}{1 + \phi_0} + I_{(z_i=0)} \left(\frac{1}{\phi_0 + e^{-\tilde{\theta}_i}} \right) \right\},$$

$$S_{LR} = 2 \sum_{i=1}^n \left\{ -\log \left(\frac{1 + \hat{\phi}}{1 + \phi_0} \right) + I_{(z_i=0)} \log \left(\frac{\hat{\phi} + e^{-\tilde{\theta}_i}}{\phi_0 + e^{-\tilde{\theta}_i}} \right) + I_{(z_i>0)} [z_i \mathbf{x}_i (\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}) - (\hat{\theta}_i - \tilde{\theta}_i)] \right\},$$

$$S_R = \left[\sum_{i=1}^n \left\{ \frac{-1}{1 + \phi_0} + I_{(z_i=0)} \left(\frac{1}{\phi_0 + e^{-\tilde{\theta}_i}} \right) \right\} \right]^2 \tilde{K}^{\phi\phi}(\boldsymbol{\beta}, \phi)$$

e

$$S_W = (\hat{\phi} - \phi_0)^2 \hat{K}^{\phi\phi}(\boldsymbol{\beta}, \phi)^{-1},$$

sendo $\tilde{\boldsymbol{\beta}}$ é o EMV de $\boldsymbol{\beta}$ sob a hipóteses nula e, $\hat{\boldsymbol{\beta}}$ e $\hat{\phi}$ são os EMVs no espaço paramétrico completo e, $\tilde{K}^{22}(\boldsymbol{\beta}, \phi)$ e $\hat{K}^{22}(\boldsymbol{\beta}, \phi)^{-1}$ são obtidos de (2.47).

Agora, no caso do modelo de regressão ZI-Geométrica com covariáveis no parâmetro θ_i em (2.17), considerando a função de ligação logística associada à função linear em (2.23) dada por

$$\log \left(\frac{\theta_i}{1 - \theta_i} \right) = \mathbf{x}_i \boldsymbol{\beta} \quad \text{ou} \quad \theta_i = \frac{\exp(\mathbf{x}_i \boldsymbol{\beta})}{1 + \exp(\mathbf{x}_i \boldsymbol{\beta})},$$

os elementos do vetor escore definidos em (2.25), correspondentes ao modelo são dados por

$$U_1(\boldsymbol{\vartheta}) = \sum_{i=1}^n \left\{ I_{(z_i=0)} \frac{1}{\phi + (1 - \theta_i)} + I_{(z_i>0)} \left[\frac{z_i}{\theta_i} - \frac{1}{(1 - \theta_i)} \right] \right\} \frac{\theta_i}{1 + \exp(\mathbf{x}_i \boldsymbol{\beta})} \mathbf{x}_i^\top \quad (2.48)$$

e

$$U_2(\boldsymbol{\vartheta}) = \frac{-n}{1 + \phi} + \sum_{i=1}^n \left\{ I_{(z_i=0)} \frac{1}{\phi + (1 - \theta_i)} \right\}. \quad (2.49)$$

A matriz de informação de Fisher dada em (2.5) para os parâmetros $\boldsymbol{\beta}$ e ϕ no modelo de regressão ZI-Geométrica tem como componentes

$$K_{\boldsymbol{\beta}\boldsymbol{\beta}} = \frac{1}{1 + \phi} \sum_{i=1}^n \left[\frac{-2}{(1 - \theta_i)} - \frac{(1 + 2(1 - \theta_i)^{-1})}{\phi + (1 - \theta_i)} - \frac{1 - 3\theta_i^2(1 - \theta_i)^{-1}}{\theta_i(1 - \theta_i)} \right] \frac{\theta_i^2}{(1 + e^{\mathbf{x}_i \boldsymbol{\beta}})^2} \mathbf{x}_i^\top \mathbf{x}_i, \quad (2.50)$$

$$K_{\beta\phi} = \frac{1}{1+\phi} \sum_{i=1}^n \left[\frac{\theta_i}{(\phi + (1-\theta_i))(1 + \exp \mathbf{x}_i \boldsymbol{\beta})} \right] \mathbf{x}_i^\top \quad (2.51)$$

e

$$K_{\phi\phi} = \sum_{i=1}^n \frac{\phi + (1-\theta_i)}{(1+\phi)^3} + \frac{1}{1+\phi} \sum_{i=1}^n \frac{\theta_i}{(\phi + (1-\theta_i))^2}, \quad (2.52)$$

as quais expressas na forma matricial podem ser escritas como

$$K_{\beta\beta} = \sum_{i=1}^n \mathbf{c}_i \mathbf{x}_i^\top \mathbf{x}_i = \mathbf{X}^\top \text{diag}(\mathbf{c}) \mathbf{X} \quad \text{e} \quad K_{\beta\phi} = \sum_{i=1}^n \mathbf{h}_i \mathbf{x}_i^\top = \mathbf{X}^\top \mathbf{H},$$

em que $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^\top$ e $\text{diag}(\mathbf{c})$ é uma matriz diagonal $n \times n$ com (i, i) -ésima entrada dada por

$$\mathbf{c}_i(\boldsymbol{\vartheta}) = \frac{1}{1+\phi} \left(\frac{-2}{(1-\theta_i)} - \frac{1+2(1-\theta_i)^{-1}}{\phi + (1-\theta_i)} - \frac{1-3\theta_i^2(1-\theta_i)^{-1}}{\theta_i(1-\theta_i)} \right) \frac{\theta_i^2}{(1+e^{\mathbf{x}_i \boldsymbol{\beta}})^2},$$

e $\mathbf{H} = \text{diag}(\mathbf{H}) \mathbf{X} \mathbf{e}_p$ em que $\mathbf{e}_p$ é um vetor que tem como primeiro valor igual a 1 e têm os outros elementos iguais a 0 e $\text{diag}(\mathbf{H})$ é uma matriz diagonal $n \times n$ com (i, i) -ésima componente dada por

$$\mathbf{h}_i(\boldsymbol{\vartheta}) = \frac{1}{1+\phi} \left(\frac{\theta_i}{(\phi + (1-\theta_i))(1+e^{\mathbf{x}_i \boldsymbol{\beta}})} \right).$$

De (2.6) tem-se

$$K^{\phi\phi}(\boldsymbol{\vartheta})^{-1} = \left(\sum_{i=1}^n \frac{\phi + (1-\theta_i)}{(1+\phi)^3} + \frac{1}{1+\phi} \sum_{i=1}^n \frac{\theta_i}{(\phi + (1-\theta_i))^2} \right) - \mathbf{H}^\top \mathbf{X} [\mathbf{X}^\top \text{diag}(\mathbf{c}) \mathbf{X}]^{-1} \mathbf{X}^\top \mathbf{H}. \quad (2.53)$$

As estatísticas gradiente, da razão de verossimilhança, escore e Wald dadas em (2.36) - (2.39) para testar H_0 em (2.35), correspondentes ao modelo de regressão ZI-Geométrica, são dadas, respectivamente, por

$$\begin{aligned} S_G &= (\hat{\phi} - \phi_0) \left(\frac{-n}{1+\phi_0} + \sum_{i=1}^n \left\{ I_{(z_i=0)} \frac{1}{\phi_0 + (1-\tilde{\theta}_i)} \right\} \right), \\ S_{LR} &= 2 \left(-n \log \left(\frac{1+\hat{\phi}}{1+\phi_0} \right) + \sum_{i=1}^n I_{(z_i=0)} \log \left(\frac{\hat{\phi} + (1-\hat{\theta}_i)}{\phi_0 + (1-\tilde{\theta}_i)} \right) \right. \\ &\quad \left. + \sum_{i=1}^n z_i I_{(z_i>0)} \log \left(\frac{\hat{\theta}_i}{\tilde{\theta}_i} \right) + \sum_{i=1}^n I_{(z_i>0)} \log \left(\frac{1-\hat{\theta}_i}{1-\tilde{\theta}_i} \right) \right), \end{aligned}$$

$$S_R = \left(\frac{-n}{1 + \phi_0} + \sum_{i=1}^n I_{(z_i=0)} \frac{1}{\phi_0 + (1 - \tilde{\theta}_i)} \right) \tilde{K}^{\phi\phi}(\boldsymbol{\beta}, \phi)$$

e

$$S_W = (\hat{\phi} - \phi_0)^2 \hat{K}^{\phi\phi}(\boldsymbol{\beta}, \phi)^{-1},$$

em que $\tilde{K}^{22}(\boldsymbol{\beta}, \phi)$ e $\hat{K}^{\phi\phi}(\boldsymbol{\beta}, \phi)^{-1}$ são obtidas de (2.53).

Alternativamente, os diferentes modelos podem ser comparados, penalizando os ajustes usando o critério de informação de Akaike (1973) dado por $AIC = -2\ell(\boldsymbol{\vartheta}) + \dim(\boldsymbol{\vartheta})$ e o critério de Schwarz (1978) definido por $SBC = -2\ell(\boldsymbol{\vartheta}) + \dim(\boldsymbol{\vartheta}) \log(n)$, em que $\dim(\boldsymbol{\vartheta})$ é o número de parâmetros. Portanto, para os modelos em avaliação com o mesmo conjunto de dados, será selecionado o melhor modelo, aquele que possui o menor valor de critério de seleção.

2.3 Estudo de simulação

Nesta seção apresentamos um estudo de simulação com e sem covariáveis com a finalidade de avaliar a performance da distribuição da estatística gradiente (S_G) com as distribuições das estatísticas da razão de verossimilhança (S_{LR}), score (S_R) e Wald (S_W), nos modelos ZIP e ZIG, em testar as hipóteses: $H_0 : \phi = 0$ contra $H_1 : \phi \neq 0$, para amostras pequenas, moderadas e grandes.

2.3.1 Simulações sem covariáveis no modelo ZI-Poisson

Neste caso, sem covariáveis, o estudo de simulação é baseado na estimação da taxa de rejeição das distribuições empíricas e do poder do teste das quatro estatísticas assumindo a hipótese nula.

Consideramos 10000 réplicas de Monte Carlo utilizadas na comparação do tamanho do teste e os níveis nominais dos testes foram $\alpha = 1\%, 5\%$ e 10% . Esta configuração foi realizada em cinco diferentes tamanhos de amostras $n = 30, 50, 100, 200, 400$ e 600 . Em cada caso, consideramos uma variação de $\theta = 0.5, 1, 2$. Os valores de θ foram escolhidos para a geração de zeros, pois quanto menor for o valor de θ , maior será o número de zeros gerados. A implementação das simulações foi desenvolvida no *software* R (R Development Core Team, 2009) e os resultados da taxa de rejeição estão apresentados na Tabela 2.1.

Tabela 2.1: Taxas de rejeição da hipótese $H_0 : \phi = 0$, no modelo ZI-Poisson, para diferentes valores de θ .

n	$(\alpha\%)$	$\theta = 0,5$				$\theta = 1,0$				$\theta = 2,0$			
		S_G	S_R	S_{LR}	S_W	S_G	S_R	S_{LR}	S_W	S_G	S_R	S_{LR}	S_W
30	1	1,0	0,7	0,5	2,9	1,0	0,8	1,0	1,9	0,8	0,6	0,6	1,0
	5	3,2	3,4	4,1	5,9	5,0	4,6	4,7	5,3	3,9	3,8	4,5	5,6
	10	7,2	9,0	8,7	9,7	9,7	10,1	10,3	10,4	8,6	8,4	9,5	9,8
50	1	0,9	0,8	0,9	2,2	0,8	0,9	1,0	1,7	0,9	0,8	1,0	2,0
	5	4,7	4,4	4,9	5,9	4,9	4,6	5,3	5,5	4,7	4,8	5,6	6,2
	10	9,7	9,7	10,3	10,5	10,1	10,1	10,6	10,3	10,2	10,1	10,8	10,4
100	1	1,1	0,9	1,1	2,0	1,0	0,9	1,0	1,6	1,0	0,9	1,0	1,6
	5	4,8	5,0	5,0	6,5	5,0	5,0	5,0	5,7	4,9	4,9	5,1	5,7
	10	10,0	10,2	10,0	12,0	10,1	10,3	10,2	11,2	10,0	10,2	10,2	10,8
200	1	1,0	1,0	1,1	1,5	1,1	1,1	1,1	1,4	1,0	1,1	1,1	1,4
	5	4,8	5,0	5,0	5,4	4,9	5,0	5,1	5,3	5,0	5,1	5,1	5,3
	10	9,9	10	10,0	10,6	10,0	10,0	10,0	10,3	10	10,0	10,1	10,3
400	1	1,0	1,0	1,0	1,2	1,0	1,0	1,0	1,1	1,0	1,0	1,0	1,0
	5	4,9	4,9	5,1	5,3	4,9	4,9	5,0	5,2	5,0	5,0	5,0	5,1
	10	9,9	10,0	10,1	10,4	9,8	9,9	10,0	10,0	9,8	9,8	10,0	10,0
600	1	0,9	0,9	0,9	1,1	1,2	1,2	1,2	1,0	1,0	1,0	1,0	1,1
	5	4,9	4,9	5,0	5,0	4,9	5,1	5,0	5,0	4,9	4,8	4,9	4,9
	10	9,8	9,8	9,8	9,9	10,2	10,2	10,2	10,0	9,9	9,9	9,8	9,8

Da Tabela 2.1 observa-se que as taxas de rejeição nula do teste gradiente apresenta uma boa aproximação aos valores nominais para tamanhos amostrais pequenos ($n = 30, 50, 100$) comparados com os resultados dos testes usuais. À medida que aumenta o tamanho das amostras o teste gradiente é equivalente aos testes de escore, Wald e da razão de verossimilhança para testes de hipóteses em relação ao parâmetro de inflação do modelo ZI-Poisson. Note que as taxas do teste de Wald são superestimadas em alguns casos.

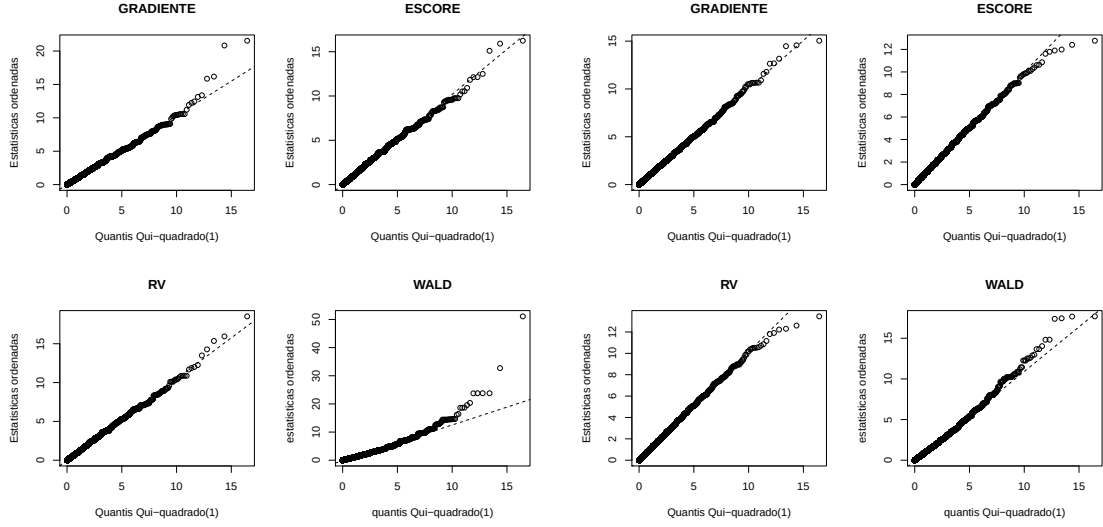
Tabela 2.2: *Percentis estimados com amostra de tamanho $n = 50$ sob $H_0 : \phi = 0$.*

θ		$P_{,7} = 1,07$	$P_{,8} = 1,64$	$P_{,9} = 2,71$	$P_{,95} = 3,84$	$P_{,99} = 6,63$
0,5	S_G	1,07	1,66	2,54	3,66	6,93
	S_R	1,07	1,63	2,61	3,60	6,21
	S_{LR}	1,09	1,71	2,75	3,73	6,32
	S_W	1,08	1,76	3,08	5,16	11,60
1,0	S_G	1,11	1,67	2,77	3,91	6,35
	S_R	1,10	1,71	2,73	3,77	6,53
	S_{LR}	1,10	1,70	2,72	3,80	6,72
	S_W	1,10	1,66	2,78	3,78	7,63
2,0	S_G	1,11	1,71	2,76	3,78	6,60
	S_R	1,12	1,70	2,74	3,82	6,46
	S_{LR}	1,15	1,73	2,88	3,97	6,87
	S_W	1,12	1,71	2,87	4,36	8,33

Tabela 2.3: *Percentis estimados com amostra de tamanho $n = 200$ sob $H_0 : \phi = 0$.*

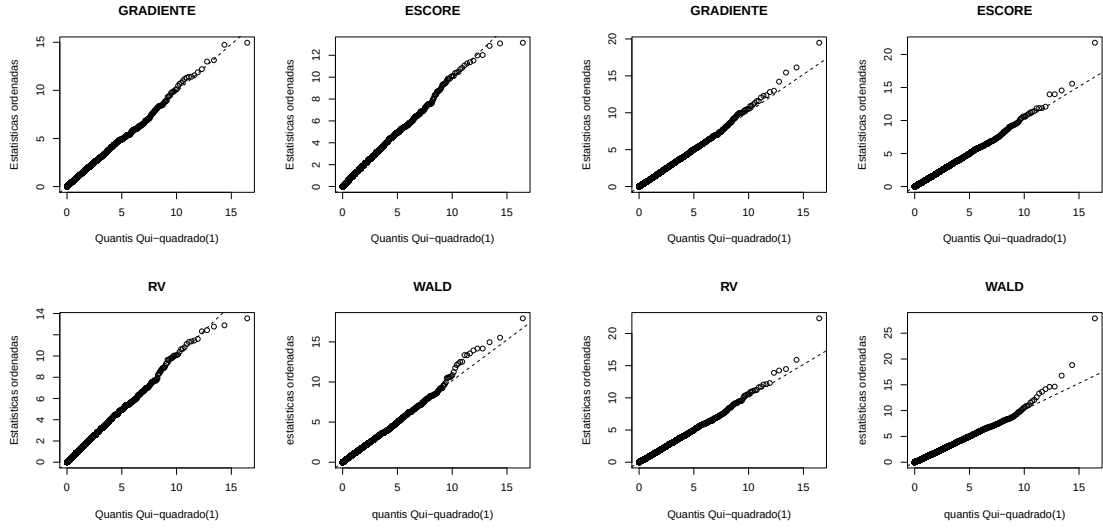
θ		$P_{,7} = 1,07$	$P_{,8} = 1,64$	$P_{,9} = 2,71$	$P_{,95} = 3,84$	$P_{,99} = 6,63$
0,5	S_G	1,05	1,64	2,68	3,80	6,55
	S_R	1,08	1,66	2,73	3,83	6,52
	S_{LR}	1,09	1,65	2,72	3,84	6,77
	S_W	1,08	1,66	2,83	3,96	7,68
1,0	S_G	1,05	1,62	2,75	3,83	6,89
	S_R	1,06	1,62	2,71	3,86	6,84
	S_{LR}	1,06	1,62	2,72	3,88	6,85
	S_W	1,04	1,61	2,71	3,92	6,96
2,0	S_G	1,07	1,65	2,69	3,89	6,64
	S_R	1,08	1,66	2,71	3,88	6,69
	S_{LR}	1,08	1,66	2,73	3,89	6,76
	S_W	1,08	1,65	2,75	3,90	7,40

Nas Tabelas 2.2-2.3 são apresentados os quantis da distribuição empírica da estatística gradiente, razão de verossimilhança, escore e Wald, conjuntamente com os quantis da distribuição Qui-quadrado com *um* grau de liberdade, em que observamos que os quantis são bem aproximados aos quantis teóricos. Esses resultados podem ser observados na Figura 2.1. Mas, em alguns casos os quantis empíricos diferem do quantil teórico para a estatística de Wald.



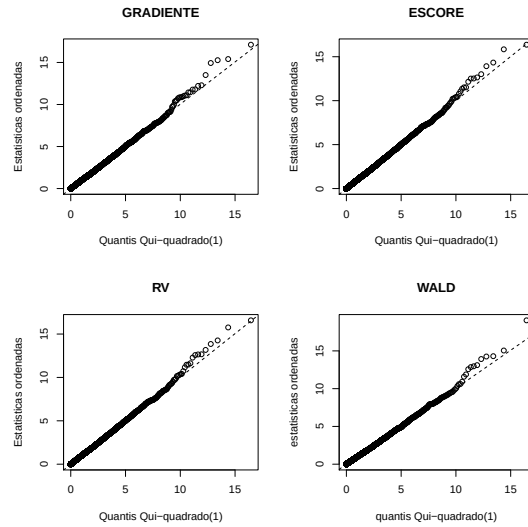
(a) $n = 30$

(b) $n = 50$



(c) $n = 100$

(d) $n = 200$



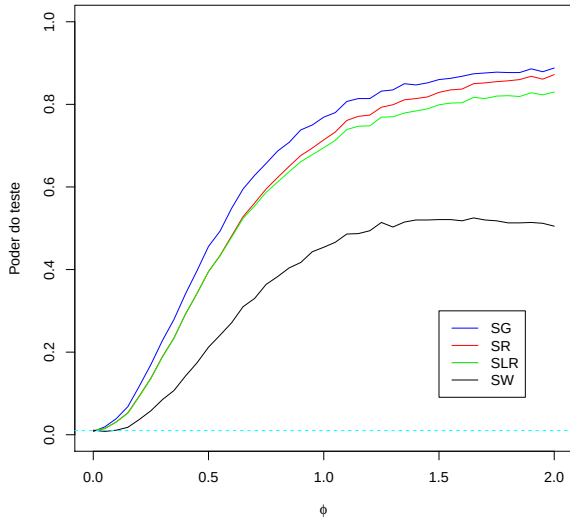
(e) $n = 400$

Figura 2.1: Quantis empíricas das distribuições S_G, S_R, S_{LR} e S_W quando $\theta = 1$.

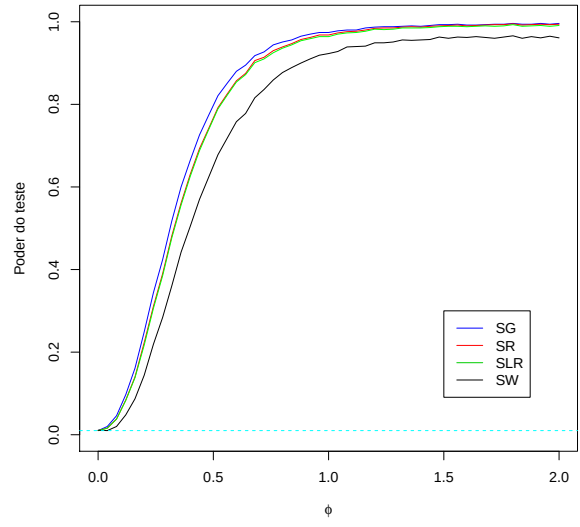
O poder do teste das quatro estatísticas é apresentado na Tabela 2.4 para $n = 100$ e 200, em que é considerado uma pequena variação de ϕ para observar a variação do poder da estatística gradiente. Os resultados, indicam que a estatística gradiente é competitivo às estatísticas de escore, da razão de verossimilhança e Wald, pois o poder do teste é similar.

Tabela 2.4: *Poder do teste de ϕ quando $\theta = 2$.*

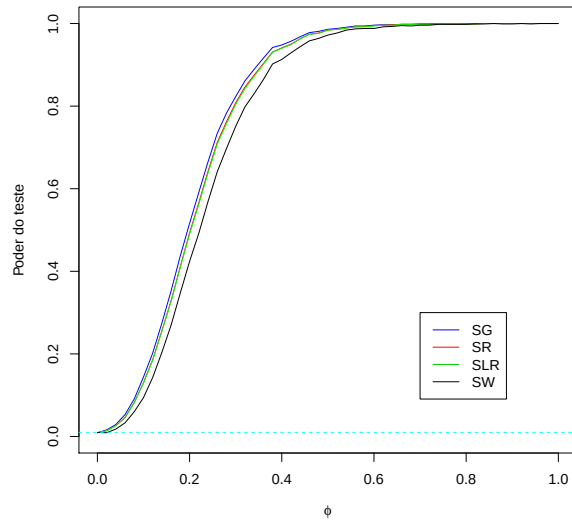
n	ϕ	$\alpha = 0,01$				$\alpha = 0,05$				$\alpha = 0,10$			
		S_G	S_R	S_{LR}	S_W	S_G	S_R	S_{LR}	S_W	S_G	S_R	S_{LR}	S_W
100	0,00	0,01	0,01	0,01	0,02	0,05	0,05	0,06	0,06	0,10	0,10	0,10	0,11
	0,04	0,04	0,04	0,03	0,01	0,13	0,12	0,11	0,07	0,20	0,20	0,18	0,14
	0,08	0,15	0,13	0,12	0,04	0,31	0,30	0,29	0,19	0,42	0,40	0,39	0,32
	0,12	0,30	0,28	0,26	0,11	0,52	0,49	0,48	0,37	0,63	0,61	0,60	0,53
	0,16	0,49	0,46	0,43	0,24	0,69	0,67	0,66	0,56	0,78	0,77	0,76	0,70
	0,20	0,65	0,62	0,60	0,38	0,83	0,81	0,81	0,72	0,89	0,88	0,88	0,84
	0,24	0,78	0,75	0,73	0,53	0,91	0,90	0,89	0,83	0,95	0,94	0,94	0,92
	0,28	0,87	0,85	0,84	0,68	0,95	0,95	0,95	0,91	0,97	0,97	0,97	0,96
	0,32	0,92	0,90	0,90	0,78	0,97	0,97	0,97	0,95	0,99	0,98	0,98	0,98
	0,36	0,96	0,95	0,94	0,85	0,99	0,99	0,99	0,97	0,99	0,99	0,99	0,99
	0,40	0,97	0,97	0,96	0,90	0,99	0,99	0,99	0,98	1,00	1,00	1,00	0,99
	0,44	0,98	0,98	0,98	0,94	0,99	0,99	0,99	0,99	1,00	1,00	1,00	1,00
	0,48	0,99	0,99	0,99	0,96	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
	0,52	1,00	0,99	0,99	0,98	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
	0,56	1,00	0,99	0,99	0,98	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
	0,60	1,00	1,00	1,00	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
	0,64	1,00	1,00	1,00	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
	0,68	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
200	0,00	0,01	0,01	0,01	0,01	0,05	0,05	0,05	0,06	0,10	0,10	0,10	0,10
	0,04	0,08	0,07	0,06	0,03	0,20	0,19	0,18	0,13	0,29	0,29	0,28	0,23
	0,08	0,32	0,29	0,28	0,17	0,53	0,52	0,50	0,43	0,64	0,63	0,62	0,57
	0,12	0,61	0,58	0,56	0,43	0,80	0,79	0,78	0,72	0,87	0,86	0,86	0,83
	0,16	0,83	0,82	0,80	0,70	0,94	0,93	0,93	0,90	0,97	0,96	0,96	0,95
	0,20	0,94	0,93	0,92	0,87	0,98	0,98	0,98	0,97	0,99	0,99	0,99	0,99
	0,24	0,98	0,98	0,98	0,95	1,00	1,00	0,99	0,99	1,00	1,00	1,00	1,00
	0,28	0,99	0,99	0,99	0,98	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
	0,32	1,00	1,00	1,00	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
	0,36	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00



(a) $n = 100$

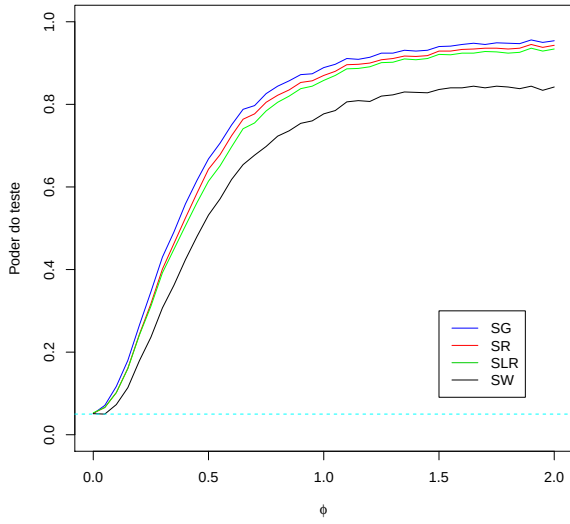


(b) $n = 200$

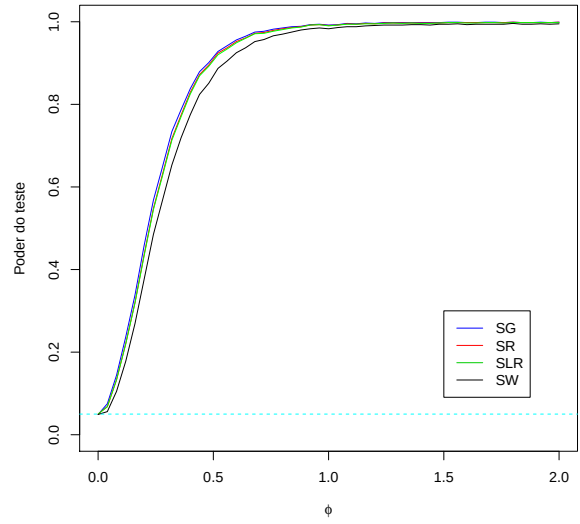


(c) $n = 400$

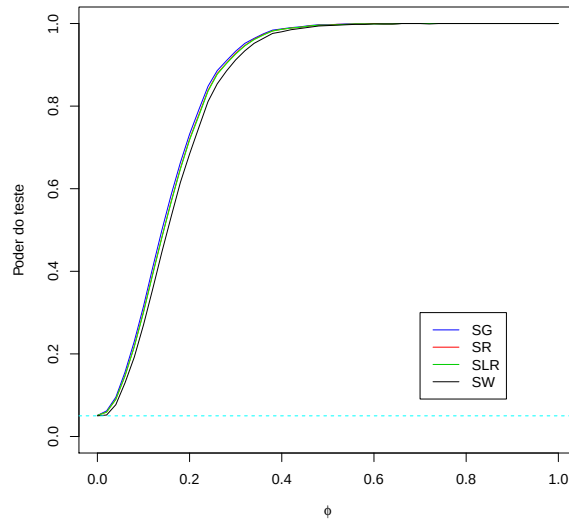
Figura 2.2: Poder do teste S_G, S_R, S_{LR} e S_W para diferentes valores de ϕ e tamanhos de amostra $n = 100, 200$ e 400 quando $\theta = 1$, ao nível nominal de 1%.



