



IZABELA REGINA CARDOSO DE OLIVEIRA

**MÉTODOS PARA COMPUTAÇÃO DA FUNÇÃO DE
DISTRIBUIÇÃO GAMA INCOMPLETA
NÃO-CENTRAL, DE SUA INVERSA E DO
PARÂMETRO DE NÃO-CENTRALIDADE**

LAVRAS - MG

2011

IZABELA REGINA CARDOSO DE OLIVEIRA

**MÉTODOS PARA COMPUTAÇÃO DA FUNÇÃO DE DISTRIBUIÇÃO
GAMA INCOMPLETA NÃO-CENTRAL, DE SUA INVERSA E DO
PARÂMETRO DE NÃO-CENTRALIDADE**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

Orientador
Dr. Daniel Furtado Ferreira

**LAVRAS - MG
2010**

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca da UFLA**

Oliveira, Izabela Regina Cardoso de.

Métodos para computação da função de distribuição gama incompleta não-central, de sua inversa e do parâmetro de não-centralidade / Izabela Regina Cardoso de Oliveira. – Lavras : UFLA, 2010.

60 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2010.

Orientador: Daniel Furtado Ferreira.

Bibliografia.

1. Qui-quadrado não-central. 2. Poisson. 3. Algoritmo. 4. R. 5. Poder. I. Universidade Federal de Lavras. II. Título.

CDD – 519.24

IZABELA REGINA CARDOSO DE OLIVEIRA

**MÉTODOS PARA COMPUTAÇÃO DA FUNÇÃO DE DISTRIBUIÇÃO
GAMA INCOMPLETA NÃO-CENTRAL, DE SUA INVERSA E DO
PARÂMETRO DE NÃO-CENTRALIDADE**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

APROVADA em 28 de dezembro de 2010.

Dr. Daniel Furtado Ferreira

UFLA

Dr. José Airton Rodrigues Nunes

UFLA

Dr. Renato Ribeiro de Lima

UFLA

Dr. Daniel Furtado Ferreira
Orientador

LAVRAS - MG

2010

*Aos meus pais Tadeu e Maria José,
ao meu irmão Delano,
pelo amor incondicional.*

DEDICO

AGRADECIMENTOS

Ao amigo e orientador Daniel Furtado Ferreira, por todos os conselhos, pela atenção e pelas palavras sábias que sempre me incentivaram e iluminaram meus caminhos. Sua dedicação e profissionalismo serão minha inspiração por toda a vida.

Aos meus pais, Tadeu e Maria José; ao meu irmão, Delano, pelo amor, apoio e respeito em relação às minhas decisões, sempre caminhando ao meu lado. Vocês são meu alicerce e fonte de força!

Ao Marcelo, pela paciência, por me escutar, por ser meu grande amigo e companheiro de todas as horas.

Aos professores do Departamento de Ciências Exatas da UFLA, pelos ensinamentos, pela amizade e excelente convivência.

Aos professores da banca, Renato Ribeiro de Lima e José Airton Rodrigues Nunes, pela disponibilidade, compreensão e pelas valiosas contribuições que muito engrandeceram este trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro eficiente e constante durante o curso.

Às amigas do peito, Helena, Paula e Raquel, pela amizade doce e sincera de longos anos. Obrigada pela presença em mais uma etapa da minha vida!

A todos os funcionários do Departamento de Ciências Exatas, em especial a Josi (secretária da Pós-graduação), pela prontidão e suporte.

Aos colegas do curso de Pós-graduação em Estatística e Experimentação Agropecuária, em especial, Carol, Hernani, Felipe, Danilo, Siomara e Andressa. Foi um imenso prazer conhecê-los e conviver com todos durante esse tempo.

Ao meu primo Fabyano, pelos conselhos e pelo *insight* que mudou a minha vida.

Muito obrigada a todos!

RESUMO

A distribuição gama incompleta não-central pode ser vista como a forma geral da distribuição qui-quadrado não-central e é expressa como uma mistura de uma Poisson com a função gama incompleta. Knüsel e Bablok (1996) e Ruben (1974) propuseram aproximações para a obtenção da função gama incompleta não-central, sem, no entanto, que os procedimentos numéricos tivessem sido implementados. Chen (2005) apresenta métodos para a obtenção da função de distribuição gama incompleta não-central e de sua inversa, mas os códigos de suas soluções não estão disponíveis. Sendo assim, a distribuição gama incompleta não-central não está disponível nos programas estatísticos convencionais como SAS, R, Statistica, entre outros. Para o caso particular da distribuição qui-quadrado não-central, existem soluções numéricas de alta qualidade e implementações em programas como o R. Entretanto, um dos problemas ainda remanescentes é o da obtenção do parâmetro de não-centralidade, dados os valores da função de distribuição, de x e de α , tanto para a distribuição gama, quanto para a qui-quadrado não-central. Este trabalho teve como objetivo propor um algoritmo adaptando o método originalmente proposto para a função beta incompleta não-central de Benton e Krishnamoorthy (2003) e um outro que combine o método desses autores com o método da inversão da função de distribuição em relação ao parâmetro de não-centralidade, utilizando para isso o método de Newton-Raphson. Os algoritmos foram disponibilizados em pseudocódigos e implementados como rotinas do programa R. Para avaliar a acurácia e a velocidade de processamento dos algoritmos implementados em R, foram obtidos resultados da função de distribuição, função densidade, quantis e parâmetro de não-centralidade da gama incompleta não-central e de seu caso particular, a qui-quadrado não-central, utilizando para isso as configurações dos argumentos utilizadas nos trabalhos de Benton e Krishnamoorthy (2003) e Chen (2005). Os resultados obtidos por esses autores em suas aproximações foram utilizados para fins de comparação. As rotinas implementadas apresentaram bom desempenho e, em geral, foram mais acuradas que as rotinas concorrentes. Quanto ao tempo de processamento, se implementadas em uma linguagem de mais baixo nível, as rotinas propostas são mais eficientes do que o algoritmo disponível no programa R para o cálculo da função de distribuição qui-quadrado não-central.

Palavras-chave: Qui-quadrado não-central. Poisson. Algoritmo. R. Poder.

ABSTRACT

The noncentral incomplete gamma distribution can be viewed as a generalization of the noncentral chi-squared distribution and it can be expressed as a mixture of a Poisson function with a incomplete gamma function. Approximations to the noncentral gamma distribution were proposed by Knüsel and Bablok (1996) and Ruben (1974), but they did not implement numerical procedures. Chen (2005) presents methods for obtaining the noncentral incomplete gamma function and its inverse, but the codes of his solutions are not available. The noncentral gamma distribution is not available in conventional statistical programs like SAS, R, Statistica and other. For the particular case of the noncentral chi-squared distribution high-quality numerical solutions are implemented in programs like R . However, one problem still remaining: obtaining the noncentrality parameter given the values of the distribution function, x and α , for both the noncentral gamma and the noncentral chi-squared distributions. This paper aimed to propose an algorithm by combining the method originally proposed by Benton and Krishnamoorthy (2003) for the noncentral beta distribution with the method of inversion of the distribution function with respect to the noncentrality parameter using Newton-Raphson. The algorithms were available in pseudocodes and implemented as routines of the R program. To evaluate the accuracy and speed of computation of the algorithms implemented in R, results of the distribution function, density function, quantiles and noncentrality parameter of the noncentral incomplete gamma and its particular case, the noncentral chi-squared, were obtained for the arguments settings used in the papers of Benton and Krishnamoorthy (2003) and Chen (2005) . The results obtained by these authors in their approaches were used for comparison purposes. The implemented routines performed well and, in general, were more accurate than other approximations. If the proposed routines were implemented in a lower level language, they would be more efficient, in terms of speed of computation, than the algorithm available in R program for the noncentral chi-squared distribution.

Keywords: Noncentral chi-squared. Poisson. Algorithm. R. Power.

LISTA DE FIGURAS

Figura 1	Função gama completa para valores reais do argumento $\alpha > 0$. . .	14
Figura 2	Função densidade gama padrão	19
Figura 3	Função densidade gama não-central para $\alpha = 5$ (a) e para $\alpha = 20$ (b), considerando diferentes valores de δ	47
Figura 4	Função densidade gama não-central para $\delta = 5$ (a) e para $\delta = 20$ (b), considerando diferentes valores de α	48

LISTA DE TABELAS

Tabela 1	Valores exatos e computados de $P(\chi_{\nu}^2(\delta) \leq x)$, usando o método proposto (MP), o método de Benton e Krishnamoorthy (2003) (BK) e a computação intervalar (Intervalar), para as configurações consideradas por Wang e Kennedy (1994 citados por BENTON; KRISHNAMOORTHY, 2003, p. 257).	42
Tabela 2	Valores exatos e computados de $P(\chi_{\nu}^2(\delta) \leq x)$, usando o método proposto (MP) e o método disponível no programa R (R), para as configurações consideradas por Wang e Kennedy (1994 citados por BENTON; KRISHNAMOORTHY, 2003, p. 257).	43
Tabela 3	Valores exatos e computados de $I_x(\alpha, \delta)$ para algumas configurações apresentadas por Chen (2005) usando o método sugerido por este autor, baseado em equações diferenciais de segunda ordem a partir da função gama (GB), e o método proposto (MP).	44
Tabela 4	Valores exatos e computados dos três quartis de $I_x(\alpha, \delta)$ para algumas configurações apresentadas por Chen (2005) usando o método sugerido por este autor, baseado em equações diferenciais de segunda ordem a partir da função gama (GB), e o método proposto (MP).	45
Tabela 5	Valores exatos e computados do parâmetro de não-centralidade δ para algumas configurações, usando o método proposto (MP).	46
Tabela 6	Valores da função densidade qui-quadrado não-central $\chi_{\nu}^2(\delta)$ para diferentes configurações dos parâmetros, usando o método proposto (MP) e o método disponível no programa R (R).	46

SUMÁRIO

1	INTRODUÇÃO	11
2	REFERENCIAL TEÓRICO	14
2.1	A função gama	14
2.2	A distribuição gama	16
2.3	Método de Newton-Raphson e método da secante	20
2.4	Distribuições não-centrais	22
2.5	Algoritmos para computação de funções de distribuições não-centrais	25
3	MATERIAL E MÉTODOS	28
3.1	A função de distribuição gama incompleta não-central	28
3.2	Inversa da função de distribuição em relação a δ	32
3.3	Inversa da função de distribuição em relação a x	35
3.4	A função densidade da distribuição gama incompleta não-central	38
3.5	Implementação das rotinas e comparação dos resultados	39
4	RESULTADOS E DISCUSSÃO	41
5	CONCLUSÕES	49
	REFERÊNCIAS	50
	APÊNDICE A	53

1 INTRODUÇÃO

A distribuição gama incompleta não-central pode ser vista como a forma geral da distribuição qui-quadrado não-central. As distribuições não-centrais são vistas como misturas de uma distribuição discreta com outra contínua. A distribuição gama incompleta não-central pode ser expressa dessa forma como uma mistura de uma Poisson com a função gama incompleta. Assim, a função de distribuição gama incompleta não-central é dada por

$$I_x(\alpha, \delta) = P(X \leq x) = \sum_{i=0}^{\infty} \frac{e^{-\delta/2} (\delta/2)^i}{i!} I_x(\alpha + i), \quad (1)$$

em que $\delta \geq 0$ é o parâmetro de não-centralidade, $x > 0$, $I_x(\alpha)$ é a função de distribuição gama incompleta central

$$I_x(\alpha) = \frac{1}{\Gamma(\alpha)} \int_0^x e^{-t} t^{\alpha-1} dt$$

$\Gamma(\alpha)$ e a função gama completa e $\alpha > 0$ é o parâmetro de forma da gama. Se $\delta = 0$, a função gama incompleta não-central simplifica-se no caso central. Se $x = 0$, o valor retornado é igual a 0 e se $x = 1$, o valor retornado é igual a 1.

Knüsel e Bablok (1996) e Ruben (1974) propuseram aproximações para a obtenção da função gama incompleta não-central, sem, no entanto, que os procedimentos numéricos tivessem sido implementados. Segundo Chen (2005), as aproximações para a densidade e inversa da função de distribuição gama não-central não estão disponíveis na literatura. Esse mesmo autor apresenta soluções para esses problemas usando equações diferenciais não-lineares do Sistema S. Infelizmente, os códigos utilizados pelo autor não estão disponíveis, ao contrário do que é afirmado em seu trabalho. Sendo assim, a distribuição gama incompleta não-central não está disponível nos programas estatísticos convencionais como SAS, R, Statistica, entre outros.

Para o caso particular da distribuição qui-quadrado não-central, existem soluções numéricas de alta qualidade (BENTON; KRISHNAMOORTHY, 2003) e implementações em programas como SAS e R. Entretanto, um dos problemas

ainda remanescentes é o da obtenção do parâmetro de não-centralidade, dados os valores da função de distribuição, de x e de α , tanto para a distribuição gama, quanto para a qui-quadrado não-central.

