



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
DEPARTAMENTO DE ESTATÍSTICA



Comparação entre algoritmos geradores das Distribuições Normal, Qui-Quadrado, F de Snedecor e t de Student através de simulação

Autor: Denis Camargo Schutz
Orientador: Professor Dr. Lori Viali

Porto Alegre, 09 de Julho de 2012.

Universidade Federal do Rio Grande do Sul
Instituto de Matemática
Departamento de Estatística

Comparação entre algoritmos geradores das Distribuições Normal, Qui-Quadrado, F de Snedecor e t de Student através de simulação

Autor: Denis Camargo Schutz

Monografia apresentada para obtenção
do grau de Bacharel em Estatística.

Banca Examinadora:
Professor Dr. Lori Viali (orientador)
Professor Dr. Hubert Ahlert (convidado)

Porto Alegre, 09 de Julho de 2012.

Dedico este trabalho a meus pais, Aldemira Fátima de Camargo Schutz e João Carlos Schutz, por terem me dado a vida e sempre batalharem ao máximo para que eu pudesse estudar.

RESUMO

O presente trabalho faz uma comparação, através de Simulação Monte Carlo, entre diferentes algoritmos para geração de números aleatórios das distribuições Normal, Qui-Quadrado, F de Snedecor e t de Student. Estas quatro distribuições são muito empregadas em Estatística e têm como característica em comum não serem integráveis analiticamente. Foram utilizadas 100 replicações, com diferentes tamanhos de amostra para cada distribuição. As análises descritivas foram realizadas mediante a comparação das estimativas médias de cada algoritmo ao longo das 100 amostras com os valores “exatos” da distribuição. Também foi estimado o erro padrão, e gráficos foram gerados para algumas medidas-resumo, tais como a média, o desvio-padrão e a mediana. De forma geral, os resultados foram bastante equilibrados, pois, para um mesmo tamanho de amostra em uma dada distribuição, alguns algoritmos apresentaram um comportamento semelhante.

Palavras-chave: Simulação Monte Carlo. Algoritmos. Geração de números aleatórios. Distribuição Normal. Distribuição Qui-Quadrado. Distribuição F de Snedecor. Distribuição t de Student. Amostra.

LISTA DE FIGURAS

Figura 1 – Gráfico da distribuição normal padrão, com média $\mu = 0$ e variância $\sigma^2 = 1$	21
Figura 2 – Gráfico da função gama no domínio dos números reais positivos.....	23
Figura 3 – Gráfico da Distribuição t de Student para alguns valores de ν comparada à Distribuição Normal (em preto)	24
Figura 4 – Gráfico da Distribuição Qui-Quadrado para alguns valores de ν	26
Figura 5 – Gráfico da Distribuição F de Snedecor para $m = 5$ e $n = 20$	28
Figura 6 – Diagrama ilustrativo das relações entre as distribuições de probabilidade.....	30
Figura 7 – Histograma de 1.000.000 números aleatórios gerados pelo algoritmo do Mersenne Twister	34
Figura 8 – Gráfico de dispersão com 2000 pares de números aleatórios gerados em sequência	35
Figura 9 – Gráfico de dispersão com 32500 pares de números aleatórios gerados em sequência	35
Figura 10 – Histograma de 100 valores gerados da distribuição $N(0,1)$ pelo algoritmo de Kinderman-Monahan.....	61
Figura 11 – Histograma de 1000 valores gerados da distribuição $N(0,1)$ pelo algoritmo de Rejeição pela Distribuição Exponencial	70
Figura 12 – Histograma de 10000 valores gerados da distribuição $N(0,1)$ pelo algoritmo de Box-Muller	80

LISTA DE QUADROS

Quadro 1 – Gráficos dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 100.....	53
Quadro 2 – Gráficos dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 100.....	55
Quadro 3 – Gráficos dos valores estimados para o 2º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 100.....	58
Quadro 4 – Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 100	60
Quadro 5 – Gráficos dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 1000.....	62
Quadro 6 – Gráficos dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 1000.....	65
Quadro 7 – Gráficos dos valores estimados para o 2º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 1000.....	67
Quadro 8 – Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 1000	69
Quadro 9 – Gráficos dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 10000.....	72
Quadro 10 – Gráficos dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	74
Quadro 11 – Gráficos dos valores estimados para o 2º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 10000.....	77
Quadro 12 – Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 10000	79
Quadro 13 – Gráficos dos valores estimados para a esperança da distribuição $\chi^2(1)$ em 100 amostras de tamanho 100.....	86
Quadro 14 - Gráficos dos valores estimados para a esperança da distribuição $\chi^2(17)$ em 100 amostras de tamanho 100.....	86
Quadro 15 – Gráficos dos valores estimados para a esperança da distribuição $\chi^2(50)$ em 100 amostras de tamanho 100.....	87
Quadro 16 – Gráficos dos valores estimados para o desvio-padrão da distribuição $\chi^2(1)$ em 100 amostras de tamanho 100.....	89
Quadro 17 – Gráficos dos valores estimados para o desvio-padrão da distribuição $\chi^2(17)$ em 100 amostras de tamanho 100.....	89
Quadro 18 – Gráficos dos valores estimados para o desvio-padrão da distribuição $\chi^2(50)$ em 100 amostras de tamanho 100.....	90
Quadro 19 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2(1)$ em 100 amostras de tamanho 100	93
Quadro 20 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2(17)$ em 100 amostras de tamanho 100	93
Quadro 21 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2(50)$ em 100 amostras de tamanho 100	94
Quadro 22 – Histogramas de 100 valores gerados da Distribuição Qui-Quadrado pelo algoritmo Logaritmo Produto Uniformes	95
Quadro 23 - Gráficos dos valores estimados para a esperança da distribuição $\chi^2(1)$ em 100 amostras de tamanho 1000.....	96

Quadro 24 - Gráficos dos valores estimados para a esperança da distribuição χ^2 (17) em 100 amostras de tamanho 1000.....	97
Quadro 25 - Gráficos dos valores estimados para a esperança da distribuição χ^2 (50) em 100 amostras de tamanho 1000.....	97
Quadro 26 - Gráficos dos valores estimados para o desvio-padrão da distribuição χ^2 (1) em 100 amostras de tamanho 1000.....	99
Quadro 27 - Gráficos dos valores estimados para o desvio-padrão da distribuição χ^2 (17) em 100 amostras de tamanho 1000.....	100
Quadro 28 - Gráficos dos valores estimados para o desvio-padrão da distribuição χ^2 (50) em 100 amostras de tamanho 10.....	100
Quadro 29 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição χ^2 (1) em 100 amostras de tamanho 1000.....	103
Quadro 30 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição χ^2 (17) em 100 amostras de tamanho 1000.....	104
Quadro 31 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição χ^2 (50) em 100 amostras de tamanho 1000.....	104
Quadro 32 – Histogramas de 1000 valores gerados da Distribuição Qui-Quadrado pelo algoritmo Soma Quadrados Normal.....	106
Quadro 33 - Gráficos dos valores estimados para a esperança da distribuição χ^2 (1) em 100 amostras de tamanho 10000.....	107
Quadro 34 - Gráficos dos valores estimados para a esperança da distribuição χ^2 (17) em 100 amostras de tamanho 10000.....	108
Quadro 35 - Gráficos dos valores estimados para a esperança da distribuição χ^2 (50) em 100 amostras de tamanho 10000.....	108
Quadro 36 - Gráficos dos valores estimados para o desvio-padrão da distribuição χ^2 (1) em 100 amostras de tamanho 10000.....	110
Quadro 37 - Gráficos dos valores estimados para o desvio-padrão da distribuição χ^2 (17) em 100 amostras de tamanho 10000.....	111
Quadro 38 - Gráficos dos valores estimados para o desvio-padrão da distribuição χ^2 (50) em 100 amostras de tamanho 10000.....	111
Quadro 39 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição χ^2 (1) em 100 amostras de tamanho 10000.....	114
Quadro 40 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição χ^2 (17) em 100 amostras de tamanho 10000.....	115
Quadro 41 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição χ^2 (50) em 100 amostras de tamanho 10000.....	115
Quadro 42 – Histogramas de 10000 valores gerados da Distribuição Qui-Quadrado pelo algoritmo Logaritmo Produto Uniformes.....	117
Quadro 43 - Gráficos dos valores estimados para a esperança da distribuição F(1,5) em 100 amostras de tamanho 10000.....	120
Quadro 44 - Gráficos dos valores estimados para a esperança da distribuição F(5,5) em 100 amostras de tamanho 10000.....	120
Quadro 45 - Gráficos dos valores estimados para o desvio-padrão da distribuição F(1,5) em 100 amostras de tamanho 10000.....	122
Quadro 46 - Gráficos dos valores estimados para o desvio-padrão da distribuição F(5,5) em 100 amostras de tamanho 10000.....	122

Quadro 47 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição F(1,5) em 100 amostras de tamanho 10000	124
Quadro 48 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição F(5,1) em 100 amostras de tamanho 10000	125
Quadro 49 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição F(5,5) em 100 amostras de tamanho 10000	125
Quadro 50 – Histogramas de 10000 valores gerados da Distribuição F pelo algoritmo F pela Beta	126
Quadro 51 – Gráficos dos valores estimados para a esperança da distribuição t(3) em 100 amostras de tamanho 100.....	131
Quadro 52 - Gráficos dos valores estimados para a esperança da distribuição t(15) em 100 amostras de tamanho 100.....	131
Quadro 53 – Gráficos dos valores estimados para a esperança da distribuição t(30) em 100 amostras de tamanho 100.....	132
Quadro 54 – Gráficos dos valores estimados para o desvio-padrão da distribuição t(3) em 100 amostras de tamanho 100.....	133
Quadro 55 – Gráficos dos valores estimados para o desvio-padrão da distribuição t(15) em 100 amostras de tamanho 100.....	134
Quadro 56 – Gráficos dos valores estimados para o desvio-padrão da distribuição t(30) em 100 amostras de tamanho 100.....	134
Quadro 57- Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição t(3) em 100 amostras de tamanho 100.....	137
Quadro 58 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição t(15) em 100 amostras de tamanho 100.....	137
Quadro 59 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição t(30) em 100 amostras de tamanho 100.....	138
Quadro 60 – Histogramas de 100 valores gerados da Distribuição t de Student pelo algoritmo Normal/Qui-Quadrado	139
Quadro 61 - Gráficos dos valores estimados para a esperança da distribuição t(3) em 100 amostras de tamanho 10000.....	140
Quadro 62 - Gráficos dos valores estimados para a esperança da distribuição t(15) em 100 amostras de tamanho 10000.....	141
Quadro 63 - Gráficos dos valores estimados para a esperança da distribuição t(30) em 100 amostras de tamanho 10000.....	141
Quadro 64 - Gráficos dos valores estimados para o desvio-padrão da distribuição t(3) em 100 amostras de tamanho 10000.....	143
Quadro 65 - Gráficos dos valores estimados para o desvio-padrão da distribuição t(15) em 100 amostras de tamanho 10000.....	143
Quadro 66 - Gráficos dos valores estimados para o desvio-padrão da distribuição t(30) em 100 amostras de tamanho 10000.....	144
Quadro 67 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição t(3) em 100 amostras de tamanho 10000.....	146
Quadro 68 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição t(15) em 100 amostras de tamanho 10000.....	147
Quadro 69 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição t(30) em 100 amostras de tamanho 10000.....	147
Quadro 70 – Histogramas de 10000 valores gerados da Distribuição t de Student pelo algoritmo Normal/Qui-Quadrado	149

LISTA DE TABELAS

Tabela 1 – Medidas estatísticas das distribuições Normal, t de Student, Gama, Qui-Quadrado e F de Snedecor	31
Tabela 2 – Média e erro padrão dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 100	52
Tabela 3 – Média e erro padrão dos valores estimados para o máximo da distribuição $N(0,1)$ em 100 amostras de tamanho 100	53
Tabela 4 – Média e erro padrão dos valores estimados para o mínimo da distribuição $N(0,1)$ em 100 amostras de tamanho 100.....	54
Tabela 5 – Média e erro padrão dos valores estimados para a amplitude da distribuição $N(0,1)$ em 100 amostras de tamanho 100	54
Tabela 6 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 100	55
Tabela 7 – Média e erro padrão dos valores estimados para a assimetria da distribuição $N(0,1)$ em 100 amostras de tamanho 100	56
Tabela 8 – Média e erro padrão dos valores estimados para a curtose da distribuição $N(0,1)$ em 100 amostras de tamanho 100.....	56
Tabela 9 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 100	57
Tabela 10 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 100	57
Tabela 11 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 100	58
Tabela 12 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição $N(0,1)$ em 100 amostras de tamanho 100	59
Tabela 13 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 100	59
Tabela 14 – Estimativas médias do algoritmo de Kinderman-Monahan em 100 amostras de tamanho 100 da distribuição $N(0,1)$	61
Tabela 15 – Média e erro padrão dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 1000	61
Tabela 16 – Média e erro padrão dos valores estimados para o máximo da distribuição $N(0,1)$ em 100 amostras de tamanho 1000	62
Tabela 17 – Média e erro padrão dos valores estimados para o mínimo da distribuição $N(0,1)$ em 100 amostras de tamanho 1000	63
Tabela 18 – Média e erro padrão dos valores estimados para a amplitude da distribuição $N(0,1)$ em 100 amostras de tamanho 1000	63
Tabela 19 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 1000	64
Tabela 20 – Média e erro padrão dos valores estimados para a assimetria da distribuição $N(0,1)$ em 100 amostras de tamanho 1000	65
Tabela 21 – Média e erro padrão dos valores estimados para a curtose da distribuição $N(0,1)$ em 100 amostras de tamanho 1000	65
Tabela 22 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 1000	66
Tabela 23 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 1000	66
Tabela 24 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 1000	67
Tabela 25 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição $N(0,1)$ em 100 amostras de tamanho 1000	68
Tabela 26 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 1000	68

Tabela 27 – Estimativas médias do algoritmo de Rejeição pela Distribuição Exponencial em 100 amostras de tamanho 1000 da distribuição $N(0,1)$	70
Tabela 28 – Média e erro padrão dos tempos de execução, em segundos, medidos em 100 amostras de tamanho 10000 da distribuição $N(0,1)$	71
Tabela 29 – Média e erro padrão dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	71
Tabela 30 – Média e erro padrão dos valores estimados para o máximo da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	72
Tabela 31 – Média e erro padrão dos valores estimados para o mínimo da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	73
Tabela 32 – Média e erro padrão dos valores estimados para a amplitude da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	73
Tabela 33 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	74
Tabela 34 – Média e erro padrão dos valores estimados para a assimetria da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	75
Tabela 35 – Média e erro padrão dos valores estimados para a curtose da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	75
Tabela 36 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	76
Tabela 37 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	76
Tabela 38 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	77
Tabela 39 – Média e erro padrão dos valores estimados para a distância interquartílica da distribuição $N(0,1)$ em 100 amostras de tamanho 10000	78
Tabela 40 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 10000	78
Tabela 41 – Estimativas médias do algoritmo de Box-Muller em 100 amostras de tamanho 10000 da distribuição $N(0,1)$	80
Tabela 42 – Média e erro padrão dos valores estimados para a esperança da distribuição Qui-Quadrado em 100 amostras de tamanho 100.....	85
Tabela 43 – Média e erro padrão dos valores estimados para o máximo da distribuição Qui-Quadrado em 100 amostras de tamanho 100.....	87
Tabela 44 – Média e erro padrão dos valores estimados para o mínimo da distribuição Qui-Quadrado em 100 amostras de tamanho 100.....	87
Tabela 45 - Média e erro padrão dos valores estimados para a amplitude da distribuição Qui-Quadrado em 100 amostras de tamanho 100.....	88
Tabela 46 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição Qui-Quadrado em 100 amostras de tamanho 100	88
Tabela 47 – Média e erro padrão dos valores estimados para a assimetria da distribuição Qui-Quadrado em 100 amostras de tamanho 100.....	90
Tabela 48 – Média e erro padrão dos valores estimados para a curtose da distribuição Qui-Quadrado em 100 amostras de tamanho 100.....	91
Tabela 49 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 100.....	91
Tabela 50 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 100.....	91
Tabela 51 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 100.....	92
Tabela 52 – Média e erro padrão dos valores estimados para a distância interquartílica da distribuição Qui-Quadrado em 100 amostras de tamanho 100	92
Tabela 53 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição Qui-Quadrado em 100 amostras de tamanho 100	92

Tabela 54 – Estimativas médias do algoritmo Logaritmo Produto Uniformes em 100 amostras de tamanho 100 da Distribuição Qui-Quadrado.....	95
Tabela 55 – Média e erro padrão dos valores estimados para a esperança da distribuição Qui-Quadrado em 100 amostras de tamanho 1000.....	96
Tabela 56 – Média e erro padrão dos valores estimados para o máximo da distribuição Qui-Quadrado em 100 amostras de tamanho 1000.....	98
Tabela 57 – Média e erro padrão dos valores estimados para o mínimo da distribuição Qui-Quadrado em 100 amostras de tamanho 1000.....	98
Tabela 58 – Média e erro padrão dos valores estimados para a amplitude da distribuição Qui-Quadrado em 100 amostras de tamanho 1000.....	98
Tabela 59 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição Qui-Quadrado em 100 amostras de tamanho 1000	99
Tabela 60 – Média e erro padrão dos valores estimados para a assimetria da distribuição Qui-Quadrado em 100 amostras de tamanho 1000.....	101
Tabela 61 – Média e erro padrão dos valores estimados para a curtose da distribuição Qui-Quadrado em 100 amostras de tamanho 1000.....	101
Tabela 62 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 1000.....	101
Tabela 63 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 1000.....	102
Tabela 64 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 1000.....	102
Tabela 65 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição Qui-Quadrado em 100 amostras de tamanho 1000	102
Tabela 66 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição Qui-Quadrado em 100 amostras de tamanho 1000	103
Tabela 67 – Estimativas médias do algoritmo Soma Quadrados Normal em 100 amostras de tamanho 1000 da Distribuição Qui-Quadrado.....	105
Tabela 68 – Média e erro padrão dos tempos de execução, em segundos, medidos em 100 amostras de tamanho 10000 da Distribuição Qui-Quadrado.....	106
Tabela 69 - Média e erro padrão dos valores estimados para a esperança da distribuição Qui-Quadrado em 100 amostras de tamanho 10000	107
Tabela 70 - Média e erro padrão dos valores estimados para o máximo da distribuição Qui-Quadrado em 100 amostras de tamanho 10000	109
Tabela 71 - Média e erro padrão dos valores estimados para o mínimo da distribuição Qui-Quadrado em 100 amostras de tamanho 10000	109
Tabela 72 - Média e erro padrão dos valores estimados para a amplitude da distribuição Qui-Quadrado em 100 amostras de tamanho 10000	109
Tabela 73 - Média e erro padrão dos valores estimados para o desvio-padrão da distribuição Qui-Quadrado em 100 amostras de tamanho 10000	110
Tabela 74 – Média e erro padrão dos valores estimados para a assimetria da distribuição Qui-Quadrado em 100 amostras de tamanho 10000	112
Tabela 75 – Média e erro padrão dos valores estimados para a curtose da distribuição Qui-Quadrado em 100 amostras de tamanho 10000	112
Tabela 76 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 10000	112
Tabela 77 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 10000	113
Tabela 78 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 10000	113
Tabela 79 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição Qui-Quadrado em 100 amostras de tamanho 10000	113
Tabela 80 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição Qui-Quadrado em 100 amostras de tamanho 10000	114

Tabela 81 – Estimativas médias do algoritmo Logaritmo Produto Uniformes em 100 amostras de tamanho 10000 da Distribuição Qui-Quadrado.....	116
Tabela 82 – Média e erro padrão dos tempos de execução, em segundos, medidos em 100 amostras de tamanho 10000 da Distribuição F	119
Tabela 83 - Média e erro padrão dos valores estimados para a esperança da distribuição F em 100 amostras de tamanho 10000	119
Tabela 84 - Média e erro padrão dos valores estimados para o máximo da distribuição F em 100 amostras de tamanho 10000	121
Tabela 85 - Média e erro padrão dos valores estimados para o mínimo da distribuição F em 100 amostras de tamanho 10000.....	121
Tabela 86 - Média e erro padrão dos valores estimados para a amplitude da distribuição F em 100 amostras de tamanho 10000	121
Tabela 87 - Média e erro padrão dos valores estimados para o desvio-padrão da distribuição F em 100 amostras de tamanho 10000	121
Tabela 88 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição F em 100 amostras de tamanho 10000	123
Tabela 89 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição F em 100 amostras de tamanho 10000	123
Tabela 90 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição F em 100 amostras de tamanho 10000	123
Tabela 91 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição F em 100 amostras de tamanho 10000	123
Tabela 92 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição F em 100 amostras de tamanho 10000.....	124
Tabela 93 – Estimativas médias do algoritmo F pela Beta em 100 amostras de tamanho 10000 da Distribuição F	126
Tabela 94 – Média e erro padrão dos valores estimados para a esperança da distribuição t de Student em 100 amostras de tamanho 100.....	130
Tabela 95 – Média e erro padrão dos valores estimados para o máximo da distribuição t de Student em 100 amostras de tamanho 100	132
Tabela 96 – Média e erro padrão dos valores estimados para o mínimo da distribuição t de Student em 100 amostras de tamanho 100	132
Tabela 97- Média e erro padrão dos valores estimados para a amplitude da distribuição t de Student em 100 amostras de tamanho 100.....	133
Tabela 98 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição t de Student em 100 amostras de tamanho 100.....	133
Tabela 99 – Média e erro padrão dos valores estimados para a assimetria da distribuição t de Student em 100 amostras de tamanho 100.....	135
Tabela 100 – Média e erro padrão dos valores estimados para a curtose da distribuição t de Student em 100 amostras de tamanho 100.....	135
Tabela 101 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição t de Student em 100 amostras de tamanho 100.....	135
Tabela 102 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição t de Student em 100 amostras de tamanho 100.....	135
Tabela 103 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição t de Student em 100 amostras de tamanho 100.....	136
Tabela 104 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição t de Student em 100 amostras de tamanho 100.....	136
Tabela 105 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição t de Student em 100 amostras de tamanho 100	136
Tabela 106 – Estimativas médias do algoritmo Normal/Qui-Quadrado em 100 amostras de tamanho 100 da Distribuição t de Student	139
Tabela 107 – Média e erro padrão dos tempos de execução, em segundos, medidos em 100 amostras de tamanho 10000 da Distribuição t de Student	140

Tabela 108 - Média e erro padrão dos valores estimados para a esperança da distribuição t de Student em 100 amostras de tamanho 10000	140
Tabela 109 - Média e erro padrão dos valores estimados para o máximo da distribuição t de Student em 100 amostras de tamanho 10000	142
Tabela 110 - Média e erro padrão dos valores estimados para o mínimo da distribuição t de Student em 100 amostras de tamanho 10000	142
Tabela 111- Média e erro padrão dos valores estimados para a amplitude da distribuição t de Student em 100 amostras de tamanho 10000	142
Tabela 112 - Média e erro padrão dos valores estimados para o desvio-padrão da distribuição t de Student em 100 amostras de tamanho 10000	142
Tabela 113 – Média e erro padrão dos valores estimados para a assimetria da distribuição t de Student em 100 amostras de tamanho 10000	144
Tabela 114 – Média e erro padrão dos valores estimados para a curtose da distribuição t de Student em 100 amostras de tamanho 10000	144
Tabela 115 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição t de Student em 100 amostras de tamanho 10000	145
Tabela 116 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição t de Student em 100 amostras de tamanho 10000	145
Tabela 117 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição t de Student em 100 amostras de tamanho 10000	145
Tabela 118 – Média e erro padrão dos valores estimados para a distância interquartílica da distribuição t de Student em 100 amostras de tamanho 10000	145
Tabela 119 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição t de Student em 100 amostras de tamanho 10000	146
Tabela 120 – Estimativas médias do algoritmo Normal/Qui-Quadrado em 100 amostras de tamanho 10000 da Distribuição t de Student	148
Tabela 121 – Algoritmos “vencedores” para cada distribuição	151

SUMÁRIO

1. INTRODUÇÃO	18
1.1 Tema e objetivos	19
1.2 Justificativas	19
2. REVISÃO TEÓRICA	20
2.1 A Distribuição Normal.....	20
2.2 A Função Gama	22
2.3 A Distribuição t de Student.....	23
2.4 A Distribuição Qui-Quadrado	25
2.5 A Distribuição F de Snedecor.....	27
2.6 A Distribuição Gama	28
2.7 Relações entre as distribuições	30
2.8 Considerações sobre as distribuições	31
3. GERAÇÃO DE NÚMEROS ALEATÓRIOS.....	32
3.1 Geração de Variáveis Aleatórias Uniformes	32
3.1.1 O Gerador de Números Pseudo-Aleatórios Mersenne Twister	33
3.2 Geração de Variáveis Aleatórias Não-Uniformes	36
3.2.1 O Método de Inversão.....	36
3.2.2 O Método de Aceitação-Rejeição	37
3.2.3 O Método do Quociente de Uniformes.....	38
3.2.4 Outros Métodos.....	38
4. METODOLOGIA PARA A ANÁLISE COMPARATIVA DOS ALGORITMOS.....	40
5. GERADORES DE NÚMEROS ALEATÓRIOS DA DISTRIBUIÇÃO NORMAL	42
5.1 Geradores baseados no Método de Inversão.....	42
5.1.1 Inversão Aproximada	42
5.1.2 Inversão pelas aproximações de Wichura	42
5.2 Geradores baseados no Método de Aceitação-Rejeição	43
5.2.1 Gerador Normal a partir da Distribuição de Cauchy	43
5.2.2 Gerador Normal a partir da Distribuição Exponencial.....	44
5.2.3 Gerador de Marsaglia e Bray (1964).....	44
5.2.4 Gerador de Kinderman e Ramage (1976).....	45
5.3 Geradores baseados no Método do Quociente de Uniformes.....	47
5.3.1 Quociente de Uniformes de Kinderman e Monahan (1977)	47
5.3.2 Quociente de Uniformes de Leva (1992)	48
5.3.3 Quociente de Uniformes exposto por Insúa, Jiménez e Martín (2009)	48
5.4 Geradores baseados em outros métodos.....	49
5.4.1 Soma de Doze Uniformes ou Método da Convolução.....	49
5.4.2 Método de Box-Muller	49
5.4.3 Variante de Marsaglia	50
5.4.4 Método Ahrens-Dieter	50
6. ESTIMATIVAS DAS AMOSTRAS DA DISTRIBUIÇÃO NORMAL.....	52
6.1 Estimativas com amostras de tamanho $n = 100$	52
6.1.1 Esperança	52
6.1.2 Máximo e Mínimo	53
6.1.3 Amplitude.....	54
6.1.4 Desvio-Padrão.....	55
6.1.5 Assimetria.....	56
6.1.6 Curtose	56

6.1.7 Quartis	57
6.1.8 Distância Interquartílica.....	59
6.1.9 Teste de Aderência	59
6.1.10 Conclusões.....	60
6.2. Estimativas com amostras de tamanho $n = 1000$	61
6.2.1 Esperança	61
6.2.2 Máximo e Mínimo	62
6.2.3 Amplitude.....	63
6.2.4 Desvio-Padrão.....	64
6.2.5 Assimetria.....	65
6.2.6 Curtose	65
6.2.7 Quartis	66
6.2.8 Distância Interquartílica.....	68
6.2.9 Teste de Aderência	68
6.2.10 Conclusões.....	69
6.3. Estimativas com amostras de tamanho $n = 10000$	70
6.3.1 Tempo de Execução	71
6.3.2 Esperança	71
6.3.3 Máximo e Mínimo	72
6.3.4 Amplitude.....	73
6.3.5 Desvio-Padrão.....	74
6.3.6 Assimetria.....	75
6.3.7 Curtose	75
6.3.8 Quartis	76
6.3.9 Distância Interquartílica.....	78
6.3.10 Teste de Aderência	78
6.3.11 Conclusões.....	79
7. GERADORES DE NÚMEROS ALEATÓRIOS DA DISTRIBUIÇÃO QUI-QUADRADO	81
7.1 Geradores baseados na Distribuição Gama	81
7.2 Geradores baseados na Distribuição Normal	82
7.2.1 Soma dos Quadrados de Normais.....	82
7.2.2 Aproximação Normal.....	83
7.3 Gerador baseado na Distribuição Uniforme e na Distribuição Normal	83
7.3.1 Logaritmo do Produto de Uniformes	83
7.4 Gerador baseado na Distribuição Exponencial e na Distribuição Normal	83
8. ESTIMATIVAS DAS AMOSTRAS DA DISTRIBUIÇÃO QUI-QUADRADO.....	85
8.1 Estimativas com amostras de tamanho $n = 100$	85
8.1.1 Esperança	85
8.1.2 Máximo e Mínimo	87
8.1.3 Amplitude.....	88
8.1.4 Desvio-Padrão.....	88
8.1.5 Assimetria.....	90
8.1.6 Curtose	91
8.1.7 Quartis	91
8.1.8 Distância Interquartílica.....	92
8.1.9 Teste de Aderência	92
8.1.10 Conclusões.....	94
8.2 Estimativas com amostras de tamanho $n = 1000$	95
8.2.1 Esperança	96
8.2.2 Máximo e Mínimo	98

8.2.3 Amplitude.....	98
8.2.4 Desvio-Padrão.....	99
8.2.5 Assimetria.....	101
8.2.6 Curtose.....	101
8.2.7 Quartis.....	101
8.2.8 Distância Interquartílica.....	102
8.2.9 Teste de Aderência.....	103
8.2.10 Conclusões.....	105
8.3. Estimativas com amostras de tamanho $n = 10000$	106
8.3.1 Tempo de Execução.....	106
8.3.2 Esperança.....	107
8.3.3 Máximo e Mínimo.....	109
8.3.4 Amplitude.....	109
8.3.5 Desvio-Padrão.....	110
8.3.6 Assimetria.....	112
8.3.7 Curtose.....	112
8.3.8 Quartis.....	112
8.3.9 Distância Interquartílica.....	113
8.3.10 Teste de Aderência.....	114
8.3.11 Conclusões.....	116
9. GERADORES DE NÚMEROS ALEATÓRIOS DA DISTRIBUIÇÃO F.....	118
9.1 Gerador baseado na Distribuição Qui-Quadrado.....	118
9.2 Gerador baseado na Distribuição Beta.....	118
10. ESTIMATIVAS DAS AMOSTRAS DA DISTRIBUIÇÃO F.....	119
10.1 Estimativas com amostras de tamanho $n = 10000$	119
10.1.1 Tempo de Execução.....	119
10.1.2 Esperança.....	119
10.1.3 Máximo e Mínimo.....	121
10.1.4 Amplitude.....	121
10.1.5 Desvio-Padrão.....	121
10.1.6 Quartis.....	123
10.1.7 Distância Interquartílica.....	123
10.1.8 Teste de Aderência.....	124
10.1.9 Conclusões.....	126
11. GERADORES DE NÚMEROS ALEATÓRIOS DA DISTRIBUIÇÃO t DE STUDENT.....	127
11.1 Método de Rejeição Polar.....	127
11.2 Rejeição da densidade $t(3)$	127
11.3 Gerador baseado na Distribuição Gama.....	128
11.4 Gerador baseado na Distribuição Normal e na Distribuição Qui-Quadrado.....	128
12. ESTIMATIVAS DAS AMOSTRAS DA DISTRIBUIÇÃO t DE STUDENT.....	130
12.1 Estimativas com amostras de tamanho $n = 100$	130
12.1.1 Esperança.....	130
12.1.2 Máximo e Mínimo.....	132
12.1.3 Amplitude.....	133
12.1.4 Desvio-Padrão.....	133
12.1.5 Assimetria.....	135
12.1.6 Curtose.....	135
12.1.7 Quartis.....	135
12.1.8 Distância Interquartílica.....	136
12.1.9 Teste de Aderência.....	136

12.1.10 Conclusões.....	138
12.2. Estimativas com amostras de tamanho $n = 10000$	139
12.2.1 Tempo de Execução	140
12.2.2 Esperança	140
12.2.3 Máximo e Mínimo	142
12.2.4 Amplitude.....	142
12.2.5 Desvio-Padrão.....	142
12.2.6 Assimetria.....	144
12.2.7 Curtose	144
12.2.8 Quartis	145
12.2.9 Distância Interquartílica.....	145
12.2.10 Teste de Aderência	146
12.2.11 Conclusões.....	148
13. ALGORITMOS DISPONÍVEIS NO SOFTWARE R.....	150
14. ALGORITMOS “VENCEDORES” PARA CADA DISTRIBUIÇÃO	151
15. CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS	152
REFERÊNCIAS	153

1. INTRODUÇÃO

Muitas vezes, profissionais da área de Ciências Exatas fazem uso de modelos probabilísticos para representar situações reais, ou então para descrever um experimento aleatório. Entretanto, mesmo com um modelo probabilístico, certas questões não podem ser resolvidas analiticamente e teremos de recorrer a estudos de simulação para obter aproximações de quantidades de interesse. De modo bastante amplo, estudos de simulação tentam reproduzir num ambiente controlado o que se passa com um problema real. Para tal propósito, a solução de um problema real envolverá a simulação de variáveis aleatórias de um ou mais modelos probabilísticos de interesse. A simulação de variáveis aleatórias supõe que o pesquisador disponha de um bom gerador de números aleatórios. Um número aleatório representa um valor da Distribuição Uniforme no intervalo $[0; 1]$.

Originalmente, os números aleatórios eram gerados manualmente (simulação manual) ou mecanicamente, usando dados, roletas, etc. Isso se tornava muito trabalhoso ou mesmo impraticável quando era necessário gerar uma quantidade muito grande de números aleatórios, da ordem de 1.000 ou 10.000. A solução moderna foi substituir esses métodos de geração por simulação em computadores digitais, utilizando números pseudo-aleatórios em vez de números aleatórios.

Os números pseudo-aleatórios são obtidos por meio de algoritmos matemáticos recursivos determinísticos. Logo, um número pseudo-aleatório gerado numa iteração dependerá do número gerado na iteração anterior e, portanto, não será realmente aleatório, originando o nome pseudo-aleatório. (BUSSAB e MORETTIN, 2006). Além disso, toda sequência, por mais longa que seja, será finita, e o conjunto de valores gerados tenderá a repetir-se.

O propósito dos geradores de números pseudo-aleatórios é produzir uma sequência de números que aparentam ser gerados aleatoriamente de uma distribuição de probabilidade previamente especificada, tais como a Uniforme, a Normal, a Exponencial, etc. Tratam-se de programas de computador cujo objetivo é imitar ou simular o comportamento típico de uma sequência de variáveis aleatórias independentes. São algoritmos específicos, sequenciais e determinísticos (se inicializados em computadores ou momentos diferentes com o mesmo estado inicial ou semente, produzem a mesma sequência de números pseudo-aleatórios).

Os geradores de números pseudo-aleatórios são ingredientes cruciais para uma grande faixa de aplicações que envolvem a simulação de variáveis aleatórias, tais como experimentos estatísticos, sistemas estocásticos, análises numéricas com métodos de Monte Carlo, algoritmos probabilísticos, jogos de computador e criptografia, entre outros. (VIEIRA, RIBEIRO e CASTRO E SOUZA, 2004). Nesses estudos, espera-se que o(s) gerador(es) de números pseudo-aleatórios envolvido(s) seja(m) o mais eficiente possível, fornecendo valores muito próximos aos do modelo de interesse, com a maior rapidez e o menor custo computacional.

1.1 Tema e objetivos

Este trabalho tem por objetivo comparar diferentes algoritmos geradores de números pseudo-aleatórios das distribuições Normal, t de Student, Qui-Quadrado e F de Snedecor através de simulação. Neste estudo, uma simulação pode ser entendida como uma particular realização de cada um desses modelos. Assim, os valores simulados podem ser considerados uma amostra aleatória de cada uma das distribuições. Tentar-se-á responder questões como, por exemplo: dentre os vários algoritmos existentes para a geração de uma distribuição normal padrão, qual o mais eficiente? Qual o mais simples e fácil de implementar? Qual o mais rápido? Que algoritmo converge mais rapidamente para a distribuição proposta?

1.2 Justificativas

As distribuições Normal, t de Student, Qui-Quadrado e F de Snedecor são largamente utilizadas em diversos procedimentos estatísticos, tais como testes de hipóteses (paramétricos e não-paramétricos), intervalos de confiança, Inferência Bayesiana, etc. Entretanto, nenhuma delas é integrável analiticamente. Todos os valores dessas distribuições estão dispostos em tabelas, construídas por meio de técnicas de cálculo numérico. Daí, decorre o interesse e a necessidade de gerar números pseudo-aleatórios dessas distribuições quando se realiza um estudo de simulação, pois as amostras obtidas são supostas apresentar esses comportamentos por hipótese. Assim, seria conveniente gerar dados que, de fato, apresentem esses tipos de distribuições.

2. REVISÃO TEÓRICA

A seguir, será feita uma revisão teórica da função gama e das distribuições de probabilidade Normal, t de Student, Qui-Quadrado, F de Snedecor e Gama. A função gama está presente em várias funções densidade de probabilidade, tais como nas distribuições t de Student, Qui-Quadrado, F de Snedecor e Gama. A Distribuição Gama será abordada pois, dependendo dos valores de seus parâmetros, esta poderá dar origem a uma Distribuição Qui-Quadrado.

2.1 A Distribuição Normal

A Distribuição Normal é uma das distribuições de probabilidade que tem maiores aplicações na Estatística. É também conhecida como Distribuição de Gauss-Moivre-Laplace.

A Distribuição Normal foi introduzida pela primeira vez por Abraham de Moivre (1667-1754) em um artigo no ano de 1733, que foi reproduzido na segunda edição de seu *The Doctrine of Chances* (1738) no contexto da aproximação de distribuições binomiais para grandes valores de n. Seu resultado foi estendido por Pierre Simon Laplace (1749-1827), em seu livro *Analytical Theory of Probabilities* (1812), no que hoje é conhecido como Teorema de Moivre-Laplace. (WIKIPEDIA, 2012).