(a) $n = 100$

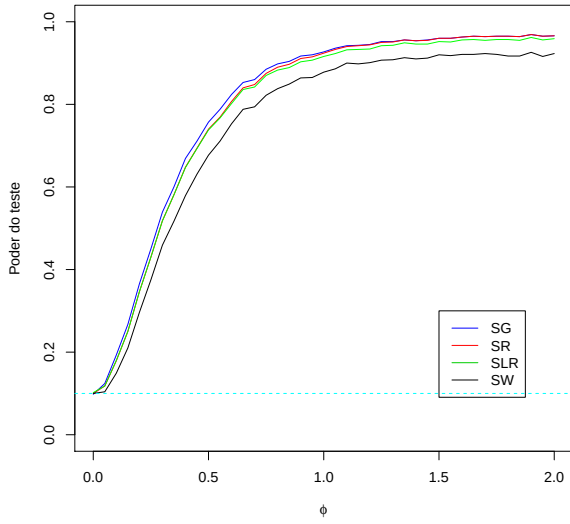


(b) $n = 200$

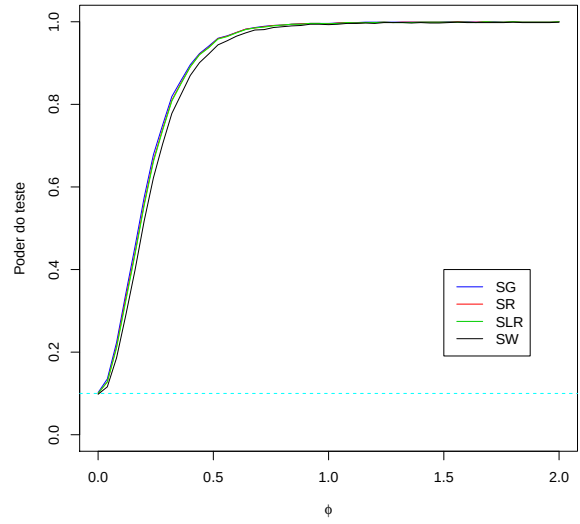


(c) $n = 400$

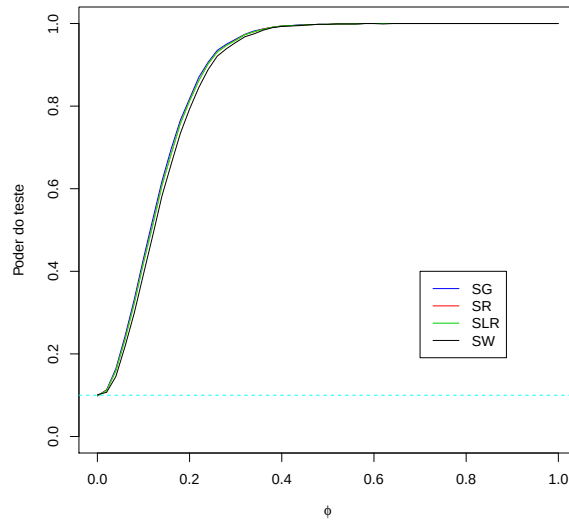
Figura 2.3: Poder do teste S_G, S_R, S_{LR} e S_W para diferentes valores de ϕ e tamanhos de amostra $n = 100, 200$ e 400 quando $\theta = 1$, ao nível nominal de 5%.



(a) $n = 100$



(b) $n = 200$



(c) $n = 400$

Figura 2.4: Poder do teste S_G, S_R, S_{LR} e S_W para diferentes valores de ϕ e tamanhos de amostra $n = 100, 200$ e 400 quando $\theta = 1$, ao nível nominal de 10%.

As Figuras 2.2 - 2.4 mostram o poder do teste para as estatística gradiente, escore, da razão de verossimilhança e Wald, em que observa-se que as quatro estatísticas têm o mesmo poder, para amostras grandes. Mas, a estatística Wald tem menos poder para amostras de tamanho $n = 100$.

2.3.2 Simulações sem covariáveis no modelo ZI-Geométrica

O estudo de simulação é baseado na estimação das taxas de rejeição das distribuições empíricas e do poder do teste das quatro estatísticas assumindo a hipótese nula para o modelo ZI-Geométrica.

Foram consideradas 10000 réplicas de Monte Carlo utilizadas na comparação do tamanho do teste e os níveis nominais dos testes foram $\alpha = 1\%, 5\%$ e 10% . Esta configuração foi realizada em cinco diferentes tamanhos de amostras $n = 30, 50, 100, 200, 400$ e 600 . Em cada caso consideramos uma variação de $\theta = 0,3; 0,5$ e $0,7$. Os resultados das taxas de rejeição estão apresentados na Tabela 2.5 e observa-se que os valores estimados do teste gradiente apresentam melhor aproximação comparado com as taxas dos testes de escore, Wald e razão de verossimilhança. Os valores das taxas de rejeição são aproximados ao valor nominal do teste na medida que aumenta o tamanho das amostras. O teste gradiente é equivalente aos testes de escore, Wald e razão de verossimilhança, para os testes de hipóteses em relação ao parâmetro de inflação do modelo ZI-Geométrica. Neste caso também acontece que, quanto menor for o valor de θ , maior será número de zeros gerados.

Tabela 2.5: Taxas de rejeição da hipótese para testar $H_o : \phi = 0$, no modelo ZIG, para diferentes valores de θ .

n	$(\alpha\%)$	$\theta = 0,3$				$\theta = 0,5$				$\theta = 0,7$			
		S_G	S_{LR}	S_R	S_W	S_G	S_{LR}	S_R	S_W	S_G	S_{LR}	S_R	S_W
30	1	1,50	0,62	5,92	4,64	0,90	0,94	5,05	3,05	0,94	0,87	4,85	2,68
	5	4,47	4,47	9,73	9,12	4,49	5,20	10,11	7,10	4,36	5,29	9,16	6,63
	10	7,89	9,71	12,28	11,40	9,38	11,04	13,85	11,40	9,65	10,19	13,50	10,74
50	1	1,18	0,98	4,96	4,54	0,94	0,94	2,92	1,64	1,06	1,06	3,50	2,20
	5	4,76	4,86	9,20	8,16	4,92	4,92	7,36	5,96	4,98	5,66	7,84	6,20
	10	8,66	10,24	13,84	11,98	9,70	9,80	12,24	10,06	9,98	10,42	12,68	10,78
100	1	1,18	1,10	3,36	2,70	0,90	0,80	1,74	1,24	1,02	1,20	1,98	1,60
	5	5,04	5,02	7,92	6,88	4,56	4,98	6,46	5,54	4,80	4,94	6,02	5,20
	10	9,62	10,30	12,40	11,78	9,62	9,68	11,56	10,68	9,60	9,60	10,76	9,84
200	1	1,01	0,96	1,98	1,71	0,97	0,90	1,53	1,27	1,20	1,20	1,81	1,59
	5	4,75	5,09	5,99	5,69	4,93	5,08	5,52	5,15	5,29	5,27	5,66	5,34
	10	9,80	10,27	11,02	10,82	9,49	9,64	10,12	9,84	9,77	9,79	10,46	10,08
400	1	0,94	0,93	1,46	1,36	1,00	0,89	1,06	0,95	0,91	0,89	1,10	0,94
	5	4,97	4,92	5,31	5,15	4,72	4,71	4,85	4,73	4,65	4,64	4,95	4,69
	10	10,03	9,98	10,48	10,25	9,42	9,55	9,68	9,53	9,79	9,71	10,05	9,87
600	1	1,04	1,03	1,38	1,28	1,08	0,94	1,06	0,93	1,26	1,22	1,34	1,24
	5	4,94	5,02	5,29	5,14	4,92	5,00	5,35	5,16	5,15	5,27	5,42	5,27
	10	9,76	9,87	10,15	10,05	10,24	10,19	10,24	10,11	9,95	9,96	10,30	10,11

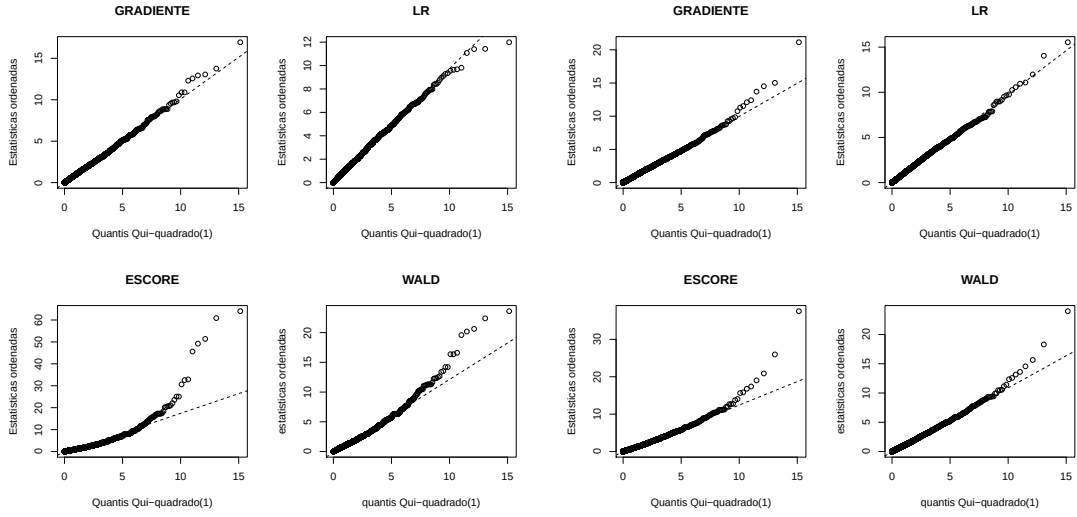
Tabela 2.6: *Percentis estimados com amostra de tamanho $n = 100$ sob $H_0 : \phi = 0$.*

θ		$P_{,7} = 1,07$	$P_{,8} = 1,64$	$P_{,9} = 2,71$	$P_{,95} = 3,84$	$P_{,99} = 6,63$
0,3	S_G	1,10	1,66	2,67	3,86	6,98
	S_{LR}	1,06	1,66	2,87	3,85	6,76
	S_R	1,17	1,78	3,23	5,11	14,02
	S_W	1,14	1,73	3,01	4,73	11,72
0,5	S_G	1,07	1,64	2,64	3,67	6,40
	S_{LR}	1,08	1,64	2,69	3,84	6,32
	S_R	1,11	1,71	2,97	4,36	8,23
	S_W	1,08	1,66	2,82	4,01	7,20
0,7	S_G	1,09	1,62	2,65	3,77	6,68
	S_{LR}	1,11	1,67	2,66	3,83	6,92
	S_R	1,13	1,68	2,83	4,25	9,33
	S_W	1,10	1,62	2,68	3,90	7,98

Tabela 2.7: *Percentis estimados com amostra de tamanho $n = 400$ sob $H_0 : \phi = 0$.*

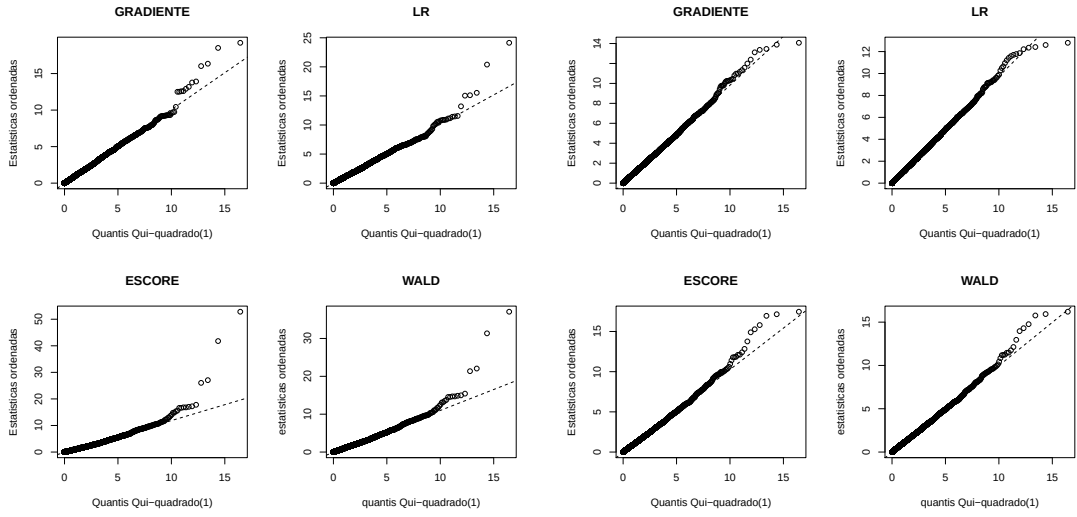
θ		$P_{,7} = 1,07$	$P_{,8} = 1,64$	$P_{,9} = 2,71$	$P_{,95} = 3,84$	$P_{,99} = 6,63$
0,3	S_G	1,08	1,68	2,71	3,82	6,52
	S_{LR}	1,10	1,66	2,70	3,81	6,58
	S_R	1,08	1,68	2,81	4,01	7,41
	S_W	1,07	1,67	2,77	3,93	7,19
0,5	S_G	1,05	1,60	2,63	3,75	6,60
	S_{LR}	1,05	1,61	2,64	3,73	6,42
	S_R	1,05	1,60	2,64	3,77	6,79
	S_W	1,05	1,59	2,60	3,71	6,58
0,7	S_G	1,03	1,62	2,68	3,72	6,53
	S_{LR}	1,03	1,63	2,66	3,72	6,41
	S_R	1,03	1,63	2,71	3,80	6,76
	S_W	1,03	1,62	2,68	3,73	6,55

Nas Tabelas 2.6-2.7 são apresentados os quantis da distribuição empírica da estatística gradiente, da razão de verossimilhança, escore e Wald, conjuntamente com os quantis da distribuição Qui-quadrado com *um* grau de liberdade, em que observamos que os quantis são bem aproximados aos valores quantis teóricos. Esses resultados podem ser observados nas Figura 2.5. Note que a estatística gradiente e da razão de verossimilhança são melhores comportadas em relação às outras estatísticas mesmo par uma amostra pequena.



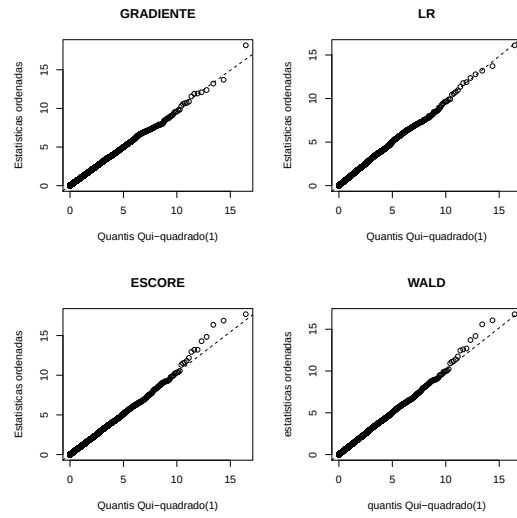
(a) $n = 50$

(b) $n = 100$



(c) $n = 200$

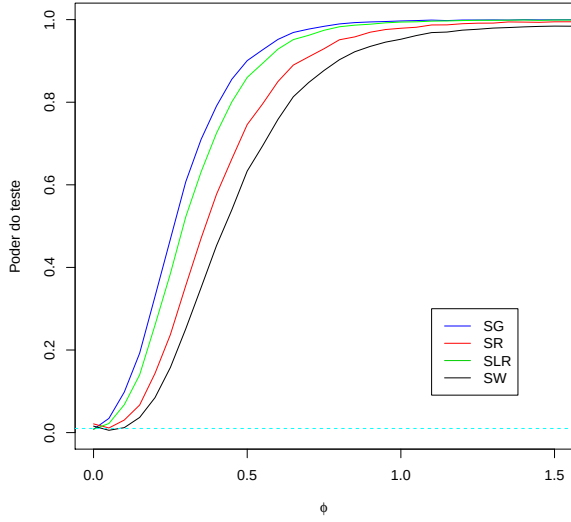
(d) $n = 400$



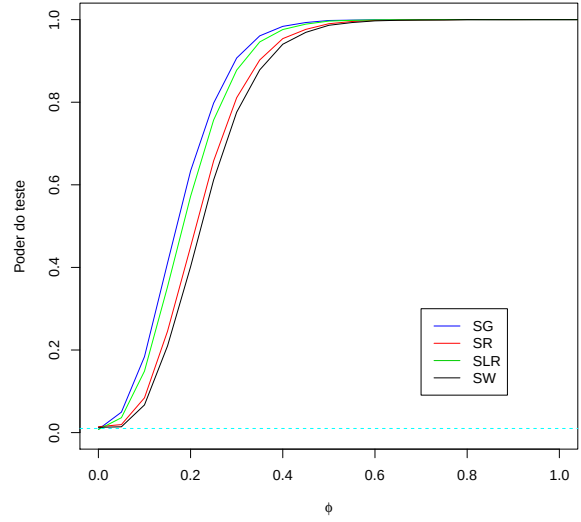
(e) $n = 600$

Figura 2.5: Quantis empíricos das distribuições S_G, S_{LR}, S_R e S_W quando $\theta = 0,5$.

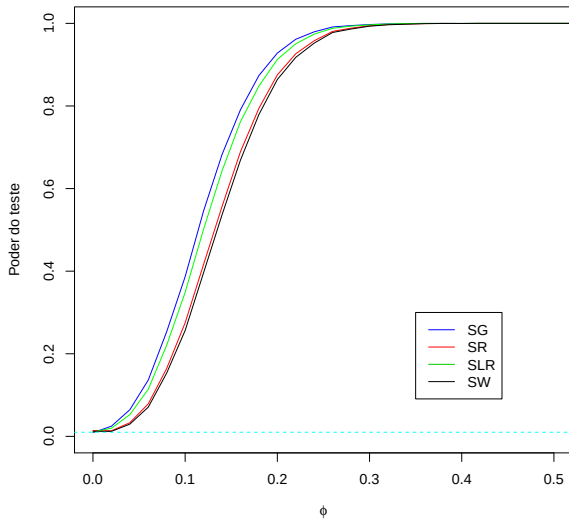
O poder do teste das quatro estatísticas, são apresentadas graficamente nas Figuras 2.6-2.8 em que tem-se considerado uma pequena variação de ϕ para observar a variação do poder da estatística gradiente. Os resultados indicam que a estatística gradiente é competitivo às estatísticas de escore, da razão de verossimilhança e Wald pois o poder do teste é similar. Note que o poder da estatística gradiente é maior aos outros quando se trata de amostra pequena e moderada ($n = 100, 200$).



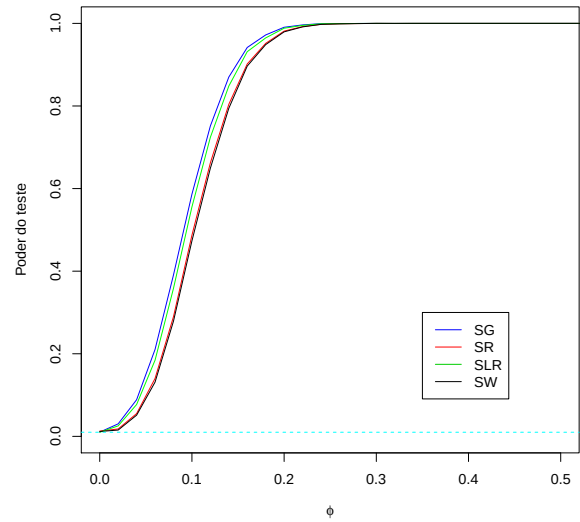
(a) $n = 100$



(b) $n = 200$

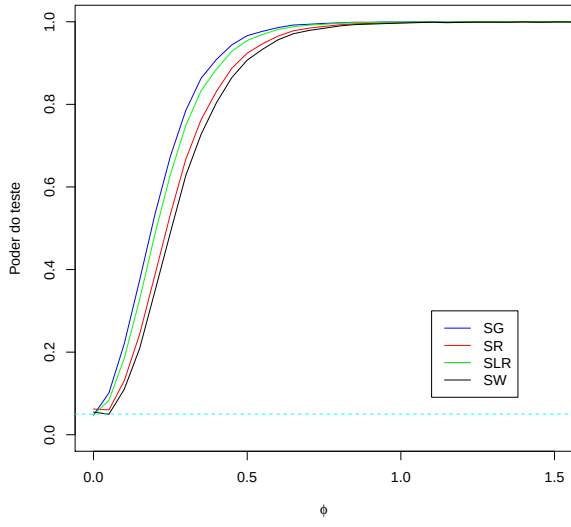


(c) $n = 400$

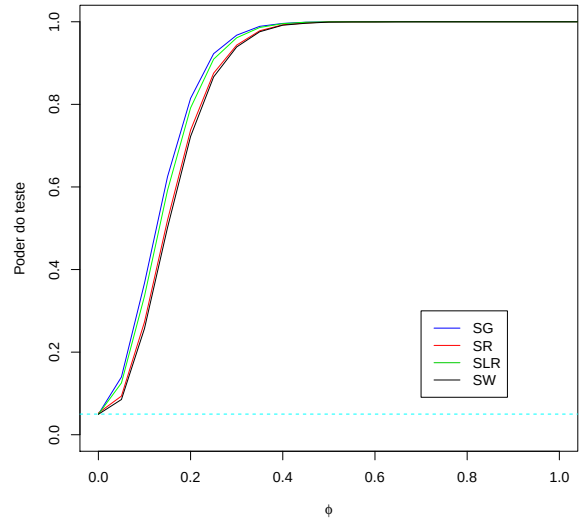


(d) $n = 600$

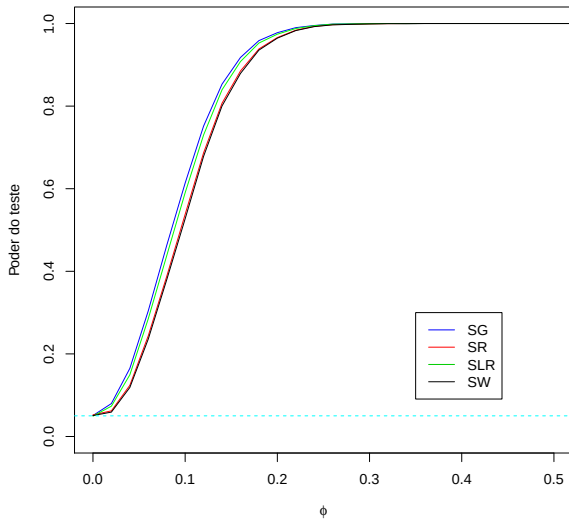
Figura 2.6: Poder do teste S_G, S_R, S_{LR} e S_W para diferentes valores de ϕ e tamanhos de amostra $n = 100, 200, 400$ e 600 quando $\theta = 0, 7$, ao nível nominal de 1%.



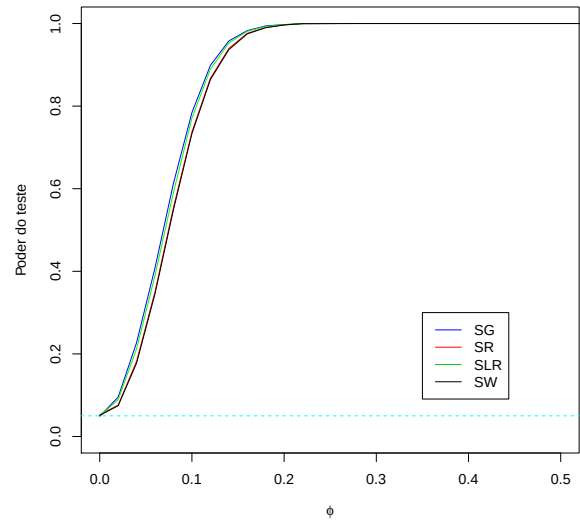
(a) $n = 100$



(b) $n = 200$

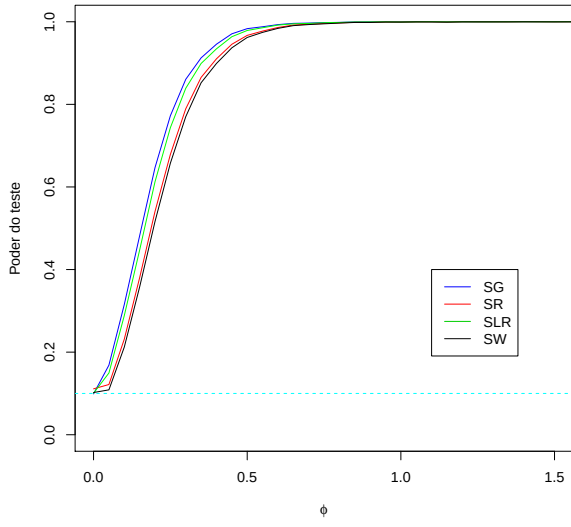


(c) $n = 400$

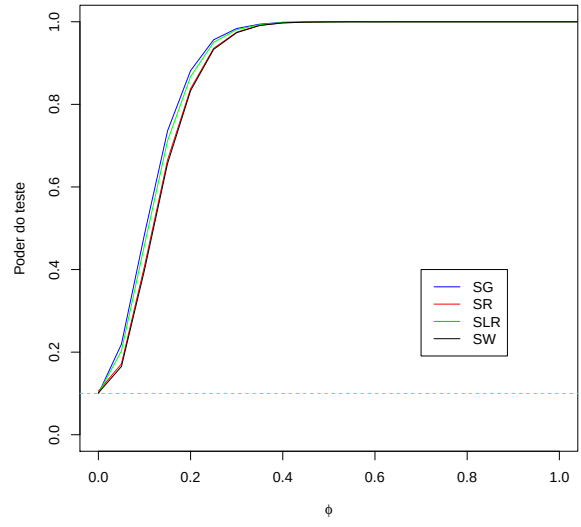


(d) $n = 600$

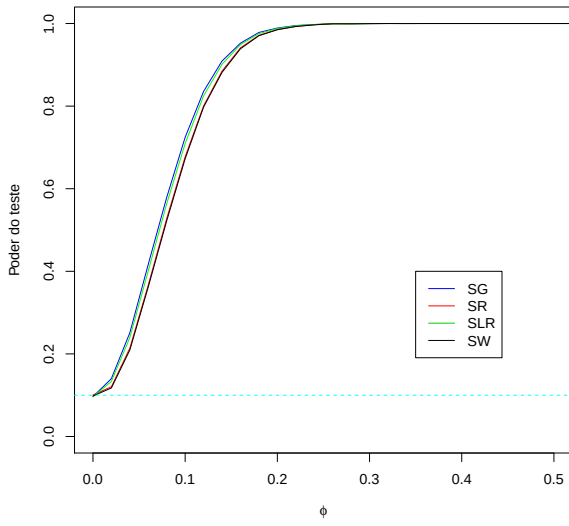
Figura 2.7: Poder do teste S_G, S_R, S_{LR} e S_W para diferentes valores de ϕ e tamanhos de amostra $n = 100, 200, 400$ e 600 quando $\theta = 0,7$, ao nível nominal de 5%.



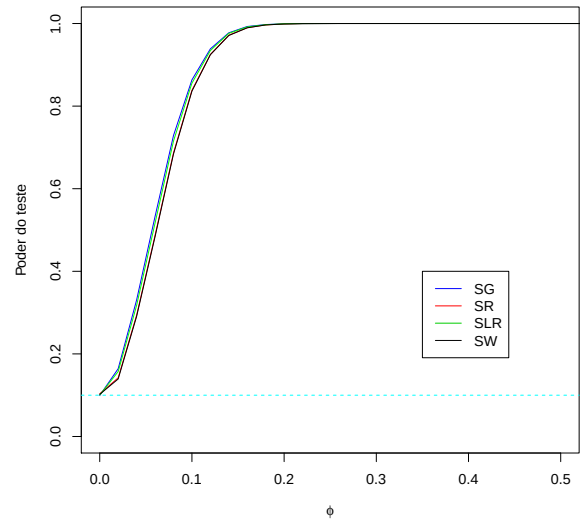
(a) $n = 100$



(b) $n = 200$



(c) $n = 400$



(d) $n = 600$

Figura 2.8: Poder do teste S_G, S_R, S_{LR} e S_W para diferentes valores de ϕ e tamanhos de amostra $n = 100, 200, 400$ e 600 quando $\theta = 0,7$, ao nível nominal de 10%.

2.3.3 Simulação com covariáveis no modelo ZI-Poisson

Para avaliar o desempenho da estatística gradiente (S_G) com a estatística da razão de verossimilhança (S_{LR}), escore (S_R) e Wald (S_W) foram simulados valores do modelo de regressão ZI-Poisson (2.13) com parâmetros $\phi = 0$ e $\theta_i = \exp(\beta_1 + \beta_2 \mathbf{x}_{i1} + \beta_3 \mathbf{x}_{i2})$, $i = 1, \dots, n$, em que $\beta_1 = -1$, $\beta_2 = -0,5$, $\beta_3 = 1$ e a covariável $\mathbf{x}_{i1}$ foi gerada de uma distribuição Bernoulli com probabilidade de sucesso igual a 0,5 e a covariável $\mathbf{x}_{i2}$ gerada de uma distribuição Uniforme no intervalo (0,1). A avaliação é em termos da estimação das taxas de rejeição nula, e comparação da distribuição empírica das estatística gradiente com a distribuição teórica. O estudo é baseada em 1000 amostras de Monte Carlo de tamanhos $n = 30, 50, 100, 200$ e 400 do modelo regressão ZI-Poisson com $\phi = 0$.

Tabela 2.8: *Taxas de rejeição da hipótese para testar $H_0 : \phi = 0$ no modelo de regressão ZI-Poisson.*

n	(%)	S_G	S_{LR}	S_R	S_W
30	1	1,50	0,80	0,50	0,00
	5	6,30	3,90	2,20	0,30
	10	10,80	8,00	5,60	1,50
50	1	1,30	0,70	0,30	0,00
	5	4,40	2,90	1,30	0,50
	10	9,60	6,90	3,90	2,30
100	1	0,90	0,30	0,10	0,00
	5	3,70	2,70	1,10	0,80
	10	8,40	6,50	3,70	3,00
200	1	1,10	0,80	0,40	0,20
	5	4,80	3,80	2,70	1,90
	10	9,00	8,20	6,30	5,60
400	1	1,00	0,80	0,40	0,20
	5	5,50	4,70	2,80	2,70
	10	10,60	9,70	7,70	7,30

Na Tabela 2.8 é apresentada as taxas de rejeição nula para os valores nominais de 1%, 5% e 10%. No caso quando a amostra é de tamanho $n = 30$, as taxas das estatísticas escore e Wald respectivas apresentam resultados subestimados aos valores nominais. E as taxas de rejeição empíricas do teste gradiente apresenta melhor aproximação aos valores nominais na medida que aumenta o tamanho amostral. A Tabela 2.9 mostra os percentis empíricos das estatísticas de teste que são satisfatoriamente aproximados com o percentil teórico, na medida que aumenta o tamanho das amostras, indicando que a distribuição assintótica é uma Qui-quadrado com *um* grau de liberdade.