As distribuições não-centrais são utilizadas no cálculo do poder de alguns testes como, por exemplo, o de χ^2 aplicado a tabelas de contingência e em testes de razão de variâncias. Nesses casos, as distribuições frequentemente utilizadas são a qui-quadrado e F não-centrais. As três distribuições centrais e não-centrais comumente usadas - qui-quadrado, F e t - são casos particulares das distribuições gama e beta centrais e não-centrais. Destaca-se, então, a importância do desenvolvimento e aperfeiçoamento de métodos computacionais para a aproximação das funções gama e beta não-centrais.

Os algoritmos tradicionais utilizados no cálculo das distribuições não-centrais iniciam o cômputo da função em $i = 0$ e as probabilidades discreta e contínua para $i = 1, 2, \dots$ são computadas recursivamente. Além de esses algoritmos serem ineficazes em relação ao tempo de processamento, quando a média da variável aleatória discreta é grande, pode ocorrer um problema computacional conhecido como *underflow error*.

Duas alternativas foram propostas na literatura para contornar esse problema. A primeira utiliza o cálculo da função de distribuição de forma recursiva, a partir do valor inteiro mais próximo da média da distribuição de probabilidade discreta na mistura (BENTON; KRISHNAMOORTHY, 2003). A segunda faz o uso de um cálculo recursivo da função de distribuição de probabilidade acumulada (fdpa) a partir de um intervalo delimitado por k_1 e k_2 , obtidos a partir do modelo probabilístico discreto da mistura (BAHAREV; KEMÉNY, 2008). Essa segunda abordagem foi usada por Knüsel e Bablok (1996) para a obtenção da função gama não-central, mas possui a grande restrição de ter que estabelecer os limites k_1 e k_2 . Para isso é necessário inverter a função de distribuição de probabilidade do modelo discreto da mistura. Associado a esse segundo algoritmo, Baharev e Kemény (2008) propuseram o uso do método de Newton-Raphson para a obtenção do parâmetro de não-centralidade δ , especificamente para a distribuição beta não-central.

Assim, os objetivos desse trabalho são:

1. Propor um algoritmo usando o método de Benton e Krishnamoorthy (2003)

em seu pseudocódigo para obter a função de distribuição gama incompleta não-central;

2. Propor um algoritmo que combine o método de Benton e Krishnamoorthy (2003) com o método de Newton-Raphson em seu pseudocódigo para obter a inversa e o parâmetro de não-centralidade da função gama incompleta não-central;
3. Implementar os algoritmos propostos como rotinas do programa R;
4. Comparar o desempenho dos algoritmos propostos com os resultados apresentados na literatura de outros algoritmos e com os resultados obtidos pelo programa R da distribuição qui-quadrado não-central.

2 REFERENCIAL TEÓRICO

2.1 A função gama

A função gama - denotada por $\Gamma(\cdot)$ e frequentemente utilizada na estatística para definir algumas funções densidades – é expressa por

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \quad (2)$$

para $\alpha > 0$. A função gama completa está ilustrada na Figura 1 em função de α .

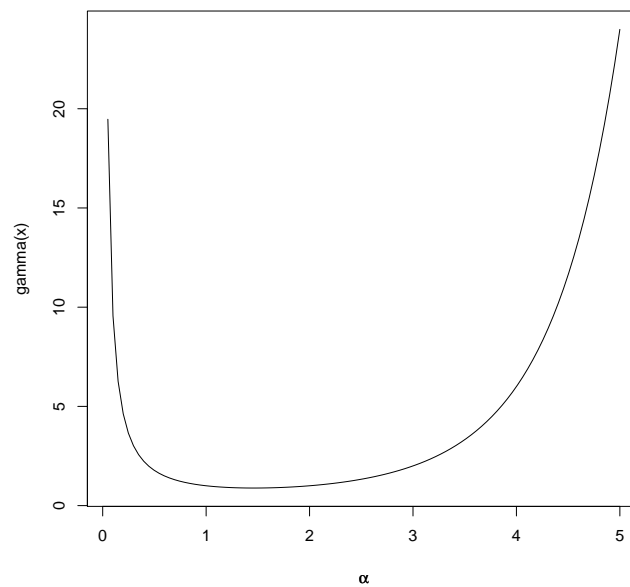


Figura 1 Função gama completa para valores reais do argumento $\alpha > 0$

De acordo com Mood, Graybill e Boes (1974), $\Gamma(\alpha)$ é simplesmente uma notação para a integral definida da equação (2), que não possui solução explícita. Entretanto, a partir de uma relação de recorrência e fazendo a integração por partes

para $dv = e^{-x}dx$ e $u = x^{\alpha-1}$, tem-se

$$\begin{aligned}\Gamma(\alpha) &= \int_0^{\infty} x^{\alpha-1} e^{-x} dx \\ &= -x^{\alpha-1} e^{-x} \Big|_0^{\infty} + (\alpha-1) \int_0^{\infty} x^{\alpha-2} e^{-x} dx \\ &= (\alpha-1)\Gamma(\alpha-1)\end{aligned}$$

e o resultado

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha).$$

Ainda, ao considerar $\alpha = n$ e n um número inteiro qualquer, verifica-se

$$\Gamma(n+1) = n!.$$

Outro importante resultado deve ser destacado ao considerar n como um número inteiro:

$$\Gamma\left(n + \frac{1}{2}\right) = \frac{1.3.5 \dots (2n-1)}{2^n} \sqrt{\pi}$$

e, em particular,

$$\Gamma\left(\frac{1}{2}\right) = 2\Gamma\left(\frac{3}{2}\right) = \sqrt{\pi}.$$

Aproximações numéricas baseadas em expansões em série são encontradas na literatura para calcular os valores da função gama (PRESS et al., 1992). A fórmula de Lanczos (1964) é uma das melhores aproximações e pode ser apresentada por

$$\begin{aligned}\Gamma(z+1) &= \left(z + \gamma + \frac{1}{2}\right)^{(z+1/2)} e^{-(z+\gamma+1/2)} \sqrt{2\pi} \left[\rho_0 + \frac{\rho_1}{z+1} + \frac{\rho_2}{z+2} + \dots \right. \\ &\quad \left. \dots + \frac{\rho_N}{z+N} + \epsilon \right],\end{aligned}\tag{3}$$

para $z > 0$, γ e ρ_i constantes, para $i = 0, 1, \dots, N$. Para $N = 6$, o autor mostrou que o erro é menor do que 2×10^{-10} , em todo o domínio real da função ($z > 0$), sendo a $\gamma = 5$ e as demais constantes dadas por:

$$\left\{ \begin{array}{l} \rho_0 = 1,000000000190015 \\ \rho_1 = 76,18009172947146 \\ \rho_2 = -86,50532032941677 \\ \rho_3 = 24,01409824083091 \\ \rho_4 = -1,231739572450155 \\ \rho_5 = 1,208650973866179 \times 10^{-3} \\ \rho_6 = -5,395239384953 \times 10^{-6}. \end{array} \right.$$

O argumento da função gama em (3) deve ser maior que 1 para que a aproximação seja aplicável. Entretanto, pode-se utilizar a fórmula da reflexão, apresentada a seguir, para obter valores da função gama para argumentos reais negativos

$$\Gamma(1 - z) = \frac{\pi z}{\Gamma(z + 1) \sin(\pi z)}.$$

2.2 A distribuição gama

A distribuição gama é uma generalização da distribuição exponencial com média $1/\beta$, para $\beta > 0$. Uma variável aleatória gama pode ser considerada uma variável aleatória contínua de tempo de espera, em que os eventos são gerados por um processo de Poisson. Além do fato de a distribuição exponencial ser um caso particular da gama, a soma de variáveis aleatórias exponenciais independentes possui distribuição gama, o que faz com que essas duas distribuições sejam frequentemente discutidas juntas (MOOD; GRAYBILL; BOES, 1974).

Ao assumir que Y_1, \dots, Y_α são variáveis aleatórias exponenciais independentes, sendo a função geradora de momentos para Y_i

$$m_{Y_i}(t) = \frac{\beta}{\beta - t},$$

tem-se que a função geradora de momentos para $\sum_{i=1}^{\alpha} Y_i$ é dada por

$$m_{\sum Y_i}(t) = \prod_{i=1}^{\alpha} m_{Y_i}(t) = \left(\frac{\beta}{\beta - t} \right)^{\alpha},$$

que é a função geradora de momentos de uma distribuição gama com parâmetros α e β .

A função densidade de uma variável aleatória X com distribuição gama é

$$f(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} e^{-x/\beta} x^{\alpha-1}, \quad (4)$$

para $x > 0$, $\alpha > 0$ e $\beta > 0$. Seus parâmetros α e β são chamados de parâmetros de forma e de escala, respectivamente. Com a variação de α há grandes alterações na forma da distribuição, enquanto que com a de β a distribuição assume formas semelhantes, mas em diferentes intervalos (BEARZOTI, 1998). Se $\alpha = 1$ em (4), a função densidade gama especializa-se na função densidade exponencial e se α é um número inteiro positivo, a distribuição é conhecida como distribuição de Erlang.

Pode-se definir, ainda, a distribuição gama de três parâmetros, cuja função densidade é dada por

$$f(x|\alpha, \beta, \gamma) = \frac{1}{\beta\Gamma(\alpha)} e^{-(x-\gamma)/\beta} \left(\frac{x-\gamma}{\beta} \right)^{\alpha-1}$$

para $\alpha > 0$, $\beta > 0$ e $x > \gamma$, em que γ é o parâmetro de locação.

A distribuição gama padrão ocorre quando $\beta = 1$ e $\gamma = 0$ e possui função densidade

$$f(x|\alpha) = \frac{1}{\Gamma(\alpha)} e^{-x} x^{\alpha-1}$$

e função de distribuição acumulada

$$F(x|\alpha) = I_x(\alpha) = \frac{1}{\Gamma(\alpha)} \int_0^x e^{-t} t^{\alpha-1} dt. \quad (5)$$

De acordo com Knüsel e Bablok (1996), a função $\Gamma(\alpha)$ é a função gama completa,

enquanto a integral, $F(x|\alpha)$, é conhecida como função gama incompleta.

A obtenção da gama padrão a partir de uma variável aleatória Y com distribuição gama com três parâmetros se dá de maneira simples pelo método Jacobiano da transformação. Seja

$$f_X(x) = \frac{1}{\beta\Gamma(\alpha)} e^{-(x-\gamma)/\beta} \left(\frac{x-\gamma}{\beta} \right)^{\alpha-1},$$

fazendo a transformação

$$y = \frac{x-\gamma}{\beta},$$

tem-se que

$$x = \beta y + \gamma.$$

O jacobiano da transformação é o determinante da derivada de primeira ordem de x em relação a y e é dado por

$$J = \left| \frac{\partial x}{\partial y} \right| = |\beta|$$

para $\beta > 0$. Assim, a distribuição de Y é

$$\begin{aligned} f_Y(y) &= f_X(x) \times \text{abs}(J) \\ &= \frac{1}{\beta\Gamma(\alpha)} e^{-(\beta y + \gamma - \gamma)/\beta} \left(\frac{\beta y + \gamma - \gamma}{\beta} \right)^{\alpha-1} \times \beta \\ &= \frac{1}{\Gamma(\alpha)} e^{-y} y^{\alpha-1} \end{aligned}$$

para $y > 0$ e $\alpha > 0$, que é a função densidade gama padrão.

Krishnamoorth (2006) afirma que o grau de assimetria da distribuição gama diminui à medida que α aumenta, como mostrado na Figura 2. Além disso, para grandes valores de α , a quantidade $(x - \alpha)/\sqrt{\alpha}$ é aproximadamente distribuída como uma variável aleatória normal padrão.

As propriedades da função gama incompleta são apresentadas a seguir,

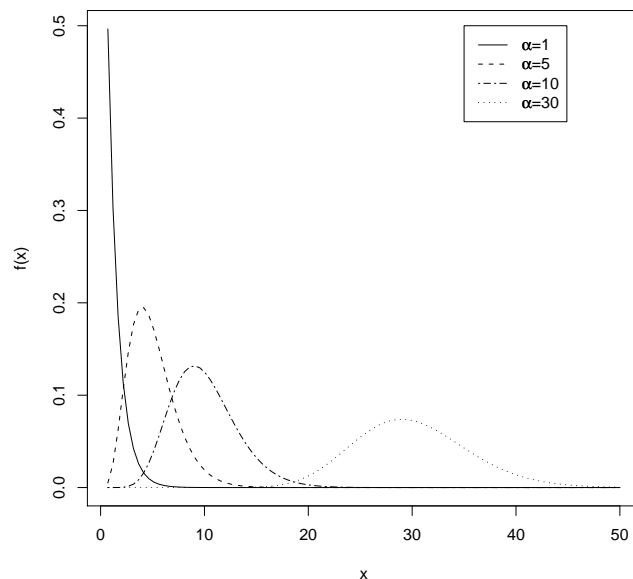


Figura 2 Função densidade gama padrão

considerando n um número inteiro qualquer. Para isso, é necessário definir as seguintes notações a partir da função gama completa

$$\begin{aligned}\Gamma(\alpha) &= \gamma(\alpha, x) + \Gamma(\alpha, x) \\ &= \int_0^x e^{-t} t^{\alpha-1} dt + \int_x^{\infty} e^{-t} t^{\alpha-1} dt,\end{aligned}$$

sendo $\gamma(\alpha, x)$ a função gama incompleta não normalizada e $\Gamma(\alpha, x)$ o complemento da função gama incompleta não normalizada.

