

Regressão para Dados Censurados sob Mistura da Distribuição

Gaussiana Inversa com sua Recíproca Complementar

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pelo Sr. Víctor Manuel Maehara Oyata ^{7/16} e aprovada pela Comissão Julgadora.

Campinas, 19 de agosto de 1994

Jonathan Biele

Prof. Dr. Jonathan Biele *t*

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência de Computação, UNICAMP, como requisito parcial para obtenção do Título de MESTRE em Estatística

A minha esposa Doris
e
filhos: Rocío e Yoshitomi.

Agradecimentos

Gostaria de expressar meus agradecimentos

- Aos professores do departamento de estatística pela contribuição à minha formação.
- Ao professor Jonathan, pela orientação, dedicação e amizade, que foram essenciais para a realização deste trabalho.
- À professora Clarice pela amizade, importantes sugestões e ajuda na correção da presente tese.
- A Rissa pela amizade e ajuda em algumas correções do português
- Aos funcionários do departamento pela colaboração.
- Aos amigos que conquistei durante o mestrado, pelo companheirismo e apoio.
- A minha amada esposa Doris pelo apoio e compreensão, principalmente nas horas mais difíceis.
- A CAPES, pelo apoio financeiro.

SUMÁRIO

Introdução	viii
1. Distribuição Gaussiana Inversa	1
1.1 Introdução	1
1.2 Funções de densidade	2
1.3 Algumas propriedades importantes	6
1.4 Modelo Exponencial com Dispersão	10
1.5 Divisibilidade Infinita	12
2. Distribuição da Recíproca Complementar de uma Variável Gaus-	
siana Inversa	16
2.1 Introdução	16
2.2 Função de densidade e sua relação com a densidade da Gaussiana Inversa	17
2.3 Algumas propriedades importantes	19
3. Mistura da Distribuição Gaussiana Inversa Com sua Recíproca	
Complementar	26
3.1 Introdução	26
3.2 Função de densidade	26
3.3 Função de Distribuição	29
3.4 Propriedades desta Família de Mistura	30

3.5	Aspectos relacionados aos Modelos Exponenciais com Dispersão e Divisibilidade Infinita	41
4.	Inferência sobre os Parâmetros da Mistura de Distribuições Gaussianas Inversas com sua Recíproca Complementar	50
4.1	Introdução	50
4.2	Estimação Pontual de Parâmetros para Dados Sem Censura	50
4.3	Estimação por Intervalo e Testes de Hipóteses para Dados sem Censura	58
4.3.1	Para os parâmetros da distribuição Gaussiana Inversa	58
4.3.2	Para os parâmetros da distribuição da Recíproca Complementar de uma variável Gaussiana Inversa	61
4.3.3	Teste de hipóteses e estimação por intervalo para os parâmetros da <i>M-IG</i>	61
4.4	Exemplo 4.1-Resistência de rolamentos	66
4.5	Dados Censurados	69
4.6	Estimação de Parâmetros com Dados Censurados	71
4.7	Testes de Hipóteses e Estimação por Intervalo com Dados Censurados	89
4.7.1	Teste de Hipóteses e Estimação por Intervalo Para os Parâmetros da IG ($p = 0$)	90
4.7.2	Teste de Hipóteses para os Parâmetros da R-IG ($p = 1$)	92
4.7.3	Teste de Hipóteses e Estimação por Intervalos sobre os Parâmetros da M-IG	95
4.8	Exemplo 4.2. Tempo de remissão de pacientes com Leucemia	99
5.	Um Modelo de Regressão para Dados da Mistura de Gaussianas Inversas com sua Recíproca Complementar	104

5.1	Introdução	104
5.2	Modelo de Regressão	104
5.3	Estimação de Parâmetros para Dados sem Censura	107
5.4	Estimação com dados censurados	117
5.5	Adequação do modelo	136
5.6	Teste de Hipóteses e Estimação por Intervalo para Dados sem Censura	137
5.6.1	Teste de Hipóteses e Estimação por Intervalos para os Parâmetros da Regressão Gaussiana Inversa	137
5.6.2	Teste de Hipóteses e Estimação por Intervalo para os Parâmetros da Regressão da R-IG	138
5.6.3	Teste de Hipóteses sobre os Parâmetros da Regressão para da- dos da M-IG	141
5.7	Testes de Hipóteses e Estimação por Intervalo com Dados Censurados	143
5.7.1	Teste de Hipóteses e Estimação por Intervalo para os Parâmetros da Regressão IG	143
5.7.2	Testes de Hipóteses para os Parâmetros da R-IG	146
5.7.3	Teste de Hipóteses e Estimação por Intervalos sobre os Parâmetros da M-IG	149
5.8	Exemplos	153
5.8.1	Exemplo 1-Teste de Vida de Isolantes de Fluido Elétrico . . .	153
5.8.2	Exemplo 2- Teste de Vida de Isolantes Elétricos	157
6.	Considerações Finais	162
	APÊNDICE	164
	BIBLIOGRAFIA	178

BIBLIOGRAFIA 183

LISTA DE FIGURAS

1.1	Gráfico da função de densidade f_4 para $\lambda = 0.2$ fixo e valores de ψ diferentes.	4
1.2	Gráfico da função de densidade f_4 para $\psi = 1$ fixo e diferentes valores de λ	5
1.3	Gráfico da função risco para alguns valores de μ e ν	9
2.1	Gráfico da função de densidade g para $\nu = 2$ fixo e diferentes valores de μ	18
2.2	Gráfico da função de densidade g para $\mu = 5$ fixo e diferentes valores de ν	19
2.3	Gráfico da função risco para alguns valores de μ e de ν	23
3.1	Gráfico da função de densidade g para $\nu = 0.2$ e $p = 0.6$ e vários valores de μ	28
3.2	Gráfico da função de densidade g para $\mu = 0.2$ e $p = 0.6$ e vários valores de ν	29
3.3	Gráfico da função risco para alguns valores de μ , ν e de p	32
3.4	Gráfico da função de densidade g para $\mu = 20$ e $\nu = 15$ e três valores de p	37

3.5	Gráfico da função de densidade g para $\mu = 20$ e $\nu = 5$ e três valores de p	38
3.6	Gráfico da função de densidade g para $\mu = 20$ e $\nu = 0.1$ e dois valores de p	39
4.1	Gráfico da função de sobrevivência estimada por Kaplan Meier e de máxima verossimilhança sob IG., para dados de resistência de rolamentos.	68
4.2	Gráfico da função de sobrevivência estimada por Kaplan Meier e de máxima verossimilhança sob R-IG., para dados de resistência de rolamentos.	69
4.3	Gráfico da função de sobrevivência estimada por Kaplan Meier e de máxima verossimilhança sob IG., para dados de tempos de remissão de pacientes com leucemia.	101
4.4	Gráfico da função de sobrevivência estimada por Kaplan Meier e de máxima verossimilhança sob M-IG com $p = 0.824$, para dados de tempos de remissão de pacientes com leucemia. . .	102
4.5	Gráfico da função de sobrevivência estimada por Kaplan Meier e de máxima verossimilhança sob R-IG., para dados de tempos de remissão de pacientes com leucemia.	103
5.1	Gráfico dos resíduos, (a) regressão IG, (b) regressão R-IG. . .	156
5.2	Gráfico dos resíduos, (a) regressão IG, (b) regressão R-IG. . .	160

LISTA DE TABELAS

4.1	Número de revoluções (em milhões) até a falha ocorra em teste de rolamento.	66
4.2	Estimativas de máxima verossimilhança para dados da resistência de rolamentos, para diferentes valores de p	67
4.3	Tempo de remissão (em semanas) de pacientes com leucemia, (+) obsevarções censuradas.	99
4.4	Estimativas de máxima verossimilhança para dados de tempos de remissão de pacientes com leucemia, para diferentes valores de p	100
5.1	Dados de tempos de vida de isolantes de fluido elétrico	154
5.2	Estimativas de máxima veross. para vários valores de p	155
5.3	Tempo de vida de isolantes (horas) para cada temperatura, em graus centígrados, (+) observações censuradas.	158
5.4	Estimativas de máxima verossimilhança para vários valores de p	159

Introdução

A mistura de distribuições Gaussiana Inversa com sua Recíproca Complementar (M-IG) foi apresentada por Jørgensen, Seshadri e Whitmore (1991). Esta mistura de distribuições possui propriedades interessantes, algumas delas semelhantes à distribuição Gaussiana Inversa ou à Recíproca Complementar de uma variável Gaussiana Inversa, casos particulares desta mistura. Algumas destas propriedades estão relacionadas a distribuição Normal, tais como a função de distribuição acumulada pode ser expressa em termos da função de distribuição acumulada da Normal Padrão, tem uma relação assintótica com à distribuição normal, assim como cumprem em forma semelhantes com algumas de suas propriedades.

Devido a estas propriedades interessantes, satisfeitas por esta mistura de distribuições e da mesma forma que a distribuição Gaussiana Inversa ou a Recíproca Complementar de uma variável Gaussiana Inversa, constituem uma alternativa para modelar o tempo de falha, assim como Weibull, Lognormal, Gama, etc, distribuições úteis na engenharia e medicina. Esta mistura é uma alternativa para modelar tempo de falha nas seguintes situações: (1) quando existe evidência de que as taxas de falha não são proporcionais, (2) quando o logaritmo dos tempos não se ajustam a um modelo de locação-escala, (3) quando existe evidência de que os tempos de falha possam ser interpretado como a soma de duas variáveis independentes, uma Bernoulli

Composta e a outra Gaussiana Inversa, (4) quando por experiência ou natureza de fenômeno parece que os dados possam ser modelados por esta mistura.

Pelas razões expostas, neste trabalho é feito um estudo desta mistura de distribuições e realizada uma extensão do trabalho apresentado por Jørgensen, Seshadri e Whitmore (1991) para caso de dados censurados e com covariáveis. Assim, no Capítulo 1, da presente dissertação, faz-se uma revisão da distribuição Gaussiana Inversa, apresentam-se as diferentes reparametrizações de sua função de densidade e suas propriedades mais importantes. No Capítulo 2, define-se a distribuição da Recíproca Complementar de uma variável Gaussiana Inversa; apresenta-se sua função de densidade e sua relação com as diferentes reparametrizações da densidade da Gaussiana Inversa e suas principais propriedades, as quais foram derivadas das propriedades da Mistura de Distribuições da Gaussiana Inversa com sua Recíproca Complementar e da distribuição da Recíproca de uma variável Gaussiana Inversa. No Capítulo 3, obtém-se a função de densidade da mistura e sua função de distribuição acumulada, discute-se suas propriedades mais importantes, demonstra-se que é uma convolução de duas variáveis com Modelos Exponenciais com Dispersão e que esta mistura é infinitamente divisível. No Capítulo 4, apresentam-se procedimentos de inferência para uma amostra com e sem censura; neste capítulo foi usada uma reparametrização diferente à utilizada por Jørgensen, Seshadri e Whitmore (1991), quem só apresenta procedimentos de inferência para uma amostra não censurada. No Capítulo 5, propõe-se um modelo de regressão para dados da M-IG, onde assume-se que a recíproca do parâmetro μ depende linearmente das covariáveis. Desenvolvem-se ainda, procedimentos de inferência para dados com e sem censura.

1. Distribuição Gaussiana Inversa

1.1 Introdução

Schrodinger (1915) apresentou pela primeira vez a distribuição Gaussiana Inversa, ao obter a função de densidade do tempo da primeira passagem em Movimento Browniano (“Brownian Motion”) com tendência positiva. Esta distribuição foi posteriormente estudada por Tweedie e Wald. Assim, Tweedie(1956) propôs o nome de Gaussiana Inversa pela relação existente entre as funções geratrizes de cumulantes do tempo para percorrer uma unidade de distância e da distância percorrida por uma unidade, em Movimento Browniano com tendência positiva. Também, Wald (1947) obteve a mesma função de densidade como limite da distribuição do tamanho da amostra no teste de razão de verossimilhança sequencial, por isso na literatura russa, também se conhece como a distribuição de Wald.

O primeiro estudo detalhado desta distribuição foi apresentado por Tweedie (1957a, 1957b), onde estabelece muitas propriedades estatísticas importantes. Posteriormente, Folks e Chhikara (1978) fazem uma revisão sobre esta distribuição, ao apresentar suas principais propriedades e aplicações. Um estudo mais completo pode ser encontrado em Chhikara e Folks (1989).

A família paramétrica de distribuições Gaussiana Inversa é uma alternativa para modelar tempos de falha, assim como Weibull, Lognormal, Gama, etc, distribuições

comumente usadas na engenharia e medicina. Esta distribuição pode ser útil quando os dados ou natureza do fenômeno em estudo, mostram evidência de que não se cumprem uma ou ambas das duas situações: (1) O logaritmo dos tempos não se ajustam a um modelo locação-escala, por exemplo a distribuição Valor Extremo (distribuição do log de uma variável Weibull) e Lognormal, (ver Lawless (1982)); (2) Riscos proporcionais, suposição fundamental do modelo de Cox (ver Miller (1982)). Assim, num estudo com dados reais, tendo evidência que a condição (1) e/ou (2) não são satisfeitas, baseados na natureza dos dados ou considerações empíricas do fenômeno em estudo, então a distribuição Gaussiana Inversa pode ser considerada como um modelo alternativo. Para a distribuição Gaussiana Inversa existem muitas aplicações, algumas são baseadas em dados de tempos relacionados à existência de movimento Browniano. Outras aplicações, simplesmente se justificam pela qualidade de ajuste dos dados. Exemplos de ambos tipos de aplicações podem ser encontrado em Folks e Chhikara (1989).

Neste capítulo, apresenta-se um pequeno resumo sobre esta distribuição. Assim, na Seção 1.1 são mostradas as diferentes parametrizações da função de densidade e as relações entre elas, na Seção 1.2 as suas principais propriedades, na Seção 1.3 se demonstra que a distribuição é um Modelo Exponencial com Dispersão; e por último, na Seção 1.4 se prova que uma variável Gaussiana Inversa é infinitamente divisível.

1.2 Funções de densidade

De acordo com Tweedie (1957), uma variável aleatória contínua Y tem uma distribuição Gaussiana Inversa com dois parâmetros reais positivas se tem função de densidade escrita em qualquer uma das formas seguintes

$$f_1(y; \gamma, \lambda) = \left(\frac{\lambda}{2\pi y^3} \right)^{1/2} \exp \left[-\gamma \lambda y + \lambda (2\gamma)^{1/2} - \frac{\lambda}{2y} \right] \quad (1.1)$$

$$f_2(y; \mu, \lambda) = \left(\frac{\lambda}{2\pi y^3} \right)^{1/2} \exp \left[-\frac{\lambda(y - \mu)^2}{2\mu^2 y} \right] \quad (1.2)$$

$$f_3(y; \mu, \psi) = \left(\frac{\mu\psi}{2\pi y^3} \right)^{1/2} \exp \left[-\frac{\psi y}{2\mu} + \psi - \frac{\mu\psi}{2y} \right] \quad (1.3)$$

$$f_4(y; \psi, \lambda) = \left(\frac{\lambda}{2\pi y^3} \right)^{1/2} \exp \left[-\frac{\psi^2 y}{2\lambda} + \psi - \frac{\lambda}{2y} \right], \quad (1.4)$$

onde $y > 0$ e os valores $\gamma, \lambda, \mu, \psi$ satisfazem as relações

$$(2\gamma)^{-1/2} = \mu = \frac{\lambda}{\psi}.$$

As funções de densidades dadas em (1.2), (1.3) e (1.4) satisfazem as seguintes relações

$$f_2(y; \mu, \lambda) = \mu^{-1} f_3\left(\frac{y}{\mu}; 1, \psi\right) = \lambda^{-1} f_4\left(\frac{y}{\lambda}; \psi, 1\right).$$

Cada função é utilizada para propósitos específicos na teoria de Movimento Browniano. Pode-se observar, na função f_4 que a forma da distribuição depende de ψ e a escala λ . Assim ψ e λ são os parâmetros de forma e escala, respectivamente. Os gráficos das funções de densidade para alguns valores de ψ e λ estão nas Figuras 1.1 e 1.2.

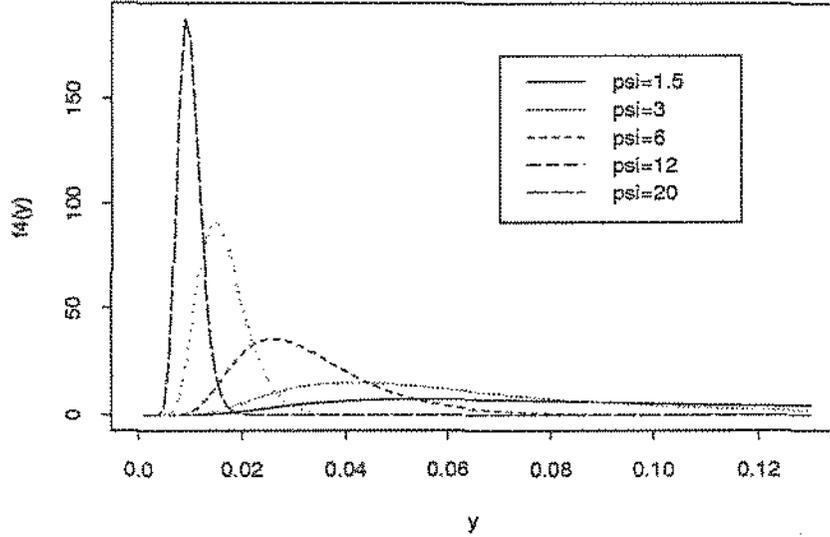


Figura 1.1: Gráfico da função de densidade f_4 para $\lambda = 0.2$ fixo e valores de ψ diferentes.

Ao considerar a seguinte reparametrização

$$\delta = \mu^{-1} \text{ e } \nu = \lambda^{-1},$$

a função de densidade pode ser escrita como

$$f_5(y; \delta, \nu) = (2\pi y^3 \nu)^{-1/2} \exp \left[-\frac{(1 - \delta y)^2}{2\nu y} \right]. \quad (15)$$

Esta reparametrização dada por Whitmore (1983), representa a função de densidade do tempo da primeira passagem para absorver uma barreira, localizada a uma unidade de distância da origem num processo de Wiener com parâmetro de tendência $\delta > 0$ e parâmetro de inconstância $\nu > 0$.

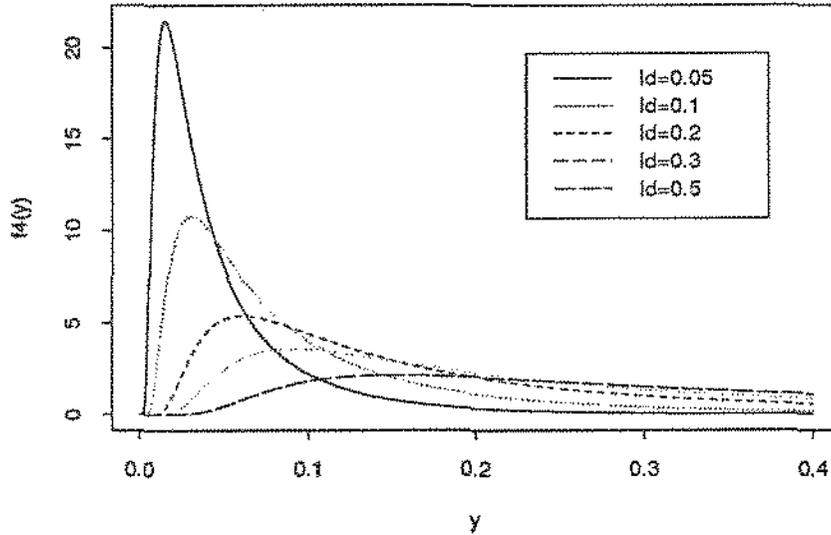


Figura 1.2: Gráfico da função de densidade f_4 para $\psi = 1$ fixo e diferentes valores de λ .

Para o desenvolvimento da presente dissertação, será considerada a seguinte função de densidade

$$f_6(y; \mu, \nu) = (2\pi y^3 \nu)^{-1/2} \exp \left[-\frac{(y - \mu)^2}{2\nu \mu^2 y} \right]. \quad (1.6)$$

Esta função é relacionada com a função de densidade f_2 da seguinte forma

$$f_6(y; \mu, \nu) = f_2(y, \mu, \nu^{-1}).$$

Para indicar que uma variável aleatória Y tem uma distribuição Gaussiana Inversa, com a função de densidade dada em (1.6) será usada a seguinte notação: $Y \sim IG(\mu, \nu)$.

1.3 Algumas propriedades importantes

Nesta seção considera-se algumas das propriedades básicas da distribuição Gaussiana Inversa. Estas e outras propriedades podem ser encontradas de forma resumida em Johnson e Kotz (1970) e também em Chhikara e Folks (1978).

Teorema 1.1

Seja $Y \sim IG(\mu, \nu)$. Então a função geratriz de momentos pode ser escrita como

$$M_Y(t) = E(e^{Yt}) = \exp\left\{(\mu\nu)^{-1} \left[1 - (1 - 2\mu^2\nu t)^{1/2}\right]\right\}. \quad (1.7)$$

□

Teorema 1.2

Seja $Y \sim IG(\mu, \nu)$. Então

i) A média de Y é

$$E(Y) = \mu. \quad (1.8)$$

ii) A variância é

$$Var(Y) = \mu^3\nu. \quad (1.9)$$

ii) A moda é

$$y_m = -\frac{3}{2}\mu^2\nu + \mu \left(1 + \frac{9}{4}\mu^2\nu^2\right)^{1/2}. \quad (1.10)$$

Prova

A média pode ser obtida facilmente ao derivar (1.7) e fazendo $t=0$. De maneira similar, ao derivar duas vezes (1.7) e fazer $t=0$, resulta em

$$E(Y^2) = \mu^2 + \mu^3\nu. \quad (1.11)$$

Logo a variância é obtida ao substituir (1.8) e (1.11), na equação seguinte

$$Var(Y) = E(Y^2) - [E(Y)]^2.$$

A moda da distribuição é obtida ao resolver a equação resultante da derivada da função de densidade dada em (1.6) com respeito a y e igualando a zero. Isso mostra que a distribuição é unimodal, desde que $\frac{\partial f_8(y;\mu,\nu)}{\partial y} = 0$ tenha uma única solução. \square

A função de distribuição acumulada, em termos da distribuição acumulada da Normal Padrão, apresenta-se no seguinte teorema:

Teorema 1.3

Seja $Y \sim IG(\mu, \nu)$. Então, a função de distribuição acumulada pode ser escrita:

$$F(y; \mu, \nu) = \Phi \left\{ (\nu y)^{-\frac{1}{2}} \left(\frac{y - \mu}{\mu} \right) \right\} + \exp \left(\frac{2}{\mu\nu} \right) \Phi \left\{ -(\nu y)^{-\frac{1}{2}} \left(\frac{y + \mu}{\mu} \right) \right\}, \quad (1.12)$$

onde $y > 0$ e

$$\Phi(z) = \int_{-\infty}^z (2\pi)^{-\frac{1}{2}} \exp \left(-\frac{t^2}{2} \right) dt.$$

A dedução da expressão dada em (1.12), em função da distribuição acumulada da Normal Padrão, foi feita por Shuster (1968) e Chhikara-Folks (1974).

A distribuição Gaussiana Inversa, assim como a Gama e Lognormal, converge assintoticamente à Normal em função de $\mu\nu$. Este resultado, pode ser verificado no seguinte teorema demonstrado por Wald (1947)

Teorema 1.4

Seja $Y \sim IG(\mu, \nu)$, então, quando $\nu\mu \rightarrow 0$ a distribuição de Y converge assintoticamente a uma distribuição Normal com média μ e variância $\mu^3\nu$. \square

A distribuição Gaussiana Inversa tem algumas propriedades análogas à distribuição Normal, apresentadas no seguinte teorema:

Teorema 1.5

i) Sejam $Y \sim IG(\mu, \nu)$ e c uma constante real positiva. Então

$$cY \sim IG(c\mu, \nu/c). \quad (1.13)$$

ii) Seja $Y \sim IG(\mu, \nu)$. Então

$$\frac{(Y - \mu)^2}{\nu\mu^2 Y} \sim \chi_{(1)}^2. \quad (1.14)$$

iii) Sejam Y_1, Y_2, \dots, Y_m variáveis aleatórias Gaussiana Inversas distribuídas independentemente com parâmetros μ_i e ν_i , para $i = 1, \dots, m$. Sejam $c_i, i = 1, \dots, m$ constantes reais positivas tais que para todo índice i , $\xi = c_i\nu_i\mu_i^2$ é uma constante positiva. Então a combinação linear

$$\sum_{i=1}^m c_i Y_i \sim IG \left(\sum_{i=1}^m c_i \mu_i, \xi \left(\sum_{i=1}^m c_i \mu_i \right)^{-2} \right). \quad (1.15)$$

A constante ξ é a condição de constância desta propriedade. \square

A função de Sobrevivência (Confiabilidade) $S(y) = 1 - F(y)$, pode ser facilmente obtida de (1.12) podendo ser escrita da seguinte forma

$$S(y; \mu, \nu) = \Phi \left\{ -\frac{y - \mu}{\mu(\nu y)^{\frac{1}{2}}} \right\} - \exp \left[2(\nu\mu)^{-1} \right] \Phi \left\{ -\frac{y + \mu}{\mu(\nu y)^{\frac{1}{2}}} \right\} \quad (1.16)$$

A função Risco (Taxa de Falha) é dada por

$$r(y; \mu, \nu) = \frac{(2\pi y^3 \nu)^{-\frac{1}{2}} \exp\left\{-\frac{(y-\mu)^2}{2\mu^2 \nu y}\right\}}{\Phi\left\{-\frac{y-\mu}{\mu(\nu y)^{\frac{1}{2}}}\right\} - \exp\left[2(\nu\mu)^{-1}\right] \Phi\left\{-\frac{y+\mu}{\mu(\nu y)^{\frac{1}{2}}}\right\}}. \quad (1.17)$$

Chhikara e Folks (1977) fizeram uma análise desta função. Desta análise, eles concluíram que $r(y)$ é crescente até algum valor que pertence ao intervalo (y_m, y_o) , onde $y_o > y_m$ e y_m moda da distribuição, logo decresce aproximando-se assintoticamente a $\frac{1}{2}(\nu\mu^2)^{-1}$. Além disso, eles também observaram que $r(y)$ pode ser virtualmente não decrescente para todo y , quando ν^{-1} é relativamente maior que μ , já que, quando $\nu^{-1} \rightarrow \infty$ e $y \rightarrow \infty$, então $r(y) \rightarrow \infty$. A Figura 1.3 Pode ilustrar uma idéia desta propriedade.

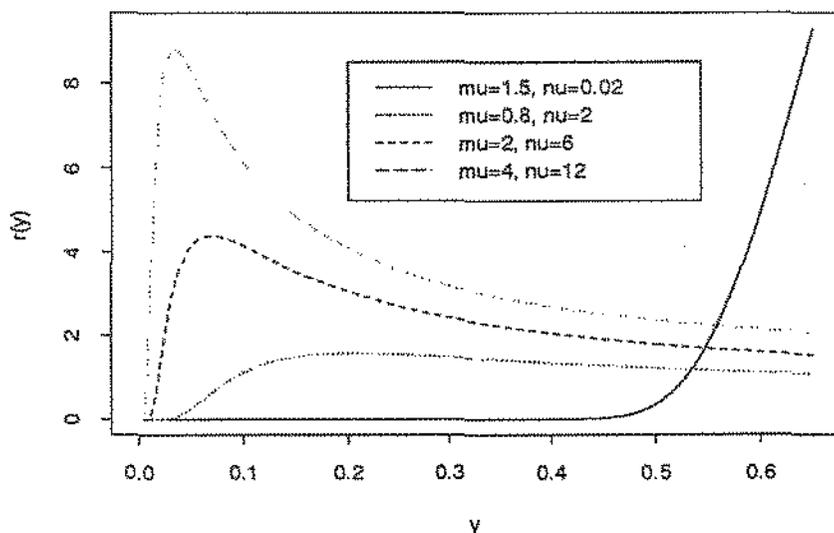


Figura 1.3: Gráfico da função risco para alguns valores de μ e ν .

Outra propriedade da distribuição Gaussiana Inversa, dada por Whitmore (1983),

importante na estimação e inferência dos parâmetros com dados censurados é

$$E(Y^{s+1}|Y > a) = \mu^2 \left[E(Y^{s-1}|Y > a) + \nu(2s-1)E(Y^s|Y > a) + \frac{2\nu a^{s+1} f_s(a; \mu, \nu)}{1 - F(a)} \right], \quad (1.18)$$

para todo s inteiro, com

$$E(Y^0|Y > a) = 1 \quad (1.19)$$

e

$$E(Y|Y > a) = \frac{\mu F\left(\frac{\mu^2}{a}\right)}{1 - F(a)}. \quad (1.20)$$

1.4 Modelo Exponencial com Dispersão

De acordo com Jørgensen (1987,1992), um *Modelo Exponencial com Dispersão* (caso contínuo) é uma distribuição com a seguinte função de densidade

$$p(y; \theta, \lambda) = a(\lambda, y) \exp \left\{ \lambda \left[y^T \theta - \kappa(\theta) \right] \right\}, \quad y \in \mathfrak{R}^k, \quad (1.21)$$

onde a e κ são funções conhecidas, λ varia em um subconjunto de \mathfrak{R}^+ e θ varia em um subconjunto de \mathfrak{R}^k . Nesta definição, pode-se observar que para λ conhecido, (1.21) é uma família exponencial, então a Família de Modelos Exponenciais com Dispersão é uma generalização da família exponencial. Para denotar que um vetor Y segue um Modelo Exponencial com Dispersão e com função densidade dada em (1.21), será adotada a notação $Y \sim ED(\mu, \nu)$, onde: $\mu = E(Y)$ é a esperança e $\nu = 1/\lambda$ o parâmetro de dispersão. A variância de Y é $Var(Y) = \nu V(\mu)$, onde $V(\mu)$ é a função

variância. Seja $\tau(\theta) = \kappa'(\theta)$. É fácil mostrar que $\mu = \tau(\theta)$ e $V(\mu) = \kappa''(\tau^{-1}(\mu))$, onde κ é a função geratriz dos cumulantes.

Para o caso discreto, Jørgensen (1987,1992) definiu o *Modelo Exponencial com Dispersão* como a distribuição de um vetor Z , com função de probabilidade:

$$p(z; \theta, \lambda) = a^*(\lambda, z) \exp \left[z^T \theta - \lambda \kappa(\theta) \right], \quad z \in Z^k, \quad (1.22)$$

onde a^* e κ são funções conhecidas. Jørgensen (1987, 1992) usa a notação $Z \sim ED^*(\theta, \lambda)$, para denotar que um vetor Z segue um Modelo Exponencial com Dispersão e com função de densidade (1.22). Também, pode deduzir-se que $E(Z) = \lambda \mu$ e $Var(Z) = \lambda V(\mu)$, onde $\mu = \tau(\theta) = \kappa'(\theta)$ e $V(\mu) = \kappa''(\tau^{-1}(\mu))$; isto é, o mesmo que para o caso contínuo.

As funções dadas em (1.21) e (1.22) estão relacionadas pela seguinte transformação: $y \rightarrow z = \lambda y$; isto é, existe uma relação fechada entre estas funções. Logo, um Modelo Exponencial com Dispersão pode ser expresso usando como densidade a função (1.21) ou (1.22), quando o modelo é contínuo. Da mesma forma, (1.21) ou (1.22) pode ser usadas como função de probabilidade quando o modelo é discreto.

Os Modelos Exponenciais com Dispersão são úteis em Modelos Lineares Generalizados (MLG), constitui uma generalização da distribuição dos erros em MLG. Estes modelos têm certa analogia com os modelos de locação-escala, onde $\mu = E(Y)$ faz o papel de locação e ν o papel de parâmetro de escala. Estes modelos têm as vantagens da elegante teoria assintótica, que generaliza a análise de desvio para os Modelos Lineares Generalizados, pois originam a versão assintótica dos testes T, F, χ^2 a partir da teoria de modelos lineares normais.

Jørgensen (1987,1992) também demonstrou que a distribuição Gaussiana Inversa

é um Modelo Exponencial com Dispersão. Isso pode ser mostrado facilmente, porque ao substituir $\gamma = -\theta$ em (1.1), tem-se uma função de densidade da forma (1.21)

$$f(y; \lambda, \theta) = a(\lambda, y) \exp\{\lambda[\theta y - \kappa(\theta)]\}, \quad \forall y > 0$$

com

$$a(\lambda, y) = \left(\frac{\lambda}{2\pi y^3}\right)^{\frac{1}{2}} \exp\left(-\frac{\lambda}{2y}\right)$$

e

$$\kappa(\theta) = -\sqrt{-2\theta},$$

para $\theta < 0$, $\lambda = \nu^{-1}$ e $\mu = (-2\theta)^{-\frac{1}{2}}$.

1.5 Divisibilidade Infinita

Steutel (1979) define Divisibilidade Infinita de uma variável Y da seguinte forma:

Definição 1

Uma variável aleatória Y é *infinitamente divisível* se para cada $n \in N$ existem variáveis aleatórias i.i.d., Y_{1n}, \dots, Y_{nn} , tais que:

$$Y \sim Y_{1n} + \dots + Y_{nn}.$$

A importância deste conceito, radica-se em que alguns modelos só podem ser definidos em termos de Divisibilidade Infinita. Assim, Steutel (1979) apresenta dois exemplos, o primeiro foi dado por Katti e o segundo por Thorim, ambos publicados em 1977.

Exemplo 1. Seja X o número de insetos coletados em certa área de terra num determinado período de tempo. A área é subdividida em 50 pequenos lotes aproximadamente iguais. Logo, a variável X pode ser representada no modelo da forma seguinte

$$X = Y_1 + \cdots + Y_{50},$$

onde é suposto que Y_1, \dots, Y_{50} são i.i.d. Também é conhecido que X pode ser aproximado pela distribuição logarítmica, que é infinitamente divisível.

Exemplo 2. Seja X a quantidade de dinheiro (considerada como variável) que uma companhia de seguros paga durante um ano. Sejam Y_1, \dots, Y_{52} as quantidades correspondentes a cada semana. Então, a variável X pode ser representada no modelo como

$$X = Y_1 + \cdots + Y_{52}$$

onde é suposto Y_1, \dots, Y_{52} são i.i.d. Neste caso, muitas distribuições são compatíveis a X , tais como a distribuição de Pareto e Lognormal, que são infinitamente divisíveis.

Stentel (1979), duvida que o segundo exemplo possa ser formulado desta forma, devido a que a distribuição da quantidade de dinheiro que a companhia paga por semana pode mudar de um período do ano a outro. Além disso, ele apresenta o seguinte teorema, que será útil para demonstrar que uma variável é infinitamente divisível.

Teorema 1.6

A divisibilidade infinita é preservada pela convolução e limite em distribuição. \square

Em alguns teoremas relacionados com divisibilidade infinita, se precisa do conceito de função completamente monótona, apresentado em Feller (1971).

Definição 2

Uma função Φ definida sobre $\mathbb{R}_0^+ = \mathbb{R}^+ \cup \{0\}$ é *completamente monótona* se possui derivada $\Phi^{(n)}$ de todas as ordens tal que $(-1)^n \Phi^{(n)}(t) \geq 0$, para todo $t > 0$.

Além disso, Feller (1971) apresenta os seguintes teoremas relacionado com a definição anterior:

Teorema 1.7

Se Φ é completamente monótona e Ψ uma função com derivada completamente monótona, então, a composição $\Phi(\Psi)$ é completamente monótona. \square

Teorema 1.8

A função $\mathfrak{Z}(t) = E(e^{-Yt})$ é a transformada de Laplace de uma distribuição infinitamente divisível se e só se existe uma função $\Phi(\cdot)$, tal que $\mathfrak{Z}(t) = \exp\{\Phi(t)\}$, onde $-\Phi(t)$ tem derivada completamente monótona e $\Phi(0) = 0$. \square

A seguir, mediante o uso da Definição 2 e o Teorema 1.8 prova-se que $Y \sim IG(\mu, \nu)$ é infinitamente divisível.

Teorema 1.9

$Y \sim IG(\mu, \nu)$ é infinitamente divisível.

Prova

Para o caso de $Y \sim IG(\mu, \nu)$, é fácil demonstrar que a transformada de Laplace é

$$\mathfrak{Z}(t) = \exp\{\Phi(t)\},$$

onde

$$\Phi(t) = \lambda [\kappa(\theta - t) - \kappa(\theta)], \text{ com } \Phi(0) = 0,$$

$$\lambda = \nu^{-\frac{1}{2}} > 0, \theta = -(2\nu\mu^2)^{-1} \text{ e } \kappa(\theta) = -(-2\theta)^{-1/2}.$$

Mas, como $-\Phi^{(1)}(t) \geq 0$, para todo $t > \theta$ e $\Phi(t) = \lambda\kappa(\theta - t) + c$, onde $\lambda > \theta$ e c é uma constante, então é preciso mostrar que $g(u) = \kappa(-u) = -(2u)^{1/2}$ é completamente monótona, para que a distribuição de Y seja infinitamente divisível. Mas, como

$$(-1)^n g^{(n)}(u) = \{1 \cdot 3 \cdot \dots \cdot (2n - 1)\} (2u)^{-\frac{2n-1}{2}} \geq 0,$$

para todo $u > 0$, temos que $g(u)$ é completamente monótona. Portanto, a distribuição de $Y \sim IG(\mu, \nu)$ é infinitamente divisível. \square

2. Distribuição da Recíproca Complementar de uma Variável Gaussiana Inversa

2.1 Introdução

Uma variável aleatória Y se distribui como a Recíproca Complementar de $W \sim IG(\mu, \nu)$, se $Y^{-1} \sim IG(\mu^{-1}, \nu\mu^2)$. A distribuição desta variável foi apresentada por Jørgensen, Seshadri e Whitmore (1991). A importância da distribuição, é que inclui como um caso particular a distribuição da Recíproca de uma variável Gaussiana Inversa, distribuição que também é útil para modelar tempo de falha. Um exemplo de aplicação desta última distribuição foi dada por Jørgensen (1982). Propriedades importantes da distribuição da Recíproca de uma variável Gaussiana Inversa pode ser encontradas em Tweedie (1957a), Wasan (1968), Johnson e Kotz (1970), e Chhikara e Folks (1989).

A distribuição da Recíproca Complementar de uma variável Gaussiana Inversa, da mesma forma que a distribuição Gaussiana Inversa, é útil para modelar tempo de falha. Esta distribuição é uma alternativa para modelar tempo de falha nas seguintes situações: (1) Quando existe evidência de que as taxas de falhas não são proporcionais. (2) Quando o logaritmo dos tempos não se ajustam a um modelo de locação-escala. (3) Quando existe evidência de que os tempos até acontecer a falha possam ser interpretados como a soma de duas variáveis, uma χ^2 , com um grau de

liberdade e a outra Gaussiana Inversa. (4) Quando, por experiência ou natureza do fenômeno, pareça que os dados possam ser modelados por esta distribuição.

Pelos motivos apresentados, e esta distribuição satisfaz propriedades semelhantes à distribuição Gaussiana Inversa, neste capítulo se faz um pequeno estudo desta distribuição. Assim, na Seção 2.2 apresenta-se a função de densidade e seus relações com as diferentes representações da densidade da IG. Na Seção 2.3, mostra-se suas propriedades mais importantes, que serviram para demonstrar algumas propriedades da mistura de distribuições Gaussiana Inversa e sua Recíproca Complementar.

2.2 Função de densidade e sua relação com a densidade da Gaussiana Inversa

A variável Y se distribui como a Recíproca Complementar de uma variável Gaussiana Inversa com função de densidade (1.6), se tem a seguinte função de densidade:

$$g(y; \mu, \nu) = (2\pi\mu^2\nu y)^{-\frac{1}{2}} \exp\left\{-\frac{(y-\mu)^2}{2\nu\mu^2 y}\right\}. \quad (2.1)$$

A função de densidade desta variável se relaciona com as funções de densidade da Gaussiana Inversa, dadas em (1.2), (1.3), (1.4), (1.5) e (1.6), da seguinte forma:

$$\begin{aligned} g(y; \mu, \nu) &= \mu^{-1} y f_2(y; \mu, \lambda) \\ &= \mu^{-2} y f_3(\mu^{-1} y; 1, \psi) \\ &= \psi \lambda^{-2} y f_4(\lambda^{-1} y; \psi, 1) \\ &= \delta y f_5(y; \delta, \nu) \\ &= \mu^{-1} y f_6(y; \mu, \nu), \end{aligned}$$

onde μ , δ , λ , e ψ satisfazem as seguintes relações:

$$\mu = \delta^{-1} = \lambda\psi^{-1} = (\nu\psi)^{-1}.$$

A Figura 2.1 mostra o gráficos da funções de densidade para $\nu = 2$ e alguns valores de μ . A Figura 2.2 mostra os gráficos da funções de densidade para $\mu = 5$ e alguns valores de ν . Nestes gráficos pode observar-se que parâmetro μ joga um papel parecido ao de um parâmetro de locação, e a escala do gráfico depende do parâmetro ν . Logo μ é uma especie de parâmetro de locação e ν parâmetros de escala da distribuição, respectivamente.

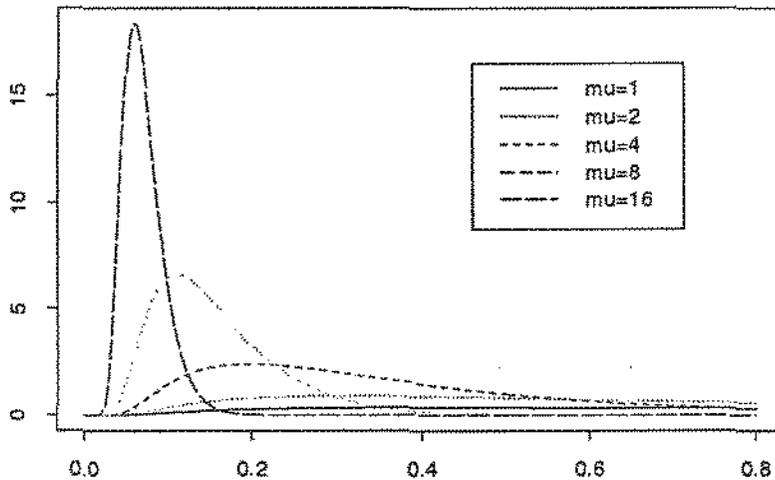


Figura 2.1: Gráfico da função de densidade g para $\nu = 2$ fixo e diferentes valores de μ .

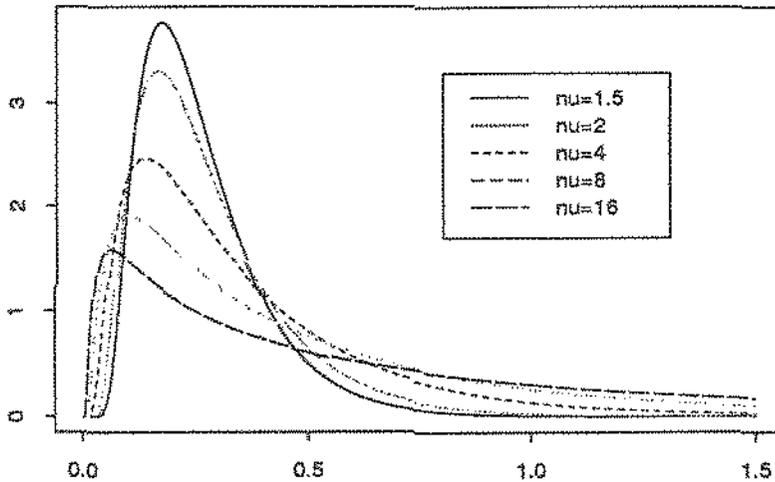


Figura 2.2: Gráfico da função de densidade g para $\mu = 5$ fixo e diferentes valores de ν .

Para denotar que uma variável aleatória Y se distribui como Recíproca Complementar de uma variável Gaussiana Inversa, com função de densidade dada em (2.1), usa-se a seguinte notação: $Y \sim R - IG(\mu, \nu)$.

2.3 Algumas propriedades importantes

A distribuição $R - IG(\mu, \nu)$ tem muitas propriedades interessantes, a maior parte é semelhantes às propriedades da Recíproca de uma variável Gaussiana Inversa. Assim, nesta seção apresentam-se algumas destas propriedades.

Teorema 2.1

Se $Y \sim R - IG(\mu, \nu)$, então a função geratriz de momentos é dada por

$$M_Y(s) = (1 - 2s\nu\mu^2)^{-\frac{1}{2}} \exp \left\{ (\nu\mu)^{-1} \left[1 - (1 - 2s\mu^2\nu)^{\frac{1}{2}} \right] \right\}. \quad (2.2)$$

Também, esta função pode ser escrita como

$$M_Y(s) = M_{Y_1}(s) M_{Y_2}(s), \quad (2.3)$$

onde

$$M_{Y_1}(s) = \exp \left\{ (\nu\mu)^{-1} \left[1 - (1 - 2s\mu^2\nu)^{\frac{1}{2}} \right] \right\} \quad (2.4)$$

é a função geratriz de momentos de uma variável $Y_1 \sim IG(\mu, \nu)$, e

$$M_{Y_2}(s) = (1 - 2s\nu\mu^2)^{-\frac{1}{2}} \quad (2.5)$$

é a função geratriz de momentos de uma variável $Y_2 \sim \nu\mu^2\chi_{(1)}^2$. \square

Deste teorema, pode-se observar que Y é uma convolução de duas variáveis independentes, a primeira variável é Gaussiana Inversa e a outra $\nu\mu^2$ vezes $\chi_{(1)}^2$. Este resultado foi demonstrado por Tweedie (1957) para a Recíproca de uma variável Gaussiana Inversa.

Teorema 2.2

Se $Y \sim R - IG(\mu, \nu)$, então

i) a média de Y é

$$E(Y) = \mu + \mu^2\nu,$$

ii) a variância é

$$Var(Y) = \mu^3\nu + 2\mu^4\nu^2,$$

iii) a moda da distribuição é

$$y_m = \mu \left\{ \left(1 + \frac{\nu^2 \mu^2}{4} \right)^{\frac{1}{2}} - \frac{\mu \nu}{2} \right\}.$$

Prova

i) e ii) Por (2.3), Y é uma convolução de duas variáveis independentes Y_1 e Y_2 .

Então, a média de Y é

$$E(Y) = E(Y_1) + E(Y_2) = \mu + \mu^2 \nu,$$

e a variância é

$$Var(Y) = Var(Y_1) + Var(Y_2) = \mu^3 \nu + 2\mu^4 \nu^2.$$

Este resultado é obtido de (2.4) e (2.5), substituindo $s = 0$ nas derivadas de $M_{Y_i}(s)$, $i = 1, 2$.

iii) Ao derivar a função de densidade dada em (2.1) com respeito a y e igualando a zero resulta

$$y^2 + \nu \mu^2 y - \mu^2 = 0.$$

Esta equação só tem uma raiz positiva para μ e ν positivos, dado por

$$y_m = \mu \left\{ \left(1 + \frac{\nu^2 \mu^2}{4} \right)^{\frac{1}{2}} - \frac{\mu \nu}{2} \right\}.$$

Isto mostra que a distribuição é unimodal. □

É fácil deduzir de (1.12), a função de distribuição acumulada de Y em termos da função de distribuição acumulada Normal Padrão. Logo, para todo $y > 0$, temos que

$$G(y; \mu, \nu) = \Phi \left\{ \frac{y - \mu}{\mu (\nu y)^{\frac{1}{2}}} \right\} - \exp \left\{ 2(\nu \mu)^{-1} \right\} \Phi \left\{ -\frac{y + \mu}{\mu (\nu y)^{\frac{1}{2}}} \right\} \quad (2.6)$$

onde

$$\Phi(a) = \int_{-\infty}^a \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} dz.$$

A função de Sobrevivência (Confiabilidade) $S(y) = 1 - G(y)$ é

$$S(y; \mu, \nu) = \Phi\left\{-\frac{y - \mu}{\mu(\nu y)^{\frac{1}{2}}}\right\} + \exp\{2(\nu\mu)^{-1}\} \Phi\left\{-\frac{y + \mu}{\mu(\nu y)^{\frac{1}{2}}}\right\} \quad (2.7)$$

para todo $y > 0$.

A função Risco (Taxa de Falha) $r(y) = g(y)/S(y)$ pode ser obtida ao substituir as expressões dadas em (2.1) e (2.7), na definição da seguinte função

$$r(y; \mu, \nu) = \frac{(2\pi\mu^2\nu y)^{-\frac{1}{2}} \exp\left\{-\frac{(y-\mu)^2}{2\nu\mu^2 y}\right\}}{\Phi\left\{-\frac{y-\mu}{\mu(\nu y)^{\frac{1}{2}}}\right\} + \exp\{2(\nu\mu)^{-1}\} \Phi\left\{-\frac{y+\mu}{\mu(\nu y)^{\frac{1}{2}}}\right\}} \quad (2.8)$$

para todo $y > 0$.

Chhikara e Folks (1977) mostraram para a distribuição Gaussiana Inversa, que sea função risco $r(y)$ é crescente até algum valor que pertence ao intervalo (y_m, y_o) , onde $y_o > y_m$ e y_m é a moda, logo decresce aproximando-se assintoticamente ao valor de $\frac{1}{2}(\nu\mu^2)^{-1}$. Pode ser mostrada, da mesma forma, que $r(y; \mu, \nu)$ tem o mesmo comportamento. Além disso, $r(y; \mu, \nu)$ de (2.8) pode ser virtualmente não decrescente para todo $y > 0$, já que, quando $\nu^{-1} \rightarrow \infty$ e $y \rightarrow \infty$, então $r(y) \rightarrow \infty$. Uma idéia desta conclusão pode ser apreciada na Figura 2.3.

Da mesma forma que a distribuição Gaussiana Inversa, a $R - IG(\mu, \nu)$ é aproximadamente Normal com média μ e variância $\nu\mu^3$, para μ fixo e $\mu\nu$ pequeno. Desta propriedade pode observar-se que a forma da distribuição depende de $\nu\mu$, então $\nu\mu$ é o parâmetro de forma da distribuição. Além disso, esta distribuição possui algumas propriedades semelhantes da distribuição Normal, que são apresentadas no seguinte teorema.

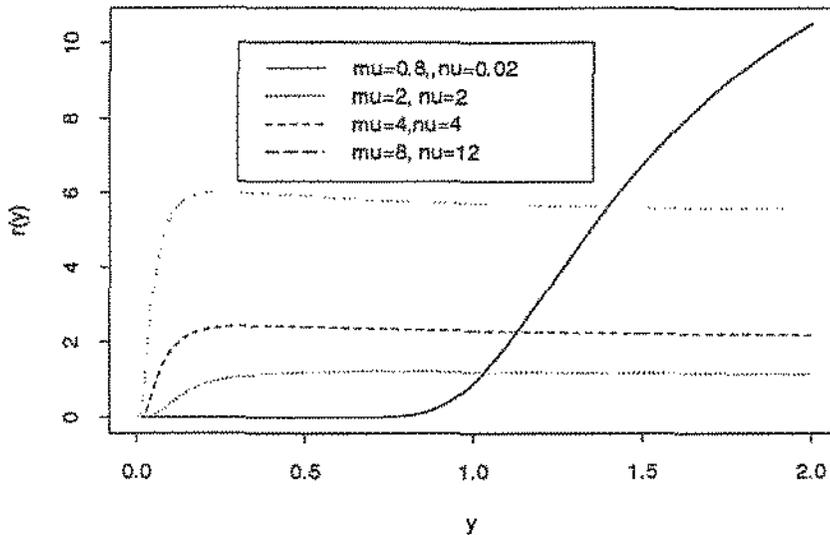


Figura 2.3: Gráfico da função risco para alguns valores de μ e de ν .

Teorema 2.3

i) Seja $Y \sim R - IG(\mu, \nu)$. Então

$$\frac{(Y - \mu)^2}{\nu \mu^2 Y} \sim \chi_{(1)}^2.$$

ii) Seja $Y \sim R - IG(\mu, \nu)$ e c uma constante real positiva. Então

$$cY \sim R - IG\left(c\mu, \frac{\nu}{c}\right).$$

□

A distribuição da Recíproca Complementar de uma variável Gaussiana Inversa satisfaz uma propriedade interessante de convolução, que é apresentada no seguinte teorema

Teorema 2.4

Sejam Y_1, \dots, Y_m variáveis aleatórias independentes, tais que $Y_i \sim R-IG(\mu_i, \nu_i)$, c_i constantes reais positiva tal que $\xi = c_i \nu_i \mu_i^2$, para todo $i = 1, \dots, m$. Então a função geratriz dos momentos de $\sum_{i=1}^m c_i Y_i$ é

$$M_{(\sum_{i=1}^m c_i Y_i)}(s) = M_{V_1}(s) M_{V_2}(s) \quad (2.9)$$

onde

$$M_{V_1}(s) = (1 - 2s\xi)^{-\frac{m}{2}}$$

e

$$M_{V_2}(s) = \exp \left\{ \xi^{-1} \left(\sum_{i=1}^m c_i \mu_i \right) \left[1 - (1 - 2s\xi)^{\frac{1}{2}} \right] \right\}.$$

□

Neste teorema, pode-se observar que $M_{V_1}(s)$ e $M_{V_2}(s)$ são as funções geratrizes de momentos das variáveis $V_1 \sim \chi_{(m)}^2$ e $V_2 \sim IG\left(\sum_{i=1}^m c_i \mu_i, \xi \left(\sum_{i=1}^m c_i \mu_i\right)^{-2}\right)$, respectivamente. Logo, a combinação linear de variáveis Recíprocas Complementares, sob condição de constância, é uma convolução de duas variáveis independentes $V_1 \sim \chi_{(m)}^2$ e $V_2 \sim IG\left(\left(\sum_{i=1}^m c_i \mu_i\right), \xi \left(\sum_{i=1}^m c_i \mu_i\right)^{-2}\right)$.