Laplace utilizou a Distribuição Normal na análise de erros de experimentos. Em 1805, Adrien-Marie Legendre (1752-1833) introduziu o método dos mínimos quadrados. Carl Friedrich Gauss (1777-1855), que alegava ter utilizado o método desde 1794, justificou-o rigorosamente em seus trabalhos de 1809, assumindo uma Distribuição Normal para os erros de observações astronômicas. Daí, deriva o motivo pelo qual esta distribuição é também chamada de distribuição dos erros. O nome “Distribuição Normal” foi atribuído independentemente por Charles S. Peirce, Francis Galton e Wilhelm Lexis, por volta de 1875.

Além de descrever uma série de fenômenos físicos e financeiros, a Distribuição Normal possui grande uso no campo da Estatística Inferencial, sendo imprescindível para o desenvolvimento da Amostragem, Estimação por intervalo e Testes de Hipóteses. É inteiramente descrita por seus parâmetros média e desvio padrão (ou, de forma equivalente, a variância), ou seja, conhecendo-se estes, consegue-se determinar qualquer probabilidade em uma Distribuição Normal.

Um uso conhecido da Distribuição Normal é que ela serve de aproximação para o cálculo de outras distribuições de probabilidade quando o número de observações se torna grande. Essa propriedade provém do Teorema Central do Limite, que afirma que “toda soma de variáveis aleatórias independentes de média finita e variância limitada é aproximadamente Normal, desde que o número de termos da soma seja suficientemente grande”.

A função densidade de probabilidade da Distribuição Normal com média μ e variância σ^2 (de forma equivalente, desvio padrão σ) é dada por:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)} \quad \mu \in \mathfrak{R}, \sigma > 0 \quad x \in \mathfrak{R}$$

Se a variável aleatória X segue esta distribuição, escreve-se $X \sim N(\mu, \sigma^2)$. Se $\mu = 0$ e $\sigma = 1$, a distribuição é chamada de distribuição normal padrão (ou normal standard ou normal reduzida ou normal centrada) e a função de densidade de probabilidade reduz-se a:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{z^2}{2}\right)} \quad z \in \mathfrak{R}$$

Esta função densidade é tão especial que recebe uma notação própria, a qual utiliza a letra grega φ e a letra z .

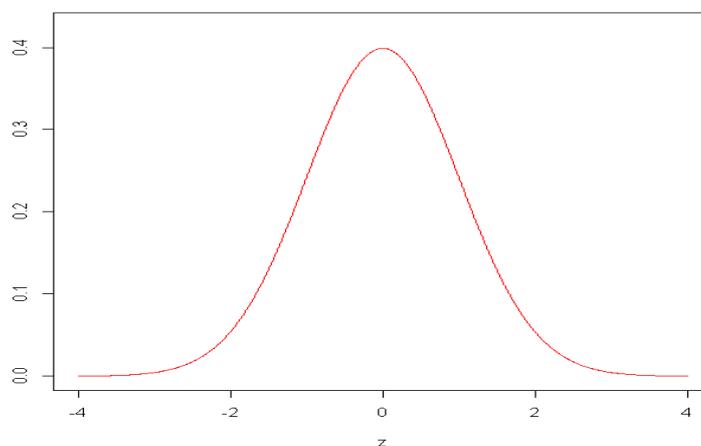


Figura 1 – Gráfico da distribuição normal padrão, com média $\mu = 0$ e variância $\sigma^2 = 1$

Fonte: elaborada pelo autor.

Uma Distribuição Normal possui as seguintes propriedades:

- Média = Mediana = μ .
- Variância = σ^2 .
- Assimetria = Curtose = 0.
- Sua função densidade de probabilidade não é integrável analiticamente, sendo necessário integração numérica ou o uso de tabelas para encontrar o valor das probabilidades.
- Os limites de sua função densidade $f(x)$ tendem a zero quando x tende a infinito.

Assim, $\lim_{x \rightarrow -\infty} f(x) = 0 = \lim_{x \rightarrow +\infty} f(x)$. Além disso, seu ponto de máximo $(x, f(x))$ é

$$\left(\mu, \frac{1}{\sigma\sqrt{2\pi}} \right).$$

- É simétrica ao redor da média.

- Cerca de 68,26% dos valores da distribuição estão compreendidos no intervalo $[\mu - \sigma, \mu + \sigma]$; 95,44%, no intervalo $[\mu - 2\sigma, \mu + 2\sigma]$; e 99,73%, no intervalo $[\mu - 3\sigma, \mu + 3\sigma]$. Este fato é conhecido como “regra 68-95-99,7” ou a “regra empírica” ou a “regra dos 3-sigmas”.

- De acordo com o Teorema Central do Limite (TCL), a soma de uma grande quantidade de variáveis aleatórias (com algumas restrições) tende a uma Distribuição Normal.

- A idéia acima resulta que a soma de variáveis aleatórias Normais é ainda Normal com média igual à soma das médias. Se as variáveis forem independentes, a variância é igual à soma das variâncias.

- Se a e b são constantes conhecidas e se X segue uma Distribuição Normal ($X \sim N(\mu, \sigma^2)$), então $aX + b \sim N(a\mu + b, a^2\sigma^2)$.

- Se X e Y são variáveis aleatórias independentes que seguem Distribuição Normal, então a soma $U = X + Y$, a diferença $V = X - Y$ ou qualquer combinação linear $W = aX + bY$ também são variáveis aleatórias com Distribuição Normal.

- Se X é uma variável aleatória $N(\mu, \sigma^2)$, então $(X - \mu) / \sigma$ é uma variável aleatória normal padrão. Reciprocamente, se X é uma variável aleatória $N(0,1)$, então $\sigma X + \mu$ é uma variável aleatória $N(\mu, \sigma^2)$.

2.2 A Função Gama

A função gama foi introduzida por Leonhard Paul Euler (1707-1783), por volta de 1730, como resultado de uma pesquisa sobre uma forma de interpolação da função fatorial. Posteriormente, foi estudada por outros matemáticos, incluindo Adrien-Marie Legendre (1752-1833), que, em 1809, a denominou de “função gama” e introduziu a notação $\Gamma(\cdot)$, utilizada atualmente.

Matematicamente, a função gama é considerada uma extensão do fatorial para o domínio dos números complexos, com exceção dos números inteiros negativos, sendo definida por:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad t \in [0; \infty) \quad x \in \mathfrak{R}^+$$

Esta expressão é também conhecida como Segunda Integral de Euler.

Graficamente, a função gama tem um comportamento estranho, especialmente para os números negativos devido aos pontos de descontinuidade. A figura abaixo mostra o gráfico da função gama apenas para os números reais positivos, os quais são de maior interesse neste trabalho.

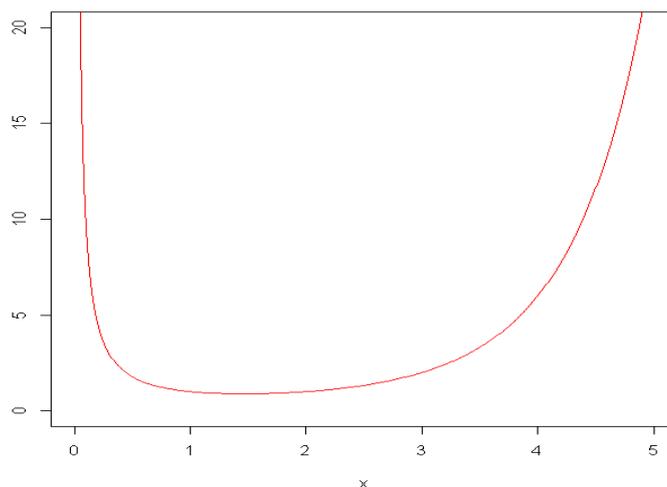


Figura 2 – Gráfico da função gama no domínio dos números reais positivos

Fonte: elaborada pelo autor.

Ela apresenta algumas propriedades peculiares:

- Se $n \in \mathbb{N}$, então $\Gamma(n) = (n-1)!$
- A função gama é recursiva, isto é, $\Gamma(x) = x\Gamma(x-1)$
- Outro resultado é dado pela fórmula de reflexão de Euler:

$$\Gamma(1-x)\Gamma(x) = \frac{\pi}{\text{sen}(\pi x)}$$

2.3 A Distribuição t de Student

A Distribuição t foi desenvolvida por William Sealey Gosset (1876-1937) em um artigo publicado na revista *Biometrics* em março de 1908, enquanto trabalhava como químico da cervejaria Guinness de Dublin, Irlanda. Como a empresa não permitia publicações com o nome real dos funcionários, ele utilizou o pseudônimo de “Student”.

Gosset descobriu que o comportamento da variável $Z = \frac{(\bar{X} - \mu_x)}{\sigma_x}$, quando o valor de σ fosse desconhecido e estimado por meio de pequenas amostras, distribuía-se simetricamente, com média zero, porém não normalmente, pois a variabilidade dependia do tamanho da amostra utilizada. Quanto menor fosse a amostra, mais variável seriam os resultados.

O modelo t de Student foi, então, caracterizado por um único parâmetro: o tamanho amostral. Em geral, esse parâmetro é especificado genericamente pela letra ν , sendo denominado de “graus de liberdade”. Quando é necessário estabelecer a relação com o tamanho da amostra, determina-se $\nu = n - 1$, ou seja, o número de elementos amostrais subtraído de uma unidade.

Todavia, a função densidade de probabilidade da que hoje é conhecida como Distribuição t de Student foi determinada por Fisher, e é definida por:

$$t_{\nu}(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu} \Gamma\left(\frac{\nu}{2}\right) \left(1 + \frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}} \quad \nu = 1, 2, 3, \dots \quad x \in \mathfrak{R}$$

na qual Γ é a função gama e ν são os graus de liberdade (VIALI e BITTENCOURT, 2007).

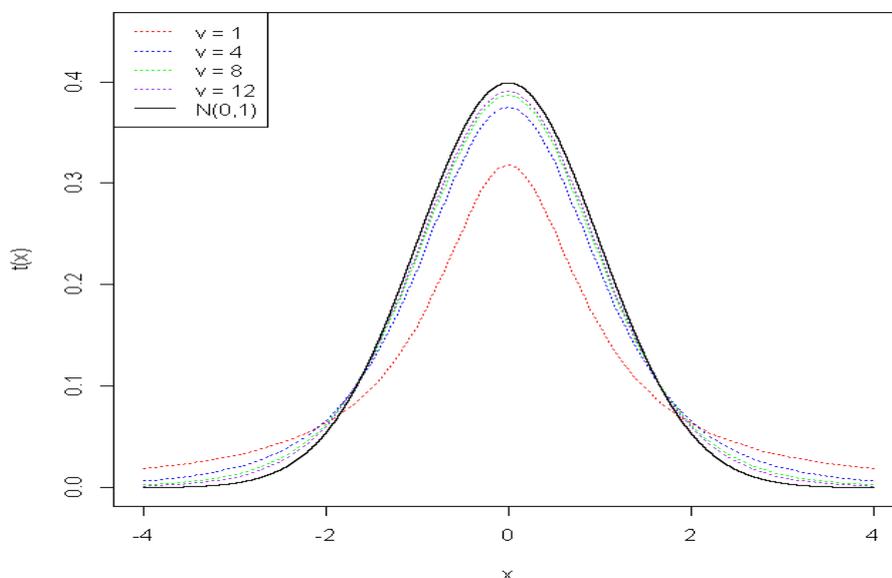


Figura 3 – Gráfico da Distribuição t de Student para alguns valores de ν comparada à Distribuição Normal (em preto)

Fonte: elaborada pelo autor.

A Distribuição t é simétrica e semelhante à curva normal padrão, porém com caudas mais altas, ou seja, uma simulação da t de Student pode gerar valores mais extremos que uma simulação da normal. Quanto maior for o valor de ν , melhor será a aproximação com a distribuição normal padrão. A função densidade da distribuição t também não é integrável analiticamente.

A Distribuição t aparece naturalmente nos casos em que o verdadeiro valor do desvio padrão da população (que supõe-se seguir a Distribuição Normal) é desconhecido, e sua estimativa a partir dos dados amostrais, dada por $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$, é utilizada para substituir o valor de σ_x (desconhecido) na equação do erro padrão $\sigma_x = \sigma_x / \sqrt{n}$.

Supondo que o tamanho da amostra n seja extraído de uma população normal, temos que a amostra é formada por n variáveis aleatórias normais independentes X_1, \dots, X_n , cuja média

amostral $\bar{X} = (X_1 + \dots + X_n)/n$ é considerada o melhor estimador para a média μ da população. Considerando $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ como a variância amostral, temos o seguinte

resultado:

A variável aleatória t dada por:

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}} \text{ ou } t = \sqrt{n} \frac{\bar{X} - \mu}{S}, \text{ tem uma Distribuição } t \text{ de Student com } \nu = n - 1 \text{ graus}$$

de liberdade.

O resultado acima também é utilizado no teste paramétrico t para comparação de duas médias populacionais, supondo que ambas as populações são normalmente distribuídas.

A Distribuição t possui as seguintes propriedades:

- Média = 0, para $\nu > 1$.
- Mediana = 0.
- Variância = $\frac{\nu}{\nu - 2}$, para $\nu > 2$.
- Assimetria = 0, para $\nu > 3$.
- Curtose = $\frac{6}{\nu - 4}$, para $\nu > 4$.

2.4 A Distribuição Qui-Quadrado

A Distribuição Qui-Quadrado ou χ^2 foi desenvolvida inicialmente, de acordo com Upton e Cook (2002), pelo físico alemão Ernst Carl Abbe (1840-1905) em 1863 e, de forma independente, pelo engenheiro geodesta alemão Friedrich Robert Helmert (1843-1917) em 1875. O modelo foi batizado e popularizado pelo estatístico britânico Karl Pearson (1857-1936) em 1900, ano no qual ele desenvolveu uma das aplicações mais populares do modelo: o teste de aderência.

A Distribuição Qui-Quadrado também é definida por um único parâmetro ν que está relacionado ao tamanho amostral ($\nu = n - 1$). Outra maneira de caracterizá-la é como a soma dos quadrados de ν normais padrão independentes (VIALI e BITTENCOURT, 2007). Por definição, se Z_1, Z_2, \dots, Z_ν forem ν distribuições normais padronizadas (ou seja, com média zero e variância unitária) independentes, então a soma de seus quadrados é uma Distribuição Qui-Quadrado com ν graus de liberdade:

$$\chi_\nu^2 = Z_1^2 + Z_2^2 + \dots + Z_\nu^2$$

Um corolário imediato da definição acima é que a soma de duas Qui-Quadrado independentes também é uma Qui-Quadrado:

$$\chi_a^2 + \chi_b^2 = \chi_{a+b}^2$$

William Gemmell Cochran (1909-1980) também mostrou que a Distribuição Qui-Quadrado é igual, a menos de uma constante, a distribuição da variância amostral.

Sua função densidade de probabilidade está definida apenas para valores não-negativos de x e, assim como a Distribuição t , depende dos graus de liberdade ν conforme segue:

$$\chi_\nu^2(x) = \frac{x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} \quad \nu = 1, 2, 3, \dots \quad x \in [0; \infty)$$

na qual Γ é a função gama definida por:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad t \in [0; \infty) \quad x \in \mathfrak{R}^+$$

e ν são os graus de liberdade (VIALI e BITTENCOURT, 2007).

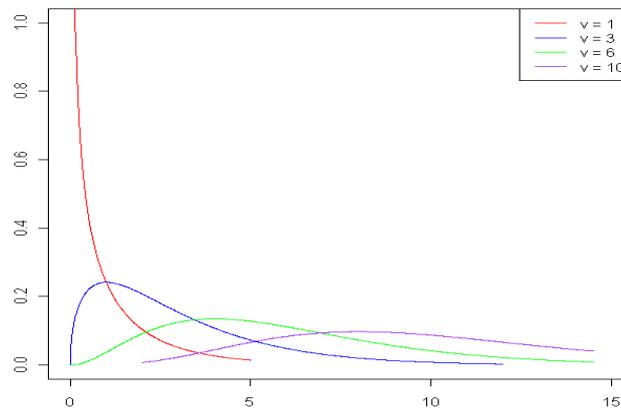


Figura 4 – Gráfico da Distribuição Qui-Quadrado para alguns valores de ν

Fonte: elaborada pelo autor.

A Distribuição Qui-Quadrado possui as seguintes propriedades:

- Média = ν .
- Mediana $\approx \nu \left(1 - \frac{2}{9\nu}\right)^3$.
- Variância = 2ν .
- Assimetria = $\sqrt{8/\nu}$.

$$\bullet \text{ Curtose} = \frac{12}{\nu}.$$

A Distribuição Qui-Quadrado é muito utilizada em testes de aderência e de independência e sua função densidade, assim como a Distribuição Normal e a Distribuição t de Student, também não é integrável analiticamente.

2.5 A Distribuição F de Snedecor

O modelo F de Snedecor foi inicialmente desenvolvido por Ronald Aylmer Fisher (1890-1962) em 1922 e, por isso, ele é também conhecido por Distribuição de Fisher ou por Distribuição de Fisher-Snedecor. Em 1934, foi tabelado por George Waddel Snedecor (1881-1974), que também introduziu a letra F para representá-lo, homenageando dessa forma o seu real criador.

A Distribuição F de Snedecor também depende de dois parâmetros denominados “graus de liberdade”. O primeiro (m) é o grau de liberdade do numerador e o segundo (n) do denominador. Na estatística, ela é caracterizada como o quociente de duas variâncias e, portanto, de duas distribuições Qui-Quadrado. Cada parâmetro, da mesma forma que nos modelos anteriores, é associado ao tamanho amostral subtraído de uma unidade. Assim, $m = n_a - 1$ e $n = n_b - 1$, onde n_a representa o número de unidades amostrais.

A função densidade de probabilidade da F não é integrável analiticamente e uma de suas possíveis representações, em termos da função gama, é dada por:

$$F_{m,n}(x) = \frac{\Gamma\left(\frac{m+n}{2}\right) m^{\frac{m}{2}} n^{\frac{n}{2}} x^{\frac{m}{2}-1}}{\Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right) (mx+n)^{\frac{m+n}{2}}} \quad m, n = 1, 2, 3, \dots \quad x \in [0; \infty)$$

na qual Γ é a função gama definida por:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad t \in [0; \infty) \quad x \in \mathfrak{R}^+$$

e m e n são os graus de liberdade.

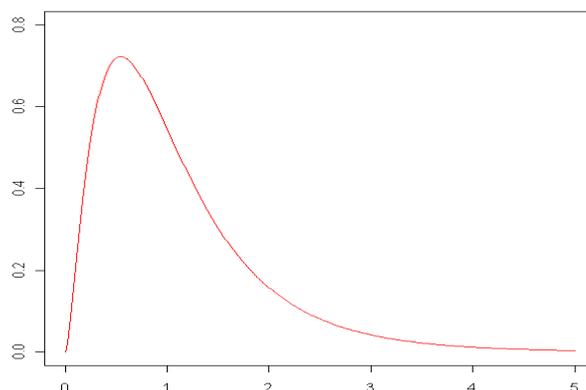


Figura 5 – Gráfico da Distribuição F de Snedecor para $m = 5$ e $n = 20$

Fonte: elaborada pelo autor.

A Distribuição F possui as seguintes propriedades:

- Média = $\frac{n}{n-2}$, para $n > 2$.
- Variância = $\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$, para $n > 4$.
- Assimetria = $\frac{(2m+n-2)\sqrt{8(n-4)}}{(n-6)\sqrt{m(m+n-2)}}$, para $n > 6$.
- Se X é uma variável aleatória $F(m,n)$, então $1/X$ é uma variável aleatória $F(n,m)$.

O modelo F é fundamental, em Estatística, para as áreas de Regressão, Planejamento de Experimentos e Análise de Variância (VIALI e BITTENCOURT, 2007).

2.6 A Distribuição Gama

A Distribuição Gama é uma distribuição de probabilidade contínua, com dois parâmetros. Há duas maneiras diferentes de representá-la:

(1) Com um parâmetro de forma k e um parâmetro de escala θ .

(2) Com um parâmetro de forma $\alpha = k$ e um parâmetro de escala inversa $\beta = 1/\theta$,

chamado de parâmetro de proporção.

Caracterizando a Distribuição Gama usando a forma k e a escala θ , temos que uma variável aleatória X , seguindo Distribuição Gama, é denotada por:

$$X \sim G(k, \theta) \equiv \text{Gama}(k, \theta)$$

Analogamente, pela caracterização (2), temos que uma variável aleatória X , seguindo Distribuição Gama, é denotada por:

$$X \sim G(\alpha, \beta) \equiv \text{Gama}(\alpha, \beta)$$

Sua função densidade de probabilidade pode ser expressa em termos da função gama parametrizada pela forma k e pela escala θ . Ambos os parâmetros são valores positivos.

A equação que define a função densidade da Distribuição Gama pela forma e pela escala é:

$$f(x; k, \theta) = \frac{1}{\theta^k} \frac{1}{\Gamma(k)} x^{k-1} e^{-\frac{x}{\theta}} \quad x \geq 0 \text{ e } k, \theta > 0$$

Já pela caracterização (2), a função densidade da Distribuição Gama é dada por:

$$f(x; \alpha, \beta) = \beta^\alpha \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad x \geq 0 \text{ e } \alpha, \beta > 0$$

Ambas as representações podem ser utilizadas no software estatístico R, e são comumente encontradas, pois uma ou outra pode ser mais conveniente dependendo da situação.

Algumas propriedades da Distribuição Gama:

- Se X_i tem uma distribuição $G(k_i, \theta)$, para $i = 1, 2, \dots, n$ (isto é, todas as distribuições têm o mesmo parâmetro de escala θ), então $\sum_{i=1}^n X_i \sim G\left(\sum_{i=1}^n k_i, \theta\right)$, se todos X_i são independentes.

- Se $X \sim G(k, \theta)$, então $cX \sim G(k, c\theta)$.

- Se $Y \sim G(\alpha, 1)$, então $X = Y/\beta$ tem distribuição $G(\alpha, \beta)$.

2.7 Relações entre as distribuições

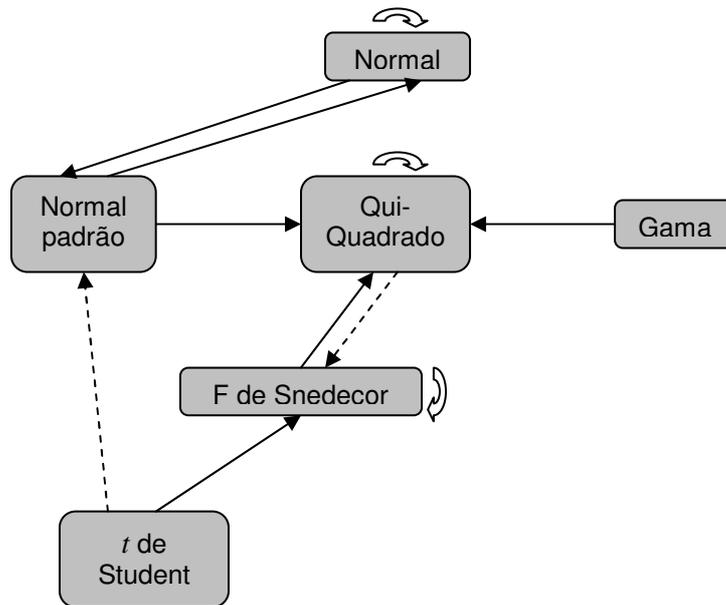


Figura 6 – Diagrama ilustrativo das relações entre as distribuições

Fonte: elaborada pelo autor.

As distribuições de probabilidade elencadas acima estão inter-relacionadas da seguinte forma:

- Normal / Normal padrão: Se X é uma variável aleatória $N(\mu, \sigma^2)$, então $(X - \mu)/\sigma$ é uma variável aleatória normal padrão. Reciprocamente, se X é uma variável aleatória normal padrão, então $\mu + \sigma X$ é uma variável aleatória $N(\mu, \sigma^2)$.

- Gama / Normal: Se X é uma variável aleatória $G(k, \theta)$ e Y é uma variável aleatória Normal com a mesma média e a mesma variância que X , então $F_X \approx F_Y$ se o parâmetro de forma k é grande em relação ao parâmetro de escala θ .

- Gama / Qui-Quadrado: Se X é uma variável aleatória $G(k, \theta)$ com $k = \nu/2$ e $\theta = 2$, então X é uma variável aleatória Qui-Quadrado com ν graus de liberdade. De forma análoga, se X é uma variável aleatória $G(\alpha, \beta)$ com $\alpha = \nu/2$ e $\beta = 1/2$, então X é uma variável aleatória Qui-Quadrado com ν graus de liberdade. Reciprocamente, se $Q \sim \chi^2(\nu)$ e c é uma constante positiva, então $c.Q \sim G(k = \nu/2, \theta = 2c)$.

- t de Student / Normal: Se X é uma variável aleatória t com um número grande de graus de liberdade ν , então $F_X \approx F_Y$, onde Y é uma variável aleatória normal padrão.

• F de Snedecor / Qui-Quadrado: Se X é uma variável aleatória $F(m, n)$ com n grande, então mX é distribuída aproximadamente como uma variável aleatória Qui-Quadrado com ν graus de liberdade.

• Qui-Quadrado / F de Snedecor: Se X_1 e X_2 são variáveis aleatórias Qui-Quadrado com ν_1 e ν_2 graus de liberdade respectivamente, então $(X_1/\nu_1)/(X_2/\nu_2)$ é uma variável aleatória $F(\nu_1, \nu_2)$.

• t de Student / F de Snedecor: Se X é uma variável aleatória t com ν graus de liberdade, então X^2 é uma variável aleatória $F(1, \nu)$; se X é uma variável aleatória t com ν graus de liberdade, então X^{-2} é uma variável aleatória $F(\nu, 1)$.

A tabela a seguir apresenta a fórmula de cálculo para algumas medidas estatísticas, baseada nos valores dos parâmetros das distribuições.

Tabela 1 – Medidas estatísticas das distribuições Normal, t de Student, Gama, Qui-Quadrado e F de Snedecor

Medidas	Normal	t de Student	Gama	Qui-Quadrado	F de Snedecor
Média	μ	$0, \nu > 1$	$k\theta$	ν	$\frac{n}{n-2}, n > 2$
Mediana	μ	0	-	$\approx \nu \left(1 - \frac{2}{9\nu}\right)^3$	-
Variância	σ^2	$\frac{\nu}{\nu-2}, \nu > 2$	$k\theta^2$	2ν	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}, n > 4$
Assimetria	0	$0, \nu > 3$	$\frac{2}{\sqrt{k}}$	$\sqrt{8/\nu}$	$\frac{(2m+n-2)\sqrt{8(n-4)}}{(n-6)\sqrt{m(m+n-2)}}, n > 6$
Curtose	0	$\frac{6}{\nu-4}, \nu > 4$	$\frac{6}{k}$	$12/\nu$	-

Fonte: elaborada pelo autor.

2.8 Considerações sobre as distribuições

Pode-se observar que a origem dos quatro modelos é diversa, mas que, de fato, apresentam algumas características comuns. Os quatro modelos apresentam aplicações na Estatística Inferencial e desempenham o principal papel num variado leque de técnicas estatísticas tanto paramétricas quanto não-paramétricas. As características mais notórias que esses quatro modelos apresentam são: possuir funções densidade de probabilidade de difícil manuseio, não integráveis analiticamente; e dependência da função gama (com exceção da Normal).

3. GERAÇÃO DE NÚMEROS ALEATÓRIOS

Para gerar realizações de uma distribuição de probabilidade específica, como a Distribuição Uniforme, por exemplo, precisamos gerar números aleatórios. Isso não pode ser realizado por máquinas, pois, na verdade, qualquer sequência produzida por uma máquina é na realidade uma sequência previsível. Daí, a denominação de sequência de números pseudo-aleatórios.

Uma sequência de números será considerada “aleatória” do ponto de vista computacional se o programa que a gerar for diferente e estatisticamente não correlacionado com o programa que a usará (FERREIRA, 2010).

Basicamente, a geração de números “aleatórios”, o que constitui o chamado Método de Monte Carlo, pode ser dividida em dois tipos: geração de números (variáveis) aleatórias uniformes e geração de números (variáveis) aleatórias não-uniformes.

3.1 Geração de Variáveis Aleatórias Uniformes

Variáveis aleatórias uniformes são aquelas que, a princípio, se situam dentro de uma determinada amplitude, geralmente entre 0 e 1, para os quais não podemos produzir uma sequência previsível de valores. Em várias linguagens estes números são gerados utilizando o comando “random” ou comandos similares. Na linguagem de programação Pascal, por exemplo, se este comando for utilizado com o argumento “n”, “random(n)”, números aleatórios inteiros U do intervalo $0 \leq U \leq n-1$ são gerados e se o argumento n não for usado, os números gerados são valores aleatórios reais do intervalo [0; 1].

Em simulação estocástica, as variáveis aleatórias com Distribuição Uniforme no intervalo [0; 1] são empregadas de muitas maneiras:

- Em forma direta.
- Para gerar outras distribuições discretas ou contínuas, pois a partir do modelo Uniforme podemos gerar realizações de variáveis aleatórias de qualquer outro modelo probabilístico, com base em transformações realizadas nos números aleatórios uniformes.
- Para gerar conjuntos de variáveis aleatórias dependentes (processos estocásticos e/ou distribuições multivariadas).

Devemos ter bem em claro que, na maioria das vezes, a performance de uma simulação estará fortemente correlacionada com a do gerador de uniformes usado. (BUSTOS e FRERY, 1992). Então, uma fonte confiável para gerar números aleatórios uniformes determina o sucesso de métodos estocásticos de inferência e de todo o processo de simulação Monte Carlo. (FERREIRA, 2010).

Segundo Costa Vieira, Ribeiro e Castro e Souza (2004), um bom gerador de números aleatórios uniformes deveria possuir as seguintes propriedades:

- Uniformidade – A sequência de números aleatórios deve passar em testes estatísticos com a finalidade de verificar a uniformidade da distribuição.

- Independência – Subseqüências da sequência completa u_0, u_1, \dots devem ser independentes. Por exemplo, membros da subseqüência par u_0, u_2, u_4, \dots devem ser independentes dos seus vizinhos ímpares u_1, u_3, \dots . Desta maneira, a sequência de pares (u_{2n}, u_{2n+1}) deve ser uniformemente distribuída no quadrado unitário.

- Período longo – O gerador deve possuir um período longo, ou seja, começar a repetir valores somente após que uma sequência muito grande de números aleatórios foi gerada. Idealmente, o gerador não deve repetir valores. Na prática, a repetição deve ocorrer somente depois da geração de um grande conjunto de valores.

- Facilidade de implementação e eficiência – Os geradores devem ser fáceis de serem implementados em uma linguagem de programação de alto nível e eficientes, isto é, utilizar poucas operações aritméticas para gerar cada número aleatório, usar todas as capacidades vetoriais/paralelas disponíveis na máquina e minimizar “overheads”, tais como chamadas a subrotinas.

- Repetição – Os geradores devem possuir a habilidade de repetir exatamente a mesma sequência de números aleatórios, pois isto é relevante em procedimentos de teste e desenvolvimento de programas.

- Portabilidade – Os geradores devem ser portáveis, isto é, gerar exatamente a mesma sequência de números aleatórios em duas máquinas diferentes, possivelmente com diferentes tamanhos de palavras.

- Subseqüências disjuntas – Os geradores devem dispor de métodos eficientes para gerar um número aleatório sem ter que passar por todos os estados intermediários. Essa característica é necessária para a utilização dos geradores de números aleatórios em ambientes de processamento paralelo, permitindo, assim, particionar a sequência em subfluxos disjuntos.

Bustos e Frery (1992) estendem o conceito de portabilidade para o âmbito das linguagens de programação: *“Portabilidade significa que, sob as mesmas condições definidoras, a sequência seja a “mesma”, independentemente da linguagem computacional usada para implementar o algoritmo de geração, e do computador usado. Na verdade, isto é bem difícil de ser atingido, particularmente com microcomputadores.”*

De fato, é sabido que os geradores de números aleatórios não conseguem atender todos os requerimentos expostos acima em sua plenitude.

3.1.1 O Gerador de Números Pseudo-Aleatórios Mersenne Twister

Segundo Ferreira (2010), um dos melhores geradores de números aleatórios é o Mersenne Twister (MT). O Mersenne Twister é um gerador de números pseudo-aleatórios

desenvolvido por Makoto Matsumoto e Takuji Nishimura, da Universidade Keio (Japão), nos anos de 1996 e 1997. Seu nome deriva do fato que o tamanho escolhido para o período do gerador é um primo de Mersenne. Primo de Mersenne é um número de Mersenne (número da forma $M_n = 2^n - 1$, com “n” sendo número natural) que também é um número primo.

O MT possui as seguintes características segundo seus desenvolvedores:

- Foi desenvolvido para eliminar as falhas dos diferentes geradores existentes.
- Apresenta o maior período e maior ordem de equi-distribuição do que de qualquer outro método implementado. Ele fornece um período que é da ordem de $2^{19.937} - 1 \approx 4,3154 \times 10^{6001}$, e uma equidistribuição 623-dimensional.
- É um dos mais rápidos geradores existentes, embora seu algoritmo seja complexo.
- Faz uso eficiente da memória.

O Mersenne Twister foi criado para simulações com Método de Monte Carlo e outras simulações estatísticas, não sendo adequado, em sua forma nativa, para tarefas de criptografia.

Existem muitas versões implementadas deste algoritmo, inclusive em linguagens como C e Fortran, e que estão disponíveis na Internet. Este algoritmo é utilizado como gerador padrão no software estatístico R, e está disponível em outros softwares do gênero, tais como SPSS (a partir da versão 13) e SAS. Também pode ser encontrado nas bibliotecas padrões das últimas versões das linguagens PHP, Python e Ruby.

A fim de testá-lo, gerou-se 1.000.000 de valores com Distribuição Uniforme [0; 1]. O resultado pode ser observado no histograma apresentado na figura 7:

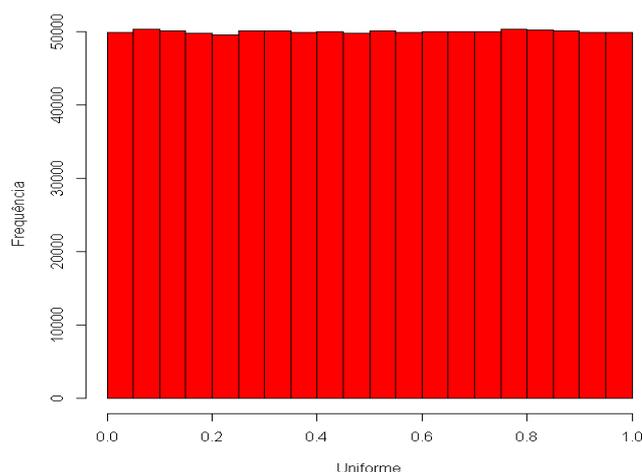


Figura 7 – Histograma de 1.000.000 números aleatórios gerados pelo algoritmo do Mersenne Twister

Fonte: elaborada pelo autor.

Como se pode verificar graficamente, não há evidências de que os dados estejam concentrados mais próximos a um determinado valor. Assim, não há indícios de falha no gerador, pois o histograma se apresentou muito próximo da Distribuição Uniforme.

Outra possível falha em geradores de números aleatórios é a de que dois valores consecutivos sejam correlacionados. Pela figura 8, podemos observar o gráfico de dispersão de 2000 pares consecutivos de números aleatórios e, como se pode observar, o diagrama está uniformemente preenchido (não existem buracos):

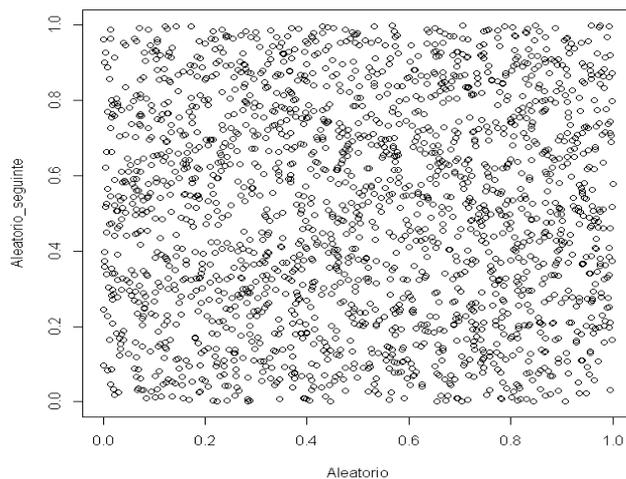


Figura 8 – Gráfico de dispersão com 2000 pares de números aleatórios gerados em sequência

Fonte: elaborada pelo autor.

Na figura 9 é apresentado o mesmo gráfico de dispersão para 32500 pares, onde se pode verificar que o diagrama está completamente preenchido, não se percebendo falhas aparentes:

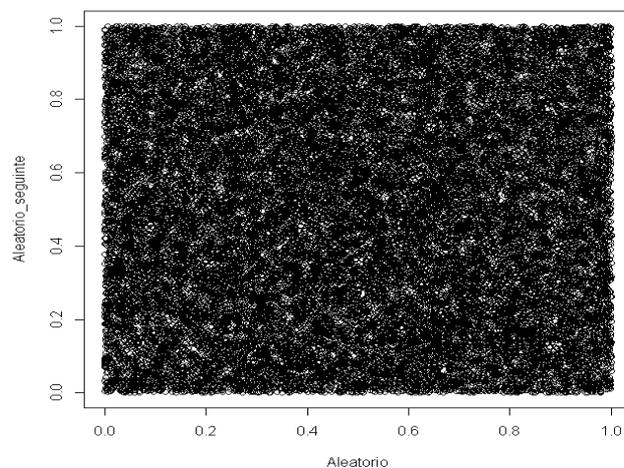


Figura 9 – Gráfico de dispersão com 32500 pares de números aleatórios gerados em sequência

Fonte: elaborada pelo autor.

3.2 Geração de Variáveis Aleatórias Não-Uniformes

É comum precisarmos de eventos provenientes de variáveis aleatórias que obedecem outras distribuições, além da Uniforme. Podemos obter variáveis aleatórias de qualquer distribuição de probabilidade a partir de números aleatórios uniformes, por essa razão a geração de números uniformes é necessária. Veremos alguns métodos de validade universal, que serão referenciados neste trabalho, os quais utilizam variáveis aleatórias com Distribuição Uniforme em $[0; 1]$ para construir variáveis aleatórias com qualquer distribuição.