Dividindo ambos os lados por $\Gamma(\alpha)$ na expressão anterior, obtém-se

$$\begin{aligned}\frac{\Gamma(\alpha)}{\Gamma(\alpha)} &= \frac{1}{\Gamma(\alpha)} \int_0^x e^{-t} t^{\alpha-1} dt + \frac{1}{\Gamma(\alpha)} \int_x^{\infty} e^{-t} t^{\alpha-1} dt \\ 1 &= P(\alpha, x) + Q(\alpha, x),\end{aligned}$$

sendo $P(\alpha, x)$ a função gama incompleta, $I_x(\alpha)$, e $Q(\alpha, x)$ o complemento da

função gama incompleta, $1 - I_x(\alpha)$. Assim, as propriedades são definidas por:

$$\begin{aligned}
 \gamma(\alpha + 1, x) &= \alpha\gamma(\alpha, x) - x^\alpha e^{-x} \\
 \gamma(n, x) &= (n - 1)! \left[1 - e^{-x} \sum_{r=0}^{n-1} \frac{x^r}{r!} \right] \\
 \Gamma(\alpha + 1, x) &= \alpha\Gamma(\alpha, x) - x^\alpha e^{-x} \\
 \Gamma(n, x) &= (n - 1)! e^{-x} \sum_{r=0}^{n-1} \frac{x^r}{r!} \\
 P(\alpha + 1, x) &= P(\alpha, x) - \frac{x^\alpha e^{-x}}{\Gamma(\alpha + 1)} = P(\alpha, x) - \frac{x^\alpha e^{-x}}{\alpha\Gamma(\alpha)} \quad (6) \\
 P(n, x) &= \frac{(n - 1)! \left[1 - e^{-x} \sum_{r=0}^{n-1} \frac{x^r}{r!} \right]}{\Gamma(n)} = 1 - e^{-x} \sum_{r=0}^{n-1} \frac{x^r}{r!}.
 \end{aligned}$$

Krishnamoorth (2006) apresenta dois métodos computacionais para encontrar probabilidades de uma variável aleatória com distribuição gama. Segundo o autor, para computar $P(X \leq x)$ quando $\alpha > 0$ e $\beta = 1$, ou seja, a gama padrão, pode-se usar a série de Pearson dada por

$$P(X \leq x) = e^{(-x)} x^\alpha \sum_{i=0}^{\infty} \frac{1}{\Gamma(\alpha + 1 + i)} x^i$$

ou usar a fração continuada

$$P(X > x) = \frac{e^{(-x)} x^\alpha}{\Gamma(\alpha)} \left(\frac{1}{x + 1 - \alpha} - \frac{1(1 - \alpha)}{x + 3 - \alpha} + \frac{2(2 - \alpha)}{x + 5 - \alpha} - \dots \right).$$

A primeira série converge mais rapidamente para $x < \alpha + 1$ e a segunda para $x \geq \alpha + 1$.

2.3 Método de Newton-Raphson e método da secante

A maioria das funções de distribuição envolve integrais complexas e o cálculo de sua inversa não possui soluções analíticas explícitas. A função de dis-

tribuição acumulada gama possui integral que pode ser calculada por artifícios especiais. Neste trabalho, métodos numéricos foram necessários para a obtenção das funções inversas da função de distribuição a fim de calcular o valor de x e do parâmetro de não-centralidade, dado que os outros parâmetros são conhecidos.

Vários métodos numéricos são apresentados por Ruggiero e Lopes (1996) para resolução de equações não lineares. Segundo as autoras, o princípio desses métodos é o de partir de uma aproximação inicial para a raiz da equação e, em seguida, refinar essa aproximação por meio de um processo iterativo.

O método iterativo de Newton-Raphson - utilizado para estimar as raízes de uma função - é basicamente uma aproximação de Taylor para a função $f(x)$, baseada nos arredores do ponto x_0 . Geometricamente, é como aproximar um arco da curva por uma reta tangente, traçada a partir de um ponto da curva. Isso faz com que o método seja conhecido também como método das tangentes (CAMPOS, 1998).

Para desenvolver o método de Newton-Raphson, toma-se um ponto da função e calcula-se sua derivada, o valor do intercepto no eixo das abscissas e o valor da função no ponto. O processo é repetido até que uma das raízes da função seja encontrada. A fórmula de recorrência do método é dada por

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)},$$

em que $f'(x_k)$ é a derivada da função em x_k e k indica a k -ésima iteração do algoritmo. Esse método é considerado um dos melhores métodos para obter aproximações de raízes de uma determinada função real, já que oferece rápida convergência, principalmente se a estimativa inicial estiver próxima da raiz da função. De acordo com Campos (1998), o valor dessa estimativa inicial (x_0) deve ser um ponto no qual a função tenha o mesmo sinal de sua derivada segunda. Barroso et al. (1987) afirmam que a condição para a convergência do método de Newton-Raphson é que $f'(x)$ e $f''(x)$ sejam não nulas e preservem o sinal no intervalo abrangido pela função e x_0 seja tal que $f'(x_0) \cdot f''(x_0) > 0$.

O método de Newton-Raphson apresenta uma grande desvantagem que é a obtenção da primeira derivada da função quando sua forma funcional for complexa e o cálculo do valor numérico dessa derivada a cada iteração. Uma alternativa para

esse problema surge com o método da secante, que estima essa primeira derivada utilizando a aproximação

$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}},$$

em que x_k e x_{k-1} são aproximações para as raízes. Nesse caso, tem-se a fórmula de recorrência

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}.$$

As condições para a convergência de ambos os métodos discutidos são praticamente as mesmas. Ruggiero e Lopes (1996) afirmam que, considerando como ideal o método em que a convergência estivesse assegurada, a ordem dessa convergência fosse alta e os cálculos por iteração fossem simples, o método de Newton-Raphson é o mais indicado sempre que for possível verificar as condições de convergência e quando o cálculo da derivada primeira não for muito elaborado. Nos casos em que a obtenção de $f'(x)$ for complexa, é aconselhável usar o método da secante.

2.4 Distribuições não-centrais

As distribuições não-centrais são vistas como misturas de uma distribuição discreta com outra contínua e suas funções de distribuição são dadas, de forma geral, por

$$P(X \leq x) = \sum_{i=0}^{\infty} P(Y = i|\delta) F_{Z_i}(x; \theta), \quad (7)$$

em que $P(Y = i|\delta)$ é a função de probabilidade discreta e $F_{Z_i}(x; \theta)$ a função de distribuição de probabilidade contínua, sendo ambas definidas de acordo com a função de distribuição não-central subjacente.

Patnaik (1949) enfatiza a importância das distribuições χ^2 e F não-centrais e suas aplicações na obtenção das funções de poder do teste χ^2 e de testes de razão de variâncias. Além disso, o autor sugere aproximações para as integrais de

probabilidades dessas distribuições.

A distribuição qui-quadrado não-central foi obtida por Fisher (1928 citado por PATNAIK, 1949) como um caso específico da distribuição do coeficiente de correlação múltiplo. Segundo Ruben (1974), embora essa distribuição - conhecida por muitos como uma generalização da distribuição Rayleigh - tenha sido bastante estudada por estatísticos e não estatísticos, algumas de suas propriedades básicas pareciam ser negligenciadas. Em seu trabalho, são apresentadas algumas relações recursivas para a densidade de probabilidade e função de distribuição de variáveis aleatórias qui-quadrado e gama não-centrais.

De acordo com Johnson, Kotz e Balakrishnan (1970), em geral, se U_1, U_2, \dots, U_ν são variáveis aleatórias normais padrão independentes e $\mu_1, \mu_2, \dots, \mu_\nu$ são constantes, a distribuição de

$$\sum_{j=1}^{\nu} (U_j + \mu_j)^2$$

é chamada de distribuição χ^2 não-central com ν graus de liberdade e parâmetro de não-centralidade $\delta = \sum_{j=1}^{\nu} \mu_j^2$. Quando $\delta = 0$, a distribuição não-central especifica-se na distribuição χ^2 . Neste trabalho, as distribuições qui-quadrado central e não-central são denotadas por χ_ν^2 e $\chi_\nu^2(\delta)$, respectivamente.

Como anteriormente mencionado, a distribuição qui-quadrado não-central é utilizada no cálculo do poder de alguns testes. Sua aplicação é geralmente útil na aproximação para o poder de testes de χ^2 aplicados a tabelas de contingência e de certos testes não-paramétricos. Johnson, Kotz e Balakrishnan (1970) ainda ressaltam as aplicações dessa distribuição em vários aspectos da teoria financeira.

Uma variável aleatória X com distribuição qui-quadrado não-central possui função densidade

$$\sum_{i=0}^{\infty} \frac{e^{-\delta/2} (\delta/2)^i}{i!} \frac{e^{-x/2} x^{\nu/2+i-1}}{2^{\nu/2+i} \Gamma(\nu/2 + i)}, \quad \text{para } x > 0$$

e função de distribuição de probabilidades

$$P(\chi_{\nu}^2(\delta) \leq x) = \sum_{i=0}^{\infty} \frac{e^{-\delta/2}(\delta/2)^i}{i!} I_{x/2}(\nu/2 + i). \quad (8)$$

Knüsel e Bablok (1996) discutem a relação entre as distribuições gama e beta e as distribuições clássicas qui-quadrado, F e t, centrais e não-centrais. A relação entre as distribuições qui-quadrado e gama não-centrais pode ser facilmente notada na equação (8), em que a quantidade $I_x(\alpha)$ foi definida em (5) como a função gama incompleta. Então, se uma variável aleatória Y possui distribuição qui-quadrado com parâmetro de não-centralidade δ , a variável aleatória $X = Y/2$ possui distribuição gama incompleta não-central com parâmetro $\alpha = \nu/2$ e o mesmo parâmetro de não-centralidade δ . Considera-se a distribuição gama incompleta não-central como uma generalização da distribuição qui-quadrado não-central.

A função de distribuição de uma variável aleatória X que possui distribuição gama incompleta não-central é definida por

$$I_x(\alpha, \delta) = P(X \leq x) = \sum_{i=0}^{\infty} \frac{e^{-\delta/2}(\delta/2)^i}{i!} I_x(\alpha + i),$$

para $\alpha > 0$ e parâmetro de não-centralidade $\delta > 0$. A distribuição gama incompleta não-central é uma mistura de uma distribuição gama central com uma Poisson $P(Y = i|\delta)$ (KNÜSEL; BABLOK, 1996).

Associações também podem ser feitas entre as distribuições F, t e beta. Todas as três distribuições centrais e não-centrais comumente usadas (qui-quadrado, F e t) são casos particulares das distribuições gama e beta centrais e não-centrais. As distribuições F e t não-centrais são casos particulares da distribuição beta não-central, enquanto a qui-quadrado não-central é um caso particular da distribuição gama não-central. Destaca-se, então, a importância do desenvolvimento e aperfeiçoamento de métodos computacionais para a aproximação das funções gama e beta não-centrais.

2.5 Algoritmos para computação de funções de distribuições não-centrais

Vários algoritmos para computação de funções de distribuições não-centrais estão disponíveis na literatura. Benton e Krishnamoorthy (2003) afirmam que a maioria desses algoritmos se baseia em dois métodos. No primeiro método, calcula-se $P(Y = 0|\delta)$ e $F_{Z_0}(x; \theta)$ em (7) e computa-se recursivamente as probabilidades discreta e contínua para $i = 1, 2, \dots$. Esse método poderia ser facilmente adotado, já que as principais funções de distribuição acumulada possuem relações recursivas. Entretanto, os autores afirmam que quando a média da variável aleatória discreta Y é grande, $P(Y = 0|\delta)$ será muito pequeno, ocasionando um problema computacional chamado de *underflow error*. Esse erro ocorre quando o resultado de uma computação é menor em magnitude do que a menor quantidade que o computador pode armazenar, fazendo-o substituir por zero o valor que seria obtido pelo cálculo exato. Além disso, o tempo computacional dos algoritmos baseados nesse método aumenta drasticamente com o aumento da média de Y e eles podem ser considerados ineficientes quando usados como algoritmos auxiliares.