Outra propriedade importante relacionada com a propriedade dada na equação (1.18) da distribuição Gaussiana Inversa, que é útil na estimação e inferência com dados censurados, é apresentada no seguinte teorema.

Teorema 2.5

Seja $Y \sim R-IG(\mu, \nu)$ e $X \sim IG(\mu, \nu)$. Então

$$E(Y^s | Y > a) = \frac{E(X^{s+1} | X > a)}{E(X | X > a)} \quad (2.10)$$

para todo s inteiro, e onde $E(X^s | X > a)$ pode ser obtidas através da fórmula (1.18).

□

A distribuição da Recíproca Complementar de uma variável Gaussiana Inversa não é um modelo exponencial com dispersão, já que não pode ser expressa na forma dada em (1.21) ou (1.22). Mas, uma variável Recíproca Complementar é infinitamente divisível. Pode-se demonstrar facilmente, já que, a variável recíproca complementar é uma convolução de duas variáveis infinitamente divisíveis, a primeira Gaussiana Inversa e a segunda Gama. Logo, pelo Teorema 1.6 do Capítulo 1, a variável Recíproca Complementar de uma Gaussiana Inversa é infinitamente divisível.

3. Mistura da Distribuição Gaussiana Inversa Com sua Recíproca Complementar

3.1 Introdução

A mistura da distribuição Gaussiana Inversa com sua Recíproca Complementar foi obtida por Jørgensen, Seshadri, e Whitmore (1991), como um caso particular da generalização da função de distribuição acumulada da Gaussiana Inversa. Eles derivaram duas representações desta distribuição, uma como uma mistura da distribuição Gaussiana Inversa com sua Recíproca Complementar e a outra como uma convolução de duas variáveis independentes, a primeira Gaussiana Inversa e segunda Binomial Composta. Também, eles discutiram propriedades importantes desta distribuição, algumas delas semelhantes à distribuição Gaussiana Inversa. Estas e outras propriedades tratadas por estes autores, serão abordadas neste capítulo.

3.2 Função de densidade

Sejam Y_1 e Y_2 variáveis aleatórias independentes, tais que

$$Y_1 \sim IG(\mu, \nu) \text{ e } Y_2 \sim R - IG(\mu, \nu). \quad (3.1)$$

Para fazer a mistura de Y_1 e Y_2 , considera-se a seguinte variável

$$Y = \begin{cases} Y_1 & \text{com probabilidade } 1 - p \\ Y_2 & \text{com probabilidade } p \end{cases}, \quad (3.2)$$

onde $0 \leq p \leq 1$. A distribuição de Y dada em (3.2) é conhecida com o nome de Mistura da Distribuição Gaussiana Inversa com sua Recíproca Complementar.

De acordo com Titterington, Smith e Makov (1985), a variável aleatória Y é uma mistura de duas variáveis Y_1 e Y_2 se tem função de densidade da forma

$$f(y|\psi) = (1 - p) f_1(y|\theta) + p f_2(y|\theta), \quad (3.3)$$

onde $0 \leq p \leq 1$, f_1 e f_2 são funções de densidade de Y_1 e Y_2 , respectivamente, e $\psi = (\theta', p)'$. Para o caso que se discute agora, f_1 e f_2 são as funções de densidade dadas em (1.6) e (2.1) respectivamente, que correspondem às densidades de Gaussiana Inversa e sua Recíproca Complementar com parâmetro $\theta = (\mu, \nu)'$. Logo, ao substituir as correspondentes funções de densidades em (3.3), resulta

$$g(y; \mu, \nu, p) = (1 - p + p\mu^{-1}y) (2\pi\nu y^3)^{-\frac{1}{2}} \exp\left\{-\frac{(y - \mu)^2}{2\nu\mu^2 y}\right\}. \quad (3.4)$$

Para denotar que uma variável aleatória se distribui como uma mistura de Gaussiana Inversa com sua Recíproca Complementar e com função de densidade dada em (3.4), será usada a notação $M - IG(\mu, \nu, p)$.

A Figura 3.1 mostra os gráficos das funções de densidades para $\nu = 0.2$, $p = 0.6$ e vários valores de μ . A Figura 3.2 mostra os gráficos das funções de densidades para $\mu = 0.2$, $p = 0.6$ e vários valores de ν . Logo, para um mesmo valor de p , pode-se observar que parâmetro μ joga o papel de um parâmetro de locação e a escala depende do parâmetro ν . Portanto, μ é uma especie de parâmetro de locação e ν o parâmetro

de escala da distribuição, respectivamente.

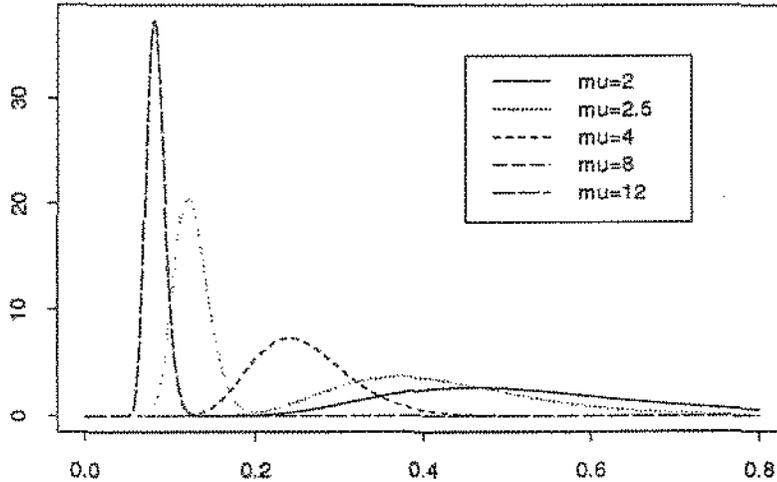


Figura 3.1: Gráfico da função de densidade g para $\nu = 0.2$ e $p = 0.6$ e vários valores de μ .

Outra representação interessante da função de densidade da Mistura de Gaussiana Inversa com sua Recíproca Complementar, consegue-se com a reparametrização

$$\gamma = \mu \left(\frac{1-p}{p} \right). \quad (3.5)$$

Logo, a função de densidade pode ser escrita

$$g_*(y; \mu, \nu, \gamma) = \left(\frac{\gamma + y}{\mu + \gamma} \right) (2\pi\nu y^3)^{-\frac{1}{2}} \exp \left\{ -\frac{(y - \mu)^2}{2\mu^2 \nu y} \right\}. \quad (3.6)$$

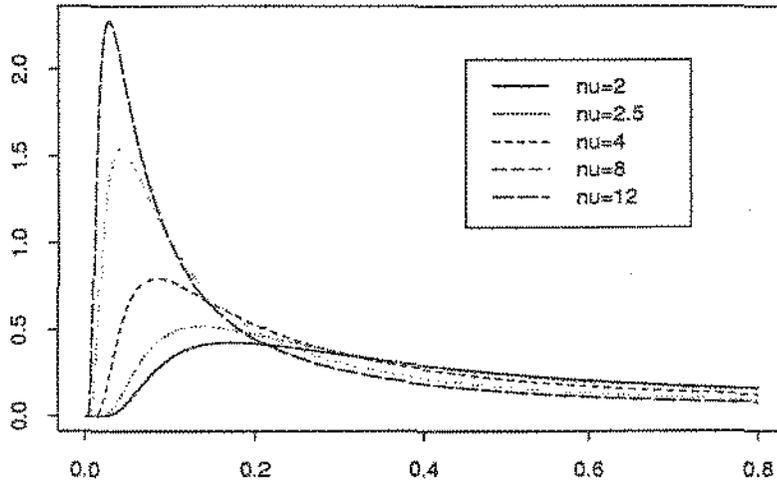


Figura 3.2: Gráfico da função de densidade g para $\mu = 0.2$ e $p = 0.6$ e vários valores de ν .

3.3 Função de Distribuição

Seja Y uma mistura das variáveis Y_1 e Y_2 . A função de distribuição acumulada de Y é dada por

$$G_p(y) = (1 - p)F_1(y) + pF_2(y), \quad (3.7)$$

onde F_1 e F_2 são funções de distribuição acumulada de Y_1 e Y_2 , respectivamente. Logo, ao substituir as expressões equivalentes dadas em (1.12) e (2.7), tem-se

$$G_p(y) = \Phi\{\alpha(y)\} + (1 - 2p) \exp\left(\frac{2}{\mu\nu}\right) \Phi\{\beta(y)\} \quad (3.8)$$

para todo $y \geq 0$, onde

$$\alpha(y) = \frac{y - \mu}{\mu(\nu y)^{\frac{1}{2}}} \quad e \quad \beta(y) = -\frac{y + \mu}{\mu(\nu y)^{\frac{1}{2}}} \quad (3.9)$$

e

$$\Phi(a) = \int_{-\infty}^a \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} dz.$$

Jørgensen, Seshadri e Whitmore (1991) demonstraram que $\alpha(y)$ e $\beta(y)$ são soluções linearmente independentes da equação diferencial de Euler

$$4y^2x'' + 4yx' - x = 0.$$

Soluções desta equação diferencial são da forma

$$x(y) = C_1y^{\frac{1}{2}} + C_2y^{-\frac{1}{2}}, \quad (3.10)$$

onde C_1 e C_2 são constantes (Hildebrand, 1976, pag. 12-15). Além disso, eles demonstraram que a função de distribuição dada em (3.8) é um caso particular da seguinte generalização da distribuição acumulada da Gaussiana Inversa

$$F(y) = A\Phi\{x_1(y)\} + B\Phi\{x_2(y)\} \quad (3.11)$$

onde $x_1(y)$ e $x_2(y)$ são soluções linearmente independentes da forma dada em (3.10), A e B são constantes escolhidas de modo que (3.11) seja função de distribuição acumulada.

3.4 Propriedades desta Família de Mistura

Esta família de distribuições tem muitas propriedades interessantes, algumas delas apresenta-se a seguir.

Se $Y \sim M - IG(\mu, \nu, p)$, pela relação (3.8), a função de Sobrevivência (Confiabilidade) $S(y) = 1 - G_p(y)$ é dada por

$$S(y; \mu, \nu, p) = \Phi \left\{ -\frac{y - \mu}{\mu(\nu y)^{\frac{1}{2}}} \right\} - (1 - 2p) \exp \left(\frac{2}{\mu\nu} \right) \Phi \left\{ -\frac{y + \mu}{\mu(\nu y)^{\frac{1}{2}}} \right\}, \quad (3.12)$$

e a função Risco (Taxa de Falha) $r(y) = g(y; \mu, \nu, p) / R(y)$ é

$$r(y; \mu, \nu, p) = \frac{(1 - p + p\mu^{-1}y)(2\pi\nu y^3)^{-\frac{1}{2}} \exp \left\{ -\frac{(y-\mu)^2}{2\nu\mu^2 y} \right\}}{\Phi \left\{ -\frac{y-\mu}{\mu(\nu y)^{\frac{1}{2}}} \right\} - (1 - 2p) \exp \left(\frac{2}{\mu\nu} \right) \Phi \left\{ -\frac{y+\mu}{\mu(\nu y)^{\frac{1}{2}}} \right\}}. \quad (3.13)$$

De uma análise desta função Risco, muito similar ao realizado por Chhikara e Folks (1977) para a função Risco da IG, chegou-se a seguinte conclusão: Existe um ponto y_0 tal que $r(y)$ é crescente até algum ponto que pertence ao intervalo (y_m, y_0) , onde $y_0 > y_m$ e y_m é a moda da $M - IG$, logo decresce aproximando-se assintoticamente ao valor $\frac{1}{2}(\nu\mu^2)^{-1}$. Além disso, também pode-se observar que $r(y)$ pode ser virtualmente não decrescente para todo y , quando ν^{-1} é relativamente maior que μ . Isto é, quando $\nu^{-1} \rightarrow \infty$ e $y \rightarrow \infty$, então $r(y) \rightarrow \infty$. A demonstração desta conclusão, encontra-se feito num relatório não publicado, Maehara (1994). Uma idéia desta propriedade pode ser observado na Figura 3.6

Também, temos os seguintes resultados:

Teorema 3.1

Seja $Y \sim M - IG(\mu, \nu, p)$, $Y_1 \sim IG(\mu, \nu)$ e $Y_2 \sim R - IG(\mu, \nu)$. Então, a função geratriz de momentos de Y é

$$M_Y(s; \mu, \nu, p) = (1 - p)M_{Y_1}(s; \mu, \nu) + pM_{Y_2}(s; \mu, \nu) \quad (3.14)$$

onde M_{Y_1} e M_{Y_2} são funções geratrizes de momentos dada em (1.7) e (2.2), respectivamente. □

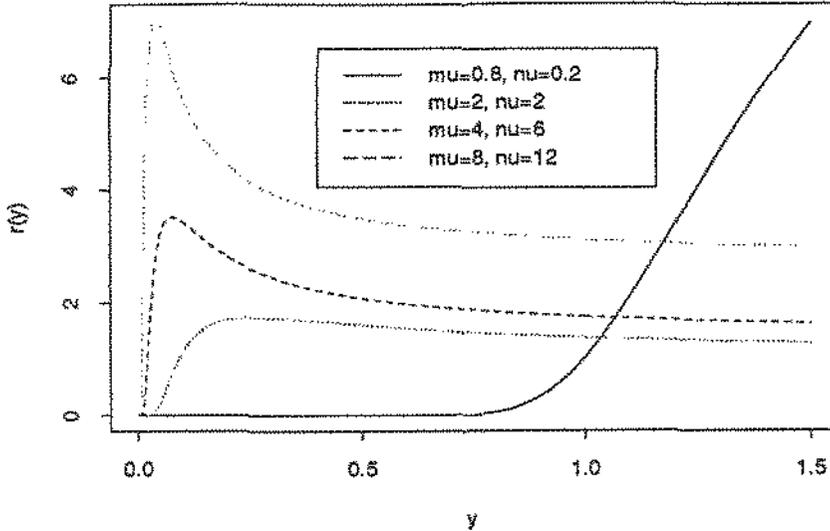


Figura 3.3: Gráfico da função risco para alguns valores de μ , ν e de p .

Teorema 3.2

Seja $Y \sim M - IG(\mu, \nu, p)$. A função geratriz de momentos de Y pode ser escrita

$$M_Y(s; \mu, \nu, p) = M_{Y_1}(s; \mu, \nu) [1 - p + pM_{Y_3}(s; \nu\mu^2)], \quad (3.15)$$

onde $M_{Y_1}(s; \mu, \nu)$ é função geratriz de momentos da variável $Y_1 \sim IG(\mu, \nu)$ e $M_{Y_3}(s; \nu\mu^2)$ é a função geratriz de momentos de uma variável $Y_3 \sim \nu\mu^2\chi_{(1)}^2$.

Prova

Ao substituir a expressão equivalente, dada na equação (2.3) em (3.14) a função geratriz de momentos de Y pode ser escrita como

$$M_Y(s; \mu, \nu, p) = (1 - p)M_{Y_1}(s; \mu, \nu) + pM_{Y_1}(s; \mu, \nu)M_{Y_3}(s; \nu\mu^2).$$

Logo, fatorando $M_{Y_1}(s; \mu, \nu)$ obtém-se (3.15). □

Esta propriedade indica que Y é uma *convolução* de uma distribuição Gaussiana Inversa com uma Bernoulli Composta. Isto é,

$$Y = Y_1 + T, \quad (3.16)$$

onde Y_1 e T são variáveis aleatórias independentes, tais que

$$Y_1 \sim IG(\mu, \nu), \quad (3.17)$$

$$T = \begin{cases} 0 & \text{com probabilidade } 1 - p \\ Y_3 & \text{com probabilidade } p, \end{cases} \quad (3.18)$$

e

$$Y_3 \sim \nu\mu^2\chi_{(1)}^2.$$

Teorema 3.3

Seja $Y \sim M - IG(\mu, \nu, p)$. Então,

i) A média de Y é

$$E(Y) = \mu + p\nu\mu^2.$$

ii) A variância de Y é

$$Var(Y) = \nu\mu^3 + \nu^2\mu^4p(3 - p).$$

iii) A mistura de distribuições da Gaussiana Inversa com sua Recíproca Complementar é sempre unimodal.

Prova

Da equação (3.16), temos que a média de uma variável $Y \sim M - IG(\mu, \nu, p)$ é:

$$E(Y) = E(Y_1) + E(T), \quad (3.19)$$

onde Y_1 e Y_2 são definidas em (3.17) e (3.18), respectivamente. Por (1.8) $E(Y_1) = \mu$ e por (3.18) $E(T) = p\nu\mu^2$. Substituindo em (3.19), tem-se a média.

Da mesma forma, a variância de Y é

$$Var(Y) = Var(Y_1) + Var(T),$$

porque Y_1 e T são independentes. Por (1.11) $Var(Y_1) = \nu\mu^3$, e para obter a $Var(T) = \nu^2\mu^4p(3-p)$, é preciso derivar o segundo fator do lado direito de (3.15). Substituindo, tem-se a variância de Y .

Ao derivar a função de densidade dada em (3.4) com respeito a y , resulta

$$- (2\pi\nu y^7)^{-\frac{1}{2}} \exp\left\{-\frac{(y-\mu)^2}{2\nu\mu^2 y}\right\} (a_0 + a_1 y - a_2 y^2 - a_3 y^3), \quad (3.20)$$

onde $a_0 = (1-p)\nu^{-1}$, $a_1 = p(\mu\nu)^{-1}$, $a_2 = p\mu^{-1} + (1-p)(\mu^2\nu)^{-1}$, e $a_3 = p(\mu^3\nu)^{-1}$.

Logo, igualando a zero resulta a seguinte equação

$$a_0 + a_1 y - a_2 y^2 - a_3 y^3 = 0,$$

uma vez que, para $y > 0$, os outros fatores de (3.20) são diferentes de zeros. Logo, pelo teorema do valor médio, esta equação tem exatamente uma raiz positiva para todo $\mu > 0$, $\nu > 0$, e $p \in [0, 1]$, dado que $a_i > 0$, $i = 1, 2, 3$. Portanto, a $M-IG$ é unimodal. \square

Assim como a distribuição Gaussiana Inversa, sua Recíproca Complementar satisfaz a propriedade apresentada no teorema a seguir.

Teorema 3.4

Seja $Y \sim M - IG(\mu, \nu, p)$ e c é uma constante positiva. Então,

$$cY \sim M - IG(c\mu, \nu/c, p).$$

Esta propriedade pode ser facilmente demonstrada usando a relação dada em (3.15). \square

A distribuição da Mistura da Gaussiana Inversa com sua Recíproca Complementar atende à seguinte propriedade de convolução.

Teorema 3.5

Sejam Y_1, \dots, Y_n variáveis aleatórias independentes, tais que $Y_i \sim M-IG(\mu_i, \nu/\mu_i^2, p)$. Então, a variável $T = \sum_{i=1}^n Y_i$ tem a seguinte função geratriz dos momentos

$$M_T(s; \mu_1, \dots, \mu_n, \nu, p) = M_{Y_1}(s; \mu, \nu/\mu^2) [1 - p + pM_{Y_3}(s; \nu)]^n, \quad (3.21)$$

onde $\mu = \sum_{i=1}^n \mu_i$, $M_{Y_1}(s; \mu, \nu/\mu^2)$ é a função geratriz dos momentos de uma variável $Y_1 \sim IG(\mu, \nu/\mu^2)$ e $M_{Y_3}(s; \nu)$ de uma variável $Y_3 \sim \nu\mu^2\chi_{(1)}^2$. Isto é, T é uma convolução de uma distribuição Gaussiana Inversa com uma Binomial Composta. \square

Teorema 3.6

Seja $Y \sim M-IG(\mu, \nu, p)$. Então $Z = \sqrt{(Y - \mu)^2 / (\nu\mu^2 Y)}$ tem a seguinte função de densidade

$$g_Z(z; \mu, \nu, p) = \left\{ 1 - (1 - 2p)z \sqrt{\frac{\mu\nu}{4 + \mu\nu z^2}} \right\} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right], \quad (3.22)$$

para todo $z \in \mathfrak{R}$. \square

Na função de densidade dada em (3.22), pode-se observar claramente que quando $p=1/2$, Z tem uma distribuição Normal Padrão. Por esta razão, se $Y \sim M-IG(\mu, \nu, 1/2)$ tem-se a família de distribuições de tempo de vida estudada por Birnbaum e Saunders (1969).

Teorema 3.7

Para qualquer $p \in [0, 1]$, a distribuição de uma variável $Y \sim M - IG(\mu, \nu, p)$ é aproximadamente Normal com média μ e variância $\nu\mu^3$ para um μ fixo e $\mu\nu$ pequeno.

Prova

Fazendo a transformação

$$Z = \frac{Y - \mu}{\mu\lambda^{\frac{1}{2}}}$$

onde $\lambda = \mu\nu$, tem-se a densidade

$$g_Z(z; \mu, \nu, p) = (2\pi)^{-\frac{1}{2}} \left\{ \frac{1 - p + p(\lambda^{\frac{1}{2}}z + 1)}{(\lambda^{\frac{1}{2}}z + 1)^{-\frac{3}{2}}} \right\} \exp \left\{ -\frac{1}{2} \left[\frac{z^2}{\lambda^{\frac{1}{2}}z + 1} \right] \right\}.$$

Logo, quando $\lambda \rightarrow 0$, tem-se a densidade da Normal Padrão. Isto implica, para μ fixo e $\nu\mu$ pequeno, $M - IG$ é aproximadamente normal com média μ e variância $\nu\mu^3$, para qualquer $p \in [0, 1]$. \square

Uma idéia desta propriedade pode ser apreciada nos gráficos dados nas Figura 3.3, Figura 3.4 e Figura 3.5. Estas figuras mostram que para um mesmo valor de μ , os gráficos das densidades para valores diferentes de p , são mais parecidas conforme diminui o valor de ν . Também, deste teorema pode observar-se que a forma da distribuição depende do valor de $\mu\nu$, então $\nu\mu$ é o parâmetro de forma da distribuição.

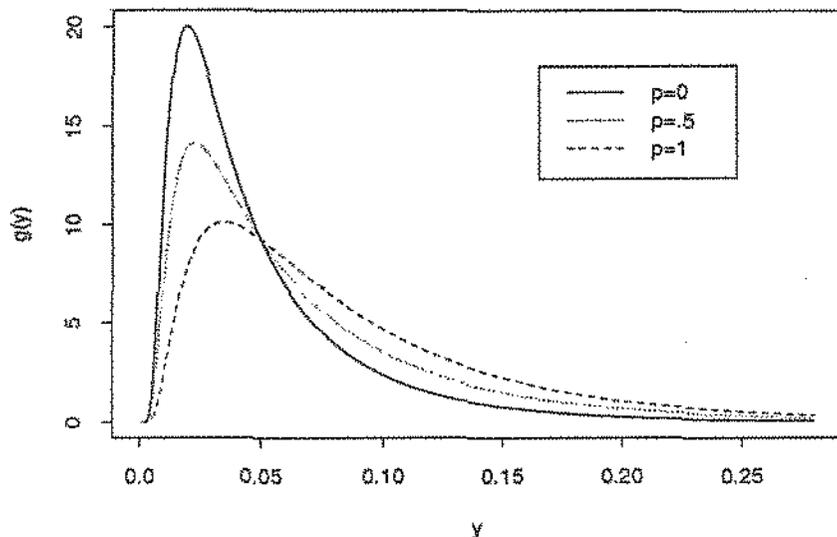


Figura 3.4: Gráfico da função de densidade g para $\mu = 20$ e $\nu = 15$ e três valores de p .

Outra propriedade interessante desta distribuição, que é semelhante a uma propriedade da distribuição Normal é a seguinte:

Teorema 3.8

Seja $Y \sim M - IG(\mu, \nu, p)$. Então,

$$W = \frac{(Y - \mu)^2}{2\nu\mu^2 Y} \sim \chi_{(1)}^2.$$

Prova

Para demonstrar esta propriedade, precisa-se fazer a transformação $z \rightarrow w = z^2$, onde Z tem função de densidade dada pelo Teorema 3.6. Desta transformação resulta a densidade de uma variável $\chi_{(1)}^2$. \square

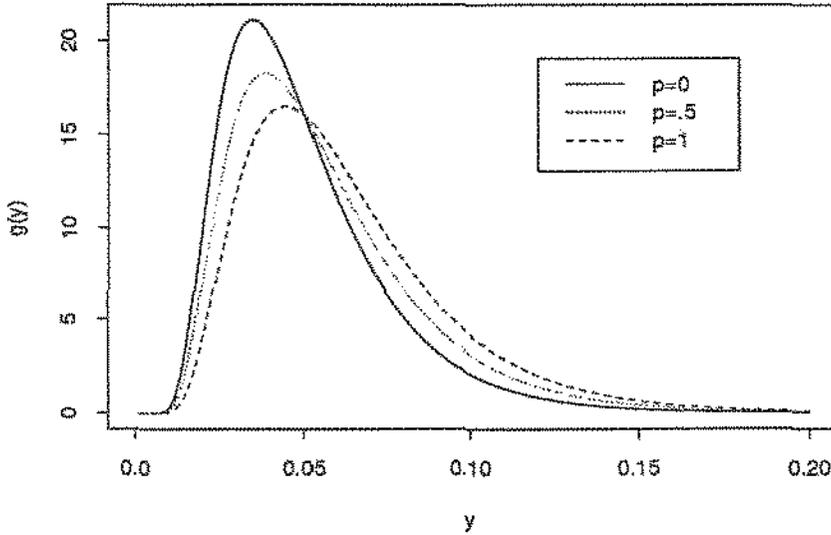


Figura 3.5: Gráfico da função de densidade g para $\mu = 20$ e $\nu = 5$ e três valores de p .

Teorema 3.9

Seja $Y \sim M - IG(\mu, \nu, \mu/(\mu + \gamma))$, onde $\gamma = \mu(1 - p)/p$ é conhecido. Então, a mistura pertence à família exponencial com dois parâmetros.

Prova

Isto pode demonstrar-se facilmente, a partir da função de densidade dada na equação (3.6), ao fazer

$$a(y, \gamma) = \frac{y + \gamma}{(2\pi y^3)^{\frac{1}{2}}}, \quad b(y) = y^{-1},$$

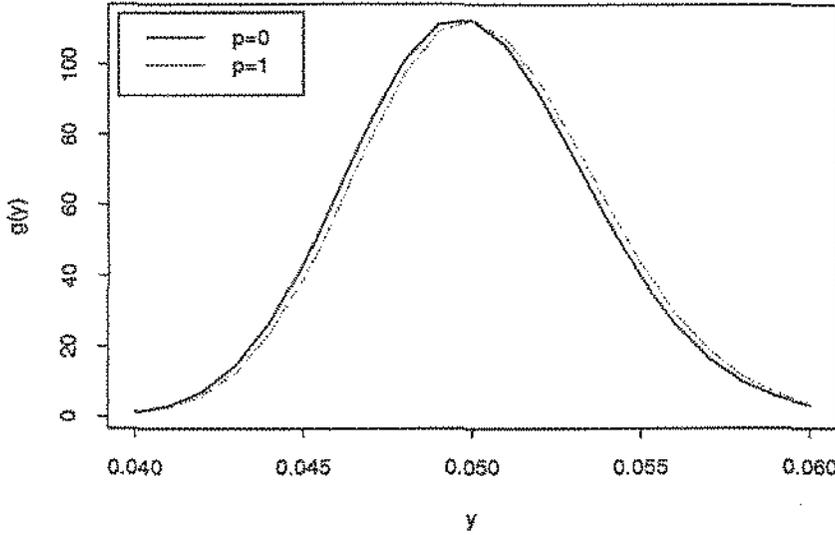


Figura 3.6: Gráfico da função de densidade g para $\mu = 20$ e $\nu = 0.1$ e dois valores de p .

$$\theta_1 = -(2\nu\mu^2)^{-1} \quad \theta_2 = -(2\nu)^{-1},$$

e

$$\kappa(\theta_1, \theta_2, \gamma) = 2(\theta_1\theta_2)^{\frac{1}{2}} - \frac{1}{2} \log(-2\theta_2) + \log \left\{ \gamma + \left(\frac{\theta_2}{\theta_1} \right)^{\frac{1}{2}} \right\}.$$

Logo, para γ conhecido, temos

$$g_*(y; \mu, \nu, p) \equiv \tilde{g}(y; \theta_1, \theta_2) = a(y, \gamma) \exp \{ \theta_1 y + \theta_2 b(y) - \kappa(\theta_1, \theta_2, \gamma) \},$$

para todo $y > 0$ e $(\theta_1, \theta_2) \in \mathbb{R}_-^2 = (-\infty, 0) \times (-\infty, 0)$. □

Deste último teorema, pode-se observar que para todo $\gamma \in [0, \infty)$, a família é regular, e para $\gamma = \infty$, temos $p = 0$ por (3.5). Neste último caso, temos a distribuição Gaussiana Inversa; a família é abrupta (Jørgensen, Seshadri e Whitmore, 1991).

Duas propriedades que serão úteis na inferência com dados censurados são expostas no seguintes teoremas:

Teorema 3.10

Seja $Y \sim M - IG(\mu, \nu, p)$. Então,

$$E \left[Y^{s+1} D_p^{-1} \mid Y > a \right] = \mu^2 \left\{ \nu (2s - 1) E \left[Y^s D_p^{-1} \mid Y > a \right] + E \left[Y^{s-1} D_p^{-1} \mid Y > a \right] + \frac{2\nu a^{s+1} f_6(a; \mu, \nu)}{S(a; \mu, \nu, p)} \right\}, \quad (3.23)$$

onde

$$D_p = 1 - p + p\mu^{-1}Y,$$

$$E \left[D_p^{-1} \mid Y > a \right] = \frac{1 - F(a; \mu, \nu)}{S(a; \mu, \nu, p)},$$

$$E \left[Y D_p^{-1} \mid Y > a \right] = \frac{\mu F(\mu^2/a; \mu, \nu)}{S(a; \mu, \nu, p)},$$

$$f_6(y; \mu, \nu) = (2\pi\nu y^3)^{-\frac{1}{2}} \exp \left[-\frac{(y - \mu)^2}{2\nu\mu^2 y} \right],$$

$F(x; \mu, \nu)$ é a função de distribuição acumulada, e $S(a; \mu, \nu, p)$ é a função de confiabilidade.

Teorema 3.11

Seja $Y \sim M - IG(\mu, \nu, p)$. Então,

$$E \left[Y^s \mid Y > a \right] = (1 - p) E \left[Y^s D_p^{-1} \mid Y > a \right] + p\mu^{-1} E \left[Y^{s+1} D_p^{-1} \mid Y > a \right], \quad (3.24)$$

onde $E \left[Y^s D_p^{-1} \mid Y > a \right]$ e $E \left[Y^{s+1} D_p^{-1} \mid Y > a \right]$ são calculados com a fórmula dada no Teorema 3.10.

As propriedades dadas nos Teoremas 3.10 e 3.11 são fáceis de demonstrar usando propriedades de integração.

3.5 Aspectos relacionados aos Modelos Exponenciais com Dispersão e Divisibilidade Infinita

Jørgensen, Seshadri e Whitmore (1991) demonstraram que a Mistura da Gaussiana Inversa com sua Recíproca Complementar é uma convolução de dois Modelos Exponencial com Dispersão, mediante técnicas desenvolvidas por Jørgensen (1987) para construir novos modelos de dispersão exponencial, chamadas Combinações e Mistura. Estas técnicas são descritas a seguir.

Sejam ED_1 e ED_2 dois modelos exponenciais com dispersão. Sejam $\tau_i(\theta_i) = E(X_i) = \mu_i$, $i = 1, 2$, as funções de valores médios com parâmetros θ_1 e θ_2 , respectivamente. Sejam κ_i e Λ_i denotando a função de cumulantes e conjunto de índices, respectivamente, para $i = 1, 2$. Suponha que

$$X_1 | X_2 = x_2 \sim ED_1^*(\lambda r + x_2 q, \theta_1), \quad X_2 \sim ED_2^*(\lambda, \theta_2'), \quad (3.25)$$

onde $r, q \in \mathfrak{R}$ são constantes tais que $\lambda r + x_2 q \in \Lambda_1 \cup \{0\}$. Aqui, interpreta-se $ED_1^*(0, \theta_1)$, como uma distribuição degenerada em zero. A função de densidade de probabilidade conjunta de $(Y_1, Y_2) = (X_1/\lambda, X_2/\lambda)$, onde as distribuições de X_1 e X_2 estão dadas em (3.25), é da forma

$$f(y_1, y_2; \lambda, \theta_1, \theta_2) = a(\lambda; y_1, y_2) \exp\{\lambda[y_1\theta_1 + y_2\theta_2 - \kappa_{12}(\theta_1, \theta_2)]\}, \quad (3.26)$$

onde

$$\theta_2 = \theta_2' - q\kappa_1(\theta_1) \quad (3.27)$$

e

$$\kappa_{12}(\theta_1, \theta_2) = r\kappa_1(\theta_1) + \kappa_2(\theta_2 + q\kappa_1(\theta_1)). \quad (3.28)$$

A distribuição dada em (3.26) é um modelo exponencial com dispersão de dimensão 2, chamado *combinação* de ED_1 e ED_2 . Esta combinação é denota por:

$$(Y_1, Y_2) \sim ED_1 \times ED_2(\mu_1, \mu_2, \nu),$$

onde $\mu_2 = \tau_2(\theta'_2)$ e $\mu_1 = \tau_1(\theta_2)(r + \mu_2 q)$ são os valores esperados de Y_1 e Y_2 , respectivamente, e $\nu = \lambda^{-1}$. As componentes ED_1 e ED_2 são conhecidas como *distribuição kernel* e *distribuição mista* da combinação, respectivamente. A distribuição marginal de Y_1 , é conhecida como *mistura da distribuição* associada com a combinação de ED_1 e ED_2 , que se denota como

$$Y_1 \sim ED_1 \wedge ED_2(\mu_1, \mu_2, \nu). \quad (3.29)$$

Por (3.26), a mistura da distribuição dada em (3.29) é um modelo exponencial com dispersão para θ_2 conhecido. Se o mesmo modelo é obtido para qualquer valor de θ_2 , então a *mistura* e a correspondente *combinação* é chamada *simples*.

A combinação definida em (3.26) e a mistura definida em (3.29) dependem das constantes r e q . O caso mais simples é obtido para $r = 0$ e $q = 1$, com a componente X_2 positiva.

Na seção anterior, mostrou-se que a mistura da Gaussiana Inversa com sua Recíproca Complementar é uma convolução de duas variáveis independentes Y_1 e T , tais que $Y_1 \sim IG(\mu, \nu)$ e T uma Bernoulli Composta, definida pela (3.18). Além disso, no Capítulo 1 se mostrou que a distribuição Gaussiana Inversa é um Modelo Exponencial com Dispersão. Também, Jørgensen (1987,1992), mostrou que a distribuição Gama ($Ga(\lambda, \theta)$) e a distribuição Binomial ($Bi(\lambda, \theta)$), são Modelos Exponenciais com Dispersão. A seguir, mediante a técnica apresentada em parágrafo

anterior, mostra-se que a variável definida em (3.18), é um modelo exponencial com dispersão.

Sejam

$$T | X_4 = x \sim Ga\left(\frac{x}{2}, \theta_3\right) \text{ e } X_4 \sim Bi(n, \theta'_4). \quad (3.30)$$

Neste caso,

$$\lambda r + xq = \frac{x}{2} \Rightarrow r = 0 \text{ e } q = \frac{1}{2}.$$

$Ga(0, \theta_3)$ é uma distribuição degenerada em zero,

$$\theta_4 = \theta'_4 - q\kappa(\theta_3) = \theta'_4 - \frac{1}{2}\kappa(\theta_3),$$

e

$$\kappa(\theta_3) = -\log(-\theta_3).$$

Logo, a distribuição $Ga \wedge Bi(n, \theta_3, \theta_4)$ é um modelo exponencial com dispersão, cuja função de densidade se deduz no seguinte parágrafo.

De acordo com (3.30), a função de densidade da distribuição condicional de T dado $X_4 = x$ é

$$f_{T|X_4}(t | x) = a_1^*(x, t) \exp\left[\theta_3 t - \frac{x}{2}\kappa_1(\theta_3)\right], \quad t \in \mathfrak{R}^+ \text{ e } x \in Z_0^+,$$

onde

$$a_1^*(x, t) = \left[\Gamma\left(\frac{x}{2}\right)\right]^{-1} t^{\left(\frac{x}{2}-1\right)} I_{(1, \dots, n)}(x), \quad \text{com } 0^0 = 1,$$

e

$$\kappa_1(\theta_3) = -\log(-\theta_3).$$

A função de densidade de X_4 é

$$f_{X_4}(x; n, \theta'_4) = a_2^*(n, x) \exp[\theta'_4 x - n\kappa_2(\theta'_4)], \quad x \in Z_0^+$$

onde

$$a_2^*(n, x) = \binom{n}{x},$$

$$\kappa_2(\theta_4) = \log(1 + e^{\theta_4}),$$

e

$$\theta_4' = \theta_4 + \log(-\theta_3)^{-\frac{1}{2}}.$$

Então, a função de densidade conjunta de T e X_4 é dada por

$$f_{T, X_4}(t, x; n, \theta_3, \theta_4) = a_3^*(t, x, n) \exp[\theta_3 t + \theta_4 x - n\kappa_3(\theta_3, \theta_4)], \quad t \in \mathfrak{R}^+ \text{ e } x \in Z_0^+,$$

onde

$$a_3^*(t, x, n) = \binom{n}{x} \left[\Gamma\left(\frac{x}{2}\right) \right]^{-1} t^{\left(\frac{x}{2}-1\right)I_{(1, \dots, n)}(x)} e^{\theta_4 x}, \quad \text{com } 0^0 = 1$$

e

$$\kappa_3(\theta_3, \theta_4) = \log\left[1 + e^{\theta_4 t - \frac{1}{2} \log(-\theta_3)}\right].$$

Logo, a densidade marginal de T é

$$f_T(t; n, \theta_3) = a_4^*(n, t) \exp[\theta_3 t - n\kappa_4(\theta_3)], \quad t \in \mathfrak{R}^+, \quad (3.31)$$

onde

$$a_4^*(n, t) = \sum_{x=0}^n \binom{n}{x} \left[\Gamma\left(\frac{x}{2}\right) \right]^{-1} t^{\left(\frac{x}{2}-1\right)I_{(1, \dots, n)}(x)} e^{\theta_4 x}$$

e

$$\kappa_4(\theta_3) = \log\left[1 + (-\theta_3)^{-\frac{1}{2}} e^{\theta_4}\right].$$

Pode observar-se que quando θ_4 é conhecido, a função dada em (3.31) é um modelo exponencial com dispersão. Sejam

$$n = 1, \theta_3 = - (2\nu\mu^2)^{-1} \text{ e } \theta_4 = \log \left(\frac{p}{1-p} \right) - \frac{1}{2} \log (2\nu\mu^2).$$

Então,

$$f_T(t; \mu, \nu, p) = (1-p) + p (2\nu\mu^2)^{-\frac{1}{2}} t^{-\frac{1}{2}} \exp \left[- (2\nu\mu^2)^{-1} t \right], \quad t > 0. \quad (3.32)$$

Esta função de densidade corresponde à mistura da distribuição Gama-Bernoulli, que é a mesma da variável T definida em (3.18).

Depois de mostrar que a variável T definida na relação (3.18) é um caso particular da mistura Gama-Binomial, vamos analisar uma representação interessante da função de densidade da convolução de uma variável Gaussiana Inversa e de uma Mistura Gama-Binomial, que inclui como caso particular, a M-IG. Para esta função de densidade, precisa-se da seguinte representação da função de densidade da Gaussiana Inversa:

$$f_{X_1}(x; \lambda_1, \theta_3) = a^*(\lambda_1, x) \exp[\theta_3 x - \lambda_1 \kappa(\theta_3)], \quad x \in \mathbb{R}^+,$$

onde

$$a^*(\lambda_1, x) = \lambda_1 (2\pi x^3)^{-\frac{1}{2}} \exp \left[-\frac{\lambda_1^2}{2x} \right],$$

$$\kappa(\theta_3) = -(-2\theta_3)^{\frac{1}{2}},$$

$$\theta_3 = - (2\nu\mu^2)^{-1} \text{ e } \lambda_1 = \nu^{-\frac{1}{2}}.$$

Além disso, seja T uma variável aleatória com função de densidade dada em (3.31). Logo, para T e X_1 independentes, $\lambda_2 = n$ e θ_4 conhecido, a convolução $T + X_1$ tem a seguinte função de densidade:

$$f_{T+X_1}(x) = \int_0^x f_{X_1}(x) f_T(x - x_1) dx_1$$

$$= a_{T+X_1}(\lambda_1, \lambda_2, x) \exp[\theta_3 x - \lambda_1 \kappa(\theta_3) - \lambda_2 \kappa_4(\theta_3)], \quad x > 0, \quad (3.33)$$

onde

$$a_{T+X_1}(\lambda_1, \lambda_2, x) = \int_0^x a^*(\lambda_1, x) a_4^*[\lambda_2, (x - x_1)] dx_1, \quad (3.34)$$

$$\kappa(\theta_3) = -(-2\theta_3)^{\frac{1}{2}}, \quad (3.35)$$

$$\kappa_4(\theta_3) = \log \left[1 + (-\theta_3)^{-\frac{1}{2}} e^{\theta_4} \right] \quad (3.36)$$

e

$$\lambda_1 = \nu^{-\frac{1}{2}}. \quad (3.37)$$

Também, pode mostrar-se que se $\lambda_2 = 1$, então, temos a distribuição Gaussiana Inversa.

Pode-se deduzir, facilmente, que a função geratriz da convolução $T + X_1$ é

$$M_{T+X_1}(s) = \exp \{ \lambda_1 [\kappa(\theta_3 + s) - \kappa(\theta_3)] + \lambda_2 [\kappa_4(\theta_3 + s) - \kappa_4(\theta_3)] \}. \quad (3.38)$$

Agora, sejam Y_1, \dots, Y_m variáveis aleatórias independentes e identicamente distribuídas com função de densidade dada por (3.33). Então, a função geratriz dos momentos de $\sum_{i=1}^m Y_i$ é:

$$M_{\sum_{i=1}^m Y_i}(s) = \exp \{ m \lambda_1 [\kappa(\theta_3 + s) - \kappa(\theta_3)] + m \lambda_2 [\kappa_4(\theta_3 + s) - \kappa_4(\theta_3)] \}. \quad (3.39)$$

Pode-se observar que (3.39) tem a mesma forma que a função geratriz dos momentos dada em (3.38). Isto indica que (3.33) é fechada com respeito à propriedade de convolução de variáveis independentes e idênticamente distribuídas com função de densidade dada em (3.33). Outro caso particular da propriedade de convolução dada em (3.38) é a propriedade de convolução dada em (3.21). Este fato pode ser mostrado ao substituir as expressões equivalentes dadas em (3.35), (3.36), mais as expressões

$$\theta_3 = -(2\nu)^{-1}, \quad \theta_4 = \log \left(\frac{p}{1-p} \right) - \frac{1}{2} \log(2\nu)$$

e

$$\lambda_1 = \mu\nu^{-\frac{1}{2}} \text{ e } \lambda_2 = n$$

em (3.38), resultando:

$$M_{T+X_1}(s) = M_{Y_1}(s; \mu, \nu/\mu^2) [1 - p + pM_{Y_3}(s; \nu)]^n.$$

Pode-se observar que o segundo fator da expressão do lado direito corresponde à mistura Gama-Binomial. Veja (3.15).

Finalmente, a variável aleatória distribuída como uma mistura de Gaussiana Inversa com sua Recíproca Complementar é infinitamente divisível. Para demonstrar isto, é preciso demonstrar que a convolução $T + X_1$ com função de densidade dada em (3.33), é infinitamente divisível, já que M-IG é um caso particular desta distribuição. Mas, pelo Teorema 1.6, só é necessário demonstrar que T e X_1 são infinitamente divisíveis, pois a divisibilidade é preservada pela convolução, e como já se mostrou, que a variável IG é infinitamente divisível (Capítulo 1). Logo, é suficiente demonstrar que a mistura Gama-Binomial é infinitamente divisível, e esta demonstração se apresenta no próximo parágrafo.

Seja $\mathfrak{F}(t)$ a transformada de Laplace da variável T e $\Phi_T(t) = \log(\mathfrak{F}(t))$. De acordo com o Teorema 1.8, a variável T é infinitamente divisível se $-\Phi_T(t)$ tem derivada completamente monótona e se $\Phi_T(0) = 0$. Mas, para o caso da mistura Gama-Binomial,

$$\mathfrak{F}(t) = \exp \{n [\kappa_4(\theta_3 - t) - \kappa_4(\theta_3)]\},$$

e logo,

$$\Phi_T(t) = n [\kappa_4(\theta_3 - t) - \kappa_4(\theta_3)]. \quad (3.40)$$

Pode observar-se de (3.40) que $\Phi_T(0) = 0$. Portanto, falta mostrar que $-\Phi_T(t)$ tem derivada completamente monótona. Pela forma da função $\Phi_T(t)$,

$$\Phi_T(t) = n\kappa_4(\theta_3 - t) + c = n \log \left[1 + (t - \theta_3)^{-\frac{1}{2}} e^{\theta_4} \right] - n\kappa_4(\theta_3).$$

Temos

$$-\frac{\partial \Phi_T(t)}{\partial t} = \frac{n}{2} \left\{ \frac{(t - \theta_3)^{-\frac{1}{2}} \exp(\theta_4)}{1 + (t - \theta_3)^{-\frac{1}{2}} \exp(\theta_4)} \right\} \geq 0, \quad \forall t > \theta_3.$$

Então, para demonstrar que a derivada $\Phi_T(t)$ é completamente monótona, só precisa mostrar que

$$h(u) = \kappa_4(-u)$$

é completamente monótona, para $u > 0$. Para tal, considere

$$h_1(x) = \log \left(\frac{x + c_1}{x} \right),$$

com

$$c_1 = \exp(\theta_4) \quad e \quad x = h_2(u) = u^{\frac{1}{2}}.$$

Então,

$$h(u) = h_1(h_2(u)).$$

Agora, de acordo com o Teorema 1.7, é preciso demonstrar que h_1 é completamente monótona e h_2 é uma função com derivada completamente monótona, para que $h(u)$ seja completamente monótona. Temos

$$(-1)^n h_1^{(n)}(x) = c_1 (x + c_1)^{-n} \geq 0, \quad \forall c_1 > 0 \text{ e } x > 0.$$

Então $h_1(x)$ é completamente monótona. Também, seja $h_3(u) = h_2^{(1)}(u) = \frac{1}{2}u^{-\frac{1}{2}}$.

Logo,

$$(-1)^n h_3^{(n)}(x) = \left(\frac{1}{2}\right)^{n+1} \left[\prod_{i=1}^n (2i-1) \right] u^{-\frac{2n+1}{2}} \geq 0, \quad \forall u > 0.$$

Então, $h_2(u)$ tem derivada completamente monótona, e assim, $h(u)$ é completamente monótona, que implica que $\Phi_T(t)$ é completamente monótona. Portanto, a variável aleatória T é infinitamente divisível, da mesma forma que a variável M-IG.

4. Inferência sobre os Parâmetros da Mistura de Distribuições Gaussianas Inversa com sua Recíproca Complementar

4.1 Introdução

Jørgensen, Seshadri, e Whitmore (1991) obtiveram o sistema de equações e a forma da matriz de informação observada que permitem encontrar as estimativas de máxima verossimilhança dos parâmetros da *M-IG*. Estas estimativas podem ser obtidas mediante procedimentos iterativos, como por exemplo Newton Raphson. Eles apresentam resultados para o caso de dados sem censura e usam a função de densidade dada na relação (3.6). Neste capítulo, se apresentam os procedimentos para realizar inferências sobre os parâmetros μ , ν e p , para o caso de dados da *M-IG* com e sem censura.

4.2 Estimação Pontual de Parâmetros para Dados Sem Censura

Seja Y_1, \dots, Y_n uma amostra aleatória de tamanho n , extraída de uma distribuição *M-IG* (μ, ν, p). Assuma que os valores observados da amostra são y_1, \dots, y_n . Logo, a função de verossimilhança é

$$L(\theta) = \prod_{i=1}^n g(y_i; \mu, \nu, p),$$

$$= \prod_{i=1}^n (1 - p + p\mu^{-1}y_i) (2\pi\nu y_i^3)^{-\frac{1}{2}} \exp \left\{ -\frac{(y_i - \mu)^2}{2\nu\mu^2 y_i} \right\}, \quad (4.1)$$

onde $\theta = (\mu, \nu, p)'$.

A log-verossimilhança é

$$\begin{aligned} \mathfrak{S}(\theta) &= \log L(\theta) \\ &= \sum_{i=1}^n \log [(1-p)\mu + py_i] - n \log \mu - \frac{1}{2} \sum_{i=1}^n \log (2\pi y_i^3) - \frac{n}{2} \log \nu - \sum_{i=1}^n \frac{(y_i - \mu)^2}{2\nu\mu^2 y_i}. \end{aligned} \quad (4.2)$$

Ao derivar a log-verossimilhança com respeito aos parâmetros, tem-se o seguinte sistema de equações:

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \mu} = \sum_{i=1}^n \frac{1-p}{(1-p)\mu + py_i} - \frac{n}{\mu} + \frac{1}{\nu\mu^3} \sum_{i=1}^n y_i - \frac{n}{\nu\mu^2} = 0, \quad (4.3)$$

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\frac{1}{\mu^2} \sum_{i=1}^n y_i - \frac{2n}{\mu} + \sum_{i=1}^n y_i^{-1} \right] = 0, \quad (4.4)$$

e

$$\frac{\partial \mathfrak{S}(\theta)}{\partial p} = \sum_{i=1}^n \frac{y_i - \mu}{(1-p)\mu + py_i} = 0. \quad (4.5)$$

Para p conhecido, o sistema de equações se reduz a

$$\sum_{i=1}^n \frac{1-p}{(1-p)\mu + py_i} + \frac{n\bar{Y}}{\nu\mu^3} - \frac{n}{\nu\mu^2} - \frac{n}{\mu} = 0, \quad (4.6)$$

e

$$-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\frac{n\bar{Y}}{\mu^2} - \frac{2n}{\mu} + n\bar{Y}_- \right] = 0 \quad (4.7)$$

onde

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n y_i \text{ e } \bar{Y}_- = \frac{1}{n} \sum_{i=1}^n y_i^{-1}.$$

O sistema de equações dado de (4.6) a (4.7) pode ser expresso como

$$\bar{Y} = \mu + \nu\mu^2 \left[1 - \frac{1}{n} \sum_{i=1}^n \frac{1}{1 + \left(\frac{p}{1-p}\right) \left(\frac{y_i}{\mu}\right)} \right] \quad (4.8)$$

e

$$\frac{1}{\bar{Y}_-} = \mu \left[\frac{\mu}{\mu^2 \nu + 2\mu - \bar{Y}} \right] < \mu, \quad (4.9)$$

desde que $\bar{Y} \geq 0$.

De (4.8) e (4.9), pode observar-se que para todo y_1, \dots, y_n positivos, a estimativa de máxima verossimilhança de μ se encontra no intervalo

$$\frac{1}{\bar{Y}_-} \leq \hat{\mu} \leq \bar{Y}.$$

$p = 0$:

Agora, se $p = 0$, (4.1) se reduz a distribuição Gaussiana Inversa. Os estimadores de máxima verossimilhança de μ e ν , para este caso são

$$\hat{\mu} = \bar{Y} \quad (4.10)$$

e

$$\hat{\nu} = \bar{Y}_- - \frac{1}{\bar{Y}}. \quad (4.11)$$

$p = 1$:

Se $p = 1$, temos para (4.1) o caso da estimação de parâmetros da distribuição da Recíproca Complementar de uma variável IG. Os estimadores de máxima verossimilhança de μ e ν , para este caso, são

$$\hat{\mu} = \frac{1}{\bar{Y}_-} \quad (4.12)$$

e

$$\hat{\nu} = \bar{Y}_- (\bar{Y}_- \bar{Y} - 1). \quad (4.13)$$

Para achar as estimativas de máxima verossimilhança de μ e ν , para $p \in (0, 1)$, pode-se usar o método de Newton Raphson. O problema geral deste método é encontrar o ponto inicial, o qual deve estar próximo das estimativas de máxima verossi-

milhança. Mas, as estimativas para o caso da distribuição Gaussiana Inversa e para a Recíproca Complementar podem ser facilmente achadas, com as fórmulas dadas de (4.10) a (4.13). Logo, pode-se aproveitar esta propriedade para construir um procedimento para achar uma solução do sistema de equações dado em (4.6) e (4.7), para qualquer $p \in (0, 1)$. Este procedimento será uma modificação do método geral que se usa para encontrar as estimativas de máxima verossimilhança quando p não é fixo.