Tabela 2.9: *Percentis estimados sob $H_0 : \phi = 0$.*

n	Estatísticas	$P_{,7} = 1,07$	$P_{,8} = 1,64$	$P_{,9} = 2,71$	$P_{,95} = 3,84$	$P_{,99} = 6,63$
30	S_G	0,96	1,56	2,87	4,46	9,23
	S_{LR}	0,72	1,08	1,84	2,57	5,25
	S_R	0,86	1,36	2,33	3,47	6,89
	S_W	0,67	0,95	1,54	2,01	3,13
50	S_G	0,88	1,47	2,68	3,95	6,97
	S_{LR}	0,71	1,15	1,93	2,79	4,99
	S_R	0,82	1,34	2,37	3,37	6,07
	S_W	0,68	1,07	1,76	2,38	3,75
100	S_G	1,10	1,75	2,78	3,84	5,84
	S_{LR}	0,90	1,38	2,09	2,93	4,09
	S_R	1,02	1,60	2,52	3,50	5,10
	S_W	0,87	1,33	1,99	2,58	3,62
200	S_G	0,94	1,52	2,44	3,42	6,93
	S_{LR}	0,83	1,29	2,01	2,73	5,18
	S_R	0,90	1,44	2,26	3,15	6,25
	S_W	0,82	1,27	1,96	2,67	4,92
400	S_G	0,92	1,48	2,65	3,85	6,69
	S_{LR}	0,84	1,32	2,27	3,23	5,28
	S_R	0,89	1,42	2,51	3,60	6,18
	S_W	0,84	1,31	2,24	3,13	5,08

Também é de interesse estimar as probabilidades de rejeição das estatísticas gradiente, razão de verossimilhança, score e Wald em relação ao teste $H_0 : \beta_2 = \beta_3 = 0$ definido em (2.40) para três valores nominais $\alpha = 1\%, 5\%$ e 10% considerando os tamanhos de amostra $n = 100, 200, 400$ e 600 . Para diferentes valores de $\phi = 1$ e 2 . Os resultados apresentam-se na Tabela 2.10 e observamos que a medida que o tamanho de amostra aumenta os valores empíricos das taxas de rejeição das estatísticas gradiente e razão de verossimilhança são melhores aproximados a seus valores nominais.

Tabela 2.10: Taxas de rejeição da hipótese $H_0 : \beta_2 = \beta_3 = 0$ para S_G, S_{LR}, S_R e S_W para diferentes tamanhos de amostra no modelo de regressão ZI-Poisson

n	$(\alpha\%)$	$\phi = 1, 0$				$\phi = 2, 0$			
		S_G	S_{LR}	S_R	S_W	S_G	S_{LR}	S_R	S_W
100	1	2,20	2,30	4,00	1,30	3,20	1,50	4,40	0,50
	5	6,50	6,30	8,40	5,50	7,40	6,30	9,30	3,50
	10	10,30	10,30	12,50	9,40	13,50	13,00	15,40	7,30
200	1	1,10	1,20	1,30	0,70	0,90	0,90	1,30	0,30
	5	5,20	5,40	5,70	4,40	5,80	5,60	6,60	4,40
	10	9,90	9,90	11,50	9,10	10,70	10,50	11,50	9,10
400	1	1,20	1,20	1,60	1,10	1,10	1,10	1,30	1,00
	5	5,40	5,40	5,80	5,10	4,90	4,90	5,30	4,50
	10	10,90	10,90	11,30	10,50	9,80	10,00	10,30	9,40
600	1	1,00	1,00	1,10	1,00	1,20	1,20	1,40	1,10
	5	5,00	4,90	5,30	4,80	6,20	6,20	6,50	5,80
	10	9,90	9,90	10,20	9,90	10,90	11,00	11,10	10,50

2.3.4 Simulações com covariáveis no modelo ZI-Geométrica

Para avaliar o desempenho da estatística gradiente (S_G) com a estatística da razão de verossimilhança (S_{LR}), escore (S_R) e Wald (S_W) foram simulados valores do modelo de regressão ZI-Geométrica (2.17) com parâmetros

$$\phi = 0 \text{ e } \theta_i = \frac{\exp(\beta_1 + \beta_2 \mathbf{x}_{i1} + \beta_3 \mathbf{x}_{i2})}{(1 + \exp(\beta_1 + \beta_2 \mathbf{x}_{i1} + \beta_3 \mathbf{x}_{i2}))}, \quad i = 1, \dots, n,$$

em que $\beta_1 = 1, \beta_2 = -1, \beta_3 = -0,5$ e a covariável $\mathbf{x}_{i1}$ foi gerada de uma distribuição Bernoulli com probabilidade de sucesso igual a $0,5$ e a covariável $\mathbf{x}_{i2}$ foi gerada de uma distribuição Uniforme no intervalo $(0, 1)$. A avaliação é em termos da estimação das taxas de rejeição nula, e comparação da distribuição empírica das estatística gradiente com a distribuição teórica. O estudo é baseado em 1000 amostras de Monte Carlo de tamanhos $n = 30, 50, 100, 200, 400$ e 600 do modelo regressão ZI-Geométrica com $\phi = 0$.

Na Tabela 2.11 é apresentado as taxas de rejeição nula para os valores nominais de 1%, 5% e 10% em que é observado as taxas de rejeição empíricas do teste gradiente são melhor aproximadas aos valores nominais comparadas com os testes usuais que estão subestimando os níveis nominais, em todos os casos. Na Tabela 2.12 temos os percentis os empíricos das quatro estatísticas de teste aproximados aos valores do percentil teórico, indicando que distribuição assintótica é uma Qui-quadrado com *um* grau de liberdade.

Tabela 2.11: *Taxas de rejeição da hipótese $H_0 : \phi = 0$ no modelo de regressão ZI-Geométrica.*

n	(%)	S_G	S_{LR}	S_R	S_W
30	1	1,40	0,30	0,00	0,00
	5	4,90	2,70	0,60	0,10
	10	10,50	6,10	2,60	0,60
50	1	1,50	0,90	0,00	0,00
	5	5,60	2,90	1,20	0,50
	10	10,10	7,20	3,40	2,10
100	1	1,20	0,40	0,00	0,00
	5	4,30	3,00	1,30	1,10
	10	8,70	7,30	3,70	3,10
200	1	1,20	0,70	0,10	0,10
	5	4,40	3,60	2,10	1,90
	10	10,00	8,50	5,70	5,10
400	1	1,30	0,50	0,10	0,00
	5	4,30	3,90	2,40	2,40
	10	9,30	8,20	6,40	6,20
600	1	1,60	1,10	0,60	0,60
	5	6,00	5,20	3,70	3,50
	10	11,80	10,90	8,70	8,40

Tabela 2.12: *Percentis estimados sob $H_0 : \phi = 0$.*

n	Estatísticas	$P_{,7} = 1,07$	$P_{,8} = 1,64$	$P_{,9} = 2,71$	$P_{,95} = 3,84$	$P_{,99} = 6,63$
30	S_G	0,92	1,46	2,77	3,83	7,44
	S_{LR}	0,80	1,24	2,24	2,97	5,21
	S_R	0,66	0,95	1,59	2,04	3,16
	S_W	0,61	0,88	1,38	1,74	2,48
50	S_G	0,92	1,52	2,72	4,09	7,66
	S_{LR}	0,83	1,34	2,26	3,29	5,87
	S_R	0,69	1,06	1,71	2,40	4,26
	S_W	0,66	1,00	1,56	2,12	3,34
100	S_G	0,94	1,43	2,52	3,65	7,17
	S_{LR}	0,86	1,29	2,21	3,12	5,86
	S_R	0,75	1,09	1,79	2,48	4,35
	S_W	0,73	1,05	1,70	2,32	3,88
200	S_G	0,94	1,46	2,70	3,73	6,76
	S_{LR}	0,89	1,36	2,45	3,35	5,89
	S_R	0,81	1,20	2,09	2,81	4,72
	S_W	0,80	1,18	2,04	2,71	4,43
400	S_G	1,03	1,48	2,56	3,64	6,82
	S_{LR}	0,99	1,41	2,40	3,37	6,16
	S_R	0,91	1,29	2,15	2,95	5,19
	S_W	0,91	1,28	2,12	2,89	5,02
600	S_G	1,09	1,78	2,94	4,12	7,22
	S_{LR}	1,05	1,71	2,79	3,86	6,64
	S_R	0,98	1,58	2,53	3,44	5,71
	S_W	0,98	1,57	2,50	3,39	5,57

Da mesma forma que o modelo anterior foram estimadas as taxas de rejeição das estatísticas gradiente, da razão de verossimilhança, score e Wald em relação ao teste $H_0 : \beta_2 = \beta_3 = 0$ para três valores nominais $\alpha = 1\%, 5\%$ e 10% considerando os tamanhos de amostra $n = 100, 200, 400$ e 600 , para diferentes valores de $\phi = 1$ e 2 . Os resultados são apresentados na Tabela 2.13 e observamos que a medida que o tamanho de amostra aumenta os valores empíricos das taxas de rejeição são melhor aproximados a seus valores nominais. Note que as taxas do teste escore apresenta valores superestimados enquanto que do Wald têm resultados subestimados aos valores nominais.

Tabela 2.13: Taxas de rejeição nula de $H_0 : \beta_2 = \beta_3 = 0$ para S_G, S_{LR}, S_R e S_W para diferentes tamanhos de amostra no modelo de regressão ZI-Geométrica

n	$(\alpha\%)$	$\phi = 1, 0$				$\phi = 2, 0$			
		S_G	S_{LR}	S_R	S_W	S_G	S_{LR}	S_R	S_W
100	1	1,60	1,50	5,40	0,50	3,00	1,50	4,80	0,40
	5	6,70	6,90	11,20	3,90	7,30	6,10	10,10	2,90
	10	11,70	12,50	15,90	10,00	12,00	12,00	15,30	6,80
200	1	0,80	0,80	2,20	0,60	1,40	1,40	3,60	0,70
	5	6,10	6,30	7,90	5,90	5,90	6,20	8,10	5,10
	10	10,40	10,90	13,10	10,30	10,70	11,00	14,30	9,80
400	1	0,50	0,50	0,90	0,50	1,20	1,30	1,70	1,30
	5	5,60	5,60	6,10	5,60	5,50	5,70	6,70	5,40
	10	9,80	9,90	10,80	9,80	10,40	11,10	12,10	10,60
600	1	1,30	1,40	1,60	1,10	1,40	1,40	1,60	1,20
	5	4,40	4,50	5,10	4,50	4,50	4,60	5,50	4,40
	10	10,00	10,20	10,90	10,20	10,00	10,10	10,70	9,80

2.4 Aplicação

Aplicamos os resultados obtidos neste capítulo em dois conjuntos de dados reais. O primeiro conjunto de dados foram introduzidos por Carrivick *et al.* (2003) num estudo sobre lesões sofridas por empregados de limpeza em um hospital na Austrália. O segundo conjunto dados, analisado por Ridout *et al.* (2001), refere-se ao número de raízes produzidas por brotos de macieira colunar Trajan micropropagadas.

2.4.1 Aplicação 1: Lesões em empregados da limpeza.

Os dados referem-se ao número de lesões sofrido por empregados da limpeza de um hospital na Austrália. Nota-se neste conjunto de dados que a proporção de empregados que não sofreram nenhuma lesão foi 78,07% de um total de 342 empregados.

Número de lesões	0	1	2	3	4	5	6	7	8
Número de empregados	267	52	16	4	2	0	0	0	1

Nesse conjunto de dados ajustamos o modelo ZI-Poisson e ZI-Geométrica. As estimativas de máxima verossimilhança (EMV) dos parâmetros dos modelos são apresentados na Tabela 2.14.

Tabela 2.14: EMV para os parâmetros com ajuste dos modelos ZI-Poisson e ZI-Geométrica para os dados de lesões em empregados da limpeza.

Modelo	$\hat{\theta}$ (erro padrão)	$\hat{\omega}$ (erro padrão)
ZIP	0,8638 (0,1341)	0,6209 (0,0536)
ZIG	0,6696 (0,0706)	0,6725 (0,0740)

As quantidades em parenteses são os erros padrões das estimativas de máxima verossimilhança. Os modelos diferem bastante na estimativa do zero determinístico, pois o modelo ZI-Poisson estima 62,09 % de zeros que são determinísticos nos dados, enquanto que o modelo ZI-Geométrica estima, aproximadamente, 67,25% de zeros determinísticos e 32,75 % de zeros provêm de forma aleatória. A EMV da proporção de empregados da limpeza que não sofreram acidentes no trabalho resultou em 78,07%. As estimativas são aproximadas com a proporção amostral.

Os resultados dos testes de hipóteses nula $H_0 : \omega = 0$ contra $H_1 : \omega \neq 0$ são apresentados na Tabela 2.15. Podemos observar que as quatro estatísticas, avaliadas pelos modelos ZI-Poisson e ZI-Geométrica, indicam que a proporção de empregados que não sofreram lesões é significativamente diferente de zero.

Tabela 2.15: Estimativas das estatísticas e nível descritivo do teste de hipóteses $H_0 : \omega = 0$.

Estatísticas	Modelo ZI-Poisson		Modelo ZI-Geométrica	
	Estimativas	Nível descritivo	Estimativas	Nível descritivo
S_G	46,60	$8,6 \times 10^{-12}$	6,30	0,01
S_{LR}	34,00	$5,51 \times 10^{-9}$	5,37	0,02
S_R	39,47	$3,35 \times 10^{-10}$	4,08	0,04
S_W	19,31	$1,11 \times 10^{-5}$	3,96	0,05

2.4.2 Aplicação 2: Cultivos de maçãs

Nesta segunda aplicação consideramos o conjunto de dados extraído de [Ridout et al. \(2001\)](#). Esses dados se referem ao número de raízes produzidas por 270 brotos de macieira colunar Trajan micropropagadas. Os brotos foram produzidos em oito diferentes culturas, onde cada um possui uma das quatro concentrações distintas de citocinina BAP sob um fotoperíodo de 8 ou 16 horas. Em cada cultura foram utilizados grupos de 30 ou 40 brotos. Nos 140 brotos cultivados sob fotoperíodo de 8 horas, apenas dois falharam na produção de raízes. Já nos 130 brotos sob fotoperíodo de 16 horas, observaram-se 62 amostras falhas. O foco do estudo é evidenciar a influência da concentração de citocinina BAP em dois fotoperíodos no crescimento de raízes de macieiras Trajan. O tamanho amostral corresponde a $n = 270$ e a porcentagem de zeros observados foi de 23,7%. Os dados foram coletados de cada broto micropropagado: y_i : número de raízes; x_{i1} : concentração de citocinina BAP e x_{i2} : fotoperíodo (0=8-horas , 1=16-horas) , $i = 1, \dots, 270$. Na Figura 2.9 mostra a distribuição do número de raízes dos 270 brotos.

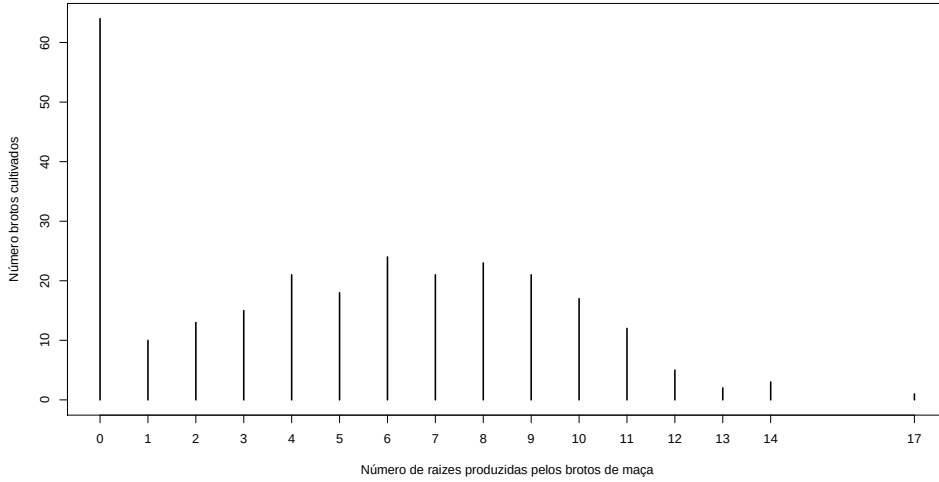


Figura 2.9: Distribuição do número de raízes produzidas pelos brotos de maçã.

Ajustamos para este conjunto de dados os modelos de regressão ZI-Poisson e ZI-Geométrica considerando as funções de ligação logarítmica e logística,

$$\log(\theta_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}, \quad i = 1, \dots, n$$

e

$$\log\left(\frac{\theta_i}{1 - \theta_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2},$$

respectivamente.

Também, calculamos os critérios AIC e SBC para os dois modelos, os resultados obtidos estão na Tabela 2.16.

Tabela 2.16: Valores de máximo da log-verossimilhança e os critérios AIC e SBC com ajuste dos modelos ZI-Poisson e ZI-Geométrica para dados de cultivos de maçã.

Modelo	$\max \ell(\cdot)$	AIC	SBC
ZIP	-677,5669	1363,134	1377,528
ZIG	-710,9049	1429,81	1444,204

Considerando os critérios estabelecidos na Tabela 2.16, selecionamos o modelo ZI-Poisson como um modelo apropriado para estes dados. Os EMVs são apresentados na Tabela 2.17 em que o símbolo * indica o nível descritivo da covariável X_i é não significativo ao nível de significância nominal de 5%.

Tabela 2.17: EMV para os parâmetros do modelo ZI-Poisson para os dados de cultivos de maçã.

Covariáveis	Parâmetros	Estimativas	E.P	$\max \ell(\cdot)$	Nível descritivo
x_1, x_2	β_0	1,969	0,051	-677,6	2×10^{-16}
	β_1	0,001	0,005		0,895*
	β_2	-0,289	0,062		$3,88 \times 10^{-6}$
x_1	β_0	1,890	0,049	-688,8	2×10^{-16}
	β_1	$3,4 \times 10^{-5}$	0,005		0,994*
x_2	β_0	1,974	0,032	-677,6	2×10^{-16}
	β_2	-0,288	0,062		$3,91 \times 10^{-6}$

Verificamos por meio das estimativas das estatísticas que a proporção de brotos sem raízes é diferente de zero. Os resultados são apresentados na Tabela 2.18, em que as estatísticas rejeitam a hipótese $H_0 : \omega = 0$ colocando-se em favor do modelo ZI-Poisson para modelar os dados.

Tabela 2.18: Estimativas das estatísticas de teste e nível descritivo para o teste $H_0 : \omega = 0$ do modelo ZI-Poisson.

Estatísticas	Estimativas	Nível descritivo
S_G	962,533	$2,5 \times 10^{-211}$
S_{LR}	216,766	$4,6 \times 10^{-49}$
S_R	4,861	$2,7 \times 10^{-2}$
S_W	47,380	$5,8 \times 10^{-12}$

Agora avaliamos a influência das covariáveis (concentração de citocinina BAP e fotoperíodo) na produção de raízes. Para isto testamos $H_0 : \beta_1 = \beta_2 = 0$ contra H_1 : Ao menos um é diferente de zero. As Tabelas 2.19 e 2.20 mostram que a covariável com informação de concentração de citocinina BAP não tem influencia no crescimento de raízes e que a produção delas depende altamente do fotoperíodo, sendo que β_2 é diferente do zero.

Tabela 2.19: Estimativas das estatísticas e nível descritivo para o teste $H_0 : \beta_1 = \beta_2 = 0$ do modelo ZI-Poisson.

Estatísticas	Estimativas	Nível descritivo
S_G	23,103	$9,62 \times 10^{-6}$
S_{LR}	22,537	$1,28 \times 10^{-5}$
S_R	21,592	$2,05 \times 10^{-5}$
S_W	21,324	$2,34 \times 10^{-5}$

Tabela 2.20: Testando a significância das hipóteses, $H_0 : \beta_1 = 0$ e $H_0 : \beta_2 = 0$.

Estatísticas	$H_0 : \beta_1 = 0$		$H_0 : \beta_2 = 0$	
	Estimativas	Nível descritivo	Estimativas	Nível descritivo
S_G	0,017	0,895	23,103	$1,535 \times 10^{-6}$
S_{LR}	0,018	0,895	22,537	$2,061 \times 10^{-6}$
S_R	0,017	0,895	21,591	$3,373 \times 10^{-6}$
S_W	0,018	0,895	21,323	3.879×10^{-6}

2.5 Conclusões

Neste capítulo foi estudada a performance da estatística gradiente, em problemas de teste de hipóteses para o parâmetro de inflação (ou deflação) de zeros e a influência das covariáveis no parâmetro de potência do modelo zero-inflacionado Série de Potência. A estatística gradiente foi comparada com as estatísticas da razão de verossimilhança, escore e Wald em termos da taxa de rejeição e do poder da estatística de teste. Além disso, mostramos os gráficos dos quantis empíricos e as curvas do poder do teste das distribuições das estatísticas, onde mostrou-se que a estatística gradiente tem um comportamento equivalente às estatísticas tradicionais: da razão de verossimilhança, escore e Wald. Também, foi observado que a vantagem da estatística gradiente é a simplicidade do seu cálculo, quando comparada com as estatísticas de Wald e de escore, pois não envolve o cálculo da matriz de informação de Fisher (ou observada). Esses resultados também foram confirmadas quando aplicados a dois conjuntos de dados de contagem com excesso de zeros da literatura. Portanto, concluímos que a estatística gradiente é uma alternativa às estatísticas tradicionais em problemas de testes de hipóteses em modelos zero-inflacionados Série de Potência.

Capítulo 3

Um Modelo de Sobrevivência Induzida por Fragilidade Discreta: Enfoque Clássico

3.1 Introdução

A análise de sobrevivência é composta por um conjunto de técnicas para a análise de dados que envolve tempos ou mais exatamente, tempos até a ocorrência de um evento de interesse em indivíduos (ou unidades experimentais). O evento de interesse pode ser por exemplo, a morte de um indivíduo de uma determinada doença, reincidência de um tumor, quebra de uma componente eletrônica, falha de um equipamento industrial, resposta a um inquérito no âmbito da psicologia, inadimplentes de clientes de banco, cancelamento de cartão de crédito de uma financeira entre outros. Assim, a análise estatística que envolvem dados de tempo de um evento de interesse (ou tempos de sobrevivência) desempenha um papel importante na medicina, epidemiologia, biologia, demografia, economia, engenharia e outras áreas do conhecimento.

A análise de sobrevivência implicitamente assume que as populações são homogêneas, ou seja, todos os indivíduos têm o mesmo risco de morte. Como mencionado no Capítulo 1, é frequentemente importante considerar a população heterogênea, ou seja, uma mistura de indivíduos com diferentes riscos. Um modelo de fragilidade fornece uma maneira conveniente de introduzir heterogeneidade e associações não observadas em modelos para dados de sobrevivência. Em sua forma mais simples, a fragilidade é um fator de proporcionalidade aleatório não observado que modifica a função de risco de um indivíduo, ou de indivíduos relacionados. Em essência, o conceito remonta ao trabalho de [Greenwood & Yule \(1920\)](#) em "propensão a acidentes". O termo de fragilidade foi introduzido por [Vaupel *et al.* \(1979\)](#) em modelos de sobrevivência univariadas. Aplicações a problemas na análise de dados

de sobrevivência multivariada a partir do trabalho de Clayton (1978).

Estudos sobre modelos de fragilidade geralmente assumem uma variável aleatória contínua e não negativa. Distribuições de fragilidade frequentemente utilizadas incluem a distribuição Gama (Hougaard, 1984, Vaupel *et al.*, 1979), Inversa Gaussiana (Hougaard, 1984), Log-Normal (dos Santos *et al.*, 1995, Vu & Knuiman, 2002) ou a distribuição Estável Positiva (Hougaard, 1986). Por meio da perspectiva computacional essas distribuições são convenientes já que é simples obter de forma fechada as funções de sobrevivência, densidade e risco, a partir da função transformada de Laplace. Por outro lado, as distribuições de fragilidade contínuas não permitem ter risco zero. Aalen & Hjort (2002) propuseram o processo de Poisson Composto Contínuo (PPCC) para a distribuição de fragilidade que considera um subgrupo de indivíduos não susceptíveis ao evento de interesse. Uma das propriedades matemáticas desse modelo é que, o ponto de massa em zero corresponde à inclusão de um subgrupo dos indivíduos não susceptíveis. Isso muitas vezes é biologicamente razoável, uma vez que apenas um subgrupo de indivíduos pode ter a composição genética que o torna vulnerável a determinadas doenças. Existem algumas situações nas quais é apropriado considerar a fragilidade distribuída discretamente, por exemplo, quando a heterogeneidade dos tempos de sobrevivência surge por causa da presença de um número aleatório de falhas por unidade ou pela causa da exposição a danos em um número aleatório de ocasiões (Aalen & Tretli, 1999, Moger *et al.*, 2004).

Na literatura são encontrados dois tipos de modelos de fragilidade discretos. O primeiro tipo é construído por meio da separação da variável de fragilidade, Z , em aqueles com números fixos ou aleatórios com pontos de massa de probabilidade. A maioria de tais referências acabam por considerar misturas finitas (Xue & Brookmeyer, 1997), em que Z é um grupo ou indicador de estrato assumindo poucos valores, ao vez de considerar uma distribuição de probabilidade sobre um suporte mais amplo. Estes autores afirmaram que seu principal resultado também é válido com uso das distribuições de fragilidade discretas, bem como as contínuas, mas não prosseguiram essa linha de investigação. Li & Zhong (2002) referiram ao modelo de riscos proporcionais com fragilidade discreta ao modelo de mistura finita como em Xue & Brookmeyer (1997). Recentemente, Wienke (2010) considerou o Processo de Poisson Composto Discreto (PPCD) para os modelos de fragilidade, e afirma que pode ser útil para um modelo de sobrevivência que contenha uma proporção de unidades experimentais para os quais o evento de interesse não ocorre mesmo após um período longo de observação (modelo de sobrevivência com fração de cura). O segundo tipo de modelos de fragilidade discreta considera que a variável de fragilidade é um número aleatório de componentes com massa de probabilidade nos números inteiros não negativos. Embora fora mencionada a possibilidade desse modelo, não foi investigada em detalhe. Recentemente, Caroni *et al.* (2010) desenvolveram os modelos de fragilidade discreta utilizando as distribuições Geométrica, Poisson e Binomial e Ata & Özel (2013) propuseram o modelo de fragilidade discreta baseado no PPCD. Os modelos propostos por esses autores podem ser

considerados como caso particular do modelo de sobrevivência com fração de cura proposto por Tsodikov *et al.* (2003).

Em muitas aplicações, o modelo de fragilidade assume um modelo de risco proporcional condicionado ao efeito aleatório (fragilidade), em que uma variável aleatória não negativa (fragilidade) Z , indica o nível de risco individual e atua multiplicativamente na função de risco. Mais especificamente, a função do risco condicionada à fragilidade pode ser escrita como

$$h(t|Z) = Z h_0(t), \quad (3.1)$$

em que $h_0(t)$ é a função de risco de base (ou basal). Note que a função de risco de base $h_0(t)$ pode ser escolhida não parametricamente ou parametricamente (como por exemplo, as distribuições Weibull, Exponencial, Gama, Gompertz e Log-Logístico). O modelo de fragilidade pode também ser representado por sua função de sobrevivência condicional

$$S(t|Z) = Pr(T > t|Z) = \exp\left(-Z \int_0^t h_0(u)du\right) = \exp(-ZH_0(t)), \quad (3.2)$$

em que $H_0(t) = \int_0^t h_0(u)du$ é uma função de risco acumulado.

A função de sobrevivência incondicional (ou marginal), $S(t)$, pode ser obtida por integração sobre o suporte da distribuição da variável aleatória Z . Assim, a função de sobrevivência incondicional é dada por

$$S(t) = \int S(t|z)dF(z)dz = E(S(t|Z)) = \mathcal{L}[H_0(t)],$$

em que $F(z)$ é a função de distribuição de Z e $\mathcal{L}(s)$ é a transformada de Laplace de Z . As distribuições mais comuns em muitas aplicações são as distribuições Gama, Inversa Gaussiana, Positiva Estável, entre outros.

Em (3.2) se a fragilidade for discreta com suporte nos inteiros não negativos, $\{0, 1, 2, \dots\}$ e a distribuição de probabilidades de Z especificada por $Pr[Z = z] = p_z$, para $z = 0, 1, \dots$, então a função de sobrevivência incondicional pode ser escrita como

$$S(t) = \sum_{z=0}^{\infty} S(t|z)p_z = E_Z(S(t|Z)) = G_Z[S_0(t)], \quad (3.3)$$

em que $G_Z(s)$ é a função geradora de probabilidade de Z e $S_0(t) = \exp(-H_0(t))$ é a função de sobrevivência basal. Note que $S(t)$ em (3.3) é a função de sobrevivência proposta por Rodrigues *et al.* (2009b). Também é fácil demonstrar que se $Pr(Z = 0) > 0$, a função de sobrevivência em (3.3) é imprópria, ou seja, $\lim_{t \rightarrow \infty} S(t) = G_Z(0) = Pr(Z = 0) > 0$. Essas propriedades caracterizam os modelos de sobrevivência com fração de cura (Maller & Zhou, 1996). No caso que $Pr(Z = 0) = 0$, a função de sobrevivência é própria, ou seja, $\lim_{t \rightarrow \infty} S(t) = G_Z(0) = Pr(Z = 0) = 0$ e $\lim_{t \rightarrow 0} S(t) = G_Z(1) = 1$.

A literatura sobre modelos de sobrevivência com fração de cura é extensa e está em rápida evolução pois, estes modelos descrevem cada vez melhor os mecanismos biológicos dos pacientes frente a uma doença. Os livros de Maller & Zhou (1996) e Ibrahim *et al.* (2001), como também artigos de Cooner *et al.* (2006), Cooner *et al.* (2007), Tournoud & Ecochard (2007), Zhao *et al.* (2009), de Castro *et al.* (2009), Rodrigues *et al.* (2009a), Rodrigues *et al.* (2009b), de Castro *et al.* (2010), Cancho *et al.* (2010), Cancho *et al.* (2011) e recentes publicações de Eudes *et al.* (2013), Milani *et al.* (2015), Ortega *et al.* (2015), Rodrigues *et al.* (2015), Macera *et al.* (2015), Morita *et al.* (2016), Cordeiro *et al.* (2016), Cancho *et al.* (2016) poderiam ser citados como alguns exemplos para desenvolvimento desta pesquisa de trabalho.

Neste Capítulo, propomos um novo modelo de sobrevivência induzido por fragilidade discreta com distribuição ZISP dada em (2.1) para dados de sobrevivência. A distribuição do tempo para o acontecimento do evento de interesse neste tipo de modelos é escolhida pelos próprios conjuntos de dados por meio do comportamento da função de risco.

Este Capítulo está organizado como segue. Na Seção 3.2 formulamos o modelo de sobrevivência com fração de cura, baseada na distribuição zero-inflacionado Série de Potência. O procedimento inferencial por meio do método de máxima verossimilhança é tratado na Seção 3.2.2. Os resultados de um estudo de simulação são apresentados na Seção 3.3. Uma aplicação com dados reais é apresentada na Seção 3.4. Finalmente, na Seção 3.5 é apresentado algumas conclusões do Capítulo.