O segundo método apresenta uma alternativa para esses problemas ao iniciar a computação da série no ponto k , que é a média da variável Y tomando-se o número inteiro mais próximo. Isso é feito porque, em geral, a série dominante nas distribuições não-centrais é aquela relacionada à probabilidade discreta $P(Y = i|\delta)$, que possui o máximo perto da média de Y . As quantidades em (7) também são computadas recursivamente a partir dos valores iniciais de $P(Y = k|\delta)$ e $F_{Z_k}(x; \theta)$.

Benton e Krishnamoorthy (2003) apresentam um algoritmo geral para esse método. Segundo esses autores, primeiramente são computados $P(Y = k|\delta)$ e $F_{Z_k}(x; \theta)$ em (7). Em seguida, computam-se $P(Y = k - i|\delta)$ e $P(Y = k + i|\delta)$ e $F_{Z_{k-i}}(x; \theta)$ e $F_{Z_{k+i}}(x; \theta)$, $i = 1, 2, \dots, k$, recursivamente, utilizando os valores iniciais $P(Y = k|\delta)$ e $F_{Z_k}(x; \theta)$. Nesse passo, a série pode convergir nas k primeiras iterações, quando k é grande. Nesses casos, a série deve ser terminada em (7) quando a soma das probabilidades associadas à distribuição discreta é próxima de 1. Se a convergência não é alcançada nesse passo, os valores de $P(Y = 2k + i|\delta)$ e $F_{Z_{2k+i}}(x; \theta)$ devem ser computados recursivamente para $i = 1, 2, \dots$ até que

$[1 - \sum_{j=0}^{2k+i} P(Y = j|\delta)F_{Z_{2k+i+1}}(x; \theta)]$ seja menor que uma tolerância de erro especificada.

Essa abordagem alternativa foi sugerida por vários autores, dentre eles Chattamvelli e Shanmugam (1997), que propuseram um método para computação da função de distribuição acumulada beta não-central, usando relações recursivas *backward* e *forward* e iniciando o cômputo em $\delta/2$, em que δ é o parâmetro de não-centralidade da função. Seus algoritmos são alternativas aos algoritmos *AS226* e *ASR84*, propostos por Frick (1990) e Lenth (1987), respectivamente, e, quando a eles comparados, mostrou-se mais eficiente em relação a acurácia e tempo de processamento.

Benton e Krishnamoorthy (2003) também se baseiam nesse segundo método na proposição de algoritmos para obtenção das funções de distribuição t e qui-quadrado não-centrais e quadrado do coeficiente de correlação múltipla amostral.

Baharev e Kemény (2008) discutem o cálculo recursivo da função de distribuição acumulada (fdpa) a partir de um intervalo delimitado por k_1 e k_2 (obtidos a partir do modelo probabilístico discreto da mistura) como forma de evitar problemas computacionais como o *underflow error*. Esses limites são estabelecidos utilizando-se os critérios dados por $P(X < k_1) \leq \varepsilon$ e $P(X > k_2) \leq \varepsilon$, em que ε é um valor pequeno tal que o erro do cômputo da fdpa seja menor ou igual a 2ε e k_1 e k_2 o maior e o menor inteiro, respectivamente, que tornam verdadeiras as afirmativas probabilísticas anteriores. Ainda assim, os autores assumem que uma grande dificuldade desse método é que os primeiros termos da série precisam ser negligenciados se o parâmetro de não-centralidade for muito grande.

Esse método intervalar foi utilizado por Knüsel e Bablok (1996) para a obtenção da função gama não-central, mas possui a grande restrição de ter que estabelecer os limites k_1 e k_2 . Para isso, é necessário inverter a função de distribuição de probabilidade do modelo discreto da mistura.

Baharev e Kemény (2008) também propuseram a obtenção do parâmetro de não-centralidade das distribuições F e beta não-centrais de forma iterativa baseada no método de Newton-Raphson. Seus cálculos foram feitos de forma eficiente, já que a função de distribuição e sua derivada foram computadas simultaneamente no mesmo *loop*. Para aplicar o método de Newton-Raphson, os

autores sugerem duas formas de se obter os valores iniciais de δ , sendo uma delas a aproximação normal proposta por Severo e Zelen (1960 citados por BAHAREV; KEMÉNY, 2008). Segundo Baharev e Kemény (2008), essas aproximações fornecem bons valores iniciais para a iteração e a convergência pode ser alcançada rapidamente.

Knüsel e Bablok (1996) e Ruben (1974) propuseram aproximações para a obtenção da função gama incompleta não-central, sem, no entanto, que os procedimentos numéricos tivessem sido implementados. Segundo Chen (2005), as aproximações para a densidade e inversa da função de distribuição gama não-central não estão disponíveis na literatura. Esse mesmo autor apresenta soluções para esses problemas usando equações diferenciais não-lineares do Sistema S. Infelizmente, os códigos utilizados pelo autor não estão disponíveis, ao contrário do que é afirmado em seu trabalho.

Sendo assim, a distribuição gama incompleta não-central não está disponível nos programas estatísticos convencionais como SAS, R, Statistica, entre outros. Para o caso particular da distribuição qui-quadrado não-central existem soluções numéricas de alta qualidade (BENTON; KRISHNAMOORTHY, 2003) e implementações em programas como R (R DEVELOPMENT CORE TEAM, 2008).

3 MATERIAL E MÉTODOS

3.1 A função de distribuição gama incompleta não-central

Como Benton e Krishnamoorthy (2003) não apresentaram um algoritmo específico para a função gama incompleta não-central essa é uma das finalidades deste trabalho, a partir do algoritmo geral apresentado pelos autores. Fórmulas recursivas foram utilizadas para obtenção das probabilidades associadas à variável aleatória discreta e à variável aleatória contínua da função apresentada em (1).

Para o caso discreto, pôde-se expressar as probabilidades $P(Y = k + 1|\delta)$ e $P(Y = k - 1|\delta)$ em função de $P(Y = k|\delta)$ facilmente por

$$P(Y = k + 1|\delta) = \frac{\delta/2}{k + 1} P(Y = k|\delta), \quad (9)$$

$$P(Y = k - 1|\delta) = \frac{k}{\delta/2} P(Y = k|\delta), \quad (10)$$

para $k \geq 0$ em (9) e $k > 0$ em (10), sendo $k = \lceil \delta/2 \rceil$ o menor inteiro que não é menor que $\delta/2$, correspondendo ao inteiro mais próximo da média da distribuição Poisson.

Definindo-se k como um ponto qualquer da soma infinita, a computação recorrente progressiva foi iniciada a partir de um ponto dado pelo valor no qual a probabilidade associada a variável discreta é máxima. Assim, o valor de k é dado pelo inteiro mais próximo à metade do valor do parâmetro de não-centralidade δ .

Para obter a computação recorrente progressiva da função gama incompleta central - ou seja, $I_x(\alpha + i)$ - foi utilizado o resultado apresentado em (6), que é dado por

$$I_x(\alpha + 1) = I_x(\alpha) - \frac{x^\alpha e^{-x}}{\Gamma(\alpha + 1)}$$

em que pôde ser aplicada a transformação de α para $\alpha - 1$ para obter a computação

recorrente regressiva

$$I_x(\alpha - 1) = I_x(\alpha) + \frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}.$$

A função gama incompleta central foi computada diretamente apenas uma vez no ponto k por $I_x(\alpha + k)$. Assim, a computação recorrente progressiva a partir desse ponto é dada por

$$I_x(\alpha + k + 1) = I_x(\alpha + k) - \frac{x^{\alpha+k}e^{-x}}{\Gamma(\alpha + k + 1)}$$

e a computação recorrente regressiva por

$$I_x(\alpha + k - 1) = I_x(\alpha + k) + \frac{x^{\alpha+k-1}e^{-x}}{\Gamma(\alpha + k)}.$$

Definindo-se a função $g(x|\alpha)$ por

$$g(x|\alpha) = \frac{x^\alpha e^{-x}}{\Gamma(\alpha + 1)},$$

pôde-se definir a fórmula recorrente progressiva por

$$g(x|\alpha + 1) = g(x|\alpha) \frac{x}{\alpha + 1}$$

e a fórmula recorrente regressiva por

$$g(x|\alpha - 1) = g(x|\alpha) \frac{\alpha}{x}.$$

No k -ésimo ponto, tem-se a função $g(x|\alpha + k)$ dada por

$$g(x|\alpha + k) = \frac{x^{\alpha+k}e^{-x}}{\Gamma(\alpha + k + 1)} \quad (11)$$

que foi computada inicialmente. A partir de (11), foi possível definir as funções

progressiva e regressiva, $g_x^p(\alpha + k)$ e $g_x^r(\alpha + k)$, por

$$g_x^p(\alpha + k) = g(x|\alpha + k) \frac{\alpha + k}{x} \quad (12)$$

e

$$g_x^r(\alpha + k) = g(x|\alpha + k), \quad (13)$$

e obter as seguintes fórmulas recorrentes no i -ésimo ponto da soma

$$I_x(\alpha + k + i) = I_x(\alpha + k + i - 1) - g_x^p(\alpha + k + i - 1) \frac{x}{\alpha + k + i - 1} \quad (14)$$

e

$$I_x(\alpha + k - i) = I_x(\alpha + k - i + 1) + g_x^r(\alpha + k - i + 1) \frac{\alpha + k + 1 - i}{x}.$$

Para aplicar o algoritmo, inicialmente especificou-se k como o valor inteiro $k = \lceil \delta/2 \rceil$. Em seguida, a partir do ponto k , computou-se a soma (14) de forma recursiva crescente utilizando (9) e (12) e decrescente utilizando (10) e (13). A soma foi computada de forma iterativa até que não houvesse alterações no valor computado em relação ao valor obtido no passo anterior, para uma tolerância de convergência pré-estabelecida. A convergência pode ser alcançada nos k primeiros passos. Caso não seja atingida nos k primeiros passos, deve-se continuar a soma além do passo $2k$ até que não haja mais modificações no valor computado para uma dada precisão pré-estipulada. Nesse último caso, utiliza-se somente a recorrência progressiva, pois não há mais termos além do termo 0. É necessário enfatizar que foram necessários algoritmos auxiliares para o cômputo da função gama incompleta central e função gama completa.

Em pseudocódigo, o algoritmo combinado proposto é dado por

1. Entrar com valores $x > 0$, $\alpha > 0$ e $d = \delta > 0$;
2. especificar $errtol = 1 \times 10^{-15}$ para a tolerância do erro e $maxitr$ para o número máximo de iterações;
3. obter $k = \lceil \delta/2 \rceil$, ou seja, parte inteira mais próxima de $d/2$;

4. fazer $a = \alpha + k$;
5. calcular o valor da função gama incompleta central para o k -ésimo termo da soma e denominá-lo de "gamac" e "gamad" para computação recorrente progressiva e regressiva, respectivamente:

$$gamac = I_x(a) \text{ e } gamad = gamac;$$
6. obter a inicialização da função $g(x|a)$ de (11) e modificá-la pelas expressões (12) e (13) para computação recorrente progressiva e regressiva e denominá-las de gxc e gxd:

$$gxc = x^a e^{-x} / \Gamma(a + 1) \times a/x \text{ e } gxd = gxc \times x/a;$$
7. calcular o k -ésimo termo da probabilidade Poisson e denominá-la "ppoic" e "ppoid" para computação recursiva crescente e decrescente por:

$$ppoic = \exp(-d/2 + k \ln(d/2) - \ln \Gamma(k + 1)) \text{ e } ppoid = ppoic;$$
8. calcular o resto das probabilidades Poisson acumuladas por $remain = 1 - ppoic$;
9. calcular o valor da função gama incompleta não-central por:

$$cdf = ppoic \times gamac;$$
10. fazer $i = 1$;
11. repetir;
 11. a. realizar a soma crescente:

$$gxc = gxc \times x/(a + i - 1);$$
 11. b. $gamac = gamac - gxc$;
 11. c. $ppoic = ppoic \times (d/2)/(k + i)$;
 11. d. $cdf = cdf + ppoic \times gamac$;
 11. e. $error = remain \times gamac$;
 11. f. $remain = remain - ppoic$;
 11. g. realizar a soma regressiva a partir de k , se houver termos remanescentes:
 se $(i > k)$ então faça:
 se $(error \leq errortol)$ ou $(i > maxitr)$ vá ao passo 12;
 faça $i = i + 1$ e vá para o passo 11;
 fim de: se $(i > k)$;
 11. h. se, $i \leq k$, execute até o passo 11p; caso contrário vá para o passo 11o.
 11. i. $gxd = gxd \times (a + 1 - i)/x$;
 11. j. $gamad = gamad + gxd$;
 11. k. $ppoid = ppoid \times (k - i + 1)/(d/2)$;
 11. l. $cdf = cdf + ppoid \times gamad$;
 11. m. $remain = remain - ppoid$;

- 11. n. se ($remain \leq errortol$) ou ($i > maxitr$) vá para o passo 12;
- 11. o. $i = i + 1$;
- 11. p. vá para o passo 11;
- 12. retorne cdf.