Seja

$$Y = \text{diag} \{y_1, y_2, \dots, y_n\}, \quad (4.14)$$

$$D_p = \text{diag} \{1 - p + p\mu^{-1}y_1, \dots, 1 - p + p\mu^{-1}y_n\} \quad (4.15)$$

e

$$\underset{\sim}{1}' = (1, \dots, 1)_{1 \times n}. \quad (4.16)$$

Então, o sistema de equações dado de (4.3) a (4.5) pode ser escrito como

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \mu} = \left(\frac{1-p}{\mu} \right) \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} + \frac{1}{\nu \mu^3} \underset{\sim}{1}' Y \underset{\sim}{1} - \frac{n}{\nu \mu^2} - \frac{n}{\mu} = 0, \quad (4.17)$$

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\frac{1}{\mu^2} \underset{\sim}{1}' Y \underset{\sim}{1} - \frac{2n}{\mu} + \underset{\sim}{1}' Y^{-1} \underset{\sim}{1} \right] = 0 \quad (4.18)$$

e

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \mu} = \frac{1}{\mu} \underset{\sim}{1}' Y D_p^{-1} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} = 0. \quad (4.19)$$

Pode-se observar que este sistema de equações tem soluções múltiplas. Usando a propriedade de que para cada p pode-se encontrar uma solução, e como é fácil achar as estimativas de máxima verossimilhança para $p = 0$, pode-se usar esta propriedade para construir um procedimento para achar o máximo global ou o maior dos máximo locais no interior do espaço paramétrico da família de distribuições M-IG. Baseado nestas propriedades, propõe-se o seguinte procedimento:

I. Encontrar as estimativas de máxima verossimilhança de μ e ν para $p = 0$. Tomar estes valores como ponto inicial no método de Newton Raphson para encontrar as estimativas para $p = 0,1$. Achar as estimativas com a precisão desejada.

II. Encontrar as estimativas de μ e ν para p incrementado em 0,1, mediante o método de Newton Raphson. Usar como ponto inicial os valores de μ e ν obtidos para p sem o incremento.

III. Repetir a etapa anterior até chegar a $p = 1$.

IV. Para cada valor de p , encontrar o valor de log-verossimilhança com a fórmula dada em (4.2).

V. Usar como ponto inicial, método de Newton Raphson, as estimativas de μ , ν e p que atinge o maior valor de log-verossimilhança na etapa III. Encontrar o máximo global ou o maior dos máximos locais com a precisão desejada.

Para encontrar as estimativas de máxima verossimilhança quando p é fixo, seguir da mesma forma às etapas I e II, logo repetir a etapa II até chegar ao valor de p fixado.

Para usar o procedimento acima, se precisa da matriz de informação observada, a qual é derivada no próximo teorema:

Teorema 4.1

A matriz de informação observada é

$$I_0 = \begin{bmatrix} I_{\hat{\mu}\hat{\mu}} & I_{\hat{\mu}\hat{\nu}} & I_{\hat{\mu}\hat{p}} \\ I_{\hat{\nu}\hat{\mu}} & I_{\hat{\nu}\hat{\nu}} & I_{\hat{\nu}\hat{p}} \\ I_{\hat{p}\hat{\mu}} & I_{\hat{p}\hat{\nu}} & I_{\hat{p}\hat{p}} \end{bmatrix}, \quad (4.20)$$

onde

$$I_{\hat{\mu}\hat{\mu}} = \left(\frac{1 - \hat{p}}{\hat{\mu}} \right)^2 \underset{\sim}{1}' \hat{D}_p^{-2} \underset{\sim}{1} + \frac{3}{\hat{\nu} \hat{\mu}^4} \underset{\sim}{1}' Y \underset{\sim}{1} - \frac{2n}{\hat{\nu} \hat{\mu}^3} - \frac{n}{\hat{\mu}^2},$$

$$I_{\hat{\mu}\hat{\nu}} = I'_{\hat{\nu}\hat{\mu}} = \frac{1}{\hat{\nu}^2 \hat{\mu}^3} \underset{\sim}{1}' Y \underset{\sim}{1} - \frac{n}{\hat{\nu}^2 \hat{\mu}^2},$$

$$\begin{aligned}
I_{\hat{\mu}\hat{p}} &= I_{\hat{p}\hat{\mu}} = \frac{1}{\hat{\mu}} \underset{\sim}{1}' \hat{D}_p^{-1} \underset{\sim}{1} + \frac{1-\hat{p}}{\hat{\mu}^2} \underset{\sim}{1}' Y \hat{D}_p^{-2} \underset{\sim}{1} - \frac{1-\hat{p}}{\hat{\mu}} \underset{\sim}{1}' \hat{D}_p^{-2} \underset{\sim}{1}, \\
I_{\hat{\nu}\hat{\nu}} &= \frac{n}{2\hat{\nu}^2}, \\
I_{\hat{\nu}\hat{p}} &= I'_{\hat{p}\hat{\nu}} = 0
\end{aligned}$$

e

$$I_{\hat{p}\hat{p}} = \frac{1}{\hat{\mu}^2} \underset{\sim}{1}' Y^2 \hat{D}_p^{-2} \underset{\sim}{1} - \frac{2}{\hat{\mu}} \underset{\sim}{1}' Y \hat{D}_p^{-2} \underset{\sim}{1} + \underset{\sim}{1}' \hat{D}_p^{-2} \underset{\sim}{1}. \quad (4.21)$$

onde $\hat{\mu}$, $\hat{\nu}$ e \hat{p} são as estimativas de máxima verossimilhança de μ , ν e p , respectivamente, e

$$\hat{D}_P = \text{diag} \{1 - \hat{p} + \hat{p}\hat{\mu}^{-1}y_1, \dots, 1 - \hat{p} + \hat{p}\hat{\mu}^{-1}y_n\}.$$

Prova

Pela definição da matriz de informação observada, seus componentes se encontram aplicando as seguintes fórmulas:

$$I_{\hat{\mu}\hat{\mu}} = -\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \mu^2} \Big|_{\theta=\hat{\theta}}, \quad (4.22)$$

$$I_{\hat{\mu}\hat{\nu}} = I'_{\hat{\nu}\hat{\mu}} = -\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \mu \partial \nu} \Big|_{\theta=\hat{\theta}}, \quad (4.23)$$

$$I_{\hat{\mu}\hat{p}} = I'_{\hat{p}\hat{\mu}} = -\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \mu \partial p} \Big|_{\theta=\hat{\theta}}, \quad (4.24)$$

$$I_{\hat{\nu}\hat{\nu}} = -\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \nu^2} \Big|_{\theta=\hat{\theta}}, \quad (4.25)$$

$$I_{\hat{\nu}\hat{p}} = I'_{\hat{p}\hat{\nu}} = -\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \nu \partial p} \Big|_{\theta=\hat{\theta}} \quad (4.26)$$

e

$$I_{\hat{p}\hat{p}} = -\frac{\partial^2 \mathfrak{S}(\theta)}{\partial p^2} \Big|_{\theta=\hat{\theta}}. \quad (4.27)$$

Temos, de (4.3) a (4.5), que

$$\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \mu^2} = -\sum_{i=1}^n \frac{(1-p)^2}{[(1-p)\mu + py_i]^2} + \frac{n}{\mu^2} - \frac{3}{\nu\mu^4} \sum_{i=1}^n y_i + \frac{2n}{\nu\mu^3}, \quad (4.28)$$

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \mu \partial \nu} = -\frac{1}{\nu^2 \mu^3} \sum_{i=1}^n y_i + \frac{n}{\nu^2 \mu^2} \quad (4.29)$$

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \mu \partial p} = -\sum_{i=1}^n \frac{1}{[(1-p)\mu + py_i]} - (1-p) \sum_{i=1}^n \frac{y_i - \mu}{[(1-p)\mu + py_i]^2}, \quad (4.30)$$

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \nu^2} = \frac{n}{2\nu^2} - \frac{1}{\nu^3} \left(\frac{1}{\mu^2} \sum_{i=1}^n y_i - \frac{2n}{\mu} + \sum_{i=1}^n y_i^{-1} \right), \quad (4.31)$$

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \nu \partial p} = 0 \quad (4.32)$$

e

$$\begin{aligned} \frac{\partial^2 \mathfrak{Z}(\theta)}{\partial p^2} = & -\sum_{i=1}^n \frac{y_i^2}{[(1-p)\mu + py_i]^2} + 2\mu \sum_{i=1}^n \frac{y_i}{[(1-p)\mu + py_i]^2} \\ & - \mu^2 \sum_{i=1}^n \frac{1}{[(1-p)\mu + py_i]^2} \end{aligned} \quad (4.33)$$

Usando as notações dadas em (4.14) e (4.15), obtém-se

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \mu^2} = -\left(\frac{1-p}{\mu}\right)^2 \underset{\sim}{1} D_p^{-2} \underset{\sim}{1} - \frac{3}{\nu \mu^4} \underset{\sim}{1}' Y \underset{\sim}{1} + \frac{2n}{\nu \mu^3} + \frac{n}{\mu^2}, \quad (4.34)$$

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \mu \partial \nu} = -\frac{1}{\nu^2 \mu^3} \underset{\sim}{1}' Y \underset{\sim}{1} + \frac{n}{\nu^2 \mu^2}, \quad (4.35)$$

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \mu \partial p} = -\frac{1}{\mu} \underset{\sim}{1} D_p^{-1} \underset{\sim}{1} - \frac{1-p}{\mu^2} \underset{\sim}{1}' Y D_p^{-2} \underset{\sim}{1} + \frac{1-p}{\mu} \underset{\sim}{1}' D_p^{-2} \underset{\sim}{1}, \quad (4.36)$$

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \nu^2} = \frac{n}{2\nu^2} - \frac{1}{\nu^3} \left[\frac{1}{\mu^2} \underset{\sim}{1}' Y \underset{\sim}{1} - \frac{2n}{\mu} + \underset{\sim}{1}' Y^{-1} \underset{\sim}{1} \right], \quad (4.37)$$

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \nu \partial p} = 0 \quad (4.38)$$

e

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial p^2} = -\frac{1}{\mu^2} \underset{\sim}{1}' Y^2 D_p^{-2} \underset{\sim}{1} + \frac{2}{\mu} \underset{\sim}{1}' Y D_p^{-2} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-2} \underset{\sim}{1}. \quad (4.39)$$

Logo, substituindo as expressões equivalentes dadas de (4.34) a (4.39) nas relações dadas de (4.22) a (4.27), temos as expressões (4.20) e (4.21). \square

Teorema 4.2

A matriz de informação observada, para o caso de p fixo, é dada por:

$$I_{p0} = \begin{bmatrix} I_{\hat{\mu}\hat{\mu}} & I_{\hat{\mu}\hat{\nu}} \\ I_{\hat{\nu}\hat{\mu}} & I_{\hat{\nu}\hat{\nu}} \end{bmatrix},$$

onde

$$I_{\hat{\mu}\hat{\mu}} = \left(\frac{1-p}{\hat{\mu}} \right)^2 \mathbf{1}' \hat{D}_p^{-2} \mathbf{1} + \frac{3}{\hat{\nu}\hat{\mu}^4} \mathbf{1}' Y \mathbf{1} - \frac{2n}{\hat{\nu}\hat{\mu}^3} - \frac{n}{\hat{\mu}^2},$$

$$I_{\hat{\mu}\hat{\nu}} = I'_{\hat{\nu}\hat{\mu}} = \frac{1}{\hat{\nu}^2\hat{\mu}^3} \mathbf{1}' Y \mathbf{1} - \frac{n}{\hat{\nu}^2\hat{\mu}^2}$$

e

$$I_{\hat{\nu}\hat{\nu}} = \frac{n}{2\hat{\nu}^2}$$

onde

$$\hat{D}_p = \text{diag} \{ 1 - p + p\hat{\mu}^{-1}y_1, \dots, 1 - p + p\hat{\mu}^{-1}y_n \},$$

sendo p o valor conhecido deste parâmetro, e $\hat{\mu}$ e $\hat{\nu}$ as estimativas de máxima verossimilhança de μ e ν , obtidas ao solucionar o sistema de equações dado em (4.6) e (4.7). \square

Uma alternativa para encontrar o máximo global ou o maior dos máximos locais é o algoritmo EM, mas este método tem a desvantagem de convergir lentamente a este valor. Maiores informações sobre este algoritmo são encontradas em Titterington, Smith e Makov (1986), e Little e Rubin (1987).

4.3 Estimação por Intervalo e Testes de Hipóteses para Dados sem Censura

4.3.1 Para os parâmetros da distribuição Gaussiana Inversa

Para a distribuição Gaussiana Inversa, existem diversos artigos e inclusive um livro com procedimentos para determinar intervalos de confiança e testes exatos para os parâmetros desta distribuição. Para maiores informações ver Chhikara e Folks (1976), Folks e Chhikara (1978), Miura (1978), Davis (1980), e Chhikara e Folks (1989, o livro). Estes procedimentos se baseiam nas distribuições exatas dos estimadores. Assim, pela propriedade dada em (1.15), pode-se demonstrar que a distribuição de \bar{Y} também é Gaussiana Inversa, mas com parâmetros μ e ν/n , isto é,

$$\bar{Y} = IG(\mu, \nu/n). \quad (4.40)$$

Outra propriedade relacionada com \bar{Y} pode ser obtida da propriedade (1.14), da qual tem-se

$$\frac{n(\bar{Y} - \mu)^2}{\nu\mu^2\bar{Y}} \sim \chi_{(1)}^2. \quad (4.41)$$

Por outro lado, Tweedie (1957), demonstrou que a distribuição amostral relacionada com a estatística $\hat{\nu}$ é

$$\frac{n\hat{\nu}}{\nu} = \frac{1}{\nu} \sum_{i=1}^n \left(\frac{1}{Y_i} - \frac{1}{\bar{Y}} \right) \sim \chi_{(n-1)}^2. \quad (4.42)$$

e que \bar{Y} e $\frac{1}{\nu} \sum_{i=1}^n \left(\frac{1}{Y_i} - \frac{1}{\bar{Y}} \right)$ estão independentemente distribuídas.

Também, pode-se mostrar que

$$\frac{1}{\nu} \sum_{i=1}^n \frac{(Y_i - \mu)^2}{\mu^2 Y_i} = \frac{1}{\nu} \sum_{i=1}^n \left(\frac{1}{Y_i} - \frac{1}{\bar{Y}} \right) + \frac{n (\bar{Y} - \mu)^2}{\nu \mu^2 \bar{Y}}. \quad (4.43)$$

Estas propriedades, análogas às propriedades de amostras de distribuições normais, fazem a distribuição Gaussiana Inversa útil em Modelos Lineares Generalizados.

A distribuição Gaussiana Inversa pertence à família exponencial com dois parâmetros. Logo

$$\left(\bar{Y}, \sum_{i=1}^n \frac{1}{Y_i} \right)$$

são estatísticas conjuntamente suficientes e completas para μ e ν . Logo, pelo teorema de Lehmann-Scheffe, se demonstra que $\left(\bar{Y}, \frac{1}{n-1} \sum_{i=1}^n \left(\frac{1}{Y_i} - \frac{1}{\bar{Y}} \right) \right)$ são os estimadores não viciados de mínima variância para μ e ν , respectivamente.

A partir destas propriedades, pode-se construir os seguintes testes de hipóteses e estimadores por intervalos:

Para testar $H_0 : \mu \leq \mu_0$ vs $H_1 : \mu > \mu_0$, quando ν é desconhecido, Folks e Chhikara (1979) propuseram a estatística

$$t = \frac{n^{\frac{1}{2}} (\bar{Y} - \mu_0)}{\mu_0 (\bar{Y} \hat{V})^{\frac{1}{2}}}, \quad (4.44)$$

como a estatística do teste uniformemente mais poderoso, onde $\hat{V} = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{1}{Y_i} - \frac{1}{\bar{Y}} \right)$.

O nível de significância, α , do teste pode ser achado com

$$\alpha = G_{\nu, (n-1)}(-t) + \left(\frac{w+2n}{w-2n} \right)^{\frac{(n-2)}{2}} G_{\nu, (n-1)} \left\{ \left[-(4n + (w+2n)t^2) \right]^{\frac{1}{2}} \right\}, \quad (4.45)$$

onde G é a função de distribuição acumulada da T de Student com $(n-1)$ graus de

liberdade, e

$$w = \sum_{i=1}^n \left(\frac{y_i}{\mu_0} + \frac{\mu_0}{y_i} \right). \quad (4.46)$$

Para testar $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$, não existe um teste uniformemente mais poderoso. Mas, existe um teste uniformemente não viciado. Assim, a estatística do teste é a mesma que aquela dada em (4.44) e o nível de significância, α , pode ser achado:

$$\alpha = P(|T_{(n-1)}| > t), \quad (4.47)$$

onde $T_{(n-1)}$ é uma variável com distribuição T de Student com $n-1$ graus de liberdade.

O intervalo de confiança para μ com $(1 - \alpha)\%$ de confiança é dado por

$$\left(\frac{\bar{Y}}{1 + t_{(n-1, 1-\frac{\alpha}{2})}(\bar{Y}\hat{V})^{\frac{1}{2}}}, \frac{\bar{Y}}{1 - t_{(n-1, 1-\frac{\alpha}{2})}(\bar{Y}\hat{V})^{\frac{1}{2}}} \right), \quad (4.48)$$

onde $t_{(n-1, 1-\frac{\alpha}{2})}$ é o valor da distribuição T de Student tal que:

$$P\left(T \geq t_{(n-1, 1-\frac{\alpha}{2})}\right) = \frac{\alpha}{2}$$

e

$$\hat{V} = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{1}{Y_i} - \frac{1}{\bar{Y}} \right).$$

A partir da propriedade dada em (4.42) pode-se construir-se teste de hipóteses e estimadores por intervalo para o parâmetro ν , semelhantes à variância de distribuições Normais.

4.3.2 Para os parâmetros da distribuição da Recíproca Complementar de uma variável Gaussiana Inversa

Testes de hipóteses e estimação por intervalos para os parâmetros da distribuição Recíproca Complementar de uma variável IG, baseados na distribuição exata de alguma estatística relacionada com esta distribuição, ainda não foram desenvolvidas. Mas, pode-se aproveitar propriedades da distribuição para construir estimadores por intervalo e testes de hipóteses para os parâmetros, já que, se $Y \sim R-IG(\mu, \nu)$ então $X = Y^{-1} \sim IG(\mu^{-1}, \nu\mu^2)$. Logo, ao fazer $\mu' = \mu^{-1}$ e $\nu' = \nu\mu^2$, pode-se realizar testes de hipóteses e estimação por intervalos para μ , com as fórmulas dadas para a distribuição Gaussiana Inversa. Mas, para o parâmetro de dispersão ν , não é possível construir testes e intervalo de confiança exatos, exceto se μ é conhecido. Neste último caso, pode usar-se a propriedade dada no Teorema 2.3, inciso (i), e mais a propriedade de aditividade das variáveis χ^2 , para obter-se

$$\sum_{i=1}^n \frac{(Y_i - \mu)^2}{\nu\mu^2 Y_i} \sim \chi^2_{(n)}.$$

Com esta estatística, pode-se realizar testes de hipóteses e estimação por intervalo, de maneira semelhante à inferência realizada para a variância de distribuição Normal com média conhecida.

4.3.3 Teste de hipóteses e estimação por intervalo para os parâmetros da *M-IG*

Para testar hipóteses e construir intervalos de confiança para os parâmetros da *M-IG*, pode usar-se a distribuição assintótica da estatística da razão de verossimilhança $L(\theta_0)/L(\hat{\theta})$. Assim, para testar a hipóteses $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$, a

estatística da razão de verossimilhança (Λ) está definida por

$$\Lambda(\theta_0) = -2 \left\{ \mathfrak{S}(\theta_0) - \mathfrak{S}(\hat{\theta}) \right\}, \quad (4.49)$$

onde $\mathfrak{S}(\theta)$ denota log-verossimilhança (dado o vetor de observações \underline{y}). Pode mostrar-se que sob condições de regularidade da teoria de verossimilhança, e sob a $H_0: \theta = \theta_0$, a distribuição assintótica de $\Lambda(\theta_0)$ é χ^2 com k graus de liberdade (k dimensão de θ). De maneira semelhante, se o vetor de parâmetros θ é particionado da forma $\theta' = (\theta'_1, \theta'_2)$ e $\tilde{\theta}_2$ é o estimador de máxima verossimilhança de θ_2 dado $\theta_1 = \theta_1^0$, então

$$\Lambda(\theta_1^0, \tilde{\theta}_2) = -2 \left\{ \mathfrak{S}(\theta_1^0, \tilde{\theta}_2) - \mathfrak{S}(\hat{\theta}_1, \hat{\theta}_2) \right\}, \quad (4.50)$$

tem distribuição χ^2 com graus de liberdade igual a dimensão de θ_1 , sob a hipótese $H_0: \theta_1 = \theta_1^0$. Neste caso, o vetor θ_2 é chamada parâmetro de perturbação.

Os testes de hipóteses baseados na estatística da razão de verossimilhança possuem a propriedade de invariância sob reparametrização (Kalbfleisch e Prentice, 1980), propriedade que não possuem procedimentos baseados na teoria assintótica normal dos estimadores de máxima verossimilhança (Lawless, 1980). Esta propriedade de invariância sob reparametrização é útil quando apresenta-se a possibilidade de reparametrizar de maneira que a distribuição de $\hat{\theta}$ se aproxime mais rapidamente à normalidade. Por estas razões, para testar hipóteses dos parâmetros da $M - IG$, usamos testes assintóticos baseados na razão de verossimilhança.

Nesta dissertação, não se considera testes de hipóteses para o parâmetro p das formas

$$H_0: p = 0 \text{ versus } H_1: p > 0, \text{ ou } H_0: p = 1 \text{ versus } H_1: p < 1.$$

Testar hipóteses destes tipos constitui um problema que até agora não há sido completamente resolvido, já que os testes assintóticos da razão de verossimilhança

estão baseados sobre as condições de regularidade da teoria de verossimilhança, que nestes casos não são satisfeitas, porque $\{0, 1\}$ não pertence ao interior do intervalo $[0, 1]$. Por este motivo, só considera-se testes de hipóteses para $p_0 \in (0, 1)$ e não para p_0 na fronteira do espaço paramétrico de p .

Os intervalos de confiança calculados mediante a distribuição assintótica da estatística da razão de verossimilhança (4.49) tem melhor aproximação, em comparação aos calculados mediante a distribuição assintótica dos estimadores de máxima verossimilhança. Mas computacionalmente são mais difíceis de determinar que aqueles baseado em normalidade assintótica. Além disso, existem pouco pacotes e algoritmos implementados para calcular os intervalos de confiança baseados na estatística da razão de verossimilhança. Alguns dos poucos algoritmos foram propostos por Venzon e Moolgavkar (1988) para descobrir os valores de θ_0 tal que (4.49) atinge o ponto crítico de $\chi^2_{(k)}$. Nesta dissertação, só apresentamos intervalos de confiança baseados na normalidade assintótica dos estimadores de máxima verossimilhança, e se deixa como tema de pesquisa o desenvolvimento de procedimento que permitam encontrar intervalos de confiança baseados na estatística da razão de verossimilhança.

Para realizar os testes de hipóteses para os parâmetros da $M-IG$, consideramos que o log-verossimilhança é

$$\mathfrak{F}(\mu, \nu, p) = \sum_{i=1}^n \log [(1-p)\mu + py_i] - n \log \mu + \frac{1}{2} \sum_{i=1}^n \log (2\pi y_i^3) + \frac{n}{2} \log \nu - \sum_{i=1}^n \frac{(y_i - \mu)^2}{2\nu \mu^2 y_i}.$$

Então, apresenta-se os seguintes casos:

i) Teste de hipóteses para μ

Considere as hipóteses $H_0 : \mu = \mu_0$ e $H_1 : \mu \neq \mu_0$. Para testar estas hipóteses,

usa-se a seguinte estatística

$$\Lambda(\mu_0, \tilde{\nu}, \tilde{p}) = -2 \{ \mathfrak{S}(\mu_0, \tilde{\nu}, \tilde{p}) - \mathfrak{S}(\hat{\mu}, \hat{\nu}, \hat{p}) \},$$

onde $\tilde{\nu}$ e \tilde{p} são os estimadores de máxima verossimilhança de ν e p dado μ_0 , respectivamente, e $\hat{\mu}$, $\hat{\nu}$, e \hat{p} são os estimadores de máxima verossimilhança de μ , ν e p , respectivamente. A estatística $\Lambda(\mu_0, \tilde{\nu}, \tilde{p})$ tem uma distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : \mu = \mu_0$.

ii) Teste de hipóteses para ν

Considere a hipóteses $H_0 : \nu = \nu_0$ e $H_1 : \nu \neq \nu_0$. Para testar estas hipóteses, usa-se a seguinte estatística

$$\Lambda(\tilde{\mu}, \nu_0, \tilde{p}) = -2 \{ \mathfrak{S}(\tilde{\mu}, \nu_0, \tilde{p}) - \mathfrak{S}(\hat{\mu}, \hat{\nu}, \hat{p}) \},$$

onde $\tilde{\mu}$ e \tilde{p} são os estimadores de máxima verossimilhança de μ e p dado ν_0 , respectivamente, e $\hat{\mu}$, $\hat{\nu}$, e \hat{p} são os estimadores de máxima verossimilhança de μ , ν e p , respectivamente. A estatística $\Lambda(\tilde{\mu}, \nu_0, \tilde{p})$ tem uma distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : \nu = \nu_0$.

iii) teste de hipóteses para p

Considere a hipóteses $H_0 : p = p_0$ versus $H_1 : p \neq p_0$. Para testar esta hipóteses usa-se a seguinte estatística

$$\Lambda(\tilde{\mu}, \tilde{\nu}, p_0) = -2 \{ \mathfrak{S}(\tilde{\mu}, \tilde{\nu}, p_0) - \mathfrak{S}(\hat{\mu}, \hat{\nu}, \hat{p}) \},$$

onde $\tilde{\mu}$ e $\tilde{\nu}$ são os estimadores de máxima verossimilhança de μ e ν dado p_0 , respectivamente. $\hat{\mu}$, $\hat{\nu}$ e \hat{p} são os estimadores de máxima verossimilhança de μ , ν e p , respectivamente. A estatística $\Lambda(\tilde{\mu}, \tilde{\nu}, p_0)$ tem uma distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : p = p_0$.

iv) Intervalos de confiança

Para encontrar os intervalos de confiança, usamos as propriedades assintóticas dos estimadores de máxima verossimilhança. Assim

$$\begin{pmatrix} \hat{\mu} \\ \hat{\nu} \\ \hat{p} \end{pmatrix} \underset{\sim}{a.} N \left[\begin{pmatrix} \mu \\ \nu \\ p \end{pmatrix}, I_0^{-1} \right], \quad (4.51)$$

onde I_0 é a matriz de informação dada em (4.20).

Com esta propriedade, os limites de confiança (LC) para μ , ν e p com $(1 - \alpha)\%$ de confiança são dados por

$$LC(\mu) : \hat{\mu} \mp Z_0 \sqrt{C_{11}},$$

$$LC(\nu) : \hat{\nu} \mp Z_0 \sqrt{C_{22}}$$

e

$$LC(p) : \hat{p} \mp Z_0 \sqrt{C_{33}}$$

respectivamente, onde $Z \sim N(0, 1)$, C_{jj} é j-ésimo elemento da diagonal de I_0^{-1} e $P(Z \leq Z_0) = 1 - \frac{\alpha}{2}$. Também, pode-se definir uma região (elipsoide) de confiança simultânea de (μ, ν, p) , com (4.51).

4.4 Exemplo 4.1-Resistência de rolamentos

Liebleim e Zelen (1956) fizeram um estudo sobre resistência de rolamentos. Cada um dos valores da Tabela 4.4 representa o número de revoluções (em milhões) até a falha ocorra. Eles assumiram que os dados provém de uma distribuição Weibull.

Tabela 4.1: Número de revoluções (em milhões) até a falha ocorra em teste de rolamento.

17.88	28.92	33.00	41.52	42.12	45.60
48.40	51.84	51.96	54.12	55.56	67.80
68.64	68.64	68.88	84.12	93.12	98.64
105.12	105.84	127.92	128.04	173.40	

Lawless(1982) fez uma análise através de gráfico de probabilidade, chegando à conclusão que também pode ser modelado pela distribuição Log-normal. Para ilustrar a aplicação desta distribuição, vamos considerar inicialmente que os dados provém de M-IG, para depois determinar a qualidade de ajuste desta distribuição. As estimativas de máxima verossimilhança são obtidas seguindo o procedimento proposto na Seção 4.2 . Este procedimento está programado em S-plus, programas que estão no Apêndice em forma geral (dados censurados ou não, com covariáveis ou sem elas). Assim, o programa 2 produz a Tabela 4.2, nela podemos observar que o máximo valor da log-verossimilhança é atingido quando o valor de $p = 1$, na fronteira do espaço paramétrico de p , onde as condições de regularidade da teoria da verossimilhança não são satisfeitas. Neste caso, vamos avaliar a qualidade de ajuste por métodos gráficos.

Assim, nas Figuras 4.1 e 4.2, pode-se observar que a curva de sobrevivência estimada, assumindo que os dados se ajustam a uma distribuição Gaussiana Inversa

Tabela 4.2: Estimativas de máxima verossimilhança para dados da resistência de rolamentos, para diferentes valores de p

μ	ν	p	log-verossim
72.220870	0.004317	0.0	-113.204648
70.085610	0.004330	0.1	-113.199260
68.142702	0.004367	0.2	-113.185923
66.342192	0.004426	0.3	-113.167899
64.645238	0.004507	0.4	-113.147543
63.020382	0.004612	0.5	-113.126591
61.440723	0.004743	0.6	-113.106401
59.881516	0.004905	0.7	-113.088160
58.317764	0.005104	0.8	-113.073077
56.721126	0.005351	0.9	-113.062585
55.055053	0.005663	1.0	-113.058567

e à Recíproca Complementar, respectivamente, comparada cada uma com curva de sobrevivência estimada pelo método de Kaplan Meier, são parecidas. Este resultado pode ser devido a que os produtos das estimativas dos parâmetro μ e ν , é muito pequeno para cada valor de p . Quando isto acontece, as estimativas das densidades, para cada valor de p é aproximadamente normal. Além disso, pode ser verificado que as estimativas das médias e variâncias para o número de revoluções até a falha acontecer, são muitos parecidos entre os diferentes valores de p . Logo, para cada diferentes valores de p , obtemos funções de sobrevivência estimadas muitos parecidas. Também, podemos observar que gráficos da função de sobrevivência estimada (em ambos casos, IG e R-IG), em comparação com a função de sobrevivência estimada obtida por o método de Kaplan Meier, obtemos em ambos casos, um bom ajuste. Logo, para qualquer valor de $p \in [0, 1]$, o ajuste dos dados é razoável.

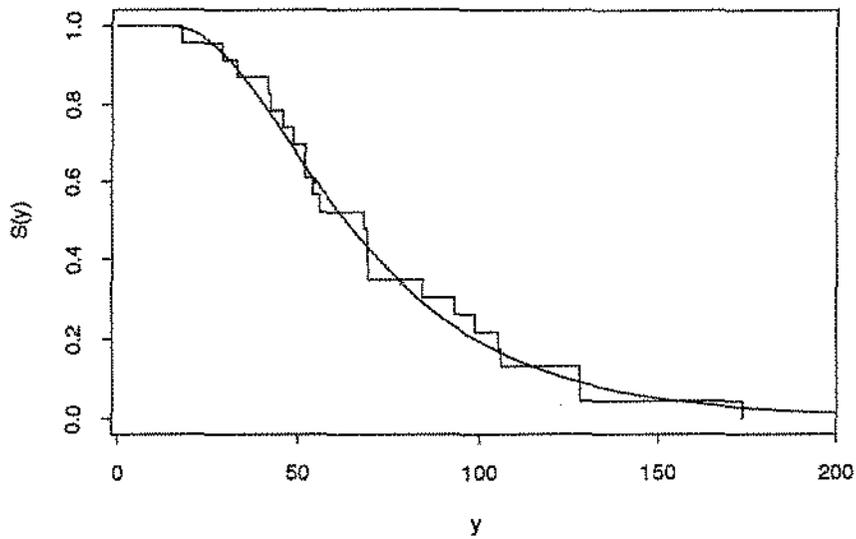


Figura 4.1: Gráfico da função de sobrevivência estimada por Kaplan Meier e de máxima verossimilhança sob IG., para dados de resistência de rolamentos.

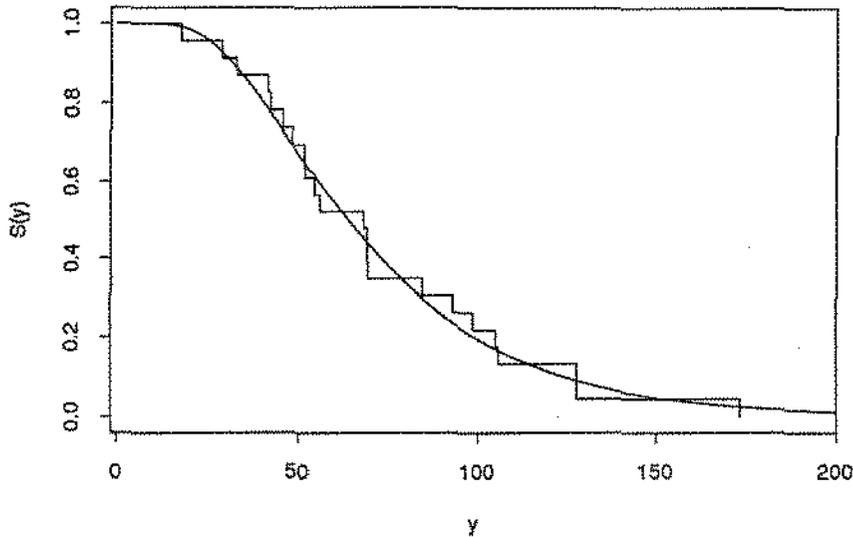


Figura 4.2: Gráfico da função de sobrevivência estimada por Kaplan Meier e de máxima verossimilhança sob R-IG., para dados de resistência de rolamentos.

4.5 Dados Censurados

Uma *censura* ocorre quando por acidente ou plano experimental, o valor da variável sob investigação não é observado para alguma unidade da amostra. Os dados censurados surgem naturalmente quando a variável aleatória de interesse é tempo de vida. Neste contexto, um dado censurado é o tempo de vida parcialmente observado para alguma unidade da amostra, da qual só conhece alguma informação parcial do evento.

Existem vários tipos de censura, para mais informação ver Kotz e Johnson (1982, 1986). Aqui, considera-se alguns deles, só para o caso de censura à direita.

Censura tipo I

Um experimento é realizado num período prefixado de tempo. A censura tipo I ocorre devido às limitações de tempo, custo do experimento e outras considerações. Isto é, seja t^* algum número préfixado, t_0 o tempo de início do experimento, t_1, t_2, \dots, t_n os tempos de entrada dos indivíduos (com $t_0 \leq t_i < t^*$, $i = 1, \dots, n$), logo o intervalo de observação para o indivíduo i é $(t_i, t^*]$. Se o indivíduo está vivo no tempo t^* , então ocorre censura tipo I. Neste caso, o número de falhas é uma variável aleatória. Além disso, se $t_1 = t_2 = \dots = t_n = t_0$, a censura se chama simples, no outro caso, ocorre censura múltipla.

Censura tipo II

Esta ocorre quando o experimentador decide restringir o experimento até obter um número determinado de falhas. Mais formalmente, sejam o_1, o_2, \dots, o_n , os tempos de falha dos n indivíduos. Então, os tempos de vida destes indivíduos são $T_1 = o_1 - t_1, T_2 = o_2 - t_2, \dots, T_n = o_n - t_n$. Sejam $T_{(1)} \leq T_{(2)} \leq \dots \leq T_{(n)}$ as estatísticas de ordem de T_1, T_2, \dots, T_n . O experimento é suspenso após da r -ésima falha e só observa-se $T_{(1)}, T_{(2)}, \dots, T_{(r)}$. Logo a amostra completa é

$$Y_1 = T_{(1)}, Y_2 = T_{(2)}, \dots, Y_r = T_{(r)}, Y_{r+1} = T_{(r)}, \dots, Y_n = T_{(R)},$$

em conjunto com um vetor de indicadores de falha observada ou censura.

Note-se que neste caso, o tempo de censura t^* é aleatório. Também, da mesma forma que para censura tipo I, se $t_1 = t_2 = \dots = t_n = t_0$, ocorre censura simples, do contrário censura múltipla.

Ambos tipos de censura apresentam-se em aplicações na engenharia, tais como em teste de vida de um grupo de transistores. Se todos são colocados no teste num tempo $t_0 = 0$ e o experimento se suspende num tempo t^* , então, pode acontecer

censura tipo I. Em outro caso, se o experimentador decide parar o experimento até que ocorram r falhas ($r < n$), então acontece censura tipo II.

Censura aleatória

Ocorre quando os tempos de censura são aleatórios. Mais especificamente, quando T_i^* , $i = 1, 2, \dots, n$, os tempos de censura, são variáveis aleatórias estocasticamente independentes, distribuídas com a mesma densidade.

Pode observar-se que em censura tipo I, quando os tempos de entrada t_1, t_2, \dots, t_n , são variáveis aleatórias, as censuras são aleatórias. Além disso, se $t_0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, tem-se o caso de *esquema de censura progressiva*. Existem várias interpretações de censura progressiva, neste caso a relevância está na entrada das unidades (para mais detalhes ver Kotz e Johnson, 1986).

4.6 Estimação de Parâmetros com Dados Censurados

Para realizar a estimação dos parâmetros com dados censurados, considera-se uma amostra com censura aleatória à direita. Assim, seja Y_1, Y_2, \dots, Y_n uma amostra aleatória de $M - IG(\mu, \nu, p)$, com função de densidade da forma (1.6). Assuma, para facilitar os cálculos matemáticos, que os r primeiros valores são observados (isto é, $Y_1 = y_1, \dots, Y_r = y_r$) e os $(n - r)$ restantes são censurados nos pontos a_{r+1}, \dots, a_n (isto é, $Y_{r+1} > a_{r+1}, \dots, Y_n > a_n$). Esta suposição facilita os cálculos, porque obtém-se fórmulas menos complicadas e, além do mais, não se perde generalidade. Seja $g(\cdot)$ a função de verossimilhança de $M - IG(\mu, \nu, p)$, dada em (4.1). Logo, a função de verossimilhança para a amostra censurada é

$$C(\theta) = \prod_{i=1}^r g(y_i; \mu, \nu, p) \prod_{i=r+1}^n S(a_i; \mu, \nu, p)$$

$$\begin{aligned}
&= \prod_{i=1}^r g(y_i; \mu, \nu, p) \prod_{i=r+1}^n \left[\int_{a_i}^{\infty} g(y_i; \mu, \nu, p) dy_i \right] \\
&= \int_c \prod_{i=1}^n g(y_i; \mu, \nu, p) dY \\
&= \int_c L(\theta) dY, \tag{4.52}
\end{aligned}$$

onde $\theta = (\mu, \nu, p)'$, $dY = \prod_{i=r+1}^n dy_i$, $L(\theta)$ é a função de verossimilhança para amostra sem censura e c o espaço dos resultado para as observações amostrais censuradas:

$$c = \{Y_1 = y_1, \dots, Y_r = y_r, Y_{r+1} > a_{r+1}, \dots, Y_n > a_n\}.$$

As estimativas de máxima verossimilhança são achada maximizando a log-verossimilhança, isto é

$$\frac{\partial \log C(\theta)}{\partial \theta} = \frac{1}{C(\theta)} \frac{\partial C(\theta)}{\partial \theta} = 0. \tag{4.53}$$

Logo, sob certas condições de regularidade,

$$\frac{\partial \log C(\theta)}{\partial \theta} = \int_c \frac{1}{L(\theta)} \frac{\partial L(\theta)}{\partial \theta} \frac{L(\theta)}{C(\theta)} dY = 0. \tag{4.54}$$

Seja $h(\theta, Y)$ uma função de (θ, Y) , e se definimos

$$E_*[h(\theta, Y)] = \int_c h(\theta, Y) \frac{L(\theta)}{C(\theta)} dY.$$

Então,

$$E_* \left[\frac{1}{L(\theta)} \frac{\partial L(\theta)}{\partial \theta} \right] = \int_c \frac{1}{L(\theta)} \frac{\partial L(\theta)}{\partial \theta} \frac{L(\theta)}{C(\theta)} dY \tag{4.55}$$

Logo,

$$E_* \left[\frac{1}{L(\theta)} \frac{\partial L(\theta)}{\partial \theta} \right] = 0.$$

Mas

$$\frac{1}{L(\theta)} \frac{\partial L(\theta)}{\partial \theta} = \frac{\partial \log L(\theta)}{\partial \theta} = \frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \tag{4.56}$$

onde $\mathfrak{S}(\theta) = \log L(\theta)$. Então,

$$E_* \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \right] = 0. \quad (4.57)$$

Pode-se observar que E_* é uma esperança tomada com respeito ao espaço de resultados c . Esta esperança tem propriedades semelhantes às da esperança comum, as quais pode ser facilmente verificadas aplicando propriedades de integração. Logo, as estimativas de máxima verossimilhança de θ são obtidas ao resolver o seguinte sistema de equações:

$$E_* \begin{bmatrix} \frac{\partial \mathfrak{S}(\theta)}{\partial \mu} \\ \frac{\partial \mathfrak{S}(\theta)}{\partial \nu} \\ \frac{\partial \mathfrak{S}(\theta)}{\partial p} \end{bmatrix} = 0. \quad (4.58)$$

Ao substituir as expressões dadas em (4.17), (4.18) e (4.19), o sistema de equações de (4.58) se transforma em

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \mu} = \frac{1-p}{\mu} \underset{\sim}{1}' E_* (D_p^{-1}) \underset{\sim}{1} + \frac{1}{\nu \mu^3} \underset{\sim}{1}' E_* (Y) \underset{\sim}{1} - \frac{n}{\nu \mu^2} - \frac{n}{\mu} = 0, \quad (4.59)$$

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\frac{1}{\mu^2} \underset{\sim}{1}' E_* (Y) \underset{\sim}{1} - \frac{2n}{\mu} + \underset{\sim}{1}' E_* (Y^{-1}) \underset{\sim}{1} \right] = 0 \quad (4.60)$$

e

$$\frac{\partial \mathfrak{S}(\theta)}{\partial p} = \frac{1}{\mu} \underset{\sim}{1}' E_* (Y D_p^{-1}) \underset{\sim}{1} - \underset{\sim}{1}' E_* (D_p^{-1}) \underset{\sim}{1} = 0. \quad (4.61)$$

Seja

$$\begin{aligned} E_* (Y^s D_p^{-1}) &= \text{diag} \{ y_1^s / (1-p + p\mu^{-1}y_1), \dots, y_r^s / (1-p + p\mu^{-1}y_r), \\ &E [Y_{r+1}^s / (1-p + p\mu^{-1}Y_{r+1}) \mid Y_{r+1} > a_{r+1}], \\ &\dots, E [Y_n^s / (1-p + p\mu^{-1}Y_n) \mid Y_n > a_n] \} \end{aligned}$$

e

$$E_* (Y^s) = \text{diag} \{ y_1, \dots, y_r, E (Y_{r+1}^s \mid Y_{r+1} > a_{r+1}), \dots, E (Y_n^s \mid Y_n > a_n) \}.$$

Para todo $s \in Z$, as expressões $E[Y_i^s / (1 - p + p\mu^{-1}Y_i) \mid Y_i > a_i]$ e $E(Y_i^s \mid Y_i > a_i)$ se acham mediante as relações equivalentes dadas nas (3.23) e (3.24), respectivamente.

Assim como no caso de dados sem censura, se p é conhecido o sistema de equações se reduz a

$$\frac{1-p}{\mu} \underset{\sim}{1}' E_* (D_p^{-1}) \underset{\sim}{1} + \frac{1}{\nu\mu^3} \underset{\sim}{1}' E_* (Y) \underset{\sim}{1} - \frac{n}{\nu\mu^2} - \frac{n}{\mu} = 0 \quad (4.62)$$

e

$$-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\frac{1}{\mu^2} \underset{\sim}{1}' E_* (Y) \underset{\sim}{1} - \frac{2n}{\mu} + \underset{\sim}{1}' E_* (Y^{-1}) \underset{\sim}{1} \right] = 0. \quad (4.63)$$

$p = 0$:

Se $p = 0$, estas equações se reduzem ao caso das equações de log-verossimilhança da distribuição Gaussiana Inversa. Se $p = 0$, temos que $E_* (D_0^{-1}) = I_n$. Com restrição $p = 0$, o sistema se reduz em

$$\begin{aligned} \frac{1}{\nu\mu^3} \underset{\sim}{1}' E_* (Y) \underset{\sim}{1} - \frac{n}{\nu\mu^2} &= 0 \\ -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\underset{\sim}{1}' E_* (Y^{-1}) \underset{\sim}{1} - \frac{n}{\mu} \right] &= 0. \end{aligned}$$

A primeira destas equações é equivalente a $\underset{\sim}{1}' E_* (Y) \underset{\sim}{1} = n\mu$, e substituindo em (4.63) gera a segunda delas. Destas equações, temos

$$\hat{\mu} = \frac{1}{n} \underset{\sim}{1}' E_* (Y) \underset{\sim}{1}, \quad (4.64)$$

e

$$\hat{\nu} = \frac{1}{n} \left[\underset{\sim}{1}' E_* (Y^{-1}) \underset{\sim}{1} - \frac{n}{\hat{\mu}} \right]. \quad (4.65)$$

Estas equações, dadas em (4.64) e (4.65) são as mesmas que se obtém ao aplicar o algoritmo EM, dado por Dempster, Laird e Rubin (1977). Para encontrar as estimativas de máximos verossimilhança, pode-se empregar o procedimento seguinte:

I. Escolher as estimativas iniciais μ e ν , as quais podem ser as estimativas de máxima verossimilhança obtidos com os dados não censurados (Thisted, 1988), calculados com as fórmulas dadas em (4.10) e (4.11).

II. Calcular $E(Y_i|Y_i > a_i)$ e $E(Y_i^{-1}|Y_i > a_i)$, para $i = r + 1, \dots, n$, de acordo com as fórmulas dadas em (1.18).

III. Substituir estes valores no lado direito das equações dadas em (4.64) e (4.65).

IV. Repetir as etapas II e III até obter as estimativas de μ e ν com a precisão desejada.

$p = 1$:

Agora, ao fazer $p = 1$ nas equações dadas em (4.62) e (4.63), temos o caso das equações de máxima verossimilhança para dados da distribuição Recíproca Complementar de uma variável IG. O sistema se reduz a

$$\underset{\sim}{1}' E_*(Y) \underset{\sim}{1} - \frac{n}{\nu \mu^2} - \frac{n}{\mu} = 0,$$

e

$$-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\frac{1}{\mu^2} \underset{\sim}{1}' E_*(Y) \underset{\sim}{1} - \frac{2n}{\mu} + \underset{\sim}{1}' E_*(Y^{-1}) \underset{\sim}{1} \right] = 0. \quad (4.66)$$

Ao resolver este sistema, temos

$$\hat{\mu} = \frac{n}{\underset{\sim}{1}' E_*(Y^{-1}) \underset{\sim}{1}} \quad (4.67)$$

e

$$\hat{\nu} = \frac{1}{n} \left[\frac{1}{\hat{\mu}^2} \underset{\sim}{1}' E_*(Y) \underset{\sim}{1} - \frac{n}{\hat{\mu}} \right]. \quad (4.68)$$

Igualmente ao caso da distribuição Gaussiana Inversa, estas equações são as mesmas que se obtém ao aplicar o algoritmo EM. Logo, para encontrar as estimativas, pode-se seguir um procedimento semelhante ao descrito para encontrar as estimativas de máxima verossimilhança dos parâmetros da Gaussiana Inversa:

I. Escolher as estimativas iniciais de μ e ν , as quais pode ser as estimativas de máxima verossimilhança obtidos com os dados não censurados, calculados com as fórmulas dadas em (4.12) e (4.13).

II. Calcular $E(Y_i|Y_i > a_i)$ e $E(Y_i^{-1}|Y_i > a_i)$, para $i = r + 1, \dots, n$, de acordo com a fórmula dada em (2.10).

III. Substituir os valores no lado direito das equações dadas em (4.67) e (4.68).

IV. Repetir as etapas II e III até obter as estimativas de μ e ν com a precisão desejada.

Para o caso de $0 < p < 1$, onde p é um valor conhecido, pode empregar-se o método de Newton Raphson. Mas, o problema deste método é encontrar o ponto inicial, o qual deve estar perto das estimativas de máxima verossimilhança. Para encontrar o ponto inicial, pode-se seguir um procedimento similar ao do caso de amostra sem censura que é uma modificação do procedimento para p não conhecido.

Voltando ao caso geral, as equações dadas de (4.59) a (4.61), como caso de dados não censurados, tem soluções múltipla. Então, para encontrar o máximo global ou o maior dos máximos, pode-se seguir o mesmo procedimento para dados não censurados. Para usar este procedimento se precisa da matriz de informação observada, a qual se define no seguinte teorema:

Teorema 4.3

A matriz de informação para amostra censuradas da M-IG é definida como

$$I_0 = \begin{bmatrix} I_{\hat{\mu}\hat{\mu}} & I_{\hat{\mu}\hat{\nu}} & I_{\hat{\mu}\hat{p}} \\ I_{\hat{\nu}\hat{\mu}} & I_{\hat{\nu}\hat{\nu}} & I_{\hat{\nu}\hat{p}} \\ I_{\hat{p}\hat{\mu}} & I_{\hat{p}\hat{\nu}} & I_{\hat{p}\hat{p}} \end{bmatrix}, \quad (4.69)$$

onde

$$I_{\hat{\mu}\hat{\nu}} = -\left(\frac{2n}{\hat{\nu}\hat{\mu}^2} + \frac{n}{\hat{\mu}^2}\right) + \frac{1}{\hat{\nu}^2\hat{\mu}^3} \underset{\sim}{1}' \left[3\hat{\nu}\hat{\mu}^2\hat{M}_1 - \hat{M}_2 + \hat{M}_1^2\right] \underset{\sim}{1} \\ - \frac{2(1-\hat{p})}{\hat{\nu}\hat{\mu}^4} \underset{\sim}{1}' \left[\hat{N}_1 - \hat{N}_0\hat{M}_1 - \frac{(1-\hat{p})\hat{\nu}\hat{\mu}^2}{2}\hat{N}_0^2\right] \underset{\sim}{1} ,$$

$$I_{\hat{\mu}\hat{p}} = \frac{1}{\hat{\nu}^2\hat{\mu}^3} \underset{\sim}{1}' \hat{M}_1 \underset{\sim}{1} - \frac{n}{\hat{\nu}^2\hat{\mu}^2} - \frac{1-\hat{p}}{2\hat{\nu}^2\hat{\mu}^3} \underset{\sim}{1}' \left[\hat{N}_1 - \hat{N}_0\hat{M}_1\right] \underset{\sim}{1} \\ - \frac{1}{2\hat{\nu}^3\hat{\mu}^5} \underset{\sim}{1}' \left[\hat{M}_2 - \hat{M}_1^2\right] \underset{\sim}{1} - \frac{1-\hat{p}}{2\hat{\nu}^2\hat{\mu}} \underset{\sim}{1}' \left[\hat{N}_{-1} - \hat{N}_0\hat{M}_{-1}\right] \underset{\sim}{1} \\ - \frac{1}{2\hat{\nu}^3\hat{\mu}^3} \underset{\sim}{1}' \left[I - \hat{M}_1\hat{M}_{-1}\right] \underset{\sim}{1} ,$$

$$I_{\hat{\nu}\hat{p}} = \frac{1}{\hat{\mu}} \underset{\sim}{1}' \left[\hat{N}_0 + \frac{(1-\hat{p})}{\hat{\mu}}\hat{N}_0\hat{N}_1 - (1-\hat{p})\hat{N}_0^2\right] \underset{\sim}{1} \\ - \frac{1}{\hat{\nu}\hat{\mu}^4} \underset{\sim}{1}' \left[\hat{N}_2 - \hat{M}_1\hat{N}_1\right] \underset{\sim}{1} + \frac{1}{\hat{\nu}\hat{\mu}^3} \underset{\sim}{1}' \left[\hat{N}_1 - \hat{M}_1\hat{N}_0\right] \underset{\sim}{1} ,$$

$$I_{\hat{\nu}\hat{\nu}} = \frac{n}{2\hat{\nu}^2} - \frac{1}{4\hat{\nu}^4\hat{\mu}^4} \underset{\sim}{1}' \left[\hat{M}_2 - \hat{M}_1^2\right] \underset{\sim}{1} - \frac{1}{2\hat{\nu}^4\hat{\mu}^2} \underset{\sim}{1}' \left[I - \hat{M}_1\hat{M}_{-1}\right] \underset{\sim}{1} \\ - \frac{1}{4\hat{\nu}^4} \underset{\sim}{1}' \left[\hat{M}_{-2} - \hat{M}_{-1}^2\right] \underset{\sim}{1} ,$$

$$I_{\hat{\nu}\hat{p}} = -\frac{1}{2\hat{\nu}^2\hat{\mu}^3} \underset{\sim}{1}' \left[\hat{N}_2 - \hat{M}_1\hat{N}_1\right] \underset{\sim}{1} - \frac{1}{2\hat{\nu}^2\hat{\mu}} \underset{\sim}{1}' \left[\hat{N}_0 - \hat{M}_{-1}\hat{N}_1\right] \underset{\sim}{1} \\ + \frac{1}{2\hat{\nu}^2\hat{\mu}^2} \underset{\sim}{1}' \left[\hat{N}_1 - \hat{M}_1\hat{N}_0\right] \underset{\sim}{1} + \frac{1}{2\hat{\nu}^2} \underset{\sim}{1}' \left[\hat{N}_{-1} - \hat{M}_{-1}\hat{N}_0\right] \underset{\sim}{1} ,$$

e

$$I_{\hat{p}\hat{p}} = \frac{1}{\hat{\mu}^2} \underset{\sim}{1}' \hat{N}_1^2 \underset{\sim}{1} - \frac{2}{\hat{\mu}} \underset{\sim}{1}' \hat{N}_1\hat{N}_0 \underset{\sim}{1} + \underset{\sim}{1}' \hat{N}_0^2 \underset{\sim}{1} ,$$

onde

$$M_s = E_*(Y^s) \text{ e } N_s = E_*(Y^s D_p^{-1}) , \quad \forall s \in Z. \quad (4.70)$$

$\hat{\mu}$, $\hat{\nu}$ e \hat{p} são as estimativas de máxima verossimilhança de μ , ν e p , respectivamente; e \hat{M}_s e \hat{N}_s são calculados com estas estimativas.