3.2 O modelo

O novo modelo resulta ao considerar que a distribuição da variável de fragilidade Z definida em (3.3) é a distribuição ZISP dada em (2.1) reproduzida novamente como

$$P(Z; \theta, \phi) = \begin{cases} \frac{\phi + a_0 A(\theta)^{-1}}{1 + \phi} & , z = 0 \\ \frac{a_z \theta^z A(\theta)^{-1}}{1 + \phi} & , z = 1, 2, 3, \dots \end{cases} \quad (3.4)$$

A função geradora de probabilidade para a fragilidade Z com distribuição dada em (3.4) é dada por

$$G_Z(\xi) = \frac{\phi}{1 + \phi} + \frac{A(\theta\xi)}{(1 + \phi)A(\theta)}, \quad |\xi| \leq 1, \quad (3.5)$$

em que $\theta \in (0, s)$ (s pode ser ∞) e $-a_0/A(\theta) \leq \phi < \infty$. Sob essas condições, de (3.5) temos que a função de sobrevivência incondicional dada em (3.3) resulta em

$$S(t) = \frac{\phi + A(\theta S_0(t))A(\theta)^{-1}}{1 + \phi}. \quad (3.6)$$

Já que $\lim_{t \rightarrow \infty} S(t) = \frac{\phi}{1+\phi} + \frac{a_0}{(1+\phi)A(\theta)} > 0$, a função de sobrevivência (3.6) é imprópria. Note que, se $\phi = 0$ em (3.6) então $S(t) = A(\theta S_0(t))/A(\theta)$, que resulta no modelo proposto por Cancho *et al.* (2013). Este modelo inclui como caso particular os modelos propostos por Caroni *et al.* (2010).

De (3.6) também é obtido a fração de risco zero (curados), que é dada por

$$p_0 = \lim_{t \rightarrow \infty} S(t) = \frac{\phi}{1 + \phi} + \frac{a_0}{(1 + \phi)A(\theta)}. \quad (3.7)$$

Note que, quando $\theta \rightarrow s$ a fração de risco zero tende para o zero determinístico, $\phi/(1 + \phi)$, enquanto que $\theta \rightarrow 0$ a fração de risco zero tende para um .

Nesta modelagem a fração de risco zero em (3.7) modela os indivíduos que não estão em risco porque eles não apresentaram o evento de interesse. A fração apresenta duas componentes. A primeira é uma fração de risco zero determinístico referida aos indivíduos imunes que são os pacientes curados por características genéticas e possuem uma forte resistência imunológica que tornam eles ser livres da doença. A segunda componente é uma fração de risco zero aleatória referida à indivíduos que inicialmente estavam em risco e pelo tratamento ou exposição à radiação resultaram curados. Assim ressaltamos que a decomposição dessa fração é importante porque permite determinar que proporção de pacientes tornam-se curados devido à fatores genéticos e que proporção é devido ao tratamento.

Portanto, a fração dos curados por tratamento é dado por

$$p_0^* = \frac{a_0}{(1 + \phi)A(\theta)}. \quad (3.8)$$

A correspondente função de densidade de (3.6) é dada pela expressão

$$f(t) = -S'(t) = \frac{\theta f_0(t)}{1 + \phi} \frac{A'(\theta S_0(t))}{A(\theta)}, \quad (3.9)$$

em que $A'(\theta S(t)) = A'(\theta) |_{\theta=\theta S_0(t)}$ e $f_0(t) = -dS_0(t)/dt$ é a função de densidade basal. Consequentemente a função de risco correspondente é dada por

$$h(t) = \frac{\theta f_0(t)A(\theta)^{-1}A'(\theta S_0(t))}{\phi + A(\theta)^{-1}A(\theta S_0(t))}. \quad (3.10)$$

A função de densidade $f(t)$ e a função de risco $h(t)$ são impróprias, uma vez que a função

de sobrevivência $S(t)$ é imprópria. O modelo de sobrevivência com fração de cura (3.6), de agora em diante, será denominado modelo de sobrevivência zero-inflacionada Série de Potência com fração de cura (ZISP-FC). A função de sobrevivência dos indivíduos em risco ou indivíduos suscetíveis ao evento de interesse, denotada por $S_{nc}(t)$, é dada por

$$S_{nc}(t) = P(T > t | Z \geq 1) = \frac{A(\theta S_0(t)) - a_0}{A(\theta) - a_0}. \quad (3.11)$$

Note que, $S_{nc}(0) = 1$ e $S_{nc}(\infty) = 0$, ou seja, S_{nc} é uma função de sobrevivência própria. A partir do modelo dado em (3.11), com diferentes escolhas para a distribuição da variável de fragilidade Z e para a distribuição basal, pode resultar em várias famílias de distribuições propostas recentemente. Por exemplo, a família Weibull-Geométrica proposta por Barreto-Souza *et al.* (2010) é obtida se considerarmos que a variável Z segue uma distribuição Geométrica e a distribuição basal é uma distribuição Weibull. Agora se a distribuição de base tem distribuição Exponencial com taxa de falha $\theta > 0$ obtemos a distribuição proposta por Chahkandi & Ganjali (2009), membros desta família de distribuição tem função taxa de falha decrescente e inclui as famílias de distribuição proposta por Adamidis & Loukas (1998), Kus (2007) e Tahmasbi & Rezaei (2008). Para as propriedades desta família de distribuição veja o trabalho de Chahkandi & Ganjali (2009).

Existe uma relação matemática entre o modelo (3.6) e o modelo de mistura padrão (Berkson & Gage, 1952, Boag, 1949), que pode ser escrito como

$$S(t) = p_0 + (1 - p_0) S_{nc}(t), \quad (3.12)$$

em que p_0 dada em (3.7). Assim temos a fração dos não curados que é dada por $1 - p_0 = \frac{1 - a_0 A(\theta)^{-1}}{1 + \phi}$.

Baseado nos resultados de Li *et al.* (2001), os modelos ZISP-FC representados como modelo de mistura em (3.12) são identificáveis desde que $S_{nc}(t)$ seja uma função parametricamente especificada, que no caso é dada em (3.11).

3.2.1 Alguns casos particulares

Nesta subseção apresentamos alguns modelos que resulta de nossa formulação. Particularmente, é considerado quando Z segue uma distribuição ZIP, ZIG e ZIL.

Se Z é uma variável com distribuição ZIP, então de (3.6), a função de sobrevivência é dada por

$$S(t) = \frac{\phi}{1 + \phi} + \frac{e^{-\theta(1 - S_0(t))}}{(1 + \phi)}. \quad (3.13)$$

Note que, quando $\phi = 0$ em (3.13), o modelo se reduz ao modelo de tempo de promoção

introduzida por [Chen et al. \(1999\)](#). De (3.7), a proporção de risco zero é dada por

$$p_0 = \lim_{t \rightarrow \infty} S(t) = \frac{\phi}{(1 + \phi)} + \frac{e^{-\theta}}{(1 + \phi)}.$$

De (3.9), a correspondente função densidade é dada por

$$f(t) = \frac{\theta f_0(t) e^{-\theta(1-S_0(t))}}{(1 + \phi)},$$

e a função de risco (3.10) correspondente é expressa por

$$h(t) = \frac{\theta f_0(t) e^{-\theta(1-S_0(t))}}{\phi + e^{-\theta(1-S_0(t))}}.$$

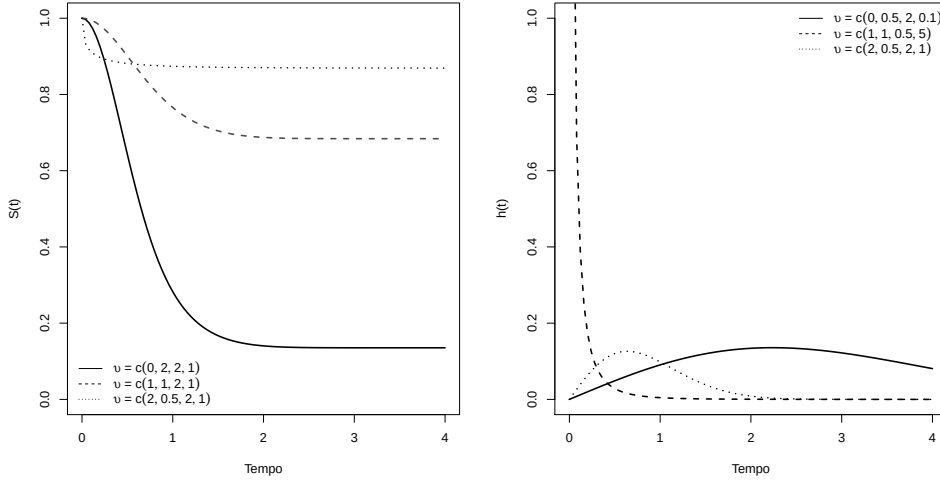


Figura 3.1: Funções de sobrevivência e risco do modelo ZIP-FC para alguns parâmetros selecionados.

Note que a função de risco decresce quando $\gamma_1 < 1$ indicando que a taxa de falha diminui ao longo do tempo, isto é, os acontecimentos do evento de interesse são baixos.

Agora, se Z tem distribuição ZIG, então em (3.6), a função de sobrevivência é dada pela seguinte expressão

$$S(t) = \frac{\phi}{1 + \phi} + \frac{(1 - \theta)}{(1 + \phi)(1 - \theta S_0(t))}. \quad (3.14)$$

A fração de risco zero é dada por

$$p_0 = \frac{\phi}{1 + \phi} + \frac{(1 - \theta)}{(1 + \phi)}.$$

As correspondentes funções de densidade e risco são dadas, respectivamente, por

$$f(t) = \frac{\theta(1 - \theta)f_0(t)}{(1 + \phi)(1 - \theta S_0(t))^2},$$

e a função de risco (3.10) correspondente é expressa por

$$h(t) = \frac{\theta(1 - \theta)f_0(t)(1 - \theta S_0(t))^2}{\phi + (1 - \theta)(1 - \theta S_0(t))^{-1}}.$$

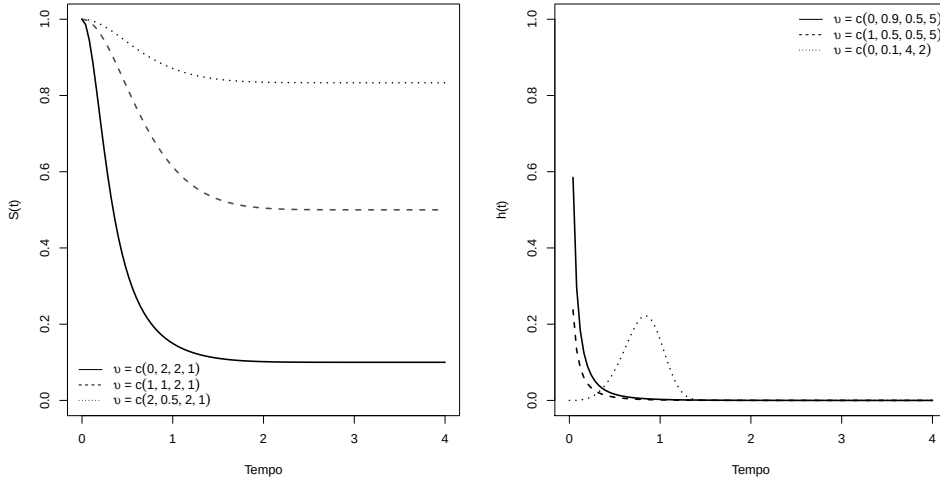


Figura 3.2: Funções de sobrevivência e risco do modelo ZIG-FC para alguns parâmetros selecionados.

Finalmente, quando Z tem distribuição ZIL, a função de sobrevivência incondicional é dada por

$$S(t) = \frac{\phi}{1 + \phi} + \frac{\log[1 - S_0(t)\theta]}{(1 + \phi)S_0(t) \log[1 - \theta]}. \quad (3.15)$$

A fração de risco zero é dada por

$$p_0 = \frac{\phi}{1 + \phi} - \frac{\theta}{(1 + \phi) \log[1 - \theta]}.$$

A função densidade associada ao modelo (3.15) com fração de risco zero é expressa por

$$f(t) = \frac{-f_0(t)}{1 + \phi} \frac{1}{S_0(t)^2 \log(1 - \theta)} \left[\frac{\theta S_0(t)}{1 - \theta S_0(t)} + \log(1 - \theta S_0(t)) \right],$$

e a correspondente função de risco dada em (3.10) é dada por

$$h(t) = \frac{\frac{-f_0(t)}{S_0(t)^2 \log(1-\theta)} \left[\frac{\theta S_0(t)}{1-S_0(t)} + \log(1-\theta S_0(t)) \right]}{\phi + \frac{\log(1-\theta S_0(t))}{S_0(t) \log(1-\theta)}}.$$

Os modelos de sobrevivência dados em (3.13)-(3.15) são denominados, respectivamente, de modelos de sobrevivência ZIP, ZIG e ZIL com fração de cura. Apresentamos as Figuras 3.1, 3.2 e 3.3 que mostram os diferentes comportamentos das funções de sobrevivência e risco dos modelos ZIP-FC, ZIG-FC e ZIL-FC, respectivamente, para alguns valores selecionados,

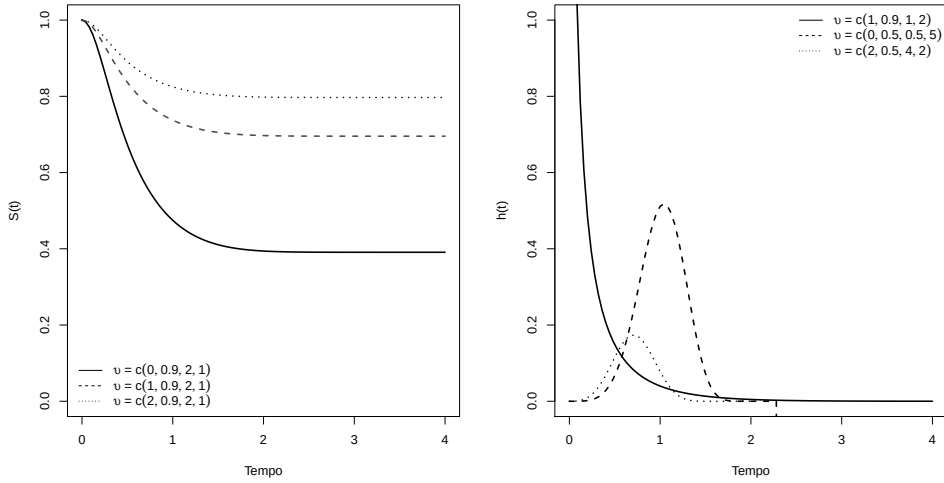


Figura 3.3: Funções de sobrevivência e risco do modelo ZIL-FC para alguns parâmetros selecionados.

3.2.2 Inferência

Considere os tempos de sobrevivência de n indivíduos, $T_1, \dots, T_n$ e $C_1, \dots, C_n$ os respectivos tempos de censura. Suponha que é observado $t_i = \min\{T_i, C_i\}$ e $\delta_i = I(T_i \leq C_i)$, em que $\delta_i = 1$ se t_i é um tempo de sobrevivência e $\delta_i = 0$ se é o tempo de censura a direita. Seja $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})$ o vetor de covariáveis para o i -ésimo indivíduo. Assumindo que o parâmetro θ do modelo (3.6) se relaciona com as covariáveis por meio de

$$g(\theta_i) = \eta_i = \eta(\mathbf{x}_i, \boldsymbol{\beta}) = \mathbf{x}_i \boldsymbol{\beta}, \quad i = 1, \dots, n, \quad (3.16)$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ é o vetor de parâmetros de regressão ($\boldsymbol{\beta} \in R^p$) associado ao vetor de covariável $\mathbf{x}_i$. Aqui $\eta(\cdot, \cdot)$ é uma função contínua duas vezes diferenciável. Finalmente, $g(\cdot)$ é a função de ligação contínua monótona e duas vezes diferenciável com domínio R^+ . Seja X

a matriz de derivadas de $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)$ com respeito a $\boldsymbol{\beta}^\top$. Assume-se que o $\text{posto}(X) = p$. Uma possível função de ligação para o parâmetro θ_i do modelo ZIP é $g(\theta_i) = \log(\theta_i)$ (ligação logarítmica) e para o parâmetro θ_i dos modelos ZIG e ZIL é $g(\theta_i) = \log(\theta_i/(1 - \theta_i))$ (ligação logística). Com esta função de ligação os modelos são identificáveis, como pode ser visto em Li *et al.* (2001).

Com a expressão (3.16), o logaritmo da função de verossimilhança de $\boldsymbol{\vartheta} = (\phi, \boldsymbol{\gamma}, \boldsymbol{\beta}^\top)$ sob censura não informativa pode ser escrito como

$$\begin{aligned} \ell(\boldsymbol{\vartheta}) = & -n \log(1 + \phi) + \sum_{i=1}^n \{ -\log(A(\theta_i)) + \delta_i [\log(\theta_i) + \log(f_0(t_i; \gamma)) \\ & + \log(A'(\theta_i S_0(t_i, \gamma)))] + (1 - \delta_i) \log[\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma))] \}. \end{aligned} \quad (3.17)$$

A estimativa de máxima verossimilhança (EMV) do parâmetro $\boldsymbol{\vartheta}$ pode ser obtida maximizando a função de log-verossimilhança.

Diferenciando (3.17) com respeito a $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ e ϕ , obtemos o vetor escore $\mathbf{U}^\top(\boldsymbol{\vartheta}) = (U_1^\top(\boldsymbol{\vartheta}), U_2^\top(\boldsymbol{\vartheta}), U_3(\boldsymbol{\vartheta}))$ cujos elementos são

$$\begin{aligned} U_1(\boldsymbol{\vartheta}) = \frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\beta}} = & \sum_{i=1}^n \left\{ \frac{-A'(\theta_i)}{A(\theta_i)} + \delta_i \left[\frac{1}{\theta_i} + \frac{A''(\theta_i S_0(t_i, \gamma))}{A'(\theta_i S_0(t_i, \gamma))} S_0(t_i, \gamma) \right] \right. \\ & \left. + (1 - \delta_i) \frac{\phi A'(\theta_i) + A'(\theta_i S_0(t_i, \gamma)) S_0(t_i, \gamma)}{\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma))} \right\} \frac{\partial \theta_i}{\partial \boldsymbol{\beta}}, \end{aligned} \quad (3.18)$$

$$\begin{aligned} U_2(\boldsymbol{\vartheta}) = \frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\gamma}} = & \sum_{i=1}^n \left\{ \delta_i \left[\frac{1}{f_0(t; \gamma)} \frac{\partial f_0(t; \gamma)}{\partial \gamma} + \frac{\theta_i A''(\theta_i S_0(t_i, \gamma))}{A'(\theta_i S_0(t_i, \gamma))} \frac{\partial S_0(t_i, \gamma)}{\partial \gamma} \right] \right. \\ & \left. + (1 - \delta_i) \frac{\theta_i A'(\theta_i S_0(t_i, \gamma))}{\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma))} \frac{\partial S_0(t_i, \gamma)}{\partial \gamma} \right\}, \end{aligned} \quad (3.19)$$

e

$$U_3(\boldsymbol{\vartheta}) = \frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \phi} = \frac{-n}{1 + \phi} + \sum_{i=1}^n \left\{ (1 - \delta_i) \frac{A(\theta_i)}{\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma))} \right\}, \quad (3.20)$$

em que $\frac{\partial \theta_i}{\partial \boldsymbol{\beta}} = \frac{\partial g^{-1}(\mathbf{x}_i \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \mathbf{x}_i^\top$.

Os EMV de $\hat{\boldsymbol{\vartheta}}$ também podem ser obtidos resolvendo o sistema de equações não lineares nos parâmetros, $U(\boldsymbol{\vartheta}) = 0$. Como as equações não lineares não têm solução explícita, consideramos o algoritmo do Newton-Rapson, implementada no *optim* no *software* R

(R Development Core Team, 2009).

Sob certas condições de regularidade (Cox & Hinkley, 1979), pode ser mostrado que a distribuição assintótica do estimador de máxima verossimilhança $\hat{\boldsymbol{\vartheta}}$ é Normal multivariada com vetor de médias $\boldsymbol{\vartheta}$ e matriz de covariâncias $\boldsymbol{\Sigma}(\hat{\boldsymbol{\vartheta}})$, a qual pode ser estimada por $\hat{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\vartheta}}) = \{J(\hat{\boldsymbol{\vartheta}})\}^{-1}$, avaliada em $\boldsymbol{\vartheta} = \hat{\boldsymbol{\vartheta}}$, em que

$$J(\boldsymbol{\vartheta}) = - \left[\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}^\top \partial \boldsymbol{\vartheta}} \right] = \begin{bmatrix} J_{\beta\beta} & J_{\beta\gamma} & J_{\beta\phi} \\ J_{\beta\gamma}^\top & J_{\gamma\gamma} & J_{\gamma\phi} \\ J_{\beta\phi}^\top & J_{\gamma\phi}^\top & J_{\phi\phi} \end{bmatrix}, \quad (3.21)$$

$J_{\beta\beta}$ é uma submatriz de ordem $p \times p$, $J_{\beta\gamma}$ de ordem $(p+s) \times (p+s)$, $J_{\beta\phi}$ de $(p+s+1) \times 1$, $J_{\gamma\gamma}$ de $s \times s$ e $J_{\phi\phi}$ é escalar. Os elementos da matriz de informação observada são apresentados no Apêndice C.2.

As estimações intervalares dos parâmetros do modelo são baseados na normalidade assintótica dos estimadores de máxima verossimilhança (EMV). O nível de confiança $100 \times (1 - \alpha)\%$ é assumido para calcular o intervalo de confiança de cada componente ϑ do vetor de parâmetros $\boldsymbol{\vartheta}$ dada por

$$\hat{\vartheta} \pm \mathcal{Z}_{\alpha/2} \sqrt{J^{-1}(\hat{\vartheta})} \quad (3.22)$$

em que $\mathcal{Z}_{\alpha/2}$ é o valor do $\alpha/2$ quantil superior da distribuição Normal padrão e $J^{-1}(\hat{\vartheta})$ é o elemento da diagonal da inversa da matriz $J(\boldsymbol{\vartheta})$ e corresponde ao estimador da variância do parâmetro ϑ .

Além da estimação dos parâmetros, o teste de hipóteses é outra questão importante. Sejam $\boldsymbol{\vartheta}_1$ e $\boldsymbol{\vartheta}_2$ subconjuntos disjuntos do vetor de parâmetros $\boldsymbol{\vartheta}$ do modelo de sobrevivência ZISP-FC. A partição associada no vetor escore e na matriz de informação observada são dadas por

$$U(\boldsymbol{\vartheta}) = \begin{bmatrix} U_1(\boldsymbol{\vartheta}) \\ U_2(\boldsymbol{\vartheta}) \end{bmatrix} \text{ e } J(\boldsymbol{\vartheta}) = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}. \quad (3.23)$$

Nosso objetivo é testar $H_0 : \boldsymbol{\vartheta}_2 = \boldsymbol{\vartheta}_{02}$ contra $H_1 : \boldsymbol{\vartheta}_2 \neq \boldsymbol{\vartheta}_{02}$, em que $\boldsymbol{\vartheta}_1$ não é especificado. Seja $\tilde{\boldsymbol{\vartheta}}$ o vetor que maximiza $\ell(\boldsymbol{\vartheta})$ restrita a H_0 , a estatística da razão verossimilhança é dada por

$$S_{LR} = 2 \left(\ell(\hat{\boldsymbol{\vartheta}}) - \ell(\tilde{\boldsymbol{\vartheta}}) \right).$$

As estatística gradiente, escore e Wald são dadas, respectivamente, por

$$S_G = U_2(\tilde{\boldsymbol{\vartheta}})^\top (\hat{\boldsymbol{\vartheta}}_2 - \boldsymbol{\vartheta}_{20}), \quad (3.24)$$

$$S_R = U_2(\tilde{\boldsymbol{\vartheta}})^\top \tilde{J}^{22}(\boldsymbol{\vartheta}) U_2(\tilde{\boldsymbol{\vartheta}}) \quad (3.25)$$

e

$$S_W = (\hat{\boldsymbol{\vartheta}}_2 - \boldsymbol{\vartheta}_{20})^\top \hat{J}^{22}(\boldsymbol{\vartheta})^{-1} (\hat{\boldsymbol{\vartheta}}_2 - \boldsymbol{\vartheta}_{20})., \quad (3.26)$$

em que

$$J^{22} = (J_{22} - J_{21} J_{11}^{-1} J_{12})^{-1}.$$

As quatro estatísticas sob H_0 e sob algumas condições de regularidade, converge em distribuição para uma distribuição Qui-quadrado com $\dim(\boldsymbol{\vartheta}_1)$ graus de liberdade.

Em especial temos interesse em desenvolver procedimentos para testar sobre o parâmetro zero determinístico do modelo ZISP-FC dada em (3.6), ou seja, testar a hipótese nula, $H_0 : \phi = \phi_0$ contra $H_1 : \phi \neq \phi_0$, em que ϕ_0 é uma quantidade especificada para ϕ . Note que, quando $\phi_0 = 0$, as hipóteses permitem avaliar a adequabilidade do modelo de sobrevivência proposta por Cancho *et al.* (2013).

A estatística gradiente para testar a hipótese H_0 é dada por

$$S_G = \left\{ \sum_{i=1}^n \frac{(1 - \delta_i) A(\tilde{\theta}_i)}{\phi_0 A(\tilde{\theta}_i) + A(\tilde{\theta}_i S_0(t_i; \tilde{\gamma}))} - \frac{n}{1 + \phi_0} \right\} (\hat{\phi} - \phi_0). \quad (3.27)$$

sob H_0 a estatística dada em (3.27) e em certas condições de regularidade tem distribuição Qui-quadrado com *um* grau de liberdade para amostras suficientemente grandes.

Alternativamente, os diferentes modelos podem ser comparados, penalizando os ajustes usando os critérios AIC e SBC como descritos no Capítulo 2.

3.2.3 Método Bootstrap

Outra técnica de inferência sobre os parâmetros do modelo ZISP-FC é o uso do método Bootstrap (Efron 1979). Esse método possui vantagens por não utilizar resultados assintóticos para a distribuição dos estimadores $\boldsymbol{\vartheta}$ e fornece uma outra forma de construção dos intervalos de confiança baseado em estimativas *bootstrap* quando trata-se de uma amostra "pequena", ou quando for crítica a suposição de normalidade. Uma adaptação desta técnica no modelo ZISP-FC seguem as sugestões de Efron (1981) e Efron & Tibshirani (1986) para dados de sobrevivência.

Os intervalos de confiança *bootstrap* não paramétrico são obtidos ao replicar B amostras com reposição de tamanho n , $(t^{*1}, \delta^{*1}, x^{*1}), \dots, (t^{*B}, \delta^{*B}, x^{*B})$ de uma amostra original (t, δ, x) . Para cada reamostra $(t^{*b}, \delta^{*b}, x^{*b})$, $b = 1, \dots, B$ são calculados os estimadores de máxima verossimilhança dos parâmetros. Os intervalos com nível $100 \times (1 - \alpha)\%$ de confiança para cada um dos parâmetros são obtidos ao calcular os quantis $(1 - \alpha/2)$ e $(\alpha/2)$ dos respectivos B estimadores de máxima verossimilhança.

3.3 Estudo de simulação

Os resultados inferenciais obtidos na Seção 3.2.2 são assintóticos. Sendo assim, desenvolvemos um estudo de simulação com diferentes tamanhos amostrais ($n = 50, 100, 200, 400$ e 600) a fim de avaliar as propriedades dos estimadores de máxima verossimilhança dos parâmetros do modelo ZIP-FC com distribuição basal Weibull, cuja função de risco é dada por $h_0(t) = \gamma_1 \gamma_2 t^{\gamma_1 - 1}$. Em nosso estudo de simulação consideramos que $\gamma_1 = 2$, $\gamma_2 = 1$, $\phi = 0,5; 2,0$ e o parâmetro θ é relacionado com a covariável $\mathbf{x}$ por meio da ligação logarítmica como em (3.16), isto é, $\log(\theta_i) = \beta_0 + \beta_1 x_i$, $i = 1, \dots, n$. A covariável x_i é gerado da distribuição Bernoulli com probabilidade de sucesso igual a $0,5$ e $\beta_0 = 2$ e $\beta_1 = -2$. As observações censuradas C_i foram gerados da distribuição Uniforme, $U(0, \tau)$, em que τ controla a porcentagem de censura. Para cada parâmetro $(\phi, \gamma_1, \gamma_2, \beta_0, \beta_1)$ e para cada tamanho amostral n , consideramos 65% de observações censuradas.

Realizamos 1000 simulações em cada tamanho amostral e foram obtidas as EMV dos parâmetros do modelo de sobrevivência ZIP-FC; Na Tabela 3.1 apresentamos a média das EMV, o desvio padrão, o viés e a raiz quadrada do erro quadrático médio (REQM) das EMVs nas diversas simulações. Observa-se que o viés e o REQM tendem para zero a medida que o tamanho da amostra aumenta em consequência as EMV são próximas dos verdadeiros valores. A Figura 3.4 mostra a distribuição empírica dos EMV, para $n = 600$, onde observamos que a distribuição dos EMV são aproximadamente simétricas ao redor da média das EMVs (linha vertical em vermelho), o que indica que a normalidade assintótica dos EMV está satisfeita. Mas esta propriedade não é verificada para amostras menores ou iguais que $n = 100$, como mostra a Figura 3.5.

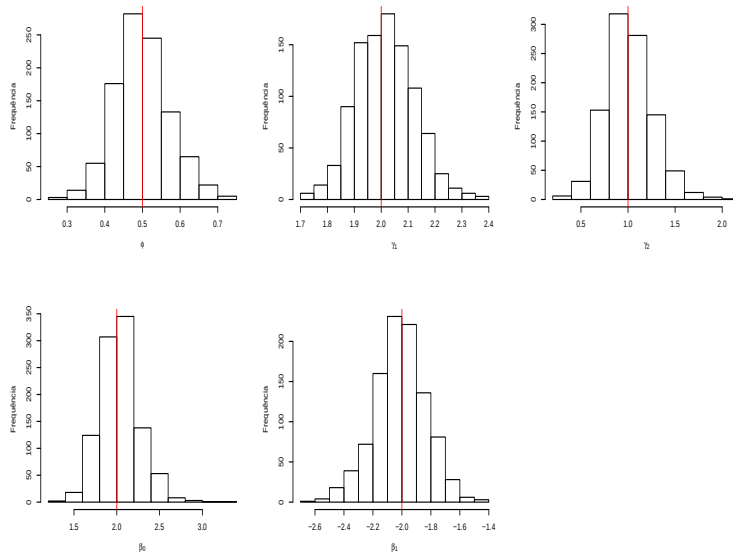


Figura 3.4: Histograma das estimativas de máxima verossimilhança dos parâmetros do modelo ZIP-FC para $n = 600$.