3.2 Inversa da função de distribuição em relação a δ

Como discutido inicialmente, aproximações para a função densidade e inversa da função de distribuição gama incompleta não-central não estão disponíveis na literatura nem nos programas de análise estatística convencionais. Além disso, não existem aproximações para o cálculo do parâmetro de não-centralidade δ dados x , α e $0 \leq p \leq 1$ a partir da solução de $I_x(\alpha, \delta) = p$. A implementação desse algoritmo foi realizada combinando os algoritmos de Baharev e Kemény (2008) e Benton e Krishnamoorthy (2003) para se determinar o parâmetro de não-centralidade δ .

Como a base do algoritmo é o método de Newton-Raphson, foi necessário obter a derivada de primeira ordem da função gama não-central em relação a δ . Esse processo foi necessário para computar no mesmo processo iterativo o valor da função gama incompleta e da sua derivada de primeira ordem, otimizando o algoritmo. A derivada de primeira ordem de $I_x(\alpha, \delta)$ em relação a δ é

$$\frac{dI_x(\alpha, \delta)}{d\delta} = -\frac{1}{2} \sum_{i=0}^{\infty} \frac{e^{-\delta/2} (\delta/2)^i}{i!} \frac{x^{\alpha+i} e^{-x}}{\Gamma(\alpha + i + 1)}. \quad (15)$$

Para a computação recorrente progressiva e regressiva dessa derivada, tomou-se o k -ésimo termo da soma (15) e, considerando apenas o segundo fator, definiu-se a função auxiliar $h(x|k, \alpha)$ por

$$h(x|k, \alpha) = \frac{x^{\alpha+k} e^{-x}}{\Gamma(\alpha + k + 1)}.$$

Nos cálculos recorrentes progressivos de $h(x|k, \alpha)$ obteve-se

$$\begin{aligned} h(x|k+1, \alpha) &= \frac{x^{\alpha+k+1}e^{-x}}{\Gamma(\alpha+k+2)} = \frac{x^{\alpha+k}e^{-x}}{\Gamma(\alpha+k+1)} \frac{x}{(\alpha+k+1)} \\ &= h(x|k, \alpha) \frac{x}{(\alpha+k+1)} \end{aligned}$$

e nos cálculos recorrentes regressivos

$$\begin{aligned} h(x|k-1, \alpha) &= \frac{x^{\alpha+k-1}e^{-x}}{\Gamma(\alpha+k)} = \frac{x^{\alpha+k}e^{-x}}{\Gamma(\alpha+k+1)} \frac{(\alpha+k)}{x} \\ &= h(x|k, \alpha) \frac{(\alpha+k)}{x}. \end{aligned}$$

Para utilizar o método de Newton-Raphson, foi necessário obter um valor inicial do parâmetro de não-centralidade, ou seja, δ_0 , por meio de uma aproximação da função gama incompleta não-central. Para isso, foi utilizada uma aproximação dos quantis x_p da função gama incompleta não-central a partir de uma modificação de uma aproximação normal para a distribuição qui-quadrado não-central dada por

$$x_p = \frac{(\alpha + \delta) \left[z_p + \sqrt{(2\alpha + \delta)^2 / (\alpha + \delta) - 1} \right]^2}{2\alpha + \delta}. \quad (16)$$

A solução da equação (16) foi feita pelo método da secante, já que a obtenção da sua derivada é muito complexa. O valor inicial δ_0 de δ foi obtido a partir dessa solução. Dada a probabilidade acumulada p , pôde-se obter z_p por $z_p = \Phi^{-1}(p)$. Seja a função $f(\delta)$ a partir de (16)

$$f(\delta) = \frac{(\alpha + \delta) \left[z_p + \sqrt{(2\alpha + \delta)^2 / (\alpha + \delta) - 1} \right]^2}{2\alpha + \delta} - x_p, \quad (17)$$

desejou-se encontrar a raiz δ , tal que $f(\delta) = 0$, dados x_p , z_p e α . Na sequência, é apresentado o algoritmo para a obtenção do valor inicial δ_0 .

1. Iniciar $dn0 = 0,5$ e $dn1 = 1,0$;
2. fixar $errortol = 1 \times 10^{-15}$ e $N_{max} = 1000$;
3. calcular $f_{dn0} = f(dn0)$ e $f_{dn1} = f(dn1)$ utilizando (17);

4. fazer $it = 1$ e repetir os passos 5 a 9 N_{max} vezes;
5. $dn2 = dn1 - fdn1 \times (dn1 - dn0)/(fdn1 - fdn0)$; se $dn2 < 0$, então $dn2 = dn1/2$;
6. $dn0 = dn1$ e $fdn0 = fdn1$;
7. $dn1 = dn2$ e $fdn1 = f(dn1)$
8. $it = it + 1$;
9. se $|dn2 - dn0| \leq \text{errortol} \times dn0$, então retorne $dn2$ e saia; senão verifique se $it \leq N_{max}$.
Se $it \leq N_{max}$ retorne ao passo 5; senão vá para o passo 10;
10. imprima a mensagem de erro: "processo iterativo não convergiu em N_{max} passos."

O algoritmo combinado proposto, em pseudocódigo, para se determinar o parâmetro de não-centralidade δ da função gama não-central é dado por

1. Entrar com $x > 0$, $\alpha > 0$ e $0 < \text{prob} < 1$;
2. especificar $\text{errortol} = 1 \times 10^{-15}$ para a tolerância do erro e maxitr para o máximo de iterações;
3. obter uma estimativa inicial de δ , denominada $\delta_{\text{new}} = \delta_0$; utilizar o método da secante na aproximação normal, conforme apresentado na sequência;
4. repetir os passos 5 a 17 N_{max} vezes; iniciar $it = 1$;
5. fazer $\delta = \delta_{\text{new}}$ e $k = \lceil \delta/2 \rceil$;
6. fazer $a = \alpha + k$;
7. $\text{gamac} = I_x(a)$ e $\text{gamad} = \text{gamac}$;
8. $\text{gxd} = \frac{1}{\Gamma(a+1)} x^a e^{-x}$ e $\text{gxc} = \text{gxd} \times a/x$;
9. $\text{hxc} = \text{gxd}$ e $\text{hxd} = \text{hxc}$;
10. $\text{ppoic} = \exp(-\delta/2 + k \ln(\delta/2) - \ln \Gamma(k+1))$ e $\text{ppoid} = \text{ppoic}$;
11. $\text{remain} = 1 - \text{ppoic}$;
12. $\text{cdf} = \text{ppoic} \times \text{gamac}$;
13. derivada no k -ésimo termo da soma: $g = \text{ppoic} \times \text{hxc}$;
14. fazer $i = 1$;
15. repetir passos 15a a 15t até a convergência;
 15. a. iniciar soma progressiva
 $\text{gxc} = \text{gxc} \times x/(a + i - 1)$;
 15. b. $\text{gamac} = \text{gamac} - \text{gxc}$;
 15. c. $\text{hxc} = \text{hxc} \times x/(a + i)$;
 15. d. atualizar as probabilidades Poisson por $\text{ppoic} = \text{ppoic} \times (\delta/2)/(k + i)$;

15. e. $cdf = cdf + ppoic \times gamac$;
15. f. $g = g + ppoic \times hxc$;
15. g. $error = remain \times gamac$;
15. h. $remain = remain - ppoic$;
15. i. realizar a soma regressiva, se houver termos remanescentes. Assim,
se $(i > k)$ então faça:
se $(error \leq errortol)$ ou $(i > maxitr)$ vá para passo 16;
faça $i = i + 1$ e vá para o passo 15;
fim de: se $(i > k)$;
15. j. senão, $(i \leq k)$, há termos remanescentes na soma regressiva e os passos 15k a 15t devem ser executados;
15. k. $gxd = gxd \times (a - i + 1)/x$;
15. l. $gamad = gamad + gxd$;
15. m. $hxd = hxd \times (a + 1 - i)/x$;
15. n. $ppoid = ppoid \times (k - i + 1)/(\delta/2)$;
15. o. $cdf = cdf + ppoid \times gamad$;
15. p. $g = g + ppoid \times hxd$;
15. q. $remain = remain - ppoid$;
15. r. se $(remain \leq errortol)$ ou $(i > maxitr)$ vá para 16;
15. s. $i = i + 1$;
15. t. vá para o passo 15;
16. se $\delta + 2(cdf - prob)/g \leq 0$, então $\delta_{new} = \delta/2$; senão $\delta_{new} = \delta + 2(cdf - prob)/g$ (passo do método de Newton-Raphson);
17. se $|\delta_{new} - \delta| \leq \delta \times tol$, em que $tol = 1 \times 10^{-15}$, então retorne δ_{new} e saia; caso contrário retorne para o passo 5 atualizando o contador de iterações ($it = it + 1$) do método de Newton-Raphson antes de retornar. Verifique se este contador ultrapassou o limite N_{max} . Se isso ocorreu vá para 18;
18. imprima a mensagem de erro: “processo iterativo não convergiu em N_{max} passos.”

3.3 Inversa da função de distribuição em relação a x

O processo para obter essa inversa da função de distribuição é muito similar ao apresentado anteriormente, para a obtenção do parâmetro de não-centralidade δ . A princípio, obteve-se a derivada de primeira ordem da função gama

incompleta não-central em relação a x , que é a sua função densidade, dada por

$$\begin{aligned} \frac{dI_x(\alpha, \delta)}{dx} &= \frac{d}{dx} \sum_{i=0}^{\infty} \frac{e^{-\delta/2} (\delta/2)^i}{i!} \left[\frac{1}{\Gamma(\alpha + i)} \int_0^x e^{-t} t^{\alpha+i-1} dt \right] \\ &= \sum_{i=0}^{\infty} \frac{e^{-\delta/2} (\delta/2)^i}{i!} \left[\frac{1}{\Gamma(\alpha + i)} e^{-x} x^{\alpha+i-1} \right]. \end{aligned} \quad (18)$$

Como no caso do parâmetro de não-centralidade, definiu-se uma função correspondente ao segundo fator de (18), no k -ésimo ponto da soma, por $h(x|\alpha + k)$. Expressando-a em termos da função $g(x)$ em (11), obteve-se

$$h(x|\alpha + k) = g(x|\alpha + k) \frac{(\alpha + k)}{x}.$$

As computações recorrentes progressiva e regressiva foram definidas, respectivamente, por

$$h(x|\alpha + k + 1) = h(x|\alpha + k) \frac{x}{\alpha + k}$$

e

$$h(x|\alpha + k - 1) = h(x|\alpha + k) \frac{(\alpha + k - 1)}{x}.$$

O valor inicial de x , x_0 , pôde ser obtido diretamente utilizando (16). O algoritmo para obtenção da inversa da função de distribuição gama incompleta não-central é dado por

1. Entrar com $d = \delta > 0$, $\alpha > 0$ e $0 < prob < 1$;
2. especificar $errortol = 1 \times 10^{-15}$ para a tolerância do erro e $maxitr$ para o máximo de iterações;
3. obter uma estimativa inicial de x , denominada $x_{new} = x_0$; utilizar a expressão da aproximação normal (16) para obter esse valor inicial x_0 ; especificar $k = \lceil \delta/2 \rceil$;
4. fazer $a = \alpha + k$;
5. repetir os passos 6 a 17 N_{max} vezes; iniciar $it = 1$;
6. fazer $x = x_{new}$;
7. $gamac = I_x(a)$ e $gamad = gamac$;

8. $gxd = \frac{1}{\Gamma(a+1)} e^{-x} x^a$ e $gxc = gxd \times a/x$;
9. $hxc = gxd \times \frac{a}{x}$ e $hxd = hxc$;
10. $ppoic = \exp(-d/2 + k \ln(d/2) - \ln \Gamma(k+1))$ e $ppoid = ppoic$;
11. $remain = 1 - ppoic$;
12. $cdf = ppoic \times gamac$;
13. derivada no k -ésimo termo da soma: $g = ppoic \times hxc$;
14. fazer $i = 1$;
15. repetir passos 15a a 15t até a convergência;
 15. a. iniciar soma progressiva
 $gxc = gxc \times x/(a+i-1)$;
 15. b. $gamac = gamac - gxc$;
 15. c. $hxc = hxc \times x/(a+i-1)$;
 15. d. atualizar as probabilidades Poisson por $ppoic = ppoic \times (d/2)/(k+i)$;
 15. e. $cdf = cdf + ppoic \times gamac$;
 15. f. $g = g + ppoic \times hxc$;
 15. g. $error = remain \times gamac$;
 15. h. $remain = remain - ppoic$;
 15. i. realizar a soma regressiva, se houver termos remanescentes. Assim,
se $(i > k)$ então faça:
se $(error \leq errortol)$ ou $(i > maxitr)$ vá para passo 16;
faça $i = i + 1$ e vá para o passo 15;
fim de: se $(i > k)$;
 15. j. senão, $(i \leq k)$, há termos remanescentes na soma regressiva e os passos 15k a 15t devem ser executados;
 15. k. $gxd = gxd \times (a-i+1)/x$;
 15. l. $gamad = gamad + gxd$;
 15. m. $hxd = hxd \times (a-i)/x$;
 15. n. $ppoid = ppoid \times (k-i+1)/(d/2)$;
 15. o. $cdf = cdf + ppoid \times gamad$;
 15. p. $g = g + ppoid \times hxd$;
 15. q. $remain = remain - ppoid$;
 15. r. se $(remain \leq errortol)$ ou $(i > maxitr)$ vá para 16;

15. s. $i = i + 1$;
15. t. vá para o passo 15;
16. se $x - (cdf - prob)/g \leq 0$, então $x_{\text{new}} = x/2$; senão $x_{\text{new}} = x - (cdf - prob)/g$ (passo do método de Newton-Raphson);
17. se $|x_{\text{new}} - x| \leq x \times tol$, em que $tol = 1 \times 10^{-15}$, então retorne x_{new} e saia; caso contrário retorne para o passo 6 atualizando o contador de iterações ($it = it + 1$) do método de Newton-Raphson antes de retornar. Verifique se este contador ultrapassou o limite N_{max} . Se isso ocorreu vá para 18;
18. imprima a mensagem de erro: “processo iterativo não convergiu em N_{max} passos.”