Prova

Por definição, a matriz de informação observada é dada por

$$I_0 = -\frac{\partial \log C(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta = \hat{\theta} = (\hat{\mu}, \hat{\nu}, \hat{p})'}$$

onde $\hat{\mu}$, $\hat{\nu}$, e \hat{p} são as estimativas de máxima verossimilhança de μ , ν e p , respectivamente. Por (4.52) e (4.53), temos que

$$\frac{\partial \log C(\theta)}{\partial \theta \partial \theta'} = - \left\{ \int_c \frac{1}{L(\theta)} \frac{\partial L(\theta)}{\partial \theta} \frac{L(\theta)}{C(\theta)} dY \right\} \cdot \left\{ \int_c \frac{1}{L(\theta)} \frac{\partial L(\theta)}{\partial \theta} \frac{L(\theta)}{C(\theta)} dY \right\}' + \int_c \frac{1}{L(\theta)} \frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} \frac{L(\theta)}{C(\theta)} dY.$$

Então, por (4.56), temos que

$$\frac{\partial \log C(\theta)}{\partial \theta \partial \theta'} = -E_* \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \right] E_* \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta'} \right] + E_* \left[\frac{1}{L(\theta)} \frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} \right].$$

Logo,

$$I_0 = \left\{ E_* \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \right] E_* \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta'} \right] \right\}_{\theta=\hat{\theta}} - E_* \left[\frac{1}{L(\theta)} \frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} \right]_{\theta=\hat{\theta}}.$$

Mas

$$E_* \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \right]_{\theta=\hat{\theta}} = 0$$

para $\hat{\theta}$ solução do sistema de equações dado desde (4.59) a (4.61), então

$$I_0 = - \left\{ E_* \left[\frac{1}{L(\theta)} \frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} \right] \right\}_{\theta=\hat{\theta}}.$$

Mas

$$\frac{1}{L(\theta)} \frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} = \frac{\partial^2 \mathfrak{S}(\theta)}{\partial \theta \partial \theta'} + \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \right] \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \right]'$$

Logo

$$I_0 = -E_* \left[\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \theta \partial \theta'} + \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \right] \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \right]' \right]_{\theta=\hat{\theta}}. \quad (4.71)$$

Ao fazer a partição da matriz I_0 em seus componentes, resulta

$$I_0 = \begin{bmatrix} I_{\hat{\mu}\hat{\mu}} & I_{\hat{\mu}\hat{\nu}} & I_{\hat{\mu}\hat{p}} \\ I_{\hat{\nu}\hat{\mu}} & I_{\hat{\nu}\hat{\nu}} & I_{\hat{\nu}\hat{p}} \\ I_{\hat{p}\hat{\mu}} & I_{\hat{p}\hat{\nu}} & I_{\hat{p}\hat{p}} \end{bmatrix},$$

onde

$$I_{\hat{\mu}\hat{\mu}} = -E_* \left[\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \mu^2} + \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \mu} \right]^2 \right]_{\theta=\hat{\theta}}, \quad (4.72)$$

$$I_{\hat{\mu}\hat{\nu}} = I_{\hat{\nu}\hat{\mu}} = -E_* \left[\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \mu \partial \nu} + \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \mu} \right] \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} \right] \right]_{\theta=\hat{\theta}}, \quad (4.73)$$

$$I_{\hat{\mu}\hat{p}} = I_{\hat{p}\hat{\mu}} = -E_* \left[\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \mu \partial p} + \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \mu} \right] \left[\frac{\partial \mathfrak{S}(\theta)}{\partial p} \right] \right]_{\theta=\hat{\theta}}, \quad (4.74)$$

$$I_{\hat{\nu}\hat{\nu}} = -E_* \left[\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \nu^2} + \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} \right]^2 \right]_{\theta=\hat{\theta}}, \quad (4.75)$$

$$I_{\hat{\nu}\hat{p}} = I_{\hat{p}\hat{\nu}} = -E_* \left[\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \nu \partial p} + \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} \right] \left[\frac{\partial \mathfrak{S}(\theta)}{\partial p} \right] \right]_{\theta=\hat{\theta}}, \quad (4.76)$$

e

$$I_{\hat{p}\hat{p}} = -E_* \left[\frac{\partial^2 \mathfrak{S}(\theta)}{\partial p^2} + \left[\frac{\partial \mathfrak{S}(\theta)}{\partial p} \right]^2 \right]_{\theta=\hat{\theta}}. \quad (4.77)$$

Calculo de $I_{\hat{\mu}\hat{\mu}}$

Sustituindo (4.17) e (4.34) em (4.72), resulta:

$$I_{\hat{\mu}\hat{\mu}} = -E_* \left\{ -\frac{(1-p)^2}{\mu^2} \underline{\underline{1}}' D_p^{-2} \underline{\underline{1}} - \frac{3}{\nu\mu^4} \underline{\underline{1}}' Y \underline{\underline{1}} + \left(\frac{2n}{\nu\mu^2} + \frac{n}{\mu^2} \right) \right. \\ \left. + \left[\frac{(-p)}{\mu} \underline{\underline{1}}' D_p^{-1} \underline{\underline{1}} + \frac{1}{\nu\mu^3} \underline{\underline{1}}' Y \underline{\underline{1}} - \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right) \right]^2 \right\}_{\theta=\hat{\theta}}.$$

Então,

$$I_{\hat{\mu}\hat{\mu}} = \left\{ \frac{(1-p)^2}{\mu^2} \underline{\underline{1}}' E_* (D_p^{-2}) \underline{\underline{1}} + \frac{3}{\nu\mu^4} \underline{\underline{1}}' E_* (Y) \underline{\underline{1}} - \left(\frac{2n}{\nu\mu^2} + \frac{n}{\mu^2} \right) \right. \\ - \frac{(1-p)^2}{\mu^2} \underline{\underline{1}}' E_* (D_p^{-1} \underline{\underline{1}} \underline{\underline{1}}' D_p^{-1}) \underline{\underline{1}} - \frac{2(1-p)}{\nu\mu^4} \underline{\underline{1}}' E_* (D_p^{-1} \underline{\underline{1}} \underline{\underline{1}}' Y) \underline{\underline{1}} \\ - \frac{1}{\nu^2 \mu^6} \underline{\underline{1}}' E_* (Y \underline{\underline{1}} \underline{\underline{1}}' Y) \underline{\underline{1}} - \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right)^2 \\ \left. + 2 \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right) \left[\frac{(1-p)}{\mu} \underline{\underline{1}}' E_* (D_p^{-1}) \underline{\underline{1}} + \frac{1}{\nu\mu^3} \underline{\underline{1}}' E_* (Y) \underline{\underline{1}} \right] \right\}_{\theta=\hat{\theta}}. \quad (4.78)$$

Também pode demonstrar-se que são satisfeitas as seguintes relações:

$$\begin{aligned}
 E_* \left(Y^s D_p^{-1} \underset{\sim}{1} \underset{\sim}{1}' Y^t D_p^{-1} \right) &= E_* \left(Y^{s+t} D_p^{-2} \right) - N_s N_t + N_s \underset{\sim}{1} \underset{\sim}{1}' N_t, \\
 E_* \left(Y^s D_p^{-1} \underset{\sim}{1} \underset{\sim}{1}' Y^t \right) &= N_{s+t} - N_s M_t + N_s \underset{\sim}{1} \underset{\sim}{1}' M_t, \\
 E_* \left(Y^s \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1} Y^t \right) &= N_{s+t} - M_s N_t + M_s \underset{\sim}{1} \underset{\sim}{1}' N_t, \quad e \\
 E_* \left(Y^s \underset{\sim}{1} \underset{\sim}{1}' Y^t \right) &= M_{s+t} - M_s M_t + M_s \underset{\sim}{1} \underset{\sim}{1}' M_t.
 \end{aligned} \tag{4.79}$$

Então, ao substituir (5.57) e (4.79) em (4.78), resulta em

$$\begin{aligned}
 I_{\hat{\mu}\hat{\mu}} &= \left\{ \frac{(1-p)^2}{\mu^2} \underset{\sim}{1}' E_* \left(D_p^{-2} \right) \underset{\sim}{1} + \frac{3}{\nu\mu^4} \underset{\sim}{1}' M_s \underset{\sim}{1} - \left(\frac{2n}{\nu\mu^2} + \frac{n}{\mu^2} \right) \right. \\
 &\quad - \frac{(1-p)^2}{\mu^2} \underset{\sim}{1}' \left[E_* \left(D_p^{-2} \right) - N_0^2 + N_0 \underset{\sim}{1} \underset{\sim}{1}' N_0 \right] \underset{\sim}{1} \\
 &\quad - \frac{2(1-p)}{\nu\mu^4} \underset{\sim}{1}' \left[N_1 - N_0 M_1 + N_0 \underset{\sim}{1} \underset{\sim}{1}' M_1 \right] \underset{\sim}{1} \\
 &\quad - \frac{1}{\nu^2\mu^6} \underset{\sim}{1}' \left[M_2 - M_1^2 + M_1 \underset{\sim}{1} \underset{\sim}{1}' M_1 \right] \underset{\sim}{1} - \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right)^2 \\
 &\quad \left. + 2 \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right) \left[\frac{(1-p)}{\mu} \underset{\sim}{1}' N_0 \underset{\sim}{1} + \frac{1}{\nu\mu^3} \underset{\sim}{1}' M_1 \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}}
 \end{aligned} \tag{4.80}$$

$$\begin{aligned}
 &= \left\{ \frac{3}{\nu\mu^4} \underset{\sim}{1}' M_1 \underset{\sim}{1} - \left(\frac{2n}{\nu\mu^2} + \frac{n}{\mu^2} \right) + \frac{(1-p)^2}{\mu^2} \underset{\sim}{1}' N_0^2 \underset{\sim}{1} \right. \\
 &\quad - \frac{2(1-p)}{\nu\mu^4} \underset{\sim}{1}' \left[N_1 - N_0 M_1 \right] \underset{\sim}{1} - \frac{1}{\nu^2\mu^6} \underset{\sim}{1}' \left[M_2 - M_1^2 \right] \underset{\sim}{1} \\
 &\quad - \left[\frac{(1-p)}{\mu} \underset{\sim}{1}' N_0 \underset{\sim}{1} + \frac{1}{\nu\mu^3} \underset{\sim}{1}' M_1 \underset{\sim}{1} \right]^2 - \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right)^2 \\
 &\quad \left. + 2 \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right) \left[\frac{(1-p)}{\mu} \underset{\sim}{1}' N_0 \underset{\sim}{1} + \frac{1}{\nu\mu^3} \underset{\sim}{1}' M_1 \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}}.
 \end{aligned} \tag{4.81}$$

Mas, substituindo (5.57) em (4.59), se obtém a expressão seguinte:

$$\frac{(1-p)}{\mu} \underset{\sim}{1}' N_0 \underset{\sim}{1} + \frac{1}{\nu\mu^3} \underset{\sim}{1}' M_1 \underset{\sim}{1} = \frac{n}{\nu\mu^2} + \frac{n}{\mu} \tag{4.82}$$

para $\theta = (\mu, \nu, p)'$ solução do sistema de equações dado de (4.59) a (4.61). Logo substituindo (4.82) em (4.81), dá:

$$\begin{aligned}
 I_{\hat{\mu}\hat{\mu}} &= \left\{ - \left(\frac{2n}{\nu\mu^2} + \frac{n}{\mu^2} \right) + \frac{1}{\nu^2\mu^6} \underset{\sim}{1}' \left[3\nu\mu^2 M_1 - M_2 + M_1^2 \right] \underset{\sim}{1} \right. \\
 &\quad - \frac{2(1-p)}{\nu\mu^4} \underset{\sim}{1}' \left[N_1 - N_0 M_1 - \frac{(1-p)\nu\mu^2}{2} N_0^2 \right] \underset{\sim}{1} \\
 &\quad \left. - \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right)^2 - \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right)^2 + 2 \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right)^2 \right\}_{\theta=\hat{\theta}}
 \end{aligned}$$

$$\begin{aligned}
&= -\left(\frac{2n}{\nu\mu^2} + \frac{n}{\mu^2}\right) + \frac{1}{\nu^2\mu^5} \underset{\sim}{1}' \left[3\nu\hat{\mu}^2\hat{M}_1 - \hat{M}_2 + \hat{M}_1^2\right] \underset{\sim}{1} \\
&\quad - \frac{2(1-p)}{\nu\mu^4} \underset{\sim}{1}' \left[\hat{N}_1 - \hat{N}_0\hat{M}_1 - \frac{(1-p)\nu\hat{\mu}^2}{2}\hat{N}_0^2\right] \underset{\sim}{1}
\end{aligned} \tag{4.83}$$

Cálculo de $I_{\hat{\mu}\hat{\nu}}$

Para obter $I_{\hat{\mu}\hat{\nu}}$, substitui-se (4.35), (4.17) e (4.18) em (4.73), resultando:

$$\begin{aligned}
I_{\hat{\mu}\hat{\nu}} &= -E_* \left\{ -\frac{1}{\nu^2\mu^3} \underset{\sim}{1}' Y \underset{\sim}{1} + \frac{n}{\nu^2\mu^2} + \left[\frac{(1-p)}{\mu} \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} + \frac{1}{\nu\mu^3} \underset{\sim}{1}' Y \underset{\sim}{1} - \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right) \right] \right. \\
&\quad \left. \cdot \left[-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left(\frac{1}{\mu^2} \underset{\sim}{1}' Y \underset{\sim}{1} - \frac{2n}{\mu} + \underset{\sim}{1}' Y^{-1} \underset{\sim}{1} \right) \right] \right\}_{\theta=\hat{\theta}} \\
&= \left\{ \frac{1}{\nu^2\mu^3} \underset{\sim}{1}' E_*(Y) \underset{\sim}{1} - \frac{n}{\nu^2\mu^2} + \left[\frac{(1-p)}{\mu} \underset{\sim}{1}' E_*(D_p^{-1}) \underset{\sim}{1} + \frac{1}{\nu\mu^3} \underset{\sim}{1}' E_*(Y) \underset{\sim}{1} - \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right) \right] \right. \\
&\quad \cdot \left(\frac{n}{2\nu} + \frac{n}{\nu^2\mu} \right) - \frac{1-p}{2\nu^2\mu^3} \underset{\sim}{1}' E_*(D_p^{-1} \underset{\sim}{1}\underset{\sim}{1}' Y) \underset{\sim}{1} - \frac{1}{2\nu^3\mu^3} \underset{\sim}{1}' E_*(Y \underset{\sim}{1}\underset{\sim}{1}' Y) \underset{\sim}{1} \\
&\quad - \frac{1-p}{2\nu^2\mu} \underset{\sim}{1}' E_*(D_p^{-1} \underset{\sim}{1}\underset{\sim}{1}' Y^{-1}) \underset{\sim}{1} - \frac{1}{2\nu^3\mu^3} \underset{\sim}{1}' E_*(Y \underset{\sim}{1}\underset{\sim}{1}' Y^{-1}) \underset{\sim}{1} \\
&\quad \left. + \frac{1}{2\nu^2} \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right) \cdot \left[\frac{1}{\mu^2} \underset{\sim}{1}' E_*(Y) \underset{\sim}{1} + \underset{\sim}{1}' E_*(Y^{-1}) \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}} .
\end{aligned} \tag{4.84}$$

Ao substituir as expressões dadas em (5.57) e (4.79) em (4.84), temos a seguinte expressão:

$$\begin{aligned}
I_{\hat{\mu}\hat{\nu}} &= \left\{ \frac{1}{\nu^2\mu^3} \underset{\sim}{1}' M_1 \underset{\sim}{1} - \frac{n}{\nu^2\mu^2} + \left(\frac{n}{2\nu} + \frac{n}{\nu^2\mu} \right) \right. \\
&\quad \cdot \left[\frac{(1-p)}{\mu} \underset{\sim}{1}' N_0 \underset{\sim}{1} + \frac{1}{\nu\mu^3} \underset{\sim}{1}' M_1 \underset{\sim}{1} - \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right) \right] \\
&\quad - \frac{1-p}{2\nu^2\mu^3} \underset{\sim}{1}' \left[N_1 - N_0 M_1 + N_0 \underset{\sim}{1}\underset{\sim}{1}' M_1 \right] \underset{\sim}{1} \\
&\quad - \frac{1}{2\nu^3\mu^5} \underset{\sim}{1}' \left[M_2 - M_1^2 + M_1 \underset{\sim}{1}\underset{\sim}{1}' M_1 \right] \underset{\sim}{1} \\
&\quad - \frac{1-p}{2\nu^2\mu} \underset{\sim}{1}' \left[N_{-1} - N_0 M_{-1} + N_0 \underset{\sim}{1}\underset{\sim}{1}' M_{-1} \right] \underset{\sim}{1} \\
&\quad - \frac{1}{2\nu^3\mu^3} \underset{\sim}{1}' \left[I - M_1 M_{-1} + M_1 \underset{\sim}{1}\underset{\sim}{1}' M_{-1} \right] \underset{\sim}{1} \\
&\quad \left. + \frac{1}{2\nu^2} \left(\frac{n}{\nu\mu^2} + \frac{n}{\mu} \right) \cdot \left[\frac{1}{\mu^2} \underset{\sim}{1}' M_1 \underset{\sim}{1} + \underset{\sim}{1}' M_{-1} \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}}
\end{aligned}$$

$$\begin{aligned}
&= \left\{ \frac{1}{\nu^2 \mu^3} \underset{\sim}{1}' M_1 \underset{\sim}{1} - \frac{n}{\nu^2 \mu^2} + \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right) \right. \\
&\quad \cdot \left[\frac{(1-p)}{\mu} \underset{\sim}{1}' N_0 \underset{\sim}{1} + \frac{1}{\nu \mu^3} \underset{\sim}{1}' M_1 \underset{\sim}{1} - \left(\frac{n}{\nu \mu^2} + \frac{n}{\mu} \right) \right] \\
&\quad - \frac{1-p}{2\nu^2 \mu^3} \underset{\sim}{1}' [N_1 - N_0 M_1] \underset{\sim}{1} - \frac{1}{2\nu^3 \mu^5} \underset{\sim}{1}' [M_2 - M_1^2] \underset{\sim}{1} \\
&\quad - \frac{1-p}{2\nu^2 \mu} \underset{\sim}{1}' [N_{-1} - N_0 M_{-1}] \underset{\sim}{1} - \frac{1}{2\nu^3 \mu^3} \underset{\sim}{1}' [I - M_1 M_{-1}] \underset{\sim}{1} \\
&\quad - \frac{1}{2\nu^2} \left[\frac{1}{\mu^2} \underset{\sim}{1}' M_1 \underset{\sim}{1} + \underset{\sim}{1}' M_{-1} \underset{\sim}{1} \right] \\
&\quad \cdot \left. \left[\frac{(1-p)}{\mu} \underset{\sim}{1}' N_0 \underset{\sim}{1} + \frac{1}{\nu \mu^3} \underset{\sim}{1}' M_1 \underset{\sim}{1} - \left(\frac{n}{\nu \mu^2} + \frac{n}{\mu} \right) \right] \right\}_{\theta=\hat{\theta}}. \tag{4.85}
\end{aligned}$$

Logo, substituindo (4.82) em (4.85), se obtém

$$\begin{aligned}
I_{\hat{\mu}\hat{\nu}} &= \left\{ \frac{1}{\nu^2 \mu^3} \underset{\sim}{1}' M_1 \underset{\sim}{1} - \frac{n}{\nu^2 \mu^2} - \frac{1-p}{2\nu^2 \mu^3} \underset{\sim}{1}' [N_1 - N_0 M_1] \underset{\sim}{1} \right. \\
&\quad - \frac{1}{2\nu^3 \mu^5} \underset{\sim}{1}' [M_2 - M_1^2] \underset{\sim}{1} - \frac{1-p}{2\nu^2 \mu} \underset{\sim}{1}' [N_{-1} - N_0 M_{-1}] \underset{\sim}{1} \\
&\quad \left. - \frac{1}{2\nu^3 \mu^3} \underset{\sim}{1}' [I - M_1 M_{-1}] \underset{\sim}{1} \right\}_{\theta=\hat{\theta}} \\
&= \frac{1}{\nu^2 \mu^3} \underset{\sim}{1}' \hat{M}_1 \underset{\sim}{1} - \frac{n}{\nu^2 \mu^2} - \frac{1-p}{2\nu^2 \mu^3} \underset{\sim}{1}' [\hat{N}_1 - \hat{N}_0 \hat{M}_1] \underset{\sim}{1} \\
&\quad - \frac{1}{2\nu^3 \mu^5} \underset{\sim}{1}' [\hat{M}_2 - \hat{M}_1^2] \underset{\sim}{1} - \frac{1-p}{2\nu^2 \mu} \underset{\sim}{1}' [\hat{N}_{-1} - \hat{N}_0 \hat{M}_{-1}] \underset{\sim}{1} \\
&\quad - \frac{1}{2\nu^3 \mu^3} \underset{\sim}{1}' [I - \hat{M}_1 \hat{M}_{-1}] \underset{\sim}{1}. \tag{4.86}
\end{aligned}$$

Cálculo de $I_{\hat{\mu}\hat{p}}$

Para obter este componente da matriz de informação observada, substitui-se (4.36), (4.17) e (4.19) em (4.74), se tem:

$$\begin{aligned}
I_{\hat{\mu}\hat{p}} &= -E_* \left\{ -\frac{1}{\mu} \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} - \frac{(1-p)}{\mu^2} \underset{\sim}{1}' Y D_p^{-2} \underset{\sim}{1} + \frac{(1-p)}{\mu} \underset{\sim}{1}' D_p^{-2} \underset{\sim}{1} \right. \\
&\quad + \left[\frac{(1-p)}{\mu} \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} + \frac{1}{\nu \mu^3} \underset{\sim}{1}' Y \underset{\sim}{1} - \left(\frac{n}{\nu \mu^2} + \frac{n}{\mu} \right) \right] \\
&\quad \cdot \left. \left[\frac{1}{\mu} \underset{\sim}{1}' Y D_p^{-1} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}}
\end{aligned}$$

$$\begin{aligned}
&= \left\{ \frac{1}{\mu} \underset{\sim}{1}' E_* \left(D_p^{-1} \right) \underset{\sim}{1} + \frac{(1-p)}{\mu^2} \underset{\sim}{1}' E_* \left(Y D_p^{-2} \right) \underset{\sim}{1} - \frac{(1-p)}{\mu} \underset{\sim}{1}' E_* \left(D_p^{-2} \right) \underset{\sim}{1} \right. \\
&\quad - \frac{(1-p)}{\mu^2} \underset{\sim}{1}' E_* \left(D_p^{-1} \underset{\sim}{1} \underset{\sim}{1}' Y D_p^{-1} \right) \underset{\sim}{1} - \frac{1}{\nu \mu^3} \underset{\sim}{1}' E_* \left(Y \underset{\sim}{1} \underset{\sim}{1}' Y D_p^{-1} \right) \underset{\sim}{1} \\
&\quad + \frac{(1-p)}{\mu} \underset{\sim}{1}' E_* \left(D_p^{-1} \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1} \right) \underset{\sim}{1} + \frac{1}{\nu \mu^3} \underset{\sim}{1}' E_* \left(Y \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1} \right) \underset{\sim}{1} \\
&\quad \left. + \left(\frac{n}{\nu \mu^2} + \frac{n}{\mu} \right) \cdot \left[\frac{1}{\mu} \underset{\sim}{1}' E_* \left(Y D_p^{-1} \right) \underset{\sim}{1} - \underset{\sim}{1}' E_* \left(D_p^{-1} \right) \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}}. \tag{4.87}
\end{aligned}$$

Usando as notações dadas em (5.57) e as expressões equivalentes dadas em (4.79), a relação dada em (4.87) se transforma em:

$$\begin{aligned}
I_{\hat{\mu}\hat{p}} &= \left\{ \frac{1}{\mu} \underset{\sim}{1}' N_0 \underset{\sim}{1} + \frac{(1-p)}{\mu^2} \underset{\sim}{1}' E_* \left(Y D_p^{-2} \right) \underset{\sim}{1} - \frac{(1-p)}{\mu} \underset{\sim}{1}' E_* \left(D_p^{-2} \right) \underset{\sim}{1} \right. \\
&\quad - \frac{(1-p)}{\mu^2} \underset{\sim}{1}' \left[E_* \left(Y D_p^{-2} \right) - N_0 N_1 + N_0 \underset{\sim}{1} \underset{\sim}{1}' N_1 \right] \underset{\sim}{1} \\
&\quad - \frac{1}{\nu \mu^3} \underset{\sim}{1}' \left[N_2 - M_1 N_1 + M_1 \underset{\sim}{1} \underset{\sim}{1}' N_1 \right] \underset{\sim}{1} \\
&\quad + \frac{(1-p)}{\mu} \underset{\sim}{1}' \left[E_* \left(D_p^{-2} \right) - N_0^2 + N_0 \underset{\sim}{1} \underset{\sim}{1}' N_0 \right] \underset{\sim}{1} \\
&\quad + \frac{1}{\nu \mu^3} \underset{\sim}{1}' \left[N_1 - M_1 N_0 + M_1 \underset{\sim}{1} \underset{\sim}{1}' N_0 \right] \underset{\sim}{1} \\
&\quad \left. + \left(\frac{n}{\nu \mu^2} + \frac{n}{\mu} \right) \cdot \left[\frac{1}{\mu} \underset{\sim}{1}' N_1 \underset{\sim}{1} - \underset{\sim}{1}' N_0 \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}} \\
I_{\hat{\mu}\hat{p}} &= \left\{ \frac{1}{\mu} \underset{\sim}{1}' N_0 \underset{\sim}{1} + \frac{(1-p)}{\mu^2} \underset{\sim}{1}' N_0 N_1 \underset{\sim}{1} - \frac{(1-p)}{\mu} \underset{\sim}{1}' N_0^2 \underset{\sim}{1} \right. \\
&\quad - \frac{1}{\nu \mu^3} \underset{\sim}{1}' \left[N_2 - M_1 N_1 \right] \underset{\sim}{1} + \frac{1}{\nu \mu^3} \underset{\sim}{1}' \left[N_1 - M_1 N_0 \right] \underset{\sim}{1} \\
&\quad - \left[\frac{(1-p)}{\mu} \underset{\sim}{1}' N_0 \underset{\sim}{1} + \frac{1}{\nu \mu^3} \underset{\sim}{1}' M_1 \underset{\sim}{1} - \left(\frac{n}{\nu \mu^2} + \frac{n}{\mu} \right) \right] \\
&\quad \left. \cdot \left[\frac{1}{\mu} \underset{\sim}{1}' N_1 \underset{\sim}{1} - \underset{\sim}{1}' N_0 \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}}. \tag{4.88}
\end{aligned}$$

O último somando é um produto de dois fatores. Por (4.82) o primeiro fator é zero.

Então, (4.88) se reduz a:

$$\begin{aligned}
I_{\hat{\mu}\hat{p}} &= \frac{1}{\mu} \underset{\sim}{1}' \left[\hat{N}_0 + \frac{(1-\hat{p})}{\mu} \hat{N}_0 \hat{N}_1 - (1-\hat{p}) \hat{N}_0^2 \right] \underset{\sim}{1} \\
&\quad - \frac{1}{\nu \hat{\mu}^3} \underset{\sim}{1}' \left[\hat{N}_2 - \hat{M}_1 \hat{N}_1 \right] \underset{\sim}{1} + \frac{1}{\nu \hat{\mu}^3} \underset{\sim}{1}' \left[\hat{N}_1 - \hat{M}_1 \hat{N}_0 \right] \underset{\sim}{1}. \tag{4.89}
\end{aligned}$$

Cálculo de $I_{\hat{\nu}\hat{\nu}}$

Ao substituir as expressões equivalentes dadas em (4.37) e (4.18) na (4.75), dá a seguinte expressão:

$$\begin{aligned}
 I_{\hat{\nu}\hat{\nu}} &= -E_* \left\{ \frac{n}{2\nu^2} - \frac{1}{\nu^3} \left(\frac{1}{\mu^2} \underset{\sim}{1}' Y \underset{\sim}{1} - \frac{2n}{\mu} + \underset{\sim}{1}' Y^{-1} \underset{\sim}{1} \right) \right. \\
 &\quad \left. + \left[-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left(\frac{1}{\mu^2} \underset{\sim}{1}' Y \underset{\sim}{1} - \frac{2n}{\mu} + \underset{\sim}{1}' Y^{-1} \underset{\sim}{1} \right) \right]^2 \right\}_{\theta=\hat{\theta}} \\
 &= \left\{ -\frac{n}{2\nu^2} + \frac{1}{\nu^3} \left(\frac{1}{\mu^2} \underset{\sim}{1}' E_*(Y) \underset{\sim}{1} - \frac{2n}{\mu} + \underset{\sim}{1}' E_*(Y^{-1}) \underset{\sim}{1} \right) \right. \\
 &\quad - \frac{1}{4\nu^4 \mu^4} \underset{\sim}{1}' E_* \left(Y \underset{\sim}{1} \underset{\sim}{1}' Y \right) \underset{\sim}{1} - \frac{1}{2\nu^4 \mu^2} \underset{\sim}{1}' E_* \left(Y \underset{\sim}{1} \underset{\sim}{1}' Y^{-1} \right) \underset{\sim}{1} \\
 &\quad - \frac{1}{4\nu^4} \underset{\sim}{1}' E_* \left(Y^{-1} \underset{\sim}{1} \underset{\sim}{1}' Y^{-1} \right) \underset{\sim}{1} - \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right)^2 \\
 &\quad \left. + \frac{1}{\nu^2} \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right) \cdot \left[\frac{1}{\mu^2} \underset{\sim}{1}' E_*(Y) \underset{\sim}{1} + \underset{\sim}{1}' E_*(Y^{-1}) \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}}. \tag{4.90}
 \end{aligned}$$

Usando as notações dadas em (5.57) e as expressões equivalentes dadas em (4.79), a relação dada em (4.90) se transforma em:

$$\begin{aligned}
 I_{\hat{\nu}\hat{\nu}} &= \left\{ -\frac{n}{2\nu^2} - \frac{2n}{\nu^3 \mu} + \frac{1}{\nu^3} \left(\frac{1}{\mu^2} \underset{\sim}{1}' M_1 \underset{\sim}{1} + \underset{\sim}{1}' M_{-1} \underset{\sim}{1} \right) \right. \\
 &\quad - \frac{1}{4\nu^4 \mu^4} \underset{\sim}{1}' \left[M_2 - M_1^2 + M_1 \underset{\sim}{1} \underset{\sim}{1}' M_1 \right] \underset{\sim}{1} \\
 &\quad - \frac{1}{2\nu^4 \mu^2} \underset{\sim}{1}' \left[I - M_1 M_{-1} + M_1 \underset{\sim}{1} \underset{\sim}{1}' M_{-1} \right] \underset{\sim}{1} \\
 &\quad - \frac{1}{4\nu^4} \underset{\sim}{1}' \left[M_{-2} - M_{-1}^2 + M_{-1} \underset{\sim}{1} \underset{\sim}{1}' M_{-1} \right] \underset{\sim}{1} \\
 &\quad - \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right)^2 + \frac{1}{\nu^2} \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right) \\
 &\quad \left. \cdot \left(\frac{1}{\mu^2} \underset{\sim}{1}' M_1 \underset{\sim}{1} + \underset{\sim}{1}' M_{-1} \underset{\sim}{1} \right) \right\}_{\theta=\hat{\theta}} \\
 &= \left\{ -\frac{n}{2\nu^2} - \frac{2n}{\nu^3 \mu} + \frac{1}{\nu^3} \left(\frac{1}{\mu^2} \underset{\sim}{1}' M_1 \underset{\sim}{1} + \underset{\sim}{1}' M_{-1} \underset{\sim}{1} \right) - \frac{1}{4\nu^4 \mu^4} \underset{\sim}{1}' [M_2 - M_1^2] \underset{\sim}{1} \right. \\
 &\quad - \frac{1}{2\nu^4 \mu^2} \underset{\sim}{1}' [I - M_1 M_{-1}] \underset{\sim}{1} - \frac{1}{4\nu^4} \underset{\sim}{1}' [M_{-2} - M_{-1}^2] \underset{\sim}{1} \\
 &\quad - \frac{1}{4\nu^4} \left(\frac{1}{\mu^2} \underset{\sim}{1}' M_1 \underset{\sim}{1} + \underset{\sim}{1}' M_{-1} \underset{\sim}{1} \right)^2 - \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right)^2 \\
 &\quad \left. + \frac{1}{\nu^2} \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right) \cdot \left(\frac{1}{\mu^2} \underset{\sim}{1}' M_1 \underset{\sim}{1} + \underset{\sim}{1}' M_{-1} \underset{\sim}{1} \right) \right\}_{\theta=\hat{\theta}}. \tag{4.91}
 \end{aligned}$$

Mas, ao substituir (5.57) em (4.60), resulta:

$$\frac{1}{2\nu^2} \left[\frac{1}{\mu^2} \underset{\sim}{1}' M_1 \underset{\sim}{1} + \underset{\sim}{1}' M_{-1} \underset{\sim}{1} \right] = \frac{n}{2\nu} + \frac{n}{\nu^2 \mu}, \quad (4.92)$$

para todo μ , ν e p solução do sistema de equações dado de (4.59) a (4.61).

Mediante a expressão equivalente dada na relação (4.92), a relação dada em (4.91), se transforma em:

$$\begin{aligned} I_{\nu\hat{\nu}} &= \left\{ -\frac{n}{2\nu^2} - \frac{2n}{\nu^3 \mu} + \frac{2}{\nu} \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right) - \frac{1}{4\nu^4 \mu^4} \underset{\sim}{1}' [M_2 - M_1^2] \underset{\sim}{1} \right. \\ &\quad - \frac{1}{2\nu^4 \mu^2} \underset{\sim}{1}' [I - M_1 M_{-1}] \underset{\sim}{1} - \frac{1}{4\nu^4} \underset{\sim}{1}' [M_{-2} - M_{-1}^2] \underset{\sim}{1} \\ &\quad - \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right)^2 - \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right)^2 \\ &\quad \left. + 2 \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right)^2 \right\}_{\theta=\hat{\theta}} \\ &= \frac{n}{2\hat{\nu}^2} - \frac{1}{4\hat{\nu}^4 \hat{\mu}^4} \underset{\sim}{1}' [\hat{M}_2 - \hat{M}_1^2] \underset{\sim}{1} - \frac{1}{2\hat{\nu}^4 \hat{\mu}^2} \underset{\sim}{1}' [I - \hat{M}_1 \hat{M}_{-1}] \underset{\sim}{1} \\ &\quad - \frac{1}{4\hat{\nu}^4} \underset{\sim}{1}' [\hat{M}_{-2} - \hat{M}_{-1}^2] \underset{\sim}{1}. \end{aligned} \quad (4.93)$$

Cálculo de $I_{\nu\hat{p}}$

Substituindo as expressões equivalentes dadas em (4.38), (4.18) e (4.19) em (4.76), temos a seguinte expressão:

$$\begin{aligned} I_{\nu\hat{p}} &= -E_* \left\{ \left[-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left(\frac{1}{\mu^2} \underset{\sim}{1}' Y \underset{\sim}{1} - \frac{2n}{\mu} + \underset{\sim}{1}' Y^{-1} \underset{\sim}{1} \right) \right] \left(\frac{1}{\mu} \underset{\sim}{1}' Y D_p^{-1} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \right) \right\} \\ &= \left\{ \left(\frac{n}{2\nu} + \frac{n}{\nu^2 \mu} \right) \cdot \left[\frac{1}{\mu} \underset{\sim}{1}' E_* (Y D_p^{-1}) \underset{\sim}{1} - \underset{\sim}{1}' E_* (D_p^{-1}) \underset{\sim}{1} \right] \right. \\ &\quad - \frac{1}{2\nu^2 \mu^3} \underset{\sim}{1}' E_* \left(Y \underset{\sim}{1} \underset{\sim}{1}' Y D_p^{-1} \right) \underset{\sim}{1} - \frac{1}{2\nu^2 \mu} \underset{\sim}{1}' E_* \left(Y^{-1} \underset{\sim}{1} \underset{\sim}{1}' Y D_p^{-1} \right) \underset{\sim}{1} \\ &\quad \left. + \frac{1}{2\nu^2 \mu^2} \underset{\sim}{1}' E_* \left(Y \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1} \right) \underset{\sim}{1} + \frac{1}{2\nu^2} \underset{\sim}{1}' E_* \left(Y^{-1} \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1} \right) \underset{\sim}{1} \right\}_{\theta=\hat{\theta}}. \end{aligned} \quad (4.94)$$

Usando as notações dadas em (5.57) e as expressões equivalentes dadas em (4.79), a relação dada em (4.94) se transforma em:

$$\begin{aligned}
I_{\hat{\nu}\hat{p}} &= \left\{ \left(\frac{n}{2\nu} + \frac{n}{\nu^2\mu} \right) \cdot \left(\frac{1}{\mu} \underset{\sim}{1}' N_1 \underset{\sim}{1} - \underset{\sim}{1}' N_0 \underset{\sim}{1} \right) \right. \\
&\quad - \frac{1}{2\nu^2\mu^3} \underset{\sim}{1}' \left[N_2 - M_1 N_1 + M_1 \underset{\sim}{1}\underset{\sim}{1}' N_1 \right] \underset{\sim}{1} \\
&\quad - \frac{1}{2\nu^2\mu} \underset{\sim}{1}' \left[N_0 - M_{-1} N_1 + M_{-1} \underset{\sim}{1}\underset{\sim}{1}' N_1 \right] \underset{\sim}{1} \\
&\quad + \frac{1}{2\nu^2\mu^2} \underset{\sim}{1}' \left[N_1 - M_1 N_0 + M_1 \underset{\sim}{1}\underset{\sim}{1}' N_0 \right] \underset{\sim}{1} \\
&\quad \left. + \frac{1}{2\nu^2} \underset{\sim}{1}' \left[N_{-1} - M_{-1} N_0 + M_{-1} \underset{\sim}{1}\underset{\sim}{1}' N_0 \right] \underset{\sim}{1} \right\}_{\theta=\hat{\theta}} \\
&= \left\{ - \left[\frac{1}{2\nu^2} \left(\frac{1}{\mu^2} \underset{\sim}{1}' M_1 \underset{\sim}{1} + \underset{\sim}{1}' M_{-1} \underset{\sim}{1} \right) - \left(\frac{n}{2\nu} + \frac{n}{\nu^2\mu} \right) \cdot \left(\frac{1}{\mu} \underset{\sim}{1}' N_1 \underset{\sim}{1} - \underset{\sim}{1}' N_0 \underset{\sim}{1} \right) \right. \right. \\
&\quad - \frac{1}{2\nu^2\mu^3} \underset{\sim}{1}' [N_2 - M_1 N_1] \underset{\sim}{1} - \frac{1}{2\nu^2\mu} \underset{\sim}{1}' [N_0 - M_{-1} N_1] \underset{\sim}{1} \\
&\quad \left. \left. + \frac{1}{2\nu^2\mu^2} \underset{\sim}{1}' [N_1 - M_1 N_0] \underset{\sim}{1} + \frac{1}{2\nu^2} \underset{\sim}{1}' [N_{-1} - M_{-1} N_0] \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}}. \tag{4.95}
\end{aligned}$$

Mas, pela expressão equivalente dado na (4.92),

$$\begin{aligned}
I_{\hat{\nu}\hat{p}} &= -\frac{1}{2\nu^2\mu^3} \underset{\sim}{1}' [\hat{N}_2 - \hat{M}_1 \hat{N}_1] \underset{\sim}{1} - \frac{1}{2\nu^2\mu} \underset{\sim}{1}' [\hat{N}_0 - \hat{M}_{-1} \hat{N}_1] \underset{\sim}{1} \\
&\quad + \frac{1}{2\nu^2\mu^2} \underset{\sim}{1}' [\hat{N}_1 - \hat{M}_1 \hat{N}_0] \underset{\sim}{1} + \frac{1}{2\nu^2} \underset{\sim}{1}' [\hat{N}_{-1} - \hat{M}_{-1} \hat{N}_0] \underset{\sim}{1}. \tag{4.96}
\end{aligned}$$

Cálculo de $I_{\hat{p}\hat{p}}$

Substituindo as expressões equivalentes dadas em (4.39) e (4.19) em (4.77),

temos:

$$I_{\hat{p}\hat{p}} = -E_* \left\{ -\frac{1}{\mu^2} \underset{\sim}{1}' Y^2 D_p^{-2} \underset{\sim}{1} + \frac{2}{\mu} \underset{\sim}{1}' Y D_p^{-2} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-2} \underset{\sim}{1} + \left(\frac{1}{\mu} \underset{\sim}{1}' Y D_p^{-1} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \right)^2 \right\}_{\theta=\hat{\theta}}$$

$$\begin{aligned}
I_{\hat{p}\hat{p}} &= \left\{ \frac{1}{\mu^2} \underset{\sim}{1}' E_* \left(Y^2 D_p^{-2} \right) \underset{\sim}{1} - \frac{2}{\mu} \underset{\sim}{1}' E_* \left(Y D_p^{-2} \right) \underset{\sim}{1} + \underset{\sim}{1}' E_* \left(D_p^{-2} \right) \underset{\sim}{1} \right. \\
&\quad - \frac{1}{\mu^2} \underset{\sim}{1}' E_* \left(Y D_p^{-1} \underset{\sim}{1}\underset{\sim}{1}' Y D_p^{-1} \right) \underset{\sim}{1} + \frac{2}{\mu} \underset{\sim}{1}' E_* \left(Y D_p^{-1} \underset{\sim}{1}\underset{\sim}{1}' D_p^{-1} \right) \underset{\sim}{1} \\
&\quad \left. - \underset{\sim}{1}' E_* \left(D_p^{-1} \underset{\sim}{1}\underset{\sim}{1}' D_p^{-1} \right) \underset{\sim}{1} \right\}_{\theta=\hat{\theta}}. \tag{4.97}
\end{aligned}$$

Usando as notações dadas em (5.57) e as expressões equivalentes dadas em (4.79), a relação dada em (4.97), se transforma em:

$$\begin{aligned}
 I_{\hat{p}\hat{p}} &= \left\{ \frac{1}{\hat{\mu}^2} \underset{\sim}{1}' E_* (Y^2 D_p^{-2}) \underset{\sim}{1} - \frac{2}{\hat{\mu}} \underset{\sim}{1}' E_* (Y D_p^{-2}) \underset{\sim}{1} + \underset{\sim}{1}' E_* (D_p^{-2}) \underset{\sim}{1} \right. \\
 &\quad - \frac{1}{\hat{\mu}^2} \underset{\sim}{1}' \left[E_* (Y^2 D_p^{-2}) - N_1^2 + N_1 \underset{\sim}{1} \underset{\sim}{1}' N_1 \right] \underset{\sim}{1} \\
 &\quad + \frac{2}{\hat{\mu}} \underset{\sim}{1}' \left[E_* (Y D_p^{-2}) - N_1 N_0 + N_1 \underset{\sim}{1} \underset{\sim}{1}' N_0 \right] \underset{\sim}{1} \\
 &\quad \left. - \underset{\sim}{1}' \left[E_* (D_p^{-2}) - N_0^2 + N_0 \underset{\sim}{1} \underset{\sim}{1}' N_0 \right] \underset{\sim}{1} \right\}_{\theta=\hat{\theta}} \\
 &= \left\{ \frac{1}{\hat{\mu}^2} \underset{\sim}{1}' N_1^2 \underset{\sim}{1} - \frac{2}{\hat{\mu}} \underset{\sim}{1}' N_1 N_0 \underset{\sim}{1} + \underset{\sim}{1}' N_0^2 \underset{\sim}{1} - \left(\frac{1}{\hat{\mu}} \underset{\sim}{1}' N_1 \underset{\sim}{1} - \underset{\sim}{1}' N_0 \underset{\sim}{1} \right)^2 \right\}_{\theta=\hat{\theta}}.
 \end{aligned}$$

Mas, ao substituir a notação dada em (5.57) na (4.61), temos:

$$\frac{1}{\hat{\mu}} \underset{\sim}{1}' N_1 \underset{\sim}{1} - \underset{\sim}{1}' N_0 \underset{\sim}{1} = 0, \quad (4.98)$$

para todo μ, ν e p solução do sistema de equações dado de (4.59) a (4.61).

Logo, por (4.98), $I_{\hat{p}\hat{p}}$ pode ser escrita como:

$$I_{\hat{p}\hat{p}} = \frac{1}{\hat{\mu}^2} \underset{\sim}{1}' \hat{N}_1^2 \underset{\sim}{1} - \frac{2}{\hat{\mu}} \underset{\sim}{1}' \hat{N}_1 \hat{N}_0 \underset{\sim}{1} + \underset{\sim}{1}' \hat{N}_0^2 \underset{\sim}{1}. \quad (4.99)$$

□

Do teorema anterior, podemos deduzir que para p conhecido a matriz de informação é

$$I_{f_0} = \begin{bmatrix} I_{\hat{\mu}\hat{\mu}} & I_{\hat{\mu}\hat{\nu}} \\ I_{\hat{\nu}\hat{\mu}} & I_{\hat{\nu}\hat{\nu}} \end{bmatrix},$$

onde

$$\begin{aligned}
 I_{\hat{\mu}\hat{\mu}} &= - \left(\frac{2n}{\hat{\nu}\hat{\mu}^2} + \frac{n}{\hat{\mu}^2} \right) + \frac{1}{\hat{\nu}^2 \hat{\mu}^6} \underset{\sim}{1}' \left[3\hat{\nu}\hat{\mu}^2 \hat{M}_1 - \hat{M}_2 + \hat{M}_1^2 \right] \underset{\sim}{1} \\
 &\quad - \frac{2(1-p)}{\hat{\nu}\hat{\mu}^4} \underset{\sim}{1}' \left[\hat{N}_1 - \hat{N}_0 \hat{M}_1 - \frac{(1-p)\hat{\nu}\hat{\mu}^2}{2} \hat{N}_0^2 \right] \underset{\sim}{1}, \quad (4.100)
 \end{aligned}$$

$$\begin{aligned}
I_{\hat{\mu}\hat{\nu}} = I'_{\hat{\nu}\hat{\mu}} &= \frac{1}{\hat{\nu}^2 \hat{\mu}^3} \underset{\sim}{1}' \hat{M}_1 \underset{\sim}{1} - \frac{n}{\hat{\nu}^2 \hat{\mu}^2} \\
&\quad - \frac{1-p}{2\hat{\nu}^2 \hat{\mu}^2} \underset{\sim}{1}' [\hat{N}_1 - \hat{N}_0 \hat{M}_1] \underset{\sim}{1} - \frac{1}{2\hat{\nu}^2 \hat{\mu}^2} \underset{\sim}{1}' [\hat{M}_2 - \hat{M}_1^2] \underset{\sim}{1} \\
&\quad - \frac{1-p}{2\hat{\nu}^2 \hat{\mu}} \underset{\sim}{1}' [\hat{N}_{-1} - \hat{N}_0 \hat{M}_{-1}] \underset{\sim}{1} - \frac{1}{2\hat{\nu}^2 \hat{\mu}^3} \underset{\sim}{1}' [I - \hat{M}_1 \hat{M}_{-1}] \underset{\sim}{1}
\end{aligned} \tag{4.101}$$

e

$$\begin{aligned}
I_{\hat{\nu}\hat{\nu}} &= \frac{n}{2\hat{\nu}^2} - \frac{1}{4\hat{\nu}^4 \hat{\mu}^4} \underset{\sim}{1}' [\hat{M}_2 - \hat{M}_1^2] \underset{\sim}{1} - \frac{1}{2\hat{\nu}^4 \hat{\mu}^2} \underset{\sim}{1}' [I - \hat{M}_1 \hat{M}_{-1}] \underset{\sim}{1} \\
&\quad - \frac{1}{4\hat{\nu}^4} \underset{\sim}{1}' [\hat{M}_{-2} - \hat{M}_{-1}^2] \underset{\sim}{1}.
\end{aligned} \tag{4.102}$$

Se $p = 0$, temos que

$$\begin{aligned}
\underset{\sim}{1}' M_1 \underset{\sim}{1} &= n\mu, \\
\underset{\sim}{1}' M_{-1} \underset{\sim}{1} &= n\nu + \frac{n}{\mu}, \\
N_s &= M_s \text{ e} \\
N_0 &= I_n.
\end{aligned}$$

Então, com $p=0$,

$$I_{\hat{\mu}\hat{\mu}} = -\frac{2n}{\hat{\nu}\hat{\mu}^2} + \frac{1}{\hat{\nu}^2 \hat{\mu}^6} \underset{\sim}{1}' [3\hat{\nu}\hat{\mu}^2 \hat{M}_1 - \hat{M}_2 + \hat{M}_1^2] \underset{\sim}{1}, \tag{4.103}$$

$$I_{\hat{\mu}\hat{\nu}} = I'_{\hat{\nu}\hat{\mu}} = -\frac{1}{2\hat{\nu}^3 \hat{\mu}^5} \underset{\sim}{1}' [\hat{M}_2 - \hat{M}_1^2] \underset{\sim}{1} - \frac{1}{2\hat{\nu}^3 \hat{\mu}^3} \underset{\sim}{1}' [I - \hat{M}_1 \hat{M}_{-1}] \underset{\sim}{1} \tag{4.104}$$

e

$$\begin{aligned}
I_{\hat{\nu}\hat{\nu}} &= \frac{n}{2\hat{\nu}^2} - \frac{1}{4\hat{\nu}^4 \hat{\mu}^4} \underset{\sim}{1}' [\hat{M}_2 - \hat{M}_1^2] \underset{\sim}{1} - \frac{1}{2\hat{\nu}^4 \hat{\mu}^2} \underset{\sim}{1}' [I - \hat{M}_1 \hat{M}_{-1}] \underset{\sim}{1} \\
&\quad - \frac{1}{4\hat{\nu}^4} \underset{\sim}{1}' [\hat{M}_{-2} - \hat{M}_{-1}^2] \underset{\sim}{1}.
\end{aligned} \tag{4.105}$$

Se $p = 1$, temos que

$$\begin{aligned}
\underset{\sim}{1}' M_1 \underset{\sim}{1} &= n\mu + n\nu\mu^2, \\
\underset{\sim}{1}' M_1 \underset{\sim}{1} + \mu^2 \underset{\sim}{1}' M_{-1} \underset{\sim}{1} &= n\nu\mu^2 + 2n\mu, \text{ e} \\
\underset{\sim}{1}' M_{-1} \underset{\sim}{1} &= \frac{n}{\mu}.
\end{aligned}$$

Então com $p = 1$, temos que

$$I_{\hat{\mu}\hat{\mu}} = - \left(\frac{2n}{\hat{\nu}\hat{\mu}^2} + \frac{n}{\hat{\mu}^2} \right) + \frac{1}{\hat{\nu}^2\hat{\mu}^6} \underset{\sim}{1}' \left[3\hat{\nu}\hat{\mu}^2\hat{M}_1 - \hat{M}_2 + \hat{M}_1^2 \right] \underset{\sim}{1}, \quad (4.106)$$

$$I_{\hat{\mu}\hat{\nu}} = I'_{\hat{\nu}\hat{\mu}} = \frac{1}{\hat{\nu}^2\hat{\mu}^3} \underset{\sim}{1}' \hat{M}_1 \underset{\sim}{1} - \frac{n}{\hat{\nu}^2\hat{\mu}^2} - \frac{1}{2\hat{\nu}^3\hat{\mu}^5} \underset{\sim}{1}' \left[\hat{M}_2 - \hat{M}_1^2 \right] \underset{\sim}{1} \\ - \frac{1}{2\hat{\nu}^3\hat{\mu}^3} \underset{\sim}{1}' \left[I - \hat{M}_1\hat{M}_{-1} \right] \underset{\sim}{1} \quad (4.107)$$

e

$$I_{\hat{\nu}\hat{\nu}} = \frac{n}{2\hat{\nu}^2} - \frac{1}{4\hat{\nu}^4\hat{\mu}^4} \underset{\sim}{1}' \left[\hat{M}_2 - \hat{M}_1^2 \right] \underset{\sim}{1} - \frac{1}{2\hat{\nu}^4\hat{\mu}^2} \underset{\sim}{1}' \left[I - \hat{M}_1\hat{M}_{-1} \right] \underset{\sim}{1} \\ - \frac{1}{4\hat{\nu}^4} \underset{\sim}{1}' \left[\hat{M}_{-2} - \hat{M}_{-1}^2 \right] \underset{\sim}{1}. \quad (4.108)$$

4.7 Testes de Hipóteses e Estimação por Intervalo com Dados Censurados

Nesta seção, consideramos testes de hipóteses aproximados, baseados na estatística da razão de verossimilhança para os parâmetros da $M - IG$ e seus casos particulares (IG e $R - IG$). A justificativa da escolha deste teste foi dado na Seção 4.3. Também, nesta seção não se considera testes de hipóteses para a fronteira do espaço paramétrico de p , já que as condições de regularidade não são satisfeitas e assim não temos os requisitos necessários para o teste baseado na estatística da razão de verossimilhança.