3.2.1 O Método de Inversão

Em sua forma mais simples, o Método de Inversão se baseia no seguinte teorema, chamado de Teorema Fundamental da Transformação de Probabilidades:

Teorema: *Sejam U uma variável uniforme $U[0; 1]$ e X uma variável aleatória com densidade f e função de distribuição F contínua e invertível, então $X = F^{-1}(U)$ possui densidade f . Sendo F^{-1} a função inversa da função de distribuição F .*

Este teorema sugere que para gerar amostras de uma variável aleatória X da qual se conhece F^{-1} , se podem gerar números U uniformes em $[0; 1]$ e fazer $X = F^{-1}(U)$. Temos, então, o seguinte algoritmo geral de inversão:

1. Gerar $U \sim U[0;1]$.
2. Fazer $X = F^{-1}(U)$.
3. Sair X .

Exemplo: Geração de valores com Distribuição Exponencial.

A função densidade de probabilidade da Distribuição Exponencial é dada por:

$$f(x; \lambda) = \lambda e^{-\lambda x} \quad x \geq 0 \text{ e } \lambda > 0$$

E sua função de distribuição acumulada é dada por:

$$F(x; \lambda) = 1 - e^{-\lambda x} \quad x \geq 0 \text{ e } \lambda > 0$$

Então, para gerar números aleatórios que seguem uma Distribuição Exponencial, pelo Teorema Fundamental da Transformação de Probabilidades, podemos substituir $F(x; \lambda)$ por U . Assim:

$$U = 1 - e^{-\lambda x}$$

Isolando x , teremos:

$$X = \frac{-\ln(1-U)}{\lambda}$$

Devido a Distribuição Uniforme ser simétrica, podemos substituir $1-U$ por U , e o algoritmo definitivo fica:

1. Gerar $U \sim U[0;1]$.
2. Fazer $X = \frac{-\ln(U)}{\lambda}$.

3. Sair X .

Também há outras distribuições conhecidas para as quais o Método de Inversão pode ser aplicado, tais como a Distribuição Weibull, a Distribuição de Cauchy e a Distribuição de Pareto.

Para variáveis aleatórias discretas, devemos modificar o teorema para podermos contemplar funções de distribuições F em escada, como são as funções de distribuição de probabilidades associadas a essas variáveis aleatórias (FERREIRA, 2010).

Uma condição mínima para a aplicação deste método é conhecer a forma explícita de F^{-1} , através de boas implementações. Entretanto, quando não conhecemos explicitamente a forma de F^{-1} , mas dispomos de uma boa aproximação da mesma, podemos utilizar o método por aproximação.

Computacionalmente, a dificuldade é obtermos analiticamente uma expressão para a função F^{-1} para muitos modelos probabilísticos, tais como a Distribuição Normal, a Distribuição t , a Distribuição Qui-Quadrado e a Distribuição F , dentre outras. Em geral, suas expressões analíticas não existem e métodos numéricos são requeridos para inverter a função de distribuição.

3.2.2 O Método de Aceitação-Rejeição

Para o Método de Inversão é conveniente conhecer a função de distribuição. Em algumas ocasiões, conhecemos a função densidade, mas não a função de distribuição em forma tratável, como ocorre, por exemplo, com a Distribuição Normal. Em outras situações, a inversa da função de distribuição não é tratável. Um método muito geral para lidar com alguns desses casos é o que se baseia na aceitação e rejeição, chamado de Método de Aceitação-Rejeição ou, simplesmente, de Método de Rejeição. Esse método foi introduzido por Von Neumann (1951).

Suponhamos que desejemos gerar variáveis aleatórias X de uma função de densidade f , mas não sabemos como fazer isso diretamente. Porém, dispomos de um procedimento para gerar variáveis aleatórias de uma função de densidade g tal que $f(x) \leq ag(x)$ para todo x , sendo $a < \infty$. O método de rejeição é assim definido:

Enquanto $U > f(X)/(ag(X))$

Gerar $X \sim g$, $U \sim U[0;1]$.

Sair X .

Observamos que o método de rejeição equivale a gerar valores $Y \sim U(0; ag(X))$ e aceitá-los se $Y \leq f(X)$. Então, tem-se que o algoritmo de rejeição proporciona saídas X tais que $P(X \leq x | X \text{ aceito}) = F(x)$, onde F é a função de distribuição de X (INSÚA, JIMÉNEZ e MARTÍN, 2009).

O problema deste método é que devemos encontrar g e, em consequência, a . Segundo Bustos e Frery (1992), “a arte no uso dos métodos de rejeição consiste em encontrar a g conveniente”.

Este método também pode ser aplicado à geração de variáveis aleatórias discretas, mas, nesses casos, é bem difícil achar funções g adequadas (BUSTOS e FRERY, 1992).

3.2.3 O Método do Quociente de Uniformes

O método de rejeição apresenta o inconveniente de ter que ser extremamente cuidadoso na seleção da densidade g no caso de distribuições com caudas pesadas. Esse problema não se dá com o Método do Quociente de Uniformes.

É conhecido que quando (U, V) se distribuem uniformemente no círculo unitário, V/U segue uma Distribuição de Cauchy. Surge, então, de forma natural, a questão de que se poderia ser possível gerar variáveis aleatórias de outras distribuições como o quociente de variáveis distribuídas uniformemente sobre certo subconjunto. Tem-se, de fato, o seguinte resultado:

Proposição: *Seja h uma função não negativa com $0 < \int h < \infty$. Seja*

$$C_h = \{(u, v) : 0 \leq u \leq \sqrt{h(v/u)}\}$$

Se (U, V) se distribui uniformemente sobre C_h , então $X = V/U$ tem função de densidade $h / (\int h)$.

Podemos fornecer o seguinte algoritmo:

Enquanto $(U, V) \notin C_h$

Gerar $U_1, U_2 \sim U[0;1]$.

Fazer $U = mU_1, V = p^i + (p^s - p^i)U_2$.

Sair $X = V/U$.

3.2.4 Outros Métodos

Na literatura, são relatados vários outros métodos para geração de números pseudo-aleatórios de um modelo probabilístico com base na Distribuição Uniforme $[0; 1]$. Um dos

métodos mais conhecidos é o Método de Composição, utilizado no caso em que a distribuição da qual desejamos gerar valores é uma mistura, que envolve uma família de densidades dependentes de um parâmetro y e sua função de distribuição H da seguinte forma:

$$f(x) = \int g(x | y) dH(y)$$

Porém, o Método de Composição é aplicável somente para modelos discretos.

4. METODOLOGIA PARA A ANÁLISE COMPARATIVA DOS ALGORITMOS

Serão geradas amostras de tamanhos diferentes: 100, 1000 e 10000 para as distribuições Normal e Qui-Quadrado; 100 e 10000 para a distribuição t de Student; e 10000 para a distribuição F. A razão pela escolha destes valores foi por serem os mais referenciados na bibliografia de Estatística Computacional e Simulação. Cada um destes tamanhos amostrais será replicado 100 vezes. Assim, teremos 100 amostras de tamanho 100, 100 amostras de tamanho 1000 e sucessivamente.

Para o caso da Normal só serão gerados valores para a Normal padrão, pois todos os algoritmos, inclusive a implementação no R, utilizam a relação $X = \mu + \sigma Z$ para gerar valores para as distribuições não padronizadas. Assim, a geração de uma $N(-2,4)$ é diretamente dependente do algoritmo para geração de uma $N(0,1)$, da mesma forma que a geração da $N(0,1)$ depende do gerador da uniforme no intervalo $[0; 1]$. Para as distribuições χ^2 , F e t serão utilizados os seguintes valores, respectivamente: $\chi^2(1)$, $\chi^2(17)$, $\chi^2(50)$; F(1,5), F(5,1), F(5,5); t(3), t(15), t(30).

Para cada amostra, serão calculadas as seguintes estatísticas: tempo de execução, esperança, máximo e mínimo, amplitude, desvio-padrão, assimetria, curtose, quartis, distância interquartilica e valor-p para o teste de aderência. Lembrando que o desvio-padrão é a raiz quadrada da variância e o 2º quartil é igual à mediana dos dados.

A aderência ao modelo será avaliada pelo Teste Qui-Quadrado de Aderência. Os níveis de significância utilizados serão os valores 5% e 1%.

O tempo de execução será determinado como sendo o tempo médio da geração das 100 amostras de tamanho 10000. Além do tempo médio, será determinado o erro padrão da média dos tempos. Para as amostras menores o tempo de execução não será avaliado em virtude de ele ser praticamente nulo. O algoritmo mais rápido será considerado o que apresentar o menor tempo médio, com o menor erro padrão.

As demais medidas serão avaliadas quanto à variação em torno dos valores do modelo. Assim, para cada uma das 100 amostras serão calculadas as estatísticas mencionadas acima (quando conveniente) e comparadas com os valores “exatos” do modelo sendo simulado. Estes valores “exatos” foram obtidos com auxílio do software R. As estatísticas para a esperança, a mediana (Distribuição Normal), o desvio-padrão e o valor-p também serão representadas graficamente. Um valor médio, com um erro padrão, será obtido para cada uma das estatísticas. A análise comparativa entre os algoritmos será feita com base nestes valores-resumo, os quais poderão ser apresentados com um número diferente de casas decimais ou em notação científica (conforme a precisão alcançada pelo software R).

Para o algoritmo “vencedor” em cada tamanho de amostra, será gerado um histograma e apresentado os valores calculados.

As análises serão realizadas no software R, versão 2.14.2, através do editor Tinn-R, versão 2.3.7.1, em ambiente Windows. O computador utilizado terá um processador de 2.13GHz e 2GB de memória RAM.

5. GERADORES DE NÚMEROS ALEATÓRIOS DA DISTRIBUIÇÃO NORMAL

Neste trabalho, consideraremos diferentes algoritmos para gerar números aleatórios da Distribuição Normal padrão ($X \sim N(0,1)$) disponíveis na literatura de Estatística Computacional e Simulação. Como indicamos na seção 2.1, se desejarmos gerar números da Distribuição Normal $Y \sim N(\mu, \sigma^2)$, basta fazer a transformação $Y = \mu + \sigma X$. Na maioria das vezes, a construção destes algoritmos foi baseada nos métodos para geração de variáveis aleatórias não-uniformes mencionados anteriormente.

5.1 Geradores baseados no Método de Inversão

Embora a função densidade do modelo Normal não seja integrável analiticamente, se conhecem várias aproximações para a sua função de distribuição, o que nos permite utilizar o Método de Inversão por aproximação.

5.1.1 Inversão Aproximada

Na literatura, uma fórmula proposta para a inversa da função de distribuição do modelo Normal é dada por:

$$F^{-1}(X) = \frac{X^{0.135} - (1-X)^{0.135}}{0.1975}$$

Então, pelo Método de Inversão, temos o seguinte algoritmo:

1. Gerar $U \sim U[0;1]$.
2. Fazer $X = \frac{U^{0.135} - (1-U)^{0.135}}{0.1975}$.
3. Sair X .

Embora o algoritmo seja baseado em uma fórmula aproximada e não em um cálculo exato, esse é um dos geradores para a Distribuição Normal de mais simples e fácil implementação em qualquer ambiente computacional. Quanto à eficiência, existe uma divergência entre os autores. Insúa, Jiménez e Martín (2009) afirmam que este algoritmo é mais rápido que os demais. Já Ferreira (2010) menciona que a operação é lenta, em virtude de utilizar a exponenciação.

5.1.2 Inversão pelas aproximações de Wichura

Em 1988, Wichura descreveu duas aproximações rápidas e de alta precisão para a inversa da função de distribuição do modelo Normal utilizando polinômios racionais. Para valores de x no intervalo $[0,075; 0,925]$, um polinômio racional em $(x-0,5)^2$ foi utilizado,

enquanto que para valores fora desse intervalo, um dos dois polinômios racionais em $\sqrt{-\ln x}$ foi utilizado. Como a maioria dos valores de x está no primeiro intervalo, a raiz quadrada e o logaritmo só precisam ser calculados em 15% das vezes. A primeira aproximação, chamada de PPND7, fornece 7 casas decimais de acurácia no intervalo $[10^{-316}; 1-10^{-316}]$, e a segunda, chamada de PPND16, fornece cerca de 16 casas decimais de acurácia no mesmo intervalo. O método de precisão mais baixa, PPND7, utiliza polinômios racionais de graus 2 e 3, enquanto PPND16 utiliza polinômios racionais de grau 7.

O software estatístico R, por padrão, usa estas aproximações propostas por Wichura para gerar números aleatórios da Distribuição Normal baseados no Método de Inversão. No R, este algoritmo está implementado em linguagem C.

Por ser extenso e complexo, uma vez que deve-se ter bastante atenção no momento de digitar os polinômios racionais, não forneceremos aqui o pseudocódigo para este algoritmo. Sua implementação, em linguagem C, pode ser encontrada no código-fonte do software R, disponível no endereço eletrônico <http://www.r-project.org>

5.2 Geradores baseados no Método de Aceitação-Rejeição

Na literatura, podem ser encontrados três algoritmos distintos para gerar números aleatórios da Distribuição Normal tendo como base o Método de Aceitação-Rejeição. O primeiro utiliza a Distribuição de Cauchy; o segundo, a Distribuição Exponencial; e o terceiro não utiliza nenhuma distribuição de probabilidade específica, baseando-se em aceitações e rejeições da própria Distribuição Uniforme $[0; 1]$.

5.2.1 Gerador Normal a partir da Distribuição de Cauchy

Para implementar este algoritmo é necessário gerar ocorrências da Distribuição de Cauchy. A densidade da Distribuição de Cauchy é dada por:

$$f(x) = \frac{1}{\pi(1+x^2)} \quad x \in \mathfrak{R}$$

Assim, podemos gerar ocorrências desta distribuição pelo Método de Inversão, em que teremos $X = \tan(\pi(U - 1/2))$.

Assumindo que $\varphi(X)$ é a densidade da distribuição normal padrão no ponto X , uma proposição para este algoritmo é a seguinte:

$$\text{Fazer } M = \sqrt{2\pi/e}.$$

Repita

$$\text{Gerar } U, V \sim U[0;1].$$

$$\text{Fazer } X = \tan(\pi(U - 1/2)).$$

Calcular $\varphi(X), f(X)$.

Até que $U \leq \varphi(X)/Mf(X)$

Sair X .

5.2.2 Gerador Normal a partir da Distribuição Exponencial

Na seção 3.2.1, vimos como gerar números aleatórios da Distribuição Exponencial através do Método de Inversão. Outro algoritmo proposto para gerar variáveis aleatórias normais padrão baseado no Método de Rejeição é aquele que utiliza a Distribuição Exponencial. Assumindo que $\varphi(X)$ é a densidade da distribuição normal padrão no ponto X , temos:

1 Gerar $U \sim U[0;1]$ e $Y \sim U[0;U\sqrt{2e/\pi}]$.

2. Fazer $X = -\ln(U)$.

3. (a). Se $Y < \varphi(X)/2$ então sair $Z = -X$.

(b). Senão se $\varphi(X)/2 < Y < \varphi(X)$ então sair $Z = X$.

(c). Senão volte para o passo 1.

Estes dois últimos algoritmos apresentam três desvantagens na sua construção: é preciso saber gerar ou possuir um bom gerador de ocorrências do modelo Cauchy ou Exponencial; é necessário calcular o valor da densidade da normal padrão no ponto X , o que é expresso por $\varphi(X)$; e ambos os algoritmos são iterativos, pois possuem laços de repetição, o que pode comprometer o desempenho.

5.2.3 Gerador de Marsaglia e Bray (1964)

Este algoritmo foi proposto inicialmente por Marsaglia e Bray em 1964.

Gerar $U \sim U[0;1]$.

Se $0 \leq U \leq 0.8638$ então

Gerar $V, W \sim U[-1;1]$.

Sair $X = 2.3153508 * U - 1 + V + W$.

Se $0.8638 < U \leq 0.9745$ então

Gerar $V \sim U[0;1]$.

Sair $X = \frac{3}{2} * (V - 1 + 9.0334237 * (U - 0.8638))$.

Se $0.9973002 < U \leq 1$ então

Repita

Gerar $V, W \sim U[0;1]$.

Fazer $X = \frac{9}{2} - \ln(W)$.

Até que $XV^2 \leq \frac{9}{2}$

Sair $X = \sqrt{2X} \sin(U - 0.9986501)$.

Se $0.9745 < U \leq 0.9973002$ então

Repita

Gerar $X \sim U[-3;3], U \sim U[0;1]$.

Fazer $V = |X|$.

Fazer $W = 6.6313339 * (3 - V)^2$.

Fazer $Soma = 0$.

Se $V < \frac{3}{2}$ então

Fazer $Soma = 6.0432809 * \left(\frac{3}{2} - V\right)$.

Se $V < 1$ então

Fazer $Soma = Soma + 13.2626678 * (3 - V^2) - W$.

Até que $U \leq 49.0024445e^{-\frac{V^2}{2}} - Soma - W$

Sair X .

5.2.4 Gerador de Kinderman e Ramage (1976)

Este algoritmo foi proposto por Kinderman e Ramage em 1976, em um artigo intitulado “*Computer Generation of Normal Random Variables*”, e encontra-se implementado no software R em linguagem C utilizando uma versão corrigida de Josef Leydold. Trata-se de um algoritmo diferenciado dos dois últimos por utilizar aceitação-rejeição nos valores da própria Distribuição Uniforme [0; 1], e não em outra distribuição de probabilidade específica. Também não necessita do cálculo da densidade da normal padrão ($\phi(X)$).

Fazer $C_1 = 0.398942280401433$.

Fazer $C_2 = 0.180025191068563$.

Fazer $A = 2.216035867166471$.

Criar a função $g(x) = C_1 * e^{(-x*x/2)} - C_2 * (A - x)$.

Gerar $U_1 \sim U[0;1]$.

Se $U_1 < 0.884070402298758$ então

Gerar $U_2 \sim U[0;1]$.

Sair $X = A * (1.131131635444180 * U_1 + U_2 - 1)$.

Se $U_1 \geq 0.973310954173898$ então

Repita

Gerar $U_2 \sim U[0;1]$.

Gerar $U_3 \sim U[0;1]$.

Fazer $tt = (A * A - 2 * \ln(U_3))$.

Até que $(U_2 * U_2 < (A * A) / tt)$

Se $U_1 < 0.986655477086949$ então

Sair $X = \sqrt{tt}$.

Senão

Sair $X = -\sqrt{tt}$.

Se $U_1 \geq 0.958720824790463$ então

Repita

Gerar $U_2 \sim U[0;1]$.

Gerar $U_3 \sim U[0;1]$.

Fazer $tt = A - 0.630834801921960 * \min(U_2, U_3)$.

Até que $(\max(U_2, U_3) \leq 0.755591531667601)$ ou $(0.034240503750111 * |U_2 - U_3| \leq g(tt))$

Se $U_2 < U_3$ então

Sair $X = tt$.

Senão

Sair $X = -tt$.

Se $U_1 \geq 0.911312780288703$ então

Repita

Gerar $U_2 \sim U[0;1]$.

Gerar $U_3 \sim U[0;1]$.

Fazer $tt = 0.479727404222441 + 1.105473661022070 * \min(U_2, U_3)$.

Até que $(\max(U_2, U_3) \leq 0.872834976671790)$ ou $(0.049264496373128 * |U_2 - U_3| \leq g(tt))$

Se $U_2 < U_3$ então

Sair $X = tt$.

Senão

Sair $X = -tt$.

Repita

(a) Gerar $U_2 \sim U[0;1]$.

(b) Gerar $U_3 \sim U[0;1]$.

(c) Fazer $tt = 0.479727404222441 - 0.595507138015940 * \min(U_2, U_3)$.

Se $tt < 0$ então volte para (a).

Até que $(\max(U_2, U_3) \leq 0.805577924423817)$ ou $(0.053377549506886 * |U_2 - U_3| \leq g(tt))$

Se $U_2 < U_3$ então

Sair $X = tt$.

Senão

Sair $X = -tt$.

Estes dois últimos algoritmos são os mais extensos para geração de ocorrências da Distribuição Normal. No momento de suas implementações, deve-se conferir atentamente a digitação dos valores numéricos envolvidos para evitar erros nos resultados. Observa-se, também, que estes algoritmos fazem uso intenso de estruturas de decisão e de repetição, como as do tipo “se-então” e “repita”, respectivamente. O uso da estrutura “repita”, que cria um bloco de instruções iterativo, pode comprometer a velocidade de processamento.

5.3 Geradores baseados no Método do Quociente de Uniformes

Os algoritmos seguintes foram construídos tendo como base o Método do Quociente de Uniformes. Esse método produz resultados exatos.

5.3.1 Quociente de Uniformes de Kinderman e Monahan (1977)

Repita

Gerar $U \sim U[0;1]$.

Gerar $V \sim U[-1;1]$.

Fazer $X = (V\sqrt{2|e|})/U$.

Se $X^2 \leq 5 - 4e^{1/4}U$ então

Sair X .

Senão se $X^2 < 4e^{-1.35}/U + 1.4$ então

Se $V^2 < -4U^2 \ln(U)$ então

Sair X .

Fim do Se

Fim do Se

Fim do Repita

5.3.2 Quociente de Uniformes de Leva (1992)

Repita

Gerar $U \sim U[0;1]$.

Gerar $V \sim U[-1;1]$.

Fazer $v = V\sqrt{2/e}$.

Fazer $X = U - 0.449871$, $Y = |v| + 0.386595$.

Fazer $Q = X^2 + Y(0.19600Y - 0.25472X)$.

Se $Q < 0.27597$ então

Sair v/U .

Senão se $Q < 0.27846$ então

Se $v^2 < -4U^2 \ln(U)$ então

Sair v/U .

Fim do Se

Fim do Se

Fim do Repita

5.3.3 Quociente de Uniformes exposto por Insúa, Jiménez e Martín (2009)

O algoritmo abaixo foi exposto por Insúa, Jiménez e Martín (2009) na obra “*Simulación – Métodos y Aplicaciones*”.

Enquanto $Z > -\ln(U)$

Gerar $U, U_1 \sim U[0;1]$.

Fazer $V = (\sqrt{2/e}) * (2U_1 - 1)$, $X = V/U$, $Z = X^2/4$.

Sair X .

Como se pode observar, todos os algoritmos baseados no Método do Quociente de Uniformes tem suas instruções contidas dentro de uma estrutura de repetição do tipo “repita” ou “enquanto”. Além disso, todos também exigem o cálculo de logaritmos e utilizam operações de radiciação (raiz quadrada) e exponenciação, os quais estão sujeitos a erros de exatidão e precisão.

5.4 Geradores baseados em outros métodos

Existem algoritmos para geração de números aleatórios da Distribuição Normal que foram construídos não tendo como base nenhum dos métodos elencados acima. Esses algoritmos estão baseados em outros teoremas ou métodos, e serão apresentados a seguir.

5.4.1 Soma de Doze Uniformes ou Método da Convolução

Este procedimento se baseia no Teorema Central do Limite (TCL) que, em linhas gerais, afirma que a soma de uma grande quantidade de variáveis aleatórias (com algumas restrições) tende a uma Distribuição Normal. Este algoritmo pode ser visto como um exemplo dessa transformação. Se as variáveis U_i , $1, \dots, n$, são independentes e identicamente distribuídas (iid) como uma Uniforme $U[0;1]$, em que $E(U_i) = 1/2$ e $Var(U_i) = 1/12$, pelo

TCL a variável aleatória $X = \frac{\sum_{i=1}^n U_i - \frac{n}{2}}{\sqrt{n/12}}$ se distribui aproximadamente como uma $N(0,1)$,

para n suficientemente grande. Para $n = 12$ já se tem uma boa aproximação, com o qual X toma a forma $\left(\sum_{i=1}^{12} U_i\right) - 6$, e o algoritmo ficaria:

1. Gerar $U_1, \dots, U_{12} \sim U[0;1]$.
2. Fazer $X = \left(\sum_{i=1}^{12} U_i\right) - 6$.
3. Sair X .

Embora seja de pouca complexidade, este algoritmo apresenta três inconvenientes:

- Trata-se de uma aproximação para a Distribuição Normal.
- O somatório em questão pode estar sujeito a erros de arredondamento dependendo do ambiente computacional em que o algoritmo seja implementado.
- Pode haver uma “explosão de Uniformes”. Por exemplo: se para gerar um valor da Normal preciso gerar 12 Uniformes, então, para gerar 10.000 valores da Normal precisarei gerar 120.000 Uniformes e assim sucessivamente.

5.4.2 Método de Box-Muller

O método exato para gerar valores da Distribuição Normal mais antigo e mais conhecido é o Método de Box-Muller (Box e Muller, 1958), o qual produz um par de números aleatórios normais padrão e independentes de um par de números uniformes. Este método utiliza o fato de que a distribuição bidimensional de dois números aleatórios normais com média zero é radialmente simétrica se ambos os componentes normais tem a mesma variância. O algoritmo de Box-Muller pode ser entendido como um método no qual os números normais

de sua saída representam as coordenadas no plano bidimensional. Devido ao algoritmo produzir dois números aleatórios cada vez que é executado, é comum a sua função geradora retornar o primeiro valor para o usuário e ocultar o outro valor para retornar na próxima chamada da função. Este algoritmo encontra-se disponível no software R.

1. Gerar $U_1, U_2 \sim U[0;1]$.
2. Fazer $R = \sqrt{-2\ln(U_1)}$, $\Theta = 2\pi U_2$.
3. Fazer $X_1 = R\cos(\Theta)$, $X_2 = R\sin(\Theta)$.
4. Sair X_1, X_2 .

As equações para obter X_1 e X_2 são conhecidas como transformações de Box-Muller.

A desvantagem deste método é a exigência do cálculo de senos e cossenos, o que envolve o cômputo da expansão de séries de Taylor. Para uma boa aproximação aos valores exatos de senos e cossenos, dependemos da maior expansão possível das séries de Taylor, o que pode variar de acordo com a linguagem de programação utilizada.

5.4.3 Variante de Marsaglia

Marsaglia introduziu sua “*Variante Polar do Método de Box-Muller*”, a qual incorpora o método de rejeição para evitar as operações trigonométricas de senos e cossenos. O algoritmo é:

Enquanto $W > 1$

Gerar $U_1, U_2 \sim U[0;1]$.

Fazer $V_1 = 2U_1 - 1$, $V_2 = 2U_2 - 1$, $W = V_1^2 + V_2^2$.

Fazer $C = \sqrt{[-2\ln(W)]/W}$.

Sair $X_1 = CV_1$, $X_2 = CV_2$.

Uma vez que este algoritmo utiliza rejeição para evitar o cálculo das funções trigonométricas, temos uma complexidade adicional na sua programação.

5.4.4 Método Ahrens-Dieter

O Método Ahrens-Dieter (Ahrens e Dieter, 1988) é um gerador normal exato que transforma um par de números aleatórios independentes da Distribuição Exponencial e da Cauchy em dois números aleatórios da Normal independentes. Este método é similar ao de Box-Muller, exceto que, ao invés de aplicar uma transformação complexa para facilmente gerar números aleatórios uniformes, aplica uma transformação mais simples em duas distribuições que são mais complexas de gerar. A princípio, as distribuições Exponencial e Cauchy poderiam

ser geradas diretamente, usando $-\ln(U)$ e $\tan(\pi(U - 1/2))$, respectivamente, o que faria deste um método de transformação. Entretanto, a única razão deste método ser factível é devido aos autores terem desenvolvido dois algoritmos baseados em rejeição para gerar amostras das distribuições Exponencial e Cauchy.

Um dos inconvenientes deste método é ser complexo de entender. Deve ser cuidadosamente implementado, pois requer muitas constantes. Este algoritmo encontra-se disponível no software R. Sua implementação, em linguagem C, pode ser encontrada no código-fonte do R, disponível no endereço eletrônico <http://www.r-project.org>

6. ESTIMATIVAS DAS AMOSTRAS DA DISTRIBUIÇÃO NORMAL

Nos capítulos subsequentes, apresentaremos as estimativas com amostras de tamanho $n = 100$, $n = 1000$ e $n = 10000$ da Distribuição Normal padrão para as seguintes medidas: esperança, máximo e mínimo, amplitude, desvio-padrão, assimetria, curtose, quartis, distância interquartilica e valor-p para o teste de aderência. Os valores “exatos” das medidas para a Distribuição Normal padrão são: esperança = 0; máximo = 4.264891 e mínimo = -4.264891; amplitude = 8.529782; desvio-padrão = 1; assimetria = 0; curtose = 0; 1º quartil = -0.6744898, 2º quartil = 0 e 3º quartil = 0.6744898; distância interquartilica = 1.34898. Considera-se como máximo o valor x tal que a probabilidade $P(X \leq x) = 0.99999$, e como mínimo o valor x tal que a probabilidade $P(X \leq x) = 0.00001$.

6.1 Estimativas com amostras de tamanho $n = 100$

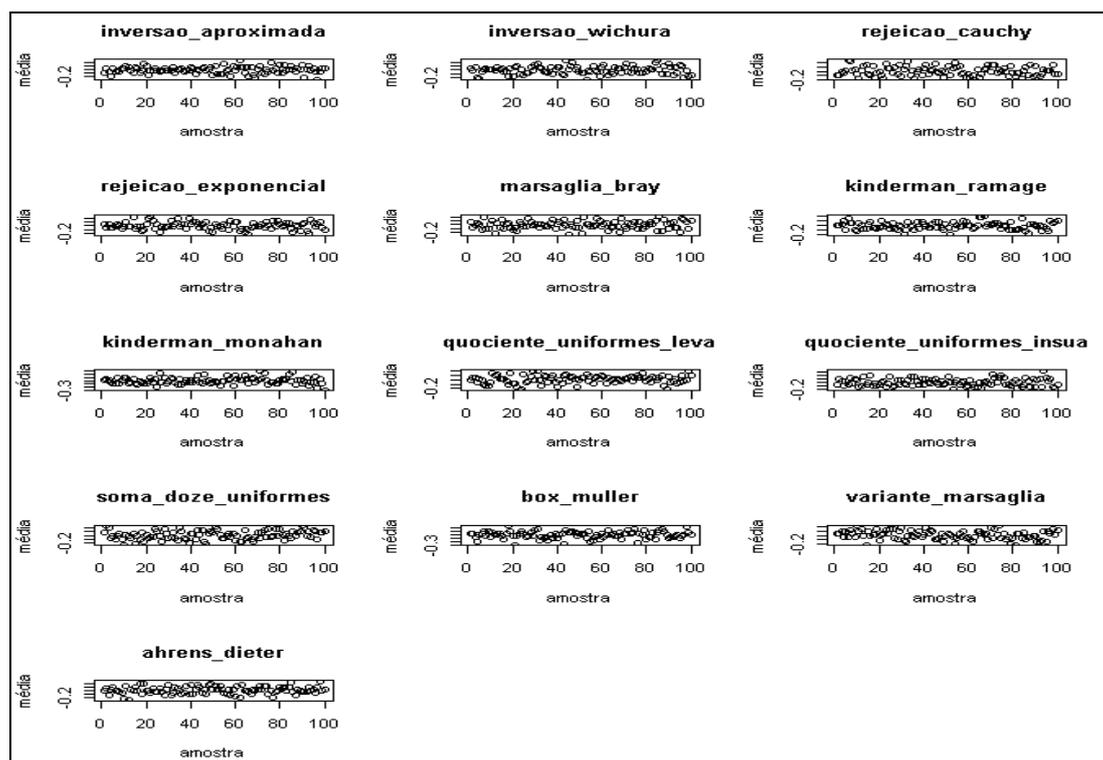
Nas tabelas abaixo, apresentamos a média e o erro padrão dos valores estimados pelos geradores para cada uma das medidas analisadas. Nos gráficos, ilustramos o comportamento das estimativas dos geradores ao longo das 100 amostras.

6.1.1 Esperança

Tabela 2 – Média e erro padrão dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Gerador	Média	Erro Padrão
Inv. Aproximada	-0.009062434	0.09543417
Inv. Wichura	-0.007578265	0.106004
Rej. Cauchy	0.007823898	0.1058975
Rej. Exponencial	0.001805649	0.09287565
Marsaglia-Bray	-0.01017365	0.1069247
Kinderman-Ramage	-0.01477516	0.08043886
Kinderman-Monahan	-0.002752763	0.09932383
Quoc. Unif. Leva	0.009079546	0.09477138
Quoc. Unif. Insua	-0.004753159	0.1032364
Soma Doze Uniformes	0.008746627	0.1104117
Box-Muller	-0.01167383	0.1024274
Variante Marsaglia	0.008910398	0.1019724
Ahrens-Dieter	0.01058423	0.09843404

Fonte: elaborada pelo autor.



Quadro 1 – Gráficos dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.

6.1.2 Máximo e Mínimo

Tabela 3 – Média e erro padrão dos valores estimados para o máximo da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Gerador	Média	Erro Padrão
Inv. Aproximada	2.532696	0.3717791
Inv. Wichura	2.517707	0.3713459
Rej. Cauchy	2.514574	0.4633389
Rej. Exponencial	2.542086	0.4041451
Marsaglia-Bray	2.42405	0.403121
Kinderman-Ramage	2.51027	0.4407196
Kinderman-Monahan	2.483085	0.3776668
Quoc. Unif. Leva	2.531315	0.4510915
Quoc. Unif. Insua	2.507519	0.4138504
Soma Doze Uniformes	2.481613	0.4073649
Box-Muller	2.490676	0.4677043
Variante Marsaglia	2.533667	0.4419769
Ahrens-Dieter	2.585604	0.4743861

Fonte: elaborada pelo autor.

Tabela 4 – Média e erro padrão dos valores estimados para o mínimo da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Gerador	Média	Erro Padrão
Inv. Aproximada	-2.473534	0.3582343
Inv. Wichura	-2.524946	0.4164138
Rej. Cauchy	-2.408515	0.3797996
Rej. Exponencial	-2.51063	0.4381672
Marsaglia-Bray	-2.49863	0.4400174
Kinderman-Ramage	-2.47332	0.3976071
Kinderman-Monahan	-2.571436	0.4578993
Quoc. Unif. Leva	-2.472008	0.4456734
Quoc. Unif. Insua	-2.539813	0.4254929
Soma Doze Uniformes	-2.411904	0.3793033
Box-Muller	-2.47362	0.4212333
Variante Marsaglia	-2.55832	0.4564264
Ahrens-Dieter	-2.46875	0.4298149

Fonte: elaborada pelo autor.

6.1.3 Amplitude

Tabela 5 – Média e erro padrão dos valores estimados para a amplitude da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Gerador	Média	Erro Padrão
Inv. Aproximada	5.00623	0.4993893
Inv. Wichura	5.042653	0.5974695
Rej. Cauchy	4.923089	0.5509244
Rej. Exponencial	5.052715	0.593337
Marsaglia-Bray	4.92268	0.6096093
Kinderman-Ramage	4.983591	0.6203069
Kinderman-Monahan	5.05452	0.594761
Quoc. Unif. Leva	5.003323	0.6467366
Quoc. Unif. Insua	5.047331	0.5992069
Soma Doze Uniformes	4.893517	0.574163
Box-Muller	4.964297	0.6359683
Variante Marsaglia	5.091987	0.6343671
Ahrens-Dieter	5.054355	0.6522928

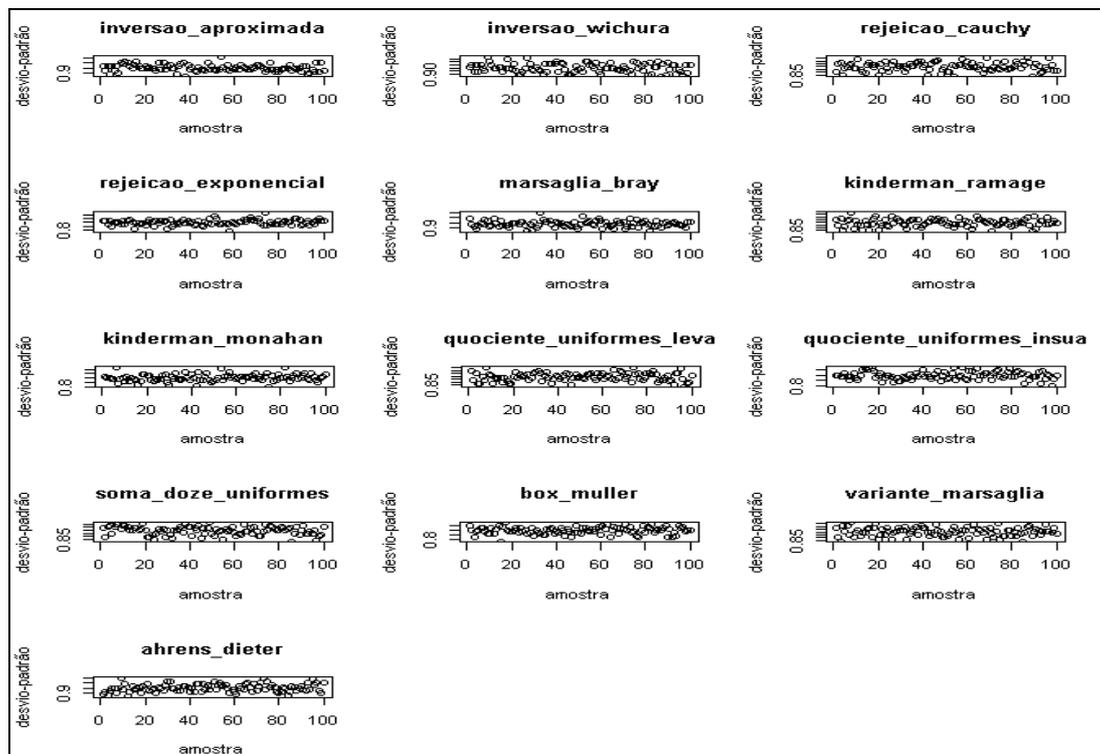
Fonte: elaborada pelo autor.

6.1.4 Desvio-Padrão

Tabela 6 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Gerador	Média	Erro Padrão
Inv. Aproximada	1.000396	0.06350335
Inv. Wichura	0.999242	0.07288895
Rej. Cauchy	0.9963197	0.0768562
Rej. Exponencial	1.000953	0.06933273
Marsaglia-Bray	0.9759807	0.06204278
Kinderman-Ramage	0.9933156	0.06913938
Kinderman-Monahan	1.007644	0.08090255
Quoc. Unif. Leva	0.9925069	0.06972294
Quoc. Unif. Insua	0.9971297	0.07582419
Soma Doze Uniformes	0.990602	0.06829935
Box-Muller	0.99713	0.07369371
Variante Marsaglia	1.002215	0.07737016
Ahrens-Dieter	1.001942	0.07607487

Fonte: elaborada pelo autor.