3.4 A função densidade da distribuição gama incompleta não-central

A função densidade gama não-central foi obtida na derivação de primeira ordem da função de distribuição gama incompleta não-central em relação a x (18). A densidade almejada é a própria derivada obtida. A função densidade resultante também é uma série, cujos valores foram computados a partir do valor inteiro mais próximo da média da distribuição discreta. O algoritmo resultante para a função densidade gama não-central é apresentado a seguir.

1. Entrar com valores $x > 0$, $\alpha > 0$ e $d = \delta > 0$;
2. especificar $errtol = 1 \times 10^{-15}$ para a tolerância do erro e $maxitr$ para o número máximo de iterações;
3. obter $k = \lceil \delta/2 \rceil$, ou seja, parte inteira mais próxima de $d/2$;
4. fazer $a = \alpha + k$;
5. $hxc = \frac{1}{\Gamma(\alpha + 1)} e^{-x} x^\alpha$;
6. $hxc = hxc \times \alpha/x$ e $hxd = hxc$;
7. $ppoic = \exp(-d/2 + k \ln(d/2) - \ln \Gamma(k + 1))$ e $ppoid = ppoic$;
8. $remain = 1 - ppoic$;
9. calcular o valor da função densidade no k -ésimo termo da soma por $g = ppoic \times hxc$;
10. fazer $i = 1$;
11. repetir passos 11a e 11n até a convergência;
 11. a. iniciar soma progressiva:
 $hxc = hxc \times x/(a + i - 1)$;
 11. b. atualizar as probabilidades Poisson por $ppoic = ppoic \times (d/2)/(k + i)$;
 11. c. $g = g + ppoic \times hxc$;

- 11. d. $error = remain \times g$;
 - 11. e. $remain = remain - ppoic$;
 - 11. f. realizar a soma regressiva a partir de k , se houver termos remanescentes:
 se $(i > k)$ então faça:
 se $(error \leq errortol)$ ou $(i > maxitr)$ vá ao passo 12;
 faça $i = i + 1$ e vá para o passo 11;
 fim de: se $(i > k)$;
 - 11. g. senão, $i \leq k$, há termos remanescentes na soma regressiva e os passos 11h a 11n devem ser executados:
 - 11. h. $hxd = hxd \times (a - i)/x$;
 - 11. i. $ppoid = ppoic \times (k - i + 1)/(d/2)$;
 - 11. j. $g = g + ppoic \times hxd$;
 - 11. k. $remain = remain - ppoic$;
 - 11. l. se $(remain \leq errortol)$ ou $(i > maxitr)$ vá para o passo 12;
 - 11. m. $i = i + 1$;
 - 11. n. vá para o passo 11;
12. retorne g , que é o valor da função densidade.

3.5 Implementação das rotinas e comparação dos resultados

Os algoritmos apresentados nas subseções anteriores foram implementados como rotinas do programa R e estão disponibilizados no apêndice A. As funções implementadas são dependentes das funções existentes no R: *pgamma*, para obtenção da função gama incompleta central, $I_x(\alpha)$; *dpois*, para obtenção da função de probabilidade Poisson; *qnorm*, para obtenção de quantis da distribuição normal padrão e *lgamma*, para obtenção do logaritmo neperiano da função gama completa, $\Gamma(\alpha)$.

Para avaliar a acurácia e a velocidade de processamento dos algoritmos implementados em R, foram obtidos resultados da função de distribuição, função densidade, quantis e parâmetro de não-centralidade da gama incompleta não-central e de seu caso particular, a qui-quadrado não-central, utilizando para isso as configurações dos argumentos utilizadas nos trabalhos de Benton e Krishnamoorthy (2003) e Chen (2005). Os resultados obtidos por esses autores foram utili-

zados para fins de comparação. Os resultados apresentados por Benton e Krishnamoorthy (2003) foram computados pelo método proposto por eles e pelo método intervalar e os apresentados por Chen (2005) foram computados pelo método baseado em equações diferenciais de segunda ordem do sistema S.

Em algumas situações, valores exatos foram apresentados. Esses valores foram obtidos por Chen (2005) em alguns casos e em outros computados por meio do *software* Maxima (MAXIMA, 2009). Para isso, os termos da série infinita (1) foram computados a partir de zero de forma recursiva. Considerando uma tolerância de 1×10^{-30} , a soma foi obtida acumulando-se os termos até que o resultados não se alterassem com a adição de cada novo termo. Todos os resultados computacionais foram obtidos utilizando um computador com processador Intel Core 2 Duo T7100, 1,8 GHz, 2 MB L2 cache e 1 GB DDR2.

As funções densidades gama não-central foram plotadas para ilustrar suas formas. Para isso, utilizou-se a função *dncgamma* implementada em R (APÊNDICE A). Foram considerados diferentes valores de α e δ , incluindo o caso central ($\delta = 0$).

4 RESULTADOS E DISCUSSÃO

Os resultados foram apresentados considerando, inicialmente, a avaliação da acurácia para as probabilidades acumuladas. Em seguida, foram apresentados os resultados para a inversa da função de distribuição em relação a x e em relação ao parâmetro de não-centralidade δ . Por fim, foram apresentados os resultados para a função densidade qui-quadrado não-central, seguidos de alguns gráficos da função densidade gama não-central.

Na Tabela 1, para as configurações de parâmetros usadas por Wang e Kennedy (1994 citados por BENTON; KRISHNAMOORTHY, 2003), são apresentados os valores exatos e computados da função de distribuição qui-quadrado não-central, caso particular da função de distribuição gama incompleta não-central. Os valores exatos foram obtidos no programa Maxima (MAXIMA, 2009). Observou-se que os valores resultantes do método proposto (MP) se aproximam mais dos valores exatos do que aqueles resultantes dos algoritmos de Benton e Krishnamoorthy (2003) (BK) e, em alguns casos, do algoritmo intervalar (Intervalar). O método intervalar foi usado como referência por Benton e Krishnamoorthy (2003). O algoritmo proposto por esses autores foi superior ao algoritmo AS 275, que resultou em *overflow error* quando os graus de liberdade considerados eram grandes. Em nenhuma situação, foram observados problemas computacionais de *overflow error* ou *underflow error* para o algoritmo proposto.

Na Tabela 2, as mesmas configurações usadas por Wang e Kennedy (1994 citados por BENTON; KRISHNAMOORTHY, 2003) e apresentadas na Tabela 1 foram utilizadas para comparar o algoritmo proposto com aquele disponível no programa R (R). Ambos algoritmos foram altamente precisos. O algoritmo proposto foi mais acurado que seu concorrente para $\nu = 31$ e $\delta = 6$ e para $\nu = 51$ e $\delta = 1$ e inferior ao mesmo nas situações em que $\nu = 1$ e $\delta = 6$, para $\nu = 100$ e $\delta = 16$ e para $\nu = 500$ e $\delta = 21$. Nesses últimos casos, o erro foi menor ou igual a 1×10^{-15} . Nas demais circunstâncias, os valores computados pelos dois métodos foram idênticos.

Ainda na Tabela 2, são apresentados os tempos médios de processamento dos dois algoritmos tomando-se a média de 200.000 repetições. Os tempos são

expressos em 1×10^{-6} segundos (μs) e apresentados entre parênteses. O algoritmo proposto foi mais eficiente que seu concorrente nas situações em que os valores de ν e δ são pequenos, ou seja, para $\nu = 1$ e $\delta = 6$ e para $\nu = 1$ e $\delta = 1$. Uma das possíveis razões para a perda de desempenho do método proposto em relação ao método do R deve-se ao fato de os algoritmos do R estarem implementados em uma linguagem de baixo nível (C ou FORTRAN) o que, reconhecidamente, os torna mais eficientes. Para averiguar tal fato, o algoritmo proposto para a obtenção da função de distribuição gama incompleta não-central foi implementado em Java e, em média, apresentou tempo de processamento 24 vezes superior ao tempo de processamento do algoritmo disponível no programa R para o cômputo da função de distribuição qui-quadrado não-central (dados não apresentados).

Tabela 1 Valores exatos e computados de $P(\chi_{\nu}^2(\delta) \leq x)$, usando o método proposto (MP), o método de Benton e Krishnamoorthy (2003) (BK) e a computação intervalar (Intervalar), para as configurações consideradas por Wang e Kennedy (1994 citados por BENTON; KRISHNAMORTHY, 2003, p. 257).

x	ν	δ	Método						
			Exato	Maxima	MP	BK	Intervalar		
0,00393	1	6	0,002498463724258036	...	58042	...	58047	...	58039
9,23636	5	1	0,8272918751175548	...	7555	...	75470	...	75548
24,72497	11	21	0,2539481822183127	...	183126	...	182580	...	183126
44,98534	31	6	0,8125198785064968	...	064968	...	064480	...	064969
38,56038	51	1	0,08519497361859119	...	85912	...	8584550	...	859118
82,35814	100	16	0,01184348822747825	...	747812	...	7441880	...	747824
331,78852	300	16	0,7355956710306710	...	0671	...	05250	...	06709
459,92612	500	21	0,02797023600800056	...	0800136	...	07884640	...	0800060
0,00016	1	1	0,006121428929881420	...	8142	...	81473	...	81423
0,00393	1	1	0,03033814229753780	...	75378	...	75378	...	75380

Na Tabela 3 estão apresentados os valores da função de distribuição gama incompleta não-central para algumas configurações de parâmetros definidas por Chen (2005). Os valores apresentados correspondem aos valores exatos e aos obtidos pelos métodos proposto e método baseado em equações diferenciais de segunda ordem a partir da função gama (GB), proposto por Chen (2005). Observou-se que o algoritmo proposto apresentou resultados idênticos aos exatos em todas as circunstâncias, enquanto seu concorrente foi inaccurado em duas situações que

Tabela 2 Valores exatos e computados de $P(\chi_{\nu}^2(\delta) \leq x)$, usando o método proposto (MP) e o método disponível no programa R (R), para as configurações consideradas por Wang e Kennedy (1994 citados por BENTON; KRISHNAMOORTHY, 2003, p. 257).

x	ν	δ	Exato	Método	
				Maxima	MP
0,00393	1	6	0,002498463724258036	...58042 (163)†	...58037 (174)
9,23636	5	1	0,8272918751175548	...7555 (282)	...7555 (121)
24,72497	11	21	0,2539481822183127	...183126 (621)	...183126 (137)
44,98534	31	6	0,8125198785064968	...064968 (503)	...06497 (153)
38,56038	51	1	0,08519497361859119	...85912 (284)	...85911 (140)
82,35814	100	16	0,01184348822747825	...747812 (595)	...747823 (161)
331,78852	300	16	0,7355956710306710	...0671 (784)	...0671 (278)
459,92612	500	21	0,02797023600800056	...0800136 (788)	...080006 (265)
0,00016	1	1	0,006121428929881420	...8142 (96)	...8142 (195)
0,00393	1	1	0,03033814229753780	...75378 (114)	...75378 (170)

† tempo médio de processamento em μs dado entre parênteses.

correspondem aos casos em que $\alpha = 3,3$, $\beta = 3$, $\delta = 11$ e $x = 33$ e $\alpha = 3,3$, $\beta = 3$, $\delta = 22$, e $x = 55$. Em ambos os casos onde ocorreram os erros, estes foram da ordem de 1×10^{-12} .