Igualmente ao caso de dados sem censura, os intervalos de confiança para os parâmetros da $M - IG$ e seus casos particulares são baseados nas propriedades assintóticas dos estimadores de máxima verossimilhança. Outros procedimentos para obter intervalos de confiança aproximados, como aqueles baseados na estatística da razão de verossimilhança (corregido ou não, Doganaksoy e Schmee, 1993), são deixados como tema de pesquisa.

4.7.1 Teste de Hipóteses e Estimação por Intervalo Para os Parâmetros da IG ($p = 0$)

Para realizar os testes de hipóteses para os parâmetros da IG, se considera a seguinte função da log-verossimilhança

$$\mathfrak{S}(\mu, \nu) = \sum_{i=1}^r \log [f_{\delta}(y_i, \mu, \nu)] + \sum_{i=r+1}^n \log [S(a_i, \mu, \nu)],$$

onde

$$f_{\delta}(y_i, \mu, \nu) = (2\pi y_i^3 \nu)^{-\frac{1}{2}} \exp \left[-\frac{(y_i - \mu)^2}{2\nu \mu^2 y_i} \right]$$

e

$$S(a_i, \mu, \nu) = \int_{a_i}^{\infty} f_{\delta}(y_i, \mu, \nu) dy_i.$$

i) Teste de hipóteses para μ

Considere as hipóteses $H_0 : \mu = \mu_0$ e $H_1 : \mu \neq \mu_0$. Para testar estas hipóteses usa-se a seguinte estatística

$$\Lambda(\mu_0, \tilde{\nu}) = -2 \{ \mathfrak{S}(\mu_0, \tilde{\nu}) - \mathfrak{S}(\hat{\mu}, \hat{\nu}) \},$$

onde $\tilde{\nu}$ é o estimador de máxima verossimilhança do parâmetro ν dado μ_0 , e $\hat{\mu}$ e $\hat{\nu}$ os estimadores de máxima verossimilhança de μ e ν , respectivamente. A estatística $\Lambda(\mu_0, \tilde{\nu})$ tem distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : \mu = \mu_0$.

A estimativa de ν dado μ_0 pode ser achada mediante o seguinte procedimento:

I. Escolher o estimativa inicial de ν , o qual, pode ser a estimativa de máxima verossimilhança obtidos com os dados não censurados, calculados com as fórmulas dadas em (4.11).

II. Calcular $E(Y_i^{-1} | Y_i > a_i)$, para $i = r + 1, \dots, n$, de acordo com as fórmulas dadas em (1.18).

III. Substituir os valores no lado direito da equação

$$\hat{\nu} = \frac{1}{n} \left[\frac{1}{\mu_0^2} \underset{\sim}{1}' E_*(Y) \underset{\sim}{1} - \frac{2n}{\mu_0} + \underset{\sim}{1}' E_*(Y^{-1}) \underset{\sim}{1} \right]$$

IV. Repetir as etapas II e III até obter a estimativa de ν com a precisão desejada.

ii) **Teste de hipóteses para ν**

Considere as hipóteses $H_0 : \nu = \nu_0$ e $H_1 : \nu \neq \nu_0$. Para testar estas hipóteses, usar a seguinte estatística

$$\Lambda(\tilde{\mu}, \nu_0) = -2 \{ \mathfrak{F}(\tilde{\mu}, \nu_0) - \mathfrak{F}(\hat{\mu}, \hat{\nu}) \}.$$

onde $\tilde{\mu}$ é o estimador de máxima verossimilhança do parâmetro μ dado ν_0 , e $\hat{\mu}$ e $\hat{\nu}$ os estimadores de máxima verossimilhança de μ e ν , respectivamente. A estatística $\Lambda(\tilde{\mu}, \nu)$ segue uma distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : \nu = \nu_0$.

A estimativa de μ dado ν_0 pode ser achada mediante o seguinte procedimento:

I. Escolher a estimativa inicial de μ , a qual pode ser a estimativa de máxima verossimilhança calculados com os dados não censurados, achadas com as fórmulas dadas em (4.10).

II. Calcular $E(Y_i | Y_i > a_i)$, para $i = r + 1, \dots, n$, de acordo com as fórmulas dadas em (1.18).

III. Substituir os valores no lado direito da equação

$$\tilde{\mu} = \frac{1}{n} \underset{\sim}{1}' E_*(Y) \underset{\sim}{1}.$$

IV. Repetir as etapas II e III até obter a estimativa de μ com a precisão desejada.

iii) **Intervalos de confiança para os parâmetros da IG**

Para encontrar os intervalos de confiança aproximados para os parâmetros da IG, usamos a propriedade de normalidade assintótica dos estimadores de máxima

verossimilhança. Assim,

$$\begin{pmatrix} \hat{\mu} \\ \hat{\nu} \end{pmatrix} \underset{a.}{\sim} N \left[\begin{pmatrix} \mu \\ \nu \end{pmatrix}, I_{I_0}^{-1} \right],$$

onde

$$I_{I_0} = \begin{bmatrix} I_{\hat{\mu}\hat{\mu}} & I_{\hat{\mu}\hat{\nu}} \\ I_{\hat{\nu}\hat{\mu}} & I_{\hat{\nu}\hat{\nu}} \end{bmatrix}$$

e as componentes da matriz I_{I_0} são dadas em (4.103), (4.104) e (4.105). Com esta propriedade, os intervalos aproximados para μ e ν , com $(1 - \alpha)\%$ de confiança, são dadas por

$$IC(\mu) : \hat{\mu} \mp Z_0 \sqrt{C_{11}}$$

e

$$IC(\nu) : \hat{\nu} \mp Z_0 \sqrt{C_{22}}$$

respectivamente, onde $Z \sim N(0, 1)$, C_{jj} é o j -ésimo elemento da diagonal da matriz I_{I_0} e $P(Z < Z_0) = 1 - \frac{\alpha}{2}$.

4.7.2 Teste de Hipóteses para os Parâmetros da R-IG ($p = 1$)

Para realizar o teste de hipóteses dos parâmetros da R-IG se considera a seguinte função da log-verossimilhança

$$\mathfrak{S}(\mu, \nu) = \sum_{i=1}^r \log [g(y_i, \mu, \nu)] + \sum_{i=r+1}^n \log [S(a_i, \mu, \nu)]$$

onde

$$g(y_i, \mu, \nu) = (2\pi y_i \mu^2 \nu)^{-\frac{1}{2}} \exp \left[-\frac{(y_i - \mu)^2}{2\nu \mu^2 y_i} \right]$$

e

$$S(a_i, \mu, \nu) = \int_{a_i}^{\infty} g(y_i, \mu, \nu) dy_i.$$

i) Teste de hipóteses para μ

Considere as hipóteses $H_0 : \mu = \mu_0$ e $H_1 : \mu \neq \mu_0$. Para testar estas hipóteses, usa-se a seguinte estatística

$$\Lambda(\mu_0, \tilde{\nu}) = -2 \{ \mathfrak{S}(\mu_0, \tilde{\nu}) - \mathfrak{S}(\hat{\mu}, \hat{\nu}) \},$$

onde $\tilde{\nu}$ é o estimador de máxima verossimilhança do parâmetro ν dado μ_0 , e $\hat{\mu}$ e $\hat{\nu}$ os estimadores de máxima verossimilhança de μ e ν , respectivamente. A estatística $\Lambda(\mu_0, \tilde{\nu})$ tem distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : \mu = \mu_0$.

A estimativa de ν dado μ_0 pode ser achada mediante o seguinte procedimento:

I. Escolher o estimativa inicial de ν , o qual, pode ser a estimativa de máxima verossimilhança obtidas com os dados não censurados, calculados com as fórmulas dadas em (4.12).

II. Calcular $E(Y_i | Y_i > a_i)$, para $i = r + 1, \dots, n$, de acordo com as fórmulas dadas em (2.10).

III. Substituir os valores no lado direito da equação

$$\tilde{\nu} = \frac{1}{n} \left[\frac{1}{\mu_0^2} \underset{\sim}{1}' E_*(Y) \underset{\sim}{1} - \frac{2n}{\mu_0} + \underset{\sim}{1}' E_*(Y^{-1}) \underset{\sim}{1} \right]$$

IV. Repetir as etapas II e III até obter a estimativa de ν com a precisão desejada.

ii) Teste de hipóteses para ν

Considere as hipóteses $H_0 : \nu = \nu_0$ e $H_1 : \nu \neq \nu_0$. Para testar estas hipóteses, usa-se a seguinte estatística

$$\Lambda(\tilde{\mu}, \nu_0) = -2 \{ \mathfrak{S}(\tilde{\mu}, \nu_0) - \mathfrak{S}(\hat{\mu}, \hat{\nu}) \},$$

onde $\tilde{\mu}$ é o estimador de máxima verossimilhança do parâmetro μ dado ν_0 , e $\hat{\mu}$ e $\hat{\nu}$ os estimadores de máxima verossimilhança de μ e ν , respectivamente. A estatística $\Lambda(\tilde{\mu}, \nu_0)$ tem distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : \nu = \nu_0$.

A estimativa de μ dado ν_0 pode ser achada mediante o seguinte procedimento:

I. Escolher a estimativa inicial de μ a qual pode ser a estimativa de máxima verossimilhança calculada com os dados não censurados, achada com a fórmula dada em (4.12).

II. Calcular $E(Y_i^{-1} | Y_i > a_i)$, para $i = r + 1, \dots, n$, de acordo com as fórmulas dadas em (2.10).

III. Substituir os valores no lado direito da equação

$$\tilde{\mu} = \frac{n + \sqrt{1 + \frac{4\nu_0}{n} \mathbf{1}' E_*(Y) \mathbf{1}}}{2\nu_0}$$

IV. Repetir as etapas II e III até obter a estimativa de μ com a precisão desejada.

iii) Intervalos de confiança para os parâmetros da R-IG

Para encontrar os intervalos de confiança aproximados para os parâmetros da R-IG, usamos a propriedade de normalidade assintótica dos estimadores de máxima verossimilhança. Assim,

$$\begin{pmatrix} \hat{\mu} \\ \hat{\nu} \end{pmatrix} \underset{\sim}{\sim} N \left[\begin{pmatrix} \mu \\ \nu \end{pmatrix}, I_{R0}^{-1} \right],$$

onde

$$I_{R0} = \begin{bmatrix} I_{\hat{\mu}\hat{\mu}} & I_{\hat{\mu}\hat{\nu}} \\ I_{\hat{\nu}\hat{\mu}} & I_{\hat{\nu}\hat{\nu}} \end{bmatrix}$$

e os componentes da matriz I_{R0} são dadas em (4.106), (4.107) e (4.108). Com esta propriedade, os intervalos aproximados para μ e ν , com $(1 - \alpha)\%$ de confiança são dadas por

$$IC(\mu) : \hat{\mu} \mp Z_0 \sqrt{C_{11}}$$

e

$$IC(\nu) : \hat{\nu} \mp Z_0 \sqrt{C_{22}},$$

onde $Z \sim N(0, 1)$, C_{jj} é o j -ésimo elemento da diagonal da matriz I_{R0} e $P(Z < Z_0) = 1 - \frac{\alpha}{2}$.

4.7.3 Teste de Hipóteses e Estimação por Intervalos sobre os Parâmetros da M-IG

Para realizar os testes de hipóteses dos parâmetros da R-IG, considera-se a seguinte função da log-verossimilhança

$$\mathfrak{S}(\mu, \nu, p) = \sum_{i=1}^r \log [g(y_i; \mu, \nu, p)] + \sum_{i=r+1}^n \log [S(a_i; \mu, \nu, p)],$$

onde

$$g(y_i; \mu, \nu, p) = [1 - p + p\mu^{-1}y_i] (2\pi y_i^3 \nu)^{-\frac{1}{2}} \exp \left[-\frac{(y_i - \mu)^2}{2\nu\mu^2 y_i} \right]$$

e

$$S(a_i, \mu, \nu) = \int_{a_i}^{\infty} g(y_i; \mu, \nu, p) dy_i.$$

i) Teste de hipóteses para μ

Considere as hipóteses $H_0 : \mu = \mu_0$ e $H_1 : \mu \neq \mu_0$. Para testar estas hipóteses usar a seguinte estatística

$$\Lambda(\mu_0, \tilde{\nu}, \tilde{p}) = -2 \{ \mathfrak{S}(\mu_0, \tilde{\nu}, \tilde{p}) - \mathfrak{S}(\hat{\mu}, \hat{\nu}, \hat{p}) \}$$

onde $\tilde{\nu}$ e \tilde{p} são os estimadores de máxima verossimilhança dos parâmetros ν e p dado μ_0 , e $\hat{\mu}$, $\hat{\nu}$ e \hat{p} os estimadores de máxima verossimilhança de μ , ν e p , respectivamente. A estatística $\Lambda(\mu_0, \tilde{\nu}, \tilde{p})$ tem distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : \mu = \mu_0$.

As estimativas de máxima verossimilhança de ν e p dado μ_0 , são encontradas resolvendo o seguinte sistema de equações

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\frac{1}{\mu_0^2} \underset{\sim}{1}' E_*(Y) \underset{\sim}{1} - \frac{2n}{\mu_0} + \underset{\sim}{1}' E_*(Y^{-1}) \underset{\sim}{1} \right] = 0$$

e

$$\frac{\partial \mathfrak{S}(\theta)}{\partial p} = \frac{1}{\mu_0} \underset{\sim}{1}' E_*(Y D_p^{-1}) \underset{\sim}{1} - \underset{\sim}{1}' E_*(D_p^{-1}) \underset{\sim}{1} = 0.$$

Para resolver este sistema, pode-se empregar o método de Newton Raphson com a seguinte matriz de informação observada, que é obtida usando o Teorema 4.3

$$I_{\mu_0} = \begin{bmatrix} I_{\tilde{\nu}\tilde{\nu}} & I_{\tilde{\nu}\tilde{p}} \\ I_{\tilde{p}\tilde{\nu}} & I_{\tilde{p}\tilde{p}}, \end{bmatrix},$$

onde

$$I_{\tilde{\nu}\tilde{\nu}} = \frac{n}{2\nu^2} - \frac{1}{4\nu^4 \mu_0^4} \underset{\sim}{1}' [\tilde{M}_2 - \tilde{M}_1^2] \underset{\sim}{1} - \frac{1}{2\nu^4 \mu_0^2} \underset{\sim}{1}' [I - \tilde{M}_1 \tilde{M}_{-1}] \underset{\sim}{1} - \frac{1}{4\nu^4} \underset{\sim}{1}' [\tilde{M}_{-2} - \tilde{M}_{-1}^2] \underset{\sim}{1},$$

$$I_{\tilde{\nu}\tilde{p}} = I_{\tilde{p}\tilde{\nu}} = -\frac{1}{2\nu^2 \mu_0^2} \underset{\sim}{1}' [\tilde{N}_2 - \tilde{M}_1 \tilde{N}_1] \underset{\sim}{1} - \frac{1}{2\nu^2 \mu_0} \underset{\sim}{1}' [\tilde{N}_0 - \tilde{M}_{-1} \tilde{N}_1] \underset{\sim}{1} + \frac{1}{2\nu^2 \mu_0^2} \underset{\sim}{1}' [\tilde{N}_1 - \tilde{M}_1 \tilde{N}_0] \underset{\sim}{1} + \frac{1}{2\nu^2} \underset{\sim}{1}' [\tilde{N}_{-1} - \tilde{M}_{-1} \tilde{N}_0] \underset{\sim}{1}$$

e

$$I_{\tilde{p}\tilde{p}} = \frac{1}{\mu_0^2} \underset{\sim}{1}' \tilde{N}_1^2 \underset{\sim}{1} - \frac{2}{\mu_0} \underset{\sim}{1}' \tilde{N}_1 \tilde{N}_0 \underset{\sim}{1} + \underset{\sim}{1}' \tilde{N}_0^2 \underset{\sim}{1}.$$

onde os componentes $E[Y_i^s (1 - p_0 + p_0 \mu_0^{-1} Y_i) | Y_i > a_i]$ e $E(Y_i^s | Y_i > a_i)$ de \tilde{N}_s e \tilde{M}_s são calculados com as estimativas de ν e p dado μ_0 , mediante as expressões dadas nas relações (3.23) e (3.24), respectivamente, para todo $s \in Z$.

ii) Teste de hipóteses para ν

Considere as hipóteses $H_0 : \nu = \nu_0$ e $H_1 : \nu \neq \nu_0$. Para testar estas hipóteses, usar a seguinte estatística

$$\Lambda(\tilde{\mu}, \nu_0, \tilde{p}) = -2 \{ \mathfrak{S}(\tilde{\mu}, \nu_0, \tilde{p}) - \mathfrak{S}(\hat{\mu}, \hat{\nu}, \hat{p}) \},$$

onde $\tilde{\mu}$ e \tilde{p} são os estimadores de máxima verossimilhança do parâmetro μ e p dado ν_0 respectivamente. Os valores $\hat{\mu}$, $\hat{\nu}$ e \hat{p} são os estimadores de máxima verossimilhança de μ , ν e p , respectivamente. A estatística $\Lambda(\tilde{\mu}, \nu_0, \tilde{p})$ tem distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : \nu = \nu_0$.

As estimativas de máxima verossimilhança de μ e p , dado ν_0 , são encontrados resolvendo o seguinte sistema de equações

$$\begin{aligned} \frac{\partial \mathfrak{S}(\theta)}{\partial \mu} &= \frac{1-p}{\mu} \underset{\sim}{1}' E_* (D_p^{-1}) \underset{\sim}{1} + \frac{1}{\nu_0 \mu^3} \underset{\sim}{1}' E_* (Y) \underset{\sim}{1} - \frac{n}{\nu_0 \mu^2} - \frac{n}{\mu} = 0, \\ \frac{\partial \mathfrak{S}(\theta)}{\partial p} &= \frac{1}{\mu} \underset{\sim}{1}' E_* (Y D_p^{-1}) \underset{\sim}{1} - \underset{\sim}{1}' E_* (D_p^{-1}) \underset{\sim}{1} = 0. \end{aligned}$$

Este sistema pode ser resolvido mediante o método de Newton Raphson com a seguinte matriz de informação observada, que é obtida usando o Teorema 4.3.

$$I_{\nu_0} = \begin{bmatrix} I_{\tilde{\mu}\tilde{\mu}} & I_{\tilde{\mu}\tilde{p}} \\ I_{\tilde{p}\tilde{\mu}} & I_{\tilde{p}\tilde{p}} \end{bmatrix},$$

onde

$$\begin{aligned} I_{\tilde{\mu}\tilde{\mu}} &= - \left(\frac{2n}{\nu_0 \mu^2} + \frac{n}{\mu^3} \right) + \frac{1}{\nu_0^2 \mu^5} \underset{\sim}{1}' \left[3\nu_0 \tilde{\mu}^2 \tilde{M}_1 - \tilde{M}_2 + \tilde{M}_1^2 \right] \underset{\sim}{1} \\ &\quad - \frac{2(1-\tilde{p})}{\nu_0 \mu^4} \underset{\sim}{1}' \left[\tilde{N}_1 - \tilde{N}_0 \tilde{M}_1 - \frac{(1-\tilde{p})\nu_0 \tilde{\mu}^2}{2} \tilde{N}_0^2 \right] \underset{\sim}{1}, \\ I_{\tilde{\mu}\tilde{p}} &= \frac{1}{\mu} \underset{\sim}{1}' \left[\tilde{N}_0 + \frac{(1-\tilde{p})}{\mu} \tilde{N}_0 \tilde{N}_1 - (1-\tilde{p}) \tilde{N}_0^2 \right] \underset{\sim}{1} \\ &\quad - \frac{1}{\nu_0 \mu^4} \underset{\sim}{1}' \left[\tilde{N}_2 - \tilde{M}_1 \tilde{N}_1 \right] \underset{\sim}{1} + \frac{1}{\nu_0 \mu^3} \underset{\sim}{1}' \left[\tilde{N}_1 - \tilde{M}_1 \tilde{N}_0 \right] \underset{\sim}{1} \end{aligned}$$

e

$$I_{\tilde{p}\tilde{p}} = \frac{1}{\tilde{\mu}^2} \tilde{1}' \tilde{N}_1^2 \tilde{1} - \frac{2}{\tilde{\mu}} \tilde{1}' \tilde{N}_1 \tilde{N}_0 \tilde{1} + \tilde{1}' \tilde{N}_0^2 \tilde{1},$$

onde os componentes $E[Y_i^s(1-p+p\mu^{-1}Y_i) | Y_i > a_i]$ e $E(Y_i^s | Y_i > a_i)$ de \tilde{N}_s e \tilde{M}_s são calculados com as estimativas de μ e p dado ν_0 , mediante as expressões dada na relações (3.23) e (3.24), respectivamente, para todo $s \in Z$.

iii) Teste de hipóteses para p

Considere as hipóteses $H_0 : p = p_0$ e $H_1 : p \neq p_0$. Para testar esta hipóteses, usar a seguinte estatística

$$\Lambda(\tilde{\mu}, \tilde{\nu}, p_0) = -2 \{ \mathfrak{S}(\tilde{\mu}, \tilde{\nu}, p_0) - \mathfrak{S}(\hat{\mu}, \hat{\nu}, \hat{p}) \},$$

onde $\tilde{\mu}$ e $\tilde{\nu}$ são os estimadores de máxima verossimilhança dos parâmetros μ e ν dado p_0 , respectivamente. Os valores $\hat{\mu}$, $\hat{\nu}$ e \hat{p} são os estimadores de máxima verossimilhança de μ , ν e p , respectivamente. A estatística $\Lambda(\tilde{\mu}, \tilde{\nu}, p_0)$ tem distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : p = p_0$.

iv) Intervalo de confiança

Para encontrar os intervalos de confiança aproximados, usamos a propriedade assintótica dos estimadores de máxima verossimilhança. Assim,

$$\begin{pmatrix} \hat{\mu} \\ \hat{\nu} \\ \hat{p} \end{pmatrix} \underset{\sim}{a.} N \left[\begin{pmatrix} \mu \\ \nu \\ p \end{pmatrix}, I_0^{-1} \right],$$

onde I_0 é a matriz de informação observada dado em (4.69).

Com estas propriedades, os limites aproximados para μ , ν e p com $(1 - \alpha)\%$ de confiança são dados por

$$LC(\mu) : \hat{\mu} \mp Z_0 \sqrt{C_{11}},$$

$$LC(\nu) : \hat{\nu} \mp Z_0 \sqrt{C_{22}}$$

e

$$LC(p) : \hat{p} \mp Z_0 \sqrt{C_{33}},$$

onde $Z \sim N(0, 1)$, C_{jj} é j -ésimo elemento da diagonal de I_0^{-1} e $P(Z \leq Z_0) = 1 - \frac{\alpha}{2}$.

4.8 Exemplo 4.2. Tempo de remissão de pacientes com Leucemia

Lawless (1982), num exercício proposto na pag. 136, apresenta dados de tempos de remissão de 30 pacientes com leucemia, tratadas sob certa terapia. Este conjunto de dados se mostra na Tabela 4.3.

Tabela 4.3: Tempo de remissão (em semanas) de pacientes com leucemia, (+) observações censuradas.

1	1	2	4	4	6
6	6	7	8	9	9
10	12	13	14	18	19
24	26	29	31+	42	45+
50+	57	60	71+	85+	91

Para ilustrar a aplicação para dados censurados desta distribuição, vamos considerar inicialmente que os dados provém de M-IG, para depois determinar a qualidade de ajuste desta distribuição. Para obter as estimativas de máxima verossimilhança, segue-se o procedimento proposto na Seção 4.4. Este procedimento está programado em S-plus, programas que estão no Apêndice em forma geral (dados censurados ou não, com covariáveis ou sem elas). Assim, o programa 2 produz a Tabela 4.4, exceto a linha dada para $p = 0.824$, que foi achada com o programa 3 do Apêndice, considerando como valor inicial no método Newton Raphson, os valores das estimativas

associadas ao máximo valor da log-verossimilhança no programa de p fixo. Logo, o

Tabela 4.4: Estimativas de máxima verossimilhança para dados de tempos de remissão de pacientes com leucemia, para diferentes valores de p .

μ	ν	p	log-verossim
46.0561	0.1265	0.000	-118.7378
30.0763	0.1285	0.100	-118.6354
23.3390	0.1340	0.200	-118.4020
19.5097	0.1423	0.300	-118.1316
16.9063	0.1534	0.400	-117.8737
14.9132	0.1681	0.500	-117.6519
13.2479	0.1878	0.600	-117.4777
11.7508	0.2157	0.700	-117.3587
10.3017	0.2589	0.800	-117.3038
9.9455	0.2731	0.824	-117.3015
8.7482	0.3376	0.900	-117.3245
6.7276	0.5536	1.000	-117.3860

máximo valor da log-verossimilhança é atingido quando o valor de $\hat{p} = 0,824$. Neste caso, vamos avaliar a qualidade de ajuste por métodos gráficos. Assim, nas Figuras 4.3, 4.4 e 4.5, pode-se observar que a curva de sobrevivência estimada, assumindo que os dados se ajustam a uma distribuição Gaussiana Inversa, a M-IG e à Recíproca Complementar, respectivamente, sendo cada uma comparada com a curva de sobrevivência estimada pelo método de Kaplan Meier.

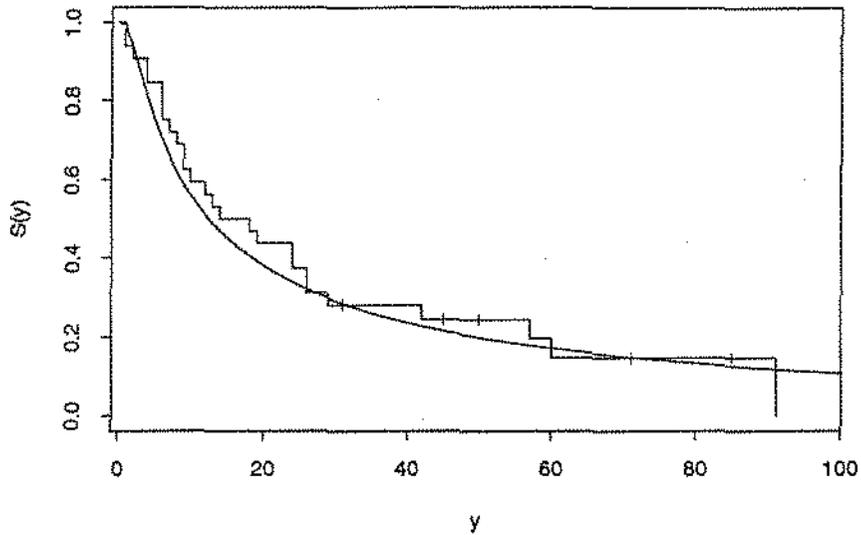


Figura 4.3: Gráfico da função de sobrevivência estimada por Kaplan Meier e de máxima verossimilhança sob IG., para dados de tempos de remissão de pacientes com leucemia.

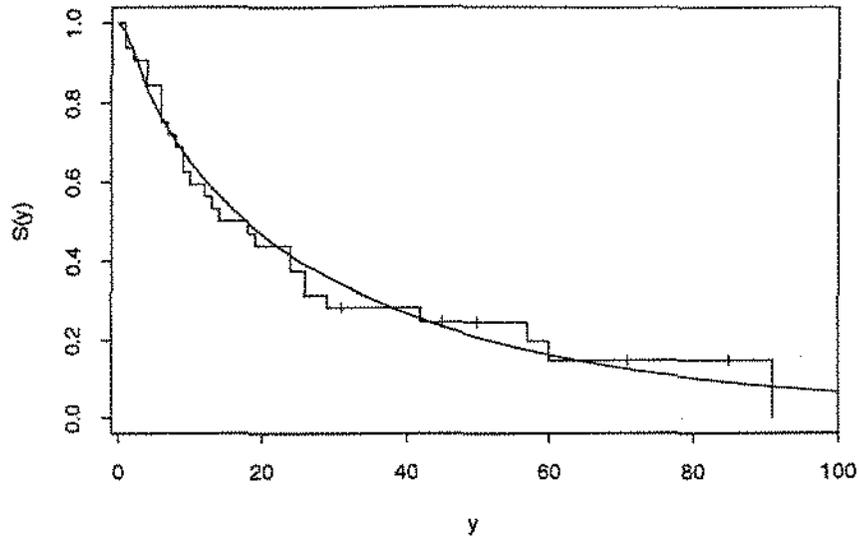


Figura 4.4: Gráfico da função de sobrevivência estimada por Kaplan Meier e de máxima verossimilhança sob M-IG com $p = 0.824$, para dados de tempos de remissão de pacientes com leucemia.

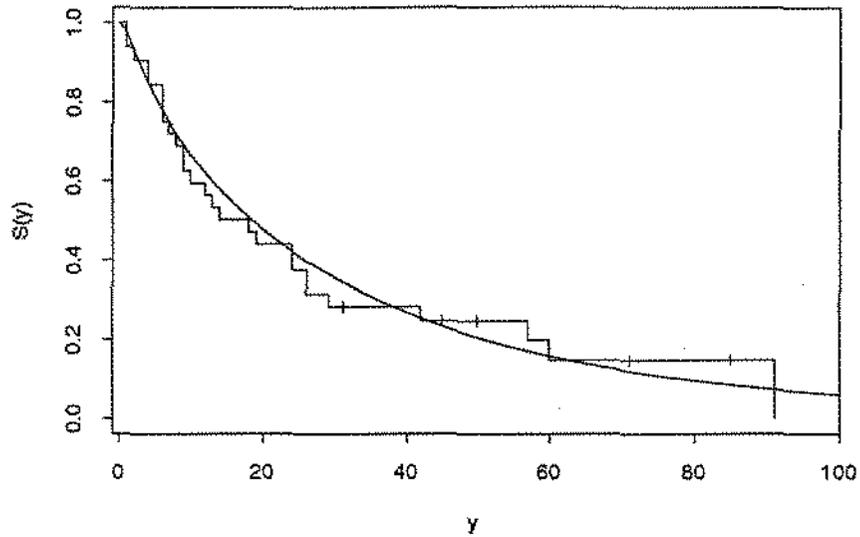


Figura 4.5: Gráfico da função de sobrevivência estimada por Kaplan Meier e de máxima verossimilhança sob R-IG., para dados de tempos de remissão de pacientes com leucemia.

Destes gráficos pode-se observar que curva de sobrevivência estimada comparado com o método de Kaplan Meier, para $\hat{p} = 0.824$ e $p = 1$ (Recíproca Complementar) obtem-se ajuste razoável.

5. Um Modelo de Regressão para Dados da Mistura de Gaussiana Inversa com sua Recíproca Complementar

5.1 Introdução

O modelo de regressão considerado neste capítulo, para dados com e sem censura da Mistura da Gaussiana Inversa com sua Recíproca Complementar, é um modelo onde assume que *a recíproca do parâmetro μ depende linearmente das covariáveis*. Para este modelo, deduz-se o sistema de equações de verossimilhança e a matriz de informação observada, que pode ser usada no método de Newton Raphson para obter as estimativas de máxima verossimilhança dos parâmetros desconhecidos. O modelo de regressão que será estudada, inclui como caso particular o modelo de regressão dado por Whitmore (1983), para dados da distribuição Gaussiana Inversa.

5.2 Modelo de Regressão

De acordo com Lawless (1982), os modelos de regressão mais usados para analisar dados de tempos de falha influenciados por covariáveis são

- modelos de riscos proporcionais, e
- modelos de locação-escala do logarítmo do tempo.

Um modelo de regressão que pertence à família de riscos proporcionais tem a propriedade de que para dois indivíduos com vetores de regressão $\underline{x} = (x_{11}, \dots, x_{1k})$ e $\underline{x}_2 = (x_{21}, \dots, x_{2k})$, respectivamente, a razão de funções de risco de $r(y|\underline{x}_1) / r(y|\underline{x}_2)$ não muda com o tempo y . Isto implica que a função de risco do tempo Y dado $\underline{x} = (x_1, \dots, x_k)$, pode ser escrita

$$r(y|\underline{x}) = r_0(y) g(\underline{x}, \underline{\beta}),$$

onde

$r_0(y)$ é a função risco básica,

$g(\underline{x}, \underline{\beta})$ é uma função que depende de \underline{x} e $\underline{\beta}$ e $\underline{\beta} = (\beta_1, \dots, \beta_k)'$ é um vetor dos coeficientes de regressão.

Um caso particular deste modelo ocorre quando

$$g(\underline{x}, \underline{\beta}) = \exp\{\underline{x}\underline{\beta}\},$$

isto é,

$$r(y|\underline{x}) = r_0(y) \exp\{\underline{x}\underline{\beta}\}.$$

Este modelo é adequado desde que $\exp\{\underline{x}\underline{\beta}\}$ é sempre positivo.

Um modelo é de locação-escala, quando o logaritmo do tempo de falha, $Y = \log T$, dado \underline{x} , onde $T = t$ Tempo, tem uma distribuição com parâmetro de locação, $\mu(\underline{x})$, e parâmetro de escala constante σ , ($\sigma > 0$), constante. Desta forma, o modelo de regressão pode ser escrito como

$$Y = \mu(\underline{x}) + \sigma\epsilon,$$

onde, ϵ tem uma distribuição que é independente de \underline{x} . Particularmente, quando

$$\mu(\underline{x}) = \underline{x}\underline{\beta},$$

tem-se o seguinte modelo

$$Y = \underset{\sim}{x}\underset{\sim}{\beta} + \sigma\epsilon.$$

Para o caso de dados da $M - IG$, considera-se um modelo diferente de riscos proporcionais ou de locação escala, assumindo-se que a recíproca do parâmetro de locação depende linearmente das covariáveis, isto é,

$$\mu \left(\underset{\sim}{x} \right) = \frac{1}{\underset{\sim}{x}\underset{\sim}{\beta}}.$$

O motivo principal pelo qual não consideramos um modelo de riscos proporcionais baseia-se na característica da função de risco ou taxa de falha da $M - IG$, descrita no Capítulo 3. Esta função é crescente até um ponto localizado entre a moda (y_m) e y_0 ($y_0 > y_m$), depois, decresce aproximando-se assintoticamente a $2\mu^2\nu$. Logo, a função risco, para valores grandes de y , é aproximadamente constante, o que dificulta o cumprimento da condição fundamental da família de riscos propocionais.

A razão pela qual não consideramos um modelo de locação-escala pode ser amparada na justificativa dada por Whitmore (1983), para o modelo de regressão proposto para dados da distribuição Gaussiana Inversa, sendo esta distribuição um caso particular da $M - IG$. A justificativa dada por Whitmore está baseada numa propriedade da distribuição Gaussiana Inversa, demonstrada por Whitmore e Yalovsky (1978). Eles demonstraram que se $Y \sim IG(\mu, \nu)$ e $(\mu\nu)^{-1}$ suficientemente grande (ao menos 10), $\log Y$ é aproximadamente normal. Logo se a função de regressão tem a forma $E(Y) = \exp \left\{ \underset{\sim}{x}\underset{\sim}{\beta} \right\}$ e $\gamma = \nu\mu$ é uma constante razoavelmente grande, então, $\log Y$ possui um modelo aproximadamente normal linear, $N \left(\underset{\sim}{x}\underset{\sim}{\beta} - \frac{1}{2}\gamma, \gamma \right)$. Com base nesta propriedade, Whitmore (1983) cita três razões pelas quais não considerou um modelo de locação-escala para a regressão de dados da distribuição Gaussiana In-

versa. Estas razões apresentam a seguir: (1) O modelo de regressão considerado para dados da IG é $E(Y) = \mu(\tilde{x}) = (\tilde{x}\tilde{\beta})^{-1}$ (modelo de regressão linear inverso), que é diferente de $E(Y) = \exp\{\tilde{x}\tilde{\beta}\}$. (2) Assume-se que o parâmetro de dispersão ν é uma constante, não ocorrendo $\gamma = (\mu\nu)^{-1}$ (quando é assumido o modelo $E(Y) = \exp\{\tilde{X}\tilde{\beta}\}$). (3) O modelo de regressão considerado para IG é exato e não aproximado, em contraste ao modelo de locação-escala para a IG . Portanto, se consideramos um modelo de locação-escala para dados da $M-IG$, este modelo serão aproximado e não exato como deveria ser.

5.3 Estimação de Parâmetros para Dados sem Censura

Para realizar a estimação dos parâmetros para dados não censurados, considere-se o seguinte. Seja Y_1, Y_2, \dots, Y_n variáveis aleatórias independentes, tais que $Y_i \sim M-IG(\mu_i, \nu, p)$, e com função de densidade dada em (3.4), onde $\mu_i^{-1} = \tilde{x}_i\tilde{\beta}$, $\tilde{\beta} = (\beta_1, \dots, \beta_k)'$ é o vetor dos parâmetros da regressão e $\tilde{x}_i = (x_{i1}, \dots, x_{ik})$ é o vetor de covariáveis, para $i = 1, \dots, n$. Além disso, sejam:

$$X = \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix}_{n \times k} \quad \mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{n \times 1} \quad Y = \begin{pmatrix} y_1 & 0 & \cdots & 0 \\ 0 & y_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & y_n \end{pmatrix}_{n \times n}. \quad (5.1)$$

Então, sabendo-se que $\tilde{x}_i\tilde{\beta} = \mu_i^{-1}$ a função de verossimilhança para a observação i , por (3.4), é dada por

$$g(y_i, \tilde{\beta}, \nu, p) = \left(1 - p + p \tilde{x}_i\tilde{\beta} y_i\right) (2\pi\nu y_i^3)^{-\frac{1}{2}} \exp\left\{-\frac{\left(1 - \tilde{x}_i\tilde{\beta} y_i\right)^2}{2\nu y_i}\right\}. \quad (5.2)$$

Isto é, pela reparametrização dada em (1.5). Então, a função de verossimilhança para a amostra pode ser escrita

$$L(\theta) = \prod_{i=1}^n g(y_i, \beta, \nu, p) \\ = \prod_{i=1}^n \left(1 - p + p \frac{x_i \beta}{y_i}\right) \prod_{i=1}^n (2\pi \nu y_i^3)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\nu} \sum_{i=1}^n \frac{(1 - \frac{x_i \beta}{y_i})^2}{y_i} \right\},$$

onde $\theta = (\beta', \nu, p)'$.

Temos que

$$\sum_{i=1}^n \frac{(1 - \frac{x_i \beta}{y_i})^2}{y_i} = \mathbf{1}' Y^{-1} \mathbf{1} - 2 \mathbf{1}' X \beta + \beta' X' Y X \beta.$$

Então, a função de verossimilhança pode ser escrita como

$$L(\theta) = \prod_{i=1}^n \left(1 - p + p \frac{x_i \beta}{y_i}\right) \prod_{i=1}^n (2\pi \nu y_i^3)^{-\frac{1}{2}} \exp \left\{ -\frac{\mathbf{1}' Y^{-1} \mathbf{1} - 2 \mathbf{1}' X \beta + \beta' X' Y X \beta}{2\nu} \right\}.$$

A log-verossimilhança é

$$\mathfrak{S}(\theta) = \log L(\theta) \\ = \sum_{i=1}^n \log \left(1 - p + p \frac{x_i \beta}{y_i}\right) - \frac{1}{2} \sum_{i=1}^n \log (2\pi y_i^3) - \frac{n}{2} \log \nu \\ - \frac{1}{2\nu} \left[\mathbf{1}' Y^{-1} \mathbf{1} - 2 \mathbf{1}' X \beta + \beta' X' Y X \beta \right]. \quad (5.3)$$

Derivando em relação aos parâmetros e igualando a zero, obtém-se o seguinte sistema de equações:

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \beta} = p \sum_{i=1}^n \frac{\frac{x_i'}{y_i}}{1 - p + p \frac{x_i \beta}{y_i}} - \frac{1}{\nu} \left(X' Y X \beta - X' \mathbf{1} \right) = 0, \quad (5.4)$$

$$\frac{\partial \mathfrak{Z}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' X \underset{\sim}{\beta} + \underset{\sim}{\beta}' X' Y X \underset{\sim}{\beta} \right] = 0, \quad (5.5)$$

e

$$\frac{\partial \mathfrak{Z}(\theta)}{\partial p} = \sum_{i=1}^n \frac{x_i \underset{\sim}{\beta} y_i}{1 - p + p x_i \underset{\sim}{\beta} y_i} - \sum_{i=1}^n \frac{1}{1 - p + p x_i \underset{\sim}{\beta} y_i} = 0. \quad (5.6)$$

Seja D_p uma matriz diagonal $n \times n$, tal que

$$D_p = \text{diag} \left\{ 1 - p + p x_1 \underset{\sim}{\beta} y_1, \dots, 1 - p + p x_n \underset{\sim}{\beta} y_n \right\}. \quad (5.7)$$

Pode demonstrar-se que as seguintes relações são satisfeitas:

$$\sum_{i=1}^n \frac{x_i' y_i}{1 - p + p x_i \underset{\sim}{\beta} y_i} = X' Y D_p^{-1} \underset{\sim}{1}, \quad (5.8)$$

$$\sum_{i=1}^n \frac{x_i \underset{\sim}{\beta} y_i}{1 - p + p x_i \underset{\sim}{\beta} y_i} = \underset{\sim}{\beta}' X' Y D_p^{-1} \underset{\sim}{1}, \quad (5.9)$$

e

$$\sum_{i=1}^n \frac{1}{1 - p + p x_i \underset{\sim}{\beta} y_i} = \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1}. \quad (5.10)$$

Com estas expressões equivalentes, o sistema de equações dado de (5.4) a (5.6), transforma-se em

$$\frac{\partial \mathfrak{Z}(\theta)}{\partial \underset{\sim}{\beta}} = p X' Y D_p^{-1} \underset{\sim}{1} - \frac{1}{\nu} \left(X' Y X \underset{\sim}{\beta} - X' \underset{\sim}{1} \right) = \underset{\sim}{0}, \quad (5.11)$$

$$\frac{\partial \mathfrak{Z}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' X \underset{\sim}{\beta} + \underset{\sim}{\beta}' X' Y X \underset{\sim}{\beta} \right] = 0, \quad (5.12)$$

e

$$\frac{\partial \mathfrak{Z}(\theta)}{\partial p} = \underset{\sim}{\beta}' X' Y D_p^{-1} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} = 0. \quad (5.13)$$

Se p é conhecido, o sistema de equações se reduz a

$$\frac{\partial \mathfrak{Z}(\theta)}{\partial \underset{\sim}{\beta}} = p X' Y D_p^{-1} \underset{\sim}{1} - \frac{1}{\nu} \left(X' Y X \underset{\sim}{\beta} - X' \underset{\sim}{1} \right) = \underset{\sim}{0}, \quad (5.14)$$

e

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' X \underset{\sim}{\beta} + \underset{\sim}{\beta}' X' Y X \underset{\sim}{\beta} \right] = 0. \quad (5.15)$$

 $p = 0$

Se $p = 0$, o sistema de equações de verossimilhança para a regressão Gaussiana Inversa é dado por

$$-\frac{1}{\nu} \left(X' Y X \underset{\sim}{\beta} - X' \underset{\sim}{1} \right) = \underset{\sim}{0}, \quad (5.16)$$

e

$$-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' X \underset{\sim}{\beta} + \underset{\sim}{\beta}' X' Y X \underset{\sim}{\beta} \right] = 0. \quad (5.17)$$

Resolvendo este sistema, tem-se

$$\underset{\sim}{\hat{\beta}} = (X' Y X)^{-1} X' \underset{\sim}{1}, \quad (5.18)$$

e substituindo (5.16) em (5.17), resulta

$$\hat{\nu} = \frac{1}{n} \left[\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - \underset{\sim}{1}' X \underset{\sim}{\hat{\beta}} \right]. \quad (5.19)$$

 $p = 1$

No caso de $p = 1$, de (5.14) e (5.15), tem-se o sistema de equações de verossimilhança para a regressão de dados da Recíproca Complementar de uma variável Gaussiana Inversa

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \underset{\sim}{\beta}} = X' D^{-1} \underset{\sim}{1} - \frac{1}{\nu} \left(X' Y X \underset{\sim}{\beta} - X' \underset{\sim}{1} \right) = \underset{\sim}{0}, \quad (5.20)$$

e

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' X \underset{\sim}{\beta} + \underset{\sim}{\beta}' X' Y X \underset{\sim}{\beta} \right] = 0, \quad (5.21)$$

onde

$$D = \text{diag} \left\{ x_{\sim 1} \beta, \dots, x_{\sim n} \beta \right\}.$$

Multiplicando (5.20) com β à esquerda tem-se os estimadores de máxima verossimilhança

$$\hat{\beta}_{\sim} = \left(X' \underset{\sim}{1} \underset{\sim}{1}' X \right)^{-1} X' \underset{\sim}{1} \underset{\sim}{1}' Y^{-1} \underset{\sim}{1}, \quad (5.22)$$

e

$$\hat{\nu} = \frac{1}{n} \left(\hat{\beta}_{\sim}' X' Y X \hat{\beta}_{\sim} - \hat{\beta}_{\sim}' X' \underset{\sim}{1} \right). \quad (5.23)$$

O estimador do parâmetro β , tem a forma do estimador de Mínimo quadrados ponderados, da regressão $Y^{-1} \underset{\sim}{1}$ sobre X , com a diferença de que a matriz de ponderação $W = \underset{\sim}{1} \underset{\sim}{1}'$ não é positiva definida.

O procedimento para encontrar as estimativas de máxima verossimilhança, para $p \in (0, 1)$ e conhecido se constitui por uma modificação do procedimento usado para resolver o sistema de equações para o caso geral, dado em (5.11) a (5.13).

O sistema de equações dado desde (5.11) a (5.13) tem múltiplas soluções. Mas, como para cada valor de p existe uma solução, então, para encontrar o máximo global, ou o maior dos máximos locais, pode-se construir um procedimento semelhante ao do caso de uma amostra sem covariáveis, baseado no método de Newton Raphson. Este procedimento é descrito a seguir:

I. Encontrar as estimativas de máxima verossimilhança de β e ν para $p = 0$. Usar estas estimativas como pontos iniciais para obter as estimativas de β e ν quando $p = 0,1$. Usar o método de Newton Raphson com a precisão desejada.

II. Aplicar o método de Newton Raphson para p , incrementado de 0,1, utilizando sempre como valores iniciais, as estimativas de β e ν obtidos sem o incremento.

III. Repetir a etapa anterior até chegar a $p = 1$.

IV. Para cada valor de p , encontrar o valor de log-verossimilhança com a fórmula dada em (5.3).

V. Para obter o máximo global, utilizar como ponto inicial no método de Newton Raphson as estimativas de $\hat{\beta}$, ν e p para as quais é atingido o maior valor de log-verossimilhança, na etapa III. Aplicar o método Newton Raphson até obter as estimativas com a precisão desejada.

Para encontrar as estimativas de máxima verossimilhança para $p \in (0, 1)$ e conhecido, seguem-se as etapa I e II do procedimento anterior. Assim, repete-se a etapa II até atingir o valor de p .

Pode acontecer que o valor máximo da log-verossimilhança ocorra na fronteira do espaço paramétrico de p , $p = 0$ ou $p = 1$, ou fora deste espaço. Quando isto ocorre, é preciso fazer análises adicionais para determinar se o modelo na fronteira do espaço paramétrico de p se ajusta bem ao conjunto de dados. Uns destes procedimentos é conferir a adequação do modelo mediante a análise de resíduos, que é descrito na Seção 5.5.

Para usar o método de Newton Raphson, necessita-se da matriz de informação esperada ou de sua estimativa. Neste trabalho, utilizamos a matriz de informação observada, a qual é derivada no seguinte teorema.

Teorema 5.1

A matriz de informação observada do modelo de regressão para dados sem cen-

sura de M-IG é

$$I_0 = \begin{bmatrix} I_{\hat{\beta}\hat{\beta}} & I_{\hat{\beta}\hat{\nu}} & I_{\hat{\beta}\hat{p}} \\ I_{\hat{\nu}\hat{\beta}} & I_{\hat{\nu}\hat{\nu}} & I_{\hat{\nu}\hat{p}} \\ I_{\hat{p}\hat{\beta}} & I_{\hat{p}\hat{\nu}} & I_{\hat{p}\hat{p}} \end{bmatrix}, \quad (5.24)$$

onde

$$\begin{aligned} I_{\hat{\beta}\hat{\beta}} &= \hat{p}^2 X'Y^2 \hat{D}_p^{-2} X + \frac{1}{\hat{\nu}} X'YX, \\ I_{\hat{\beta}\hat{\nu}} &= I'_{\hat{\nu}\hat{\beta}} = -\frac{1}{\hat{\nu}^2} \left(X'YX \hat{\beta} - X' \mathbf{1} \right), \\ I_{\hat{\beta}\hat{p}} &= I'_{\hat{p}\hat{\beta}} = -X'Y \hat{D}_p^{-1} \mathbf{1} + \hat{p} X'Y^2 \hat{D}_p^{-2} X \hat{\beta} - \hat{p} X'Y \hat{D}_p^{-2} \mathbf{1}, \\ I_{\hat{\nu}\hat{\nu}} &= \frac{n}{2\hat{\nu}^2}, \\ I_{\hat{\nu}\hat{p}} &= I_{\hat{p}\hat{\nu}} = 0, \end{aligned}$$

e

$$I_{\hat{p}\hat{p}} = \hat{\beta}' X'Y^2 \hat{D}_p^{-2} X \hat{\beta} - 2 \hat{\beta}' X'Y \hat{D}_p^{-2} \mathbf{1} + \mathbf{1}' \hat{D}_p^{-2} \mathbf{1}, \quad (5.25)$$

onde $\hat{\beta}$, $\hat{\nu}$ e \hat{p} são as estimativas de máxima verossimilhança de β , ν e p , respectivamente, e \hat{D}_p é calculado com estas estimativas.