Tabela 3.1: Média das EMV, viés e raiz quadrada do erro quadrático médio (REQM) dos parâmetros do modelo de sobrevivência ZIP-FC.

n		$\phi = 0,5$					$\phi = 2$				
		ϕ	γ_1	γ_2	β_0	β_1	ϕ	γ_1	γ_2	β_0	β_1
50	Média	0,52	2,23	1,39	2,69	-2,15	1,95	2,38	1,78	3,08	-2,13
	DP	0,32	0,45	1,50	1,68	0,72	1,13	0,70	2,18	2,28	1,08
	Viés	0,02	0,23	0,39	0,69	-0,15	-0,05	0,38	0,78	1,08	-0,14
	REQM	0,32	0,51	1,55	1,81	0,73	1,13	0,80	2,31	2,52	1,08
	PC	0,93	0,96	0,71	0,99	0,94	0,94	0,95	0,95	0,97	0,94
100	Média	0,50	2,10	1,18	2,31	-2,05	2,03	2,19	1,30	2,78	-2,14
	DP	0,21	0,29	0,84	1,11	0,46	0,84	0,47	1,40	1,84	0,76
	Viés	0,00	0,10	0,18	0,31	-0,05	0,03	0,19	0,30	0,78	-0,14
	REQM	0,21	0,31	0,86	1,15	0,47	0,84	0,51	1,43	2,00	0,77
	PC	0,93	0,94	0,92	0,99	0,95	0,94	0,95	0,79	1,00	0,94
200	Média	0,50	2,06	1,10	2,11	-2,04	2,05	2,09	1,13	2,44	-2,06
	DP	0,13	0,20	0,52	0,62	0,34	0,55	0,31	0,91	1,26	0,51
	Viés	0,00	0,06	0,10	0,11	-0,04	0,05	0,09	0,13	0,44	-0,06
	REQM	0,13	0,21	0,53	0,63	0,34	0,55	0,32	0,92	1,33	0,52
	PC	0,93	0,95	0,96	0,97	0,94	0,95	0,94	0,88	1,00	0,95
300	Média	0,50	2,03	1,05	2,06	-2,02	2,02	2,05	1,08	2,26	-2,03
	DP	0,11	0,17	0,39	0,40	0,26	0,44	0,24	0,65	0,94	0,40
	Viés	0,00	0,03	0,05	0,06	-0,02	0,02	0,05	0,08	0,26	-0,03
	REQM	0,11	0,17	0,40	0,40	0,26	0,44	0,25	0,66	0,97	0,40
	PC	0,94	0,94	0,96	0,96	0,94	0,93	0,94	0,92	0,98	0,94
400	Média	0,50	2,02	1,02	2,06	-2,02	2,02	2,04	1,05	2,21	-2,02
	DP	0,09	0,14	0,33	0,31	0,22	0,34	0,21	0,55	0,76	0,35
	Viés	0,00	0,02	0,02	0,06	-0,02	0,02	0,04	0,05	0,20	-0,02
	REQM	0,09	0,14	0,33	0,32	0,22	0,34	0,21	0,55	0,78	0,35
	PC	0,94	0,96	0,97	0,97	0,96	0,95	0,94	0,93	0,97	0,94
600	Média	0,50	2,02	1,01	2,04	-2,01	2,01	2,03	1,02	2,12	-2,02
	DP	0,07	0,11	0,25	0,23	0,18	0,29	0,16	0,41	0,53	0,28
	Viés	-0,00	0,02	0,01	0,04	-0,01	0,01	0,03	0,02	0,12	-0,02
	REQM	0,07	0,11	0,25	0,24	0,18	0,29	0,16	0,41	0,54	0,28
	PC	0,94	0,96	0,97	0,97	0,95	0,94	0,95	0,96	0,97	0,95

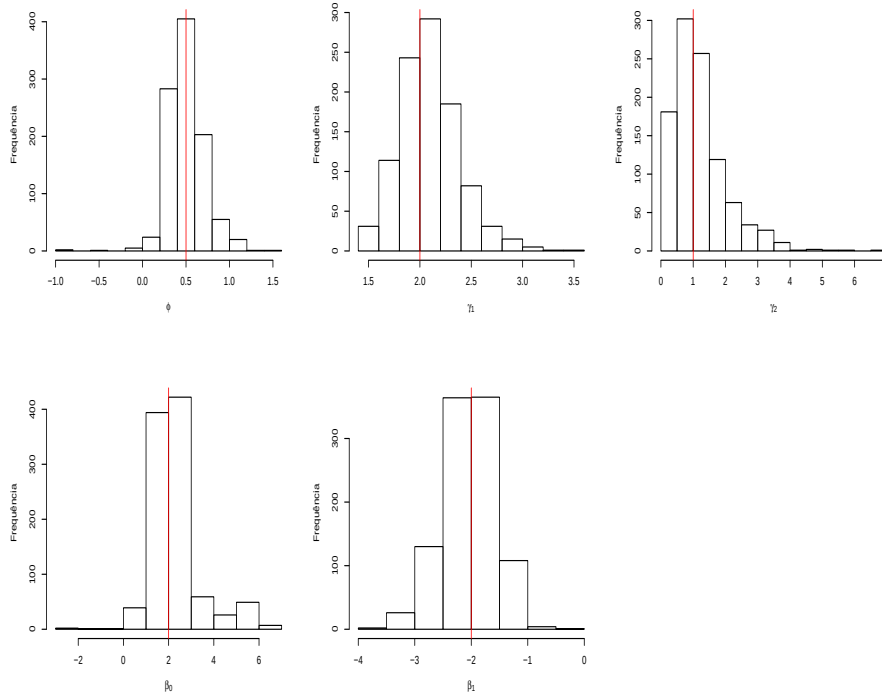


Figura 3.5: Histograma das estimativas de máxima verossimilhança dos parâmetros do modelo ZIP-FC para $n = 100$.

3.4 Aplicação

A metodologia apresentada neste capítulo será aplicada ao conjunto de dados coletados em um estudo sobre melanoma com o objetivo de avaliar o desempenho da aplicação de uma dosagem alta de interferon alfa-2b como forma de prevenir recorrência do câncer. Os pacientes foram incluídos no estudo entre 1991 e 1995, sendo acompanhados até 1998. Uma descrição mais detalhada dos dados pode ser vista em [Kirkwood *et al.* \(2000\)](#) e [Ibrahim *et al.* \(2001\)](#). A variável resposta (T) representa o tempo até a morte do paciente ou o tempo de censura. Da amostra original foram removidos 10 pacientes para os quais há valores faltantes em uma das covariáveis (x_3), restando $n = 417$ pacientes, com 56% de observações censuradas. As variáveis incluem t : tempo (em anos; média = 3,18 e desvio padrão = 1,69); x_1 : tipo de tratamento (0: sem tratamento, $n = 204$; 1: interferon, $n = 213$); x_2 : idade (em anos; média = 48,0 e desvio padrão = 13,1); x_3 : categoria do nódulo (1, $n = 82$; 2, $n = 87$; 3, $n = 137$; 4, $n = 111$); x_4 : sexo (0: masculino, $n = 263$; 1: feminino, $n = 154$); x_5 : capacidade funcional (0: ativo, $n = 363$; 1: outras, $n = 54$) e x_6 : espessura do tumor (em mm, média = 3,94 e desvio padrão = 3,20).

Na Figura 3.6(a) apresentamos a estimativa de Kaplan-Meier da função de sobrevivência para os dados descritos anteriormente, onde observamos a existência de uma aprecia-

vel fração de indivíduos com risco zero, pelo menos no que concerne ao intervalo de tempo abrangido pelo estudo. Na Figura 3.6(b) temos o gráfico TTT (Sun & Kececloglu, 1999), onde constatamos que a função de risco é crescente, sugerindo que o modelo Weibull para a distribuição basal a qual é uma alternativa adequada de modelagem destes dados. A distribuição de Weibull considerada tem função de sobrevivência dada por $S_0(t; \gamma) = \exp(-\gamma_2 t^{\gamma_1})$ e função densidade $f_0(t; \gamma) = \gamma_1 \gamma_2 t^{\gamma_1-1} \exp(-\gamma_2 t^{\gamma_1})$, para $t > 0$, $\gamma_1, \gamma_2 > 0$.

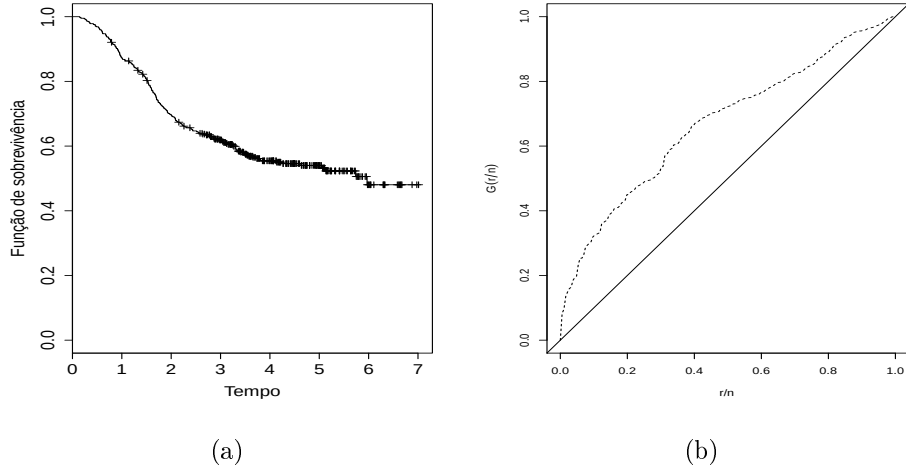


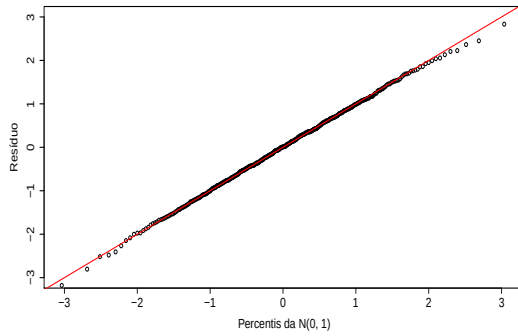
Figura 3.6: (a) Estimativa de Kaplan-Meier da função de sobrevivência e (b) gráfico TTT.

Os modelos de sobrevivência zero-inflacionado Poisson e zero-inflacionado Geométrica com fração de cura (ZIP-FC e ZIG-FC) com distribuição basal Weibull considerada para os tempos de vida, foram ajustados aos dados em que a relação entre o parâmetro θ do modelo e as covariáveis foi feita pela ligação logarítmica e logística, respectivamente, como em (3.16) com $\mathbf{x} = (1, x_1, \dots, x_6)^\top$ e $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_6)^\top$. Para obter as estimativas de máxima verossimilhança dos parâmetros em cada modelo, utilizamos o método de otimização BFGS (Lange, 1999) (uma função da linguagem R), cujos resultados estão apresentados na Tabela 3.2. Observamos que as covariáveis são não-significativas com exceção da covariável categoria do nódulo (x_3) e da covariável idade (x_2).

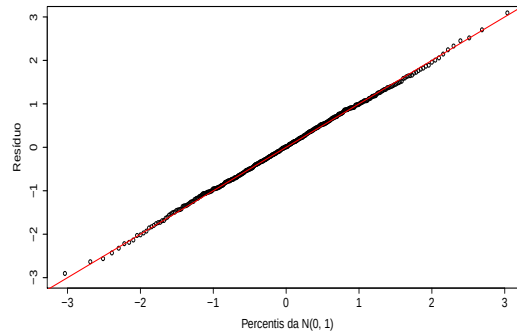
A Figura 3.7 apresenta o gráfico dos resíduos *versus* os valores esperados das estatísticas de ordem da distribuição Normal padrão, o qual indica o afastamento da suposição da normalidade para os resíduos nos modelos ZIP-FC Weibull e ZIG-FC Weibull. Observe que ambos os modelos podem ser adequados ao ajuste deste conjunto de dados.

Tabela 3.2: *EMVs com ajuste dos modelos de sobrevivência ZIP-FC Weibull e ZIG-FC Weibull para os dados de melanoma.*

Parâmetro	Modelo ZIP-FC Weibull			Modelo ZIG-FC Weibull		
	Estimativas	DP	Intervalo de confiança	Estimativas	DP	Intervalo de confiança
			assintótico 95%			assintótico 95%
ϕ	0,4609	0,1303	(0,2466; 0,6752)	0,3523	0,1660	(0,0788; 0,6251)
γ_1	1,8490	0,1263	(1,6014; 2,0965)	2,0830	0,1624	(1,7646; 2,4014)
γ_2	0,1259	0,0258	(0,0753; 0,1764)	0,0700	0,0267	(0,0176; 0,1224)
β_0	-1,7452	0,5149	(-2,7543; -0,7360)	-1,6157	0,7095	(-3,0063; -0,2251)
β_1	-0,0567	0,1922	(-0,4334; 0,3200)	-0,1733	0,2604	(-0,6836; 0,3370)
β_2	0,0147	0,0074	(0,0002; 0,0292)	0,0166	0,0090	(-0,0010; 0,0343)
β_3	0,5231	0,0855	(0,3555; 0,6907)	0,6779	0,1227	(0,4374; 0,9184)
β_4	-0,1268	0,1973	(-0,5134; 0,2599)	-0,1007	0,2586	(-0,6076; 0,4061)
β_5	0,2401	0,2711	(-0,2913; 0,7714)	0,2377	0,3541	(-0,4562; 0,9317)
β_6	0,0447	0,0296	(-0,0133; 0,1026)	0,0406	0,0369	(-0,0317; 0,1128)
$\max \ell(\cdot)$	-506,2046			$\max \ell(\cdot)$	-504,8615	
AIC	1032,409			AIC	1029,723	
SBC	1072,740			SBC	1070,054	



(a)



(b)

Figura 3.7: (a) *Estimativa dos resíduos para o modelo ZIP-FC Weibull.* (b) *Estimativa dos resíduos para o modelo ZIG-FC Weibull.*

Considerando os critérios estabelecidos na Tabela 3.2 selecionamos o modelo ZI-Geométrica como um modelo apropriado para estes dados. Prosseguimos com um modelo simplificado com apenas as covariáveis categoria do nódulo e idade, seguindo procedimentos assintóticos e Bootstrap cujos resultados compõem a Tabela 3.3. Um procedimento não paramétrico *bootstrap* foi realizado considerando $B = 5000$ réplicas obtendo assim os estimadores e intervalos *bootstrap* com 95% de confiança. É possível observar nesses resultados que os intervalos assintóticos e *bootstrap* não apresentam grandes diferenças e, em todas as

situações, esses intervalos contém o verdadeiro valor dos parâmetros.

Confrontando os modelos das Tabelas 3.2 e 3.3, de acordo com a estatística da razão de verossimilhança (S_{LR}), o modelo simplificado propicia um ajuste tão bom quanto o do modelo da Tabela 3.2 ($S_{LR} = 2,747$ com 6 graus de liberdade, nível descritivo = 0,8399). Portanto, adotamos o modelo da Tabela 3.3 e a Figura 3.8 mostra o gráfico de resíduos para o modelo simplificado.

Tabela 3.3: *EMVs assintóticas e bootstrap com ajuste do modelo de sobrevivência ZIG-FC Weibull simplificado.*

Parâmetro	Estimativas	Intervalo de confiança assintótico 95%	Estimativas	Intervalo de confiança bootstrap 95%
ϕ	0,3493	(0,0225; 0,6760)	0,3663	(0,0599; 0,6778)
γ_1	2,0668	(1,7518; 2,3818)	2,0879	(1,8004; 2,3999)
γ_2	0,0727	(0,0187; 0,1266)	0,0677	(0,0137; 0,1218)
β_0	-1,5506	(-2,7619; -0,3393)	-1,2633	(-2,3856; -0,1701)
β_2	0,0182	(0,0006; 0,0358)	0,0171	(0,0028; 0,0307)
β_3	0,6373	(0,4132; 0,8614)	0,6380	(0,4746; 0,8125)

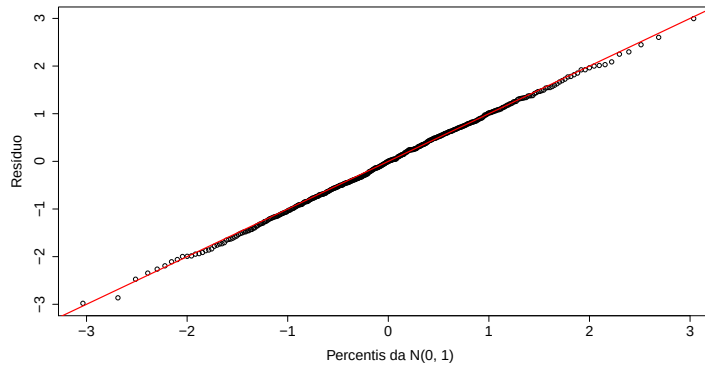


Figura 3.8: *Estimativa de resíduos para o modelo simplificado ZIG-FC Weibull.*

A estimativa de máxima verossimilhança (EMV) de proporção de risco zero determinístico, $\omega = \phi/(1 + \phi)$, é 0,259 (0,092). Esta proporção indica que 25,9% dos indivíduos possuem fatores genéticos próprios que permitiram estar livres da doença. A quantidade em parênteses é o erro padrão da EMV, obtida após aplicar o método delta (Cox & Hinkley, 1979).

Os testes de hipóteses de interesse dada na Seção 3.2.2, são equivalentes às seguintes hipóteses: $H_0 : \omega = 0$ contra $H_1 : \omega \neq 0$. Essas hipóteses permitem avaliar o modelo de

sobrevivência tempo de promoção considerado por [Yakovlev *et al.* \(1993\)](#) contra o modelo ZIG-FC. Na Tabela 3.4 são apresentados os resultados das estatísticas de teste. Pode-se observar que as quatro estatísticas, ao nível de significância de 5%, indicam que o modelo de tempo de promoção não é adequado para o ajuste dos dados de melanoma.

Tabela 3.4: *Estimativas das estatísticas e nível descritivo para o teste de hipótese $H_0 : \omega = 0$ nos dados de melanoma.*

	S_G	S_R	S_{LR}	S_W
Estatística	3,3222	3,8656	8,8907	4,4442
Nível descritivo	0,0683	0,0493	0,0029	0,0350

A Figura 3.9 mostra a função de sobrevivência estratificada por nível categoria do nódulo em pacientes com idade igual a 32 e 66 anos, que correspondem aos percentis 10 e 90. Estes gráficos destacam o impacto combinado das covariáveis sobre a fração de curados. Observamos que os tempos de sobrevivência de pacientes com melanoma diminui com a idade e o nível de categoria do nódulo.

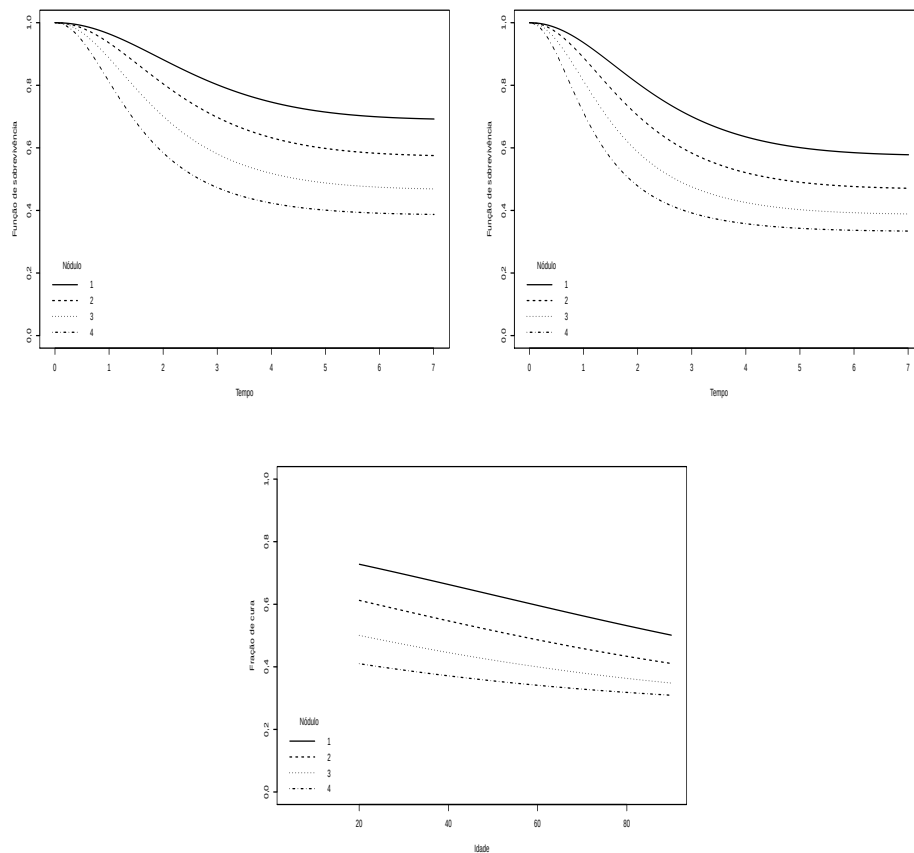


Figura 3.9: *Função de sobrevivência sob modelo ZIP-FC estratificada por categoria de nódulo para pacientes com idade igual a 32 anos (painel esquerdo) e 66 anos (painel direito) e, fração cura (painel central).*

Finalmente, a Tabela 3.5 apresenta o EMV da fração de cura para os gráficos da Figura 3.9. Os intervalos de confiança aproximado de 95% de confiança foram obtidos após a aplicação do método delta. Note que a fração de cura entre os pacientes com idade de 32 e 66 anos em cada categoria de nódulo, os intervalos não se sobrepõem. A Figura 3.9 mostra a fração de cura como função da idade, onde observamos que a fração de cura de pacientes com melanoma diminui com a idade.

Tabela 3.5: *Estimativa de máxima verossimilhança para fração de curados estratificada por categoria de nódulo e idade dos pacientes.*

Idade	Categoria de		Erro	Intervalo de confiança
	Nódulo	Estimativa	Padrão	assintótico 95%
32	1	0,6891	0,0564	(0,5786; 0,7997)
	2	0,5723	0,0479	(0,4783; 0,6663)
	3	0,4661	0,0395	(0,3886; 0,5436)
	4	0,3854	0,0443	(0,2986; 0,4722)
66	1	0,5749	0,0544	(0,4682; 0,6816)
	2	0,4683	0,0411	(0,3877; 0,5488)
	3	0,3869	0,0434	(0,3018; 0,4720)
	4	0,3329	0,0569	(0,2214; 0,4444)

3.5 Conclusões

Neste Capítulo, propusemos o modelo de sobrevivência zero-inflacionado Série de Potência com fração de cura. O modelo é flexível e inclui como casos particulares os modelos de sobrevivência propostos por [Cancho *et al.* \(2013\)](#) e [Caroni *et al.* \(2010\)](#), e o modelo de tempo de promoção proposto por [Yakovlev *et al.* \(1993\)](#). O modelo permite estimar a proporção de indivíduos sem risco, ou seja, a proporção de indivíduos imunes e curados; e avaliar de melhor forma a eficiência de um tratamento para o paciente. Procedimentos inferenciais foram desenvolvidos por meio de uma perspectiva clássica baseada no método de máxima verossimilhança e um estudo de simulação indicou que a normalidade assintótica dos EMV estão satisfeitas para tamanhos amostrais maiores que 100. Um conjunto de dados de tempos de sobrevivência da literatura foi ajustado ao modelo proposto mostrando o potencial da modelagem para esta análise de dados de sobrevivência. Aplicamos também o método bootstrap nesse conjunto de dados e observou-se que as estimativas dos parâmetros comparado com os resultados assintóticos são similares.

Capítulo 4

Modelo de sobrevivência induzida por Fragilidade Discreta: Enfoque bayesiano

No Capítulo 3 foram desenvolvidos procedimentos inferenciais desde a perspectiva clássica (ou frequentista) para o modelo de sobrevivência Zero Inflacionado Série de Potência com fração de cura (ZISP-FC) definido em (3.6), baseado em resultados assintóticos. Para amostras de tamanhos pequenos ou moderados esses resultados podem não ser confiáveis, tornando-se um problema sério uma vez que em geral, trabalha-se com amostras não muito grandes. Um procedimento alternativo que pode contornar esse problema é por meio da utilização de métodos Bayesianos. O uso de métodos Bayesianos, além de serem uma alternativa de análise, permite ainda a incorporação de conhecimento *a priori* por meio de uma distribuição *a priori*. Neste Capítulo apresenta-se o procedimento inferencial desde de uma perspectiva Bayesiana para o modelo ZISP-FC, usando o método de simulação via cadeias de Markov (Gelfand & Smith, 1990).

4.1 Inferência Bayesiana no modelo ZISP-FC

Consideramos novamente o modelo ZISP-FC definido em (3.6) com função de sobrevivência dada por

$$S(t) = \frac{\phi + A(\theta S_0(t))A(\theta)^{-1}}{1 + \phi}, \quad t > 0, \quad (4.1)$$

com $\theta > 0$ e $-a_0/A(\theta) \leq \phi < \infty$ e $S_0(t)$ a função de sobrevivência basal, a proporção de indivíduos com risco é dada por $p_0 = \frac{\phi}{1+\phi} + \frac{a_0}{(1+\phi)A(\theta)}$. A correspondente função densidade de probabilidade é dado por

$$f(t) = -S'(t) = \frac{\theta f_0(t)}{1 + \phi} \frac{A'(\theta S_0(t))}{A(\theta)}. \quad (4.2)$$

Como no Capítulo 3, considere os tempos de sobrevivência de n indivíduos, $T_1, \dots, T_n$ e $C_1, \dots, C_n$ os respectivos tempos de censura. Suponha que é observado $t_i = \min\{T_i, C_i\}$ e $\delta_i = I(T_i \leq C_i)$, em que $\delta_i = 1$ se t_i é um tempo de sobrevivência e $\delta_i = 0$ se é o tempo de censura a direita. Seja $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})$ denotando o vetor de covariáveis par o i -ésimo indivíduo. Assumindo que o parâmetro θ do modelo (3.6) se relaciona com as covariáveis por meio de

$$g(\theta_i) = \eta_i = \eta(\mathbf{x}_i, \boldsymbol{\beta}) = \mathbf{x}_i \boldsymbol{\beta}, \quad i = 1, \dots, n. \quad (4.3)$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ é o vetor de parâmetros de regressão ($\boldsymbol{\beta} \in R^p$) associado ao vetor de covariável $\mathbf{x}_i$. Aqui $\eta(\cdot, \cdot)$ é uma função contínua duas vezes diferenciável. Finalmente, $g(\cdot)$ é a função de ligação contínua monótona e duas vezes diferenciável com domínio R^+ . Uma possível função de ligação para o parâmetro θ_i do modelo ZIP θ_i é $g(\theta_i) = \log(\theta_i)$ (ligação logarítmica) e para o parâmetro θ_i dos modelos ZIG e ZIL é $g(\theta_i) = \log(\theta_i/(1 - \theta_i))$ (a ligação logística). Com a expressão (4.3), o logaritmo da função de verossimilhança de $\boldsymbol{\vartheta} = (\phi, \boldsymbol{\gamma}, \boldsymbol{\beta}^\top)$ sob censura não informativa pode ser escrito como

$$L(\boldsymbol{\vartheta}|\mathcal{D}) = \frac{1}{(1 + \phi)^n} \prod_{i=1}^n \frac{1}{A(\theta_i)} (\theta_i f_0(t_i; \gamma) A'(\theta_i S_0(t_i; \gamma)))^{\delta_i} (\phi A(\theta_i) + A(\theta_i S_0(t)))^{1-\delta_i} \quad (4.4)$$

sendo $\mathcal{D} = (\mathbf{t}, \boldsymbol{\delta}, \mathbf{x})$ com $\mathbf{t} = (t_1, \dots, t_n)^\top$, $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)^\top$ e $\mathbf{x} = (x_1, \dots, x_n)$.

Uma parte fundamental da análise Bayesiana é especificar distribuições *a priori* para todos os parâmetros desconhecidos do modelo. A fim de garantir distribuições *a posteriori* próprias, adotamos distribuições *a priori* próprias para todas as quantidade desconhecidas do modelo. Assim consideramos as seguintes distribuições *a priori* para os parâmetros do modelo ZISP-FC

- $\phi|\boldsymbol{\beta} \sim U(\max\{-1, -a_0/A(\theta_i)\}, \rho)$, com $\rho > 0$ conhecido, suficientemente grande,
- $\beta_j \sim N(\mu_j, \sigma_j^2)$, μ_j e σ_j^2 são conhecidos, com $j = 1, \dots, p$.
- $\gamma_k \sim \text{Gama}(a_k, b_k)$, $k = 1, 2$ com a_k e b_k conhecidos,

em que $N(\mu, \sigma^2)$ denota a distribuição Normal com média μ e variância σ^2 , $\text{Gama}(a, b)$ denota a distribuição Gama com média a/b e variância a/b^2 e $U(c, d)$ denota a distribuição uniforme no intervalo a e b .

A distribuição *a priori* conjunta dos parâmetros $\boldsymbol{\beta}$, ϕ e $\boldsymbol{\gamma}$ do modelo ZISP-FC é dada por

$$\pi(\boldsymbol{\vartheta}) = \pi(\phi|\boldsymbol{\beta})\pi(\boldsymbol{\beta})\pi(\boldsymbol{\gamma}) \propto \frac{\gamma_1^{a_1-1} \gamma_2^{a_2-1} \exp\left\{-\sum_{i=1}^2 b_i \gamma_i - \sum_{j=1}^p \frac{(\beta_j - \mu_j)^2}{2\sigma_j^2}\right\}}{\rho - \max\{-1, -a_0/A(\theta_i)\}}, \quad (4.5)$$

nota-se que no caso em que $a_i = b_i = 0$ ($i = 1, 2$) e $\sigma_j \rightarrow \infty$ ($j = 1, \dots, p$) a função de

densidade *a priori* conjunta é imprópria, ou seja

$$\pi(\boldsymbol{\vartheta}) \propto \frac{1}{\gamma_1 \gamma_2 (\rho - \max\{-1, -a_0/A(\theta_i)\})}.$$

Combinando a função verossimilhança (4.4) e a função de densidade *a priori* conjunta (4.5) e o Teorema de Bayes pode-se mostrar que a densidade conjunta *a posteriori* de $\boldsymbol{\vartheta} = (\phi, \boldsymbol{\gamma}, \boldsymbol{\beta}^\top)$ é dada por

$$\pi(\boldsymbol{\vartheta}|\mathcal{D}) \propto \frac{1}{(1+\phi)^n} \prod_{i=1}^n \frac{1}{A(\theta_i)} (\theta_i f_0(t_i; \gamma) A'(\theta_i S_0(t_i; \gamma)))^{\delta_i} (\phi A(\theta_i) + A(\theta_i S_0(t)))^{1-\delta_i} \pi(\boldsymbol{\vartheta}), \quad (4.6)$$

em que $\mathcal{D}$ é o conjunto de dados observados.

Notamos que a densidade *a posteriori* conjunta (4.6) não é uma densidade padrão, portanto somente podemos avaliar as densidades *a posteriori* marginais por meio de métodos de aproximação, tais como o método de Laplace (Kass *et al.*, 1991) ou usando métodos de simulação de Monte Carlo via Cadeia de Markov, tais como o amostrador de Gibbs (Casella & George, 1992) e Metropolis-Hasting (Chib & Greenberg, 1995). Nesta análise consideraremos os métodos MCMC por serem de fácil implementação computacional.