Na Tabela 4 são apresentados os valores exatos e computados dos três quartis da função gama incompleta não-central para as configurações consideradas por Chen (2005). Os valores exatos considerados foram aqueles obtidos pelo autor, da mesma forma que aqueles apresentados na Tabela 3. O algoritmo proposto foi consistente e preciso em, praticamente, todas as situações, o que não ocorreu com o seu concorrente. Houve duas situações em que os valores calculados do método proposto não foram idênticos aos exatos, quais sejam para $\alpha = 12$, $\beta = 0,5$, $\delta = 5$ e $I_x(\alpha, \delta) = 0,25$ e para $\alpha = 3,3$, $\beta = 3$, $\delta = 22$ e $I_x(\alpha, \delta) = 0,75$. No último caso, o erro do algoritmo proposto foi de 1×10^{-12} e o erro do algoritmo GB de Chen (2005) foi de 1×10^{-11} , sendo esse último erro, portanto, muito superior.

Na Tabela 5, foi destacada a principal vantagem do algoritmo proposto, que é a inversão da função de distribuição gama incompleta não central em relação ao parâmetro de não-centralidade δ . Em todos os casos considerados, os valores calculados foram idênticos aos exatos. Deve-se chamar a atenção para o fato de não haver comparações com outros métodos pelo fato de não haver implementa-

Tabela 3 Valores exatos e computados de $I_x(\alpha, \delta)$ para algumas configurações apresentadas por Chen (2005) usando o método sugerido por este autor, baseado em equações diferenciais de segunda ordem a partir da função gama (GB), e o método proposto (MP).

α	β	δ	x	Método		
				Exato Chen	MP	GB
12	0,5	7,123	6,6433	0,324771458956	...956	...956
			7,58327402	0,499999999201	...201	...201
		7	6,6311	0,327518593350	...350	...350
			7,5528735	0,499999999372	...372	...372
		5	6,4122	0,371268278420	...420	...420
			7,05926252	0,499999999850	...850	...850
20	0,05	24	1,2447	0,138101425336	...336	...336
			1,5787759	0,499999938343	...343	...343
		15	1,1526	0,234435947472	...472	...472
			1,35476627	0,499999989451	...451	...451
		9	1,0879	0,324058676276	...276	...276
			1,205770139	0,49999998959	...959	...959
3,3	3	11	11	0,059632328165	...165	...165
			22	0,389823672576	...576	...576
			33	0,746185473781	...781	...782
		22	33	0,273862083718	...718	...718
			44	0,566512702350	...350	...350
			55	0,799144186769	...769	...770
		33	55	0,434059956248	...248	...248
			66	0,667776560577	...577	...577
			77	0,837241526999	...999	...999

Tabela 4 Valores exatos e computados dos três quartis de $I_x(\alpha, \delta)$ para algumas configurações apresentadas por Chen (2005) usando o método sugerido por este autor, baseado em equações diferenciais de segunda ordem a partir da função gama (GB), e o método proposto (MP).

α	β	δ	$I_x(\alpha, \delta)$	Método		
				Exato Chen	MP	GB
12	0,5	7,123	0,25	<u>6,216794766296</u>	... 296	... 296
			0,5	<u>7,583274024321</u>	... 321	... 320
			0,75	<u>9,129920773704</u>	... 704	... 705
		7	0,25	<u>6,191013140052</u>	... 052	... 052
			0,5	<u>7,552873503384</u>	... 384	... 384
			0,75	<u>9,094583503100</u>	... 100	... 101
		5	0,25	<u>5,774383316067</u>	... 068	... 067
			0,5	<u>7,059262520763</u>	... 763	... 762
			0,75	<u>8,518169139694</u>	... 694	... 694
20	0,05	24	0,25	<u>1,366004282180</u>	... 180	... 180
			0,5	<u>1,578775950955</u>	... 955	... 955
			0,75	<u>1,810883232516</u>	... 516	... 517
		15	0,25	<u>1,165838353367</u>	... 367	... 367
			0,5	<u>1,354766277765</u>	... 765	... 765
			0,75	<u>1,562132544038</u>	... 038	... 038
		9	0,25	<u>1,034351997274</u>	... 274	... 274
			0,5	<u>1,205770139697</u>	... 697	... 697
			0,75	<u>1,394713619748</u>	... 748	... 748
3,3	3	11	0,25	<u>18,128578732317</u>	... 317	... 319
			0,5	<u>25,005835345056</u>	... 056	... 053
			0,75	<u>33,155655437095</u>	... 095	... 098
		22	0,25	<u>32,039570764631</u>	... 631	... 629
			0,5	<u>41,457823565055</u>	... 055	... 058
			0,75	<u>52,189687675863</u>	... 862	... 883
		33	0,25	<u>46,498456994365</u>	... 365	... 364
			0,5	<u>57,939764742086</u>	... 086	... 090
			0,75	<u>70,710405897668</u>	... 668	... 686

ções conhecidas nem relatos para esse algoritmo.

Tabela 5 Valores exatos e computados do parâmetro de não-centralidade δ para algumas configurações, usando o método proposto (MP).

α	x	$I_x(\alpha, \delta)$	δ	
			Exato	Calculado
0,5	0,01965	0,00249846372425804	6	6
25,5	19,28019	0,0851949736185912	1	1
250	229,96306	0,0279702360080014	21	21
12	13,2866	0,324771458955857	7,123	7,123
20	24,894	0,138101425336041	24	24
20	23,052	0,234435947471964	15	15
3,3	11/3	0,0596323281646978	11	11
3,3	44/3	0,566512702350402	22	22
3,3	22	0,66777656057718	33	33

Na Tabela 6, estão apresentados os valores da função densidade qui-quadrado não-central para algumas configurações dos parâmetros, utilizando o método proposto e o método disponível no programa R. As diferenças entre os valores computados pelos dois métodos são menores que 1×10^{-12} .

Tabela 6 Valores da função densidade qui-quadrado não-central $\chi^2_\nu(\delta)$ para diferentes configurações dos parâmetros, usando o método proposto (MP) e o método disponível no programa R (R).

x	ν	δ	Valores da função densidade	
			MP	R
0,01965	0,5	6	0,2438563792181807	0,2438563792181806
19,28019	25,5	1	0,0410169980478537	0,04101699804785309
229,96306	250	21	0,003901341402693288	0,003901341402689759
13,2866	12	7,123	0,05092503663983963	0,0509250366398392
24,894	20	24	0,00894559718063725	0,00894559718063412
23,052	20	15	0,02334498018782803	0,02334498018782735
11/3	3.3	11	0,02042930793786916	0,02042930793786706
44/3	3.3	22	0,02834319051384455	0,028343190513844
22	3.3	33	0,01926192534430278	0,01926192534429827

A partir do algoritmo para o cálculo da função densidade gama não-cen-

tral, foi possível apresentá-la nas Figuras 3 e 4. Na Figura 3, os valores de α são fixados em $\alpha = 5$ e $\alpha = 20$, considerando diversas configurações para o parâmetro δ . Na Figura 4, os valores de δ são fixados em $\delta = 5$ e $\delta = 20$, considerando diversas configurações para o parâmetro α . O comportamento dos gráficos nas situações consideradas são semelhantes. Quando são considerados valores crescentes para um dos parâmetros da distribuição, mantendo o outro fixo, a função densidade gama não-central aproxima-se da função densidade normal.

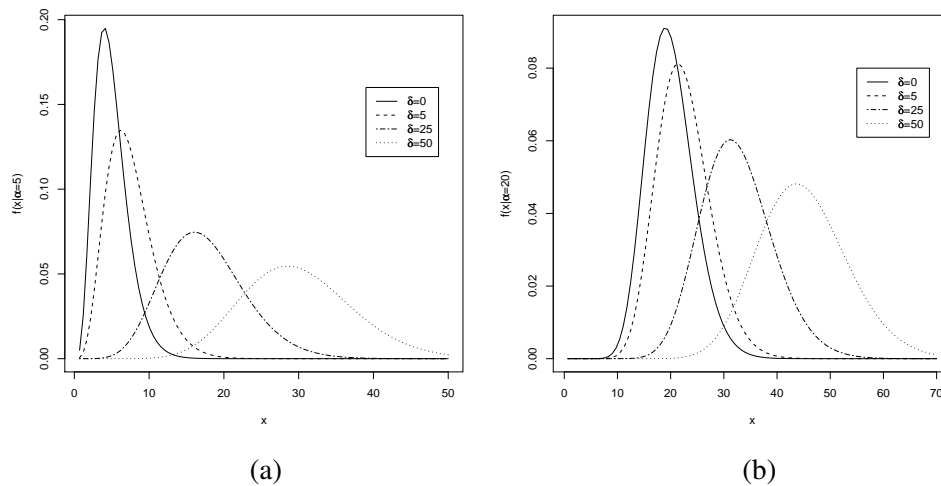


Figura 3 Função densidade gama não-central para $\alpha = 5$ (a) e para $\alpha = 20$ (b), considerando diferentes valores de δ

A característica de maior relevância do presente trabalho é a inversão da função de distribuição gama não-central em relação ao parâmetro de não-centralidade δ . Do ponto de vista prático, a obtenção de tal resultado em relação a δ , dados os valores de poder $(1 - \beta)$ e o nível de significância do teste, possibilita que diferenças mínimas significativas sejam calculadas em testes cujas estatísticas H_0 possuem distribuição qui-quadrado. Como a distribuição qui-quadrado não-central é um caso particular da gama não-central e inúmeros testes são baseados na distribuição qui-quadrado, sua aplicabilidade é imediata. Além disso, para essas mesmas situações, a função poder pode ser obtida a partir do cálculo da

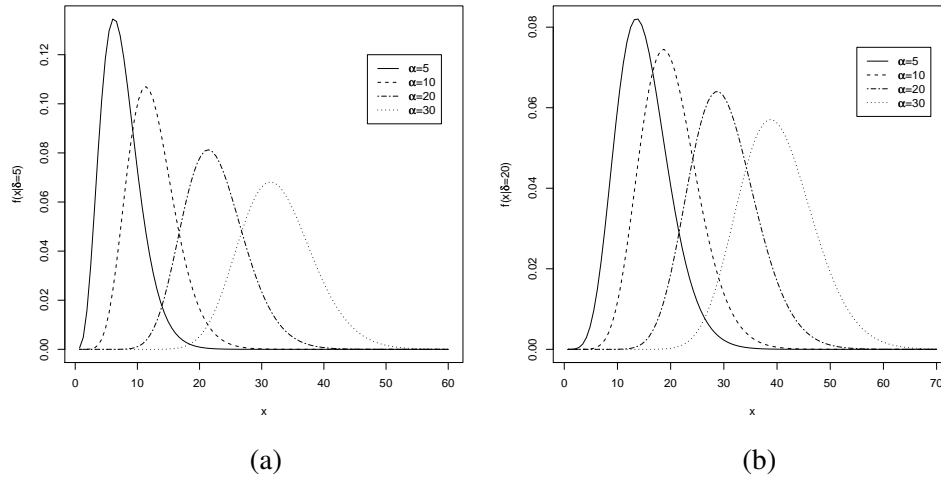


Figura 4 Função densidade gama não-central para $\delta = 5$ (a) e para $\delta = 20$ (b), considerando diferentes valores de α

função de distribuição qui-quadrado não-central. Como já salientado, uma simples transformação de variáveis possibilita que as funções relativas a gama não-central sejam utilizadas para o cômputo da qui-quadrado não-central.

5 CONCLUSÕES

Os algoritmos para a obtenção da função de distribuição gama incompleta não-central, de sua função densidade e da inversa da função de distribuição para obtenção de quantis e do parâmetro de não-centralidade foram propostos e implementados no programa R com sucesso.

As rotinas implementadas apresentaram bom desempenho e, em geral, foram mais acuradas que as rotinas concorrentes. Quanto ao tempo de processamento, se implementadas em uma linguagem de mais baixo nível, as rotinas propostas são mais eficientes do que o algoritmo disponível no *software* R para o cálculo da função de distribuição qui-quadrado não-central.

REFERÊNCIAS

BAHAREV, A.; KEMÉNY, S. On the computation of the noncentral F and noncentral beta distribution. **Statistics and Computing**, London, v. 18, n. 3, p. 333-340, Sept. 2008.

BARROSO, L. C. et al. **Cálculo numérico: com aplicações**. 2. ed. São Paulo: Harbra, 1987. 367 p.

BEARZOTI, E. **Introdução à teoria de probabilidades e à inferência estatística**. Lavras: UFLA, 1998. 288 p. Apostila.