Prova

Os componentes da matriz de informação observada para dados sem censura são

$$I_{\hat{\beta}\hat{\beta}} = -\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \beta \partial \beta'} \Big|_{\theta=\hat{\theta}}, \quad (5.26)$$

$$I_{\hat{\beta}\hat{\nu}} = I'_{\hat{\nu}\hat{\beta}} = -\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \beta \partial \nu} \Big|_{\theta=\hat{\theta}}, \quad (5.27)$$

$$I_{\hat{\beta}\hat{p}} = I'_{\hat{p}\hat{\beta}} = -\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \beta \partial p} \Big|_{\theta=\hat{\theta}}, \quad (5.28)$$

$$I_{\nu\nu} = -\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \nu^2} \Big|_{\theta=\hat{\theta}} \quad , \quad (5.29)$$

$$I_{\nu\hat{p}} = I_{\hat{p}\nu} = -\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \nu \partial p} \Big|_{\theta=\hat{\theta}} \quad (5.30)$$

e

$$I_{\hat{p}\hat{p}} = -\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial p^2} \Big|_{\theta=\hat{\theta}} \quad (5.31)$$

Mas, de (5.4) a (5.6), temos

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \tilde{\beta} \partial \tilde{\beta}'} = -p^2 \sum_{i=1}^n \frac{\tilde{x}_i' \tilde{x}_i y_i^2}{\left(1 - p + p \tilde{x}_i \tilde{\beta} y_i\right)^2} - \frac{1}{\nu} X' Y X,$$

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \tilde{\beta} \partial \nu} = \frac{1}{\nu^2} \left(X' Y X \tilde{\beta} - X' \underline{1} \right),$$

$$\begin{aligned} \frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \tilde{\beta} \partial p} &= \sum_{i=1}^n \frac{\tilde{x}_i' y_i}{1 - p + p \tilde{x}_i \tilde{\beta} y_i} - p \sum_{i=1}^n \frac{\tilde{x}_i' \tilde{x}_i \tilde{\beta} y_i^2}{\left(1 - p + p \tilde{x}_i \tilde{\beta} y_i\right)^2} \\ &\quad + p \sum_{i=1}^n \frac{\tilde{x}_i' y_i}{\left(1 - p + p \tilde{x}_i \tilde{\beta} y_i\right)^2}, \end{aligned}$$

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \nu^2} = \frac{n}{2\nu^2} - \frac{1}{\nu^3} \left[\underline{1}' Y^{-1} \underline{1} - 2 \underline{1}' X \tilde{\beta} + \tilde{\beta}' X' Y X \tilde{\beta} \right],$$

$$\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \nu \partial p} = 0$$

e

$$\begin{aligned} \frac{\partial^2 \mathfrak{Z}(\theta)}{\partial p^2} &= -\sum_{i=1}^n \frac{\tilde{\beta}' \tilde{x}_i' y_i^2 \tilde{x}_i \tilde{\beta}}{\left(1 - p + p \tilde{x}_i \tilde{\beta} y_i\right)^2} + 2 \sum_{i=1}^n \frac{\tilde{x}_i \tilde{\beta} y_i}{\left(1 - p + p \tilde{x}_i \tilde{\beta} y_i\right)^2} \\ &\quad - \sum_{i=1}^n \frac{1}{\left(1 - p + p \tilde{x}_i \tilde{\beta} y_i\right)^2}. \end{aligned}$$

Usando a notação dada em (5.6), tem-se as seguintes expressões equivalentes:

$$\sum_{i=1}^n \frac{\tilde{x}_i' \tilde{x}_i \tilde{y}_i^2}{\left(1 - p + p \tilde{x}_i \tilde{\beta} \tilde{y}_i\right)^2} = X' Y^2 D_p^{-2} X,$$

$$\sum_{i=1}^n \frac{\tilde{x}_i' \tilde{y}_i}{1 - p + p \tilde{x}_i \tilde{\beta} \tilde{y}_i} = X' Y D_p^{-1} \underline{1},$$

$$\sum_{i=1}^n \frac{\tilde{x}_i' \tilde{x}_i \tilde{\beta} \tilde{y}_i^2}{\left(1 - p + p \tilde{x}_i \tilde{\beta} \tilde{y}_i\right)^2} = X' Y^2 D_p^{-2} X \tilde{\beta},$$

$$\sum_{i=1}^n \frac{\tilde{x}_i' \tilde{y}_i}{\left(1 - p + p \tilde{x}_i \tilde{\beta} \tilde{y}_i\right)^2} = X' Y D_p^{-2} \underline{1},$$

$$\sum_{i=1}^n \frac{\tilde{\beta}' \tilde{x}_i' \tilde{y}_i^2 \tilde{x}_i \tilde{\beta}}{\left(1 - p + p \tilde{x}_i \tilde{\beta} \tilde{y}_i\right)^2} = \tilde{\beta}' X' Y^2 D_p^{-2} X \tilde{\beta},$$

$$\sum_{i=1}^n \frac{\tilde{x}_i \tilde{\beta} \tilde{y}_i}{\left(1 - p + p \tilde{x}_i \tilde{\beta} \tilde{y}_i\right)^2} = \tilde{\beta}' X' Y D_p^{-2} \underline{1}$$

e

$$\sum_{i=1}^n \frac{1}{\left(1 - p + p \tilde{x}_i \tilde{\beta} \tilde{y}_i\right)^2} = \underline{1}' D_p^{-2} \underline{1}.$$

Logo, as segundas derivadas da log-verossimilhança podem ser escrita como

$$\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \tilde{\beta} \partial \tilde{\beta}'} = -p^2 X' Y^2 D_p^{-2} X - \frac{1}{\nu} X' Y X \quad (5.32)$$

$$\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \tilde{\beta} \partial \nu} = \frac{1}{\nu^2} \left(X' Y X \tilde{\beta} - X' \underline{1} \right), \quad (5.33)$$

$$\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \tilde{\beta} \partial p} = X' Y D_p^{-1} \underline{1} - p X' Y^2 D_p^{-2} X \tilde{\beta} + p X' Y D_p^{-2} \underline{1}, \quad (5.34)$$

$$\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \nu^2} = \frac{n}{2\nu^2} - \frac{1}{\nu^3} \left[\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' X \underset{\sim}{\beta} + \underset{\sim}{\beta}' X' Y X \underset{\sim}{\beta} \right], \quad (5.35)$$

$$\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \nu \partial p} = 0 \quad (5.36)$$

e

$$\frac{\partial^2 \mathfrak{S}(\theta)}{\partial p^2} = - \underset{\sim}{\beta}' X' Y^2 D_p^{-2} X \underset{\sim}{\beta} + 2 \underset{\sim}{\beta}' X' Y D_p^{-2} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-2} \underset{\sim}{1}. \quad (5.37)$$

Por (5.15), (5.35) é

$$\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \nu^2} = -\frac{n}{2\nu^2}. \quad (5.38)$$

Então, os componentes da matriz de informação observada se obtém ao substituir as expressões equivalentes dadas de (5.32) a (5.37) no lado direito das equações dadas de (5.26) a (5.31), respectivamente. Logo, mediante a relação dada em (5.12), obtém-se as expressões dadas em (5.25). \square

Do teorema anterior, pode-se deduzir que para p conhecido a matriz de informação observada é

$$I_{p0} = \begin{bmatrix} \underset{\sim}{I}_{\hat{\beta}\hat{\beta}} & \underset{\sim}{I}_{\hat{\beta}\nu} \\ \underset{\sim}{I}_{\nu\hat{\beta}} & \underset{\sim}{I}_{\nu\nu} \end{bmatrix}, \quad (5.39)$$

onde os valores de seus componentes são achados mediante as fórmulas correspondentes dadas para $M - IG$.

$$\underline{p = 0}$$

Se $p = 0$, então pode-se mostrar por (5.16) e (5.32),

$$\underset{\sim}{I}_{\hat{\beta}\hat{\beta}} = \frac{1}{\nu} X' Y X,$$

por (5.36), temos

$$\underset{\sim}{I}_{\hat{\beta}\nu} = \underset{\sim}{I}'_{\nu\hat{\beta}} = 0,$$

e por (5.39), resulta

$$I_{\hat{\nu}\hat{\nu}} = \frac{n}{2\hat{\nu}^2},$$

onde $\hat{\nu}$ se calcula mediante a fórmula dada em (5.19).

$$\underline{p = 1}$$

Se $p = 1$, então, de (5.7) e (5.32), pode mostrar que

$$I_{\hat{\beta}\hat{\beta}} = X' \hat{D}^{-2} X + \frac{1}{\hat{\nu}} X' Y X,$$

de (5.5) e (5.34)

$$I_{\hat{\beta}\hat{\nu}} = I'_{\hat{\nu}\hat{\beta}} = -\frac{1}{\hat{\nu}} X' \hat{D}^{-1} \mathbf{1},$$

e de (5.39), resulta

$$I_{\hat{\nu}\hat{\nu}} = \frac{n}{2\hat{\nu}^2},$$

onde $\hat{\beta}$ e $\hat{\nu}$ se calculam mediante as fórmulas dadas em (5.22) e (5.23).

5.4 Estimação com dados censurados

Seja $Y_i \sim M - IG \left(\left(X_i \hat{\beta} \right)^{-1}, \nu, p \right)$, $i = 1, \dots, n$, variáveis aleatórias independentes, e com função de densidade da forma (3.4). Assumimos que os r primeiros valores são conhecidos (isto é, $Y_1 = y_1, Y_2 = y_2, \dots, Y_r = y_r$) e os restantes ($n-r$) são censurados á direita nos pontos a_{r+1}, \dots, a_n , respectivamente (isto é, $Y_{r+1} > a_{r+1}, \dots, Y_n > a_n$). Adicionalmente, assumimos $1 \leq r < n$. Logo, a função de verossimilhança para a amostra censurada é

$$\begin{aligned} C(\theta) &= \prod_{i=1}^r g(y_i, \hat{\beta}, \nu, p) \prod_{i=r+1}^n S(a_i, \hat{\beta}, \nu, p) \\ &= \prod_{i=1}^r g(y_i, \hat{\beta}, \nu, p) \prod_{i=r+1}^n \int_{a_i}^{\infty} g(y_i, \hat{\beta}, \nu, p) dy_i \end{aligned}$$

$$= \int_c \prod_{i=1}^n g(y_i, \beta, \nu, p) dY = \int_c L(\theta) dY,$$

onde $\theta = (\beta', \nu, p)'$, $L(\theta)$ é a função de verossimilhança para a amostra sem censura, $dY = \prod_{i=r+1}^n dy_i$ e c é o espaço dos resultados para as observações censuradas, isto é,

$$c = \{Y_1 = y_1, Y_2 = y_2, \dots, Y_r = y_r, Y_{r+1} > a_{r+1}, \dots, Y_n > a_n\}.$$

Agora, para obter as equações de verossimilhança, usamos o resultado dado em (4.55) do na qual demonstrou-se que

$$\frac{\partial \log C(\theta)}{\partial \theta} = E_* \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \right],$$

onde E_* é uma esperança tomada sob o espaço dos resultado c e $\mathfrak{S}(\theta) = \log L(\theta)$. Logo, as estimativas de máxima verossimilhança de θ se obtém ao resolver o seguinte sistema:

$$E_* \begin{bmatrix} \frac{\partial \mathfrak{S}(\theta)}{\partial \beta} \\ \sim \\ \frac{\partial \mathfrak{S}(\theta)}{\partial \nu} \\ \frac{\partial \mathfrak{S}(\theta)}{\partial p} \end{bmatrix} = \underset{\sim}{0}.$$

Seja

$$\begin{aligned} E_*(Y^s D_p^{-1}) = \text{diag} \left\{ y_1^s / \left(1 - p + p \frac{x_1 \beta}{\sim \sim} y_1 \right), \dots, y_r^s / \left(1 - p + p \frac{x_r \beta}{\sim \sim} y_r \right), \right. \\ \left. E \left[Y_{r+1}^s / \left(1 - p + p \frac{x_{r+1} \beta}{\sim \sim} Y_{r+1} \right) \mid Y_{r+1} > a_{r+1} \right], \right. \\ \left. \dots, E \left[Y_n^s / \left(1 - p + p \frac{x_n \beta}{\sim \sim} Y_n \right) \mid Y_n > a_n \right] \right\} \end{aligned} \quad (5.40)$$

e

$$E_*(Y^s) = \text{diag} \left\{ y_1, \dots, y_r, E \left(Y_{r+1}^s \mid Y_{r+1} > a_{r+1} \right), \dots, E \left(Y_n^s \mid Y_n > a_n \right) \right\}, \quad (5.41)$$

para todo $s \in Z$, onde $E \left[Y_i^s / \left(1 - p + p \frac{x_i \beta}{\tilde{z}} Y_i \right) \mid Y_i > a_i \right]$ e $E(Y_i^s \mid Y_i > a_i)$ acham-se mediante as expressões equivalentes dadas nas relações (3.23) e (3.24), respectivamente. Ao substituir as expressões equivalentes dadas em (5.11), (5.12) e (5.13), o sistema de equações transforma-se em:

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \beta} = p X' E_* (Y D_p^{-1}) \underline{1} - \frac{1}{\nu} \left[X' E_* (Y) X \underline{\beta} - X' \underline{1} \right] = \underline{0}, \quad (5.42)$$

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\underline{1}' E_* (Y^{-1}) \underline{1} - 2 \underline{1}' X \underline{\beta} + \beta' X' E_* (Y) X \underline{\beta} \right] = 0 \quad (5.43)$$

e

$$\frac{\partial \mathfrak{S}(\theta)}{\partial p} = \beta' X' E_* (Y D_p^{-1}) \underline{1} - \underline{1}' E_* (D_p^{-1}) \underline{1} = 0. \quad (5.44)$$

Se p é conhecido, então o sistema de equações se reduz a:

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \beta} = p X' E_* (Y D_p^{-1}) \underline{1} - \frac{1}{\nu} \left[X' E_* (Y) X \underline{\beta} - X' \underline{1} \right] = \underline{0} \quad (5.45)$$

e

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\underline{1}' E_* (Y^{-1}) \underline{1} - 2 \underline{1}' X \underline{\beta} + \beta' X' E_* (Y) X \underline{\beta} \right] = 0. \quad (5.46)$$

$$\underline{p} = 0$$

Se $p = 0$, então tem-se as equações de máxima verossimilhança para a regressão Gaussiana Inversa com dados censurados:

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \beta} = -\frac{1}{\nu} \left[X' E_* (Y) X \underline{\beta} - X' \underline{1} \right] = \underline{0} \quad (5.47)$$

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\underline{1}' E_* (Y^{-1}) \underline{1} - 2 \underline{1}' X \underline{\beta} + \beta' X' E_* (Y) X \underline{\beta} \right] = 0. \quad (5.48)$$

Este sistema de equações foi obtido por Whitmore (1983) e pode ser expresso como:

$$\hat{\beta}_{\sim} = [X' E_{\sim}(Y) X]^{-1} X' \underline{1}_{\sim}, \quad (5.49)$$

e substituindo (5.47) em (5.48), resulta

$$\hat{\nu} = \frac{1}{n} \left[\underline{1}' E_{\sim}(Y^{-1}) \underline{1}_{\sim} - \underline{1}' X \hat{\beta}_{\sim} \right]. \quad (5.50)$$

As equações dadas em (5.49) e (5.50) são as mesmas obtidas com a aplicação do algoritmo EM. Logo, para encontrar as estimativas de máxima verossimilhança pode usar-se o seguinte procedimento:

I. Escolher as estimativas iniciais de μ e ν , as quais podem ser as estimativas obtidas com os dados sem censuras mediante as formulas dadas em (5.18) e (5.19).

II. Calcular $E[Y_i | Y_i > a_i]$ e $E[Y_i^{-1} | Y_i > a_i]$, para $i = r+1, \dots, n$ com a fórmula dada em (1.18).

III. Substituir os valores achados na etapa II no lado direito (5.47) e (5.48).

IV. Repetir a etapa II e III até encontrar as estimativas com a precisão desejada.

$p = 1$

Se $p = 1$ nas equações dadas em (5.45) e (5.46), então, tem-se as equações de verossimilhança para Regressão de dados da Recíproca Complementar de uma variável Gaussiana Inversa:

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \beta_{\sim}} = X' D^{-1} \underline{1}_{\sim} - \frac{1}{\nu} \left[X' E_{\sim}(Y) X \beta_{\sim} - X' \underline{1}_{\sim} \right] = 0 \quad (5.51)$$

$$\frac{\partial \mathfrak{S}(\theta)}{\partial \nu} = -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\underline{1}' E_{\sim}(Y^{-1}) \underline{1}_{\sim} - 2 \underline{1}' X \beta_{\sim} + \beta_{\sim}' X' E_{\sim}(Y) X \beta_{\sim} \right] = 0, \quad (5.52)$$

onde

$$D = \text{diag} \left\{ x_{\sim 1} \beta, \dots, x_{\sim n} \beta \right\}.$$

Resolvendo este sistema obtém-se

$$\hat{\beta}_{\sim} = \left[\left(\underset{\sim}{1}' X \right)' \left(\underset{\sim}{1}' X \right) \right]^{-1} X \underset{\sim}{1}' E_* (Y^{-1}) \underset{\sim}{1}, \quad (5.53)$$

e

$$\hat{\nu} = \frac{1}{n} \left[X' E_* (Y) X \underset{\sim}{\beta} - X' \underset{\sim}{1} \right]. \quad (5.54)$$

Este mesmo sistema de equações é o mesmo se obtém ao aplicar o algoritmo EM dado por Dempster, Laird e Rubin (1977). Logo, pode empregar um procedimento semelhante ao caso anterior:

I. Escolher as estimativas iniciais de μ e ν , as quais podem ser as estimativas obtidas com os dados sem censuras mediante as formulas dadas em (5.22) e (5.23).

II. Calcular $E[Y_i | Y_i > a_i]$ e $E[Y_i^{-1} | Y_i > a_i]$, para $i = r + 1, \dots, n$ com a fórmula dada em (2.10).

III. Substituir os valores achado na etapa II no lado direito (5.53) e (5.54).

IV. Repetir a etapa II e III até encontrar as estimativas com a precisão desejada.

Igualmente aos casos anteriores, o sistema de equações dado em (5.40) e (5.41) pode ser resolvido mediante o uso de métodos iterativos como Newton Raphson. Este procedimento é o mesmo que usado para a regressão com p conhecido e sem censura, descrito na Seção 5.2.

O sistema de equações dado de (5.42) a (5.44) tem soluções múltiplas, mas, para cada valor de p existe uma única solução. Assim, para encontrar o máximo global ou o maior dos máximos, pode-se empregar o mesmo procedimento usado para encontrar o máximo global ou o maior dos máximos para a regressão da $M - IG$ com dados não censurados, apresentado na Seção 5.2 para p desconhecido, com a única modificação,

na etapa IV, no cálculo do valor da log-verossimilhança:

$$\log C(\theta) = \sum_{i=1}^r \log \left[g \left(y_i, \beta, \nu, p \right) \right] + \sum_{i=r+1}^n \log \left[S \left(a_i, \beta, \nu, p \right) \right].$$

Pode acontecer que o máximo valor do log-verossimilhança ocorra na fronteira do espaço paramétrico de p ($p = 0$ ou $p = 1$) ou fora deste espaço. Quando isto ocorre, é necesario fazer análises adicionais para determinar a qualidade de ajuste do modelo na fronteira do espaço paramétrico do p . Um procedimento para fazer a adequação do modelo mediante a análise de resíduos, é descrito na Seção 5.6.

Para usar o método de Newton Raphson, se precisa da matriz de informação ou de qualquer estimativa consistente. Neste caso, usamos a matriz de informação observada, definida no seguinte teorema.

Teorema 5.2

A matriz de informação observada para o modelo de regressão com dados censurados da M-IG é

$$I_0 = \begin{bmatrix} I_{\hat{\beta}\hat{\beta}} & I_{\hat{\beta}\hat{\nu}} & I_{\hat{\beta}\hat{p}} \\ I_{\hat{\nu}\hat{\beta}} & I_{\hat{\nu}\hat{\nu}} & I_{\hat{\nu}\hat{p}} \\ I_{\hat{p}\hat{\beta}} & I_{\hat{p}\hat{\nu}} & I_{\hat{p}\hat{p}} \end{bmatrix}, \quad (5.55)$$

onde

$$I_{\hat{\beta}\hat{\beta}} = X' \left\{ \frac{\hat{p}}{\hat{\nu}} \left[\hat{p}\hat{\nu}\hat{N}_1^2 + 2\hat{D}(\hat{N}_2 - \hat{M}_1\hat{N}_1) \right] + \frac{1}{\hat{\nu}^2} \left[\hat{\nu}\hat{M}_1 - \hat{D}(\hat{M}_2 - \hat{M}_1^2)\hat{D} \right] \right\} X,$$

$$I_{\hat{\beta}\hat{\nu}} = I'_{\hat{\nu}\hat{\beta}} = \frac{1}{2\hat{\nu}^3} X' \left[\hat{D}(\hat{M}_2 - \hat{M}_1^2)\hat{D} + (I - \hat{M}_1\hat{M}_{-1}) \right] \hat{D} \underline{1} \\ - \frac{\hat{p}}{2\hat{\nu}^2} X' \left[\hat{D}(\hat{N}_2 - \hat{N}_1\hat{M}_1)\hat{D} + (2\hat{\nu}\hat{N}_1 + \hat{N}_0 - \hat{N}_1\hat{M}_{-1}) \right] \underline{1},$$

$$\begin{aligned}
I_{\hat{\beta}\hat{p}} = I_{\hat{p}\hat{\beta}} &= \frac{1}{\hat{\nu}} X' \left\{ \hat{D} [\hat{N}_2 - \hat{M}_1 \hat{N}_1] - [\hat{N}_1 - \hat{M}_1 \hat{N}_0] \right\} \hat{D} \underline{1} \\
&\quad + X' \hat{N}_1 [\hat{p} \hat{N}_1 - \hat{p} \hat{N}_0 - I] \underline{1}, \\
I_{\hat{\nu}\hat{\nu}} &= \frac{n}{2\hat{\nu}^2} - \frac{1}{4\hat{\nu}^4} \underline{1}' \left\{ \hat{D}^2 [\hat{M}_2 - \hat{M}_1^2] \hat{D}^2 + 2\hat{D} [I - \hat{M}_1 \hat{M}_{-1}] \hat{D} + [\hat{M}_{-2} - \hat{M}_{-1}^2] \right\} \underline{1}, \\
I_{\hat{\nu}\hat{p}} = I_{\hat{p}\hat{\nu}} &= \frac{1}{2\hat{\nu}^2} \underline{1}' \left\{ [\hat{N}_{-1} - \hat{M}_{-1} \hat{N}_0] - [\hat{N}_0 - \hat{M}_{-1} \hat{N}_1] \hat{D} + \hat{D}^2 [\hat{N}_1 - \hat{M}_1 \hat{N}_0] \right. \\
&\quad \left. - \hat{D}^2 [\hat{N}_2 - \hat{M}_1 \hat{N}_1] \hat{D} \right\} \underline{1},
\end{aligned}$$

e

$$I_{\hat{p}\hat{p}} = \underline{1}' \hat{D} \hat{N}_1^2 \hat{D} \underline{1} - 2 \underline{1}' \hat{N}_0 \hat{N}_1 \hat{D} \underline{1} + \underline{1}' \hat{N}_0^2 \underline{1}, \quad (5.56)$$

onde

$$M_s = E_*(Y^s) \text{ e } N_s = E_*(Y^s D_p^{-1}), \quad \forall s \in Z, \quad (5.57)$$

$\hat{\beta}$, $\hat{\nu}$ e \hat{p} são estimativas de β , ν e p , respectivamente, e \hat{D} , \hat{N}_s , e \hat{M}_s são valores D , N_s e M_s calculados com estas estimativas para todo $s \in Z$.

Prova

A matriz de informação observada, para o modelo de regressão com dados censurados da M-IG se define por

$$I_0 = - \frac{\partial^2 \log C(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta = \hat{\theta} = (\hat{\beta}, \hat{\nu}, \hat{p})'}$$

onde $\hat{\beta}$, $\hat{\nu}$ e \hat{p} são as estimativas de máxima verossimilhança de β , ν e p , respectivamente. Mas, no capítulo anterior, demonstrou-se

$$I_0 = -E_* \left[\frac{\partial^2 \mathfrak{S}(\theta)}{\partial \theta \partial \theta'} + \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \right] \left[\frac{\partial \mathfrak{S}(\theta)}{\partial \theta} \right]' \right]_{\theta = \hat{\theta}}$$

Logo os componentes de I_0 são:

$$I_{\hat{\beta}\hat{\beta}} = -E_* \left[\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \hat{\beta}^2} + \left[\frac{\partial \mathfrak{Z}(\theta)}{\partial \hat{\beta}} \right] \cdot \left[\frac{\partial \mathfrak{Z}(\theta)}{\partial \hat{\beta}} \right]' \right]_{\theta=\hat{\theta}}, \quad (5.58)$$

$$I_{\hat{\beta}\hat{\nu}} = I'_{\hat{\nu}\hat{\beta}} = -E_* \left[\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \hat{\beta} \partial \hat{\nu}} + \left[\frac{\partial \mathfrak{Z}(\theta)}{\partial \hat{\beta}} \right] \left[\frac{\partial \mathfrak{Z}(\theta)}{\partial \hat{\nu}} \right] \right]_{\theta=\hat{\theta}}, \quad (5.59)$$

$$I_{\hat{\beta}\hat{p}} = I'_{\hat{p}\hat{\beta}} = -E_* \left[\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \hat{\beta} \partial \hat{p}} + \left[\frac{\partial \mathfrak{Z}(\theta)}{\partial \hat{\beta}} \right] \left[\frac{\partial \mathfrak{Z}(\theta)}{\partial \hat{p}} \right]' \right]_{\theta=\hat{\theta}}, \quad (5.60)$$

$$I_{\hat{\nu}\hat{\nu}} = -E_* \left[\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \hat{\nu}^2} + \left[\frac{\partial \mathfrak{Z}(\theta)}{\partial \hat{\nu}} \right]^2 \right]_{\theta=\hat{\theta}}, \quad (5.61)$$

$$I_{\hat{\nu}\hat{p}} = I_{\hat{p}\hat{\nu}} = -E_* \left[\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \hat{\nu} \partial \hat{p}} + \left[\frac{\partial \mathfrak{Z}(\theta)}{\partial \hat{\nu}} \right] \left[\frac{\partial \mathfrak{Z}(\theta)}{\partial \hat{p}} \right] \right]_{\theta=\hat{\theta}}, \quad (5.62)$$

e

$$I_{\hat{p}\hat{p}} = -E_* \left[\frac{\partial^2 \mathfrak{Z}(\theta)}{\partial \hat{p}^2} + \left[\frac{\partial \mathfrak{Z}(\theta)}{\partial \hat{p}} \right]^2 \right]_{\theta=\hat{\theta}}. \quad (5.63)$$

Cálculo de $I_{\hat{\beta}\hat{\beta}}$

Sustituindo (5.32) e (5.11) em (5.58), se tem

$$I_{\hat{\beta}\hat{\beta}} = -E_* \left\{ -p^2 X' Y^2 D_p^{-2} X - \frac{1}{\nu} X' Y X + \left[p X' Y D_p^{-1} \underline{1} - \frac{1}{\nu} (X' Y X \hat{\beta} - X' \underline{1}) \right] \cdot \left[p X' Y D_p^{-1} \underline{1} - \frac{1}{\nu} (X' Y X \hat{\beta} - X' \underline{1}) \right]' \right\}_{\theta=\hat{\theta}}.$$

Seja

$$D = \text{diag} \left\{ x_{\hat{\beta}_1}, \dots, x_{\hat{\beta}_n} \right\}.$$

Então,

$$X \hat{\beta} = D \underline{1} \text{ e } \hat{\beta}' X' = \underline{1}' D. \quad (5.64)$$

Logo, com esta notação, $I_{\hat{\beta}\hat{\beta}}$ pode ser escrita como:

$$\begin{aligned}
 I_{\hat{\beta}\hat{\beta}} &= -E_* \left\{ -p^2 X' Y^2 D_p^{-2} X - \frac{1}{\nu} X' Y X + \left[p X' Y D_p^{-1} \underline{1} - \frac{1}{\nu} \left(X' Y D \underline{1} - X' \underline{1} \right) \right] \right. \\
 &\quad \left. \cdot \left[p X' Y D_p^{-1} \underline{1} - \frac{1}{\nu} \left(X' Y D \underline{1} - X' \underline{1} \right) \right]' \right\}_{\theta=\hat{\theta}} \\
 &= \left\{ p^2 X' E_* \left(Y^2 D_p^{-2} \right) X + \frac{1}{\nu} X' E_* (Y) X - p^2 X' E_* \left(Y D_p^{-1} \underline{1} \underline{1}' D_p^{-1} Y \right) X \right. \\
 &\quad + \frac{2}{\nu} X' E_* \left(Y D \underline{1} \underline{1}' D_p^{-1} Y \right) X - \frac{2}{\nu} X' \underline{1} \underline{1}' E_* \left(D_p^{-1} Y \right) X \\
 &\quad + \frac{2}{\nu} X' E_* \left(Y D_p^{-1} \underline{1} \underline{1}' D Y \right) X - \frac{2}{\nu} X' E_* \left(Y D_p^{-1} \right) \underline{1} \underline{1}' X \\
 &\quad - \frac{1}{\nu^2} X' E_* \left(Y D \underline{1} \underline{1}' D Y \right) X + \frac{1}{\nu^2} X' \underline{1} \underline{1}' D E_* (Y) X \\
 &\quad \left. + \frac{1}{\nu^2} \left(X' E_* (Y) D \underline{1} - X' \underline{1} \right) \underline{1}' X \right\}_{\theta=\hat{\theta}}
 \end{aligned}$$

Mas, D e Y são matrizes diagonais, então podemos comutá-las. Logo, mediante esta propriedade, a relação anterior pode ser escrita como:

$$\begin{aligned}
 I_{\hat{\beta}\hat{\beta}} &= \left\{ p^2 X' E_* \left(Y^2 D_p^{-2} \right) X + \frac{1}{\nu} X' E_* (Y) X - p^2 X' E_* \left(Y D_p^{-1} \underline{1} \underline{1}' D_p^{-1} Y \right) X \right. \\
 &\quad + \frac{2}{\nu} X' D E_* \left(Y \underline{1} \underline{1}' D_p^{-1} Y \right) X - \frac{2}{\nu} X' \underline{1} \underline{1}' E_* \left(D_p^{-1} Y \right) X \\
 &\quad + \frac{2}{\nu} X' E_* \left(Y D_p^{-1} \underline{1} \underline{1}' Y \right) D X - \frac{2}{\nu} X' E_* \left(Y D_p^{-1} \right) \underline{1} \underline{1}' X \\
 &\quad - \frac{1}{\nu^2} X' D E_* \left(Y \underline{1} \underline{1}' Y \right) D X + \frac{1}{\nu^2} X' \underline{1} \underline{1}' D E_* (Y) X \\
 &\quad \left. + \frac{1}{\nu^2} \left(X' E_* (Y) D \underline{1} - X' \underline{1} \right) \underline{1}' X \right\}_{\theta=\hat{\theta}}.
 \end{aligned}$$

Agora, usando a notação dada em (5.57) e substituindo as expressões equivalentes

dadas em (4.79), temos o resultado:

$$\begin{aligned}
 I_{\hat{\beta}\hat{\beta}} &= \left\{ p^2 X' E_* (Y^2 D_p^{-2}) X + \frac{1}{\nu} X' M_1 X - p^2 X' \left[E_* (Y^2 D_p^{-2}) - N_1^2 + N_1 \underline{\underline{1}} \underline{\underline{1}}' N_1 \right] X \right. \\
 &\quad + \frac{p}{\nu} X' D \left[N_2 - M_1 N_1 + M_1 \underline{\underline{1}} \underline{\underline{1}}' N_1 \right] X - \frac{p}{\nu} X' \underline{\underline{1}} \underline{\underline{1}}' N_1 X \\
 &\quad + \frac{p}{\nu} X' \left[N_2 - N_1 M_1 + N_1 \underline{\underline{1}} \underline{\underline{1}}' M_1 \right] DX - \frac{p}{\nu} X' N_1 \underline{\underline{1}} \underline{\underline{1}}' X \\
 &\quad - \frac{1}{\nu^2} X' D \left[M_2 - M_1^2 + M_1 \underline{\underline{1}} \underline{\underline{1}}' M_1 \right] DX + \frac{1}{\nu^2} X' \underline{\underline{1}} \underline{\underline{1}}' D M_1 X \\
 &\quad \left. + \frac{1}{\nu^2} \left(X' M_1 D \underline{\underline{1}} - X' \underline{\underline{1}} \right) \underline{\underline{1}}' X \right\}_{\theta=\hat{\theta}} \\
 &= \left\{ \frac{1}{\nu} X' M_1 X + p^2 X' N_1^2 X - \frac{1}{\nu^2} X' D \left[M_2 - M_1^2 \right] DX \right. \\
 &\quad + \frac{2p}{\nu} X' D \left[N_2 - M_1 N_1 \right] X + \frac{p}{\nu} \left(X' D M_1 \underline{\underline{1}} - X' \underline{\underline{1}} - p\nu X' N_1 \underline{\underline{1}} \right) \underline{\underline{1}}' N_1 X \\
 &\quad \left. - \frac{1}{\nu^2} \left(X' D M_1 \underline{\underline{1}} - X' \underline{\underline{1}} - p\nu X' N_1 \underline{\underline{1}} \right) \left(\underline{\underline{1}}' D M_1 X - \underline{\underline{1}}' X \right) \right\}_{\theta=\hat{\theta}}.
 \end{aligned} \tag{5.65}$$

Mas, ao substituir a notação dada em capítulo anterior em (5.57) na equação (5.58), temos

$$p\nu X' N_1 \underline{\underline{1}} = X' D M_1 \underline{\underline{1}} - X' \underline{\underline{1}} \tag{5.66}$$

para todo β , ν e p que sejam solução do sistema de equações dado de (5.42) a (5.44).

Logo, substituindo (5.66) em (5.65), resulta em:

$$I_{\hat{\beta}\hat{\beta}} = X' \left\{ \frac{\hat{p}}{\hat{\nu}} \left[\hat{p} \hat{\nu} \hat{N}_1^2 + 2 \hat{D} (\hat{N}_2 - \hat{M}_1 \hat{N}_1) \right] + \frac{1}{\hat{\nu}^2} \left[\hat{\nu} \hat{M}_1 - \hat{D} (\hat{M}_2 - \hat{M}_1^2) \hat{D} \right] \right\} X. \tag{5.67}$$

Cálculo de $I_{\hat{\beta}\hat{\nu}}$

Substituindo (5.33), (5.11) e (5.12) em (5.59), resulta em:

$$\begin{aligned}
 I_{\hat{\beta}\hat{\nu}} &= -E_* \left\{ \frac{1}{\nu^2} \left(X' Y X \underline{\underline{\beta}} - X' \underline{\underline{1}} \right) + \left[p X' Y D_p^{-1} \underline{\underline{1}} - \frac{1}{\nu} \left(X' Y X \underline{\underline{\beta}} - X' \underline{\underline{1}} \right) \right] \right. \\
 &\quad \left. \cdot \left[-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left(\underline{\underline{1}}' Y^{-1} \underline{\underline{1}} - 2 \underline{\underline{1}}' X \underline{\underline{\beta}} + \underline{\underline{\beta}}' X' Y X \underline{\underline{\beta}} \right) \right] \right\}_{\theta=\hat{\theta}}.
 \end{aligned}$$

Agora, mediante a notação dada em (5.64), temos

$$\begin{aligned}
 I_{\hat{\beta}\hat{\nu}} &= -E_* \left\{ \frac{1}{\nu^2} \left(X' Y D \underline{\underline{1}} - X' \underline{\underline{1}} \right) + \left[p X' Y D_p^{-1} \underline{\underline{1}} - \frac{1}{\nu} \left(X' Y D \underline{\underline{1}} - X' \underline{\underline{1}} \right) \right] \right. \\
 &\quad \left. \cdot \left[-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left(\underline{\underline{1}}' Y^{-1} \underline{\underline{1}} - 2 \underline{\underline{1}}' D \underline{\underline{1}} + \underline{\underline{1}}' D Y D \underline{\underline{1}} \right) \right] \right\}_{\theta=\hat{\theta}} \\
 &= \left\{ -\frac{1}{\nu^2} \left[X' E_* (Y) D \underline{\underline{1}} - X' \underline{\underline{1}} \right] - \left[p X' E_* (Y D_p^{-1}) \underline{\underline{1}} - \frac{1}{\nu} \left(X' E_* (Y) D \underline{\underline{1}} - X' \underline{\underline{1}} \right) \right] \right. \\
 &\quad \cdot \left(-\frac{n}{2\nu} - \frac{1}{\nu^2} \underline{\underline{1}}' D \underline{\underline{1}} \right) - \frac{p}{2\nu^2} X' E_* \left(Y D_p^{-1} \underline{\underline{1}} \underline{\underline{1}}' Y^{-1} \right) \underline{\underline{1}} + \frac{1}{2\nu^3} X' E_* \left(Y D \underline{\underline{1}} \underline{\underline{1}}' Y^{-1} \right) \underline{\underline{1}} \\
 &\quad - \frac{1}{2\nu^3} X' \underline{\underline{1}} \underline{\underline{1}}' E_* (Y^{-1}) \underline{\underline{1}} - \frac{p}{2\nu^3} X' E_* \left(Y D_p^{-1} \underline{\underline{1}} \underline{\underline{1}}' D Y \right) D \underline{\underline{1}} \\
 &\quad \left. + \frac{1}{2\nu^3} X' E_* \left(Y D \underline{\underline{1}} \underline{\underline{1}}' D Y \right) D \underline{\underline{1}} - \frac{1}{2\nu^3} X' \underline{\underline{1}} \underline{\underline{1}}' D E_* (Y) D \underline{\underline{1}} \right\}_{\theta=\hat{\theta}}.
 \end{aligned}$$

Como D, Y e Y^{-1} são matrizes diagonais, então, podemos comutá-las. Usando esta propriedade, a relação anterior pode ser escrita como:

$$\begin{aligned}
 I_{\hat{\beta}\hat{\nu}} &= \left\{ -\frac{1}{\nu^2} \left[X' E_* (Y) D \underline{\underline{1}} - X' \underline{\underline{1}} \right] - \left[p X' E_* (Y D_p^{-1}) \underline{\underline{1}} - \frac{1}{\nu} \left(X' E_* (Y) D \underline{\underline{1}} - X' \underline{\underline{1}} \right) \right] \right. \\
 &\quad \cdot \left(-\frac{n}{2\nu} - \frac{1}{\nu^2} \underline{\underline{1}}' D \underline{\underline{1}} \right) - \frac{p}{2\nu^2} X' E_* \left(Y D_p^{-1} \underline{\underline{1}} \underline{\underline{1}}' Y^{-1} \right) \underline{\underline{1}} + \frac{1}{2\nu^3} X' D E_* \left(Y \underline{\underline{1}} \underline{\underline{1}}' Y^{-1} \right) \underline{\underline{1}} \\
 &\quad - \frac{1}{2\nu^3} X' \underline{\underline{1}} \underline{\underline{1}}' E_* (Y^{-1}) \underline{\underline{1}} - \frac{p}{2\nu^3} X' E_* \left(Y D_p^{-1} \underline{\underline{1}} \underline{\underline{1}}' Y \right) D^2 \underline{\underline{1}} \\
 &\quad \left. + \frac{1}{2\nu^3} X' D E_* \left(Y \underline{\underline{1}} \underline{\underline{1}}' Y \right) D^2 \underline{\underline{1}} - \frac{1}{2\nu^3} X' \underline{\underline{1}} \underline{\underline{1}}' E_* (Y) D^2 \underline{\underline{1}} \right\}_{\theta=\hat{\theta}}.
 \end{aligned} \tag{5.68}$$

Usando a notação dada em (5.57) e as expressões equivalentes dadas em (4.79) e (5.68) temos

$$\begin{aligned}
 I_{\hat{\beta}\hat{\nu}} &= \left\{ -\frac{1}{\nu^2} \left[X' M_1 D \underline{\underline{1}} - X' \underline{\underline{1}} \right] - \left[p X' N_1 \underline{\underline{1}} - \frac{1}{\nu} \left(X' M_1 D \underline{\underline{1}} - X' \underline{\underline{1}} \right) \right] \right. \\
 &\quad \cdot \left(-\frac{n}{2\nu} - \frac{1}{\nu^2} \underline{\underline{1}}' D \underline{\underline{1}} \right) - \frac{p}{2\nu^2} X' \left(N_0 - N_1 M_{-1} + N_1 \underline{\underline{1}} \underline{\underline{1}}' M_{-1} \right) \underline{\underline{1}} \\
 &\quad + \frac{1}{2\nu^3} X' D \left(I - M_1 M_{-1} + M_1 \underline{\underline{1}} \underline{\underline{1}}' M_{-1} \right) \underline{\underline{1}} - \frac{1}{2\nu^3} X' \underline{\underline{1}} \underline{\underline{1}}' M_{-1} \underline{\underline{1}} \\
 &\quad - \frac{p}{2\nu^3} X' \left(N_2 - N_1 M_1 + N_1 \underline{\underline{1}} \underline{\underline{1}}' M_1 \right) D^2 \underline{\underline{1}} \\
 &\quad \left. + \frac{1}{2\nu^3} X' D \left(M_2 - M_1^2 + M_1 \underline{\underline{1}} \underline{\underline{1}}' M_1 \right) D^2 \underline{\underline{1}} - \frac{1}{2\nu^3} X' \underline{\underline{1}} \underline{\underline{1}}' M_1 D^2 \underline{\underline{1}} \right\}_{\theta=\hat{\theta}}
 \end{aligned}$$

$$\begin{aligned}
&= \left\{ -\frac{1}{\nu^2} \left[X' M_1 D \underset{\sim}{1} - X' \underset{\sim}{1} \right] - \left[p X' N_1 \underset{\sim}{1} - \frac{1}{\nu} \left(X' M_1 D \underset{\sim}{1} - X' \underset{\sim}{1} \right) \right] \right. \\
&\quad \cdot \left(-\frac{n}{2\nu} - \frac{1}{\nu^2} \underset{\sim}{1}' D \underset{\sim}{1} \right) - \frac{p}{2\nu^2} X' (N_0 - N_1 M_{-1}) \underset{\sim}{1} \\
&\quad + \frac{1}{2\nu^3} X' D (I - M_1 M_{-1}) \underset{\sim}{1} - \frac{p}{2\nu^2} X' (N_2 - N_1 M_1) D^2 \underset{\sim}{1} \\
&\quad + \frac{1}{2\nu^3} X' D (M_2 - M_1^2) D^2 \underset{\sim}{1} - \frac{1}{2\nu^2} \left[p X' N_1 \underset{\sim}{1} - \frac{1}{\nu} \left(X' D M_1 \underset{\sim}{1} - X' \underset{\sim}{1} \right) \right] \\
&\quad \cdot \left. \left(\underset{\sim}{1}' M_{-1} \underset{\sim}{1} + \underset{\sim}{1}' M_1 D^2 \underset{\sim}{1} \right) \right\}_{\theta=\hat{\theta}}.
\end{aligned} \tag{5.69}$$

Usando a expressão equivalente dada em (5.66), tem-se:

$$\begin{aligned}
I_{\hat{\beta}\hat{\nu}} &= \frac{1}{2\nu^3} X' \left[\hat{D} (\hat{M}_2 - \hat{M}_1^2) \hat{D} + (I - \hat{M}_1 \hat{M}_{-1}) \right] \hat{D} \underset{\sim}{1} \\
&\quad - \frac{p}{2\nu^2} X' \left[\hat{D} (\hat{N}_2 - \hat{N}_1 \hat{M}_1) \hat{D} + (2\nu \hat{N}_1 + \hat{N}_0 - \hat{N}_1 \hat{M}_{-1}) \right] \underset{\sim}{1}.
\end{aligned} \tag{5.70}$$

Cálculo de $I_{\hat{\beta}\hat{p}}$

Substituindo (5.34), (5.11) e (5.13) em (5.60), tem-se

$$\begin{aligned}
I_{\hat{\beta}\hat{p}} &= -E_* \left\{ X' Y D_p^{-1} \underset{\sim}{1} - p X' Y^2 D_p^{-2} X \underset{\sim}{\beta} + p X' Y D_p^{-2} \underset{\sim}{1} + \left[p X' Y D_p^{-1} \underset{\sim}{1} - \frac{1}{\nu} \left(X' Y X \underset{\sim}{\beta} - X' \underset{\sim}{1} \right) \right] \right. \\
&\quad \cdot \left. \left[\underset{\sim}{\beta}' X' Y D_p^{-1} \underset{\sim}{1} - \underset{\sim}{1} D_p^{-1} \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}}.
\end{aligned}$$

Usando a notação dada (5.64), na relação anterior temos

$$\begin{aligned}
I_{\hat{\beta}\hat{p}} &= -E_* \left\{ X' Y D_p^{-1} \underset{\sim}{1} - p X' Y^2 D_p^{-2} D \underset{\sim}{1} + p X' Y D_p^{-2} \underset{\sim}{1} + \left[p X' Y D_p^{-1} \underset{\sim}{1} - \frac{1}{\nu} \left(X' Y D \underset{\sim}{1} - X' \underset{\sim}{1} \right) \right] \right. \\
&\quad \cdot \left. \left[\underset{\sim}{1}' D Y D_p^{-1} \underset{\sim}{1} - \underset{\sim}{1} D_p^{-1} \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}} \\
&= E_* \left\{ -X' Y D_p^{-1} \underset{\sim}{1} + p X' Y^2 D_p^{-2} D \underset{\sim}{1} - p X' Y D_p^{-2} \underset{\sim}{1} \right. \\
&\quad - \frac{1}{\nu} X' \underset{\sim}{1} \cdot \left[\underset{\sim}{1}' D Y D_p^{-1} \underset{\sim}{1} - \underset{\sim}{1} D_p^{-1} \underset{\sim}{1} \right] \\
&\quad - p X' Y D_p^{-1} \underset{\sim}{1} \underset{\sim}{1}' D Y D_p^{-1} \underset{\sim}{1} + \frac{1}{\nu} X' Y D \underset{\sim}{1} \underset{\sim}{1}' D Y D_p^{-1} \underset{\sim}{1} \\
&\quad \left. + p X' Y D_p^{-1} \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} - \frac{1}{\nu} X' Y D \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \right\}_{\theta=\hat{\theta}}.
\end{aligned}$$

Mas, D, Y e D_p^{-1} são matrizes diagonais. Então, podem comutar-se, e usando esta propriedade, a relação anterior se transforma em

$$\begin{aligned}
 I_{\hat{\beta}\hat{p}} = & \left\{ -X'E_* \left(YD_p^{-1} \right) \underline{1} + pX'E_* \left(Y^2D_p^{-2} \right) D \underline{1} - pX'E_* \left(YD_p^{-2} \right) \underline{1} \right. \\
 & - \frac{1}{\nu} X' \underline{1} \cdot \left[\underline{1}' DE_* \left(YD_p^{-1} \right) \underline{1} - \underline{1} E_* \left(D_p^{-1} \right) \underline{1} \right] \\
 & - pX'E_* \left(YD_p^{-1} \underline{1}\underline{1}' YD_p^{-1} \right) D \underline{1} + \frac{1}{\nu} X'DE_* \left(Y \underline{1}\underline{1}' YD_p^{-1} \right) D \underline{1} \\
 & \left. + pX'E_* \left(YD_p^{-1} \underline{1}\underline{1}' D_p^{-1} \right) \underline{1} - \frac{1}{\nu} X'DE_* \left(Y \underline{1}\underline{1}' D_p^{-1} \right) \underline{1} \right\}_{\theta=\hat{\theta}}.
 \end{aligned} \tag{5.71}$$

Usando as notações dada em (5.57) e as expressões equivalentes dadas em (4.79) e (5.71), temos:

$$\begin{aligned}
 I_{\hat{\beta}\hat{p}} = & \left\{ -X'N_1 \underline{1} + pX'E_* \left(Y^2D_p^{-2} \right) D \underline{1} - pX'E_* \left(YD_p^{-2} \right) \underline{1} \right. \\
 & - \frac{1}{\nu} X' \underline{1} \cdot \left[\underline{1}' DN_1 \underline{1} - \underline{1} N_0 \underline{1} \right] \\
 & - pX' \left[E_* \left(Y^2D_p^{-2} \right) - N_1^2 + N_1 \underline{1}\underline{1}' N_1 \right] D \underline{1} \\
 & + \frac{1}{\nu} X'D \left[N_2 - M_1N_1 + M_1 \underline{1}\underline{1}' N_1 \right] D \underline{1} \\
 & + pX' \left[E_* \left(YD_p^{-2} \right) - N_1N_0 + N_1 \underline{1}\underline{1}' N_0 \right] \underline{1} \\
 & \left. - \frac{1}{\nu} X'D \left(N_1 - M_1N_0 + M_1 \underline{1}\underline{1}' N_0 \right) \underline{1} \right\}_{\theta=\hat{\theta}} \\
 = & \left\{ -X'N_1 \underline{1} + pX'N_1^2 D \underline{1} - pX'N_1N_0 \underline{1} \right. \\
 & - \frac{1}{\nu} X'D \left[N_1 - M_1N_0 \right] \underline{1} + \frac{1}{\nu} X'D \left[N_2 - M_1N_1 \right] D \underline{1} \\
 & \left. - \left[pX'N_1 \underline{1} - \frac{1}{\nu} \left(X'DM_1 \underline{1} - X' \underline{1} \right) \right] \left[\underline{1}' N_1 D \underline{1} - \underline{1}' N_0 \underline{1} \right] \right\}_{\theta=\hat{\theta}}.
 \end{aligned}$$

Usando a expressão equivalente dada em (5.66), temos:

$$\begin{aligned}
 I_{\hat{\beta}\hat{p}} = & \frac{1}{\nu} X' \left\{ \hat{D} \left[\hat{N}_2 - \hat{M}_1 \hat{N}_1 \right] - \left[\hat{N}_1 - \hat{M}_1 \hat{N}_0 \right] \right\} \hat{D} \underline{1} \\
 & + X' \hat{N}_1 \left[\hat{p} \hat{N}_1 - \hat{p} \hat{N}_0 - I \right] \underline{1}.
 \end{aligned} \tag{5.72}$$

Cálculo $I_{\hat{\nu}\hat{\nu}}$

Sustituindo as expressões equivalentes dadas em (5.35) e (5.11) em (5.61), tem-se

$$I_{\hat{\nu}\hat{\nu}} = -E_* \left\{ \frac{n}{2\nu^2} - \frac{1}{\nu^3} \left(\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' X \underset{\sim}{\beta} + \underset{\sim}{\beta}' XYX \underset{\sim}{\beta} \right) \right. \\ \left. \left[-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left(\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' X \underset{\sim}{\beta} + \underset{\sim}{\beta}' XYX \underset{\sim}{\beta} \right) \right]^2 \right\}_{\theta=\hat{\theta}}.$$

Agora, mediante a notação dada em (5.64), temos o resultado

$$I_{\hat{\nu}\hat{\nu}} = -E_* \left\{ \frac{n}{2\nu^2} - \frac{1}{\nu^3} \left(\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' DYD \underset{\sim}{1} \right) \right. \\ \left. + \left[-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left(\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' DYD \underset{\sim}{1} \right) \right]^2 \right\}_{\theta=\hat{\theta}} \\ = -E_* \left\{ \frac{n}{2\nu^2} - \frac{1}{\nu^3} \left(\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' DYD \underset{\sim}{1} \right) \right. \\ \left. + \frac{n^2}{4\nu^2} - \frac{n}{2\nu^3} \left(\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' DYD \underset{\sim}{1} \right) \right. \\ \left. + \frac{1}{4\nu^4} \left(\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} \underset{\sim}{1}' Y^{-1} \underset{\sim}{1} + 2 \underset{\sim}{1}' DYD \underset{\sim}{1} \underset{\sim}{1}' Y^{-1} \underset{\sim}{1} + \underset{\sim}{1}' DYD \underset{\sim}{1} \underset{\sim}{1}' DYD \underset{\sim}{1} \right) \right. \\ \left. - \frac{1}{\nu^4} \underset{\sim}{1}' D \underset{\sim}{1} \left(\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} + \underset{\sim}{1}' DYD \underset{\sim}{1} \right) + \frac{1}{\nu^4} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' D \underset{\sim}{1} \right\}_{\theta=\hat{\theta}}.$$