Quadro 2 – Gráficos dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.

6.1.5 Assimetria

Tabela 7 – Média e erro padrão dos valores estimados para a assimetria da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Gerador	Média	Erro Padrão
Inv. Aproximada	0.02902254	0.2571842
Inv. Wichura	-0.0223686	0.2439079
Rej. Cauchy	0.02635954	0.217975
Rej. Exponencial	0.003477309	0.2376233
Marsaglia-Bray	-0.01374043	0.2320286
Kinderman-Ramage	-0.0003383674	0.2398639
Kinderman-Monahan	-0.04180047	0.2361352
Quoc. Unif. Leva	0.02170604	0.2331396
Quoc. Unif. Insua	-0.02183757	0.2363794
Soma Doze Uniformes	-0.01089978	0.2309453
Box-Muller	0.03744194	0.234206
Variante Marsaglia	-0.01625267	0.2777991
Ahrens-Dieter	0.000102327	0.2364053

Fonte: elaborada pelo autor.

6.1.6 Curtose

Tabela 8 – Média e erro padrão dos valores estimados para a curtose da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Gerador	Média	Erro Padrão
Inv. Aproximada	0.0121851	0.4554933
Inv. Wichura	0.008668426	0.4174302
Rej. Cauchy	-0.09153277	0.3994047
Rej. Exponencial	0.03596081	0.4327935
Marsaglia-Bray	0.039234	0.5269268
Kinderman-Ramage	0.02258049	0.4750382
Kinderman-Monahan	-0.006353618	0.4320223
Quoc. Unif. Leva	-0.002457271	0.4602357
Quoc. Unif. Insua	0.006453771	0.3898836
Soma Doze Uniformes	-0.1098272	0.4534674
Box-Muller	-0.03922979	0.4789778
Variante Marsaglia	0.05180219	0.4938133
Ahrens-Dieter	-0.009752955	0.478546

Fonte: elaborada pelo autor.

6.1.7 Quartis

Tabela 9 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 100

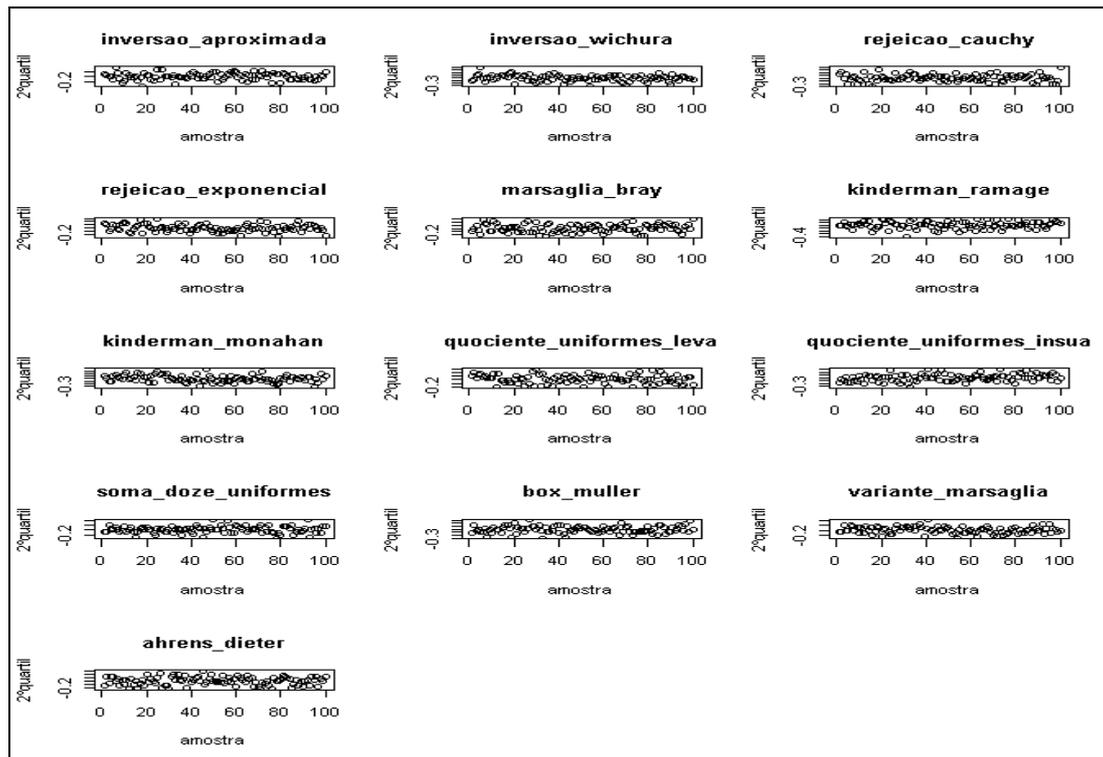
Gerador	Média	Erro Padrão
Inv. Aproximada	-0.6628656	0.1390403
Inv. Wichura	-0.6675943	0.1319938
Rej. Cauchy	-0.6790974	0.1568243
Rej. Exponencial	-0.6603809	0.1225857
Marsaglia-Bray	-0.6565412	0.1445299
Kinderman-Ramage	-0.6704237	0.1242368
Kinderman-Monahan	-0.6641848	0.1325942
Quoc. Unif. Leva	-0.65425	0.1362246
Quoc. Unif. Insua	-0.6643328	0.1391937
Soma Doze Uniformes	-0.6753588	0.1411471
Box-Muller	-0.6827325	0.1323297
Variante Marsaglia	-0.6614249	0.1394763
Ahrens-Dieter	-0.6571	0.1210521

Fonte: elaborada pelo autor.

Tabela 10 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Gerador	Média	Erro Padrão
Inv. Aproximada	-0.02281699	0.1241362
Inv. Wichura	-0.004929354	0.1317373
Rej. Cauchy	0.01026994	0.1380688
Rej. Exponencial	0.01563119	0.1225143
Marsaglia-Bray	-0.01186945	0.1224363
Kinderman-Ramage	-0.01184882	0.09706678
Kinderman-Monahan	-0.003526452	0.1194232
Quoc. Unif. Leva	0.002608566	0.1149112
Quoc. Unif. Insua	-0.003913473	0.1113667
Soma Doze Uniformes	0.02130182	0.1320383
Box-Muller	-0.02020447	0.1339905
Variante Marsaglia	0.008734944	0.1279172
Ahrens-Dieter	0.01791564	0.128293

Fonte: elaborada pelo autor.



Quadro 3 – Gráficos dos valores estimados para o 2º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.

Tabela 11 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Gerador	Média	Erro Padrão
Inv. Aproximada	0.6483198	0.1258568
Inv. Wichura	0.6714481	0.1351443
Rej. Cauchy	0.6752402	0.1349913
Rej. Exponencial	0.6489866	0.1329137
Marsaglia-Bray	0.6365014	0.1346928
Kinderman-Ramage	0.6393106	0.1269985
Kinderman-Monahan	0.6861109	0.1418966
Quoc. Unif. Leva	0.6742543	0.1240231
Quoc. Unif. Insua	0.6617501	0.1384011
Soma Doze Uniformes	0.686223	0.1544554
Box-Muller	0.6564531	0.1397859
Variante Marsaglia	0.6741447	0.1315035
Ahrens-Dieter	0.6827423	0.1360271

Fonte: elaborada pelo autor.

6.1.8 Distância Interquartilica

Tabela 12 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição $N(0,1)$ em 100 amostras de tamanho 100

Gerador	Média	Erro Padrão
Inv. Aproximada	1.311185	0.145169
Inv. Wichura	1.339042	0.1319143
Rej. Cauchy	1.354338	0.1842912
Rej. Exponencial	1.309368	0.1451868
Marsaglia-Bray	1.293043	0.1525419
Kinderman-Ramage	1.309734	0.1593492
Kinderman-Monahan	1.350296	0.1562913
Quoc. Unif. Leva	1.328504	0.1552029
Quoc. Unif. Insua	1.326083	0.1592239
Soma Doze Uniformes	1.361582	0.1515428
Box-Muller	1.339186	0.1740855
Variante Marsaglia	1.33557	0.1668602
Ahrens-Dieter	1.339842	0.1400257

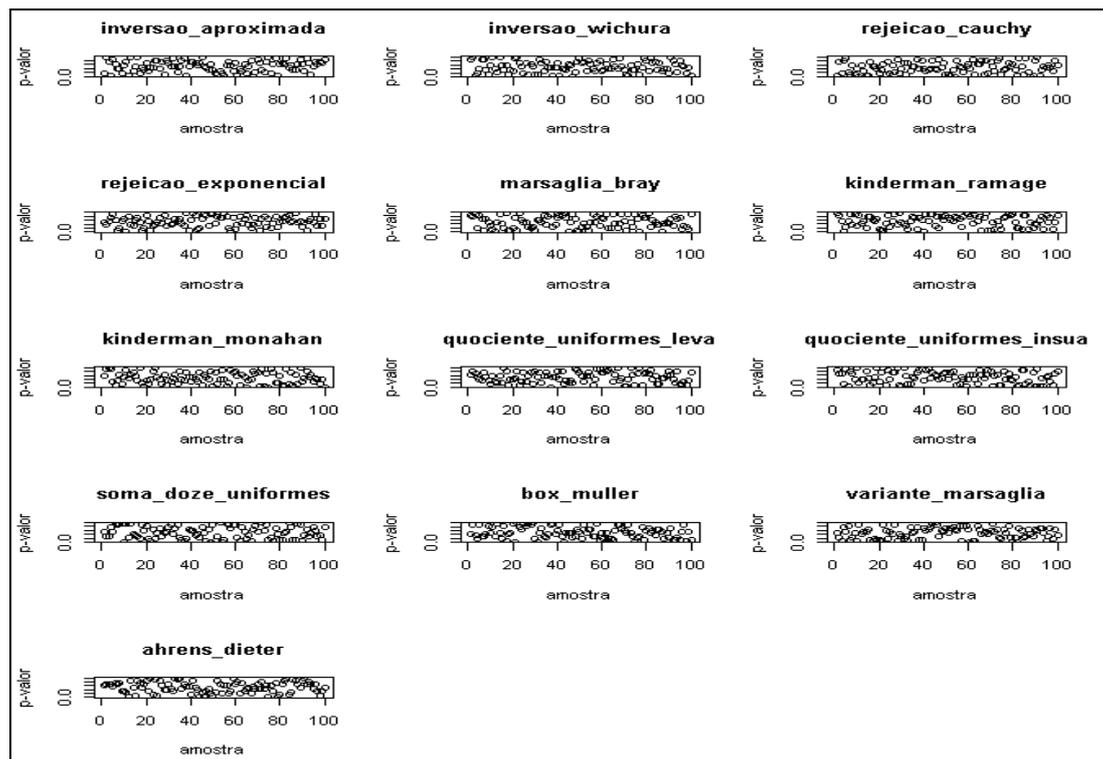
Fonte: elaborada pelo autor.

6.1.9 Teste de Aderência

Tabela 13 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 100

Gerador	Média	Erro Padrão
Inv. Aproximada	0.5026506	0.2685548
Inv. Wichura	0.5131365	0.2762698
Rej. Cauchy	0.4513709	0.2858284
Rej. Exponencial	0.5371342	0.2636209
Marsaglia-Bray	0.4932079	0.2950988
Kinderman-Ramage	0.5462355	0.2922936
Kinderman-Monahan	0.464263	0.2881989
Quoc. Unif. Leva	0.5170986	0.2759117
Quoc. Unif. Insua	0.5065225	0.290438
Soma Doze Uniformes	0.497217	0.306954
Box-Muller	0.4951024	0.2869744
Variante Marsaglia	0.4655814	0.2798491
Ahrens-Dieter	0.5253088	0.2753594

Fonte: elaborada pelo autor.



Quadro 4 – Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.

6.1.10 Conclusões

De maneira geral, para as 100 amostras com tamanho 100, os algoritmos Kinderman-Monahan e Ahrens-Dieter foram os que mais se aproximaram, em média, aos valores exatos das medidas analisadas. Entretanto, o algoritmo de Kinderman-Monahan mostrou-se um pouco mais preciso que o algoritmo de Ahrens-Dieter. Em termos de variabilidade, ambos se comportaram de forma similar.

Assim, podemos dizer que o algoritmo de Kinderman-Monahan comportou-se melhor para a geração de amostras com tamanho $n = 100$ da Distribuição Normal padrão.

Tabela 14 – Estimativas médias do algoritmo de Kinderman-Monahan em 100 amostras de tamanho 100 da distribuição $N(0,1)$

Medidas	Kinderman-Monahan
Esperança	-0.002752763
Máximo	2.483085
Mínimo	-2.571436
Amplitude	5.05452
Desvio-Padrão	1.007644
Assimetria	-0.04180047
Curtose	-0.006353618

1º Quartil	-0.6641848
2º Quartil	-0.003526452
3º Quartil	0.6861109
Dist. Interquart.	1.350296

Fonte: elaborada pelo autor.

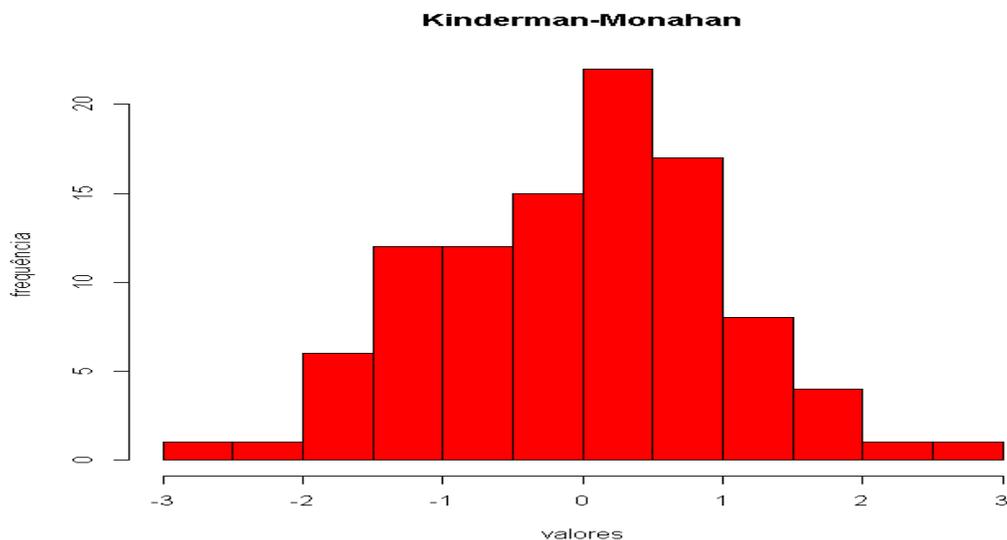


Figura 10 – Histograma de 100 valores gerados da distribuição $N(0,1)$ pelo algoritmo de Kinderman-Monahan

Fonte: elaborada pelo autor.

6.2. Estimativas com amostras de tamanho $n = 1000$

Nas tabelas abaixo, apresentamos a média e o erro padrão dos valores estimados pelos geradores para cada uma das medidas analisadas. Nos gráficos, ilustramos o comportamento das estimativas dos geradores ao longo das 100 amostras.

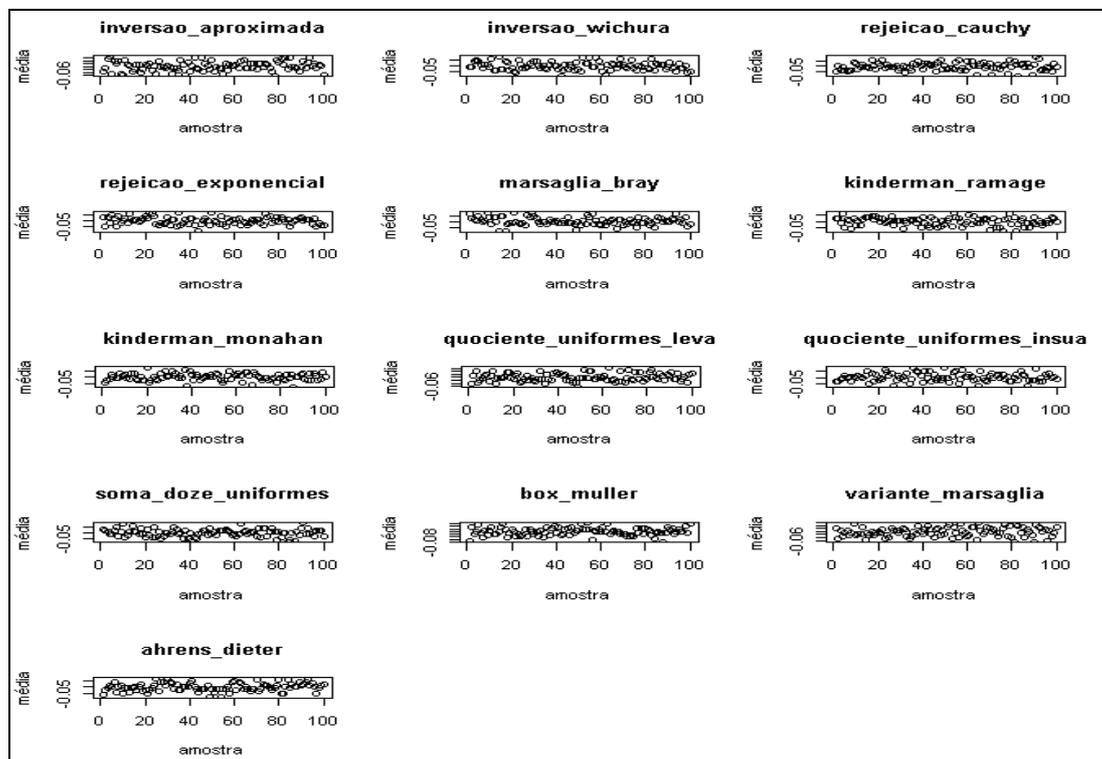
6.2.1 Esperança

Tabela 15 – Média e erro padrão dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Gerador	Média	Erro Padrão
Inv. Aproximada	0.002014318	0.03129956
Inv. Wichura	-0.002079295	0.03300581
Rej. Cauchy	-0.002815066	0.03429791
Rej. Exponencial	0.0001384719	0.02992357
Marsaglia-Bray	0.003190983	0.03452673
Kinderman-Ramage	-0.001291773	0.03202166
Kinderman-Monahan	0.001041138	0.02778461
Quoc. Unif. Leva	0.0004834135	0.02646444

Quoc. Unif. Insua	0.004022597	0.0330536
Soma Doze Uniformes	-0.0006903926	0.0318287
Box-Muller	-0.0008262409	0.02924798
Variante Marsaglia	-0.002155977	0.03027569
Ahrens-Dieter	-0.002501833	0.03127202

Fonte: elaborada pelo autor.



Quadro 5 – Gráficos dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.

6.2.2 Máximo e Mínimo

Tabela 16 – Média e erro padrão dos valores estimados para o máximo da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Gerador	Média	Erro Padrão
Inv. Aproximada	3.161852	0.3090106
Inv. Wichura	3.196863	0.3647633
Rej. Cauchy	3.283016	0.3426755
Rej. Exponencial	3.223687	0.3602605
Marsaglia-Bray	3.281389	0.4154233
Kinderman-Ramage	3.261837	0.3621373
Kinderman-Monahan	3.242655	0.3544424
Quoc. Unif. Leva	3.243226	0.3253418
Quoc. Unif. Insua	3.1706	0.3394337

Soma Doze Uniformes	3.101809	0.3155943
Box-Muller	3.215017	0.3867896
Variante Marsaglia	3.20478	0.2929196
Ahrens-Dieter	3.210632	0.2775037

Fonte: elaborada pelo autor.

Tabela 17 – Média e erro padrão dos valores estimados para o mínimo da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Gerador	Média	Erro Padrão
Inv. Aproximada	-3.169214	0.3226643
Inv. Wichura	-3.298735	0.3736771
Rej. Cauchy	-3.289649	0.4256568
Rej. Exponencial	-3.182163	0.3331722
Marsaglia-Bray	-3.227082	0.3436708
Kinderman-Ramage	-3.184129	0.3174223
Kinderman-Monahan	-3.236524	0.3363518
Quoc. Unif. Leva	-3.249184	0.3745607
Quoc. Unif. Insua	-3.120346	0.2724511
Soma Doze Uniformes	-3.147004	0.2957891
Box-Muller	-3.216268	0.3210548
Variante Marsaglia	-3.297524	0.387599
Ahrens-Dieter	-3.252783	0.3418995

Fonte: elaborada pelo autor.

6.2.3 Amplitude

Tabela 18 – Média e erro padrão dos valores estimados para a amplitude da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Gerador	Média	Erro Padrão
Inv. Aproximada	6.331065	0.4586844
Inv. Wichura	6.495598	0.5209199
Rej. Cauchy	6.572665	0.5685247
Rej. Exponencial	6.405849	0.5428715
Marsaglia-Bray	6.508471	0.5407429
Kinderman-Ramage	6.445966	0.470401
Kinderman-Monahan	6.479179	0.5048546
Quoc. Unif. Leva	6.49241	0.4873395
Quoc. Unif. Insua	6.290945	0.4391695
Soma Doze Uniformes	6.248813	0.4413202
Box-Muller	6.431285	0.4837155
Variante Marsaglia	6.502304	0.4641195
Ahrens-Dieter	6.463415	0.4245059

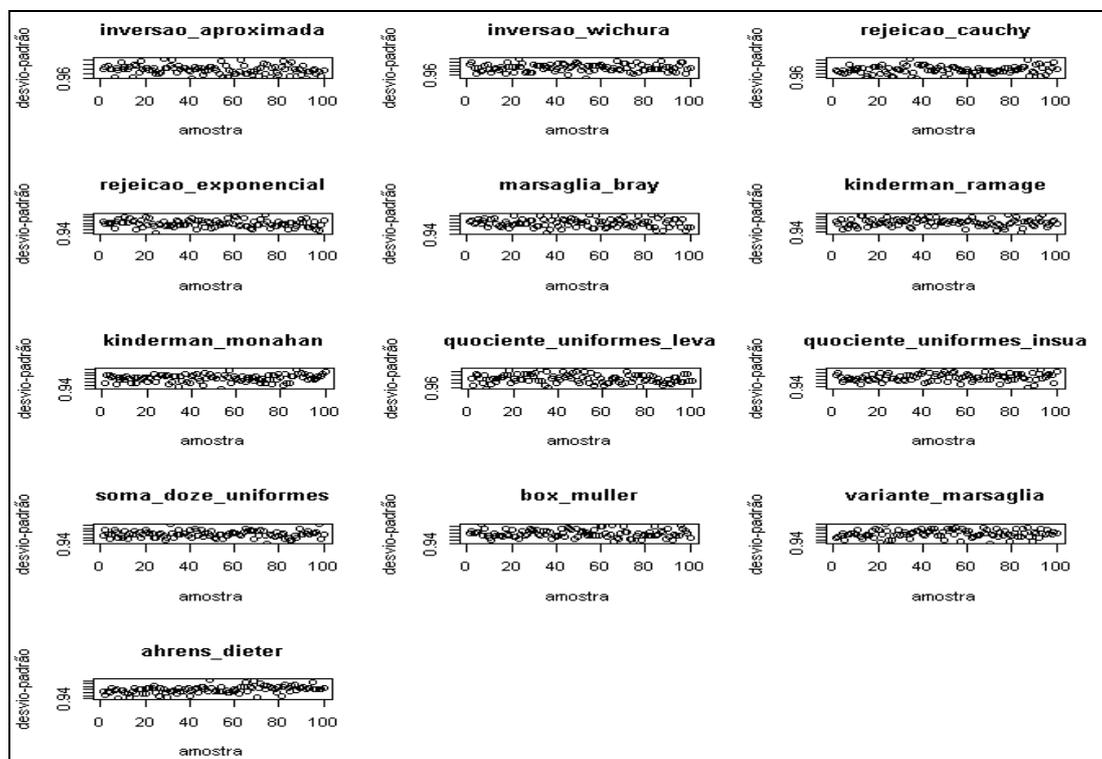
Fonte: elaborada pelo autor.

6.2.4 Desvio-Padrão

Tabela 19 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Gerador	Média	Erro Padrão
Inv. Aproximada	0.9985553	0.01926768
Inv. Wichura	1.004279	0.02111584
Rej. Cauchy	1.003426	0.02143183
Rej. Exponencial	1.000373	0.02312311
Marsaglia-Bray	0.9983488	0.02232527
Kinderman-Ramage	1.000134	0.02512456
Kinderman-Monahan	1.003499	0.02258203
Quoc. Unif. Leva	1.002539	0.02043831
Quoc. Unif. Insua	0.9941485	0.02107412
Soma Doze Uniformes	1.00111	0.02214742
Box-Muller	0.996857	0.02185337
Variante Marsaglia	0.9986953	0.02377147
Ahrens-Dieter	1.000206	0.02285446

Fonte: elaborada pelo autor.



Quadro 6 – Gráficos dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.

6.2.5 Assimetria

Tabela 20 – Média e erro padrão dos valores estimados para a assimetria da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Gerador	Média	Erro Padrão
Inv. Aproximada	-0.004506061	0.07164439
Inv. Wichura	-0.01557684	0.07423158
Rej. Cauchy	0.003774131	0.07394407
Rej. Exponencial	0.007041263	0.07951698
Marsaglia-Bray	-0.01176757	0.086832
Kinderman-Ramage	-0.009974317	0.07397464
Kinderman-Monahan	0.003023725	0.07557843
Quoc. Unif. Leva	0.0004293911	0.07242903
Quoc. Unif. Insua	-0.0006660868	0.07239461
Soma Doze Uniformes	-0.0063495	0.06945569
Box-Muller	0.001022244	0.08411176
Variante Marsaglia	-0.002885438	0.08461139
Ahrens-Dieter	-0.004838704	0.07377537

Fonte: elaborada pelo autor.

6.2.6 Curtose

Tabela 21 – Média e erro padrão dos valores estimados para a curtose da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Gerador	Média	Erro Padrão
Inv. Aproximada	-0.0255475	0.1379419
Inv. Wichura	-0.01269046	0.154017
Rej. Cauchy	-0.0009228562	0.1744196
Rej. Exponencial	-0.01568317	0.1559886
Marsaglia-Bray	0.02226501	0.1737806
Kinderman-Ramage	-0.006898075	0.1492507
Kinderman-Monahan	-0.0365112	0.1448672
Quoc. Unif. Leva	-0.01048786	0.167665
Quoc. Unif. Insua	-0.04045058	0.1433633
Soma Doze Uniformes	-0.1037695	0.1361594
Box-Muller	-0.0044863	0.1405253
Variante Marsaglia	0.003023854	0.1473875
Ahrens-Dieter	-0.01796486	0.1532614

Fonte: elaborada pelo autor.

6.2.7 Quartis

Tabela 22 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição N(0,1) em 100 amostras de tamanho 1000

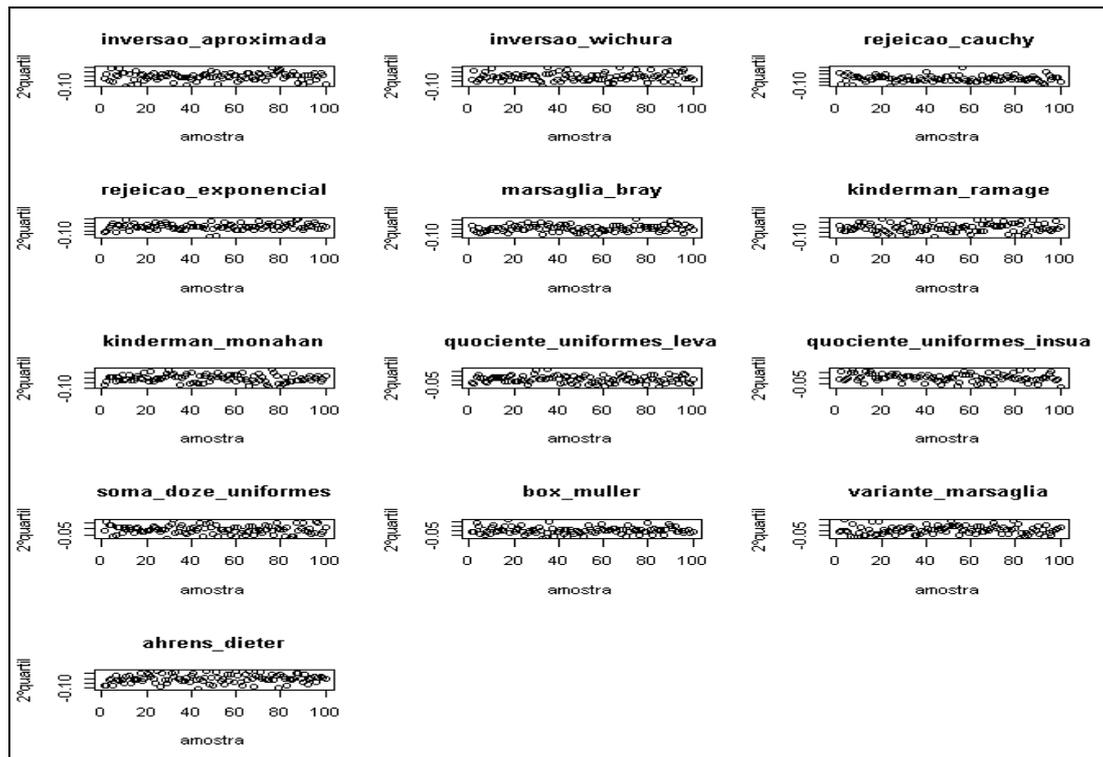
Gerador	Média	Erro Padrão
Inv. Aproximada	-0.6712368	0.04225581
Inv. Wichura	-0.6781676	0.04113612
Rej. Cauchy	-0.6792409	0.03787453
Rej. Exponencial	-0.6741933	0.04050138
Marsaglia-Bray	-0.665509	0.04226359
Kinderman-Ramage	-0.6715896	0.04703856
Kinderman-Monahan	-0.682993	0.04029391
Quoc. Unif. Leva	-0.6761248	0.04052167
Quoc. Unif. Insua	-0.6610673	0.03960869
Soma Doze Uniformes	-0.6827118	0.04688311
Box-Muller	-0.6733365	0.04329852
Variante Marsaglia	-0.6735085	0.04686752
Ahrens-Dieter	-0.6762204	0.04461841

Fonte: elaborada pelo autor.

Tabela 23 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição N(0,1) em 100 amostras de tamanho 1000

Gerador	Média	Erro Padrão
Inv. Aproximada	0.001870556	0.03771755
Inv. Wichura	0.000113983	0.03958677
Rej. Cauchy	-0.005639999	0.04216683
Rej. Exponencial	-0.0006449595	0.03999646
Marsaglia-Bray	0.004664403	0.04397486
Kinderman-Ramage	0.001184501	0.04123253
Kinderman-Monahan	0.002688699	0.03827916
Quoc. Unif. Leva	0.0008143753	0.03307258
Quoc. Unif. Insua	0.002831336	0.03898503
Soma Doze Uniformes	-0.001977901	0.03810385
Box-Muller	-0.001869335	0.0380701
Variante Marsaglia	-0.003201695	0.03986383
Ahrens-Dieter	-0.0015663	0.03989253

Fonte: elaborada pelo autor.



Quadro 7 – Gráficos dos valores estimados para o 2º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.

Tabela 24 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Gerador	Média	Erro Padrão
Inv. Aproximada	0.6736855	0.0420988
Inv. Wichura	0.6782685	0.04116969
Rej. Cauchy	0.6718305	0.04810266
Rej. Exponencial	0.672289	0.04272079
Marsaglia-Bray	0.674154	0.04426279
Kinderman-Ramage	0.6788996	0.04190076
Kinderman-Monahan	0.6797421	0.03955191
Quoc. Unif. Leva	0.6754142	0.04069196
Quoc. Unif. Insua	0.6780062	0.04446724
Soma Doze Uniformes	0.6817653	0.04430915
Box-Muller	0.6711728	0.04166654
Variante Marsaglia	0.6694676	0.03874335
Ahrens-Dieter	0.6734105	0.03971431

Fonte: elaborada pelo autor.

6.2.8 Distância Interquartilica

Tabela 25 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Gerador	Média	Erro Padrão
Inv. Aproximada	1.344922	0.04593239
Inv. Wichura	1.356436	0.04057731
Rej. Cauchy	1.351071	0.04705975
Rej. Exponencial	1.346482	0.05117137
Marsaglia-Bray	1.339663	0.04921025
Kinderman-Ramage	1.350489	0.05251921
Kinderman-Monahan	1.362735	0.05016925
Quoc. Unif. Leva	1.351539	0.05461544
Quoc. Unif. Insua	1.339074	0.04933058
Soma Doze Uniformes	1.364477	0.05853384
Box-Muller	1.344509	0.05054018
Variante Marsaglia	1.342976	0.04685666
Ahrens-Dieter	1.349631	0.04926479

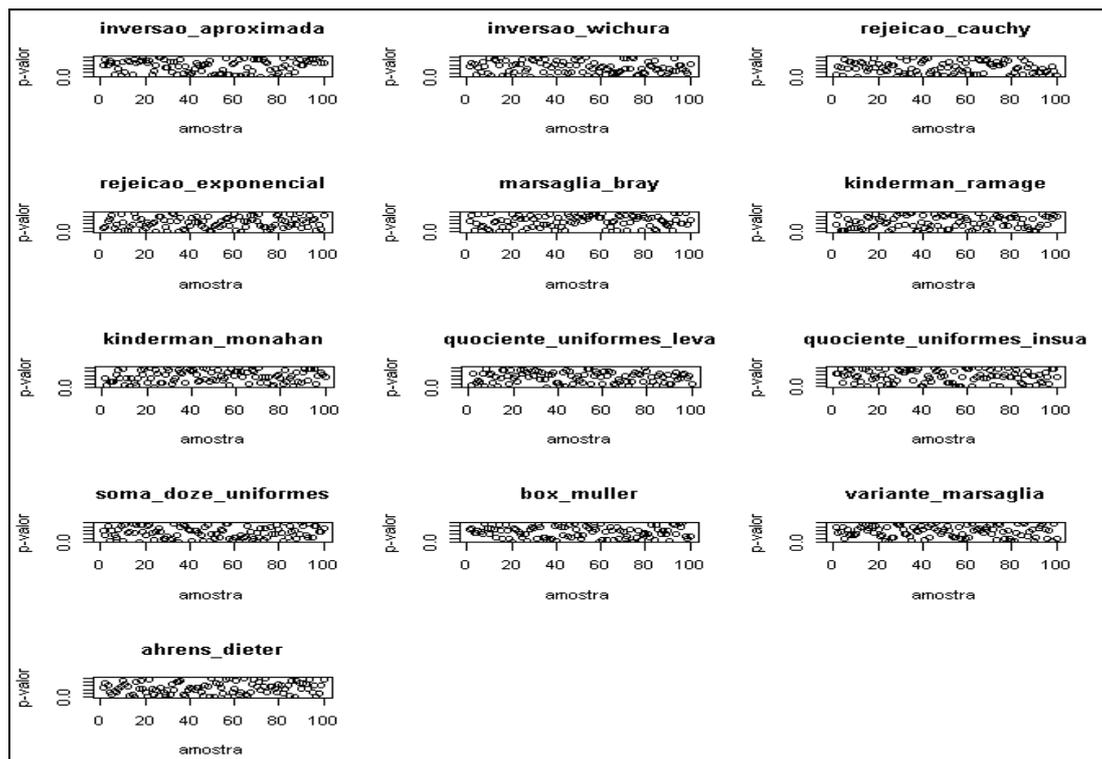
Fonte: elaborada pelo autor.

6.2.9 Teste de Aderência

Tabela 26 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Gerador	Média	Erro Padrão
Inv. Aproximada	0.5381316	0.3195284
Inv. Wichura	0.4840885	0.2940198
Rej. Cauchy	0.4564268	0.2925798
Rej. Exponencial	0.4534743	0.2880199
Marsaglia-Bray	0.5308782	0.293837
Kinderman-Ramage	0.4797575	0.3009116
Kinderman-Monahan	0.5501258	0.2868982
Quoc. Unif. Leva	0.521443	0.2685215
Quoc. Unif. Insua	0.5417912	0.3045601
Soma Doze Uniformes	0.4913324	0.2950767
Box-Muller	0.4940139	0.288086
Variante Marsaglia	0.5443138	0.2808249
Ahrens-Dieter	0.4650643	0.2863751

Fonte: elaborada pelo autor.



Quadro 8 – Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.

6.2.10 Conclusões

De maneira geral, para as 100 amostras com tamanho 1000, os algoritmos Rejeição pela Distribuição Exponencial, Rejeição pela Distribuição de Cauchy, Variante de Marsaglia, Quociente de Uniformes de Leva e Box-Muller foram os que mais se aproximaram, em média, aos valores exatos das medidas analisadas. Para este tamanho de amostra, podemos verificar que os algoritmos baseados no método de aceitação-rejeição mostraram-se mais precisos que os demais, com uma leve vantagem para o algoritmo baseado na Distribuição Exponencial, o qual, de 11 medidas calculadas, em cinco delas pode ser elencado como um dos geradores mais precisos (embora nem sempre tenha apresentado a menor variabilidade).

Assim, podemos dizer que o algoritmo de Rejeição pela Distribuição Exponencial comportou-se melhor para a geração de amostras com tamanho $n = 1000$ da Distribuição Normal padrão.

Tabela 27 – Estimativas médias do algoritmo de Rejeição pela Distribuição Exponencial em 100 amostras de tamanho 1000 da distribuição $N(0,1)$

Medidas	Rej. Exponencial
Esperança	0.0001384719
Máximo	3.223687
Mínimo	-3.182163
Amplitude	7.438032
Desvio-Padrão	1.000373
Assimetria	0.007041263
Curtose	-0.01568317
1º Quartil	-0.6741933
2º Quartil	-0.0006449595
3º Quartil	0.672289
Dist. Interquart.	1.346482

Fonte: elaborada pelo autor.