BENTON, D.; KRISHNAMOORTHY, K. Computing discrete mixtures of continuous distributions: noncentral chisquare, noncentral t and the distribution of the square of the sample multiple correlation coefficient. **Computational Statistics and Data Analysis**, New York, v. 43, n. 2, p. 249-267, Apr. 2003.

CAMPOS, F. F. **Cálculo numérico**. Belo Horizonte: UFMG, 1998. 149 p.

CHATTAMVELLI, R.; SHANMUGAN, R. Computing the non-central beta distribution function. **Journal of the Royal Statistical Society, Series C**, London, v. 46, n. 1, p. 146-156, Mar. 1997.

CHEN, Z. Y. The S -system computation of non-central gamma distribution. **Journal of Statistical Computation and Simulation**, New York, v. 75, n. 10, p. 813-829, Oct. 2005.

FISHER, R. A. The general sampling distribution of the multiple correlation coefficient. **Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character**, London, v. 121, n. 788, p. 654-673, Dec. 1928.

FRICK, H. Algorithm *ASR84*: a remark on algorithm *AS – 226*: computing non-central beta probabilities. **Journal of the Royal Statistical Society, Series C**, London, v. 39, n. 2, p. 311-312, Apr. 1990.

JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. **Distributions in statistics: continuous univariate distributions**. New York: J. Wiley, 1970. 752 p.

KNÜSEL, L.; BABLOK, B. Computation of the noncentral gamma distribution. **SIAM Journal on Scientific Computing**, Philadelphia, v. 17, n. 5, p. 1224-1231, Sept. 1996.

KRISHNAMOORTHY, K. **Handbook of statistical distributions with applications**. Boca Raton: CRC, 2006. 346 p.

LANCZOS, C. A precision approximation of the gamma function. **Journal of the Society for Industrial and Applied Mathematics, Numerical Analysis**, Philadelphia, v. 1, p. 86-96, 1964.

LENTH, R. V. Algorithm *AS226*: computing noncentral beta probabilities. **Journal of the Royal Statistical Society. Series C, Applied Statistics**, London, v. 36, n. 2, p. 241-244, 1987.

MAXIMA. **Maxima**: a computer algebra system. 2009. Disponível em: <<http://maxima.sourceforge.net/>>.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics**. Singapore: McGraw-Hill International, 1974. 480 p.

PATNAIK, P. B. The non-central χ^2 and *F*-distribution and their applications. **Biometrika**, London, v. 36, n. 1/2, p. 202-232, 1949.

PRESS, W. H. et al. **Numerical recipes in Fortran: the art of scientific computing**. Cambridge: Cambridge University, 1992. 994 p.

R DEVELOPMENT CORE TEAM. **R**: a language and environment for statistical computing. Vienna, 2008. Disponível em: <<http://www.R-project.org>>.

RUBEN, H. Non-central chi-square and gamma revisited. **Communications in Statistics**, Ontario, v. 3, n. 7, p. 607-633, 1974.

RUGGIERO, M. A. G.; LOPES, V. L. da R. **Cálculo numérico: aspectos teóricos e computacionais**. 2. ed. Rio de Janeiro: Makron Books, 1996. 406 p.

SEVERO, N.; ZELLEN, M. Normal approximation to the chi-square and noncentral F probability functions. **Biometrika**, London, v. 47, 1960.

WANG, M.; KENNEDY, W. J. Self-validating computations of probabilities for selected central and non-central probability functions. **Journal of the American Statistical Association**, v. 89, p. 878-887, 1994.

APÊNDICE A - Rotinas R

```
#####
# Algoritmo implementado em R para obter a função de distri- #
# buição gama incompleta não-central, F(x|alpha, delta)      #
#####

pncgamma <- function(x, alpha, delta)
{
  if ((alpha <= 0) | (x <= 0))
    stop("Alpha e x devem ser positivos!")
  if (delta < 0) stop("Delta deve ser não-negativo!")
  errortol <- 1e-15
  maxitr <- 5000
  d <- delta
  k <- ceiling(d / 2)
  a <- alpha + k
  gamac <- pgamma(x, a)
  gamad <- gamac
  gxd <- exp(a*log(x)-x-lgamma(a+1))
  gxc <- gxd * a / x
  ppoic <- dpois(k, d / 2)
  ppoid <- ppoic
  remain <- 1 - ppoic
  cdf <- ppoic * gamac
  i <- 1
  convergiu <- F
  while (!convergiu)
  {
    gxc <- gxc * x / (a + i - 1)
    gamac <- gamac - gxc
    ppoic <- ppoic * (d / 2) / (k + i)
    cdf <- cdf + ppoic * gamac
    error <- remain * gamac
    remain <- remain - ppoic
    if (i > k)
    {
      if ((error <= errortol) | (i > maxitr)) convergiu <- TRUE
      i <- i + 1
    }
  }
}
```

```

    } else
    {
      gxd    <- gxd * (a + 1 - i) / x
      gamad  <- gamad + gxd
      ppoid  <- ppoid * (k - i + 1) / (d / 2)
      cdf    <- cdf + ppoid * gamad
      remain <- remain - ppoid
      if ((remain <= errortol) | (i > maxitr)) convergiu <- TRUE
      i      <- i + 1
    }
  } # convergiu
  return(cdf)
}

```

```

#####
# Algoritmo implementado em R para obter a inversa de          #
# F(x|alpha, delta) em relação a delta                       #
#####

###Algoritmo para a obtenção do valor inicial d0###
d0<- function(x, p, alpha)
{
  dn0 <- 0.5
  dn1 <- 1
  errortol<- 1e-15
  nmax <- 1000
  z    <- qnorm(p)
  a    <- alpha
  fdn0<-(((a+dn0)*(z+sqrt((2*a+dn0)^2/(a+dn0)-1))^2)/(2*a+dn0))-x
  fdn1<-(((a+dn1)*(z+sqrt((2*a+dn1)^2/(a+dn1)-1))^2)/(2*a+dn1))-x
  it   <- 1
  convd <- F
  while(!convd)
  {
    dn2 <- dn1 - fdn1 * (dn1 - dn0) / (fdn1 - fdn0)
    if (dn2=="NaN") break
    if (dn2 < 0) dn2 <- dn1 / 2
    dn0 <- dn1
    fdn0 <- fdn1
    dn1 <- dn2
  }
}

```

```

fdn1 <- (((a+dn1)*(z+sqrt((2*a+dn1)^2/(a+dn1)-1))^2)/(2*a+dn1))-x
if ((abs(dn2 - dn0) <= errortol * dn0) | (it > nmax)) convd <- T
it <- it + 1
} #convd
#Proteção para a aproximação ruim
if (dn2=="NaN") return(1) else return(dn2)
}

indncgamma <- function(x, alpha, p)
{
  if ((alpha <= 0) | (x <= 0))
    stop("Alpha e x devem ser positivos!")
  if ((p <= 0) | (p >= 1)) stop("P deve estar entre 0 e 1!")
  errortol <- 1e-15
  maxitr <- 5000
  dn <- d0(x,p,alpha)
  it <- 1
  convnewton <- F
  while (!convnewton)
  {
    d <- dn
    k <- ceiling(d / 2)
    a <- alpha + k
    gamac <- pgamma(x,a)
    gamad <- gamac
    gxd <- exp(a*log(x)-x-lgamma(a+1))
    gxc <- gxd * a / x
    hxc <- gxd
    hxd <- hxc
    ppoic <- dpois(k, d / 2)
    ppoid <- ppoic
    remain <- 1 - ppoic
    cdf <- ppoid * gamac
    g <- ppoid * hxc
    i <- 1
    convergiu <- F
    while (!convergiu)
    {
      gxc <- gxc * x / (a + i - 1)
      gamac <- gamac - gxc
    }
  }
}

```

```

hxc    <- hxc * x / (a + i)
ppoic  <- ppoic * (d / 2) / (k + i)
cdf    <- cdf + ppoic * gamac
g      <- g + ppoic * hxc
error  <- remain * gamac
remain <- remain - ppoic
if (i > k)
{
  if ((error <= errortol) | (i > maxitr))
    convergiu <- TRUE
  i    <- i + 1
} else
{
  gxd   <- gxd * (a + 1 - i) / x
  gamad <- gamad + gxd
  hxd   <- hxd * (a + 1 - i) / x
  ppoid <- ppoid * (k - i + 1) / (d / 2)
  cdf   <- cdf + ppoid * gamad
  g     <- g + ppoid * hxd
  remain <- remain - ppoid
  if ((remain <= errortol) | (i > maxitr))
    convergiu <- TRUE
  i     <- i + 1
}
}#convergiu
if (d + 2*(cdf - p) / g <= 0)
{
  dn   <- d / 2
} else
{
  dn   <- d + 2*(cdf - p) / g
}
if ((abs(dn - d) <= d * errortol) | (it > maxitr))
convnewton <- TRUE
it <- it + 1
}#convnewton
return(dn)
}#geral

```

```
#####
# Algoritmo implementado em R para obter a inversa de #
# F(x|alpha, delta) em relação a x #
#####

inxncgamma <- function(p, alpha, delta)
{
  if (alpha <= 0) stop("Alpha deve ser positivo!")
  if (delta < 0) stop("Delta deve ser não-negativo!")
  if ((p <= 0) | (p >= 1)) stop("P deve estar entre 0 e 1!")
  errortol <- 1e-15
  maxitr <- 5000
  d <- delta
  k <- ceiling(d / 2)
  a <- alpha + k
  z <- qnorm(p)
  x0 <- ((a+d)*(z+sqrt((2*a+d)^2/(a+d)-1))^2)/(2*a+d)
  xn <- x0
  it <- 1
  convnewton <- F
  while (!convnewton)
  {
    x <- xn
    gamac <- pgamma(x, a)
    gamad <- gamac
    gxd <- exp(a*log(x)-x-lgamma(a+1))
    gxc <- gxd * a / x
    hxc <- gxd * a / x
    hxd <- hxc
    ppoic <- dpois(k, d / 2)
    ppoid <- ppoic
    remain <- 1 - ppoic
    cdf <- ppoic * gamac
    g <- ppoic * hxc
    i <- 1
    convergiu <- F
    while (!convergiu)
    {
      gxc <- gxc * x / (a + i - 1)
    }
  }
}
```

```

gamac <- gamac - gxc
hxc   <- hxc * x / (a + i - 1)
ppoic <- ppoic * (d / 2) / (k + i)
cdf   <- cdf + ppoic * gamac
g     <- g + ppoic * hxc
error <- remain * gamac
remain <- remain - ppoic
if (i > k)
{
  if ((error <= errortol) | (i > maxitr))
    convergiu <- TRUE
  i <- i + 1
} else
{
  gxd   <- gxd * (a + 1 - i) / x
  gamad <- gamad + gxd
  hxd   <- hxd * (a - i) / x
  ppoid <- ppoid * (k - i + 1) / (d / 2)
  cdf   <- cdf + ppoid * gamad
  g     <- g + ppoid * hxd
  remain <- remain - ppoid
  if ((remain <= errortol) | (i > maxitr))
    convergiu <- TRUE
  i <- i + 1
}
}#convergiu
if (x - (cdf - p) / g <= 0)
{
  xn <- x / 2
} else
{
  xn <- x - (cdf - p) / g
}
if ((abs(xn - x) <= x * errortol) | (it > maxitr))
convnewton <- TRUE
it <- it + 1
}#convnewton
return(xn)
}#geral

```

```
#####
# Algoritmo implementado em R para obter a função densidade gama#
# incompleta não-central, f(x|alpha, delta) #
#####

dncgamma <- function(x, alpha, delta)
{
  if ((alpha <= 0) | (x <= 0))
    stop("Alpha e x devem ser positivos!")
  if (delta < 0) stop("Delta deve ser não-negativo!")
  errortol <- 1e-15
  maxitr <- 5000
  d <- delta
  k <- ceiling(d / 2)
  a <- alpha + k
  hxc <- exp(a*log(x)-x-lgamma(a+1))
  hxc <- hxc * a/x
  hxd <- hxc
  ppoic <- dpois(k, d / 2)
  ppoid <- ppoic
  remain <- 1 - ppoic
  g <- ppoic * hxc
  i <- 1
  convergiu <- F
  while (!convergiu)
  {
    hxc <- hxc * x/(a + i -1)
    ppoic <- ppoic * (d/2)/(k + i)
    g <- g + ppoic * hxc
    error <- remain * g
    remain <- remain - ppoic
    if (i > k)
    {
      if ((error <= errortol) | (i > maxitr)) convergiu <- TRUE
      i <- i + 1
    } else
    {
      hxd <- hxd * (a - i) / x
    }
  }
}
```

```
    ppoid <- ppoid * (k - i + 1) / (d / 2)
    g      <- g + ppoid * hxd
    remain <- remain - ppoid
    if ((remain <= errortol) | (i > maxitr))convergiu <- TRUE
    i      <- i + 1
  }
} # convergiu
return(g)
}
```