Para obtermos uma amostra *a posteriori* de $\phi, \boldsymbol{\gamma}$ e $\boldsymbol{\beta}$, fazemos uso do algoritmo de Gibbs que se baseia em sucessivas gerações das distribuições condicionais de $\pi(\phi|\boldsymbol{\beta}, \boldsymbol{\gamma}, \mathcal{D})$, $\pi(\boldsymbol{\gamma}|\boldsymbol{\beta}, \phi, \mathcal{D})$ e $\pi(\boldsymbol{\beta}|\boldsymbol{\gamma}, \phi, \mathcal{D})$.

Da densidade *a posteriori* conjunta dada em (4.6) pode-se demonstrar que as densidades condicionais *a posteriori* completas para o algoritmo de Gibbs são dadas por

$$\begin{aligned} \pi(\phi|\boldsymbol{\beta}, \boldsymbol{\gamma}, \mathcal{D}) &\propto \frac{1}{(1+\phi)^n} \prod_{i=1}^n [\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma))]^{1-\delta_i} \pi(\phi|\boldsymbol{\beta}), \\ \pi(\boldsymbol{\beta}|\boldsymbol{\gamma}, \phi, \mathcal{D}) &\propto \prod_{i=1}^n \frac{1}{A(\theta_i)} [\theta_i A'(\theta_i S_0(t_i; \gamma))]^{\delta_i} [\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma))]^{1-\delta_i} \pi(\boldsymbol{\beta}), \\ \pi(\boldsymbol{\gamma}|\boldsymbol{\beta}, \phi, \mathcal{D}) &\propto \prod_{i=1}^n [f_0(t_i; \gamma) A'(\theta_i S_0(t_i; \gamma))]^{\delta_i} [\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma))]^{1-\delta_i} \pi(\boldsymbol{\gamma}) \end{aligned} \quad (4.7)$$

Observe que as densidades dadas em (4.7) não são conhecidas de modo que faremos uso do algoritmo de Metropolis-Hasting para gerar $\phi, \boldsymbol{\gamma}$ e $\boldsymbol{\beta}$ das densidades condicionais *a posteriori* completas.

Uma alternativa Bayesiana para testar as hipóteses dadas Seção 3.2.2 é por meio da determinação do fator de Bayes do modelo sob a hipótese nula (modelo M_0) com respeito a hipótese alternativa (modelo M_1). Como sugerido por Geisser & Eddy (1979) neste trabalho utiliza-se o pseudo-fator de Bayes para testar essas hipóteses considerando as verossimi-

lanças preditivas marginais do modelo M_0 e M_1 e essas verossimilhanças são obtidas como produto das densidades preditivas condicionais ordenadas (CPO) definidas por

$$f(y_r|\mathcal{D}^{(-r)}) = \int_{\Theta} f(y_r|\boldsymbol{\vartheta})\pi(\boldsymbol{\vartheta}|\mathcal{D}^{(-r)})d\boldsymbol{\vartheta}, \quad r = 1, \dots, n, \quad (4.8)$$

em que $\mathcal{D}^{(-r)}$ é o conjunto de observações com r -ésima observação excluída. O pseudo-fator de Bayes do modelo M_0 com respeito ao modelo M_1 é definido por

$$\text{PSFB}_{01} = \frac{\prod_{r=1}^n f(y_r|\mathcal{D}^{(-r)}, M_0)}{\prod_{r=1}^n f(y_r|\mathcal{D}^{(-r)}, M_1)} \quad (4.9)$$

em que $f(y_r|\mathcal{D}^{(-r)}, M_j)$, $j=0,1$ é o CPO_r para r -ésima observação do modelo sob H_j , $j = 0, 1$. Rejeita-se a hipótese nula H_0 , se $\text{PSFB}_{01} < 1$, portanto, o melhor modelo é o modelo M_1 .

Por exemplo, tem-se interesse em testar as hipóteses $H_0 : \phi = 0$ vs $H_1 : \phi \neq 0$, ou seja, deseja-se comparar o modelo ZISP-FC e o modelo Série de Potência com fração de cura (SP-FC) proposta por [Cancho et al. \(2013\)](#). Sob as hipótese H_0 (ou H_1), não é possível obter uma expressão analítica para CPO . Uma estimativa do CPO pode ser obtida por integração de Monte Carlo, com amostras *a posteriori*. Seja $\boldsymbol{\vartheta}^{(i)}, \dots, \boldsymbol{\vartheta}^{(S)}$, é uma amostra *a posteriori* dos parâmetros do modelo sob H_0 , uma estimativa do CPO para o modelo SP-FC é dado por

$$\hat{f}(y_r|\mathcal{D}^{(-r)}, M_0) = \begin{cases} S \sum_{i=1}^S \left(\theta^{(i)} f_0(t_r) \frac{A'(\theta^{(i)} S_0(t_r))}{A(\theta^{(i)})} \right)^{-1}, & \text{se } \delta_r = 1, \\ S \sum_{i=1}^S \left(\frac{A(\theta^{(i)} S_0(t_r))}{A(\theta^{(i)})} \right)^{-1}, & \text{se } \delta_r = 0 \end{cases}$$

4.1.1 Análise de diagnóstico caso a caso

Após a modelagem é importante verificar as suposições feitas para o modelo conduzido, para um estudo da robustez com o intuito de encontrar possíveis observações influentes, as quais podem causar distorções nos resultados da análise. Uma maneira comum de avaliar a influência de uma observação no ajuste de um modelo é por meio da deleção de casos [Cook & Weisberg \(1982\)](#), em que o efeito de remover cada caso completamente da análise é estudado. Baseada no mesmo raciocínio, a metodologia da análise de influência caso a caso com um abordagem Bayesiana possibilita determinar quais casos podem influenciar nos resultados da análise.

Seja $D_\psi(P, P_{(-i)})$ a medida de divergência ψ entre P e $P_{(-i)}$, em que P é a distribuição *a posteriori* de $\boldsymbol{\vartheta}$ dado dados completos, e $P_{(-i)}$ é a distribuição *a posteriori* de $\boldsymbol{\vartheta}$ sem o i -ésimo caso. Especificamente,

$$D_\psi(P, P_{(-i)}) = \int \pi(\boldsymbol{\vartheta}; \mathcal{D}) \psi\left(\frac{\pi(\boldsymbol{\vartheta}|\mathcal{D})}{\pi(\boldsymbol{\vartheta}|\mathcal{D}^{(-i)})}\right) d\boldsymbol{\vartheta}, \quad (4.10)$$

$D_\psi(P, P_{(-i)})$ é uma medida do efeito da omissão do i -ésimo caso dos dados completos na distribuição *a posteriori* de $\boldsymbol{\vartheta}$. Note que em geral, $D(P, P_{(-i)}) \neq D(P_{(-i)}, P)$ e ψ é uma função convexa com, $\psi(1) = 0$; [Dey & Birmiwal \(1994\)](#) apresenta várias escolhas da função ψ como por exemplo, se $\psi(u) = -\log(u)$ obtemos a divergência Kullback-Leibler (K-L divergência); se $\psi(u) = (u - 1)\log(u)$ temos J-distância que é uma versão simétrica da divergência Kullback-Leibler; $\psi(u) = 0,5|u - 1|$ refere-se a distância variacional ou a norma L_1 e, se $\psi(u) = (u - 1)^2$, é a divergência Qui-Quadrado (χ^2 -divergência).

Existe uma relação matemática entre o *CPO* definida em (4.8) e a medida de ψ -divergência, que pode ser escrita como

$$D_\psi(P, P_{(-i)}) = E_{\boldsymbol{\vartheta}|\mathcal{D}} \left[\psi \left(\frac{CPO_i}{L(\boldsymbol{\vartheta}|\mathcal{D}_i)} \right) \right], \quad (4.11)$$

em que o valor esperado é tomado com relação à distribuição *a posteriori* $\pi(\boldsymbol{\vartheta}|\mathcal{D})$ e $L(\boldsymbol{\vartheta}|\mathcal{D}_i)$ é a função de verossimilhança da i -ésima observação. De (4.11) a K-L divergência ([Cho et al., 2009](#)), pode ser representada por

$$D_{K-L}(P, P_{(-i)}) = \log(CPO_i) + E_{\boldsymbol{\vartheta}|\mathcal{D}} \{\log L(\boldsymbol{\vartheta}|\mathcal{D}_i)\}. \quad (4.12)$$

De (4.11) uma estimativa de Monte Carlo de $D_\psi(P, P_{(-i)})$ por amostragem a partir da distribuição *a posteriori* do $\boldsymbol{\vartheta}$ pode ser obtida. Seja $\boldsymbol{\vartheta}^{(1)}, \dots, \boldsymbol{\vartheta}^{(Q)}$ uma amostra de tamanho S de $\pi(\boldsymbol{\vartheta}|\mathcal{D})$. Então a estimativa de Monte Carlo de $D_\psi(P, P_{(-i)})$ é dada por

$$D_\psi(\widehat{P}, P_{(-i)}) = -\frac{1}{S} \sum_{s=1}^S \left[\psi \left(\frac{\widehat{CPO}_i}{L(\boldsymbol{\vartheta}^{(s)}|\mathcal{D}_i)} \right) \right]. \quad (4.13)$$

[Peng & Dey \(1995\)](#) e [Weiss \(1996\)](#) ressaltaram a dificuldade em detectar o ponto de corte da medida de divergência, na detecção de observações influentes ou não. Como em [Peng & Dey \(1995\)](#) considere o experimento de uma moeda viesada com probabilidade de sucesso p e outra moeda não viesada, então a medida de ψ -divergência entre a moeda não viesada e viesada é dada por

$$D_\psi(f_0, f_1) = \int \psi \left(\frac{f_0(w)}{f_1(w)} \right) f_1(w) dw, \quad (4.14)$$

em que $f_0(w) = p^w(1-p)^{1-w}$ e $f_1(w) = 0,5$, $w = 0, 1$. No entanto, se $D_\psi(f_0, f_1) = d_\psi(p)$, então pode ser facilmente verificado que d_ψ satisfaz a seguinte função

$$d_\psi(p) = \frac{\psi(2p) + \psi(2(1-p))}{2}. \quad (4.15)$$

Note que a função $d_\psi(\cdot)$ é simétrica em torno de $p = 0,5$ e atinge o valor mínimo neste ponto, $d_\psi(0,5) = 0$ e $f_0 = f_1$. Além disso, se $p > 0,75$ (ou $p \leq 0,25$), como uma moeda fortemente enviesada, temos que $d_{L_1}(0,75) = d_{\chi^2}(0,75) = 0,25$, portanto, o i -ésimo caso

é considerado influente quando $d_{L_1}(0, 75) > 0,25$ ou $d_{\chi^2} > 0,25$. Assim, para divergência $K - L$, podemos considerar a observação influente quando $d_{K-L} > 0,14$ e, analogamente, usando o J -distancia, a observação com $d_J > 0,27$ pode ser considerada como observação influente.

4.2 Estudo de simulação

Nesta Seção é apresentado um estudo de simulação para investigar a propriedades frequentistas das estimativas Bayesianas dos parâmetros do modelo ZISP-FC e verificar o comportamento do pseudo fator de Bayes para discriminar modelos sob H_0 e sob H_1 . Como a Seção 3.3, um estudo de simulação é baseado no modelo ZIP-FA com distribuição basal Weibull, cuja função de risco dada por $h_0(t) = \gamma_1 \gamma_2 t^{\gamma_1 - 1}$. Em nosso estudo de simulação consideramos que $\gamma_1 = 2$, $\gamma_2 = 1$, $\phi = 0,5; 2,0$ e o parâmetro θ é relacionado com a covariável x por meio da ligação logarítmica como em (3.16), isto é, $\log(\theta_i) = \beta_0 + \beta_1 x_i$. $i = 1, \dots, n$. A covariável x_i é gerada da distribuição Bernoulli com probabilidade de sucesso igual a 0,5 e, $\beta_0 = 2$ e $\beta_1 = -2$. As observações censuradas C_i foram geradas da distribuição Uniforme, $U(0, \tau)$, em que τ controla a porcentagem de censura. Neste estudo consideramos aproximadamente 65% de observações censuradas. Várias combinações de tamanhos amostrais ($n = 100, 200, 400$ e 600) foram considerados no estudo. Realizamos 1000 simulações para cada tamanho de amostra e foram obtidas a média, desvio padrão (DP), viés, raiz quadrada do erro quadrático médio (REQM) e a probabilidade de cobertura do intervalo HPD de 95%, para cada parâmetro do modelo de sobrevivência ZIP-FC.

Para cada simulação foi considerado a distribuição *a priori* dada em (4.5) com $\gamma_j \sim \text{Gama}(1, 0.01)$, $j = 0, 1$, $\beta_j \sim N(0, 100)$ e $\phi | \beta \sim U(\max\{-1, -e^{\theta_i}\}, 100)$. A partir das densidades condicionais completas foram gerados duas cadeias separadas cada uma com 25000 iterações cada uma. Como as densidades condicionais não são de famílias de distribuição conhecida, utilizamos o algoritmo de Metropolis-Hasting embutidos nos ciclos do amostrados de Gibbs. Com a finalidade de diminuir o efeito dos pontos iniciais foram descartadas as primeiras 5000 iterações e para evitar o problema de autocorrelação das séries, fixamos um espaçamento de tamanho 10, conduzindo uma amostra *a posteriori* de 1000 para cada cadeia. A convergência das amostras foi monitorada por vários critérios considerados em Cowles & Carlin (1996).

Na Tabela 4.1 reportamos as médias de Monte Carlo da média *a posteriori*, desvio padrão das médias *a posteriori* (DP), raiz quadrada do erro quadrático médio (REQM) das médias, viés e a probabilidade de cobertura (PC) do intervalo HPD de 95% de confiança para cada parâmetro do modelo ZISP-FC. Observa-se que o viés e o REQM, tende para zero a medida que o tamanho da amostra aumenta em consequência as médias *a posteriori* são próximas dos verdadeiros valores. Além disso as probabilidades de cobertura são bem

aproximadas com o nível de confiança de 95 %. Destacamos que os resultados de simulação clássicos e Bayesianos são similares.

Tabela 4.1: *As médias de Monte Carlos da média a posteriori, desvio padrão das médias a posteriori (DP), raiz quadrada do erro quadrático médio (REQM) das médias, viés e a probabilidade de cobertura (PC) do intervalo HPD de 95% de confiança para cada parâmetro do modelo ZISP-FC.*

n		$\phi = 0,5$					$\phi = 2,0$				
		ϕ	γ_1	γ_2	β_0	β_1	ϕ	γ_1	γ_2	β_0	β_1
100	Média	0,55	2,09	2,25	2,21	-1,96	2,27	2,22	2,49	2,29	-2,02
	DP	0,17	0,25	1,02	0,60	0,48	0,65	0,39	0,97	0,74	0,76
	VIES	0,05	0,09	0,25	0,21	0,04	0,27	0,22	0,49	0,29	-0,02
	REQM	0,18	0,27	1,05	0,63	0,48	0,70	0,45	1,08	0,80	0,76
	PC	0,97	0,93	0,96	0,97	0,94	0,96	0,91	0,99	0,98	0,94
200	Média	0,54	2,05	2,16	2,12	-1,99	2,19	2,08	2,23	2,22	-1,96
	DP	0,12	0,17	0,78	0,40	0,32	0,47	0,26	0,96	0,58	0,53
	VIES	0,04	0,05	0,16	0,12	0,01	0,19	0,08	0,23	0,22	0,04
	REQM	0,13	0,18	0,79	0,42	0,32	0,50	0,27	0,99	0,62	0,53
	PC	0,95	0,94	0,94	0,97	0,94	0,97	0,94	0,98	0,98	0,93
400	Média	0,53	2,02	2,06	2,06	-1,99	2,13	2,05	2,03	2,22	-1,96
	DP	0,08	0,12	0,51	0,25	0,23	0,35	0,17	0,78	0,44	0,33
	VIES	0,03	0,02	0,06	0,06	0,01	0,13	0,05	0,03	0,22	0,04
	REQM	0,09	0,13	0,51	0,26	0,23	0,38	0,17	0,78	0,49	0,33
	PC	0,97	0,94	0,95	0,97	0,92	0,95	0,95	0,92	0,97	0,94
600	Média	0,50	1,97	1,96	2,00	-1,82	2,31	2,07	1,95	2,15	-2,01
	DP	0,07	0,06	0,19	0,02	0,12	0,40	0,04	0,34	0,44	0,24
	VIES	-0,00	-0,03	-0,04	-0,00	0,18	0,31	0,07	-0,05	0,15	-0,01
	REQM	0,05	0,05	0,14	0,02	0,20	0,45	0,13	0,28	0,39	0,20
	PC	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

Para avaliar o pseudo fator de Bayes consideramos testar as hipóteses $H_0 : \phi = 0$ vs $H_1 : \phi \neq 0$. O que é equivalente o modelo ZIP-FC e o modelo tempo de promoção proposta por [Chen et al. \(1999\)](#). Amostras de diferente tamanhos ($n = 100, 200, 400$ e 600) do modelo ZIP-FC gerado com os mesmos parâmetros dadas acima e $\phi = 0; 0,5$ e 2 . Mil amostras foram geradas para cada simulação. Na Tabela 4.2, apresenta-se a porcentagem de vezes que hipóteses nula é rejeitada considerando-se os pseudos fator de Bayes do modelo sob H_0 com respeito o modelo sob H_1 . Consideramos que H_0 é rejeitada quando $PSFB_{01} < 1$. Esses resultados indicam que pseudo fator de Bayes consegue discriminar o modelo de forma adequada.

Tabela 4.2: Porcentagens de vezes que $H_0 : \phi = 0$ foi rejeitada em 1000 amostras.

ϕ	n				
	100	200	300	400	600
0,0	12,1	11,2	10,9	10,2	9,5
0,5	84,4	98,8	99,0	100,0	100,0
2,0	85,2	97,0	100,0	100,0	100,0

4.3 Aplicação

A metodologia apresentada neste Capítulo será aplicado ao conjunto de dados de melanoma descrita em (3.4). Os modelos de sobrevivência zero-inflacionado Poisson e zero-inflacionado Geométrica com fração de cura (ZIP-FC e ZIG-FC) com distribuição basal Weibull considerada para os tempos de vida, foram ajustados aos dados em que a relação entre o parâmetro θ do modelo e as covariáveis foi feita pela ligação logarítmica e logística, respectivamente, como em (3.16) com $\mathbf{x} = (1, x_1, \dots, x_6)^\top$ e $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_6)^\top$.

Na análise Bayesiana, consideramos a distribuição *a priori* dada em (4.5) com $\beta_j \sim N(0, 100)$, $\gamma_j \sim \text{Gama}(1, 0.01)$, $j = 0, 1, \dots, 6$, $\phi \sim U(\max\{-1, -e^{\theta_i}\}, 100)$ (para modelo ZIP-FC) $\phi \sim U(\max\{-1, -(1 - \theta_i)^{-1}\}, 100)$ e as amostras *a posteriori* são obtidas de mesma forma que na Seção 4.2. Na Tabela 4.3 são apresentados os sumários *a posteriori* dos parâmetros dos Modelos ZIP-FC e ZIG-FC, em que observamos que para ambos modelos com exceção da covariável categoria de nódulo (x_3), todas as covariáveis são não significativas.

Tabela 4.3: Média, mediana e desvio padrão (DP) *a posteriori* e o intervalo HPD de 95% de confiança dos parâmetros dos Modelos ZIP-FC e ZIG-FC, para os dados de melanoma.

Parâmetro	Modelo ZIP-FC Weibull					Modelo ZIG-FC Weibull				
	Média	Mediana	DP	Intervalo HPD (95%)		Média	Mediana	DP	Intervalo HPD (95%)	
				LI	LS				LI	LS
ϕ	0,471	0,470	0,145	0,203	0,769	0,425	0,421	0,170	0,128	0,781
α	1,830	1,831	0,125	1,592	2,070	2,100	2,097	0,169	1,787	2,435
β	0,126	0,124	0,027	0,077	0,182	0,061	0,060	0,026	0,013	0,110
$\hat{\beta}_0$	-1,622	-1,652	0,533	-2,582	-0,521	-1,297	-1,375	0,859	-3,016	0,283
$\hat{\beta}_1$	-0,047	-0,043	0,194	-0,464	0,297	-0,212	-0,207	0,267	-0,731	0,295
$\hat{\beta}_2$	0,014	0,014	0,007	-0,001	0,028	0,016	0,016	0,009	-0,002	0,035
$\hat{\beta}_3$	0,498	0,499	0,087	0,327	0,662	0,679	0,678	0,120	0,448	0,913
$\hat{\beta}_4$	-0,129	-0,127	0,198	-0,511	0,264	-0,063	-0,066	0,266	-0,608	0,433
$\hat{\beta}_5$	0,236	0,236	0,276	-0,342	0,745	0,235	0,236	0,373	-0,497	0,965
$\hat{\beta}_6$	0,040	0,040	0,028	-0,015	0,095	0,035	0,036	0,038	-0,036	0,111

Para detectar as possíveis observações influentes na distribuição *a posteriori* dos parâmetros do modelo, as estimativas de Monte Carlo das medidas de ψ -divergência foram obtidas com as amostras *a posteriori* dos parâmetros dos modelos ZIP-FC e ZIG-FC. Na Figura 4.1 apresenta-se as estimativas de Monte Carlo das medidas de divergência $K - L$, distancia L_1 , distância J e a divergência χ^2 , onde a observação 176 é indicado como possível observação influente, no modelo ZIP-FC, mas essa observação não é identificado no modelo ZIG-FC, indicado que o esse modelo é robusto na presença de observações influentes. A Tabela 4.4 contém as medidas de divergência para a observação 176, onde observa-se que se trata-se de uma paciente do sexo masculino, de 53 anos de idade, tratado sem interferon, com categoria de nódulo de nível 4, com capacidade funcional ativo, espessura de tumor 1,92 cm e que faleceu após 6 anos de ser detectado a doença.

Tabela 4.4: Estimativas de de Monte Carlo das medidas ψ -divergência para os modelos ZIP-FC e ZIG-FC.

Modelo	t	x_1	x_2	x_3	x_4	x_5	x_6	Medida			
								D_{K-L}	D_J	D_{L_1}	D_{χ^2}
ZIP-FC	6	0	53	4	0	0	1,92	0,262	0,561	0,287	0,608
ZIG-FC	6	0	53	4	0	0	1,92	0,161	0,341	0,224	0,341

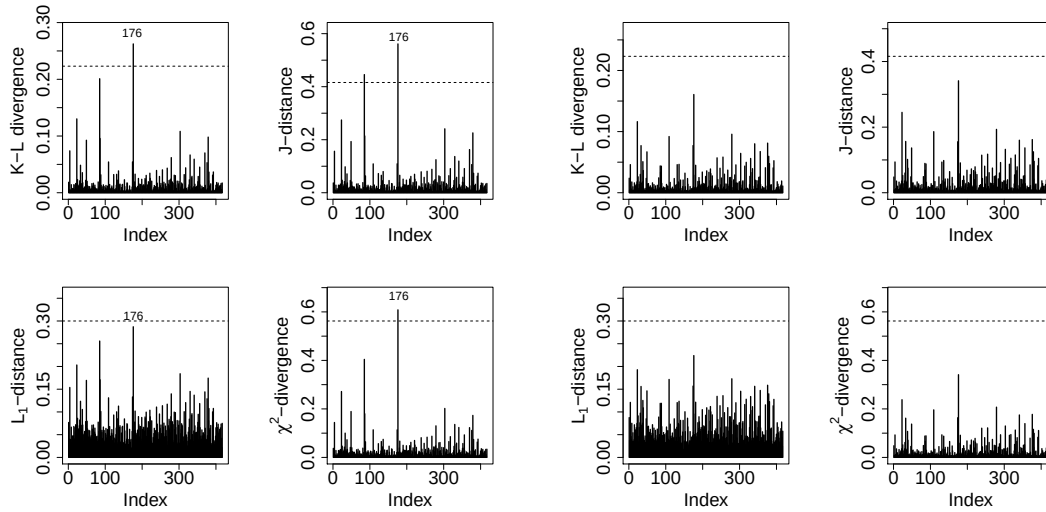


Figura 4.1: Estimativas de Monte Carlo das medidas de ψ -divergência do modelo do modelo ZIP-FC (panel direito) e ZIG-FC (panel esquerdo) para dados de melanoma.

A seguir comparamos os modelos ZIP-FC (modelo M_1) e ZIG-FC (modelo M_0), utilizando o pseudo fator de Bayes. A estimativa de Monte Carlo do pseudo fator de Bayes do modelo ZIG-FC com respeito ao modelo ZIP-FC resultou, $PSFB_{01} = 4,5$, isto indica uma forte evidência a favor do modelo ZIG-FC. Para confirmar esse resultado, consideramos

o critério Bayesiano DIC ("deviance information criteria") proposta por Spiegelhalter *et al.* (2002) e a versão Bayesiana do AIC ("Aikake information criterion") e BIC ("Bayesian information criterion") discutida por Brooks (2002). Na Tabela 4.5, apresenta-se as estimativas dos critérios, onde observa-se para todos critérios que o modelo ZIG-FC melhor para a modelagem dos dados de melanoma. Portanto retemos o modelo ZIG-FC na análise a continuação

Tabela 4.5: *Estimativas de Monte Carlo do DIC, EAIC e EBIC para o ajuste dos dados de melanoma.*

Modelo	DIC	EAIC	EBIC
ZIP-FC	1032,037	1042,295	1082,625
ZIG-FC	1029,056	1039,814	1080,145

Com a finalidade de comparar os resultados da análise clássica, do modelo ZIG-FC simplificado com análise Bayesiana, ajustamos o modelo considerando as covariáveis idade (x_2) e categoria do nódulo (x_3). Na Tabela 4.6 apresentamos as estimativas Bayesianas para o modelo simplificado, esse resultado são similares aos resultados clássicos.

Tabela 4.6: *Sumário a posteriori, média, mediana, desvio padrão e intervalo HPD de 90% de confiança, para o modelo simplificado.*

			Desvio	Intervalo HPD (90%)	
	Média	Mediana	padrão	LI	LS
ϕ	0,382	0,381	0,174	0,099	0,664
γ_1	2,056	2,053	0,157	1,808	2,317
γ_2	0,069	0,066	0,029	0,018	0,113
$\hat{\beta}_0$	-1,263	-1,341	0,744	-2,447	-0,160
$\hat{\beta}_2$	0,017	0,017	0,009	0,001	0,031
$\hat{\beta}_3$	0,622	0,621	0,117	0,427	0,808

A estimativa Bayesiana da proporção de risco zero determinístico, $\omega = \phi/(1 + \phi)$ é 0,276(0,0987) a qual indica que um 27,6% dos indivíduos curados da melanoma tem fatores genéticos próprios que permitem a cura da doença.

A seguir testa-se as hipóteses $H_0 : \omega = 0$ contra $H_1 : \omega \neq 0$, com a finalidade de avaliar se a variável de fragilidade pode ser modelada pela distribuição Geométrica. Para isso estimamos o pseudo fator de Bayes do modelo de sobrevivência Geométrica com fração de cura (modelo M_0) com respeito o modelo ZIG-FC (modelo M_1) cujo valor resultou em $PSFB_{01} = 0,312$, o qual indica que o modelo Geométrico não é adequado para modelar a

variável de fragilidade nos dados de melanoma. Esse resultado é confirmado pelo critérios DIC, AIC e BIC com mostra as estimativas dadas na Tabela 4.7.

Tabela 4.7: *Estimativas Bayesianas DIC, EAIC e EBIC para o ajuste dos dados de melanoma.*

Modelo	DIC	EAIC	EBIC
ZIG-FC	1023,458	1030,344	1054,543
Geométrica-FC	1026,301	1031,338	1051,504

A Figura 4.2 mostra a função de sobrevivência estratificada por nível categoria do nódulo em pacientes com idade igual a 32 e 66 anos, que correspondem aos percentis 10 e 90. Estes gráficos destacam o impacto combinado das covariáveis sobre a fração de curados. Nestes gráficos, observamos que os tempos de sobrevivência dos pacientes com melanoma diminui com a idade e o nível de categoria do nódulo.

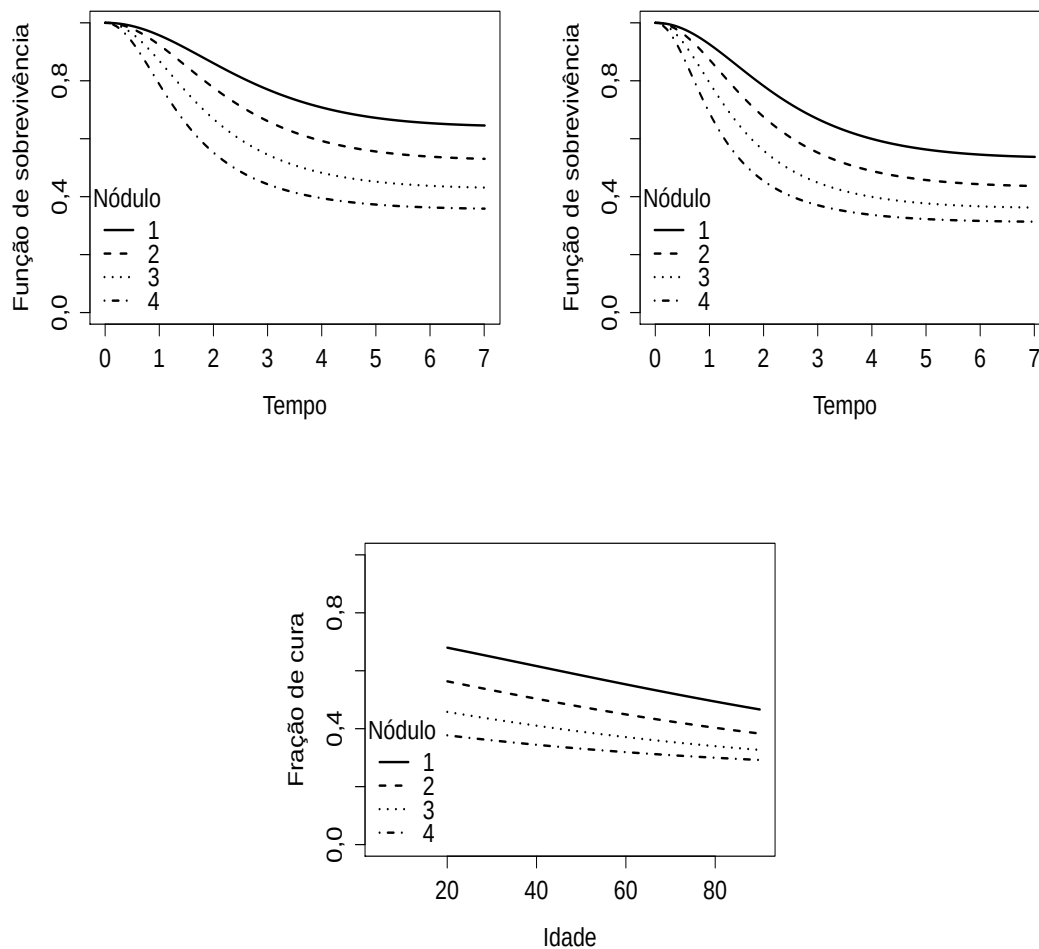


Figura 4.2: *Estimativa Bayesiana da função de sobrevivência sob modelo ZIG-FC estratificada por categoria de nódulo para pacientes com idade igual a 32 anos (painel esquerdo) e 66 anos (painel direito) e, fração cura (painel central).*

Finalmente, apresentamos a Tabela 4.8 com as estimativas Bayesianas da fração de cura e os intervalos HPD ao nível de 90% de confiança. Note que a fração de cura entre os pacientes com idade de 32 e 66 anos em cada categoria de nódulo, os intervalos não se sobrepõem. A Figura 4.2 mostra a fração de cura como função da idade, onde observamos que a fração de cura de pacientes com melanoma diminui com a idade.