Rejeição pela Distribuição Exponencial

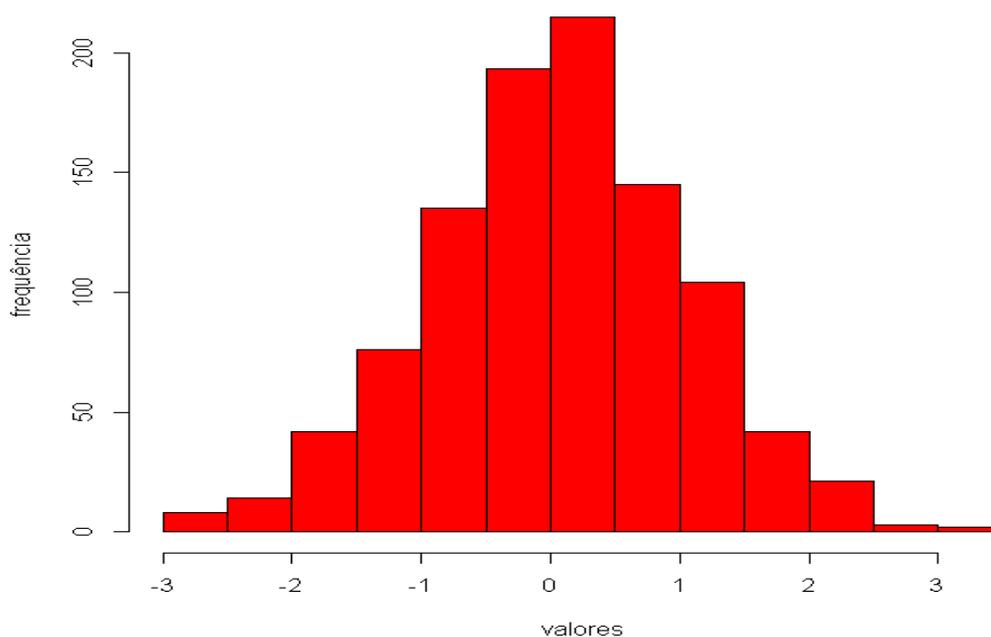


Figura 11 – Histograma de 1000 valores gerados da distribuição $N(0,1)$ pelo algoritmo de Rejeição pela Distribuição Exponencial

Fonte: elaborada pelo autor.

6.3. Estimativas com amostras de tamanho $n = 10000$

Nas tabelas a seguir, apresentamos a média e o erro padrão dos valores estimados pelos geradores para cada uma das medidas analisadas. Nos gráficos, ilustramos o comportamento das estimativas dos geradores ao longo das 100 amostras.

6.3.1 Tempo de Execução

Tabela 28 – Média e erro padrão dos tempos de execução, em segundos, medidos em 100 amostras de tamanho 10000 da distribuição $N(0,1)$

Gerador	Média	Erro Padrão
Inv. Aproximada	0.3785	0.02090744
Inv. Wichura	1.3302	0.02677931
Rej. Cauchy	1.0843	0.02690256
Rej. Exponencial	1.9955	0.03801581
Marsaglia-Bray	0.7728	0.02155941
Kinderman-Ramage	0.6618	0.02041885
Kinderman-Monahan	0.8855	0.01838835
Quoc. Unif. Leva	0.9466	0.02094823
Quoc. Unif. Insua	0.8964	0.02638181
Soma Doze Uniformes	0.3658	0.01718468
Box-Muller	0.4043	0.01451262
Variante Marsaglia	0.5529	0.0193477
Ahrens-Dieter	1.4558	0.03188299

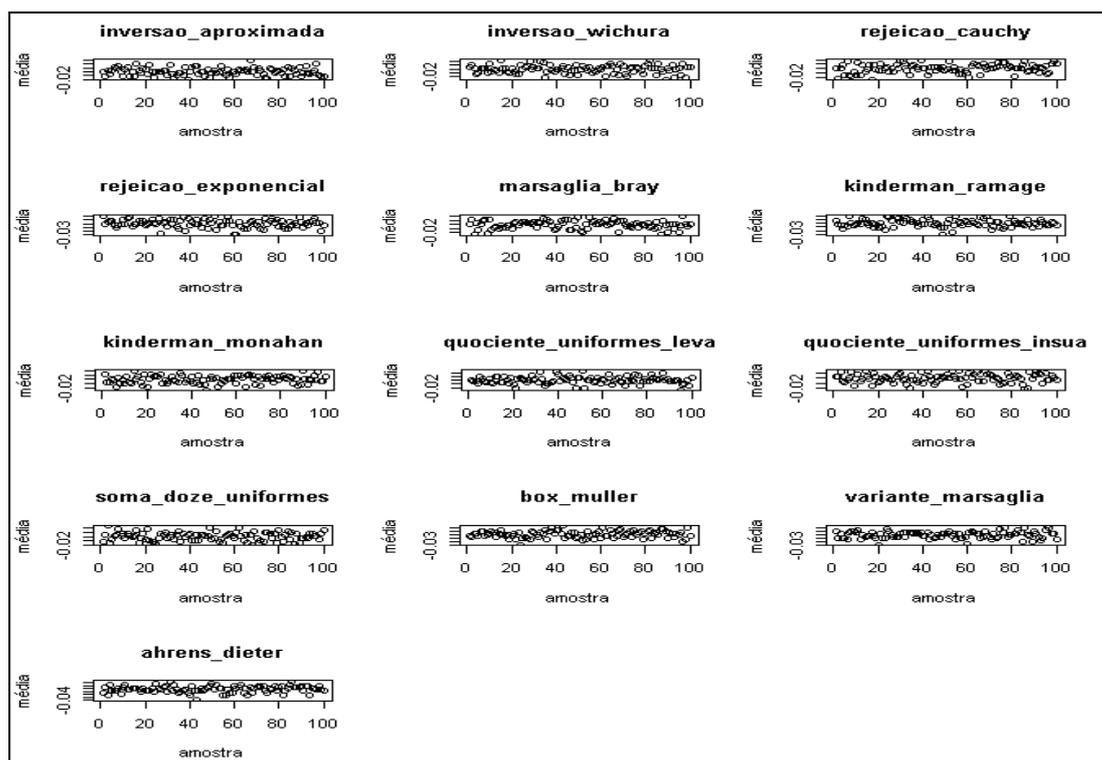
Fonte: elaborada pelo autor.

6.3.2 Esperança

Tabela 29 – Média e erro padrão dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Gerador	Média	Erro Padrão
Inv. Aproximada	4.078576e-05	0.00963597
Inv. Wichura	-0.0005912234	0.0108083
Rej. Cauchy	-2.444202e-05	0.01005797
Rej. Exponencial	0.0006174231	0.01088158
Marsaglia-Bray	-0.0006449274	0.009479491
Kinderman-Ramage	1.44414e-05	0.01013144
Kinderman-Monahan	-0.0005370181	0.008940114
Quoc. Unif. Leva	0.0003107256	0.009677595
Quoc. Unif. Insua	-0.0006837361	0.008291807
Soma Doze Uniformes	0.0002035495	0.009101711
Box-Muller	0.0004087335	0.009980175
Variante Marsaglia	-0.0002465947	0.01037074
Ahrens-Dieter	-0.001972872	0.01112516

Fonte: elaborada pelo autor.



Quadro 9 – Gráficos dos valores estimados para a esperança da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

6.3.3 Máximo e Mínimo

Tabela 30 – Média e erro padrão dos valores estimados para o máximo da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Gerador	Média	Erro Padrão
Inv. Aproximada	3.67812	0.1951931
Inv. Wichura	3.887468	0.361838
Rej. Cauchy	3.833186	0.2765049
Rej. Exponencial	3.858787	0.2889402
Marsaglia-Bray	3.803388	0.3094301
Kinderman-Ramage	3.860144	0.2643519
Kinderman-Monahan	3.88738	0.3085162
Quoc. Unif. Leva	3.835344	0.3134597
Quoc. Unif. Insua	3.911728	0.3068029
Soma Doze Uniformes	3.618569	0.2687502
Box-Muller	3.886008	0.3121628
Variante Marsaglia	3.915806	0.2765527
Ahrens-Dieter	3.791881	0.2822758

Fonte: elaborada pelo autor.

Tabela 31 – Média e erro padrão dos valores estimados para o mínimo da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Gerador	Média	Erro Padrão
Inv. Aproximada	-3.660849	0.2222916
Inv. Wichura	-3.789637	0.2614513
Rej. Cauchy	-3.864249	0.3156253
Rej. Exponencial	-3.87753	0.2795228
Marsaglia-Bray	-3.873991	0.3124315
Kinderman-Ramage	-3.882799	0.3072965
Kinderman-Monahan	-3.830847	0.3023282
Quoc. Unif. Leva	-3.842296	0.3212197
Quoc. Unif. Insua	-3.862537	0.3334432
Soma Doze Uniformes	-3.667338	0.2377975
Box-Muller	-3.827891	0.2785688
Variante Marsaglia	-3.841867	0.3099171
Ahrens-Dieter	-3.855592	0.3202348

Fonte: elaborada pelo autor.

6.3.4 Amplitude

Tabela 32 – Média e erro padrão dos valores estimados para a amplitude da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Gerador	Média	Erro Padrão
Inv. Aproximada	7.338968	0.2875556
Inv. Wichura	7.677105	0.430032
Rej. Cauchy	7.697435	0.3923243
Rej. Exponencial	7.736316	0.4072806
Marsaglia-Bray	7.67738	0.4693779
Kinderman-Ramage	7.742943	0.436601
Kinderman-Monahan	7.718227	0.421272
Quoc. Unif. Leva	7.67764	0.4750116
Quoc. Unif. Insua	7.774266	0.4411526
Soma Doze Uniformes	7.285907	0.3325534
Box-Muller	7.713899	0.441164
Variante Marsaglia	7.757674	0.4261002
Ahrens-Dieter	7.647472	0.4475289

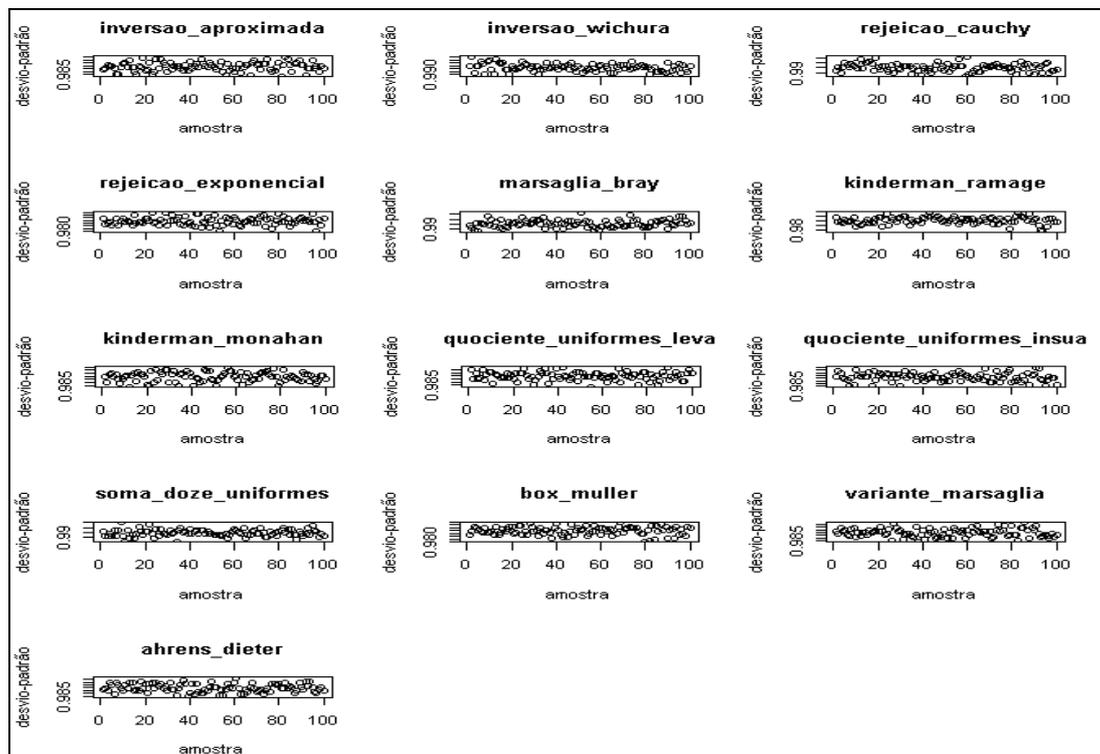
Fonte: elaborada pelo autor.

6.3.5 Desvio-Padrão

Tabela 33 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Gerador	Média	Erro Padrão
Inv. Aproximada	0.9986805	0.007287532
Inv. Wichura	1.001513	0.006440669
Rej. Cauchy	1.00055	0.00780846
Rej. Exponencial	0.9986234	0.007175574
Marsaglia-Bray	1.000342	0.007662034
Kinderman-Ramage	1.000927	0.007522393
Kinderman-Monahan	0.9992433	0.007482825
Quoc. Unif. Leva	0.9995201	0.006742008
Quoc. Unif. Insua	1.001852	0.007659301
Soma Doze Uniformes	0.9992795	0.006831979
Box-Muller	0.9995184	0.006948187
Variante Marsaglia	0.9997244	0.007335063
Ahrens-Dieter	0.9999237	0.007390292

Fonte: elaborada pelo autor.



Quadro 10 – Gráficos dos valores estimados para o desvio-padrão da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

6.3.6 Assimetria

Tabela 34 – Média e erro padrão dos valores estimados para a assimetria da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Gerador	Média	Erro Padrão
Inv. Aproximada	-0.0004667106	0.02768499
Inv. Wichura	0.0003648139	0.02760971
Rej. Cauchy	0.0003492828	0.02255564
Rej. Exponencial	0.0007948796	0.02005484
Marsaglia-Bray	-0.0004806933	0.02317302
Kinderman-Ramage	0.0009172677	0.02602432
Kinderman-Monahan	-0.002262727	0.02740893
Quoc. Unif. Leva	0.001203913	0.02335504
Quoc. Unif. Insua	-0.000310124	0.02618433
Soma Doze Uniformes	-0.001984239	0.02413896
Box-Muller	-0.002687322	0.02319569
Variante Marsaglia	0.003313385	0.02334822
Ahrens-Dieter	-0.002117759	0.02295513

Fonte: elaborada pelo autor.

6.3.7 Curtose

Tabela 35 – Média e erro padrão dos valores estimados para a curtose da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Gerador	Média	Erro Padrão
Inv. Aproximada	-0.007741045	0.04796694
Inv. Wichura	0.0001346954	0.04985316
Rej. Cauchy	-0.003503153	0.04939722
Rej. Exponencial	0.006151363	0.05392317
Marsaglia-Bray	-0.0009852633	0.04832359
Kinderman-Ramage	0.001766893	0.04906434
Kinderman-Monahan	-0.0002936447	0.05054862
Quoc. Unif. Leva	-0.00722352	0.05358064
Quoc. Unif. Insua	-0.001001362	0.05331843
Soma Doze Uniformes	-0.1006847	0.04137404
Box-Muller	-0.0006654236	0.0434142
Variante Marsaglia	0.01010244	0.04539371
Ahrens-Dieter	-0.009525276	0.05025189

Fonte: elaborada pelo autor.

6.3.8 Quartis

Tabela 36 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição N(0,1) em 100 amostras de tamanho 10000

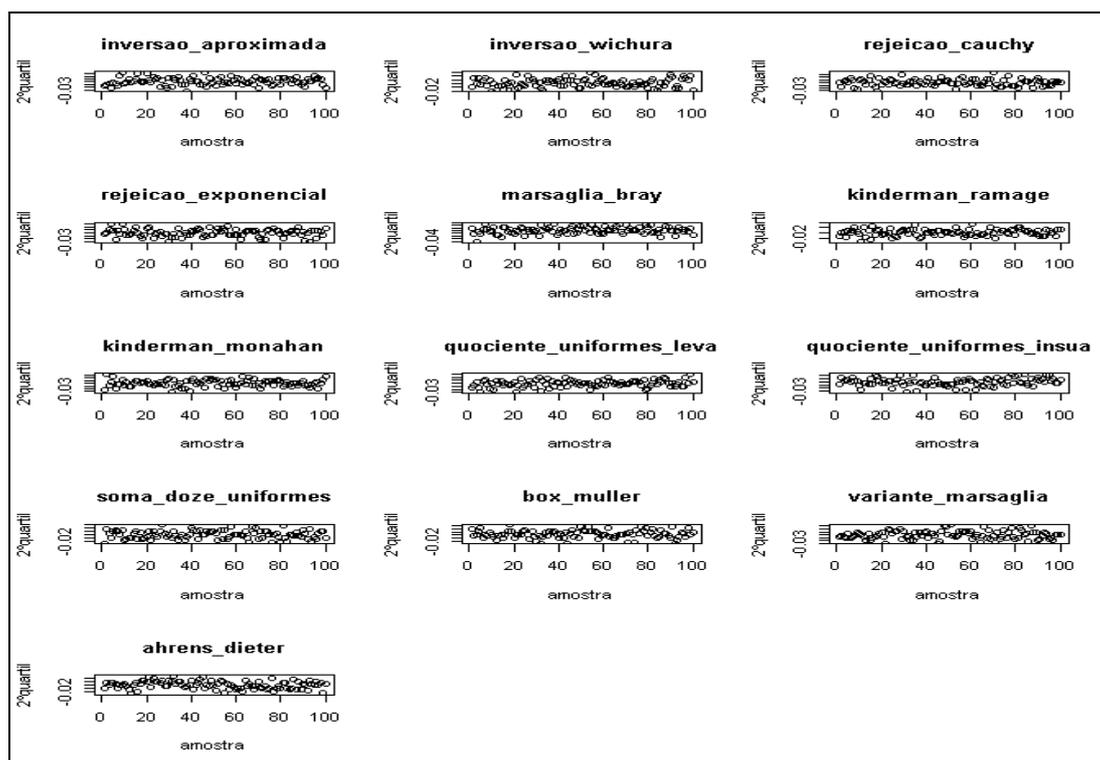
Gerador	Média	Erro Padrão
Inv. Aproximada	-0.6706447	0.01463033
Inv. Wichura	-0.6772253	0.01430168
Rej. Cauchy	-0.6750242	0.01437012
Rej. Exponencial	-0.6730028	0.01422749
Marsaglia-Bray	-0.6753228	0.0127348
Kinderman-Ramage	-0.6750868	0.01398925
Kinderman-Monahan	-0.6742666	0.01293207
Quoc. Unif. Leva	-0.6752771	0.01454069
Quoc. Unif. Insua	-0.6752463	0.01177066
Soma Doze Uniformes	-0.679744	0.01301145
Box-Muller	-0.6742513	0.01185906
Variante Marsaglia	-0.6741473	0.01427654
Ahrens-Dieter	-0.6753484	0.01407911

Fonte: elaborada pelo autor.

Tabela 37 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição N(0,1) em 100 amostras de tamanho 10000

Gerador	Média	Erro Padrão
Inv. Aproximada	0.000306218	0.01364375
Inv. Wichura	-0.00107235	0.01346843
Rej. Cauchy	-0.0008030546	0.01366323
Rej. Exponencial	0.0007396668	0.01453965
Marsaglia-Bray	0.0006881889	0.01251402
Kinderman-Ramage	-0.0001945409	0.01279031
Kinderman-Monahan	0.0004270331	0.01125215
Quoc. Unif. Leva	7.449303e-05	0.01205641
Quoc. Unif. Insua	-0.00120771	0.01086031
Soma Doze Uniformes	4.913974e-05	0.01306972
Box-Muller	0.001085232	0.01280928
Variante Marsaglia	-0.0007062951	0.01242037
Ahrens-Dieter	-0.002233441	0.01333005

Fonte: elaborada pelo autor.



Quadro 11 – Gráficos dos valores estimados para o 2º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

Tabela 38 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Gerador	Média	Erro Padrão
Inv. Aproximada	0.6709169	0.01446716
Inv. Wichura	0.673097	0.01466783
Rej. Cauchy	0.6760559	0.01329012
Rej. Exponencial	0.673573	0.01372757
Marsaglia-Bray	0.6733118	0.01387864
Kinderman-Ramage	0.6741849	0.01347749
Kinderman-Monahan	0.6727967	0.0125165
Quoc. Unif. Leva	0.6738794	0.01368002
Quoc. Unif. Insua	0.6754664	0.01372269
Soma Doze Uniformes	0.680843	0.01379811
Box-Muller	0.6748083	0.0122649
Variante Marsaglia	0.6731657	0.01245246
Ahrens-Dieter	0.6728348	0.01452161

Fonte: elaborada pelo autor.

6.3.9 Distância Interquartilica

Tabela 39 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Gerador	Média	Erro Padrão
Inv. Aproximada	1.341562	0.01668102
Inv. Wichura	1.350322	0.01662842
Rej. Cauchy	1.35108	0.01573124
Rej. Exponencial	1.346576	0.01633259
Marsaglia-Bray	1.348635	0.01622021
Kinderman-Ramage	1.349272	0.01595811
Kinderman-Monahan	1.347063	0.01539988
Quoc. Unif. Leva	1.349156	0.01516611
Quoc. Unif. Insua	1.350713	0.01569344
Soma Doze Uniformes	1.360587	0.01506888
Box-Muller	1.34906	0.01390899
Variante Marsaglia	1.347313	0.01458046
Ahrens-Dieter	1.348183	0.01578208

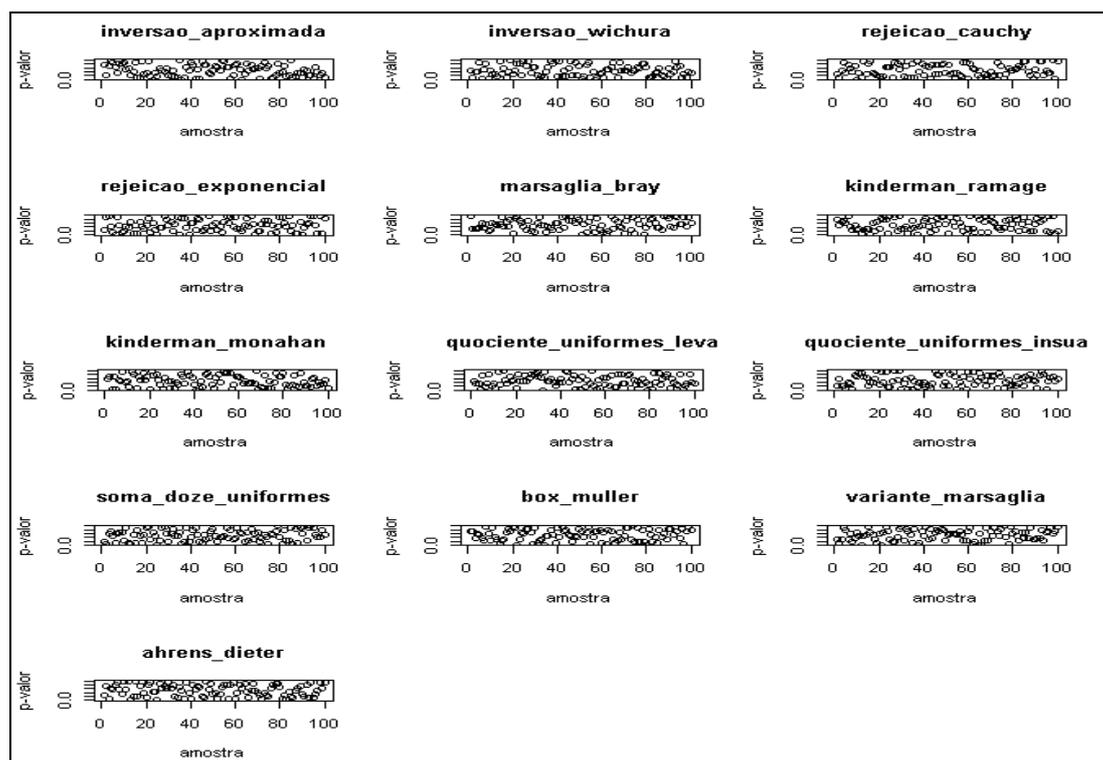
Fonte: elaborada pelo autor.

6.3.10 Teste de Aderência

Tabela 40 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Gerador	Média	Erro Padrão
Inv. Aproximada	0.4732453	0.2935678
Inv. Wichura	0.4562296	0.285868
Rej. Cauchy	0.4648547	0.3030879
Rej. Exponencial	0.4793055	0.3023391
Marsaglia-Bray	0.5396041	0.2829489
Kinderman-Ramage	0.5195344	0.2987819
Kinderman-Monahan	0.469697	0.2880658
Quoc. Unif. Leva	0.4770902	0.2782912
Quoc. Unif. Insua	0.4835915	0.2875804
Soma Doze Uniformes	0.5123736	0.2748049
Box-Muller	0.5313329	0.2923391
Variante Marsaglia	0.5197972	0.2823057
Ahrens-Dieter	0.5222597	0.2979281

Fonte: elaborada pelo autor.



Quadro 12 – Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $N(0,1)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

6.3.11 Conclusões

De maneira geral, para as 100 amostras com tamanho 10000, os algoritmos Kinderman-Ramage, Box-Muller e Variante de Marsaglia foram os que mais se aproximaram, em média, aos valores exatos das medidas analisadas. Em se tratando de precisão das estimativas, o algoritmo de Kinderman-Ramage mostrou-se um pouco melhor que os demais. Entretanto, sua variabilidade (erro padrão) sempre esteve acima dos demais. O algoritmo de Box-Muller, ao contrário, mostrou ser mais constante nas suas estimativas, pois foi o gerador que apresentou, na maioria das vezes, a menor variabilidade. Além do mais, o algoritmo de Box-Muller foi um dos mais rápidos dentre todos os algoritmos analisados.

Por tudo isso, podemos dizer que o algoritmo de Box-Muller comportou-se melhor para a geração de amostras com tamanho $n = 10000$ da Distribuição Normal padrão.

Tabela 41 – Estimativas médias do algoritmo de Box-Muller em 100 amostras de tamanho 10000 da distribuição $N(0,1)$

Medidas	Box-Muller
Tempo de Execução	0.4043 s
Esperança	0.0004087335
Máximo	3.886008
Mínimo	-3.827891
Amplitude	7.713899
Desvio-Padrão	0.9995184
Assimetria	-0.002687322
Curtose	-0.0006654236
1º Quartil	-0.6742513
2º Quartil	0.001085232
3º Quartil	0.6748083
Dist. Interquart.	1.34906

Fonte: elaborada pelo autor.

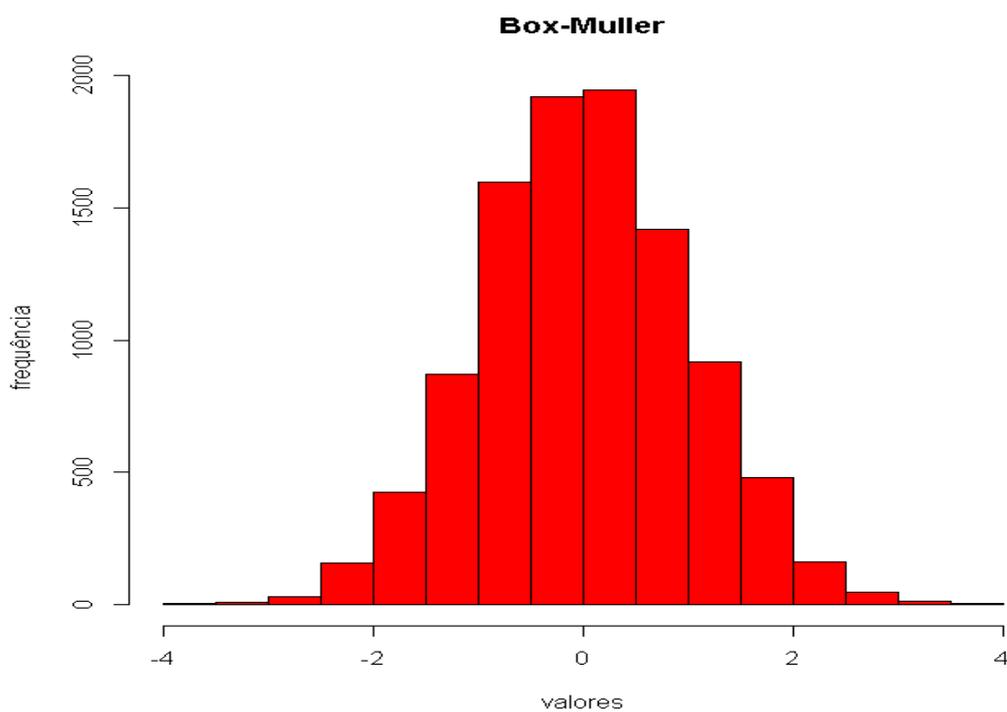


Figura 12 – Histograma de 10000 valores gerados da distribuição $N(0,1)$ pelo algoritmo de Box-Muller

Fonte: elaborada pelo autor.

7. GERADORES DE NÚMEROS ALEATÓRIOS DA DISTRIBUIÇÃO QUI-QUADRADO

Neste capítulo, apresentaremos alguns geradores de números aleatórios da Distribuição Qui-Quadrado encontrados na literatura de Estatística Computacional e Simulação. A desvantagem desses algoritmos está na sua dependência de outras distribuições de probabilidade, como a Distribuição Normal, por exemplo. Ou seja, para se implementar o algoritmo, é necessário saber gerar ou dispor de um bom gerador de alguma outra distribuição de probabilidade. No caso dos algoritmos para a Distribuição Qui-Quadrado, geralmente é necessário um bom gerador de ocorrências da Distribuição Gama ou da Distribuição Normal padrão. Em nossos estudos posteriores, onde analisamos as estimativas desses geradores, utilizamos o gerador normal disponível por padrão no software R (Inversão por Wichura) na implementação dos algoritmos da Distribuição Qui-Quadrado.

7.1 Geradores baseados na Distribuição Gama

Como vimos em nossa Revisão Teórica, a Distribuição Qui-Quadrado também pode ser expressa em termos da Distribuição Gama com parâmetros $k = \nu/2$ e $\theta = 2$ (ou, equivalentemente, $\alpha = \nu/2$ e $\beta = 1/2$), onde ν são os graus de liberdade. Assim, podemos aproveitar os algoritmos para geração de números aleatórios da Distribuição Gama para gerar ocorrências da Distribuição Qui-Quadrado.

Geralmente, na literatura especializada, os algoritmos para geração de números aleatórios da Distribuição Gama são apresentados de acordo com o valor assumido para o parâmetro α : geradores da Gama para $\alpha \leq 1$ e geradores da Gama para $\alpha > 1$. Então, podemos combinar esses dois algoritmos (um para $\alpha \leq 1$ e outro para $\alpha > 1$) e dispor de um gerador da Gama para qualquer valor de α .

Para $\alpha \leq 1$, caso da Distribuição Qui-Quadrado com grau de liberdade ν igual a 1 ou 2, podemos utilizar o algoritmo de rejeição proposto por Ahrens e Dieter (1974):

$$\text{Fazer } \beta = \frac{e + \alpha}{e}, \quad c = \frac{1}{\alpha}.$$

Repita

$$\text{Gerar } U, W \sim U[0;1].$$

$$\text{Fazer } V = \beta U.$$

Se $V \leq 1$ então

$$\text{Fazer } X = V^c.$$

$$\text{Fazer } \textit{Aceitar} = [W \leq e^{-X}].$$

senão

Fazer $X = -\log(c(\beta - V))$.

Fazer $Aceitar = [W \leq X^{\alpha-1}]$.

Até que $Aceitar = Verdadeiro$.

Sair X .

Para $\alpha > 1$, caso da Distribuição Qui-Quadrado com grau de liberdade ν maior do que 2, podemos utilizar o algoritmo de Cheng e Feast (1979), o qual é baseado no método do quociente de uniformes:

Fazer $c_1 = \alpha - 1$, $c_2 = \frac{\alpha - (1/6\alpha)}{c_1}$, $c_3 = \frac{2}{\alpha - 1}$, $c_4 = 1 + c_3$, $c_5 = \frac{1}{\sqrt{\alpha}}$.

1. Enquanto $U_1 \notin (0,1)$

Gerar $U_1, U_2 \sim U(0,1)$.

Se $\alpha > 2.5$ então

Fazer $U_1 = U_2 + c_5(1 - 1.86U_1)$.

2. Fazer $W = c_2U_2/U_1$.

3. Se $c_3U_1 + W + W^{-1} \leq c_4$ então ir a 5.

4. Se $c_3 \log U_1 - \log W + W \geq 1$ então voltar a 1.

5. Sair $X = c_1W$.

O software R utiliza um algoritmo baseado na Distribuição Gama para gerar valores da Distribuição Qui-Quadrado. Por ser muito extenso, não forneceremos aqui o pseudocódigo para este algoritmo. Entretanto, sua implementação, em linguagem C, pode ser encontrada no código-fonte do software, disponível no endereço eletrônico <http://www.r-project.org>. Neste trabalho, chamaremos este gerador de "Gama - R".

7.2 Geradores baseados na Distribuição Normal

Os algoritmos seguintes são baseados unicamente na Distribuição Normal. Lembrando que ν são os graus de liberdade da Distribuição Qui-Quadrado, a qual queremos gerar.

7.2.1 Soma dos Quadrados de Normais

Gerar $Z_1, \dots, Z_\nu \sim N(0,1)$.

Sair $X = \sum_{i=1}^{\nu} Z_i^2$.

7.2.2 Aproximação Normal

Gerar $Z \sim N(0,1)$.

$$\text{Sair } X = \frac{(Z + \sqrt{2\nu - 1})^2}{2}.$$

7.3 Gerador baseado na Distribuição Uniforme e na Distribuição Normal

7.3.1 Logaritmo do Produto de Uniformes

Se ν é par então

Gerar $U_1, \dots, U_{\nu/2} \sim U[0;1]$.

$$\text{Sair } X = -2 \log \left(\prod_{i=1}^{\nu/2} U_i \right).$$

Se ν é ímpar então

Gerar $U_1, \dots, U_{(\nu-1)/2} \sim U[0;1]$.

Gerar $Z \sim N(0,1)$.

$$\text{Sair } X = -2 \log \left(\prod_{i=1}^{(\nu-1)/2} U_i \right) + Z^2.$$

7.4 Gerador baseado na Distribuição Exponencial e na Distribuição Normal

O algoritmo seguinte é dependente de um bom gerador de ocorrências da Distribuição Exponencial e da Distribuição Normal padrão. Esse gerador pode ser obtido através da implementação de um algoritmo específico. Em nossas análises, utilizaremos o método da inversão para gerar números aleatórios da Distribuição Exponencial. Como já dissemos anteriormente, os números aleatórios da Distribuição Normal serão gerados através do algoritmo disponível por padrão no software R (Inversão por Wichura).

Se ν é par então

Gerar $Y_1, \dots, Y_{\nu/2} \sim Exp(2)$.

$$\text{Sair } X = \sum_{i=1}^{\nu/2} Y_i.$$

Se ν é ímpar então

Gerar $Y \sim \chi^2(\nu - 1)$.

Gerar $Z \sim N(0,1)$.

Sair $X = Y + Z^2$.

Note que, quando ν é ímpar, este algoritmo remete à geração, primeiramente, de uma qui-quadrado com grau de liberdade ν par.

8. ESTIMATIVAS DAS AMOSTRAS DA DISTRIBUIÇÃO QUI-QUADRADO

Nos capítulos subsequentes, apresentaremos as estimativas com amostras de tamanho $n = 100$, $n = 1000$ e $n = 10000$ da Distribuição Qui-Quadrado com graus de liberdade ν iguais a 1, 17 e 50 para as seguintes medidas: esperança, máximo e mínimo, amplitude, desvio-padrão, assimetria, curtose, quartis, distância interquartílica e valor-p para o teste de aderência. Os valores “exatos” das medidas para a Distribuição Qui-Quadrado com graus de liberdade ν iguais a 1, 17 e 50 são, respectivamente: esperança = 1, 17, 50; máximo = 19.51142, 53.97429, 104.5417 e mínimo = $1.570796e-10$, 2.303421, 18.15884; amplitude = 19.51142, 51.67087, 86.38286; desvio-padrão = 1.414214, 5.830952, 10; assimetria = 2.828427, 0.6859943, 0.4; curtose = 12, 0.7058824, 0.24; 1º quartil = 0.101531, 12.79193, 42.94208; 2º quartil = 0.4549364, 16.33818, 49.33494 e 3º quartil = 1.323304, 20.48868, 56.3336; distância interquartílica = 1.221773, 7.69675, 13.39152. Considera-se como máximo o valor x tal que a probabilidade $P(X \leq x) = 0.99999$, e como mínimo o valor x tal que a probabilidade $P(X \leq x) = 0.00001$.

8.1 Estimativas com amostras de tamanho $n = 100$

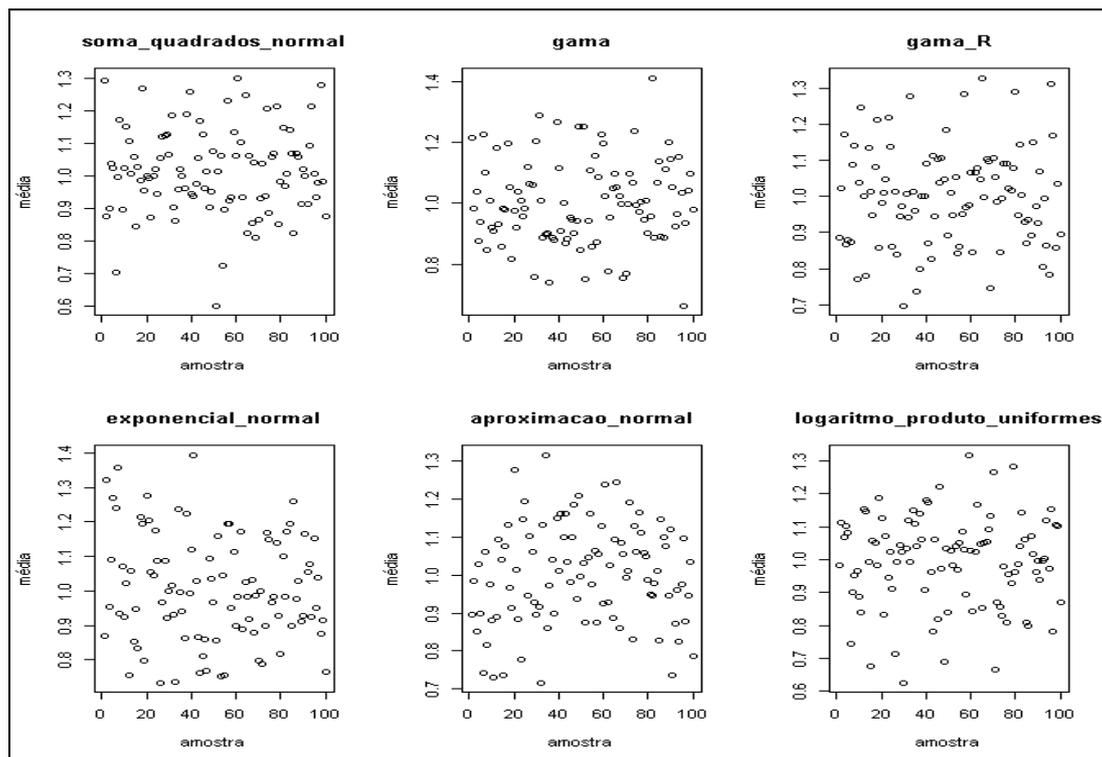
Nas tabelas abaixo, apresentamos a média e o erro padrão dos valores estimados pelos geradores para cada uma das medidas analisadas. Nos gráficos, ilustramos o comportamento das estimativas dos geradores ao longo das 100 amostras.