Como D , Y e Y^{-1} são matrizes diagonais, podemos comutar-las. Usando esta propriedade na relação anterior, temos

$$I_{\hat{\nu}\hat{\nu}} = \left\{ -\frac{n}{2\nu^2} + \frac{1}{\nu^3} \left[\underset{\sim}{1}' E_* (Y^{-1}) \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' DE_* (Y) D \underset{\sim}{1} \right] - \frac{n^2}{4\nu^2} \right. \\ \left. + \frac{n}{2\nu^3} \left[\underset{\sim}{1}' E_* (Y^{-1}) \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' DE_* (Y) D \underset{\sim}{1} \right] \right. \\ \left. - \frac{1}{4\nu^4} \underset{\sim}{1}' E_* \left(Y^{-1} \underset{\sim}{1} \underset{\sim}{1}' Y^{-1} \right) \underset{\sim}{1} - \frac{1}{2\nu^4} \underset{\sim}{1}' D^2 E_* \left(Y \underset{\sim}{1} \underset{\sim}{1}' Y^{-1} \right) \underset{\sim}{1} \right. \\ \left. - \frac{1}{4\nu^4} \underset{\sim}{1}' D^2 E_* \left(Y \underset{\sim}{1} \underset{\sim}{1}' Y \right) D^2 \underset{\sim}{1} + \frac{1}{\nu^4} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' E_* (Y^{-1}) \underset{\sim}{1} \right. \\ \left. + \frac{1}{\nu^4} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' DE_* (Y) D \underset{\sim}{1} - \frac{1}{\nu^4} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' D \underset{\sim}{1} \right\}_{\theta=\hat{\theta}}. \quad (5.73)$$

Usando a notação dada em (5.57) e a expressão equivalente dada em (4.79), a ex-

pressão dada em (5.73) se transforma em

$$\begin{aligned}
I_{\hat{\nu}\hat{\nu}} &= \left\{ -\frac{n}{2\nu^2} + \frac{1}{\nu^3} \left[\underset{\sim}{1}' M_{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' D M_1 D \underset{\sim}{1} \right] - \frac{n^2}{4\nu^2} \right. \\
&\quad + \frac{n}{2\nu^3} \left[\underset{\sim}{1}' M_{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' D M_1 D \underset{\sim}{1} \right] \\
&\quad - \frac{1}{4\nu^4} \underset{\sim}{1}' \left[M_{-2} - M_{-1}^2 + M_{-1} \underset{\sim}{1} \underset{\sim}{1}' M_{-1} \right] \underset{\sim}{1} \\
&\quad - \frac{1}{2\nu^4} \underset{\sim}{1}' D^2 \left[I - M_1 M_{-1} + M_1 \underset{\sim}{1} \underset{\sim}{1}' M_{-1} \right] \underset{\sim}{1} \\
&\quad - \frac{1}{4\nu^4} \underset{\sim}{1}' D^2 \left[M_2 - M_1^2 + M_1 \underset{\sim}{1} \underset{\sim}{1}' M_1 \right] D^2 \underset{\sim}{1} \\
&\quad + \frac{1}{\nu^4} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' M_{-1} \underset{\sim}{1} + \frac{1}{\nu^4} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' M_1 D^2 \underset{\sim}{1} \\
&\quad \left. - \frac{1}{\nu^4} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' D \underset{\sim}{1} \right\}_{\theta=\hat{\theta}} \\
&= \left\{ -\frac{n}{2\nu^2} - \frac{n^2}{4\nu^2} + \left(\frac{n}{2\nu^3} + \frac{1}{\nu^3} \right) \left[\underset{\sim}{1}' M_{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' D M_1 D \underset{\sim}{1} \right] \right. \\
&\quad - \frac{1}{4\nu^4} \underset{\sim}{1}' \left[M_{-2} - M_{-1}^2 \right] \underset{\sim}{1} - \frac{1}{2\nu^4} \underset{\sim}{1}' D^2 \left[I - M_1 M_{-1} \right] \underset{\sim}{1} \\
&\quad - \frac{1}{4\nu^4} \left[\underset{\sim}{1}' M_{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' M_1 D^2 \underset{\sim}{1} \right]^2 \\
&\quad \left. - \frac{1}{4\nu^4} \underset{\sim}{1}' D^2 \left[M_2 - M_1^2 \right] D^2 \underset{\sim}{1} \right\}_{\theta=\hat{\theta}}.
\end{aligned} \tag{5.74}$$

Mas, ao substituir a notação dada em capítulo anterior em (5.57) e a expressão equivalente dada em (5.64), na equação (5.59), temos:

$$n\nu = \underset{\sim}{1}' M_{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' D M_1 D \underset{\sim}{1}. \tag{5.75}$$

Logo, substituindo (5.75) em (5.74), temos o resultaltado:

$$I_{\hat{\nu}\hat{\nu}} = \frac{n}{2\hat{\nu}^2} - \frac{1}{4\hat{\nu}^4} \underset{\sim}{1}' \left\{ \hat{D}^2 \left[\hat{M}_2 - \hat{M}_1^2 \right] \hat{D}^2 + 2\hat{D} \left[I - \hat{M}_1 \hat{M}_{-1} \right] \hat{D} + \left[\hat{M}_{-2} - \hat{M}_{-1}^2 \right] \right\} \underset{\sim}{1}. \tag{5.76}$$

Cálculo de $I_{\hat{\nu}\hat{\beta}}$

Substituindo (5.36), (5.12) e (5.13) em (5.62), temos que:

$$\begin{aligned}
I_{\hat{\nu}\hat{\beta}} &= -E_* \left\{ -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left(\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' X \underset{\sim}{\beta} + \underset{\sim}{\beta}' X' Y X \underset{\sim}{\beta} \right) \right. \\
&\quad \left. \cdot \left[\underset{\sim}{\beta}' X' Y D_p^{-1} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}}.
\end{aligned}$$

Agora, mediante a notação dada em (5.64), temos que:

$$\begin{aligned}
I_{\nu\hat{p}} &= -E_* \left\{ -\frac{n}{2\nu} + \frac{1}{2\nu^2} \left(\underset{\sim}{1}' Y^{-1} \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' DYD \underset{\sim}{1} \right) \right. \\
&\quad \left. \cdot \left[\underset{\sim}{1}' DYD_p^{-1} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \right] \right\}_{\theta=\hat{\theta}} \\
&= E_* \left\{ \frac{n}{2\nu} \left[\underset{\sim}{1}' DYD_p^{-1} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \right] - \frac{1}{2\nu^2} \underset{\sim}{1}' Y^{-1} \underset{\sim}{1} \underset{\sim}{1}' DYD_p^{-1} \underset{\sim}{1} \right. \\
&\quad + \frac{1}{\nu^2} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' DYD_p^{-1} \underset{\sim}{1} - \frac{1}{2\nu^2} \underset{\sim}{1}' DYD \underset{\sim}{1} \underset{\sim}{1}' DYD_p^{-1} \underset{\sim}{1} \\
&\quad + \frac{1}{2\nu^2} \underset{\sim}{1}' Y^{-1} \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} - \frac{1}{\nu^2} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \\
&\quad \left. + \frac{1}{2\nu^2} \underset{\sim}{1}' DYD \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \right\}_{\theta=\hat{\theta}}.
\end{aligned}$$

Como D , D_p^{-1} , Y e Y^{-1} são matrizes diagonais, podemos comutá-las, e usando esta propriedade na relação anterior temos

$$\begin{aligned}
I_{\nu\hat{p}} &= \left\{ \frac{n}{2\nu} \left[\underset{\sim}{1}' DE_* (YD_p^{-1}) \underset{\sim}{1} - \underset{\sim}{1}' E_* (D_p^{-1}) \underset{\sim}{1} \right] - \frac{1}{2\nu^2} \underset{\sim}{1}' E_* (Y^{-1} \underset{\sim}{1} \underset{\sim}{1}' YD_p^{-1}) D \underset{\sim}{1} \right. \\
&\quad + \frac{1}{\nu^2} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' DE_* (YD_p^{-1}) \underset{\sim}{1} - \frac{1}{2\nu^2} \underset{\sim}{1}' D^2 E_* (Y \underset{\sim}{1} \underset{\sim}{1}' YD_p^{-1}) D \underset{\sim}{1} \\
&\quad + \frac{1}{2\nu^2} \underset{\sim}{1}' E_* (Y^{-1} \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1}) \underset{\sim}{1} - \frac{1}{\nu^2} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' E_* (D_p^{-1}) \underset{\sim}{1} \\
&\quad \left. + \frac{1}{2\nu^2} \underset{\sim}{1}' D^2 E_* (Y \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1}) \underset{\sim}{1} \right\}_{\theta=\hat{\theta}}.
\end{aligned} \tag{5.77}$$

Usando as notações dada em (5.57) e as expressões equivalentes dadas em (4.79), a relação dada em (5.77) transforma-se em

$$\begin{aligned}
I_{\nu\hat{p}} &= \left\{ \frac{n}{2\nu} \left[\underset{\sim}{1}' DN_1 \underset{\sim}{1} - \underset{\sim}{1}' N_0 \underset{\sim}{1} \right] - \frac{1}{2\nu^2} \underset{\sim}{1}' \left[N_0 - M_{-1}N_1 + M_{-1} \underset{\sim}{1} \underset{\sim}{1}' N_1 \right] D \underset{\sim}{1} \right. \\
&\quad + \frac{1}{\nu^2} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' DN_1 \underset{\sim}{1} - \frac{1}{2\nu^2} \underset{\sim}{1}' D^2 \left[N_2 - M_1N_1 + M_1 \underset{\sim}{1} \underset{\sim}{1}' N_1 \right] D \underset{\sim}{1} \\
&\quad + \frac{1}{2\nu^2} \underset{\sim}{1}' \left[N_{-1} - M_{-1}N_0 + M_{-1} \underset{\sim}{1} \underset{\sim}{1}' N_0 \right] \underset{\sim}{1} - \frac{1}{\nu^2} \underset{\sim}{1}' D \underset{\sim}{1} \underset{\sim}{1}' N_0 \underset{\sim}{1} \\
&\quad \left. + \frac{1}{2\nu^2} \underset{\sim}{1}' D^2 \left[N_1 - M_1N_0 + M_1 \underset{\sim}{1} \underset{\sim}{1}' N_0 \right] \underset{\sim}{1} \right\}_{\theta=\hat{\theta}}
\end{aligned}$$

$$\begin{aligned}
&= \left\{ \frac{n}{2\nu} \left[\underset{\sim}{1}' D N_1 \underset{\sim}{1} - \underset{\sim}{1}' N_0 \underset{\sim}{1} \right] - \frac{1}{2\nu^2} \underset{\sim}{1}' [N_0 - M_{-1} N_1] D \underset{\sim}{1} \right. \\
&\quad - \frac{1}{2\nu^2} \left[\underset{\sim}{1}' D^2 M_1 \underset{\sim}{1} - 2 \underset{\sim}{1}' D \underset{\sim}{1} + \underset{\sim}{1}' M_{-1} \underset{\sim}{1} \right] \left[\underset{\sim}{1}' N_1 D \underset{\sim}{1} - \underset{\sim}{1}' N_0 \underset{\sim}{1} \right] \\
&\quad - \frac{1}{2\nu^2} \underset{\sim}{1}' D^2 [N_2 - M_1 N_1] D \underset{\sim}{1} + \frac{1}{2\nu^2} \underset{\sim}{1}' [N_{-1} - M_{-1} N_0] \underset{\sim}{1} \\
&\quad \left. + \frac{1}{2\nu^2} \underset{\sim}{1}' D^2 [N_1 - M_1 N_0] \underset{\sim}{1} \right\}_{\theta=\hat{\theta}}.
\end{aligned} \tag{5.78}$$

Mas, mediante a notação dada em (5.57) e (5.64) em (5.13), temos que

$$\underset{\sim}{1}' D N_1 \underset{\sim}{1} - \underset{\sim}{1}' N_0 \underset{\sim}{1} = 0. \tag{5.79}$$

Usando as relações (5.75) e (5.64) em (5.78), temos:

$$\begin{aligned}
I_{\hat{p}\hat{p}} &= \frac{1}{2\nu^2} \underset{\sim}{1}' \left\{ [\hat{N}_{-1} - \hat{M}_{-1} \hat{N}_0] \underset{\sim}{1} - [\hat{N}_0 - \hat{M}_{-1} \hat{N}_1] \hat{D} \right. \\
&\quad \left. + \hat{D}^2 [\hat{N}_1 - \hat{M}_1 \hat{N}_0] \underset{\sim}{1} - \hat{D}^2 [\hat{N}_2 - \hat{M}_1 \hat{N}_1] \hat{D} \right\} \underset{\sim}{1}.
\end{aligned} \tag{5.80}$$

Cálculo de $I_{\hat{p}\hat{p}}$

Substituindo (5.37) e (5.13) em (5.63), temos:

$$\begin{aligned}
I_{\hat{p}\hat{p}} &= -E_* \left\{ -\underset{\sim}{\beta}' X' Y^2 D_p^{-2} X \underset{\sim}{\beta} + 2 \underset{\sim}{\beta}' X' Y D_p^{-2} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-2} \underset{\sim}{1} \right. \\
&\quad \left. + \left[\underset{\sim}{\beta}' X' Y D_p^{-1} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \right]^2 \right\}_{\theta=\hat{\theta}},
\end{aligned}$$

e mediante a notação dada em (5.64), temos:

$$\begin{aligned}
I_{\hat{p}\hat{p}} &= -E_* \left\{ -\underset{\sim}{1}' D Y^2 D_p^{-2} D \underset{\sim}{1} + 2 \underset{\sim}{1}' D Y D_p^{-2} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-2} \underset{\sim}{1} \right. \\
&\quad \left. + \left[\underset{\sim}{1}' D Y D_p^{-1} \underset{\sim}{1} - \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \right]^2 \right\}_{\theta=\hat{\theta}} \\
&= E_* \left\{ \underset{\sim}{1}' D Y^2 D_p^{-2} D \underset{\sim}{1} - 2 \underset{\sim}{1}' D Y D_p^{-2} \underset{\sim}{1} + \underset{\sim}{1}' D_p^{-2} \underset{\sim}{1} \right. \\
&\quad - \underset{\sim}{1}' D Y D_p^{-1} \underset{\sim}{1} \underset{\sim}{1}' D Y D_p^{-1} \underset{\sim}{1} + 2 \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \underset{\sim}{1}' D Y D_p^{-1} \underset{\sim}{1} \\
&\quad \left. - \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \underset{\sim}{1}' D_p^{-1} \underset{\sim}{1} \right\}_{\theta=\hat{\theta}}.
\end{aligned}$$

Como D , D_p^{-1} , Y e Y^{-1} são matrizes diagonais, então podem comutar-se e

$$I_{\hat{p}\hat{p}} = \left\{ \underset{\sim}{1}' DE_* (Y^2 D_p^{-2}) D \underset{\sim}{1} - 2 \underset{\sim}{1}' DE_* (Y D_p^{-2}) \underset{\sim}{1} + \underset{\sim}{1}' E_* (D_p^{-2}) \underset{\sim}{1} \right. \\ \left. - \underset{\sim}{1}' DE_* (Y D_p^{-1} \underset{\sim}{1}\underset{\sim}{1}' Y D_p^{-1}) D \underset{\sim}{1} + 2 \underset{\sim}{1}' E_* (D_p^{-1} \underset{\sim}{1}\underset{\sim}{1}' Y D_p^{-1}) D \underset{\sim}{1} \right. \\ \left. - \underset{\sim}{1}' E_* (D_p^{-1} \underset{\sim}{1}\underset{\sim}{1}' D_p^{-1}) \underset{\sim}{1} \right\}_{\theta=\hat{\theta}} . \quad (5.81)$$

Usando a notação dada em (5.57) e as expressões equivalentes dadas em (4.79) e (5.81), temos o resultado:

$$I_{\hat{p}\hat{p}} = \left\{ \underset{\sim}{1}' DE_* (Y^2 D_p^{-2}) D \underset{\sim}{1} - 2 \underset{\sim}{1}' DE_* (Y D_p^{-2}) \underset{\sim}{1} + \underset{\sim}{1}' E_* (D_p^{-2}) \underset{\sim}{1} \right. \\ \left. - \underset{\sim}{1}' D \left[E_* (Y^2 D_p^{-2}) - N_1^2 + N_1 \underset{\sim}{1}\underset{\sim}{1}' N_1 \right] D \underset{\sim}{1} \right. \\ \left. + 2 \underset{\sim}{1}' \left[E_* (Y D_p^{-2}) - N_0 N_1 + N_0 \underset{\sim}{1}\underset{\sim}{1}' N_1 \right] D \underset{\sim}{1} \right. \\ \left. - \underset{\sim}{1}' \left[E_* (D_p^{-2}) - N_0^2 + N_0 \underset{\sim}{1}\underset{\sim}{1}' N_0 \right] \underset{\sim}{1} \right\}_{\theta=\hat{\theta}} \\ \left\{ \underset{\sim}{1}' D N_1^2 D \underset{\sim}{1} - 2 \underset{\sim}{1}' N_0 N_1 D \underset{\sim}{1} + \underset{\sim}{1}' N_0^2 \underset{\sim}{1} \right. \\ \left. - \left[\underset{\sim}{1}' D N_1 \underset{\sim}{1} - \underset{\sim}{1}' N_0 \underset{\sim}{1} \right]^2 \right\}_{\theta=\hat{\theta}} . \quad (5.82)$$

Usando a relação (5.75) em (5.82), temos que;

$$I_{\hat{p}\hat{p}} = \underset{\sim}{1}' \hat{D} \hat{N}_1^2 \hat{D} \underset{\sim}{1} - 2 \underset{\sim}{1}' \hat{N}_0 \hat{N}_1 \hat{D} \underset{\sim}{1} + \underset{\sim}{1}' \hat{N}_0^2 \underset{\sim}{1} . \quad (5.83)$$

□

No teorema anterior, se p é conhecido, então a matriz de informação se reduz a

$$I_{pJ} = \begin{bmatrix} I_{\hat{\beta}\hat{\beta}} & I_{\hat{\beta}\hat{\nu}} \\ I_{\hat{\nu}\hat{\beta}} & I_{\hat{\nu}\hat{\nu}} \end{bmatrix}, \quad (5.84)$$

onde

$$I_{\hat{\beta}\hat{\beta}} = X' \left\{ \frac{p}{\hat{\nu}} \left[p \hat{\nu} \hat{N}_1^2 + 2 \hat{D} (\hat{N}_2 - \hat{M}_1 \hat{N}_1) \right] + \frac{1}{\hat{\nu}^2} \left[\hat{\nu} \hat{M}_1 - \hat{D} (\hat{M}_2 - \hat{M}_1^2) \hat{D} \right] \right\} X,$$

$$I_{\hat{\beta}\hat{\beta}} = I'_{\hat{\nu}\hat{\beta}} = \frac{1}{2\hat{\nu}^3} X' \left[\hat{D} (\hat{M}_2 - \hat{M}_1^2) \hat{D} + (I - \hat{M}_1 \hat{M}_{-1}) \right] \hat{D} \underline{1} \\ - \frac{\nu}{2\hat{\nu}^2} X' \left[\hat{D} (\hat{N}_2 - \hat{N}_1 \hat{M}_1) \hat{D} + (2\hat{\nu} \hat{N}_1 + \hat{N}_0 - \hat{N}_1 \hat{M}_{-1}) \right] \underline{1}$$

e

$$I_{\hat{\nu}\hat{\nu}} = \frac{n}{2\hat{\nu}^2} - \frac{1}{4\hat{\nu}^4} \underline{1}' \left\{ \hat{D}^2 [\hat{M}_2 - \hat{M}_1^2] \hat{D}^2 + 2\hat{D} [I - \hat{M}_1 \hat{M}_{-1}] \hat{D} + [\hat{M}_{-2} - \hat{M}_{-1}^2] \right\} \underline{1}. \quad (5.85)$$

Se $p = 0$ então, tem-se a matriz de informação observada para a regressão com dados censurados de Gaussiana Inversa. Os componentes de I_{pf} se reduzem

$$I_{\hat{\beta}\hat{\beta}} = X' \frac{1}{\hat{\nu}^2} \left[\hat{\nu} \hat{M}_1 - \hat{D} (\hat{M}_2 - \hat{M}_1^2) \hat{D} \right] X,$$

$$I_{\hat{\beta}\hat{\nu}} = I'_{\hat{\nu}\hat{\beta}} = \frac{1}{2\hat{\nu}^3} X' \left[\hat{D} (\hat{M}_2 - \hat{M}_1^2) \hat{D} + (I - \hat{M}_1 \hat{M}_{-1}) \right] \hat{D} \underline{1}$$

e

$$I_{\hat{\nu}\hat{\nu}} = \frac{n}{2\hat{\nu}^2} - \frac{1}{4\hat{\nu}^4} \underline{1}' \left\{ \hat{D}^2 [\hat{M}_2 - \hat{M}_1^2] \hat{D}^2 + 2\hat{D} [I - \hat{M}_1 \hat{M}_{-1}] \hat{D} + [\hat{M}_{-2} - \hat{M}_{-1}^2] \right\} \underline{1}. \quad (5.86)$$

Se $p = 1$, tem-se a matriz de informação observada para a regressão da recíproca de uma variável Gaussiana Inversa. Os componentes de I_{pf} se reduzem

$$I_{\hat{\beta}\hat{\beta}} = X' \left\{ \hat{D}^{-2} + \frac{1}{\hat{\nu}^2} \left[\hat{\nu} \hat{M}_1 - \hat{D} (\hat{M}_2 - \hat{M}_1^2) \hat{D} \right] \right\} X,$$

$$I_{\hat{\beta}\hat{\nu}} = I'_{\hat{\nu}\hat{\beta}} = \frac{1}{2\hat{\nu}^3} X' \left[\hat{D} (\hat{M}_2 - \hat{M}_1^2) \hat{D} + (I - \hat{M}_1 \hat{M}_{-1}) \right] \hat{D} \underline{1} - \frac{1}{\hat{\nu}} X' \hat{D}^{-1} \underline{1}$$

e

$$I_{\hat{\nu}\hat{\nu}} = \frac{n}{2\hat{\nu}^2} - \frac{1}{4\hat{\nu}^4} \underline{1}' \left\{ \hat{D}^2 [\hat{M}_2 - \hat{M}_1^2] \hat{D}^2 + 2\hat{D} [I - \hat{M}_1 \hat{M}_{-1}] \hat{D} + [\hat{M}_{-2} - \hat{M}_{-1}^2] \right\} \underline{1}. \quad (5.87)$$

5.5 Adequação do modelo

Lawless (1982) apresenta uma definição geral de resíduos, baseada na função de risco acumulada. Sejam Y_1, \dots, Y_n variáveis aleatórias independentes com função de risco acumulado $H(Y_i | x_{\sim i})$, $i = 1, \dots, n$, respectivamente. Então $S(Y_i | x_{\sim i}) = \exp[-H(Y_i | x_{\sim i})]$, $i = 1, \dots, n$, são variáveis aleatórias i.i.d., uniformemente distribuídas no intervalo $(0, 1)$. Logo, os $H(Y_i | x_{\sim i})$ são variáveis aleatórias exponenciais com parâmetro $\theta = 1$. A partir desta suposição, o resíduo para a observação não censurada y_i se define

$$\hat{e}_i = \hat{H}(y_i | x_{\sim i}), \quad (5.88)$$

onde $\hat{H}(y_i | x_{\sim i})$ utiliza as estimativas de máxima verossimilhança dos parâmetros desconhecidos. Logo, como primeira aproximação, para uma amostra não censurada de tamanho n , os resíduos $\hat{e}_1, \dots, \hat{e}_n$ podem ser tratados como uma amostra aleatória de uma distribuição exponencial padrão.

Para uma observação censurada, Lawless (1982), pag. 282, define o resíduo ajustado

$$\hat{e}_i = \hat{H}(y_i^* | x_{\sim i}) + 1, \quad (5.89)$$

onde y_i^* é a observação censurada.

Portanto, quando o modelo é adequado, o gráfico de $-\log[\hat{S}(\hat{e}_i)]$ versus \hat{e}_i , para $i = 1, \dots, n$, onde $\hat{S}(\hat{e}_i)$ é o estimado Kaplan Meier da função de sobrevivência dos \hat{e}_i , é altamente ajustada por uma reta de coeficiente angular 1.

5.6 Teste de Hipóteses e Estimação por Intervalo para Dados sem Censura

5.6.1 Teste de Hipóteses e Estimação por Intervalos para os Parâmetros da Regressão Gaussiana Inversa

Whitmore (1983) apresentou testes de hipóteses e estimação por intervalo para os parâmetros do modelo de Regressão Linear Inverso para dados da distribuição Gaussiana Inversa. Estes testes e a estimação por intervalo são baseados na propriedade de normalidade assintótica dos estimadores de máxima verossimilhança

$$\begin{pmatrix} \hat{\beta} \\ \hat{\nu} \end{pmatrix} \underset{\sim}{a.} N \left[\begin{pmatrix} \beta \\ \nu \end{pmatrix}, I_{I0}^{-1} \right],$$

onde

$$I_{I0} = \begin{bmatrix} \frac{1}{\nu} X' Y X & 0 \\ 0' & \frac{n}{2\hat{\nu}^2} \end{bmatrix}.$$

Por outro lado,

$$n\hat{\nu} = \left[\underline{1}' Y^{-1} \underline{1} - \underline{1}' X \hat{\beta} \right] = \left(YX \hat{\beta} - \underline{1} \right)' Y^{-1} \left(YX \hat{\beta} - \underline{1} \right)$$

é a soma total de variáveis recíproca. Esta variável tem uma distribuição $\nu\chi_{(n-k)}^2$ (Chhikara e Folks, 1989). Logo,

$$\frac{n\hat{\nu}}{\nu} \sim \chi_{(n-k)}^2. \quad (5.90)$$

Baseado nestas propriedades, Whitmore (1983) construiu a seguinte fórmula para achar os intervalos de confiança para β_j com $(1 - \alpha)\%$ de confiança:

$$IC(\beta_j) : \hat{\beta}_j \mp t \left[nC_{jj} \left(\frac{\hat{\nu}}{n-k} \right) \right],$$

onde $t = T\left(n - k, 1 - \frac{\alpha}{2}\right)$ e C_{jj} é o j -ésimo elemento da diagonal da matriz $C = (X'YX)^{-1}$. Também pode usar-se estes intervalos de confiança para testar a hipótese $H_0 : \beta_j = \beta_{j0}$ versus $H_1 : \beta_j \neq \beta_{j0}$, com um nível de significância α . Assim, se aceita a hipótese H_0 , se o intervalo contém o valor de β_{j0} .

Os testes de hipóteses e estimação de intervalo para o parâmetro $\hat{\nu}$ podem ser feitos de forma similar ao caso da variância de distribuições normais. Assim, para testar $H_0 : \nu = \nu_0$ versus $H_1 : \nu \neq \nu_0$, usa-se a seguinte estatística

$$\frac{n\hat{\nu}}{\nu_0},$$

e pela propriedade dada em (5.90), tem distribuição χ^2 com $(n - k)$ graus de liberdade sob $H_0 : \nu = \nu_0$. Os intervalos de confiança para o parâmetro ν , com $(1 - \alpha)\%$ de confiança, são calculados da seguinte forma:

$$IC(\nu) : \left(\frac{n\hat{\nu}}{\chi_{(n-k, 1-\frac{\alpha}{2})}^2}, \frac{n\hat{\nu}}{\chi_{(n-k, \frac{\alpha}{2})}^2} \right).$$

onde $\chi_{(n-k)}^2$ tem distribuição χ^2 com $(n - k)$ graus de liberdade e $P\left(\chi_{(n-k)}^2 < \chi_{(n-k, 1-\alpha)}^2\right) = 1 - \alpha$.

5.6.2 Teste de Hipóteses e Estimação por Intervalo para os Parâmetros da Regressão da R-IG

Para realizar os testes dos parâmetros da regressão R-IG, se usaram procedimentos baseados na estatística da razão de verossimilhança. Esta estatística é função do log-verossimilhança, que para o caso do modelo de regressão da R-IG, é dada por

$$\mathfrak{S}(\beta, \nu) = \sum_{i=1}^n \log(X_i \beta) - \frac{1}{2} \sum_{i=1}^n \log(2\pi y_i) - \frac{n}{2} \log \nu - \frac{1}{2\nu} \left[\mathbf{1}' Y^{-1} \mathbf{1} - 2 \mathbf{1}' X \beta + \beta' X' Y X \beta \right].$$

i) Testes de hipóteses para os coeficientes de regressão

Considere as hipóteses $H_0 : \beta = \beta_{\sim_0}$ e $H_1 : \beta \neq \beta_{\sim_0}$. Para testar H_0 versus H_1 , usamos a estatística da razão de verossimilhança

$$\Lambda(\beta_{\sim_0}, \tilde{\nu}) = -2 \left\{ \mathfrak{F}(\beta_{\sim_0}, \tilde{\nu}) - \mathfrak{F}(\hat{\beta}, \hat{\nu}) \right\},$$

onde $\tilde{\nu}$ é o estimador de máxima verossimilhança de ν dado β_{\sim_0} . As estatísticas $\hat{\beta}$ e $\hat{\nu}$ são os estimadores de máxima verossimilhança de β e ν , respectivamente. A estatística $\Lambda(\beta_{\sim_0}, \tilde{\nu})$ tem distribuição χ^2 com k graus de liberdade sob $H_0 : \beta = \beta_{\sim_0}$.

O estimador de máxima verossimilhança de ν dado β_{\sim_0} , é dado por

$$\tilde{\nu} = \frac{1}{n} \left(\mathbf{1}' Y^{-1} \mathbf{1} - 2 \mathbf{1}' X \beta_{\sim_0} + \beta_{\sim_0}' X' Y X \beta_{\sim_0} \right).$$

Com esta fórmula pode-se calcular a estimativa de máxima verossimilhança de ν dado β_{\sim_0} .

ii) Teste de hipóteses para o parâmetro ν

Considere as hipóteses $H_0 : \nu = \nu_0$ versus $H_1 : \nu \neq \nu_0$. Neste caso usamos a estatística da razão de verossimilhança

$$\Lambda(\tilde{\beta}, \nu_0) = -2 \left\{ \mathfrak{F}(\tilde{\beta}, \nu_0) - \mathfrak{F}(\hat{\beta}, \hat{\nu}) \right\},$$

onde $\tilde{\beta}$ é o estimador de máxima verossimilhança de β dado ν_0 . As estatísticas $\hat{\beta}$ e $\hat{\nu}$ são os estimadores de máxima verossimilhança de β e ν , respectivamente. A estatística $\Lambda(\tilde{\beta}, \nu_0)$ tem distribuição χ^2 com k graus de liberdade, sob $H_0 : \nu = \nu_0$.

A estimativa de máxima verossimilhança de β dado ν_0 , pode ser obtido resolvendo a seguinte equação

$$X' D^{-1} \mathbf{1} - \frac{1}{\nu_0} \left(X' Y X \tilde{\beta} - X' \mathbf{1} \right) = 0,$$

onde

$$D = \text{diag} \left\{ x_{1\sim}\beta, \dots, x_{n\sim}\beta \right\}.$$

Esta equação pode ser resolvida usando o método de Newton Raphson.

iii) Intervalos de confiança

Os intervalos de confiança para os parâmetros da regressão R-IG estão baseados na propriedade de normalidade assintótica dos estimadores de máxima verossimilhança. Isto é,

$$\begin{pmatrix} \hat{\beta} \\ \hat{\nu} \end{pmatrix} \underset{\sim}{a.} N \left[\begin{pmatrix} \beta \\ \nu \end{pmatrix}, I_{R0}^{-1} \right],$$

onde

$$I_{R0} = \begin{bmatrix} X' \hat{D}^{-2} X + \frac{1}{\hat{\nu}} X' Y X & 0 \\ 0' & \frac{n}{2\hat{\nu}^2} \end{bmatrix}.$$

Logo, os limites com $(1 - \alpha) \%$ de confiança, para o j -ésimo coeficiente de regressão e para o parâmetro de dispersão ν são dados por

$$LC(\beta_j) : \hat{\beta}_j \mp Z_0 \sqrt{C_{jj}}$$

e

$$LC(\nu) : \hat{\nu} \mp Z_0 \sqrt{\frac{2}{n} \hat{\nu}},$$

onde $Z \sim N(0, 1)$, C_{jj} é o j -ésimo elemento da matriz

$$C = \left(X' \hat{D}^{-2} X + \frac{1}{\hat{\nu}} X' Y X \right)^{-1} \text{ e } P(Z < Z_0) = 1 - \frac{\alpha}{2}.$$

5.6.3 Teste de Hipóteses sobre os Parâmetros da Regressão para dados da M-IG

Para realizar os testes de hipóteses para os parâmetros da *M-IG*, consideramos que o log-verossimilhança é

$$\mathfrak{S}(\underset{\sim}{\beta}, \nu, p) = \sum_{i=1}^n \log \left[1 - p + p \frac{X_i \underset{\sim}{\beta} y_i}{\underset{\sim}{\nu}} \right] + \frac{1}{2} \sum_{i=1}^n \log (2\pi y_i^3) + \frac{n}{2} \log \nu - \sum_{i=1}^n \frac{\left(X_i \underset{\sim}{\beta} y_i - 1 \right)^2}{2\nu y_i}.$$

Então, apresenta-se os seguintes:

i) Teste de hipóteses para $\underset{\sim}{\beta}$

Considere as hipóteses $H_0 : \underset{\sim}{\beta} = \underset{\sim}{\beta}_0$ e $H_1 : \underset{\sim}{\beta} \neq \underset{\sim}{\beta}_0$. Para testar estas hipóteses, usa-se a seguinte estatística

$$\Lambda \left(\underset{\sim}{\beta}, \tilde{\nu}, \tilde{p} \right) = -2 \left\{ \mathfrak{S} \left(\underset{\sim}{\beta}, \tilde{\nu}, \tilde{p} \right) - \mathfrak{S} \left(\hat{\underset{\sim}{\beta}}, \hat{\nu}, \hat{p} \right) \right\},$$

onde $\tilde{\nu}$ e \tilde{p} são os estimadores de máxima verossimilhança de ν e p dado $\underset{\sim}{\beta}_0$, respectivamente. As estatísticas $\hat{\underset{\sim}{\beta}}$, $\hat{\nu}$ e \hat{p} são os estimadores de máxima verossimilhança de $\underset{\sim}{\beta}$, ν e p , respectivamente. A estatística $\Lambda \left(\underset{\sim}{\beta}, \tilde{\nu}, \tilde{p} \right)$ tem distribuição assintótica χ^2 com k graus de liberdade sob $H_0 : \underset{\sim}{\beta} = \underset{\sim}{\beta}_0$.

ii) Teste de hipóteses para ν

Considere a hipóteses $H_0 : \nu = \nu_0$ versus $H_1 : \nu \neq \nu_0$. Para testar estas hipóteses usa-se a seguinte estatística

$$\Lambda \left(\tilde{\underset{\sim}{\beta}}, \nu_0, \tilde{p} \right) = -2 \left\{ \mathfrak{S} \left(\tilde{\underset{\sim}{\beta}}, \nu_0, \tilde{p} \right) - \mathfrak{S} \left(\hat{\underset{\sim}{\beta}}, \hat{\nu}, \hat{p} \right) \right\},$$

onde $\tilde{\underset{\sim}{\beta}}$ e \tilde{p} são os estimadores de máxima verossimilhança de $\underset{\sim}{\beta}$ e p dado ν_0 , respectivamente. As estatísticas $\hat{\underset{\sim}{\beta}}$, $\hat{\nu}$ e \hat{p} são os estimadores de máxima verossimilhança de

β , ν e p , respectivamente. A estatística $\Lambda(\tilde{\beta}, \nu_0, \tilde{p})$ tem distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : \nu = \nu_0$.

iii) Teste de hipóteses para p

Considere as hipóteses $H_0 : p = p_0$ versus $H_1 : p \neq p_0$. Para testar estas hipóteses, usa-se a seguinte estatística

$$\Lambda(\tilde{\beta}, \tilde{\nu}, p_0) = -2 \left\{ \mathfrak{S}(\tilde{\beta}, \tilde{\nu}, p_0) - \mathfrak{S}(\hat{\beta}, \hat{\nu}, \hat{p}) \right\},$$

onde $\tilde{\beta}$ e $\tilde{\nu}$ são os estimadores de máxima verossimilhança de β e ν dado p_0 , respectivamente. As estatísticas $\hat{\beta}$, $\hat{\nu}$ e \hat{p} são os estimadores de máxima verossimilhança de β , ν e p , respectivamente. A estatística $\Lambda(\tilde{\beta}, \tilde{\nu}, p_0)$ tem distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : p = p_0$.

iv) Intervalos de confiança

Para encontrar os intervalos de confiança, usamos as propriedades assintóticas dos estimadores de máxima verossimilhança. Assim

$$\begin{pmatrix} \hat{\beta} \\ \hat{\nu} \\ \hat{p} \end{pmatrix} \underset{\sim}{\overset{a.}{\sim}} N \left[\begin{pmatrix} \beta \\ \nu \\ p \end{pmatrix}, I_0^{-1} \right]$$

onde I_0 é a matriz de informação dada em (5.24).

Com estas propriedades, os limites para β_j , ν e p com $(1 - \alpha)\%$ de confiança são dadas por

$$LC(\beta_j) : \hat{\beta}_j \mp Z_0 \sqrt{C_{jj}},$$

para $j = 1, \dots, k$,

$$LC(\nu) : \hat{\nu} \mp Z_0 \sqrt{C_{(k+1)(k+1)}},$$

$$LC(p) : \hat{p} \mp Z_0 \sqrt{C_{(k+2)(k+2)}},$$

onde $Z \sim N(0, 1)$, C_{ii} é i -ésimo elemento da diagonal de I_0^{-1} e $P(Z \leq Z_0) = 1 - \frac{\alpha}{2}$.

5.7 Testes de Hipóteses e Estimação por Intervalo com Dados Censurados

Nesta seção, consideramos testes de hipóteses aproximados, baseados na estatística da razão de verossimilhança para os parâmetros da $M - IG$ e seus casos particulares (IG e $R - IG$). A justificativa para considerar estes testes foi dado na Seção 4.3. Também, nesta seção não se considera testes de hipóteses para a fronteira do espaço paramétrico de p , já que as condições necessárias de regularidade não são satisfeitas para testes baseados na estatística da razão de verossimilhança.

Semelhante ao caso de dados sem censura, os intervalos de confiança para os parâmetros da $M - IG$ e seus casos particulares são baseados na propriedade assintótica dos estimadores de máxima verossimilhança. Outros procedimentos para obtenção de intervalos de confiança aproximados, como aqueles baseados na estatística da razão de verossimilhança (corregido ou não, Doganaksoy e Schmee, 1993), não serão abordados neste trabalho, podendo vir ser um tema de pesquisa.

5.7.1 Teste de Hipóteses e Estimação por Intervalo para os Parâmetros da Regressão IG

Para realizar os testes de hipóteses para os parâmetros da IG se considera a seguinte função do log-verossimilhança

$$\mathfrak{S}(\beta, \nu) = \sum_{i=1}^r \log [f_s(y_i, \beta, \nu)] + \sum_{i=r+1}^n \log [S(a_i, \beta, \nu)],$$

onde

$$f_6(y_i, \beta, \nu) = (2\pi y_i^3 \nu)^{-\frac{1}{2}} \exp \left[-\frac{\left(\frac{x \beta y_i}{\tilde{y}_i} - 1 \right)^2}{2\nu y_i} \right]$$

e

$$S(a_i, \mu, \nu) = \int_{a_i}^{\infty} g(y_i, \beta, \nu) dy_i.$$

i) Teste de hipóteses para β

Considere as hipóteses $H_0: \beta = \beta_0$ e $H_1: \beta \neq \beta_0$. Para testar estas hipóteses ussa-se a seguinte estatística

$$\Lambda(\beta, \tilde{\nu}) = -2 \left\{ \mathfrak{S}(\beta, \tilde{\nu}, \tilde{p}) - \mathfrak{S}(\beta, \hat{\nu}) \right\},$$

onde $\tilde{\nu}$ e \tilde{p} são os estimadores de máxima verossimilhança dos parâmetros ν e p dado β , respectivamente. As estatística $\hat{\beta}$ e $\hat{\nu}$ são os estimadores de máxima verossimilhança de β e ν , respectivamente. A estatística $\Lambda(\beta, \tilde{\nu})$ tem distribuição assintótica χ^2 com k graus de liberdade sob $H_0: \beta = \beta_0$.

A estimativa de ν dado β pode ser achada mediante o seguinte procedimento:

I. Escolher a estimativa inicial de ν , a qual, pode ser a estimativa se máxima verossimilhança obtida com os dados não censurados, calculados com as fórmulas dadas em (5.19).

II. Calcular $E(Y_i^{-1} | Y_i > a_i)$ e $E(Y_i | Y_i > a_i)$, para $i = r + 1, \dots, n$, de acordo com as fórmulas dadas em (1.18).

III. Substituir os valores no lado direito da equação

$$\tilde{\nu} = \frac{1}{n} \left[\mathbf{1}' E_*(Y^{-1}) \mathbf{1} - 2 \mathbf{1}' X \beta_0 + \beta_0' X' E_*(Y) X \beta_0 \right].$$

IV. Repetir as etapas II e III até obter a estimativa de ν com a precisão desejada.

ii) Teste de hipóteses para ν

Considere as hipóteses $H_0 : \nu = \nu_0$ e $H_1 : \nu \neq \nu_0$. Para testar estas hipóteses, usa-se a seguinte estatística

$$\Lambda \left(\tilde{\beta}_{\sim}, \nu_0 \right) = -2 \left\{ \mathfrak{S} \left(\tilde{\beta}_{\sim}, \nu_0 \right) - \mathfrak{S} \left(\hat{\beta}_{\sim}, \hat{\nu}_{\sim} \right) \right\}$$

onde $\tilde{\beta}_{\sim}$ é o estimador de máxima verossimilhança do parâmetro β dado ν_0 . As estatísticas $\hat{\beta}_{\sim}$ e $\hat{\nu}_{\sim}$ são os estimadores de máxima verossimilhança de β e ν , respectivamente. A estatística $\Lambda \left(\tilde{\beta}_{\sim}, \nu_0 \right)$ tem distribuição assintótica χ^2 com 1 grau de liberdade $H_0 : \nu = \nu_0$.

A estimativa de β dado ν_0 pode ser achada mediante o seguinte procedimento:

I. Escolher a estimativa inicial de β , a qual, pode ser a estimativa de máxima verossimilhança calculado com os dados não censurados, achados com as fórmulas dadas em (5.18).

II. Calcular $E(Y_i | Y_i > a_i)$, para $i = r + 1, \dots, n$, de acordo com as fórmulas dadas em (1.18).

III. Substituir os valores no lado direito da equação

$$\tilde{\beta}_{\sim} = [X' E_*(Y) X]^{-1} X' \underline{1}.$$

IV. Repetir as etapas II e III até obter a estimativa de μ com a precisão desejada.

iii) Intervalos de confiança para os parâmetros da IG

Para encontrar os intervalos de confiança aproximados para os parâmetros da IG, usamos a propriedade de normalidade assintótica dos estimadores de máxima verossimilhança. Assim,

$$\begin{pmatrix} \hat{\beta}_{\sim} \\ \hat{\nu}_{\sim} \end{pmatrix} \underset{\sim}{\sim} N \left[\begin{pmatrix} \beta \\ \nu \end{pmatrix}, I_{I_0}^{-1} \right],$$

onde

$$I_{I0} = \begin{bmatrix} I_{\hat{\beta}\hat{\beta}} & I_{\hat{\beta}\hat{\nu}} \\ I_{\hat{\nu}\hat{\beta}} & I_{\hat{\nu}\hat{\nu}} \end{bmatrix}$$

e os componentes da matriz I_{I0} estão dados em (5.86). Com esta propriedade, os intervalos aproximados para β_j (j -ésimo coeficiente de regressão, $j = 1, \dots, k$) e ν , com $(1 - \alpha)$ % de confiança são dados por

$$IC(\beta_j) : \hat{\beta}_j \mp Z_0 \sqrt{C_{jj}}$$

$$IC(\nu) : \hat{\nu} \mp Z_0 \sqrt{C_{(k+1)(k+1)}}$$

onde $Z \sim N(0, 1)$, C_{ii} é o i -ésimo elemento da diagonal da matriz I_{I0}^{-1} e $P(Z < Z_0) = 1 - \frac{\alpha}{2}$.

5.7.2 Testes de Hipóteses para os Parâmetros da R-IG

Para realizar os testes de hipóteses dos parâmetros da R-IG, se considera a seguinte função do log-verossimilhança

$$\mathfrak{S}(\beta, \nu) = \sum_{i=1}^r \log [g(y_i, \beta, \nu)] + \sum_{i=r+1}^n \log [S(a_i, \beta, \nu)]$$

onde

$$g(y_i, \beta, \nu) = \frac{x_{i\beta} (2\pi y_i \nu)^{-\frac{1}{2}} \exp \left[-\frac{(x_{i\beta} y_i - 1)^2}{2\nu y_i} \right]}{2\nu y_i}$$

e

$$S(a_i, \beta, \nu) = \int_{a_i}^{\infty} g(y_i, \beta, \nu) dy_i$$

i) Teste de hipóteses para β

Considere as hipóteses $H_0 : \beta_{\sim} = \beta_{\sim_0}$ e $H_1 : \beta_{\sim} \neq \beta_{\sim_0}$. Para testar estas hipóteses, usa-se a seguinte estatística

$$\Lambda \left(\beta_{\sim}, \tilde{\nu} \right) = -2 \left\{ \mathfrak{F} \left(\beta_{\sim}, \tilde{\nu} \right) - \mathfrak{F} \left(\hat{\beta}_{\sim}, \hat{\nu} \right) \right\},$$

onde $\tilde{\nu}$ é o estimador de máxima verossimilhança do parâmetro ν dado β_{\sim} . As estatísticas $\hat{\beta}_{\sim}$ e $\hat{\nu}$ são os estimadores de máxima verossimilhança de $\hat{\beta}_{\sim}$ e ν , respectivamente. A estatística $\Lambda \left(\beta_{\sim}, \tilde{\nu} \right)$ tem distribuição assintótica χ^2 com k graus de liberdade sob $H_0 : \beta_{\sim} = \beta_{\sim_0}$.

A estimativa de ν dado β_{\sim} pode ser achado mediante o seguinte procedimento:

I. Escolher a estimativa inicial de ν , a qual, pode ser a estimativa de máxima verossimilhança obtidas com os dados não censurados, calculados com as fórmulas dadas em (5.20).

II. Calcular $E(Y_i | Y_i > a_i)$ e $E(Y_i^{-1} | Y_i > a_i)$, para $i = r + 1, \dots, n$, de acordo com as fórmulas dadas em (2.10).

III. Substituir os valores no lado direito da equação

$$\tilde{\nu} = \frac{1}{n} \left[\mathbf{1}'_{\sim} E_{\star} (Y^{-1}) \mathbf{1}_{\sim} - 2 \mathbf{1}'_{\sim} X \beta_{\sim_0} + \beta'_{\sim_0} X' E_{\star} (Y) X \beta_{\sim_0} \right].$$

IV. Repetir as etapas II e III até obter a estimativa de ν com a precisão desejada.

ii) Teste de hipóteses para ν

Considere as hipóteses $H_0 : \nu = \nu_0$ e $H_1 : \nu \neq \nu_0$. Para testar estas hipóteses, usa-se a seguinte estatística

$$\Lambda \left(\tilde{\beta}_{\sim}, \nu_0 \right) = -2 \left\{ \mathfrak{F} \left(\tilde{\beta}_{\sim}, \nu_0 \right) - \mathfrak{F} \left(\hat{\beta}_{\sim}, \hat{\nu} \right) \right\},$$

onde $\tilde{\beta}_{\sim}$ é o estimador de máxima verossimilhança do parâmetro β_{\sim} dado ν_0 . As estatísticas $\hat{\beta}_{\sim}$ e $\hat{\nu}$ são os estimadores de máxima verossimilhança de $\hat{\beta}_{\sim}$ e ν , respectiva-

mente. A estatística $\Lambda(\tilde{\beta}, \nu_0)$ tem distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0: \nu = \nu_0$.

A estimativa de β dado ν_0 se obtém resolvendo a seguinte equação

$$X'D^{-1} \underset{\sim}{1} - \frac{1}{\nu_0} [X'E_*(Y)X \underset{\sim}{\beta} - X' \underset{\sim}{1}] = 0,$$

onde

$$D = \text{diag} \left\{ \underset{\sim}{x}_1 \beta, \dots, \underset{\sim}{x}_n \beta \right\}$$

Uma solução aproximada desta equação pode ser obtida com o seguinte procedimento.

I. Escolher a estimativa inicial de β que pode ser a estimativa de máxima verossimilhança calculadas com os dados não censurados, achadas com as fórmulas dadas em (5.20).

II. Calcular $E(Y_i|Y_i > a_i)$, para $i = r + 1, \dots, n$, de acordo com as fórmulas dadas em (2.10).

III. Substituir os valores no lado direito da equação

$$\tilde{\beta}^{(i+1)} = \tilde{\beta}^{(i)} + I_{\tilde{\beta}}^{-1} U(\tilde{\beta}^{(i)}),$$

onde

$$I_{\tilde{\beta}} = X'D^{-2}X - \frac{1}{\nu_0} [X'E_*(Y)X],$$

e

$$U(\tilde{\beta}) = X'D^{-1} \underset{\sim}{1} - \frac{1}{\nu_0} [X'E_*(Y)X \underset{\sim}{\beta} - X' \underset{\sim}{1}].$$

IV. Repetir as etapas II e III até obter a estimativa de μ com a precisão desejada.

iii) Intervalos de confiança para os parâmetros da R-IG

Para encontrar os intervalos de confiança aproximados para os parâmetros da R-IG, usamos a propriedade de normalidade assintótica dos estimadores de máxima

verossimilhança. Assim

$$\begin{pmatrix} \hat{\beta} \\ \hat{\nu} \end{pmatrix} \underset{\sim}{\sim} N \left[\begin{pmatrix} \beta \\ \nu \end{pmatrix}, I_{R0}^{-1} \right],$$

onde

$$I_{R0} = \begin{bmatrix} I_{\hat{\beta}\hat{\beta}} & I_{\hat{\beta}\hat{\nu}} \\ I_{\hat{\beta}\nu} & I_{\hat{\nu}\hat{\nu}} \end{bmatrix},$$

e os componentes da matriz I_{R0} são dadas em (5.87). Com esta propriedade, os intervalos aproximados para β_j (j -ésimo coeficiente de regressão, $j = 1, \dots, k$) e ν , com $(1 - \alpha)\%$ de confiança são dados por

$$IC(\beta_j) : \hat{\beta}_j \mp Z_0 \sqrt{C_{jj}}$$

e

$$IC(\nu) : \hat{\nu} \mp Z_0 \sqrt{C_{(k+1)(k+1)}}$$

onde $Z \sim N(0, 1)$, C_{ii} é o i -ésimo elemento da diagonal da matriz I_{R0}^{-1} e $P(Z < Z_0) = 1 - \frac{\alpha}{2}$.

5.7.3 Teste de Hipóteses e Estimação por Intervalos sobre os Parâmetros da M-IG

Para realizar os testes de hipóteses dos parâmetros da R-IG, considera-se a seguinte função do log-verossimilhança

$$\mathfrak{S}(\beta, \nu, p) = \sum_{i=1}^r \log \left[g(y_i; \beta, \nu, p) \right] + \sum_{i=r+1}^n \log \left[S(a_i; \beta, \nu, p) \right],$$

onde

$$g(y_i; \beta, \nu) = \left[1 - p + p \frac{x_i \beta y_i}{\nu} \right] (2\pi y_i^3 \nu)^{-\frac{1}{2}} \exp \left[-\frac{\left(\frac{x_i \beta y_i - 1}{2\nu y_i} \right)^2}{2\nu y_i} \right]$$

e

$$S(a_i, \beta, \nu) = \int_{a_i}^{\infty} g(y_i; \beta, \nu, p) dy_i.$$

i) Teste de hipóteses para β

Considere as hipóteses $H_0: \beta = \beta_0$ e $H_1: \beta \neq \beta_0$. Para testar estas hipóteses, usa-se a seguinte estatística

$$\Lambda(\beta_0, \tilde{\nu}, \tilde{p}) = -2 \left\{ \mathfrak{S}(\beta_0, \tilde{\nu}, \tilde{p}) - \mathfrak{S}(\hat{\beta}, \hat{\nu}, \hat{p}) \right\},$$

onde $\tilde{\nu}$ e \tilde{p} são os estimadores de máxima verossimilhança dos parâmetros ν e p dado β_0 , respectivamente. As estatísticas $\hat{\beta}$, $\hat{\nu}$ e \hat{p} são os estimadores de máxima verossimilhança de β , ν e p , respectivamente. A estatística $\Lambda(\beta_0, \tilde{\nu}, \tilde{p})$ tem distribuição assintótica χ^2 com k graus de liberdade $H_0: \beta = \beta_0$.