Tabela 4.8: *Estimativa Bayesianas para fração de curados estratificada por categoria de nódulo e idade dos pacientes.*

Idade	Categoria de Nódulo	Média	Mediana	Desvio Padrão	Intervalo HPD (90%)	
					LI	LS
32	1	0,6573	0,6661	0,0732	0,5429	0,7787
	2	0,5549	0,5569	0,0553	0,4686	0,6470
	3	0,4644	0,4650	0,0421	0,3999	0,5366
	4	0,3957	0,3965	0,0459	0,3169	0,4656
66	1	0,5652	0,5667	0,0582	0,4795	0,6718
	2	0,4718	0,4708	0,0432	0,3991	0,5414
	3	0,3995	0,3998	0,0437	0,3287	0,4712
	4	0,3497	0,3515	0,0557	0,2604	0,4378

4.4 Conclusões

Neste Capítulo foi desenvolvido uma análise Bayesiana para o modelo ZISP-FC. A metodologia Bayesiana é baseada em métodos de Monte Carlo via cadeia de Markov (MCMC). O pseudo fator de Bayes foi considerado para testar hipóteses a respeito do parâmetro de interesse do modelo ZISP-FC. A fim de avaliar a robustez do modelo na presença de observações influentes, foi desenvolvida uma análise de diagnóstico Bayesiano. As propriedades frequentistas das estimativas Bayesianas são investigadas por meio de um estudo de simulação. E a potencialidade da metodologia Bayesiana é demonstrada no ajuste de um conjunto de dados coletados em estudo de melanoma.

Capítulo 5

Conclusões e Pesquisas Futuras

Neste trabalho foi formulada a distribuição da estatística gradiente para testar a inflação de zeros nos modelos zero inflacionado série de potências (ZISP). Como casos particulares, os modelos zero-inflacionado Poisson e zero-inflacionado Geométrica foram abordados para avaliar a performance da estatística gradiente em comparação das estatísticas da razão de verossimilhança, escore e Wald, para pequenas, moderadas e grandes tamanhos de amostras, baseado em um estudo de simulação. Resultados de simulação mostraram que o desempenho da estatística gradiente foi competitiva às estatísticas clássicas para os casos com e sem covariáveis. Na avaliação do poder do teste, verificamos que a estatística gradiente foi a melhor. Também, foram apresentados dois exemplos de aplicação, sendo um pertencente a área de saúde pública e a outra da agronomia em que as estatísticas identificaram a significância do excesso de zeros para cada problema. Com este estudo realizado, a estatística gradiente se torna uma metodologia alternativa e atrativa aos testes clássicos na análise do parâmetro de inflação nos modelos ZISP.

Um novo modelo para dados de sobrevivência com fração de cura foi proposto e estudadas as propriedades estruturais. O novo modelo foi obtido partir de um modelo de fragilidade com distribuição ZISP. Esse modelo tem a propriedade de classificar a proporção de indivíduos com risco zero em duas componentes (imunes e curados), uma que é devido aos fatores próprios dos indivíduos (ou fatores determinísticos) e a segunda é devido a fatores aleatórios. Além disso o modelo proposto inclui como casos particulares o modelo proposto por [Cancho *et al.* \(2013\)](#) entre outros. Também foram desenvolvidas procedimentos inferenciais para os parâmetros do modelo proposto, baseada na metodologia de máxima verossimilhança. Para estudar as propriedades assintóticas dos estimadores de máxima verossimilhança dos parâmetros do modelo, um estudo de simulação foi realizado e o estudo mostrou que a normalidade assintótica é satisfeita para amostras maiores de 100. Também a metodologia é avaliada em uma aplicação baseado em um estudo de pacientes com melanoma.

Alternativamente a metodologia clássica, foram desenvolvidas procedimentos infe-

rênciais desde uma perspectiva Bayesiana. A metodologia Bayesiana é baseada em métodos de Monte Carlo via Cadeia de Markov. Além disso, um estudo de diagnóstico por meio de uma medida de divergência geral que inclui como caso particular a medida de divergência de Kullback-Leibler foi realizado. O teste Bayesiano baseado no pseudo fator de Bayes para testar a hipóteses a respeito dos parâmetros do novo foi desenvolvido e, o estudo de simulação mostra o poder de discriminação dos modelos do pseudo fator de Bayes é adequado.

Trabalhos futuros

Como continuação da desenvolvida presente tese propomos os seguintes temas de pesquisa.

1. Investigar o teste gradiente para os modelos ZISP semiparamétricos na linha de [Li \(2012\)](#).
2. Seria interessante desenvolver um estudo da estatística gradiente nos modelos zero inflacionados Conway-Maxwell-Poisson (ZI-COM) proposto por [Barriga & Louzada \(2014\)](#).
3. Desenvolver o estudo da estatística gradiente em modelos aditivos generalizados para dados de contagem zero inflacionados na linha de [Nian \(2014\)](#).
4. Estudar os modelos ZISP-FC para dados de sobrevivência bivariados e multivariados.
5. Utilizar os modelos ZISP-FC para dados com eventos recorrentes, a qual é uma proposta que estende o trabalho de [Macera *et al.* \(2015\)](#).
6. Considerar generalizações do modelo de fragilidade como fragilidade compartilhada nas suas diversas direções ([Liang *et al.*, 1995](#)) para os modelos ZISP-FC.

Apêndice A

A.1 Obtenção da estatística Gradiente

Seja a estatística Wald modificada proposta por Hayakawa & Puri (1985) dada por

$$S_{Wm} = (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^\top K(\boldsymbol{\vartheta}_0)(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0) \quad (\text{A.1})$$

Suponha uma matriz quadrada $L_{r \times r}$ tal que $L^\top L = K(\boldsymbol{\vartheta})$. Usamos as estatísticas de escore dada em (1.3) e Wald modificada em (A.1), que podem ser rescritas na forma

$$\begin{aligned} S_R &= [L^{-1}U(\boldsymbol{\vartheta}_0)]^\top L^{-1\top} U(\boldsymbol{\vartheta}_0), \\ S_{Wm} &= [L(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)]^\top L(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0) \end{aligned}$$

em que os vetores $L^{-1}U(\boldsymbol{\vartheta}_0)$ e $L(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)$ possuem distribuição $N_r(\mathbf{0}, I_r)$ para um n suficientemente grande. I_r é uma matriz identidade de ordem r . Então pode-se afirmar que $L^{-1}U(\boldsymbol{\vartheta}_0) - L(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)$ convergem em probabilidade para zero. Além o produto desses dois vetores padronizados segue uma distribuição Qui-quadrado central com r graus de liberdade e resulta em

$$[L^{-1}U(\boldsymbol{\vartheta}_0)]^\top L(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0) = U(\boldsymbol{\vartheta}_0)^\top ((L^{-1})^\top)^\top L(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0) = U(\boldsymbol{\vartheta}_0)^\top (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0).$$

Assim a estatística gradiente é dada por

$$S_G = U(\boldsymbol{\vartheta}_0)^\top (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0).$$

A.1.1 Normalidade assintótica das estatísticas

A estatística de Wald em (1.4) possui uma forma quadrática correspondente à distribuição Normal assintótica com media $\boldsymbol{\vartheta}_0$ e matriz de covariância $K(\hat{\boldsymbol{\vartheta}})$, isto é,

$$(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0) \xrightarrow{D} N_r(0, K(\hat{\boldsymbol{\vartheta}})),$$

em que $\dim(\boldsymbol{\vartheta}_0) = r$, logo pela forma quadrática temos

$$S_W \xrightarrow{D} \chi_r^2.$$

A estatística escore em (1.3) baseia-se na normalidade assintótica da componente de interesse da função escore

$$U(\boldsymbol{\vartheta}_0) \xrightarrow{D} N_r(0, K(\boldsymbol{\vartheta}_0)^{-1}),$$

assim também pela sua forma quadrática

$$S_R \xrightarrow{D} \chi_r^2.$$

Esta propriedade assintótica também a possui a estatística S_{LR} e foi mostrado em Wilks (1938). Enquanto a normalidade assintótica da estatística gradiente é mostrada no Teorema 1 de Terrell (2002), desde que os dados cumprem as condições de regularidade para $\ell(\boldsymbol{\vartheta})$ (Bickel e Doksum ,2001 Capitulo 6.2), sendo $\hat{\boldsymbol{\vartheta}}$ a estimativa assintoticamente eficiente para $\boldsymbol{\vartheta}$. Em consequência S_G possui uma distribuição Qui-quadrado com (r) graus de liberdade.

Portanto, as estatísticas (1.2) - (1.5) são equivalentes até primeira ordem, pois todas convergem sob hipótese nula para a distribuição χ_r^2 . Disto, a hipótese nula H_0 pode ser rejeitada para grandes valores das estatísticas mencionadas, comparados com o quantil de ordem $1 - \alpha$ da distribuição Qui-quadrado com (r) graus de liberdade para um nível de significância nominal determinado.

Apêndice B

B.1 Propriedades do modelo SP

A obtenção da esperança e variância do modelo SP é baseado na transformação exponencial natural. Considere

$$Q(X, \eta) = \exp\{\eta T(X) + d_0(\eta) + e(X)\} I_{\{0,1,2,\dots\}}(x)$$

em que $d_0(\eta) = d(c^{-1}(\eta))$. Assim, a esperança e variância de $T(X)$ são dadas, respectivamente, por

$$\begin{aligned} E[T(X)] &= -d'_0(\eta) \quad e \\ Var[T(X)] &= -d''_0(\eta). \end{aligned}$$

Para $X = Z$, a forma exponencial da classe série de potências é dada por

$$Q(Z = z, \theta) = \exp\{z \log \theta - \log A(\theta) + \log a_z\},$$

em que $c(\theta) = \log \theta$, $d(\theta) = -\log A(\theta)$ e $e(z) = \log a_z$. Logo, se $T(Z) = z$ e $\eta = c(\theta)$, temos que $d_0(\eta) = d(e^\eta)$. Portanto, a forma exponencial natural do modelo SP pode ser reescrita da forma

$$Q_0(z, \eta) = \exp\{z\eta - \log A(e^\eta) + \log a_z\}$$

em consequência por redução de termos, a esperança e a variância são dadas, respectivamente, por

$$E[Z; \theta] = \frac{1}{A(e^\eta)} A'(e^\eta) \quad (\text{B.1})$$

$$\begin{aligned} &= -\frac{1}{A(e^\eta)} \frac{dA(e^\eta)}{d\theta} \frac{d\theta}{d\eta}, \quad d\theta = e^\eta d\eta \\ &= \frac{\theta}{A(\theta)} A'(\theta) \\ \text{Var}[Z; \theta] &= \frac{A(e^\eta)(A(e^\eta))''_\eta - (A(e^\eta)'_\eta)^2}{(A(e^\eta))^2} \quad (\text{B.2}) \\ &= \frac{A(\theta)A'(\theta)\theta + A(\theta)A''(\theta)\theta^2 - \theta^2(A'(\theta))^2}{(A(\theta))^2} \\ &= \frac{\theta A'(\theta)}{A(\theta)} + \theta^2 \left(\frac{A''(\theta)}{A(\theta)} - \left[\frac{A'(\theta)}{A(\theta)} \right]^2 \right). \end{aligned}$$

O segundo momento é dado por

$$E[Z^2; \theta] = \text{Var}[Z; \theta] + (E[Z; \theta])^2 = \frac{\theta}{A(\theta)} (A'(\theta) + \theta A''(\theta)). \quad (\text{B.3})$$

Também, é possível obter as expressões da esperança e da variância pelo método dos momentos.

A função geradora de probabilidade (f.g.p) para esta classe de distribuições é dada por

$$G_{z; \theta} = \frac{A(\theta z)}{A(\theta)}. \quad (\text{B.4})$$

B.2 Propriedades do modelo ZISP

A esperança e a variância são dadas, respectivamente, por

$$\begin{aligned} E[Z; \vartheta] &= \sum_{z=0}^{\infty} z P(Z; \vartheta) \quad (\text{B.5}) \\ &= 0 + \sum_{z=1}^{\infty} z \frac{a_z \theta^z}{(1+\phi)A(\theta)} \\ &= \frac{1}{(1+\phi)} E[Z; \theta] \\ &= \frac{\theta}{(1+\phi)} \frac{A'(\theta)}{A(\theta)}. \end{aligned}$$

Precisamos calcular o segundo momento, que é da forma

$$\begin{aligned}
E[Z^2; \vartheta] &= \sum_{z=0}^{\infty} z^2 P(Z; \vartheta) \\
&= 0 + \sum_{z=1}^{\infty} z^2 \frac{a_z \theta^z}{(1+\phi)A(\theta)} \\
&= \frac{1}{(1+\phi)} E[Z^2; \theta] \\
&= \frac{\theta}{(1+\phi)} \left(\frac{A'(\theta) + \theta A''(\theta)}{A(\theta)} \right).
\end{aligned} \tag{B.6}$$

Podemos determinar também o r -ésimo momento, que é dado por

$$\begin{aligned}
E[Z^r; \vartheta] &= \sum_{z=0}^{\infty} z^r P(Z; \vartheta) \\
&= 0 + \sum_{z=1}^{\infty} z^r \frac{a_z \theta^z}{(1+\phi)A(\theta)} \\
&= \frac{E[Z^r; \theta]}{1+\phi}.
\end{aligned} \tag{B.7}$$

(B.8)

Logo, a variância é dada por

$$\begin{aligned}
V[Z; \vartheta] &= E[Z^2; \vartheta] - (E[Z; \vartheta])^2 \\
&= \frac{E[Z^2; \theta]}{1+\phi} - \left(\frac{E[Z; \theta]}{1+\phi} \right)^2 \\
&= \frac{V[Z; \theta] + E[Z; \theta]^2}{1+\phi} - \left(\frac{E[Z; \theta]}{1+\phi} \right)^2 \\
&= \frac{1}{1+\phi} V[Z; \theta] + \frac{\phi}{1+\phi} E[Z; \theta]^2.
\end{aligned} \tag{B.9}$$

A função geradora de probabilidade é dada por

$$\begin{aligned}
G_{z;\vartheta} &= \sum_{z_i=0}^{\infty} y^{z_i} P(Z_i; \vartheta) \\
&= y^0 \frac{\phi + a_0 A(\theta)^{-1}}{1 + \phi} + \sum_{z_i=1}^{\infty} y^{z_i} \frac{a_{z_i} \theta^{z_i} A(\theta)^{-1}}{1 + \phi} \\
&= \frac{\phi + a_0 A(\theta)^{-1}}{1 + \phi} + \frac{A(\theta)^{-1}}{1 + \phi} \sum_{z_i=1}^{\infty} a_{z_i} (\theta y)^{z_i} \\
&= \frac{\phi}{1 + \phi} + \frac{A(\theta)^{-1}}{1 + \phi} (a_0 + \sum_{z_i=1}^{\infty} a_{z_i} (\theta y)^{z_i}) \\
&= \frac{\phi}{1 + \phi} + \frac{1}{1 + \phi} \frac{A(\theta y)}{A(\theta)} \\
&= \frac{\phi}{1 + \phi} + \frac{1}{1 + \phi} G_{z;\theta}.
\end{aligned} \tag{B.10}$$

B.2.1 Cálculo da função score e da matriz de IF

Primeiras derivadas:

$$\begin{aligned}
\frac{\partial \ell}{\partial \phi} &= \sum_{i=1}^n \left\{ \frac{-1}{1 + \phi} + I_{(z_i=0)} \frac{1}{\phi + a_0 A(\theta)^{-1}} \right\} \\
&= \frac{-n}{1 + \phi} + \frac{n_0}{\phi + a_0 A(\theta)^{-1}}.
\end{aligned}$$

Igualando a zero, resulta que

$$\phi = \frac{n_0 - n a_0 A(\hat{\theta})^{-1}}{n - n_0}.$$

Logo,

$$\begin{aligned}
\frac{\partial \ell}{\partial \theta} &= \sum_{i=1}^n \left\{ I_{(z_i=0)} \frac{-a_0}{\phi + a_0 A(\theta)^{-1}} \frac{A'(\theta)}{A(\theta)^2} + I_{(z_i>0)} \left(\frac{z_i}{\theta} - \frac{A'(\theta)}{A(\theta)} \right) \right\} \\
&= \frac{-a_0}{\phi + a_0 A(\theta)^{-1}} \frac{A'(\theta)}{A(\theta)^2} \sum_{i=1}^n I_{(z_i=0)} + \frac{1}{\theta} \sum_{i=1}^n z_i I_{(z_i>0)} - \frac{A'(\theta)}{A(\theta)} \sum_i I_{(z_i>0)}.
\end{aligned}$$

Da mesma forma, igualamos a zero, substituímos ϕ e reduzimos os termos, tem-se

$$\begin{aligned}
\frac{1}{\theta} \sum_{i=1}^n z_i I_{(z_i > 0)} &= \frac{-n_0 a_0}{\phi + a_0 A(\theta)^{-1}} \frac{A'(\theta)}{A(\theta)^2} + (n - n_0) \frac{A'(\theta)}{A(\theta)} \\
&= \frac{-n_0 a_0}{\frac{n_0 - n a_0 A'(\hat{\theta})^{-1}}{n - n_0} + a_0 A(\theta)^{-1}} \frac{A'(\theta)}{A(\theta)^2} + (n - n_0) \frac{A'(\theta)}{A(\theta)} \\
&= (n - n_0) \left[\frac{a_0}{(1 - a_0) A(\theta)^{-1}} \frac{1}{A(\theta)} + 1 \right] \frac{A'(\theta)}{A(\theta)} \\
\frac{\sum_{i=1}^n z_i I_{(z_i > 0)}}{(n - n_0)} &= \frac{\theta A'(\theta)}{A(\theta) - a_0} \\
\bar{z} &= \frac{\hat{\theta} A'(\hat{\theta})}{A(\hat{\theta}) - a_0}.
\end{aligned} \tag{B.11}$$

Segundas derivadas

$$\begin{aligned}
\frac{\partial^2 \ell}{\partial \theta^2} &= \sum_{i=1}^n \left\{ I_{(z_i=0)} \left[\frac{-a_0}{\phi + a_0 A(\theta)^{-1}} \frac{\partial}{\partial \theta} \left(\frac{A'(\theta)}{A(\theta)^2} \right) + \frac{A'(\theta)}{A(\theta)^2} \frac{\partial}{\partial \theta} \left(\frac{-a_0}{\phi + a_0 A(\theta)^{-1}} \right) \right] \right. \\
&\quad \left. - \frac{1}{\theta^2} z_i I_{(z_i > 0)} - I_{(z_i > 0)} \frac{\partial}{\partial \theta} \left(\frac{A'(\theta)}{A(\theta)} \right) \right\}
\end{aligned} \tag{B.12}$$

$$\begin{aligned}
&= -a_0 \left[\frac{A(\theta) A''(\theta) (a_0 + \phi A(\theta)) - A'(\theta)^2 (a_0 + 2\phi A(\theta))}{A(\theta)^4 (\phi + a_0 A(\theta)^{-1})} \right] \sum_{i=1}^n I_{(z_i=0)} \\
&\quad - \frac{1}{\theta^2} \sum_{i=1}^n z_i I_{(z_i > 0)} - \frac{A(\theta) A''(\theta) - (A'(\theta))^2}{A(\theta)^2} \sum_{i=1}^n I_{(z_i > 0)},
\end{aligned}$$

$$\frac{\partial^2 \ell}{\partial \phi^2} = \sum_{i=1}^n \left\{ \frac{1}{(1 + \phi)^2} + I_{(z_i=0)} \frac{-1}{(\phi + a_0 A(\theta)^{-1})^2} \right\}, \tag{B.13}$$

$$\frac{\partial^2 \ell}{\partial \phi \partial \theta} = \frac{\partial^2 \ell}{\partial \theta \partial \phi} = \sum_{i=1}^n \left\{ I_{(z_i=0)} \frac{A'(\theta)}{A(\theta)} \frac{-a_0}{(\phi + a_0 A(\theta)^{-1})^2} \right\}, \tag{B.14}$$

em que,

$$\begin{aligned}
A'(\theta) &= \partial A(\theta) / \partial \theta, \\
\frac{\partial}{\partial \theta} \left(\frac{A'(\theta)}{A(\theta)} \right) &= \frac{A(\theta) A''(\theta) - (A'(\theta))^2}{A(\theta)^2}, \\
\frac{\partial}{\partial \theta} \left(\frac{A'(\theta)}{A(\theta)^2} \right) &= \frac{A(\theta)^2 A''(\theta) - 2(A'(\theta))^2 A(\theta)}{A(\theta)^4}.
\end{aligned}$$

Precisamos aplicar a esperança nas segundas derivadas. Usamos $E[I_{(z_i=0)}] = P(z_i = 0) = \frac{\phi + a_0 A(\theta)^{-1}}{1 + \phi}$ e $E[I_{(z_i > 0)}] = P(z_i > 0) = \frac{1 - a_0 A(\theta)^{-1}}{1 + \phi}$ para encontrar os valores das componentes da matriz de informação de Fisher. Primeiramente, calcularmos a esperança em (B.11)

$$\begin{aligned}
\frac{1}{\theta} E \left[\sum_{i=1}^n z_i I_{(z_i > 0)} \right] &= \frac{-a_0}{\phi + a_0 A(\theta)^{-1}} \frac{A'(\theta)}{A(\theta)^2} \sum_{i=1}^n P(z_i = 0) + \frac{A'(\theta)}{A(\theta)} \sum_{i=1}^n P(z_i > 0) \\
&= \frac{n}{1 + \phi} \left[\frac{a_0}{A(\theta) + (1 - a_0 A(\theta)^{-1})} \right] \frac{A'(\theta)}{A(\theta)} \\
E \left[\sum_{i=1}^n z_i I_{(z_i > 0)} \right] &= \frac{n\theta}{1 + \phi} \frac{A'(\theta)}{A(\theta)}. \tag{B.15}
\end{aligned}$$

Logo,

$$\begin{aligned}
E \left[\frac{\partial^2 \ell}{\partial \theta^2} \right] &= -a_0 \left[\frac{A(\theta) A''(\theta) (a_0 + \phi A(\theta)) - A'(\theta)^2 (a_0 + 2\phi A(\theta))}{A(\theta)^4 (\phi + a_0 A(\theta)^{-1})} \right] \sum_{i=1}^n E[I_{(z_i=0)}] \\
&\quad - \frac{1}{\theta^2} \sum_{i=1}^n E[z_i I_{(z_i > 0)}] - \frac{A(\theta) A''(\theta) - (A'(\theta))^2}{A(\theta)^2} \sum_i E[I_{(z_i > 0)}] \\
&= -a_0 \left[\frac{A(\theta) A''(\theta) (a_0 + \phi A(\theta)) - A'(\theta)^2 (a_0 + 2\phi A(\theta))}{A(\theta)^4 (\phi + a_0 A(\theta)^{-1})} \right] \frac{n}{1 + \phi} \\
&\quad - \frac{n}{1 + \phi} \left[\frac{\theta A'(\theta)}{A(\theta)} + \frac{A(\theta) A''(\theta) - (A'(\theta))^2}{A(\theta)^2} (1 - a_0 A(\theta)^{-1}) \right], \\
E \left[\frac{\partial^2 \ell}{\partial \phi^2} \right] &= \frac{n}{(1 + \phi)^2} - \frac{1}{(\phi + a_0 A(\theta)^{-1})^2} \sum_{i=1}^n E[I_{(z_i=0)}] \\
&= \frac{-n(1 - a_0 A(\theta)^{-1})}{(1 + \phi)^2 (\phi + a_0 A(\theta)^{-1})}, \\
E \left[\frac{\partial^2 \ell}{\partial \theta \partial \phi} \right] &= \frac{-na_0}{(1 + \phi)(\phi + a_0 A(\theta)^{-1})} \frac{A'(\theta)}{A(\theta)^2} = E \left[\frac{\partial^2 \ell}{\partial \phi \partial \theta} \right].
\end{aligned}$$

Proposição 1. O modelo ZISP pertence a família exponencial de dois parâmetros então as condições de regularidade são válidas e os EMV são assintoticamente Normais.

Prova

Consideramos de forma geral

$$P(Z_i; \theta, \phi) = \begin{cases} \frac{\phi + P(0, \theta)}{1 + \phi} & , z_i = 0 \\ \frac{P(z_i; \theta)}{1 + \phi} & , z_i = 1, 2, 3, \dots \end{cases} \tag{B.16}$$

Podemos rescrever na seguinte forma

$$P(Z_i; \theta, \phi) = \left\{ \frac{\phi + P(0, \theta)}{1 + \phi} \right\}^{t(z_i)} \left\{ \frac{P(z_i; \theta)}{1 + \phi} \right\}^{(1-t(z_i))}$$

em que

$$t(z) = \begin{cases} 1 & , \text{se } z_i = 0 \\ 0 & , \text{se } z_i \geq 1. \end{cases}$$

Aplicando o logaritmo, obtemos

$$\begin{aligned}\log P(z_i; \theta, \phi) &= t(z_i)\{\log(\phi + P(0; \theta)) - \log(1 + \phi)\} + (1 - t(z_i))\{\log P(z; \theta) - \log(1 + \phi)\} \\ &= t(z_i)\log(\phi + P(0; \theta)) + (1 - t(z_i))\log P(z_i; \theta) - \log(1 + \phi).\end{aligned}$$

Logo, a expressão acima pode ser escrita da forma

$$P(z_i; \theta, \phi) = W(z_i)V(\theta, \phi) \exp\{t(z_i)C_1(\theta, \phi) + (1 - t(z_i))C_2(\theta, \phi)\},$$

em que $C_1 = \log(\phi + P(0; \theta))$, $C_2 = \log P(z_i; \theta)$, $W(z_i) = 1$ e $V(\theta, \phi) = (1 + \phi)^{-1}$. Assim, o modelo ZISP pertence a família exponencial de dois parâmetros. C_1 e C_2 tem interior não vazio e $T(Z_i) = (t(z_i), 1 - t(z_i))$ é uma estatística suficiente e completa.

Apêndice C

C.1 Método do gráfico TTT

Uma técnica de verificação gráfica é importante porque permite extrair informação qualitativa que pode ser utilizada na determinação empírica da forma da função de risco. Esta técnica é conhecida como TTT-plot proposto por [Sun & Kececiloglu \(1999\)](#).

A curva TTT é construída obtendo um gráfico de

$$G(r/n) = \frac{[\sum_{i=1}^r T_{i:n} + (n-r)T_{r:n}]}{(\sum_{i=1}^n T_{i:n})} \text{ versus } \frac{r}{n} \quad (\text{C.1})$$

em que $r = 1, \dots, n$ e $T_{i:n}$, $i = 1, \dots, n$ são as estatísticas de ordem da amostra.

É provado que a função de risco cresce (decresce) se o gráfico TTT é côncavo (convexo) e a distribuição de Weibull que acomoda essas formas de curva, é a candidata ao ajuste. Se a curva é convexa e então concava a função de risco tem a forma de "U", e no caso reverso é unimodal.