8.1.1 Esperança

Tabela 42 – Média e erro padrão dos valores estimados para a esperança da distribuição Qui-Quadrado em 100 amostras de tamanho 100

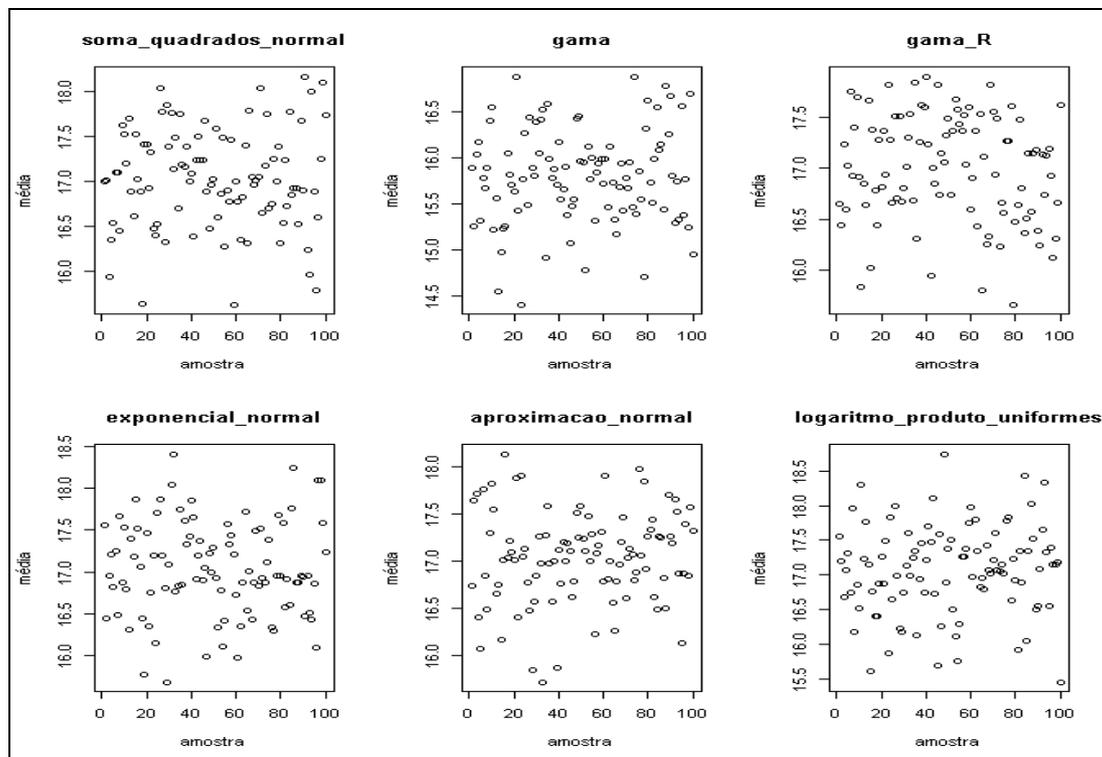
Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.013611	0.1303288	17.02971	0.552995	50.03928	1.07885
Gama	1.002445	0.1390373	15.79455	0.5109335	46.71836	1.05343
Gama - R	1.00102	0.1345539	16.99746	0.5218185	49.9462	1.033436
Exponencial-Normal	1.006839	0.1528259	17.03385	0.5591885	50.28968	1.144026
Aprox. Normal	1.002002	0.1305188	17.04111	0.4866329	50.09301	1.054753
Log. Produto Uniformes	0.9949566	0.1376038	17.07459	0.6391054	49.99594	1.067584

Fonte: elaborada pelo autor.



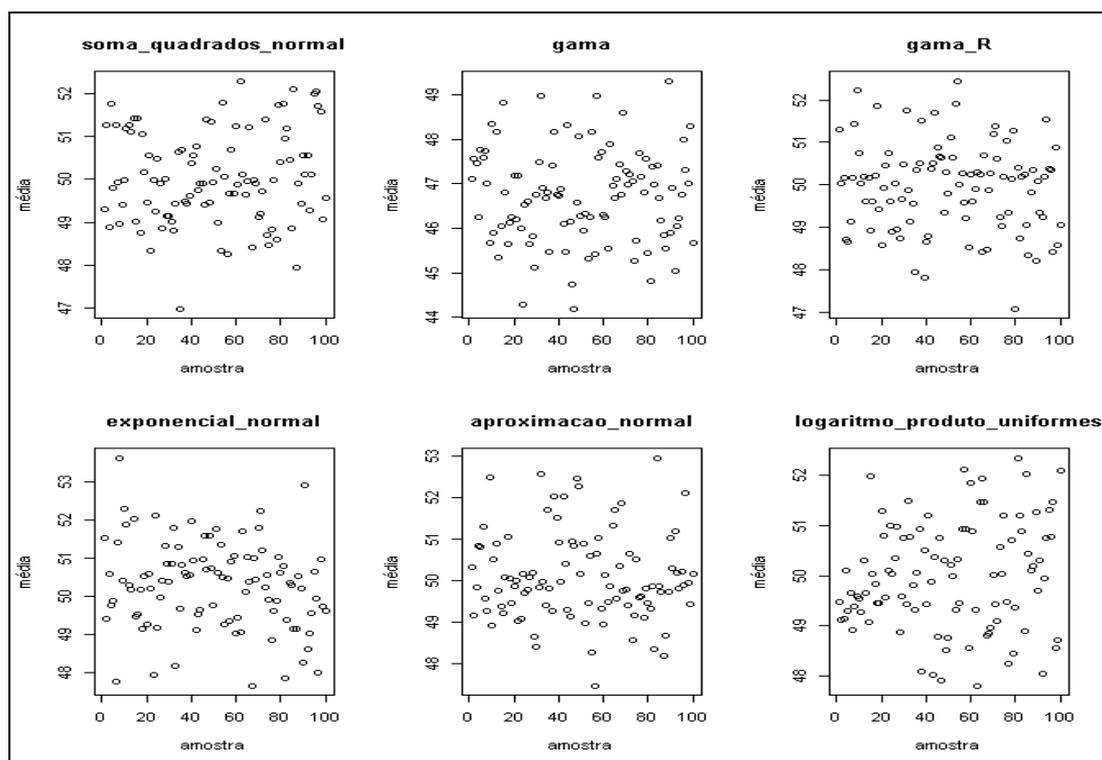
Quadro 13 – Gráficos dos valores estimados para a esperança da distribuição $\chi^2(1)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 14 - Gráficos dos valores estimados para a esperança da distribuição $\chi^2(17)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 15 – Gráficos dos valores estimados para a esperança da distribuição $\chi^2(50)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.

8.1.2 Máximo e Mínimo

Tabela 43 – Média e erro padrão dos valores estimados para o máximo da distribuição Qui-Quadrado em 100 amostras de tamanho 100

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	7.83478	2.373793	35.66113	3.878113	77.77607	5.548218
Gama	7.789801	2.701185	35.23331	3.718802	79.60354	4.842029
Gama - R	7.674068	2.103014	35.00361	3.928359	78.77379	6.05703
Exponencial-Normal	7.959177	2.765039	35.03132	3.479249	79.27873	6.22584
Aprox. Normal	6.305941	1.522088	34.18193	3.918248	78.35757	5.643764
Log. Produto Uniformes	7.220294	1.847453	36.05747	4.39571	78.5726	5.669813

Fonte: elaborada pelo autor.

Tabela 44 – Média e erro padrão dos valores estimados para o mínimo da distribuição Qui-Quadrado em 100 amostras de tamanho 100

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	0.0004084102	0.000876263	5.788777	1.203202	28.80684	2.779686
Gama	0.0003683153	0.0007983038	5.626165	1.092559	27.43229	2.715992
Gama - R	0.0002209926	0.0002939924	5.941221	1.228423	28.36079	3.158944
Exponencial-Normal	0.0002616999	0.0004857139	5.920704	1.238791	28.91423	3.056686
Aprox. Normal	0.0005140366	0.0008861322	5.257081	1.239701	28.04644	2.919457
Log. Produto Uniformes	0.0003712888	0.0007504343	5.824605	1.202796	28.91065	2.918235

Fonte: elaborada pelo autor.

8.1.3 Amplitude

Tabela 45 - Média e erro padrão dos valores estimados para a amplitude da distribuição Qui-Quadrado em 100 amostras de tamanho 100

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	7.834372	2.373783	29.87235	4.025711	48.96923	6.082347
Gama	7.789433	2.701208	29.60714	3.86343	52.17125	5.683367
Gama - R	7.673847	2.102969	29.06239	4.277951	50.413	6.859562
Exponencial-Normal	7.958915	2.765042	29.11062	3.625154	50.3645	7.024145
Aprox. Normal	6.305427	1.522081	28.92485	4.078494	50.31113	6.051686
Log. Produto Uniformes	7.219923	1.847458	30.23287	4.616039	49.66195	6.553711

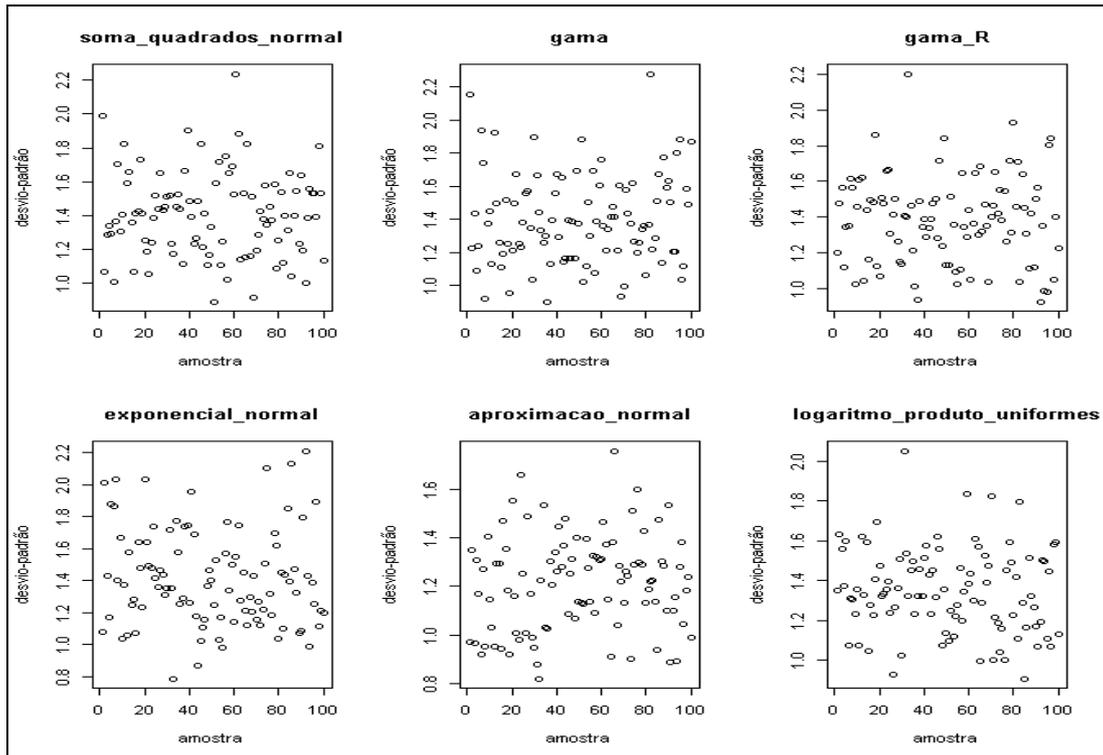
Fonte: elaborada pelo autor.

8.1.4 Desvio-Padrão

Tabela 46 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição Qui-Quadrado em 100 amostras de tamanho 100

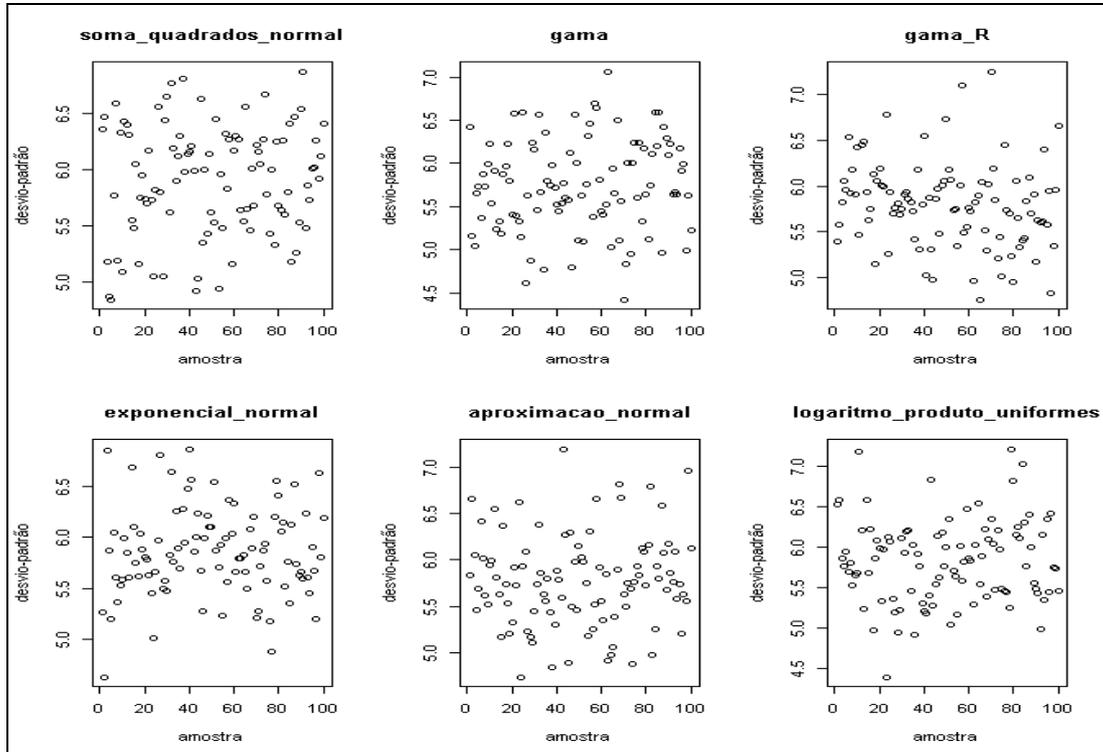
Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.405883	0.2500815	5.90586	0.492661	9.976356	0.7468697
Gama	1.40088	0.2733829	5.740065	0.5319134	9.710136	0.9146257
Gama - R	1.385505	0.2430769	5.796403	0.4647291	9.993621	0.7290067
Exponencial-Normal	1.415108	0.3011704	5.853997	0.4323223	10.02227	0.7768654
Aprox. Normal	1.212119	0.1943454	5.777463	0.4962268	10.00456	0.6611747
Log. Produto Uniformes	1.347999	0.2137367	5.829967	0.5089148	9.979114	0.6958168

Fonte: elaborada pelo autor.



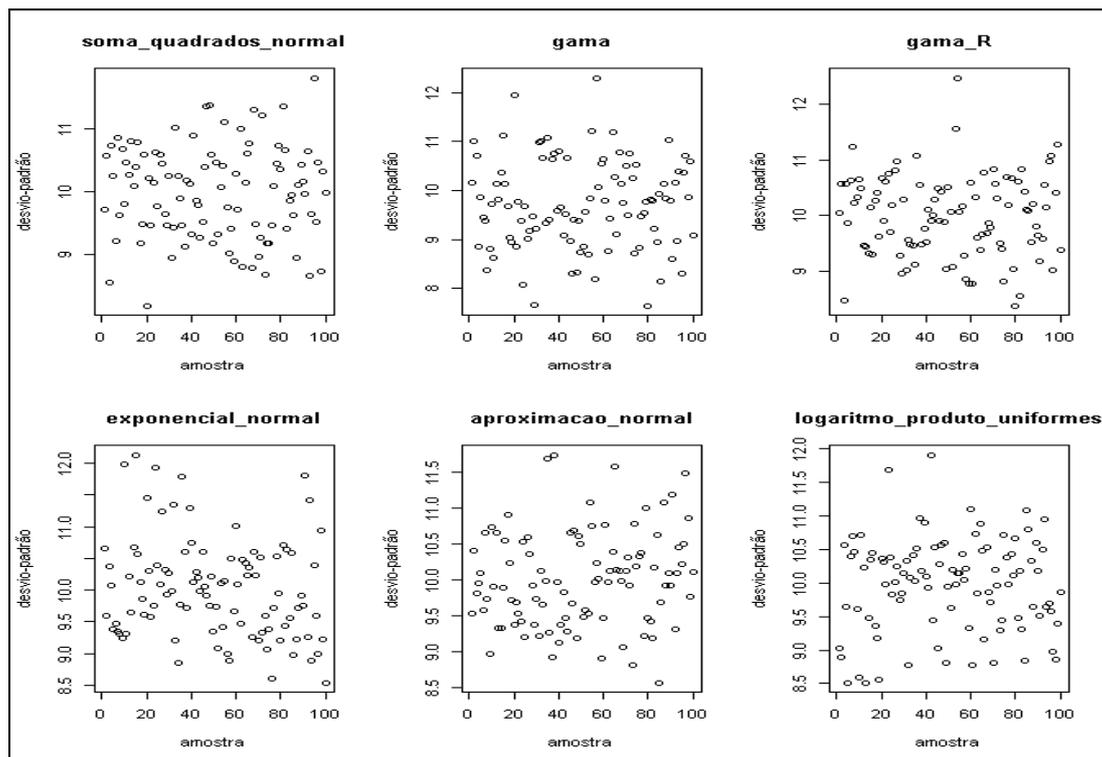
Quadro 16 – Gráficos dos valores estimados para o desvio-padrão da distribuição $\chi^2(1)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 17 – Gráficos dos valores estimados para o desvio-padrão da distribuição $\chi^2(17)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 18 – Gráficos dos valores estimados para o desvio-padrão da distribuição χ^2 (50) em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.

8.1.5 Assimetria

Tabela 47 – Média e erro padrão dos valores estimados para a assimetria da distribuição Qui-Quadrado em 100 amostras de tamanho 100

Gerador	χ^2 (1)		χ^2 (17)		χ^2 (50)	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	2.437157	0.7363119	0.6527182	0.3113891	0.3133786	0.2472258
Gama	2.501086	0.8445426	1.028708	0.3318637	0.9355545	0.294272
Gama - R	2.465459	0.6586784	0.6393768	0.2962906	0.3678915	0.2628729
Exponencial-Normal	2.52915	0.8612415	0.6265297	0.2533741	0.3711038	0.2690283
Aprox. Normal	1.985724	0.4648373	0.4940468	0.2633555	0.2783906	0.2403129
Log. Produto Uniformes	2.366051	0.6861194	0.6659516	0.2887007	0.3798524	0.2719963

Fonte: elaborada pelo autor.

8.1.6 Curtose

Tabela 48 – Média e erro padrão dos valores estimados para a curtose da distribuição Qui-Quadrado em 100 amostras de tamanho 100

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	7.934095	6.357401	0.5951687	1.024767	0.04877901	0.5709505
Gama	8.432716	7.902393	1.560152	1.325459	1.574659	0.9745901
Gama - R	7.879201	5.334487	0.5266214	0.9563342	0.2023706	0.6732064
Exponencial-Normal	8.662333	7.631208	0.4418246	0.84462	0.1826368	0.6922066
Aprox. Normal	4.948974	3.257744	0.2862528	0.8027683	0.09007922	0.6715497
Log. Produto Uniformes	7.210431	5.633407	0.7213792	1.013219	0.1640637	0.7074076

Fonte: elaborada pelo autor.

8.1.7 Quartis

Tabela 49 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 100

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	0.1047113	0.0341031	12.82703	0.6685445	43.06862	1.196437
Gama	0.1146729	0.03877371	11.95431	0.5376503	40.39959	1.021207
Gama - R	0.1173784	0.04353134	12.85218	0.6172795	43.05649	1.083455
Exponencial-Normal	0.1107677	0.03680077	12.79426	0.6485682	43.24628	1.34047
Aprox. Normal	0.1458536	0.04844308	12.94459	0.6032102	43.05994	1.303908
Log. Produto Uniformes	0.115139	0.04209123	12.94895	0.7349228	42.94988	1.264603

Fonte: elaborada pelo autor.

Tabela 50 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 100

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	0.4649426	0.100327	16.38709	0.7008467	49.54433	1.250187
Gama	0.4633268	0.1045903	14.79439	0.5714141	45.50671	1.016658
Gama - R	0.4654244	0.103415	16.29942	0.6220623	49.28501	1.245475
Exponencial-Normal	0.4651829	0.107358	16.37297	0.7046895	49.67072	1.305889
Aprox. Normal	0.5534392	0.1090092	16.5745	0.6441261	49.71388	1.342246
Log. Produto Uniformes	0.4731846	0.1050414	16.39482	0.7899035	49.32701	1.460295

Fonte: elaborada pelo autor.

Tabela 51 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 100

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.348428	0.231796	20.48554	0.8147951	56.39284	1.578335
Gama	1.319698	0.2354402	18.61027	0.8604511	51.11427	1.267399
Gama - R	1.325479	0.2314025	20.50397	0.8366538	56.18925	1.608225
Exponencial-Normal	1.321106	0.2339665	20.53985	0.7752193	56.64179	1.447134
Aprox. Normal	1.423434	0.229321	20.56278	0.8498754	56.53935	1.39651
Log. Produto Uniformes	1.330116	0.2568371	20.5356	0.8926431	56.3541	1.52539

Fonte: elaborada pelo autor.

8.1.8 Distância Interquartilica

Tabela 52 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição Qui-Quadrado em 100 amostras de tamanho 100

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.243716	0.2279235	7.658515	0.845224	13.32422	1.516947
Gama	1.205025	0.2292197	6.655958	0.8579352	10.71468	1.226004
Gama - R	1.2081	0.2173916	7.651788	0.9086198	13.13276	1.570146
Exponencial-Normal	1.210338	0.2275746	7.745586	0.8420262	13.39551	1.490358
Aprox. Normal	1.27758	0.2132516	7.618193	0.9913532	13.47941	1.497422
Log. Produto Uniformes	1.214977	0.2421293	7.586655	0.8932431	13.40422	1.372958

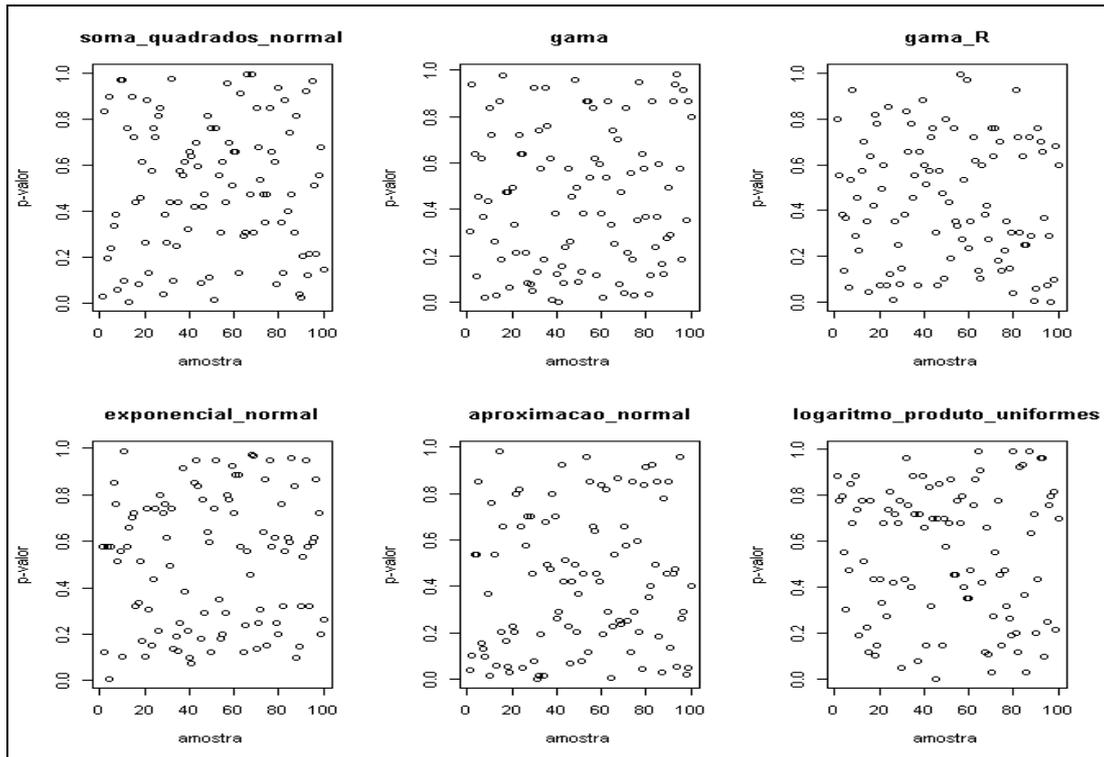
Fonte: elaborada pelo autor.

8.1.9 Teste de Aderência

Tabela 53 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição Qui-Quadrado em 100 amostras de tamanho 100

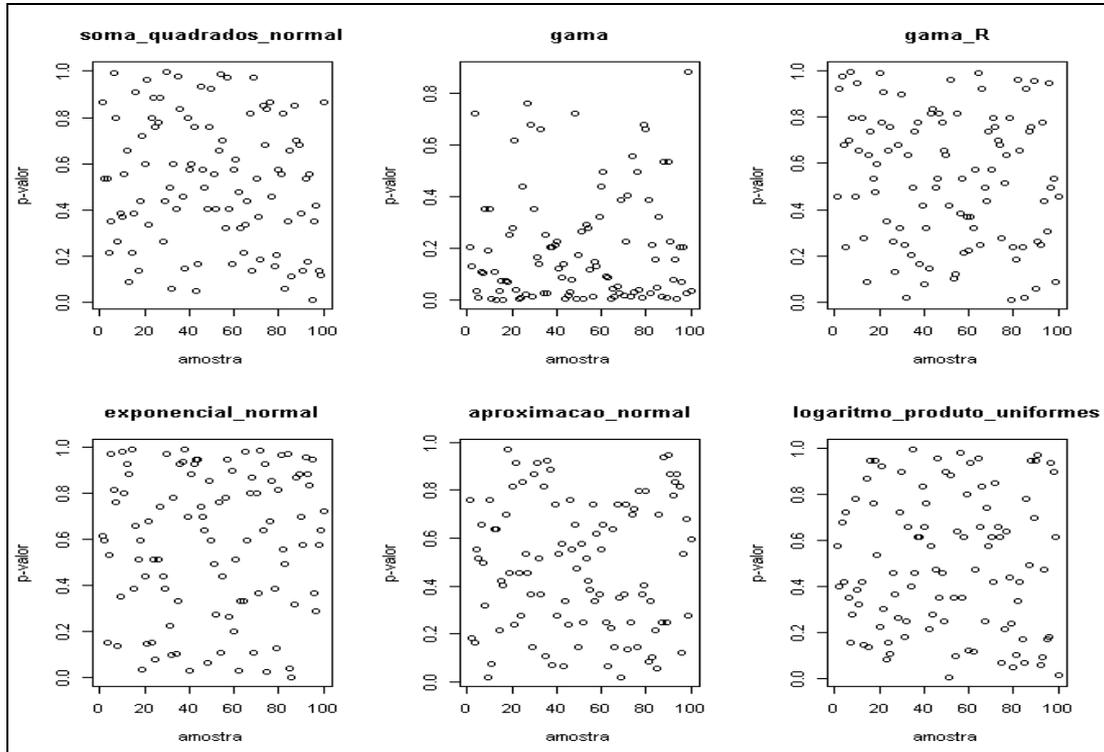
Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	0.5037142	0.2929729	0.523065	0.2781264	0.5193318	0.2931682
Gama	0.449669	0.3019771	0.1986955	0.2163475	0.03411114	0.06691684
Gama - R	0.441788	0.2710657	0.5379653	0.2825794	0.4658109	0.2913742
Exponencial-Normal	0.5141304	0.2844422	0.5806409	0.3088765	0.5233726	0.3003739
Aprox. Normal	0.4167744	0.2967426	0.4828685	0.2705549	0.5390299	0.295332
Log. Produto Uniformes	0.5436135	0.289159	0.4973817	0.2953249	0.4599109	0.2711607

Fonte: elaborada pelo autor.



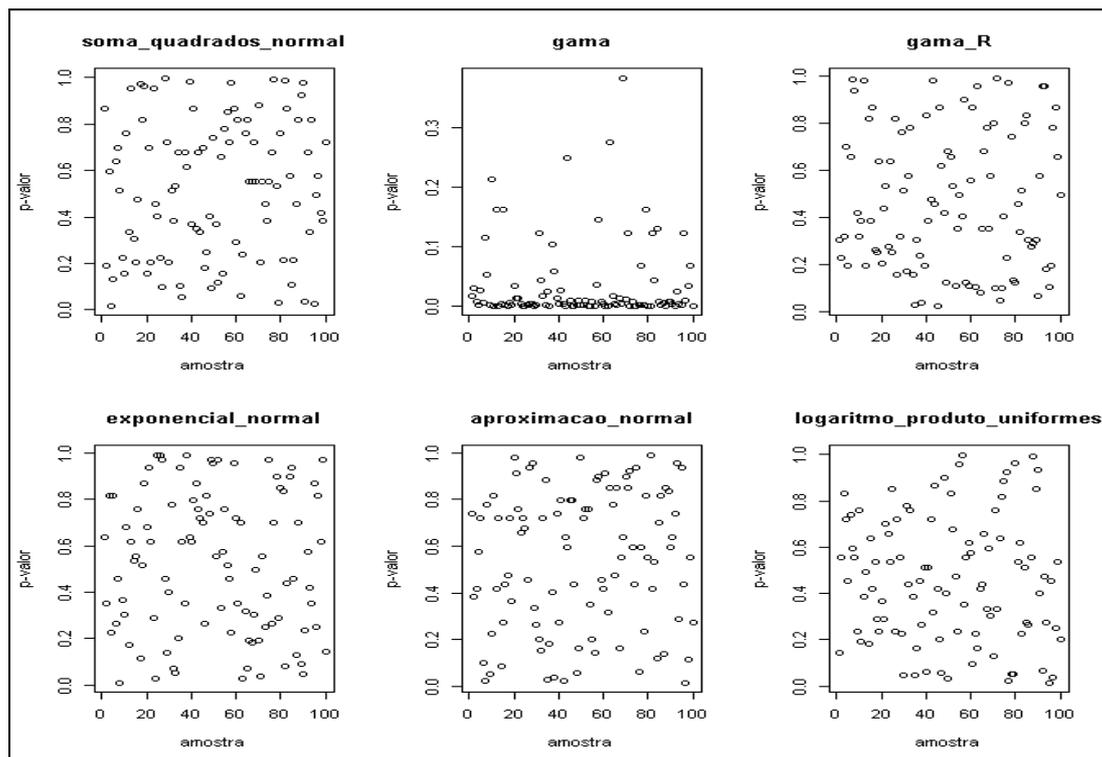
Quadro 19 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2(1)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 20 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2(17)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 21 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2(50)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.

8.1.10 Conclusões

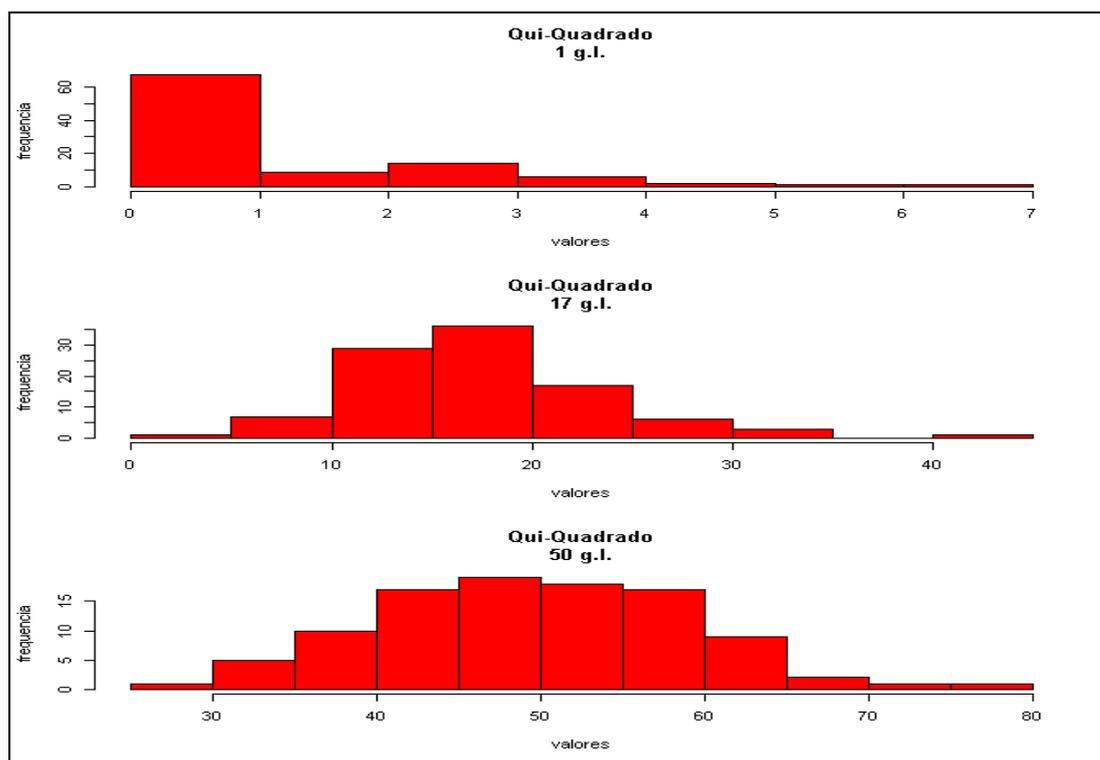
Podemos observar um resultado bastante equilibrado entre os algoritmos Soma Quadrados Normal, Gama-R, Exponencial Normal e Logaritmo Produto Uniformes, com uma leve vantagem para os dois últimos que, aparentemente, se mostraram, em média, um pouco mais precisos que os demais. Entretanto, as análises demonstraram que podemos seguramente utilizar qualquer um desses quatro algoritmos para gerar amostras com tamanho $n = 100$ da Distribuição Qui-Quadrado sem notarmos grandes diferenças entre eles.

De maneira geral, o algoritmo Logaritmo Produto Uniformes parece ter apresentado o menor erro padrão. Como este é um dos algoritmos da Distribuição Qui-Quadrado de mais fácil implementação, o indicamos para gerar amostras com tamanho $n = 100$ da Distribuição Qui-Quadrado.

Tabela 54 – Estimativas médias do algoritmo Logaritmo Produto Uniformes em 100 amostras de tamanho 100 da Distribuição Qui-Quadrado

Medidas	Logaritmo Produto Uniformes		
	$\chi^2(1)$	$\chi^2(17)$	$\chi^2(50)$
Esperança	0.9949566	17.07459	49.99594
Máximo	7.220294	36.05747	78.5726
Mínimo	0.0003712888	5.824605	28.91065
Amplitude	7.219923	30.23287	49.66195
Desvio-Padrão	1.347999	5.829967	9.979114
Assimetria	2.366051	0.6659516	0.3798524
Curtose	7.210431	0.7213792	0.1640637
1º Quartil	0.115139	12.94895	42.94988
2º Quartil	0.4731846	16.39482	49.32701
3º Quartil	1.330116	20.5356	56.3541
Dist. Interquart.	1.214977	7.586655	13.40422

Fonte: elaborada pelo autor.



Quadro 22 – Histogramas de 100 valores gerados da Distribuição Qui-Quadrado pelo algoritmo Logaritmo Produto Uniformes

Fonte: Elaborado pelo autor.

8.2 Estimativas com amostras de tamanho n = 1000

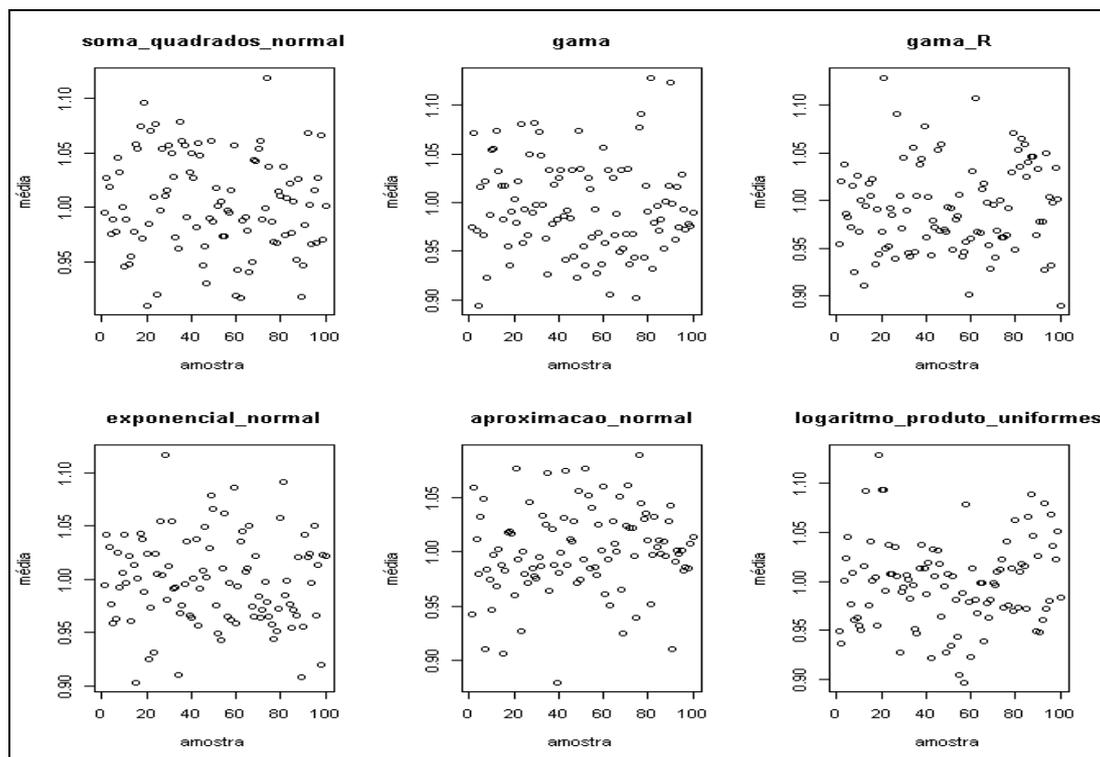
Nas tabelas a seguir, apresentamos a média e o erro padrão dos valores estimados pelos geradores para cada uma das medidas analisadas. Nos gráficos, ilustramos o comportamento das estimativas dos geradores ao longo das 100 amostras.