As estimativas de máxima verossimilhança de ν e p dado β_0 , encontra-se resolvendo o seguinte sistema de equações

$$-\frac{n}{2\nu} + \frac{1}{2\nu^2} \left[\mathbf{1}' E_* (Y^{-1}) \mathbf{1} - 2 \mathbf{1}' X \beta_0 + \beta_0' X' E_* (Y) X \beta_0 \right] = 0$$

e

$$\beta_0' X' E_* (Y D_p^{-1}) \mathbf{1} - \mathbf{1}' E_* (D_p^{-1}) \mathbf{1} = 0.$$

Para resolver este sistema, pode empregar-se o método de Newton Raphson com a seguinte matriz de informação observada

$$I_{\beta_0} = \begin{bmatrix} I_{\tilde{\nu}\tilde{\nu}} & I_{\tilde{\nu}\tilde{p}} \\ I_{\tilde{p}\tilde{\nu}} & I_{\tilde{p}\tilde{p}} \end{bmatrix},$$

onde

$$I_{\tilde{\nu}\tilde{\nu}} = \frac{n}{2\tilde{\nu}^2} - \frac{1}{4\tilde{\nu}^4} \mathbf{1}' \left\{ D^2 [\tilde{M}_2 - \tilde{M}_1^2] D^2 + 2D [I - \tilde{M}_1 \tilde{M}_{-1}] D + [\tilde{M}_{-2} - \tilde{M}_{-1}^2] \right\} \mathbf{1},$$

$$I_{\hat{\nu}\hat{p}} = I'_{\hat{p}\hat{\nu}} = \frac{1}{2\hat{\nu}^2} \underset{\sim}{1}' \left\{ \left[\tilde{N}_{-1} - \tilde{M}_{-1}\tilde{N}_0 \right] \underset{\sim}{1} - \left[\tilde{N}_0 - \tilde{M}_{-1}\tilde{N}_1 \right] D + D^2 \left[\tilde{N}_1 - \tilde{M}_1\tilde{N}_0 \right] \underset{\sim}{1} - D^2 \left[\tilde{N}_2 - \tilde{M}_1\tilde{N}_1 \right] D \right\} \underset{\sim}{1}$$

e

$$I_{\hat{p}\hat{p}} = \underset{\sim}{1}' D \tilde{N}_1^2 D \underset{\sim}{1} - 2 \underset{\sim}{1}' \tilde{N}_0 \tilde{N}_1 D \underset{\sim}{1} + \underset{\sim}{1}' \tilde{N}_0^2 \underset{\sim}{1}.$$

Os componentes $E \left[Y_i^s / \left(1 - p_0 + p_0 \underset{\sim}{x}_i \underset{\sim}{\beta}_0 Y_i \right) \mid Y_i > a_i \right]$ e $E(Y_i^s \mid Y_i > a_i)$ de \tilde{N}_s e \tilde{M}_s , respectivamente, são calculados com as estimativas de ν e p dado $\underset{\sim}{\beta}$, mediante as expressões dada nas relações (3.23) e (3.24), respectivamente, para todo $s \in Z$,

$$D = \text{diag} \left\{ \underset{\sim}{x}_1 \underset{\sim}{\beta}_0, \dots, \underset{\sim}{x}_n \underset{\sim}{\beta}_0 \right\}$$

$$\text{e } \mu_{i0}^{-1} = \underset{\sim}{x}_i \underset{\sim}{\beta}_0.$$

ii) Teste de hipóteses para ν

Considere as hipóteses $H_0 : \nu = \nu_0$ e $H_1 : \nu \neq \nu_0$. Para testar estas hipóteses, usa-se a seguinte estatística

$$\Lambda \left(\underset{\sim}{\hat{\beta}}, \nu_0, \underset{\sim}{\hat{p}} \right) = -2 \left\{ \mathfrak{S} \left(\underset{\sim}{\hat{\beta}}, \nu_0, \underset{\sim}{\hat{p}} \right) - \mathfrak{S} \left(\hat{\beta}, \hat{\nu}, \hat{p} \right) \right\},$$

onde $\underset{\sim}{\hat{\beta}}$ e $\underset{\sim}{\hat{p}}$ são os estimadores de máxima verossimilhança do parâmetro $\underset{\sim}{\beta}$ e p dado ν_0 , respectivamente. As estatísticas $\hat{\beta}$, $\hat{\nu}$ e \hat{p} são os estimadores de máxima verossimilhança de $\underset{\sim}{\beta}$, ν e p , respectivamente. A estatística $\Lambda \left(\underset{\sim}{\hat{\beta}}, \nu_0, \underset{\sim}{\hat{p}} \right)$ tem distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : \nu = \nu_0$.

As estimativas de máxima verossimilhança de $\underset{\sim}{\beta}$ e p dado ν_0 , pode ser obtidas resolvendo o seguinte sistema de equações

$$p X' E_* \left(Y D_p^{-1} \right) \underset{\sim}{1} - \frac{1}{\nu_0} \left[X' E_* (Y) X \underset{\sim}{\beta} - X' \underset{\sim}{1} \right] = 0$$

e

$$\tilde{\beta}' X' E_* (Y D_p^{-1}) \underline{1} - \underline{1}' E_* (D_p^{-1}) \underline{1} = 0.$$

Este sistema pode ser resolvido mediante o método de Newton Raphson com a seguinte matriz de informação observada

$$I_{\nu_0} = \begin{bmatrix} I_{\tilde{\beta}\tilde{\beta}} & I_{\tilde{\beta}\tilde{p}} \\ I_{\tilde{p}\tilde{\beta}} & I_{\tilde{p}\tilde{p}} \end{bmatrix},$$

onde

$$I_{\tilde{\beta}\tilde{\beta}} = X' \left\{ \frac{\tilde{P}}{\nu_0} [\tilde{p}\nu_0\tilde{N}_1^2 + 2\tilde{D}(\tilde{N}_2 - \tilde{M}_1\tilde{N}_1)] + \frac{1}{\nu_0^2} [\nu_0\tilde{M}_1 - \tilde{D}(\tilde{M}_2 - \tilde{M}_1^2)\tilde{D}] \right\} X,$$

$$I_{\tilde{\beta}\tilde{p}} = I'_{\tilde{p}\tilde{\beta}} = \frac{1}{\nu_0} X' \left\{ \tilde{D} [\tilde{N}_2 - \tilde{M}_1\tilde{N}_1] - [\tilde{N}_1 - \tilde{M}_1\tilde{N}_0] \right\} \hat{D} \underline{1} \\ + X' \tilde{N}_1 [\tilde{p}\tilde{N}_1 - \tilde{p}\tilde{N}_0 - I] \underline{1}$$

e

$$I_{\tilde{p}\tilde{p}} = \underline{1}' \tilde{D} \tilde{N}_1^2 \tilde{D} \underline{1} - 2 \underline{1}' \tilde{N}_0 \tilde{N}_1 \tilde{D} \underline{1} + \underline{1}' \tilde{N}_0^2 \underline{1}.$$

Aqui, $\mu_i^{-1} = \tilde{x}_i \tilde{\beta}$, os componentes $E [Y_i^s / (1 - p + p\mu_i^{-1}Y_i) | Y_i > a_i]$ e $E (Y_i^s | Y_i > a_i)$ de \tilde{N}_s e \tilde{M}_s , respectivamente, são calculados com as estimativas de $\tilde{\beta}$ e p dado ν_0 , mediante as expressões dadas na relações (3.23) e (3.24), respectivamente, para todo $s \in Z$, e

$$\hat{D} = \text{diag} \left\{ \tilde{x}_1 \tilde{\beta}, \dots, \tilde{x}_n \tilde{\beta} \right\}.$$

iii) Teste de hipóteses para p

Considere as hipóteses $H_0 : p = p_0$ e $H_1 : p \neq p_0$. Para testar estas hipóteses, usa-se a seguinte estatística

$$\Lambda(\tilde{\beta}, \tilde{\nu}, p_0) = -2 \left\{ \mathfrak{F}(\tilde{\beta}, \tilde{\nu}, p_0) - \mathfrak{F}(\hat{\beta}, \hat{\nu}, \hat{p}) \right\}$$

onde $\tilde{\beta}$ e $\tilde{\nu}$ são os estimadores de máxima verossimilhança do parâmetro β e ν dado p_0 respectivamente. As estatísticas $\hat{\beta}$, $\hat{\nu}$ e \hat{p} são os estimadores de máxima verossimilhança de β , ν e p , respectivamente. A estatística $\Lambda(\tilde{\beta}, \tilde{\nu}, p_0)$ tem distribuição assintótica χ^2 com 1 grau de liberdade sob $H_0 : p = p_0$.

iv) Intervalo de confiança

Para encontrar os intervalos de confiança aproximados, usamos as propriedades assintóticas dos estimadores de máxima verossimilhança. Assim

$$\begin{pmatrix} \hat{\beta} \\ \hat{\nu} \\ \hat{p} \end{pmatrix} \begin{matrix} a. \\ \sim \end{matrix} N \left[\begin{pmatrix} \beta \\ \nu \\ p \end{pmatrix}, I_0^{-1} \right],$$

onde I_0 é a matriz de informação observada dado em (4.69).

Com estas propriedades, os limites aproximados para β_j (j -ésimo coeficiente de regressão, $j = 1, \dots, k$), ν e p com $(1 - \alpha)\%$ de confiança são dados por

$$LC(\beta_j) : \hat{\beta}_j \mp Z_0 \sqrt{C_{jj}},$$

$$LC(\nu) : \hat{\nu} \mp Z_0 \sqrt{C_{(k+1)(k+1)}}$$

e

$$LC(p) : \hat{p} \mp Z_0 \sqrt{C_{(k+2)(k+2)}},$$

onde $Z \sim N(0, 1)$, C_{ii} é i -ésimo elemento da diagonal de I_0^{-1} e $P(Z \leq Z_0) = 1 - \frac{\alpha}{2}$.

5.8 Exemplos

5.8.1 Exemplo 1-Teste de Vida de Isolantes de Fluido Elétrico

Nelson(1980) apresenta alguns dados de tempos para o colapso de um tipo de isolante de fluido elétrico, submetido a diferentes níveis de voltagem. Os dados se

mostra na Tabela 5.1. Nesta tabela não se considera os dados a um nível de voltagem de 26 Kv, utilizadas no livro.

Tabela 5.1: Dados de tempos de vida de isolantes de fluido elétrico
nível de Voltagem n_i tempos de colapso

nível de Voltagem	n_i	tempos de colapso
28	5	68.85, 426.07, 110.29, 108.29, 1067.6
30	11	17.05, 22.66, 21.02, 175.88, 139.07, 144.12, 20.46, 43.40, 194.90, 47.30, 7.74
32	15	0.40, 82.85, 9.88, 89.29, 215.10, 2.75, 0.79, 15.93, 3.91, 0.27, 0.69, 100.58, 27.80, 13.95, 53.24
34	19	0.96, 4.15, 0.19, 0.78, 8.01, 31.75, 7.35, 6.50, 8.27, 33.91, 32.52, 3.16, 4.85, 2.78, 4.67, 1.31, 12.06, 36.71, 72.89
36	15	1.97, 0.59, 2.58, 1.69, 2.71, 25.50, 0.35, 0.99, 3.99, 3.67, 2.07, 0.96, 5.35, 2.90, 13.77,
38	8	0.47, 0.73, 1.40, 0.74, 0.39, 1.13, 0.09, 2.38

O modelo sugerido pelos engenheiros, para cada nível de voltagem, é o modelo Weibull. Aqui vamos assumir o modelo de regressão para a Mistura de distribuições Gaussiana Inversa com sua Recíproca Complementar. Isto é que os tempos de vida têm uma distribuição dada por (3.4), com parâmetro de escala ν e de forma μ_i , tal que

$$\mu_i = \beta_0 + \beta_1 x_i,$$

para $i = 1, \dots, 6$ e onde x_i = voltagem em Kv.

Para obter as estimativas, vamos primeiro investigar para que valor de p no intervalo $[0, 1]$ obtém-se o máximo valor da log-verossimilhança. As estimativas de β_0 , β_1 e ν para valores diferentes de $p \in [0, 1]$, são obtidas pelo programa 2, feito em S-plus para p fixo e dado no Apêndice. Com este programa obtemos a Tabela 5.2

Tabela 5.2: Estimativas de máxima veross. para vários valores de p

β_0	β_1	ν	p	log-verossim
-0.45176	0.01585	0.62596	0.0	-305.78018
-0.50546	0.01889	0.69484	0.1	-301.20804
-0.61450	0.02310	0.77640	0.2	-296.23860
-0.71818	0.02699	0.86376	0.3	-292.36829
-0.82504	0.03095	0.96464	0.4	-289.35729
-0.94194	0.03525	1.08913	0.5	-287.01099
-1.07700	0.04020	1.25330	0.6	-285.20760
-1.24402	0.04632	1.49200	0.7	-283.87412
-1.47372	0.05474	1.88758	0.8	-282.96378
-1.86015	0.06895	2.74481	0.9	-282.40792
-2.96830	0.10990	6.61030	1.0	-281.78010

Destes resultados, podemos observar que o valor de máxima verossimilhança obtém-se para $p = 1$, na fronteira do espaço paramétrico de p . Logo, como na fronteira do espaço paramétrico, não são satisfeitas as condições de regularidade da teoria de verossimilhança, não podemos fazer testes de hipóteses para determinar se a Recíproca Complementar é o melhor modelo que ajusta a este conjunto de dados em comparação com outros modelos da M-IG. Então, para avaliar a qualidade de ajuste na fronteira, fazemos a análise de resíduos, com as fórmulas dadas na Seção 5.5. Assim, na Figura 5.1, o gráfico (a) representa os resíduos considerando um modelo de regressão Gaussiana Inversa, e o gráfico (b) representa os resíduos considerando um modelo de regressão da Recíproca Complementar.

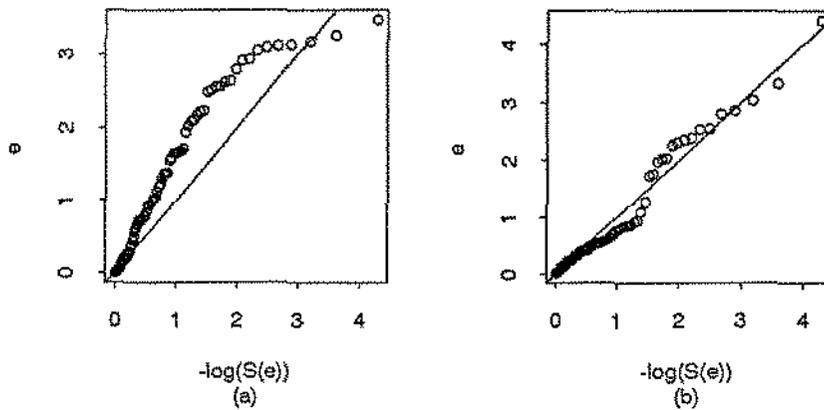


Figura 5.1: Gráfico dos resíduos, (a) regressão IG, (b) regressão R-IG.

Do gráfico dos resíduos podemos observar que os resíduos obtidos para o modelo de regressão da Recíproca complementar estão mais perto da reta com coeficiente angular 1, em comparação dos resíduos para o modelo de Regressão Gaussiana Inversa. Desta análise podemos concluir que o modelo de regressão dado para a Recíproca Complementar ajusta-se razoavelmente a este conjunto de dados.

Logo, considerando que o conjunto de dados ajusta-se a um modelo de regressão da Recíproca Complementar, temos as seguintes estimativas:

$$\hat{\beta}_0 = 2.968 (1.089)$$

$$\hat{\beta}_1 = 0.110 (0.040),$$

$$\hat{\nu} = 6.610 (4.883),$$

onde os valores entre parênteses ao lado das estimativas representam as raízes quadradas

dos elementos correspondentes de I_0^{-1} , sendo I_0 a matriz de informação observada dado em (5.39) com seus componentes dados para $p = 1$. Isto é, as estimativas dos desvios padrões das estimativas de máxima verossimilhança de β_0 , β_1 e ν .

Baseado em normalidade assintótica, os intervalos com 95% de confiança são

$$IC(\beta_0) : [-5.103, -0.833],$$

$$IC(\beta_1) : [0.031, 0.188],$$

$$IC(\nu) : [0, 16.180].$$

Os intervalos para β_0 e β_1 não contem o valor de zero, então, baseado em normalidade assintótica, podemos afirmar que com um nível de significância $\alpha = 0.05$ que estes parâmetros são diferentes de zero.

5.8.2 Exemplo 2- Teste de Vida de Isolantes Elétricos

Os dados da Tabela 5.3 são resultados de um teste acelerado de vida para isolantes elétricos de sistemas de uma nova classe H, e referem-se ao tempo em horas que os isolantes levaram para tornar-se defeituosos. Os testes foram realizados em pequenos motores, a temperaturas elevadas. Dez motores foram colocados para trabalhar a cada uma das temperaturas de 190, 220, 240 e 260 °C, e inspecionados periodicamente para detectar a ocorrência de falha. Os tempos da Tabela 5.3 referem-se à média entre o tempo em que a falha foi observada e o tempo da inspeção anterior. Os tempos entre inspeção foram 7, 4, 2, e 2 dias para as respectivas temperaturas. Para estes dados, temos 7 censuras, correspondendo ao tempo de vida dos isolantes que ainda estavam em boas condições no final do experimento.

Tabela 5.3: Tempo de vida de isolantes (horas) para cada temperatura, em graus centígrados, (+) observações censuradas.

	190	220	240	260
	7228	1764	1175	1128
	7228	2436	1521	1464
	7228	2436	1569	1512
	8448	2436 +	1617	1608
	9167	2436	1665	1632 +
	9167	2436	1665	1632 +
	9167	3108	1713	1632 +
	9167	3108	1761	1632 +
	10511	3108	1881+	1632 +
	10511	3108	1953	1896

Estes dados foram analisados por Nelson (1990), que considerou o modelo de regressão log-normal para o ajuste dos tempos de vida. Aqui vamos assumir o modelo de regressão para a Mistura de distribuições Gaussiana Inversa com sua Recíproca Complementar. Isto é, que os tempos de vida têm uma distribuição dada por (3.4), com parâmetro de escala ν e de forma μ_i , tal que

$$\mu_i = \beta_0 + \beta_1 x_i$$

para $i = 1, \dots, 4$ e onde $x_i =$ temperatura em graus centígrados. Para obter as estimativas, procuramos o valor de p no intervalo $[0, 1]$, que obtém-se o máximo valor da log-verossimilhança. As estimativas de β_0 , β_1 e ν , para diferentes valores de $p \in [0, 1]$ são obtidas pelo programa 2, feito em S-plus para p fixo, dado no Apêndice. Com este programa, obtemos a Tabela 5.4. Destes resultados, podemos observar que o valor de máxima verossimilhança obtém-se para $p = 1$, na fronteira

do espaço paramétrico de p . Logo, como na fronteira do espaço paramétrico não

Tabela 5.4: Estimativas de máxima verossimilhança para vários valores de p

β_0	β_1	ν	p	log-verosim
-1.373e-3	7.861e-6	1.608e-5	.0	-268.1
-1.371e-3	7.861e-6	1.610e-5	.1	-268.1
-1.370e-3	7.861e-6	1.613e-5	.2	-268.1
-1.368e-3	7.862e-6	1.619e-5	.3	-268.0
-1.367e-3	7.864e-6	1.625e-5	.4	-268.0
-1.366e-3	7.866e-6	1.634e-5	.5	-268.0
-1.365e-3	7.869e-6	1.645e-5	.6	-267.9
-1.365e-3	7.874e-6	1.657e-5	.7	-267.9
-1.364e-3	7.879e-6	1.672e-5	.8	-267.8
-1.364e-3	7.886e-6	1.688e-5	.9	-267.7
-1.365e-3	7.894e-6	1.706e-5	1.	-267.7

satisfaz as condições de regularidade da teoria de verossimilhança, não podemos fazer teste de hipóteses para determinar se a Recíproca Complementar é o melhor modelo que ajusta a este conjunto de dados em comparação com outros modelos da M-IG. Então, para avaliar a qualidade de ajuste na fronteira, fazemos a análise de resíduos, com as fórmulas dadas na Seção 5.5. Assim, na Figura 5.2 temos que o gráfico (a) representa os resíduos considerando um modelo de regressão Gaussiana Inversa e o gráfico (b) representa os resíduos considerando um modelo de regressão da Recíproca Complementar.

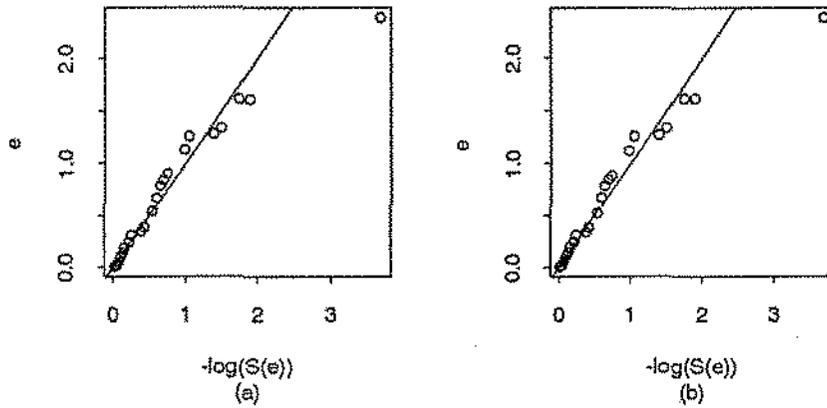


Figura 5.2: Gráfico dos resíduos, (a) regressão IG, (b) regressão R-IG.

O gráficos dos resíduos para ambos modelos são muito parecidos. Este resultado se deve a que a estimativa do parâmetro ν é muito pequena e como se observou no Capítulo 3, que para valores de μ semelhantes as densidades para $p \in [0, 1]$ são mais semelhantes quando ν é menor. Além disso, podemos observar que os resíduos estão ao redor da reta com coeficiente angular 1, exceto um resíduo que se encontra um pouco longe desta reta. Este ponto corresponde aos resíduos das 4 observações censuradas acontecida para a temperatura de 260 graus centígrados. Desta análise, podemos concluir que o modelo de regressão da M-IG para qualquer valor de $p \in [0, 1]$, se ajusta razoavelmente a este conjunto de dados. Logo, considerando que o conjunto de dados ajusta-se a um modelo de regressão da recíproca Complementar, temos as seguintes estimativas:

$$\hat{\beta}_0 = 1.365 \times 10^{-3} (9.177 \times 10^{-5}),$$

$$\hat{\beta}_1 = 7.894 \times 10^{-6} (4.389 \times 10^{-7}),$$

$$\hat{\nu} = 1.706 \times 10^{-5} (4.501 \times 10^{-6}),$$

onde os valores entre parênteses ao lado das estimativas representam as raízes quadradas dos elementos correspondentes de I_0^{-1} , sendo I_0 a matriz de informação observada dado em (5.87). Isto é, as estimativas dos desvios padrões das estimativas de máxima verossimilhança.

Baseado em normalidade assintótica, os intervalos com 95% de confiança são

$$IC(\beta_0) : [1.185 \times 10^{-3}, 1.545 \times 10^{-3}],$$

$$IC(\beta_1) : [7.034 \times 10^{-6}, 8.755 \times 10^{-6}],$$

$$IC(\nu) : [8.705 \times 10^{-6}, 2.589 \times 10^{-5}].$$

Estes intervalos não contém o valor zero, então, baseado em normalidade assintótica, podemos afirmar que com um nível de significância $\alpha = 0.05$ os parâmetros do modelo de regressão Recíproca complementar são diferentes de zero.

6. Considerações Finais

Do trabalho realizado podemos concluir que a M-IG constitui uma alternativa para modelar tempo de sobrevivência, assim como a distribuição Gaussiana Inversa e sua Recíproca Complementar.

A teoria desenvolvida nos Capítulos 4 e 5 possibilita o uso deste modelo, tanto para dados censurados como com covariáveis.

Quando o valor do parâmetro de dispersão ν é pequeno pode acontecer que seja impossível determinar para que valor de p se obtém o melhor ajuste.

Não existem testes para discriminar entre os modelos na fronteira do espaço paramétrico de p , já que nestes casos as condições de regularidade da teoria de verossimilhança não são satisfeitas.

Precisa-se fazer simulações para avaliar a qualidade de estimação de parâmetros que pode-se obter para M-IG; para uma amostra gerada aleatoriamente de mil pontos não obtivemos boas estimativas dos parâmetros fixados.

Nesta dissertação trabalhamos com $M - IG(\mu, \nu, p)$. Seria uma pesquisa interessante propor um modelo de regressão para $M - IG\left(\mu, \nu, \frac{\mu}{\gamma + \mu}\right)$, isto é, quando consideramos a seguinte reparametrização $\gamma = \mu\left(\frac{1-p}{p}\right)$. Por exemplo, ao considerar $\mu_i^{-1} = x_i \beta$, ν e γ constantes, temos um modelo diferente ao apresentado nesta dissertação.

Finalmente, para enriquecer o trabalho realizado, precisaríamos construir um algoritmo que permita determinar intervalos de confiança baseados na estatística da razão de verossimilhança.

APÊNDICE

PROGRAMAS-

Para usar estes programas elaborados em linguagem S-plus, precisa-se que o conjunto de dados constituam uma matrix, em onde a primeira columna represente a variavel tempo, a segunda coluna o estatus da observacao (falha=1, censura=0) e as outras colunas as covariaveis. Alem disso, as observacoes censuradas devem ser colocadas nas ultimas filas da matriz.

PROGRAMA 1

```
#Programas para a entrada de dados
#ler.dta.
lerdta <- function(arquivo,ncol)
  { if(ncol==2) {ler1dta(file,ncol)} else {ler2dta(file,ncol)}
  }

#ler1.dta.
ler1dta <- function(file,ncol)
  {dados<-matrix(scan(file),byrow=T, ncol=ncol)
  tempo<-dados[,1]
  vecm<-dados[,2]
  nf<-nrow(dados)
  nm<-t(vecm)%*%vecm
  nm<<-as.numeric(nm)
  xm<-rep(1,nm)
  xc<-NA
  xm<<-as.matrix(xm)}
```

```

tm<-tempo[1:nm]
if(nm==nf) {xc<-as.null(xc)} else {ncp<-nm+1; tc<-tempo[ncp:nf];
xc<-rep(1,(nf-nm)); xx<-rep(1,nf)}
tempo
}

```

#ler2.dta.

```

ler2dta <- function(file,ncol)
{dados<-matrix(scan(file),byrow=T, ncol=ncol)
tempo<-dados[,1]
vecm<-dados[,2]
nf<-nrow(dados)
nc<-ncol
if(nc==3) {covx<-dados[,3]} else {covx<-dados[,3:nc]}
xx<-cbind(intercept=1,covx)
nm<-t(vecm)%*%vecm
nm<-as.numeric(nm)
tm<-tempo[1:nm]
xm<-xx[1:nm,]
xc<-NA
if(nf==nm) {xc<-as.null(xc)} else {ncp<-nm+1;tc<-tempo[ncp:nf];
xc<-xx[ncp:nf,] }
tempo
}

```

PROGRAMA 2

#newbvpl.s

```

#programa para achar as estimativas de maxima verossimilhanca dos
#parametros da M-IG com p fixo, mais a o valor de o log da funcao de
#verossilhanca para dados com ou sem censura, tambem com ou sem covariaveis.
#xm=matriz das covariaveis incluindo intersepto para os tempo de falha.
#xc=matrix das covariaveis incluindo intersepto para os tempo censurados.
#tm=vetor de tempos de falha.
#tc=vetor de tempos de censura.
newbvpl<-function(xm,xc,tm,tc)
{if(is.null(xc)){bve<-c(bo(xm,tm),vio(xm,tm))}else
{bve<-bvem(xm,xc,tm,tc)}
pm<-0.0
bvpe<-c(bve,0)
}

```

```

lvn<-vlogv(bvpe)
uvob(xm,bve,bvpe,lvn,pm)
#   sink("d:/victor/res/leuke.out")
#   print(x=c(uvpest,lvn),digits=4)
#   print(x=c(bvpe,lvn),digits=4)
while(pm<=0.9)
  {pm<-pm+0.1
  lvi<-lvn
  bvest<-newbvpf(bve,pm)
  bvest<-as.vector(bvest)
  bvpe<-c(bvest,pm)
  lvn<-vlogv(bvpe)
  bve<-bvest
  uvob(xm,bve,bvpe,lvn,pm)
#   print(x=c(uvpest,lvn),append=T,digits=4)
#   print(x=c(bvpe,lvn),append=T,digits=4)
  }
#   sink()
#   if(ncol(xm)==1){uvpest}else{bvpest}
  bve
  }

```

PROGRAMA 3

```

#newbvp.s
#programa para achar as estimativas de maxima verossimilhanca dos parametros
#da M-IG com ou sem censura, tambem com ou sem covariaveis e para p entre 0 e 1
#bvpn=(betas,nu,p), para regressao, ou bvpn=(mu,nu,p), para uma amostra.
newbvp<-function(bvpn)
  { if(ncol(xm)==1){bet<-bvpn[1];ibet<-bet^-1;
  vp<-bvpn[2:3];bvpn<-c(ibet,vp)}
  lvi<-vlogv(bvpn)
  cat("bvpn=",bvpn,"lvi=",lvi)
  bvpi <- 0.01 * bvpn
  i<-0
  while(max(abs(bvpn-bvpi))>10e-6 && i<=500)
    { i<-i+1
    bvpi<-bvpn
    inf<-infma(bvpi)
    vscor<-fscorma(bvpi)
    }
  }

```

```

        bvpn<-bvpi+solve(inf)%*%vscor
        lvn<-vlogv(bvpn)
#       cat("i=",i,"bvpn=",bvpn,"lvn=",lvn,"\n")
        bve<-bvpn[1:(length(bvpn)-1)]
        pm<-bvpn[length(bvpn)]
        uvob(xm,bve,bvpn,lvn,pm)
        }
#     if(ncol(xm)==1){uvpest}else{bvpe}
        bve
    }

```

PROGRAMA 4

```

#newbvpf.s
#programa de Newton Raphson para a mistura da gaussiana inversa
#com seu reciproco complementar para p fixo.
newbvpf<-function(bvn,pm)
{ bvi<-0.001*bvn
  bvn<-as.vector(bvn)
  bvi<-as.vector(bvi)
  i<-0
  while(max(abs(bvn-bvi))>10e-8)
  { i<-i+1
    bvi<-bvn
    bvpi<-c(bvi,pm)
    inf<-infma(bvpi)
    vscor<-fscorma(bvpi)
    nom<-nrow(inf)-1
    inf1<-inf[i:nom,1:nom]
    vscor1<-vscor[i:nom]
    bvn<-bvi+solve(inf1)%*%vscor1
  }
  bvn
}

```

PROGRAMA 5

```

#fscorma.s
#funcao para achar a funcao scores para dados com censura ou sem ela, para

```

```
#M-IG.
```

```
#para p=0 vetor scores para gaussiana inversa.
```

```
fscorma<-function(bvp)
  {if(is.null(xc)) {fscormco(bvp)}
   else{fscormco(bvp)+fscormce(bvp)}
  }
```

```
#fscormco.s.
```

```
#funcao de scores da regressao da mistura da gaussiana inversa com seu
```

```
#complemento reciproco, para dados sem censura.
```

```
fscormco<-function(bvp)
  { bm<-bvp[1:(length(bvp)-2)]
    vm<-bvp[length(bvp)-1]
    pm<-bvp[length(bvp)]
    vm<-as.numeric(vm)
    pm<-as.numeric(pm)
    dm<-xm%*%bm
    dm<-as.vector(dm)
    idp<-(pm*(dm*tm)+(1-pm))^-1
    pfs<-pm*(t(xm)%*%(tm*idp))-t(xm)%*%((tm*dm)-1)/vm
    itm1<-tm^-1
    umo<-rep(1,nrow(xm))
    psfs<--nrow(xm)/(2*vm)
    ssfs<-(t(umo)%*%(itm1-2*dm)+t(dm)%*%(tm*dm))/(2*(vm^2))
    sfs<-psfs+ssfs
    tfs<-t(dm)%*%(tm*idp)-t(umo)%*%idp
    rbind(pfs,sfs,tfs)
  }
```

```
#fscormce.s.
```

```
#funcao de scores da regressao da mistura da gaussiana inversa com seu
```

```
#complemento reciproco, para dados com censura.
```

```
fscormce<-function(bvp)
  { bm<-bvp[1:(length(bvp)-2)]
    vm<-bvp[length(bvp)-1]
    pm<-bvp[length(bvp)]
    vm<-as.numeric(vm)
    pm<-as.numeric(pm)
    dc<-xc%*%bm
    dc<-as.vector(dc)
    en0<-no(tc,dc,vm,pm)
```

```

en1<-n1(tc,dc,vm,pm)
en2<-ens(tc,dc,vm,pm,en0,en1,1)
en0<-as.vector(en0)
en1<-as.vector(en1)
en2<-as.vector(en2)
ey<-(1-pm)*en1+pm*(dc*en2)
ein1<-eins(tc,dc,vm,pm,en0,en1,0)
ein1<-as.vector(ein1)
eyi1<-(1-pm)*ein1+pm*(dc*en0)
umo<-rep(1,nrow(xc))
ptpfs<-pm*(t(xc)%*%en1)
stpfs<-t(xc)%*%(ey*dc)
stpfs1<-t(xc)%*%umo
pfs<-ptpfs-((stpfs-stpfs1)/vm)
ptsfs<- -nrow(xc)/(2*vm)
stsfs<-t(umo)%*%eyi1
ttsfs<-2*(t(umo)%*%dc)
ctsfs<-t(dc)%*%(ey*dc)
sfs<-ptsfs+(stsfs-ttsfs+ctsfs)/(2*(vm^2))
ptfs<-t(dc)%*%en1
stfs<-t(umo)%*%en0
tfs<-ptfs-stfs
rbind(pfs,sfs,tfs)
}

```

PROGRAMA 6

```

#infma.s.
#funcao para achar a matriz de informacao observada para a regresao
#da mistura M-IG, com dados completos o com censura.
#para bvp=(b,v,0) matriz de inf. da gaussiana inversa.
infma<-function(bvp)
  {if(is.null(xc)){infmco(bvp)}
  else{infmco(bvp)+infmce(bvp)}
}

#infmco.s
#funcao para achar a matriz de informacao observada para dados sem censura.
infmco<-function(bvp)
  { bm<-bvp[1:(length(bvp)-2)]
  vm<-bvp[length(bvp)-1]

```

```

pm<-bvp[length(bvp)]
vm<-as.numeric(vm)
pm<-as.numeric(pm)
dm<-xm%*%bm
dm<-as.vector(dm)
dp<-pm*(dm*tm)+(1-pm)
idp<-dp^-1
iqdp<-dp^-2
qtm<-tm^2
qp<-pm^2
qdptm<-iqdp*qtm
umo<-rep(1,nrow(xm))
pibb<-qp*(t(xm)%*%(qdptm*xm))
sibb<-(t(xm)%*%(tm*xm))/vm
ibb<-pibb+sibb
ibv<- -pm*(t(xm)%*%(tm*idp))/vm
ibp<-t(xm)%*%(pm*(qdptm*dm-iqdp*tm)-tm*idp)
ivv<-nrow(xm)/(2*(vm^2))
#   stivv<-t(umo)%*%((tm^-1)-2*dm+(dm^2)*tm)/(vm^3)
#   ivv<-ptivv+stivv
ivp<-0
pipp<-t(dm)%*%(qdptm*dm)
sipp<-(2*t(tm*iqdp))%*%dm-t(iqdp)%*%umo
ipp<-pipp-sipp
dimb<-length(ibv)
dimv<-dimb+1
dimioc<-dimb+2
ioc<-matrix(NA,nrow=dimioc,ncol=dimioc)
ioc[1:dimb,1:dimb]<-ibb
ioc[1:dimb,dimv]<-ibv
ioc[dimv,1:dimb]<-t(ibv)
ioc[1:dimb,dimioc]<-ibp
ioc[dimioc,1:dimb]<-t(ibp)
ioc[dimv,dimv]<-ivv
ioc[dimv,dimioc]<-ivp
ioc[dimioc,dimv]<-t(ivp)
ioc[dimioc,dimioc]<-ipp
ioc
}
#infmce.s.

```

```
#matriz de informacao observada para a regressao da
#mistura da gaussiana inversa com seu complemento reciproco, so para dados
#censurados.
```

```
infmce<-function(bvp)
  { bm<-bvp[1:(length(bvp)-2)]
    vm<-bvp[length(bvp)-1]
    pm<-bvp[length(bvp)]
    vm<-as.numeric(vm)
    pm<-as.numeric(pm)
    dc<-xc%*%bm
    dc<-as.vector(dc)
    en0<-no(tc,dc,vm,pm)
    en1<-n1(tc,dc,vm,pm)
    en2<-ens(tc,dc,vm,pm,en0,en1,1)
    en3<-ens(tc,dc,vm,pm,en1,en2,2)
    eni1<-eins(tc,dc,vm,pm,en0,en1,0)
    eni2<-eins(tc,dc,vm,pm,eni1,en0,-1)
    en0<-as.vector(en0)
    en1<-as.vector(en1)
    en2<-as.vector(en2)
    en3<-as.vector(en3)
    eni1<-as.vector(eni1)
    eni2<-as.vector(eni2)
    em1<-(1-pm)*en1+pm*(dc*en2)
    em2<-(1-pm)*en2+pm*(dc*en3)
    emi1<-(1-pm)*eni1+pm*(dc*en0)
    emi2<-(1-pm)*eni2+pm*(dc*eni1)
    em1<-as.vector(em1)
    em2<-as.vector(em2)
    emi1<-as.vector(emi1)
    emi2<-as.vector(emi2)
    qdc<-dc^2
    umo<-rep(1,nrow(xc))
    ptibb<-pm*((pm*vm)*(en1^2)+2*(dc*(en2-emi1*en1)))/vm
    stibb<-(vm*em1-qdc*(em2-(em1^2)))/(vm^2)
    tibb<-ptibb+stibb
    ibb<-t(xc)%*(xc*tibb)
    ptibv<-((qdc*(em2-(em1^2))-(em1*emi1-1))*dc)/(2*(vm^3))
    stibv<-(qdc*(em2-en1*em1)-(eni1*emi1-en0-2*vm*en1))/(2*(vm^2))
    ibv<-t(xc)%*(ptibv-pm*stibv)
```

```

ptibp<-(qdc*(en2-em1*en1)-(en1-em1*en0)*dc)/vm
stibp<-en1*(pm*(en1*dc)-pm*en0-1)
ibp<-t(xc)%*%(ptibp+stibp)
ptivv<-nrow(xc)/(2*(vm^2))
stivv<-qdc*(em2-(em1^2))*qdc
ttivv<-2*(qdc*(em1*em1-1))
ctivv<-em12-(em1^2)
ivv<-ptivv-(t(stivv-ttivv+ctivv)%*%umo)/(4*(vm^4))
ptivp<-(en1-em1*en0)+qdc*(en1-em1*en0)
stivp<-(en0-em1*en1)*dc+qdc*(en2-em1*en1)*dc
tivp<-(ptivp-stivp)/(2*(vm^2))
ivp<-t(tivp)%*%umo
aipp<-qdc*(en1^2)-2*(dc*en1*en0)+(en0^2)
ipp<-t(aipp)%*%umo
dimb<-length(ibv)
dimv<-dimb+1
dimio<-dimb+2
io<-matrix(NA,nrow=dimio,ncol=dimio)
io[1:dimb,1:dimb]<-ibb
io[1:dimb,dimv]<-ibv
io[dimv,1:dimb]<-t(ibv)
io[1:dimb,dimio]<-ibp
io[dimio,1:dimb]<-t(ibp)
io[dimv,dimv]<-ivv
io[dimv,dimio]<-ivp
io[dimio,dimv]<-t(ivp)
io[dimio,dimio]<-ipp
io
}

```

PROGRAMA 7

```

#bvem.s.
#acha as estimativas de maxima verossimilhanca para os parametros
#da gaussiana inversa com dados censurados via o algoritmo EM.
bvem<-function(xm,xc,tm,tc)
{ tm<-as.vector(tm)
  tc<-as.vector(tc)
  bi<-bo(xm,tm)
  vi<-vio(xm,tm)

```

```

d<-xc%*%bi
ntc<-as.vector(n1(tc,d,vi,0))
nt<-c(tm,ntc)
ibo<-as.vector(bo(xx,nt))
ivo<-as.vector(vio(xx,nt))
bvo<-c(ibo,ivo)
bvn<-nbv(ibo,ivo)
i<-0
starttime <- proc.time()
while(max(abs(bvn-bvo))>10e-8)# 10e-8 deu certo.
  {i<-i+1
  bvo<-bvn
  ibo<-bvo[1:length(ibo)]
  ivo<-bvo[length(bvo)]
  bvn<-nbv(ibo,ivo)
  bvno<-c(bvn,0)
  lvig<-vlogv(bvno)
  cat(" i = ",i,"bvn=",bvn,"lvig=",lvig,"\n")
  }
  print(bvn)
cat(paste("elapsed time = ",proc.time()-starttime,sep=""),"\n")
  bvn
  }

```

PROGRAMA 8

```

#nbv.s
#programa para achar os novos estimados dos coeficientes de regressao e
#parametro de volatility para dados de gaussiana inversa que permita
#cumprir com a condicao de precisao ao iniciar o algoritmo EM.
nbv<-function(b1,vi)
  { b1<-as.vector(b1)
  dn<-xc%*%b1
  emn<-n1(tc,dn,vi,0)
  emn<-as.vector(emn)
  iemn<-eins(tc,dn,vi,0,i,emn,0)
  iemn<-as.vector(iemn)
  m1<-c(tm,emn)
  iy<-tm^-1
  im1<-c(iy,iemn)

```

```

bn<-bo(xx,m1)
bn<-as.vector(bn)
umo<-rep(1,length(im1))
vn<-(t(umo)%*%im1-t(bo(xx,m1))%*%t(xx)%*%umo)/length(im1)
vn<-as.vector(vn)
c(bn,vn)
}

```

PROGRAMA 9

```

#bo.s
#programa para achar os coeficientes de regressao da gaussiana inversa
#com dados completos.
bo<-function(x1,t1) #estimados dos coeficientes de regressao (b)
{ n<-length(t1) # Whitmore p.306 eqn 3.
  umo<-rep(1,n)
  invvx <- solve(t(x1)%*%(x1*t1))%*%t(x1)
  invvx%*%umo
}

```

PROGRAMA 10

```

#vio.s
#programa para achar o estimado do parametro de volatility
#na regressao gaussiana inversa.
vio<-function(z1,s1) #estimado do parametro de volatility
{ n<-length(s1) # Whitmore p.306 eqn 4.
  umo<-rep(1,n)
  y<-s1^-1
  (t(umo)%*%(umo*y)-t(umo)%*%z1)%*%bo(z1,s1))/n
}

```

PROGRAMA 11

```

#no.s.
#programa para achar  $E[(1-p+pu^{-1}Y)^{-1}/Y > a]$ .
no<-function(tpc,dp,vp,p)
{ spm<-1-Gpy(tpc,dp,vp,p)
  sg<-1-Gpy(tpc,dp,vp,0)
  sg/spm
}

```

}

#n1.s.

#programa para achar $E[Y(1-p+pdY)/Y > a]$.

```
n1<-function(tpc,dp,vp,p)
  { iqdy<-((dp^2)*tpc)^-1
    spm<-1-Gpy(tpc,dp,vp,p)
    Gpy(iqdy,dp,vp,0)/(dp*spm)
  }
```

#ens.s.

#programa para achar $E[Y^s(1-p+pdY)/Y > a]$, para $s > 1$.

```
ens<-function(tpc,dp,vp,p,nsm1,nsm2,s)
  { qd<-dp^2
    ptns<-(vp*(2*s-1))*nsm2
    ttns<-((2*vp)*((tpc^(s+1))*gpm(tpc,dp,vp,0)))/(1-Gpy(tpc,dp,vp,p))
    (ptns+nsm1+ttns)/qd
  }
```

#eins.s

#programa para achar $E[Y^s(1-p+pdY)^{-1}/Y > a]$, para $s < 0$.

```
eins<-function(tpc,dp,vp,p,ns1,ns2,s)
  { qd<-dp^2
    ptins<-qd*ns2
    stins<-(vp*(2*s-1))*ns1
    ttins<-((2*vp)*((tpc^(s+1))*gpm(tpc,dp,vp,0)))/(1-Gpy(tpc,dp,vp,p))
    ptins-stins-ttins
  }
```

PROGRAMA 12

#gpm.s

#funcao de densidade da mistura gaussiana inversa com seu reciproco

#complementar.

#funcao de densidade da gaussiana inversa=gpm(te,de,ve,0)

```
gpm<-function(te,de,ve,pe)
  { cy<-2*pi*(te^3)*ve
    ecy<-cy^-.5
    z<- -((de*te-1)^2/(2*ve*te))
    (1-pe+pe*de*te)*ecy*exp(z)
```

}

PROGRAMA 13

```

#Gpy.s
#programa para achar o valor da funcao de distribuicao da mistura
#gaussiana inversa com seu reciproco complementar.
#funcao de distribuicao gaussiana inversa evaluado no
#ponto "y"=Gpy(y,d,v,0). # Whitmore p.306 eqn 5.
Gpy<-function(y,d,v,p)
  { y<-as.vector(y)
    d <- as.vector(d)
    co<-(y*v)^.5
    z1<-(d*y-1)/co
    z2<--(d*y+1)/co
    p1<-pnorm(z1,0,1)
    p2<-pnorm(z2,0,1)
#    if(p2==0){p1} else{p1+(1-2*p)*exp(2*d/v)*p2}
    x2<-p1+(1-2*p)*exp(2*d/v)*p2
    x2[p2==0]<-p1[p2==0]
    x2
  }

```

PROGRAMA 14

```

#vlogv.s
#programa para achar o valor de log da funcao de verossimilanca
#da mistura da gaussiana inversa com sua reciproca complementar.
vlogv<-function(bvp)
  {nlog<-length(bvp)
    if(!is.vector(bvp)) bvp<-as.vector(bvp)
    bn<-bvp[1:(nlog-2)]
    bn<-as.vector(bn)
    vn<-bvp[nlog-1]
    pn<-bvp[nlog]
    if(is.null(xc)){lv<-lvtco(tm,bn,vn,pn)}
    else{lvtco(tm,bn,vn,pn)+
    lvtce(tc,bn,vn,pn)}
  }

```

```
#lvtco.s.  
#programa para achar o valor do log-verossimilhanca para a M-IG, sem censura.  
lvtco<-function(tmv,bmv,vmv,pmv)  
  { dmv<-xm%*%bmv  
    dmv<-as.vector(dmv)  
    gtm<-gpm(tmv,dmv,vmv,pmv)  
    lgtm<-log(gtm)  
    umof<-rep(1,length(gtm))  
    t(lgtm)%*%umof  
  }  
  
#lvtce.s  
#programa para achar o valor de logverossimilhanca para M-IG com censura.  
lvtce<-function(tcv,bcv,vcv,pcv)  
  { dcv<-xc%*%bcv  
    dcv<-as.vector(dcv)  
    gtc<-Gpy(tcv,dcv,vcv,pcv)  
    ligtc<-log(1-gtc)  
    umoc<-rep(1,length(gtc))  
    t(ligtc)%*%umoc  
  }
```

BIBLIOGRAFIA

- [1] Atkinson, S. E. (1992). The performance of standard and hybrid EM algorithms for ML estimates of the Normal Mixture model with censoring. *J. Statist. Comput. Simul.*, Vol. 44:105-115.
- [2] Birnbaum, Z. W. e Saunders (1969), S. C.. A new family of life distributions. *J. Appl. Probab.* 6:319-327.
- [3] Becker, R. A., Chambers, J. M. e Wilks, A. R. (1988). *The New S language - A programming Environment for Data Analysis and Graphics*. California: Wadsworth & Brooks/Cole
- [4] Chhikara, R. S., e Folks, J. L. (1974). Estimation of the inverse Gaussian distribution function. *J. Amer. Statist. Ass.*, 69:250-254.
- [5] Chhikara, R. S., e Folks, J. L. (1976). Optimum test procedures for the mean of first passage time in Brownian motion with positive drift (Inverse Gaussian distribution). *Technometrics*, 18:189-193.
- [6] Chhikara, R. S., e Folks, J. L. (1977). The Inverse Gaussian distribution as a lifetime model. *Technometrics*, 19:461-468.

- [7] Chhikara, R. S., e Folks, J. L. (1989). *The Inverse Gaussian distribution*, New York: Marcel Dekker, Inc.
- [8] Davis, A. S. (1980). Use of the likelihood ratio test on the Inverse Gaussian distribution. *Amer. Statist.*, 34:108-110.
- [9] Dempster, A. P., Laird, N. M., e Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via EM algorithm (with discussion). *J. Roy. Statist. Soc. B* 39:1-38.
- [10] Doganaksoy, N., e Schemee, J. (1993). Comparisons of approximate confidence intervals for distributions used in life-data analysis. *Technometrics*, 35:175-184
- [11] Folks, J. L., e Chhikara, R. S. (1978). The Inverse Gaussian distribution and its statistical application (with discussion). *J. Roy. Stat. Soc. Ser B* 40:263-289.
- [12] Feller, W. (1971). *An introduction to probability theory and its applications*, vol II (2nd ed.). New York: Wiley.
- [13] Hildebrand, F. B. (1976). *Advanced Calculus for Applications*, (2nd ed), Prentice Hall
- [14] Johnson, N. L., e Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions 1*. Boston: Houghton Mifflin.
- [15] Jørgensen, B. (1982). Statistical properties of the Generalized Inverse Gaussian distribution. *Lecture Notes in Statistics, Vol. 9*, New York: Springer-Verlag.
- [16] Jørgensen, B. (1987). Exponential dispersion Models (with discussion). *J. Roy. Statist. Soc. Ser. B* 49: 127-162.

- [17] Jørgensen, B. (1992). Exponential dispersion Models and extensions: a Review. *Int. Statist. Rev.* 60: 5-20.
- [18] Jørgensen, B., Seshadri, V. e Whitmore, G. A. (1991). On the mixture of the inverse Gaussian distribution with its complementary reciprocal. *Scand. J. Statist.* 18: 77-89.
- [19] Kalbfleisch, J. D., e Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*. New York: Wiley
- [20] Kotz, S., e Johnson, L. (1982). *Encyclopedia of Statistical Sciences, Vol. 1* New York: Wiley.
- [21] Kotz, S., e Johnson, L. (1986). *Encyclopedia of Statistical Sciences, Vol. 7* New York: Wiley.
- [22] Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*, New York: Wiley.
- [23] Letac, G., e Seshadri, V. (1986). On a conjecture concerning Inverse Gaussian Regression. *Int. Statist. Rev.*, 54:187-190.
- [24] Little, R. J. A., e Rubin, D. B. (1987). *Statistical Analysis with Missing Data*, New York: Wiley.
- [25] Maehara, V. M. (1994). Análise da Função Risco da Mistura de Distribuições Gaussianas Inversas com sua Recíproca Complementar. Relatório não publicado.
- [26] Miura, C. K. (1978). Tests for the mean of the inverse Gaussian distribution. *Scand. J. Statist.* 5:200-204.

- [27] Nelson, W. B. (1982). *Applied Life Data Analysis*. New York: Wiley
- [28] Nelson, W. B. (1990). *Accelerated Testing - Statistical Models Plans, and Data Analysis*. New York: Wiley
- [29] Schrödinger, E. (1915). Zur theorie der fall-und steigversuche an teilchen mit Brownscher Bewegung. *Physikalische Zeitschrift*, 16:133-134.
- [30] Steutel, F. W. (1979). Infinite Divisibility in Theory and Practice. *Scand J. Statist.*, 6: 57-64.
- [31] Shuster, J. J. (1968). On the Inverse Gaussian distribution function. *J.Amer. Statist. Ass.*, 63:1514-1516.
- [32] Thisted, R. A. (1988). *Elements of Statistical Computing*. Chapman and Hall.
- [33] Titterington, D. M., Smith, A. F. e Makov, U. E. (1985). *Statistical Analysis of Finite Mixture Distributions*. New York: Wiley.
- [34] Tweedie, M. C. K. (1956). Some statistical Properties of inverse Gaussian distributions. *Virginia J. Sci.*, 7:160-165.
- [35] Tweedie, M. C. K. (1957a). Statistical properties of Inverse Gaussian distributions I. *Ann. Math. Statist.*, 28:362-377.
- [36] Tweedie, M. C. K. (1957b). Statistical properties of Inverse Gaussian distributions II. *Ann. Math. Statist.*, 28:696-705.
- [37] Venzon, D. J. e Moolgavkar, S. H. (1988) . A method for computing profile likelihood based confidence intervals. *Appl. Statist.*, 37:87-94.

- [38] Wald, A. (1947). *Sequential Analysis*, New York: Wiley.
- [39] Wasan, M. T. (1968). First passage time distribution of Brownian motion with positive drift (Inverse Gaussian distribution). Queen's Paper in Pure and Applied Mathematics, 19, Queens University, Canada.
- [40] Whitmore, G. A. (1983). A regression method for censored Inverse Gaussian data. *Can. J. Statist.*, 11:305-315.
- [41] Whitmore, G. A., e Yalovsky, M. (1978). A Normalizing Logarithmic Transformation for Inverse Gaussian Variables. *Technometrics*, 20: 207-208.
- [42] Zelen, M. (1969). Factorial experiments in life testing. *Technometrics*, 1:269-288.