C.2 Matriz de informação observada do modelo ZISP-FC

Componentes da matriz $K(\vartheta)$

$$K_{\beta\beta} = \sum_{i=1}^n \left\{ -\frac{A(\theta_i)A'(\theta_i) - (A'(\theta_i))^2}{A(\theta_i)^2} - \delta_i \left[\frac{1}{\theta_i^2} - S_0^2(t_i; \gamma) \frac{A'(\theta_i S_0(t_i; \gamma))A'''(\theta_i S_0(t_i; \gamma)) - (A''(\theta_i S_0))^2}{A'(\theta_i S_0(t_i; \gamma))^2} \right] \right. \\ \left. + (1 - \delta_i) \frac{c_1 \dot{c}_2 - c_2 \dot{c}_1}{c_1^2} \right\} \frac{\partial \theta_i}{\partial \beta} \frac{\partial \theta_i}{\partial \beta^\top} + \mathbf{D} \frac{\partial^2 \theta_i}{\partial \beta \partial \beta^\top} \quad (\text{C.2})$$

em que

$$c_1 = \phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma)), \\ c_2 = \dot{c}_1 = \phi A'(\theta_i) + A'(\theta_i S_0(t_i; \gamma)) S_0(t_i; \gamma),$$

$$\dot{c}_2 = \phi A''(\theta_i) + A''(\theta_i S_0(t_i; \gamma)) S_0(t_i; \gamma)^2,$$

e

$$\mathbf{D} = \frac{-A'(\theta_i)}{A(\theta_i)} + \delta_i \left[\frac{1}{\theta_i} + \frac{A''(\theta_i S_0(t_i; \gamma))}{A'(\theta_i S_0(t_i; \gamma))} S_0(t_i; \gamma) \right] + (1 - \delta_i) \frac{\phi A'(\theta_i) + A'(\theta_i S_0(t_i; \gamma)) S_0(t_i; \gamma)}{\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma))}.$$

$$K_{\beta\gamma} = \sum_{i=1}^n \left\{ \delta_i \left[\frac{A''(\theta_i S_0(t_i; \gamma))}{A'(\theta_i S_0(t_i; \gamma))} + \theta_i S_0(t_i; \gamma) \left(\frac{A'(\theta_i S_0(t_i; \gamma)) A''(\theta_i S_0(t_i; \gamma)) - (A'(\theta_i S_0(t_i; \gamma)))^2}{A'(\theta_i S_0(t_i; \gamma))^2} \right) \right] \right. \\ \left. (1 - \delta_i) \frac{c_1 c_2^* - c_2 c_1^*}{c_1^2} \right\} \frac{\partial S_0(t_i; \gamma)}{\partial \gamma} \frac{\partial \theta_i}{\partial \beta}, \quad (\text{C.3})$$

em que

$$c_1^* = A(\theta_i S_0(t_i; \gamma)) \theta_i$$

e

$$c_2^* = A'(\theta_i S_0(t_i; \gamma)) + \theta_i S_0(t_i; \gamma) A''(\theta_i S_0(t_i; \gamma)).$$

$$K_{\beta\phi} = \sum_{i=1}^n \left\{ (1 - \delta_i) \frac{(\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma))) A'(\theta_i) - (\phi A'(\theta_i) + A'(\theta_i S_0(t_i; \gamma))) A(\theta_i)}{(\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma)))^2} \right\} \frac{\partial \theta_i}{\partial \beta}. \quad (\text{C.4})$$

$$K_{\gamma\gamma} = \sum_{i=1}^n \left\{ \delta_i \left[\frac{1}{f_0(t_i; \gamma)} \frac{\partial^2 f_0(t_i; \gamma)}{\partial \gamma^2} + \theta_i^2 \frac{A'(\theta_i S_0(t_i; \gamma)) A'''(\theta_i S_0(t_i; \gamma)) - (A''(\theta_i S_0(t_i; \gamma)))^2}{A'(\theta_i S_0(t_i; \gamma))^2} \left(\frac{\partial S_0(t_i; \gamma)}{\partial \gamma} \right)^2 \right. \right. \\ \left. \left. \theta_i \frac{A''(\theta_i S_0(t_i; \gamma))}{A'(\theta_i S_0(t_i; \gamma))} \frac{\partial^2 S_0(t_i; \gamma)}{\partial \gamma^2} \right] + (1 - \delta_i) \theta_i \frac{c_1 d_1^* - d_1 c_1^*}{c_1^2} \left(\frac{\partial S_0(t_i; \gamma)}{\partial \gamma} \right)^2 + \frac{d_1}{c_1} \frac{\partial^2 S_0(t_i; \gamma)}{\partial \gamma^2} \right\} \quad (\text{C.5})$$

em que

$$d_1 = A'(\theta_i S_0(t_i; \gamma)),$$

$$d_1^* = A''(\theta_i S_0(t_i; \gamma)) \theta_i$$

e

$$c_1^* = A'(\theta_i S_0(t_i; \gamma)) \theta_i$$

$$K_{\phi\gamma} = \sum_{i=1}^n \left\{ -(1 - \delta_i) \theta_i \frac{A'(\theta_i S_0(t_i; \gamma)) A(\theta_i)}{(\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma)))^2} \frac{\partial S_0(t_i; \gamma)}{\partial \gamma} \right\} \quad (C.6)$$

e

$$K_{\phi\phi} = \frac{n}{(1 + \phi)^2} - \sum_{i=1}^n \left\{ \frac{(A(\theta_i))^2}{(\phi A(\theta_i) + A(\theta_i S_0(t_i; \gamma)))^2} \right\}. \quad (C.7)$$

C.3 Vetor escore e matriz de informação observada do modelo ZIP- FC

Logaritmo da função de verossimilhança

$$\begin{aligned} \ell(\boldsymbol{\vartheta}) = -n \log(1 + \phi) + \sum_{i=1}^n \{ & -\theta_i + \delta_i [\log(\theta_i) + \log f_0(t_i; \gamma) + \theta_i S_0(t_i; \gamma)] \\ & (1 - \delta_i) \log[\phi \exp(\theta_i) + \exp(\theta_i S_0(t_i; \gamma))] \}. \end{aligned}$$

Elementos do vetor escore

$$\begin{aligned} U_1(\boldsymbol{\vartheta}) &= \frac{\partial \ell}{\partial \beta} = \sum_{i=1}^n \left\{ -1 + \delta_i \left[\frac{1}{\theta_i} + S_0(t_i; \gamma) \right] + (1 - \delta_i) \frac{\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))} S_0(t_i; \gamma)}{\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}} \right\} \theta_i \mathbf{x}_i^\top, \\ U_2(\boldsymbol{\vartheta}) &= \frac{\partial \ell}{\partial \gamma_1} = \sum_{i=1}^n \left\{ \delta_i \left[\frac{1}{f_0(t_i; \gamma)} \frac{\partial f_0(t_i; \gamma)}{\partial \gamma_1} + \theta_i \frac{\partial S_0(t_i; \gamma)}{\partial \gamma_1} \right] + \frac{(1 - \delta_i) \theta_i e^{(\theta_i S_0(t_i; \gamma))}}{\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}} \frac{\partial S_0(t_i; \gamma)}{\partial \gamma_1} \right\}, \\ U_3(\boldsymbol{\vartheta}) &= \frac{\partial \ell}{\partial \gamma_2} = \sum_{i=1}^n \left\{ \delta_i \left[\frac{1}{f_0(t_i; \gamma)} \frac{\partial f_0(t_i; \gamma)}{\partial \gamma_2} + \theta_i \frac{\partial S_0(t_i; \gamma)}{\partial \gamma_2} \right] + \frac{(1 - \delta_i) \theta_i e^{(\theta_i S_0(t_i; \gamma))}}{\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}} \frac{\partial S_0(t_i; \gamma)}{\partial \gamma_2} \right\}, \\ U_4(\boldsymbol{\vartheta}) &= \frac{\partial \ell}{\partial \phi} = \frac{-n}{1 + \phi} + \sum_{i=1}^n \left\{ \frac{(1 - \delta_i) e^{\theta_i}}{\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}} \right\}, \end{aligned} \quad (C.8)$$

em que

$$\begin{aligned}
\frac{\partial \theta_i}{\partial \boldsymbol{\beta}} &= \theta_i \mathbf{x}_i^\top, \\
\frac{\partial f_0(t_i; \gamma)}{\partial \gamma_1} &= (1 + (1 - \gamma_2) \log t) \gamma_2 \gamma_1 t^{\gamma_1 - 1} e^{-\gamma_2 t^{\gamma_1}}, \\
\frac{\partial S_0(t_i; \gamma)}{\partial \gamma_1} &= -\gamma_2 t^{\gamma_1} \log t e^{-\gamma_2 t^{\gamma_1}}, \\
\frac{\partial f_0(t_i; \gamma)}{\partial \gamma_2} &= (\gamma_1 - \gamma_1 \gamma_2 t^{\gamma_1}) e^{-\gamma_2 t^{\gamma_1}} t^{\gamma_1 - 1}, \\
\frac{\partial S_0(t_i; \gamma)}{\partial \gamma_2} &= -e^{-\gamma_2 t^{\gamma_1}} t^{\gamma_1}.
\end{aligned}$$

Componentes da matriz observada

Componentes da submatriz $K_{\boldsymbol{\beta}\boldsymbol{\gamma}}$

$$\frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} = \sum_{i=1}^n \left\{ \frac{-\delta_i}{\theta_i^2} + (1 - \delta_i) \left[\frac{c_1 \dot{c}_2 - c_2 \dot{c}_1}{c_1^2} \right] \theta_i^2 \mathbf{x}_i \mathbf{x}_i^\top + \mathbf{D} \theta_i \mathbf{x}_i \mathbf{x}_i^\top \right\} \quad (\text{C.9})$$

em que

$$\begin{aligned}
c_1 &= \phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}, \\
c_2 &= \dot{c}_1 = \phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))} S_0(t_i; \gamma), \\
\dot{c}_2 &= \phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))} S_0(t_i; \gamma)^2, \\
\mathbf{D} &= \left\{ -1 + \delta_i \left[\frac{1}{\theta_i} + S_0(t; \gamma) \right] + (1 - \delta_i) \frac{\phi e^{\theta_i} + e^{(\theta_i S_0(t; \gamma))} S_0(t; \gamma)}{\phi e^{\theta_i} + e^{(\theta_i S_0(t; \gamma))}} \right\}.
\end{aligned}$$

$$\frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \gamma_1} = \sum_{i=1}^n \left\{ \delta_i - (1 - \delta_i) \frac{\phi \theta_i e^{\theta_i(1+S_0(t_i; \gamma))} (1 - S_0(t_i; \gamma))}{(\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))})^2} \right\} \frac{\partial S_0(t_i; \gamma)}{\partial \gamma_1} \theta_i \mathbf{x}_i^\top. \quad (\text{C.10})$$

$$\frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \gamma_2} = \sum_{i=1}^n \left\{ \delta_i - (1 - \delta_i) \frac{\phi \theta_i e^{\theta_i(1+S_0(t_i; \gamma))} (1 - S_0(t_i; \gamma))}{(\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))})^2} \right\} \frac{\partial S_0(t_i; \gamma)}{\partial \gamma_2} \theta_i \mathbf{x}_i^\top. \quad (\text{C.11})$$

$$\begin{aligned}
\frac{\partial^2 \ell}{\partial \gamma_1^2} = & \sum_{i=1}^n \left\{ \delta_i \left[\frac{-1}{f_0^2(t_i; \gamma)} \left(\frac{\partial f_0(t_i; \gamma)}{\partial \gamma_1} \right)^2 + \frac{1}{f_0(t_i; \gamma)} \frac{\partial^2 f_0(t_i; \gamma)}{\partial \gamma_1^2} + \theta_i \frac{\partial^2 S_0(t_i; \gamma)}{\partial \gamma_1} \right] \right. \\
& + \theta_i (1 - \delta_i) \left[\frac{e^{(\theta_i S_0(t_i; \gamma))}}{e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}} \frac{\partial^2 S_0(t_i; \gamma)}{\partial \gamma_1^2} \right. \\
& \left. \left. + \frac{[\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}] e^{\theta_i S_0(t_i; \gamma)} - e^{(2\theta_i S_0(t_i; \gamma))}}{(\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))})^2} \theta_i \left(\frac{\partial S_0(t_i; \gamma)}{\partial \gamma_1} \right)^2 \right] \right\}. \tag{C.12}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell}{\partial \gamma_2^2} = & \sum_{i=1}^n \left\{ \delta_i \left[\frac{-1}{f_0^2(t_i; \gamma)} \left(\frac{\partial f_0(t_i; \gamma)}{\partial \gamma_2} \right)^2 + \frac{1}{f_0(t_i; \gamma)} \frac{\partial^2 f_0(t_i; \gamma)}{\partial \gamma_2^2} + \theta_i \frac{\partial^2 S_0(t_i; \gamma)}{\partial \gamma_2} \right] \right. \\
& + \theta_i (1 - \delta_i) \left[\frac{e^{(\theta_i S_0(t_i; \gamma))}}{e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}} \frac{\partial^2 S_0(t_i; \gamma)}{\partial \gamma_2^2} \right. \\
& \left. \left. + \frac{[\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}] e^{\theta_i S_0(t_i; \gamma)} - e^{(2\theta_i S_0(t_i; \gamma))}}{(\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))})^2} \theta_i \left(\frac{\partial S_0(t_i; \gamma)}{\partial \gamma_2} \right)^2 \right] \right\}. \tag{C.13}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell}{\partial \gamma_1 \partial \gamma_2} = & \sum_{i=1}^n \left\{ \delta_i \left[\frac{-1}{f_0^2(t_i; \gamma)} \frac{\partial f_0(t_i; \gamma)}{\partial \gamma_1} \frac{\partial f_0(t_i; \gamma)}{\partial \gamma_2} + \frac{1}{f_0(t_i; \gamma)} \frac{\partial^2 f_0(t_i; \gamma)}{\partial \gamma_2 \partial \gamma_1} + \theta_i \frac{\partial^2 S_0(t_i; \gamma)}{\partial \gamma_2 \partial \gamma_1} \right] \right. \\
& + \theta_i (1 - \delta_i) \left[\frac{\phi \theta_i e^{\theta_i(1+S_0(t_i; \gamma))}}{(\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))})^2} \frac{\partial S_0(t_i; \gamma)}{\partial \gamma_1} \frac{\partial S_0(t_i; \gamma)}{\partial \gamma_2} + \frac{e^{\theta_i(1+S_0(t_i; \gamma))}}{\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}} \frac{\partial^2 S_0(t_i; \gamma)}{\partial \gamma_1 \partial \gamma_2} \right] \right\}. \tag{C.14}
\end{aligned}$$

Componentes da submatriz $K_{\beta\gamma\phi}$

$$\frac{\partial^2 \ell}{\partial \beta \partial \phi} = \sum_{i=1}^n \left\{ (1 - \delta_i) \frac{(\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}) e^{\theta_i} - e^{\theta_i} (\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))}) S_0(t_i; \gamma)}{(\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))})^2} \theta_i \right\} \mathbf{x}_i \mathbf{x}_i^\top. \tag{C.15}$$

$$\frac{\partial^2 \ell}{\partial \gamma_1 \partial \phi} = \sum_{i=1}^n \left\{ (1 - \delta_i) \frac{-e^{\theta_i} e^{(\theta_i S_0(t_i; \gamma))} S_0(t_i; \gamma) \theta_i}{(\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))})^2} \frac{\partial S_0(t_i; \gamma)}{\partial \gamma_1} \right\}. \tag{C.16}$$

$$\frac{\partial^2 \ell}{\partial \gamma_2 \partial \phi} = \sum_{i=1}^n \left\{ (1 - \delta_i) \frac{-e^{\theta_i} e^{(\theta_i S_0(t_i; \gamma))} S_0(t_i; \gamma) \theta_i}{(\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))})^2} \frac{\partial S_0(t_i; \gamma)}{\partial \gamma_2} \right\}. \tag{C.17}$$

Escalar $K_{\phi\phi}$

$$\frac{\partial^2 \ell}{\partial \phi^2} = \frac{-n}{(1+\phi)^2} - \sum_{i=1}^n \left\{ (1-\delta_i) \frac{e^{-2\theta_i}}{(\phi e^{\theta_i} + e^{(\theta_i S_0(t_i; \gamma))})^2} \right\}. \quad (\text{C.18})$$

Referências Bibliográficas

- Aalen, O. O. & Hjort, N. L. (2002). Frailty models that yield proportional hazards. *Statistics & Probability Letters*, **58**(4), 335–342. [48](#)
- Aalen, O. O. & Tretli, S. (1999). Analyzing incidence of testis cancer by means of a frailty model. *Cancer Causes & Control*, **10**(4), 285–292. [48](#)
- Adamidis, K. & Loukas, S. (1998). A lifetime distribution with decreasing failure rate. *Statistics & Probability Letters*, **39**(1), 35–42. [52](#)
- Agarwal, D., Gelfand, A. & Citron-Pousty, S. (2002). Zero-inflated models with application to spatial count data. *Environmental and Ecological Statistics*, **9**(4), 341–355. [1](#)
- Akaike, H. (1973). *Information theory and an extension of the maximum likelihood principle*, volume Volume 1, pages 267–281. Akademiai Kiado. [23](#)
- Ata, N. & Özel, G. (2013). Survival functions for the frailty models based on the discrete compound poisson process. *Journal of Statistical Computation and Simulation*, **83**(11), 2105–2116. [48](#)
- Barreto-Souza, W., De Moraes, A. & Cordeiro, G. (2010). The Weibull-geometric distribution. *Journal of Statistical Computation and Simulation*, **9**(1), 1–13. [52](#)
- Barriga, G. D. & Louzada, F. (2014). The zero-inflated conway–maxwell–poisson distribution: Bayesian inference, regression modeling and influence diagnostic. *Statistical Methodology*, **21**, 23–34. [1](#), [80](#)
- Berkson, J. & Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, **47**(259), 501–515. [52](#)
- Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society B*, **11**(1), 15–53. [52](#)
- Böhning, D., Dietz, E., Schlattmann, P., Mendonça, L. & Kirchner, U. (1999). The zero-inflated poisson model and the decayed, missing and filled teeth index in dental epidemiology. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **162**(2), 195–209. [1](#)
- Brooks, S. P. (2002). Discussion on the paper by Spiegelhalter, Best, Carlin, and van der Linde (2002). *Journal of the Royal Statistical Society B*, **64**, 616–618. [76](#)
- Cancho, V., Rodrigues, J. & de Castro, M. (2011). A flexible model for survival data with a cure rate: a Bayesian approach. *Journal of Applied Statistics*, **38**(1), 57–70. [50](#)

- Cancho, V. G., Rodrigues, J. & de Castro, M. (2010). A flexible model for survival data with a cure rate: a Bayesian approach. Accepted for publication in *Journal Applied of Statistics*. [50](#)
- Cancho, V. G., Louzada, F. & Ortega, E. M. (2013). The power series cure rate model: an application to a cutaneous melanoma data. *Communications in Statistics-Simulation and Computation*, **42**(3), 586–602. [5](#), [51](#), [58](#), [66](#), [70](#), [79](#)
- Cancho, V. G., Dey, D. K. & Louzada, F. (2016). Unified multivariate survival model with a surviving fraction: an application to a brazilian customer churn data. *Journal of Applied Statistics*, **43**(3), 572–584. [50](#)
- Caroni, C., Crowder, M. & Kimber, A. (2010). Proportional hazards models with discrete frailty. *Lifetime Data Analysis*, **16**(3), 374–384. [3](#), [48](#), [51](#), [66](#)
- Carrivick, P. J., Lee, A. H. & Yau, K. K. (2003). Zero-inflated poisson modeling to evaluate occupational safety interventions. *Safety Science*, **41**(1), 53–63. [42](#)
- Casella, G. & George, E. I. (1992). Explaining the gibbs sampler. *The American Statistician*, **46**(3), 167–174. [69](#)
- Chahkandi, M. & Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Comput. Stat. Data Anal.*, **53**(12), 4433–4440. [52](#)
- Chen, M.-H., Ibrahim, J. G. & Sinha, D. (1999). A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*, **94**(447), 909–919. [53](#), [73](#)
- Chib, S. & Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The american statistician*, **49**(4), 327–335. [69](#)
- Cho, H., Ibrahim, J. G., Sinha, D. & Zhu, H. (2009). Bayesian case influence diagnostics for survival models. *Biometrics*, **65**(1), 116–124. [71](#)
- Choo-Wosoba, H., Levy, S. M. & Datta, S. (2015). Marginal regression models for clustered count data based on zero-inflated conway–maxwell–poisson distribution with applications. *Biometrics*. [1](#)
- Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65**(1), 141–151. [48](#)
- Consul, P. C. & Jain, G. C. (1973). A generalization of the poisson distribution. *Technometrics*, **15**(4), 791–799. [1](#)
- Cook, R. D. & Weisberg, S. (1982). Residuals and influence in regression. [70](#)
- Cooner, F., Banerjee, S. & McBean, A. M. (2006). Modelling geographically referenced survival data with a cure fraction. *Statistical Methods in Medical Research*, **15**, 307–324. [50](#)
- Cooner, F., Banerjee, S., Carlin, B. P. & Sinha, D. (2007). Flexible cure rate modeling under latent activation schemes. *Journal of the American Statistical Association*, **102**, 560–572. [50](#)

- Cordeiro, G. M., Cancho, V. G., Ortega, E. M. & Barriga, G. D. (2016). A model with long-term survivors: Negative binomial birnbaum-saunders. *Communications in Statistics-Theory and Methods*, **45**(5), 1370–1387. [50](#)
- Cowles, M. K. & Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association*, **91**, 883–904. [72](#)
- Cox, D. R. & Hinkley, D. V. (1979). *Theoretical statistics*. CRC Press. [8](#), [57](#), [64](#)
- de Castro, M., Cancho, V. G. & Rodrigues, J. (2009). A Bayesian long-term survival model parametrized in the cured fraction. *Biometrical Journal*, **51**, 443–455. [50](#)
- de Castro, M., Cancho, V. & Rodrigues, J. (2010). A hands-on approach for fitting long-term survival models under the gamlss framework. [50](#)
- del Castillo, J. & Pérez-Casany, M. (2005). Overdispersed and underdispersed poisson generalizations. *Journal of Statistical Planning and Inference*, **134**(2), 486–500. [1](#)
- Deng, D. & Zhang, Y. (2015). Score tests for both extra zeros and extra ones in binomial mixed regression models. *Communications in Statistics-Theory and Methods*, **44**(14), 2881–2897. [1](#)
- Dey, D. K. & Birmiwala, L. R. (1994). Robust bayesian analysis using divergence measures. *Statistics & Probability Letters*, **20**(4), 287–294. [71](#)
- dos Santos, D. M., Davies, R. B. & Francis, B. (1995). Nonparametric hazard versus nonparametric frailty distribution in modelling recurrence of breast cancer. *Journal of Statistical Planning and Inference*, **47**(1), 111–127. [48](#)
- Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association*, **76**(374), 312–319. [58](#)
- Efron, B. & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science*, pages 54–75. [58](#)
- Eudes, A. M., TOMAZELLA, V. L. D. & CALSAVARA, V. F. (2013). Modelagem de sobrevivência com frac ao de cura para dados de tempo de vida weibull modificada. *Rev. Bras. Biom*, **30**(3), 326–342. [50](#)
- Geisser, S. & Eddy, W. F. (1979). A predictive approach to model selection. *Journal of the American Statistical Association*, **74**(365), 153–160. [69](#)
- Gelfand, A. E. & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85**(410), 398–409. [67](#)
- Greenwood, M. & Yule, G. U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society*, pages 255–279. [47](#)
- Gupta, P. L., Gupta, R. C. & Tripathi, R. C. (1995). Inflated modified power series distributions with applications. *Communications in Statistics-Theory and Methods*, **24**(9), 2355–2374. [1](#), [2](#)

- Gupta, P. L., Gupta, R. C. & Tripathi, R. C. (1996). Analysis of zero-adjusted count data. *Computational Statistics & Data Analysis*, **23**(2), 207–218. [1](#)
- Gupta, P. L., Gupta, R. C. & Tripathi, R. C. (2005). Score test for zero inflated generalized poisson regression model. *Communications in Statistics-Theory and Methods*, **33**(1), 47–64. [1](#)
- Hall, D. (2000). Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics*, **56**(4), 1030–1039. [1](#)
- Hougaard, P. (1984). Life table methods for heterogeneous populations: distributions describing the heterogeneity. *Biometrika*, **71**(1), 75–83. [48](#)
- Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika*, **73**(3), 671–678. [3](#), [48](#)
- Hsu, W.-W., Todem, D., Kim, K. & Sohn, W. (2014). A wald test for zero inflation and deflation for correlated count data from dental caries research. *Statistical Modelling*, **14**(6), 471–488. [1](#)
- Ibrahim, J. G., Chen, M.-H. & Sinha, D. (2001). *Bayesian Survival Analysis*. Springer, New York. [3](#), [50](#), [61](#)
- Kass, R. E., Tierney, L. & Kadane, J. B. (1991). Laplace’s method in bayesian analysis. *Contemporary Mathematics*, **115**, 89–99. [69](#)
- Kirkwood, J. M., Ibrahim, J. G., Sondak, V. K., Richards, J., Flaherty, L. E., Ernstoff, M. S., Smith, T. J., Rao, U., Steele, M. & Blum, R. H. (2000). High- and low-dose interferon alfa-2b in high-risk melanoma: First analysis of intergroup trial E1690/S9111/C9190. *Journal of Clinical Oncology*, **18**, 2444–2458. [61](#)
- Kus, C. (2007). A new lifetime distribution distributions. *Computational Statistics & Data Analysis*, **11**, 4497–4509. [52](#)
- Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics*, **34**(1), 1–14. [1](#)
- Lange, K. (1999). *Numerical Analysis for Statisticians*. Springer, New York, NY. [62](#)
- Lemonte, A. (2016). *The Gradient Test: Another Likelihood-Based Test*. Academic Press. [4](#), [5](#)
- Lemonte, A. J. & Ferrari, S. L. (2011). Testing hypotheses in the birnbaum–saunders distribution under type-ii censored samples. *Computational Statistics & Data Analysis*, **55**(7), 2388–2399. [4](#)
- Lemonte, A. J. & Ferrari, S. L. (2012). Local power and size properties of the lr, wald, score and gradient tests in dispersion models. *Statistical Methodology*, **9**(5), 537 – 554. [4](#)
- Lemonte, A. J., Cordeiro, G. M. & Moreno-Arenas, G. (2016). Improved likelihood-based inference in birnbaum–saunders nonlinear regression models. *Applied Mathematical Modelling*, **40**(19–20), 8185 – 8200. [4](#)
- Li, C.-S. (2012). Score tests for semiparametric zero-inflated poisson models. *International Journal of Statistics and Probability*, **1**(2), 1. [80](#)

- Li, C. S., Taylor, J. & Sy, J. (2001). Identifiability of cure models. *Statistics and Probability Letters*, **54**, 389–395. [52](#), [56](#)
- Li, H. & Zhong, X. (2002). Multivariate survival models induced by genetic frailties, with application to linkage analysis. *Biostatistics*, **3**(1), 57–75. [48](#)
- Liang, K.-Y., Self, S. G., Bandeen-Roche, K. J. & Zeger, S. L. (1995). Some recent developments for regression analysis of multivariate failure time data. *Lifetime data analysis*, **1**(4), 403–415. [80](#)
- Macera, M. A., Louzada, F., Cancho, V. G. & Fontes, C. J. (2015). The exponential-poisson model for recurrent event data: An application to a set of data on malaria in brazil. *Biometrical Journal*, **57**(2), 201–214. [50](#), [80](#)
- Maller, R. A. & Zhou, X. (1996). *Survival Analysis with Long-Term Survivors*. Wiley, New York. [3](#), [49](#), [50](#)
- Milani, E. A., Tomazella, V. L., Dias, T. C., Louzada, F. *et al.* (2015). The generalized time-dependent logistic frailty model: An application to a population-based prospective study of incident cases of lung cancer diagnosed in northern ireland. *Brazilian Journal of Probability and Statistics*, **29**(1), 132–144. [50](#)
- Moger, T. A., Aalen, O. O., Halvorsen, T. O., Storm, H. H. & Tretli, S. (2004). Frailty modelling of testicular cancer incidence using scandinavian data. *Biostatistics*, **5**(1), 1–14. [48](#)
- Morita, L. H. M., Tomazella, V. L. & Louzada-Neto, F. (2016). Accelerated lifetime modelling with frailty in a non-homogeneous poisson process for analysis of recurrent events data. *Quality Technology & Quantitative Management*, pages 1–21. [50](#)
- Nian, G. (2014). *A Score Test of Homogeneity in Generalized Additive Models for Zero-Inflated Count Data*. Ph.D. thesis, Kansas State University. [80](#)
- Ortega, E. M., Cordeiro, G. M., Campelo, A. K., Kattan, M. W. & Cancho, V. G. (2015). A power series beta weibull regression model for predicting breast carcinoma. *Statistics in medicine*, **34**(8), 1366–1388. [50](#)
- Patil, M. & Shirke, D. (2007). Testing parameter of the power series distribution of a zero inflated power series model. *Statistical Methodology*, **4**(4), 393–406. [1](#)
- Peng, F. & Dey, D. (1995). Bayesian analysis of outlier problems using divergence measures. *Canadian Journal of Statistics*, **23**(2), 199–213. [71](#)
- R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. [16](#), [23](#), [57](#)
- Rao, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, **44**(1), 50–57. [i](#), [ii](#), [4](#)
- Ridout, M., Hinde, J. & Demétrio, C. (2001). A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics*, **57**(1), 219–223. [1](#), [42](#), [44](#)

- Rodrigues, J., Cancho, V. G., de Castro, M. & Louzada-Neto, F. (2009a). On the unification of the long-term survival models. *Statistics & Probability Letters*, **79**, 753–759. [3](#), [50](#)
- Rodrigues, J., de Castro, M., Cancho, V. G. & Balakrishnan, N. (2009b). COM-Poisson cure rate survival models and an application to a cutaneous melanoma data. *Journal of Statistical Planning Inference*, **139**, 3605–3611. [49](#), [50](#)
- Rodrigues, J., Cordeiro, G. M., Cancho, V. G. & Balakrishnan, N. (2015). Relaxed poisson cure rate models. *Biometrical Journal*. [50](#)
- Samani, E. B., Amirian, Y. & Ganjali, M. (2012). Likelihood estimation for longitudinal zero-inflated power series regression models. *Journal of Applied Statistics*, **39**(9), 1965–1974. [1](#)
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461–464. [23](#)
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B*, **64**, 583–639. [76](#)
- Sun, F.-B. & Kececiloglu, D. (1999). A new method for obtaining the ttt plot for a censored sample. In *Reliability and Maintainability Symposium, 1999. Proceedings. Annual*, pages 112–117. IEEE. [62](#), [90](#)
- Tahmasbi, R. & Rezaei, S. (2008). A two-parameter lifetime distribution with decreasing failure rate. *Computational Statistics & Data Analysis*, **52**(8), 3889–3901. [52](#)
- Terrell, G. R. (2002). The gradient statistic. *Computing Science and Statistics*, **34**(34), 206–215. [i](#), [ii](#), [1](#), [4](#), [82](#)
- Todem, D., Hsu, W.-W. & Kim, K. (2012). On the efficiency of score tests for homogeneity in two-component parametric models for discrete data. *Biometrics*, **68**(3), 975–982. [1](#), [3](#)
- Tournoud, M. & Ecochard, R. (2007). Application of the promotion time cure model with time-changing exposure to the study of HIV/AIDS and other infectious diseases. *Statistics in Medicine*, **26**, 1008–1021. [50](#)
- Tsodikov, A. D., Ibrahim, J. G. & Yakovlev, A. Y. (2003). Estimating cure rates from survival data: An alternative to two-component mixture models. *Journal of the American Statistical Association*, **98**, 1063–1078. [49](#)
- Van den Broek, J. (1995). A score test for zero inflation in a poisson distribution. *Biometrics*, pages 738–743. [1](#), [7](#)
- Vaupel, J. W., Manton, K. G. & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, **16**(3), 439–454. [3](#), [47](#), [48](#)
- Vu, H. T. & Knuiman, M. W. (2002). A hybrid ml-em algorithm for calculation of maximum likelihood estimates in semiparametric shared frailty models. *Computational statistics & data analysis*, **40**(1), 173–187. [48](#)
- Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, **54**(3), 426–482. [i](#), [ii](#), [4](#)

- Weiss, R. (1996). An approach to bayesian sensitivity analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 739–750. [71](#)
- Weisstein, E. W. (2002). Lambert w-function. [11](#)
- Wienke, A. (2010). *Frailty models in survival analysis*. CRC Press. [48](#)
- Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Statist.*, **9**(1), 60–62. [i](#), [ii](#), [4](#), [82](#)
- Xiang, L., Lee, A. H., Yau, K. K. & McLachlan, G. J. (2006). A score test for zero-inflation in correlated count data. *Statistics in Medicine*, **25**(10), 1660–1671. [1](#)
- Xue, X. & Brookmeyer, R. (1997). Regression analysis of discrete time survival data under heterogeneity. *Statistics in Medicine*, **16**(17), 1983–1993. [48](#)
- Yakovlev, A., Yu, A. B., Bardou, V.-J., Fourquet, A., Hoang, T., Rochefodiere, A. & Tso-dikov, A. D. (1993). A simple stochastics model of tumor recurrence an its aplications to data on premenopausal breast cancer. In S. F. de Biometrie, editor, *Biometrie et Analyse de Donnes Spatio-Temporelles No 12, B*, pages 33–82, France. [5](#), [65](#), [66](#)
- Yang, Z., Hardin, J. W., Addy, C. L. & Vuong, Q. H. (2007). Testing approaches for overdispersion in poisson regression versus the generalized poisson model. *Biometrical journal*, **49**(4), 565–584. [1](#)
- Zhao, Y., Lee, A. H., Yau, K. K. W. & Burke, V. (2009). A score test for assessing the cured proportion in the long-term survivor mixture model. *Statistics in Medicine*, **28**, 3454–3466. [50](#)
- Zhou, X. & Tu, W. (2000). Confidence intervals for the mean of diagnostic test charge data containing zeros. *Biometrics*, **56**(4), 1118–1125. [1](#)