8.2.1 Esperança

Tabela 55 – Média e erro padrão dos valores estimados para a esperança da distribuição Qui-Quadrado em 100 amostras de tamanho 1000

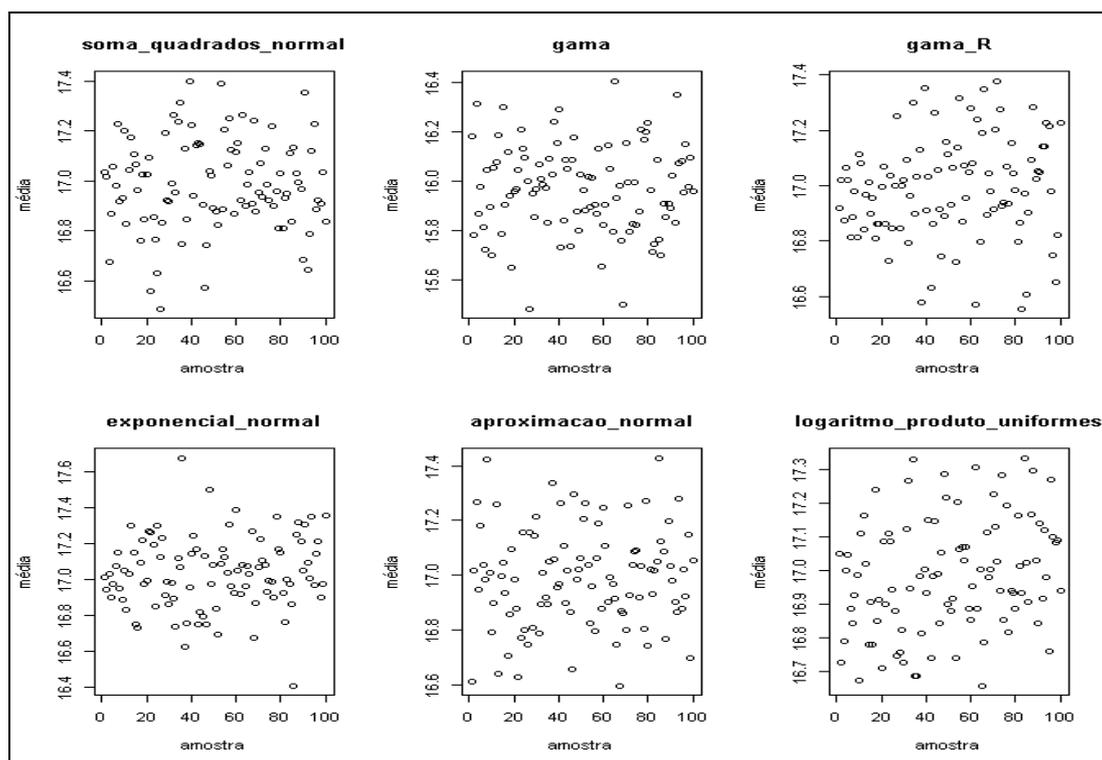
Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.00424	0.04437553	16.98659	0.1831337	50.04077	0.3444357
Gama	0.995014	0.04783313	15.97035	0.1768262	46.81807	0.2901337
Gama - R	0.9923982	0.04598366	16.98918	0.1815391	49.99648	0.3039203
Exponencial-Normal	0.9977112	0.04154762	17.03724	0.1991935	50.01545	0.3632575
Aprox. Normal	1.000872	0.03940131	16.98845	0.1808796	49.99464	0.3247249
Log. Produto Uniformes	0.9983978	0.04421942	16.98404	0.1675511	49.99014	0.3372535

Fonte: elaborada pelo autor.



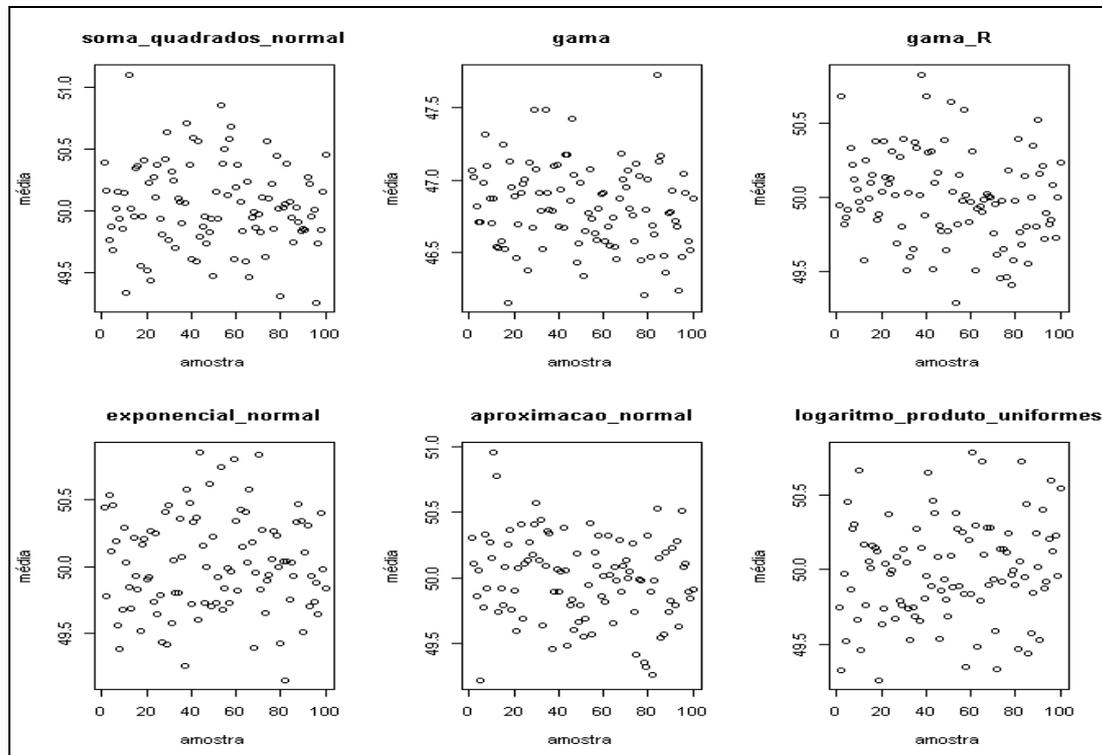
Quadro 23 - Gráficos dos valores estimados para a esperança da distribuição $\chi^2(1)$ em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.



Quadro 24 - Gráficos dos valores estimados para a esperança da distribuição $\chi^2(17)$ em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.



Quadro 25 - Gráficos dos valores estimados para a esperança da distribuição $\chi^2(50)$ em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.

8.2.2 Máximo e Mínimo

Tabela 56 – Média e erro padrão dos valores estimados para o máximo da distribuição Qui-Quadrado em 100 amostras de tamanho 1000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	12.03881	2.178756	42.49086	3.706514	88.90289	5.225067
Gama	11.89885	2.199388	42.46097	2.846838	90.928	4.855209
Gama - R	11.65273	2.024027	42.82514	4.558538	88.19078	4.228634
Exponencial-Normal	11.75759	2.217741	42.71779	3.821089	88.68212	4.938338
Aprox. Normal	9.04643	1.566842	40.7833	3.269835	87.06148	4.500872
Log. Produto Uniformes	11.9309	2.658933	42.61165	4.299816	88.25396	5.074735

Fonte: elaborada pelo autor.

Tabela 57 – Média e erro padrão dos valores estimados para o mínimo da distribuição Qui-Quadrado em 100 amostras de tamanho 1000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	3.379198e-06	5.92024e-06	4.087693	0.7359481	23.5977	2.147051
Gama	2.944007e-06	6.202337e-06	4.001263	0.6803338	23.5406	2.041333
Gama - R	2.740861e-06	4.705749e-06	4.240101	0.7179291	24.02316	1.855858
Exponencial-Normal	2.802556e-06	4.329291e-06	4.109465	0.6971601	23.87666	2.025615
Aprox. Normal	5.15536e-06	1.286584e-05	3.027657	0.8126496	22.49796	2.442721
Log. Produto Uniformes	3.02308e-06	5.406248e-06	4.180386	0.723626	23.89202	2.02322

Fonte: elaborada pelo autor.

8.2.3 Amplitude

Tabela 58 – Média e erro padrão dos valores estimados para a amplitude da distribuição Qui-Quadrado em 100 amostras de tamanho 1000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	12.03881	2.178756	38.40316	3.893352	65.3052	5.685702
Gama	11.89884	2.199388	38.45971	2.856143	67.3874	5.02528
Gama - R	11.65272	2.024028	38.58504	4.655872	64.16762	4.658971
Exponencial-Normal	11.75758	2.217742	38.60833	3.861792	64.80546	5.083443
Aprox. Normal	9.046425	1.566843	37.75564	3.390553	64.56352	5.036253
Log. Produto Uniformes	11.9309	2.658932	38.43126	4.417534	64.36194	5.610083

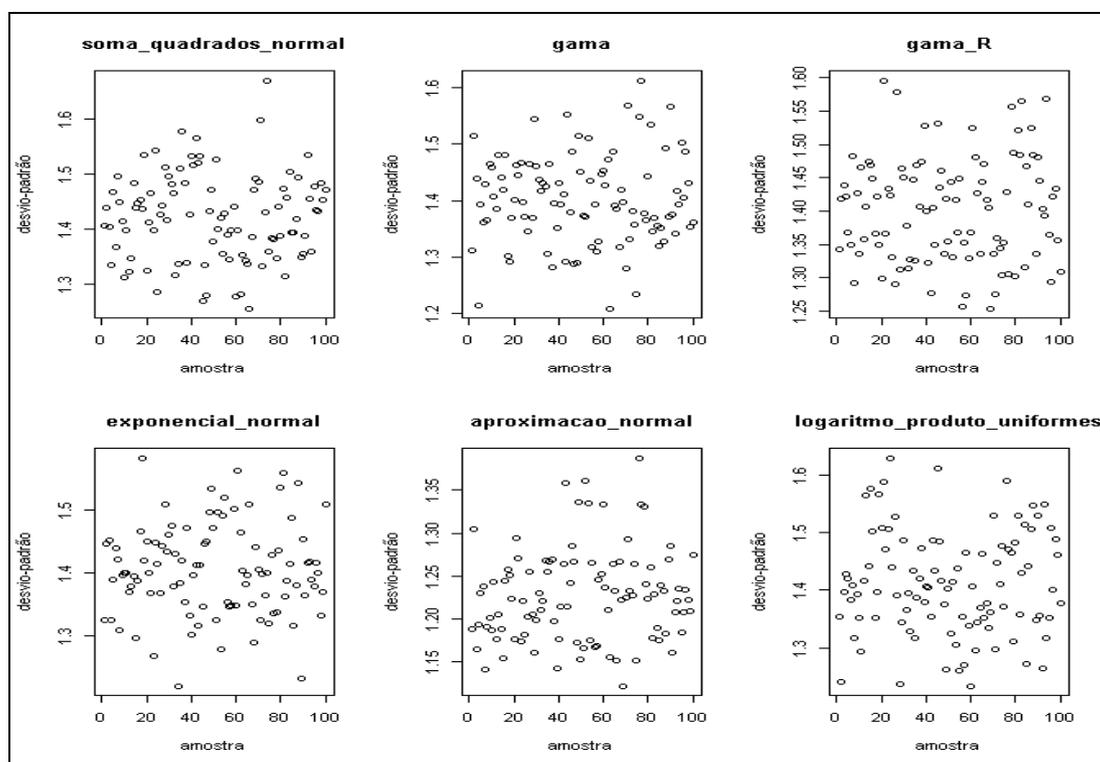
Fonte: elaborada pelo autor.

8.2.4 Desvio-Padrão

Tabela 59 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição Qui-Quadrado em 100 amostras de tamanho 1000

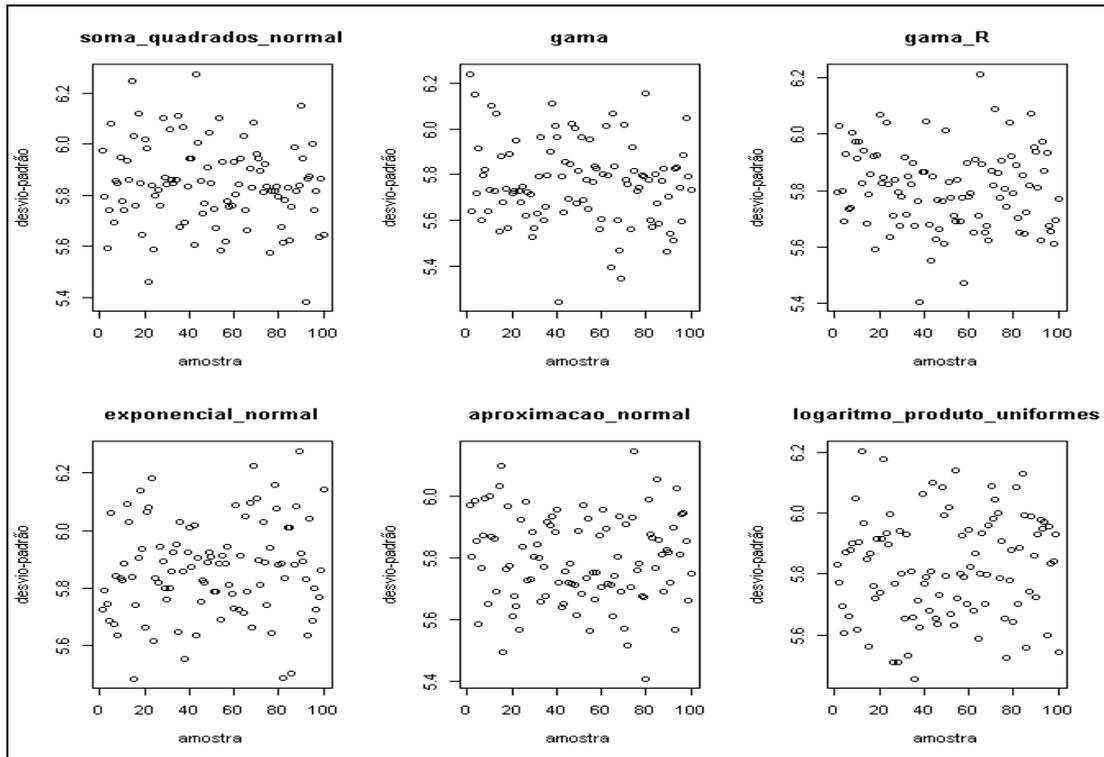
Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.421924	0.07896743	5.842718	0.1609079	10.00735	0.232679
Gama	1.403045	0.07880162	5.765446	0.1805143	9.793368	0.2941288
Gama - R	1.401075	0.07893568	5.805701	0.1409127	9.965562	0.202802
Exponencial-Normal	1.404111	0.07175905	5.865415	0.1639057	9.957347	0.2265789
Aprox. Normal	1.227181	0.05342568	5.793724	0.1417317	9.981239	0.197087
Log. Produto Uniformes	1.411868	0.09001057	5.823167	0.1687384	9.976004	0.23564

Fonte: elaborada pelo autor.



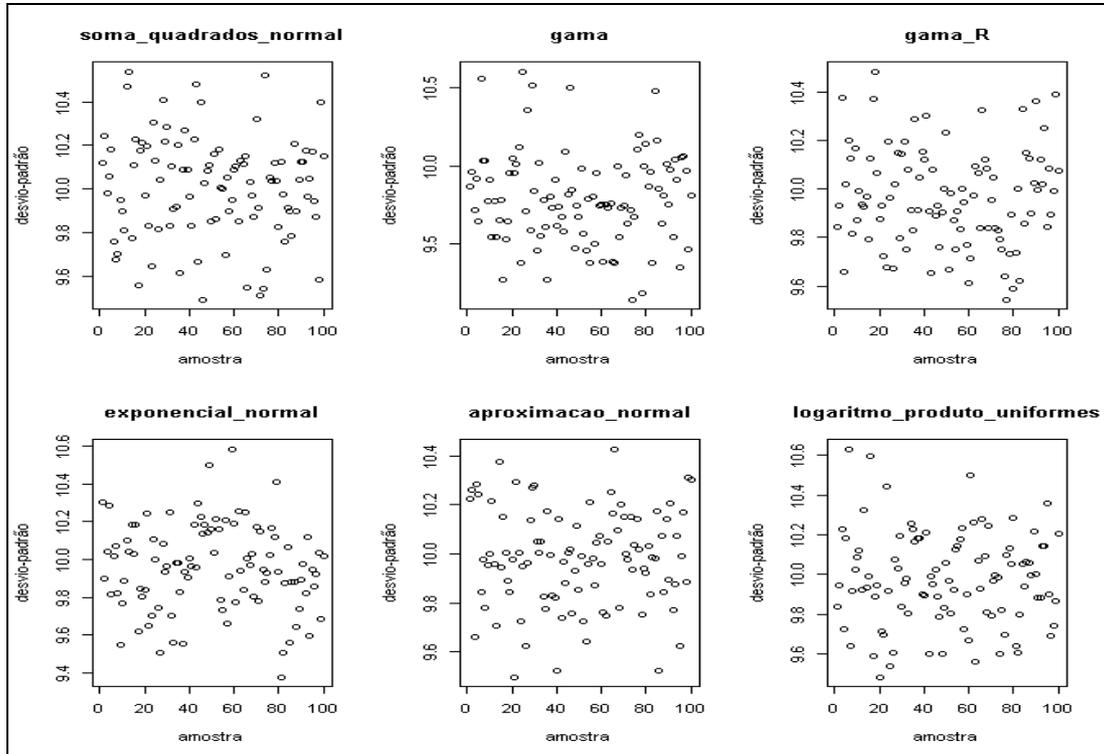
Quadro 26 - Gráficos dos valores estimados para o desvio-padrão da distribuição $\chi^2(1)$ em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.



Quadro 27 - Gráficos dos valores estimados para o desvio-padrão da distribuição χ^2 (17) em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.



Quadro 28 - Gráficos dos valores estimados para o desvio-padrão da distribuição χ^2 (50) em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.

8.2.5 Assimetria

Tabela 60 – Média e erro padrão dos valores estimados para a assimetria da distribuição Qui-Quadrado em 100 amostras de tamanho 1000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	2.814027	0.3554923	0.6913099	0.1127892	0.3831471	0.07486771
Gama	2.802112	0.3345793	1.042082	0.09317144	1.030021	0.09209641
Gama - R	2.770751	0.3533118	0.6827602	0.1187743	0.3803773	0.07723718
Exponencial-Normal	2.75016	0.3854165	0.6814897	0.1017661	0.3858478	0.09084073
Aprox. Normal	2.144612	0.2466559	0.5206336	0.09637081	0.3067995	0.07292042
Log. Produto Uniformes	2.770046	0.4374848	0.6829577	0.1227427	0.4034304	0.08992192

Fonte: elaborada pelo autor.

8.2.6 Curtose

Tabela 61 – Média e erro padrão dos valores estimados para a curtose da distribuição Qui-Quadrado em 100 amostras de tamanho 1000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	11.5264	4.472241	0.7263589	0.4782312	0.2092214	0.2183491
Gama	11.50538	4.177813	1.641141	0.434104	1.8229	0.3745492
Gama - R	11.01609	4.109657	0.7364455	0.6489397	0.1884587	0.2110019
Exponencial-Normal	10.99213	4.756035	0.7001752	0.4512081	0.2107885	0.2642397
Aprox. Normal	6.344874	2.533877	0.3913783	0.3394026	0.1104477	0.1972562
Log. Produto Uniformes	11.29971	5.371717	0.7250977	0.6046086	0.2411044	0.330943

Fonte: elaborada pelo autor.

8.2.7 Quartis

Tabela 62 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 1000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	0.1044668	0.01048598	12.78524	0.2060105	42.98451	0.3711625
Gama	0.1034144	0.01247105	12.01995	0.1962964	40.47112	0.3329982
Gama - R	0.103179	0.01108255	12.80098	0.1826826	42.92437	0.3712551
Exponencial-Normal	0.1027504	0.01187885	12.80451	0.2069324	43.00127	0.422737
Aprox. Normal	0.1324968	0.01210669	12.85337	0.2097914	43.02181	0.3798801
Log. Produto Uniformes	0.1022963	0.01049362	12.77966	0.2088025	42.93264	0.3756115

Fonte: elaborada pelo autor.

Tabela 63 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 1000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	0.4586327	0.03242938	16.31935	0.2154625	49.39096	0.4136145
Gama	0.4572804	0.03691209	15.0175	0.1725422	45.49367	0.2647553
Gama - R	0.4529179	0.0325411	16.34747	0.2266745	49.36931	0.3623708
Exponencial-Normal	0.458457	0.0365063	16.39154	0.2266728	49.37669	0.4300503
Aprox. Normal	0.5503767	0.03598974	16.4969	0.2278714	49.45569	0.3661811
Log. Produto Uniformes	0.4539841	0.03486632	16.32471	0.21808	49.31456	0.4072877

Fonte: elaborada pelo autor.

Tabela 64 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 1000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.321603	0.07218033	20.47862	0.2730428	56.38897	0.490837
Gama	1.321135	0.0835869	18.92503	0.247618	51.17745	0.4156423
Gama - R	1.309473	0.08215092	20.47525	0.3105727	56.34308	0.4461338
Exponencial-Normal	1.317646	0.07683617	20.5404	0.2841585	56.32336	0.5502078
Aprox. Normal	1.422365	0.076653	20.57523	0.26167	56.42925	0.4625401
Log. Produto Uniformes	1.312055	0.06768685	20.47264	0.2698355	56.27865	0.4544601

Fonte: elaborada pelo autor.

8.2.8 Distância Interquartilica

Tabela 65 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição Qui-Quadrado em 100 amostras de tamanho 1000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.217136	0.06944868	7.693384	0.2935874	13.40446	0.4349357
Gama	1.217721	0.07966369	6.905073	0.285838	10.70633	0.4500149
Gama - R	1.206294	0.08000485	7.674271	0.2974517	13.41871	0.4629778
Exponencial-Normal	1.214896	0.07352935	7.735889	0.2964563	13.32209	0.5389205
Aprox. Normal	1.289868	0.07402771	7.721859	0.2852833	13.40744	0.4849835
Log. Produto Uniformes	1.209759	0.06552583	7.69298	0.3208282	13.34601	0.4798559

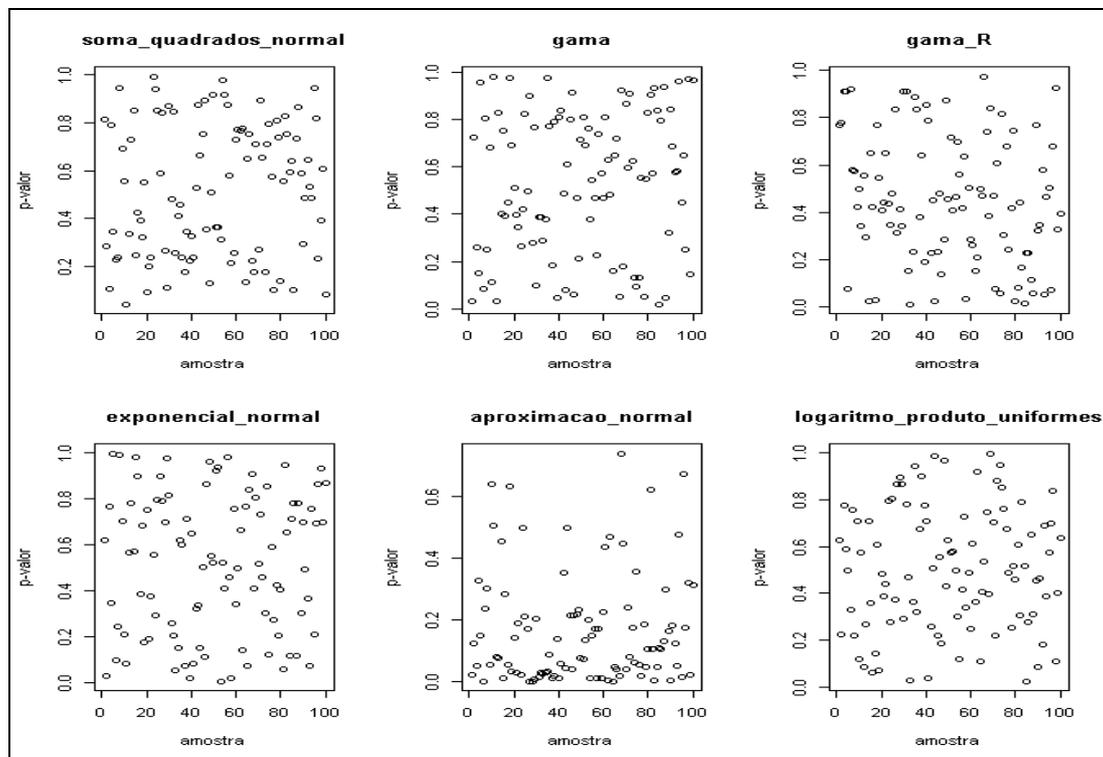
Fonte: elaborada pelo autor.

8.2.9 Teste de Aderência

Tabela 66 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição Qui-Quadrado em 100 amostras de tamanho 1000

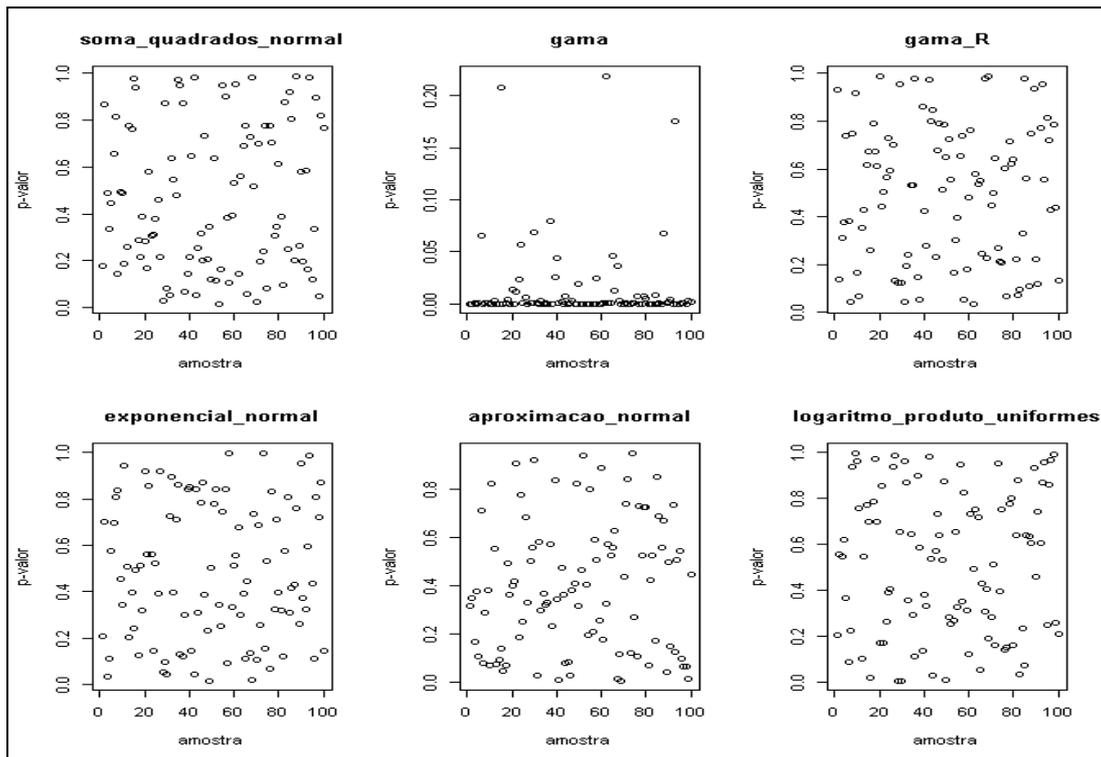
Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	0.5209141	0.2756987	0.4677519	0.3056123	0.5550917	0.2614928
Gama	0.5318049	0.2993723	0.01286402	0.03721403	2.981248e-12	2.123387e-11
Gama - R	0.4521463	0.2712331	0.497528	0.2884394	0.528615	0.2874854
Exponencial-Normal	0.5146178	0.3026659	0.4809916	0.2932923	0.5107994	0.288806
Aprox. Normal	0.164488	0.1778481	0.3994259	0.2695044	0.4792859	0.2927765
Log. Produto Uniformes	0.5008267	0.2604977	0.5143086	0.3075101	0.4949129	0.3209726

Fonte: elaborada pelo autor.



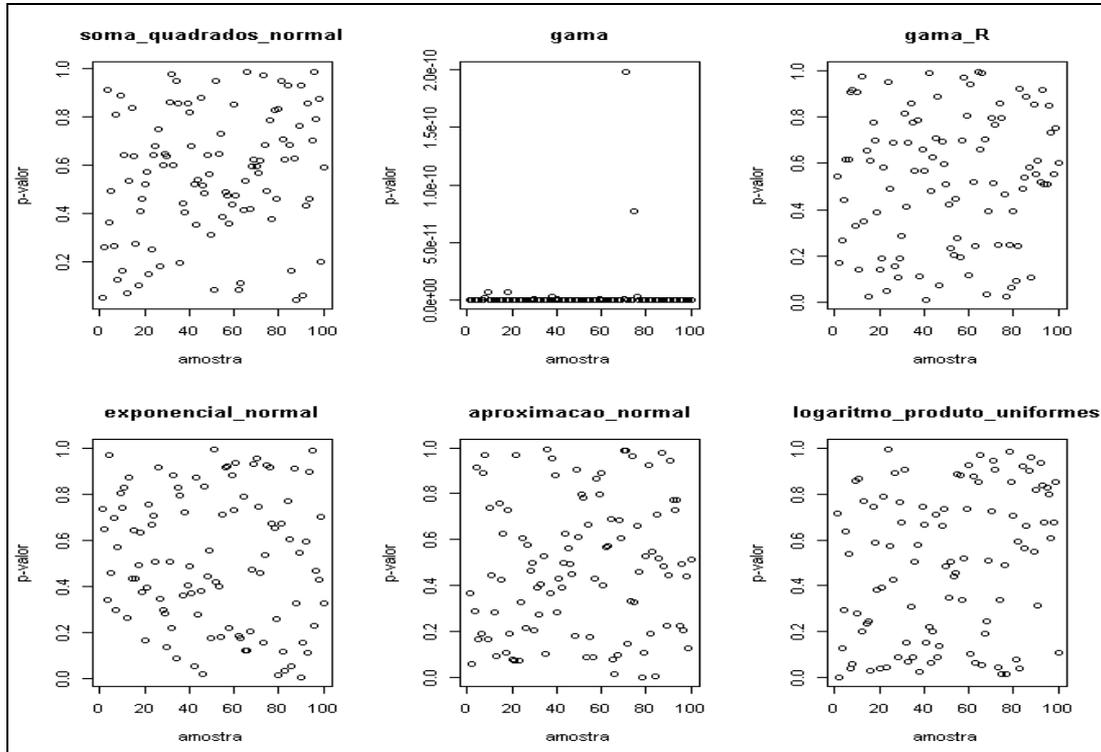
Quadro 29 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2(1)$ em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.



Quadro 30 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2(17)$ em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.



Quadro 31 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2(50)$ em 100 amostras de tamanho 1000

Fonte: Elaborado pelo autor.

8.2.10 Conclusões

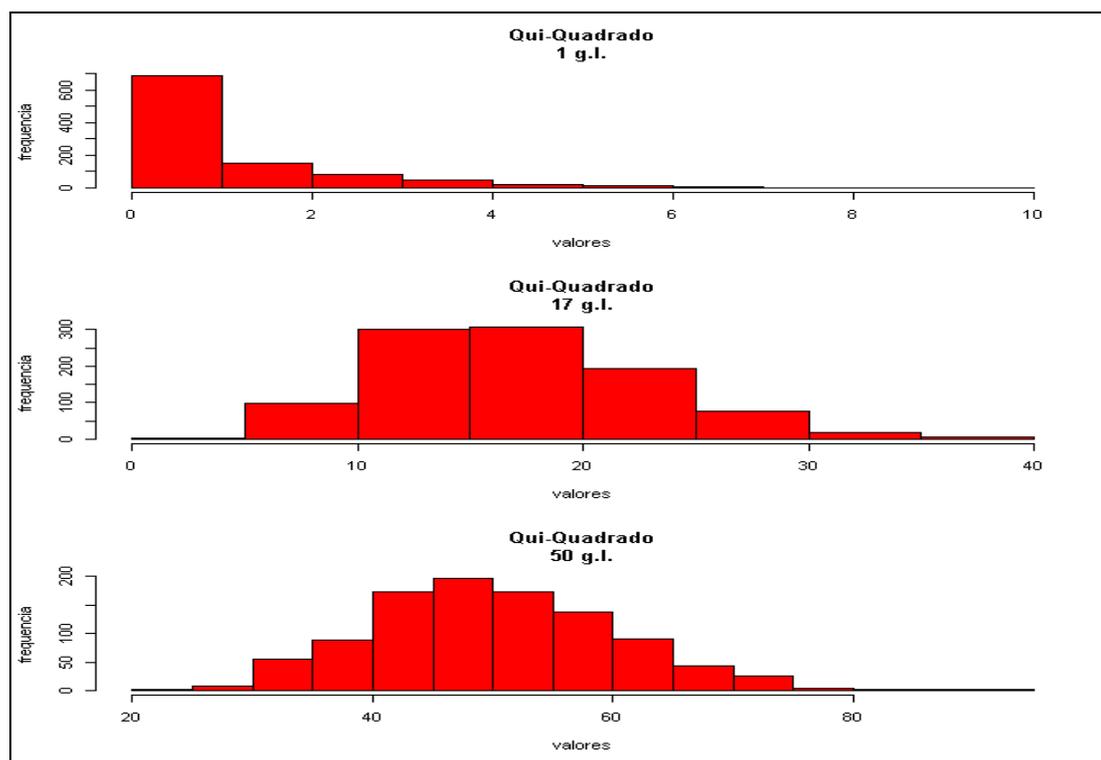
Assim como nas amostras com tamanho 100, mais uma vez, para as amostras da Distribuição Qui-Quadrado com tamanho 1000, tivemos um resultado equilibrado, com destaque para os algoritmos Soma Quadrados Normal, Logaritmo Produto Uniformes e Gama-R, os quais, em média, se mostraram mais precisos quanto às suas estimativas em relação aos demais. Dentre esses geradores, o algoritmo Soma Quadrados Normal mostrou ser, em média, o mais preciso de todos, pois quase todas as suas estimativas chegarão bem próximas aos valores exatos. Deste modo, indicamos o algoritmo Soma Quadrados Normal para gerar amostras de tamanho $n = 1000$ da Distribuição Qui-Quadrado.

Ao contrário das amostras anteriores, desta vez podemos notar que um algoritmo não passou no Teste Qui-Quadrado de Aderência: o gerador formado pela combinação de algoritmos da Gama. Assim, podemos dizer que o uso desse gerador não é aconselhável para geração de amostras com tamanho 1000 da Distribuição Qui-Quadrado.

Tabela 67 – Estimativas médias do algoritmo Soma Quadrados Normal em 100 amostras de tamanho 1000 da Distribuição Qui-Quadrado

Medidas	Soma Quadrados Normal		
	$\chi^2(1)$	$\chi^2(17)$	$\chi^2(50)$
Esperança	1.00424	16.98659	50.04077
Máximo	12.03881	42.49086	88.90289
Mínimo	3.379198e-06	4.087693	23.5977
Amplitude	12.03881	38.40316	65.3052
Desvio-Padrão	1.421924	5.842718	10.00735
Assimetria	2.814027	0.6913099	0.3831471
Curtose	11.5264	0.7263589	0.2092214
1º Quartil	0.1044668	12.78524	42.98451
2º Quartil	0.4586327	16.31935	49.39096
3º Quartil	1.321603	20.47862	56.38897
Dist. Interquart.	1.217136	7.693384	13.40446

Fonte: elaborada pelo autor.



Quadro 32 – Histogramas de 1000 valores gerados da Distribuição Qui-Quadrado pelo algoritmo Soma Quadrados Normal

Fonte: Elaborado pelo autor.

8.3. Estimativas com amostras de tamanho $n = 10000$

Nas tabelas abaixo, apresentamos a média e o erro padrão dos valores estimados pelos geradores para cada uma das medidas analisadas. Nos gráficos, ilustramos o comportamento das estimativas dos geradores ao longo das 100 amostras.

8.3.1 Tempo de Execução

Tabela 68 – Média e erro padrão dos tempos de execução, em segundos, medidos em 100 amostras de tamanho 10000 da Distribuição Qui-Quadrado

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.814	0.02670452	23.1541	0.09879164	66.0115	0.2371181
Gama	0.8261	0.02064099	1.5126	0.02517093	1.6785	0.01799972
Gama - R	1.9393	0.02383678	2.616	0.02344131	2.5761	0.02173741
Exponencial-Normal	1.8834	0.02745869	4.9227	0.03209377	8.087	0.06102657
Aprox. Normal	1.8113	0.02623341	1.8144	0.01913904	1.8084	0.01292715
Log. Produto Uniformes	2.0269	0.02232429	2.1376	0.02151438	0.4278	0.008358145

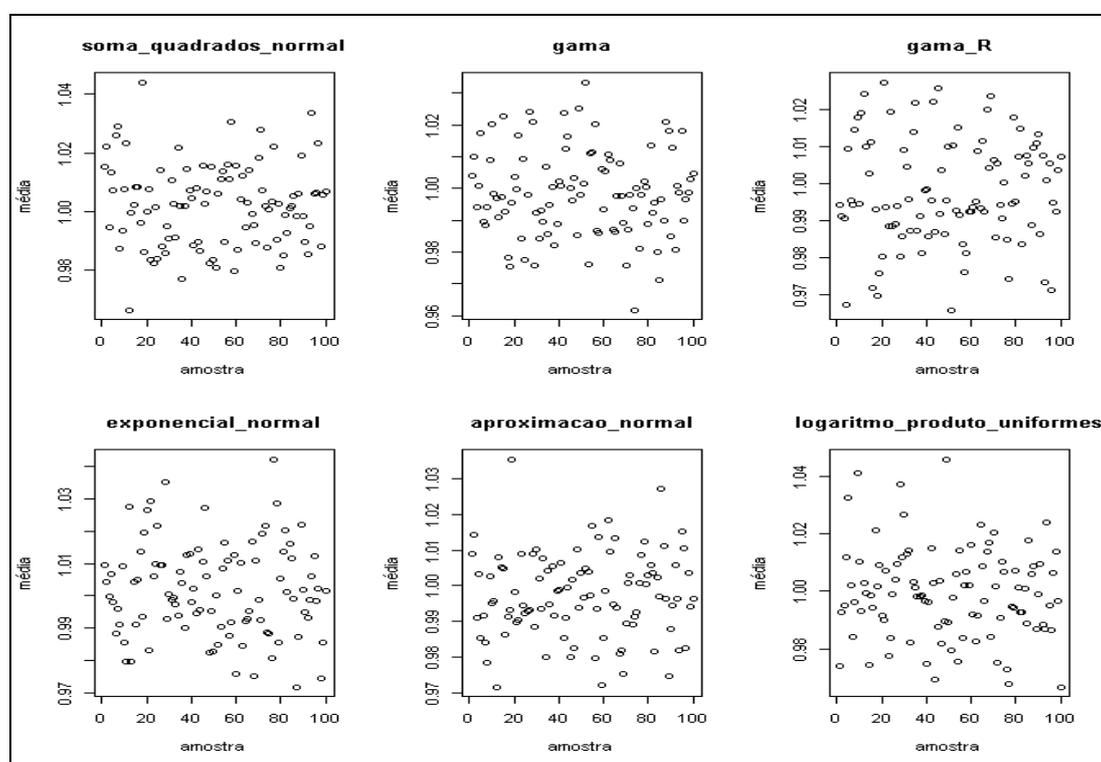
Fonte: elaborada pelo autor.

8.3.2 Esperança

Tabela 69 - Média e erro padrão dos valores estimados para a esperança da distribuição Qui-Quadrado em 100 amostras de tamanho 10000

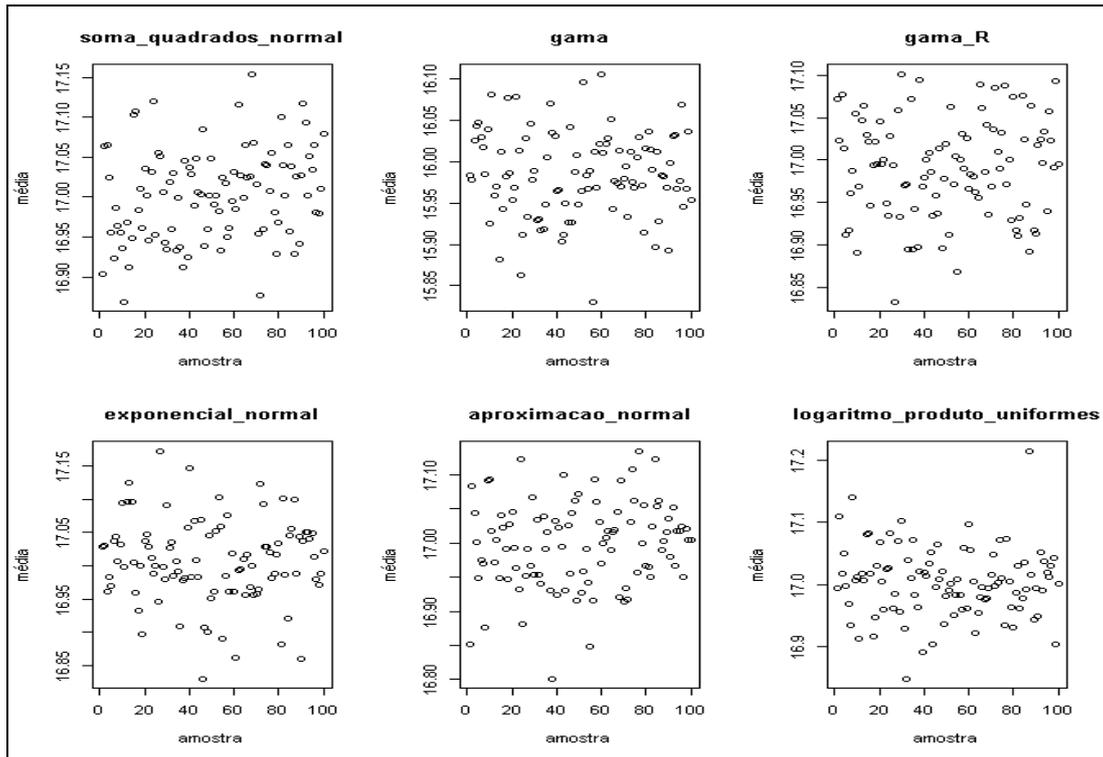
Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados						
Normal	1.002265	0.01406488	17.00356	0.05717654	50.0073	0.111365
Gama	0.9984107	0.01367806	15.9847	0.05167152	46.79615	0.0915945
Gama - R	0.9976756	0.01428686	16.99084	0.05833432	50.01021	0.09790728
Exponencial-Normal	1.00177	0.01425757	17.00686	0.06209903	50.01576	0.1077557
Aprox. Normal	0.9973358	0.01173986	16.99711	0.06244536	50.00073	0.1030472
Log. Produto Uniformes	0.9993951	0.01561921	17.00386	0.05546464	49.99514	0.1113688

Fonte: elaborada pelo autor.



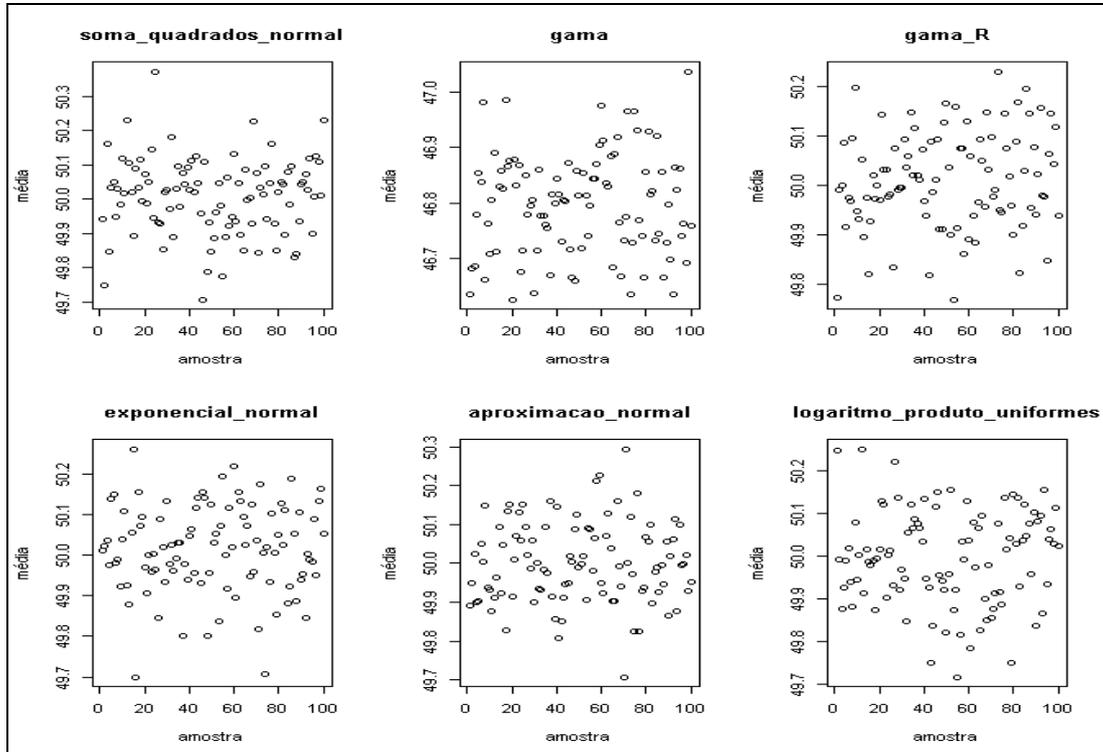
Quadro 33 - Gráficos dos valores estimados para a esperança da distribuição $\chi^2(1)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 34 - Gráficos dos valores estimados para a esperança da distribuição $\chi^2(17)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 35 - Gráficos dos valores estimados para a esperança da distribuição $\chi^2(50)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

8.3.3 Máximo e Mínimo

Tabela 70 - Média e erro padrão dos valores estimados para o máximo da distribuição Qui-Quadrado em 100 amostras de tamanho 10000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	16.48934	2.29293	48.737	3.271737	97.0787	4.485189
Gama	16.48934	2.396411	49.55455	3.464171	99.73896	4.441639
Gama - R	15.77967	2.183475	49.04764	3.50109	97.76506	4.615915
Exponencial-Normal	15.76974	1.959954	49.01826	3.281991	98.05577	4.423998
Aprox. Normal	12.03783	1.989972	46.35482	2.741745	95.44623	4.938173
Log. Produto Uniformes	15.69272	1.97937	48.75826	3.49013	98.47288	5.483772

Fonte: elaborada pelo autor.

Tabela 71 - Média e erro padrão dos valores estimados para o mínimo da distribuição Qui-Quadrado em 100 amostras de tamanho 10000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	2.862828e-08	6.439318e-08	3.003431	0.5382061	20.54746	1.48709
Gama	2.029866e-08	3.439935e-08	2.848593	0.4449955	20.05564	1.4333
Gama - R	2.437269e-08	4.134444e-08	2.931265	0.5084808	20.29962	1.779706
Exponencial-Normal	2.348853e-08	4.500242e-08	2.946182	0.4527233	20.38461	1.625204
Aprox. Normal	4.351029e-08	8.783588e-08	1.904909	0.461565	18.62333	1.919766
Log. Produto Uniformes	3.791545e-08	9.452867e-08	2.903222	0.4583714	20.26121	1.755344

Fonte: elaborada pelo autor.

8.3.4 Amplitude

Tabela 72 - Média e erro padrão dos valores estimados para a amplitude da distribuição Qui-Quadrado em 100 amostras de tamanho 10000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	16.48934	2.29293	45.73357	3.299666	76.53124	4.767536
Gama	16.48934	2.396411	46.70596	3.443595	79.68332	4.525931
Gama - R	15.77967	2.183475	46.11637	3.468019	77.46544	4.897227
Exponencial-Normal	15.76974	1.959954	46.07208	3.194558	77.67116	4.959216
Aprox. Normal	12.03783	1.989972	44.44991	2.787556	76.8229	5.270831
Log. Produto Uniformes	15.69272	1.97937	45.85503	3.494796	78.21167	5.641993

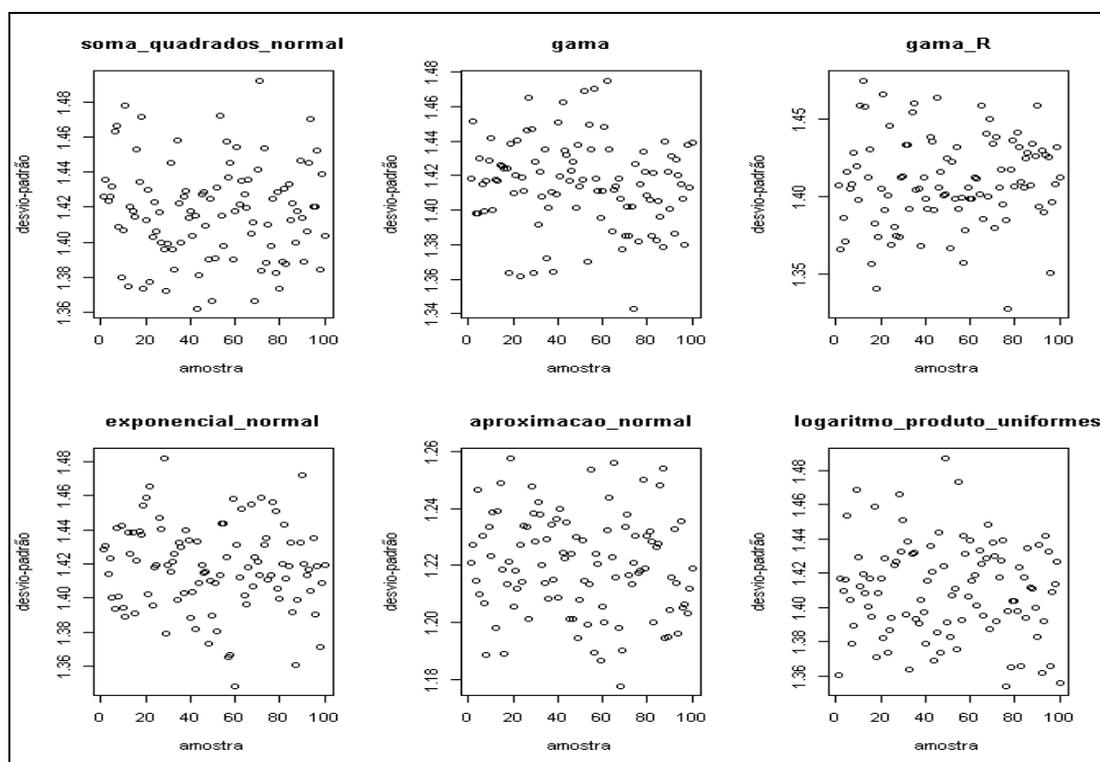
Fonte: elaborada pelo autor.

8.3.5 Desvio-Padrão

Tabela 73 - Média e erro padrão dos valores estimados para o desvio-padrão da distribuição Qui-Quadrado em 100 amostras de tamanho 10000

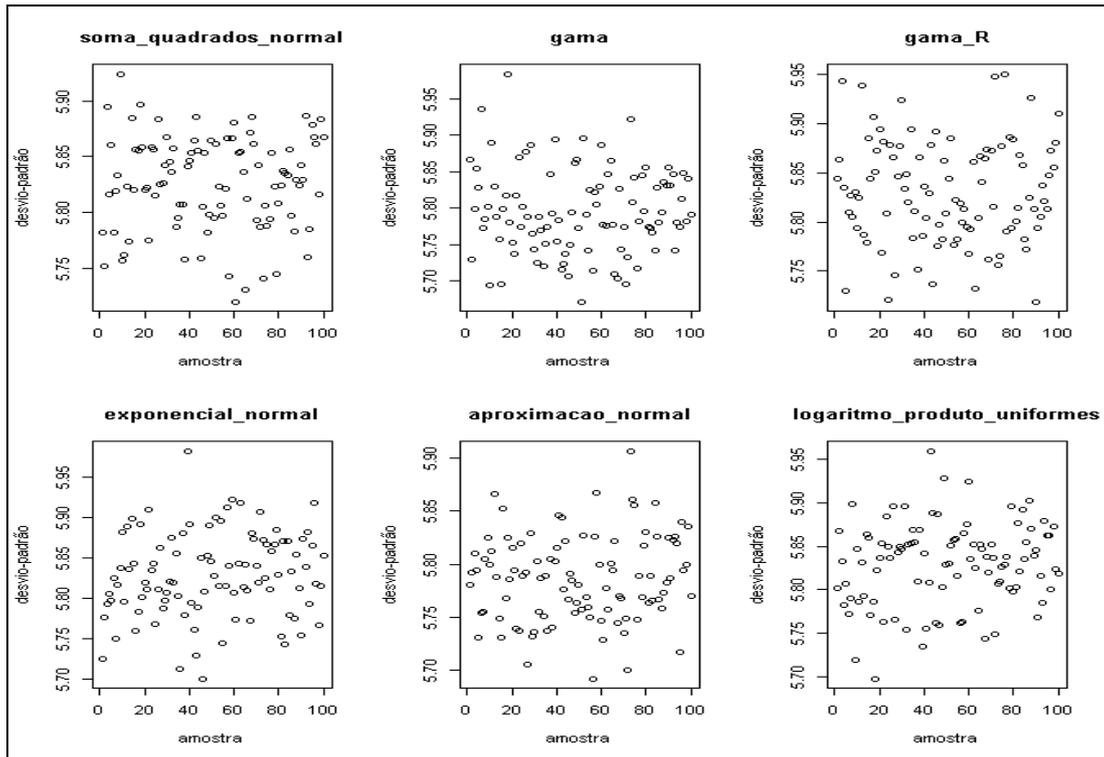
Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.41723	0.02725676	5.826287	0.04198251	9.98813	0.0715145
Gama	1.414996	0.02531607	5.795472	0.05859026	9.800763	0.0936403
Gama - R	1.409579	0.02880342	5.830417	0.05258104	10.00712	0.0744572
Exponencial-Normal	1.418003	0.02501705	5.82827	0.05226014	10.00512	0.08035139
Aprox. Normal	1.220302	0.01743595	5.786366	0.04078456	9.976783	0.07187497
Log. Produto Uniformes	1.410394	0.02759389	5.829266	0.04762828	10.00103	0.08075088

Fonte: elaborada pelo autor.



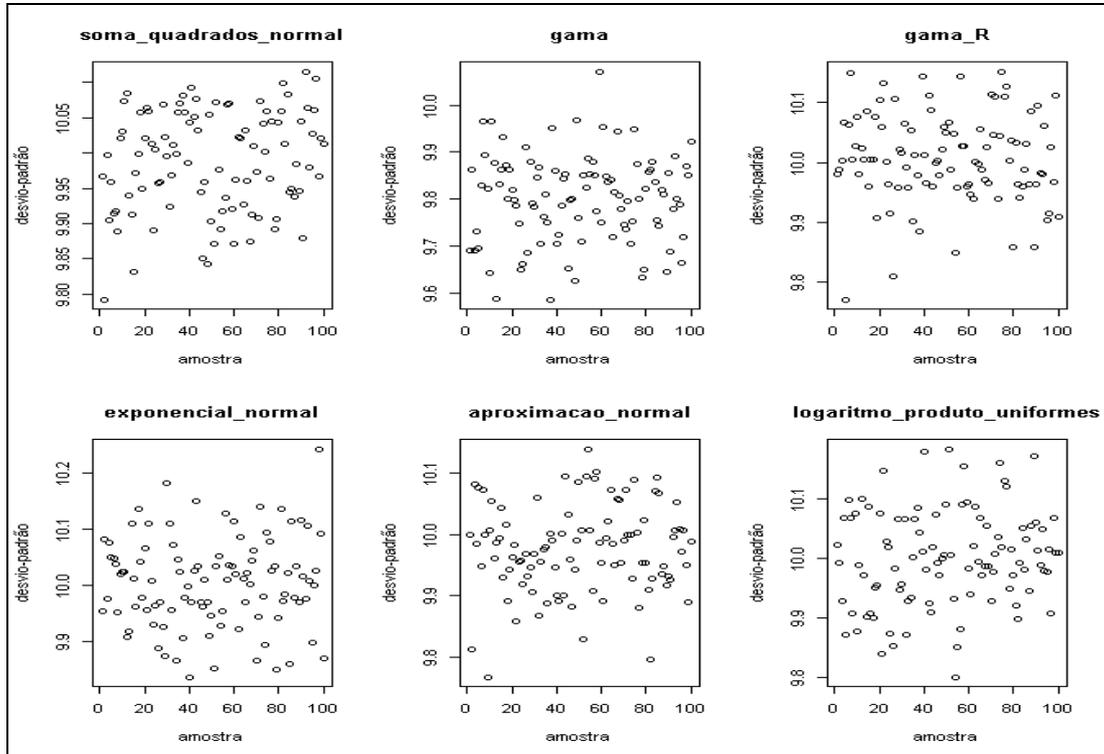
Quadro 36 - Gráficos dos valores estimados para o desvio-padrão da distribuição $\chi^2(1)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 37 - Gráficos dos valores estimados para o desvio-padrão da distribuição $\chi^2(17)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 38 - Gráficos dos valores estimados para o desvio-padrão da distribuição $\chi^2(50)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

8.3.6 Assimetria

Tabela 74 – Média e erro padrão dos valores estimados para a assimetria da distribuição Qui-Quadrado em 100 amostras de tamanho 10000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	2.831663	0.1290281	0.6801614	0.02934902	0.3943194	0.02875119
Gama	2.835818	0.1406696	1.065714	0.03800009	1.028579	0.03474946
Gama - R	2.803473	0.1205338	0.691324	0.03239232	0.4004373	0.02608753
Exponencial-Normal	2.822748	0.1325541	0.6813965	0.03042899	0.4016406	0.02365893
Aprox. Normal	2.169473	0.09495861	0.5192819	0.02860079	0.299735	0.02861736
Log. Produto Uniformes	2.78456	0.1086489	0.6846554	0.03213571	0.4029397	0.02797536

Fonte: elaborada pelo autor.

8.3.7 Curtose

Tabela 75 – Média e erro padrão dos valores estimados para a curtose da distribuição Qui-Quadrado em 100 amostras de tamanho 10000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	12.05006	1.726341	0.6774369	0.1326606	0.2155453	0.08629013
Gama	12.07778	1.950514	1.740841	0.1818896	1.833498	0.1386318
Gama - R	11.61204	1.618696	0.7284144	0.14788	0.2279506	0.07658477
Exponencial-Normal	11.75761	1.735673	0.6896507	0.1320839	0.2420802	0.07633113
Aprox. Normal	6.665851	1.143403	0.3658851	0.1006504	0.1111568	0.07213606
Log. Produto Uniformes	11.36261	1.357967	0.6990607	0.1477815	0.2503199	0.092192

Fonte: elaborada pelo autor.

8.3.8 Quartis

Tabela 76 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 10000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	0.1019406	0.003206387	12.79862	0.06730316	42.94238	0.104047
Gama	0.1009889	0.003995372	12.01681	0.05640511	40.4725	0.1024506
Gama - R	0.1014098	0.003338783	12.78128	0.06230221	42.94907	0.1284452
Exponencial-Normal	0.101537	0.003441234	12.79932	0.06255242	42.95145	0.1155342
Aprox. Normal	0.1337626	0.004240343	12.84726	0.06785608	43.00981	0.1383088
Log. Produto Uniformes	0.1012789	0.003371071	12.79817	0.05906082	42.93491	0.137136

Fonte: elaborada pelo autor.

Tabela 77 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 10000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	0.4561991	0.01019861	16.34726	0.07085798	49.33652	0.1273793
Gama	0.4525091	0.01079801	15.0019	0.05955307	45.48062	0.09431326
Gama - R	0.45295	0.008764726	16.32983	0.06063126	49.33492	0.115396
Exponencial-Normal	0.4564155	0.01096569	16.34507	0.07528235	49.35125	0.1243082
Aprox. Normal	0.5498407	0.01167911	16.49199	0.08117676	49.50313	0.1169803
Log. Produto Uniformes	0.454777	0.01150877	16.34436	0.06931472	49.33172	0.1325582

Fonte: elaborada pelo autor.

Tabela 78 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição Qui-Quadrado em 100 amostras de tamanho 10000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.32616	0.02301855	20.49486	0.0769177	56.34428	0.1726146
Gama	1.321838	0.02328378	18.95338	0.08688538	51.12572	0.1093637
Gama - R	1.32086	0.02039589	20.46898	0.07906834	56.34236	0.1443965
Exponencial-Normal	1.324601	0.02496079	20.49989	0.08904374	56.35018	0.1669024
Aprox. Normal	1.417413	0.02210073	20.60256	0.08808619	56.43978	0.148042
Log. Produto Uniformes	1.32381	0.02421657	20.48216	0.07787123	56.31924	0.155078

Fonte: elaborada pelo autor.

8.3.9 Distância Interquartilica

Tabela 79 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição Qui-Quadrado em 100 amostras de tamanho 10000

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	1.224219	0.02265025	7.696245	0.07397041	13.4019	0.1667734
Gama	1.220849	0.02199399	6.936568	0.09805816	10.65322	0.1294773
Gama - R	1.219451	0.02060459	7.687705	0.07834438	13.39329	0.1713982
Exponencial-Normal	1.223064	0.02374958	7.700567	0.09264599	13.39873	0.1669796
Aprox. Normal	1.283651	0.02088047	7.755296	0.08197058	13.42997	0.1590754
Log. Produto Uniformes	1.222531	0.02289878	7.683998	0.07770908	13.38432	0.1555941

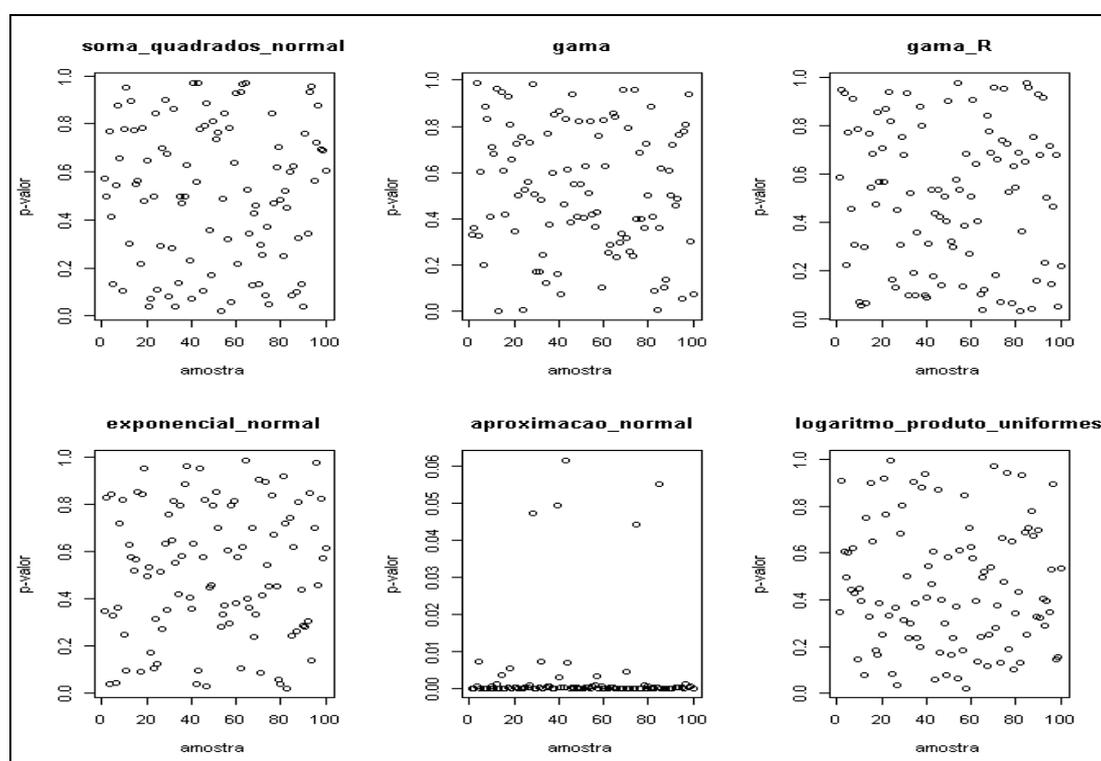
Fonte: elaborada pelo autor.

8.3.10 Teste de Aderência

Tabela 80 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição Qui-Quadrado em 100 amostras de tamanho 10000

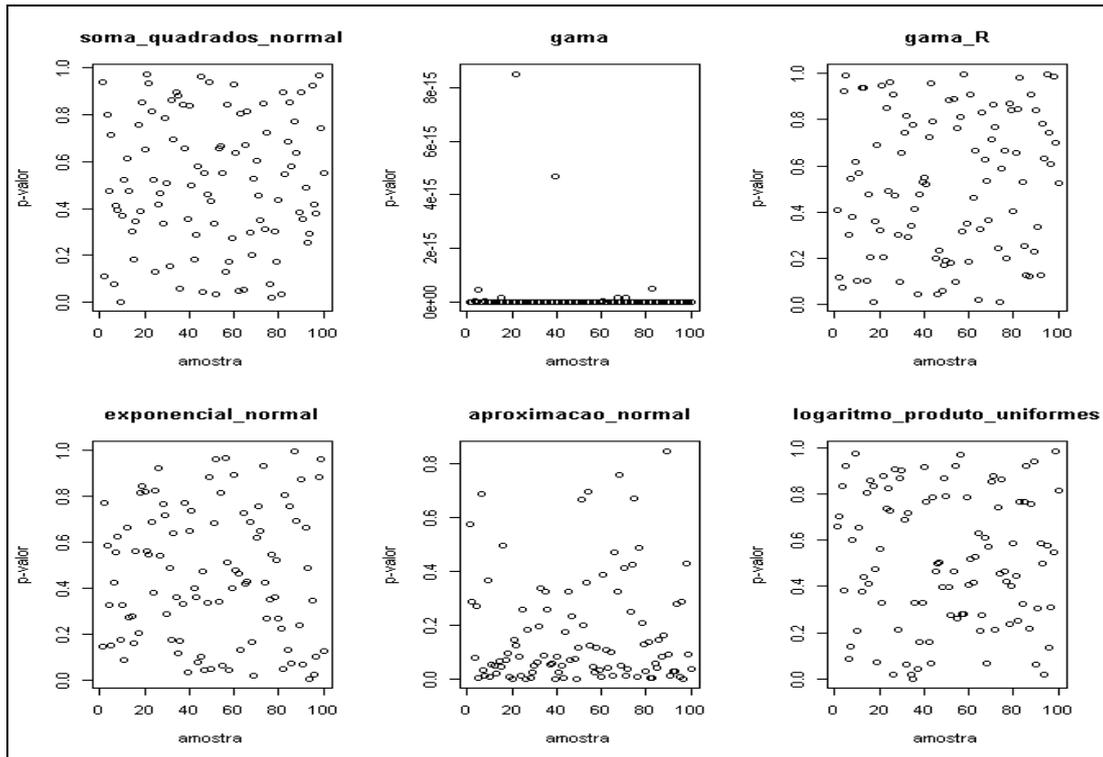
Gerador	$\chi^2_{(1)}$		$\chi^2_{(17)}$		$\chi^2_{(50)}$	
	Média	E.P.	Média	E.P.	Média	E.P.
Soma Quadrados Normal	0.5097455	0.2941802	0.510423	0.2811355	0.4561825	0.2963599
Gama	0.5249437	0.2748803	1.467026e-16	9.690749e-16	9.250326e-130	8.415125e-129
Gama - R	0.5056056	0.298097	0.5216411	0.3042426	0.4776885	0.2738906
Exponencial-Normal	0.5106824	0.2780927	0.4556484	0.2872756	0.4864259	0.2822167
Aprox. Normal	0.003118326	0.01133058	0.1656184	0.1977803	0.3720583	0.2921778
Log. Produto Uniformes	0.4538806	0.2645571	0.5115529	0.2865802	0.5127495	0.3128424

Fonte: elaborada pelo autor.



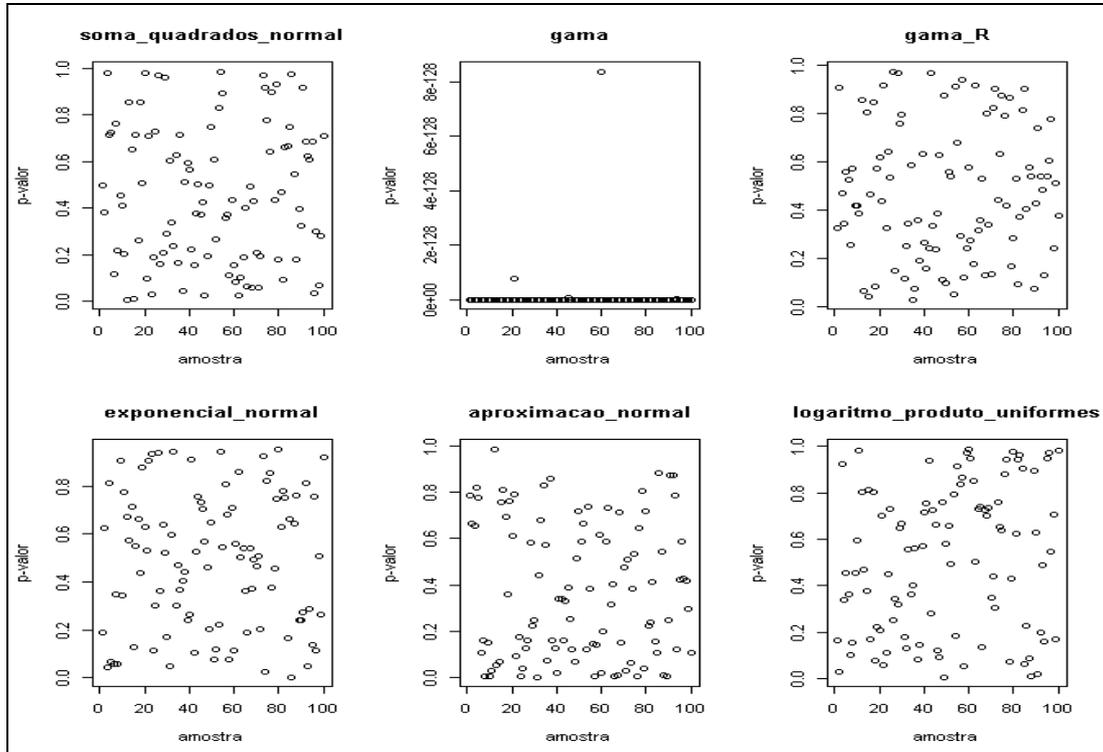
Quadro 39 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2_{(1)}$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 40 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2(17)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 41 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $\chi^2(50)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

8.3.11 Conclusões

Novamente, para as amostras com tamanho 10000 da Distribuição Qui-Quadrado, tivemos um resultado equilibrado, com destaque para os algoritmos Soma Quadrados Normal, Exponencial Normal e Logaritmo Produto Uniformes. O gerador Logaritmo Produto Uniformes, mais uma vez, mostrou-se um pouco mais preciso que os demais, e assim o indicamos para gerar amostras com tamanho $n = 10000$ da Distribuição Qui-Quadrado.

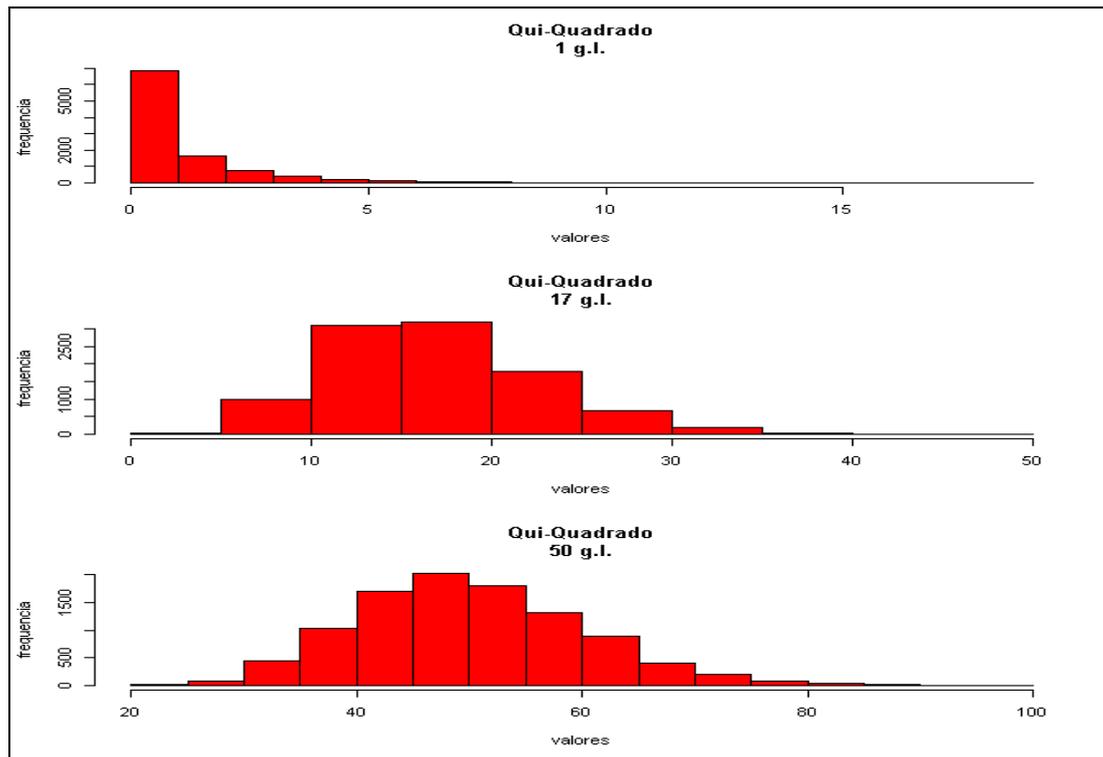
Podemos notar que, à medida que aumentamos o número de graus de liberdade da distribuição, o tempo de processamento do algoritmo Soma Quadrados Normal cresce substancialmente, chegando a ser dez vezes maior que o tempo dos demais algoritmos.

Também constata-se que dois algoritmos não passaram no Teste Qui-Quadrado de Aderência: Aproximação Normal e Gama. Sendo assim, não se recomenda o uso desses algoritmos como geradores de amostras da Distribuição Qui-Quadrado com tamanho 10000.

Tabela 81 – Estimativas médias do algoritmo Logaritmo Produto Uniformes em 100 amostras de tamanho 10000 da Distribuição Qui-Quadrado

Medidas	Logaritmo Produto Uniformes		
	$\chi^2(1)$	$\chi^2(17)$	$\chi^2(50)$
Tempo de Execução	2.0269 s	2.1376 s	0.4278 s
Esperança	0.9993951	17.00386	49.99514
Máximo	15.69272	48.75826	98.47288
Mínimo	3.791545e-08	2.903222	20.26121
Amplitude	15.69272	45.85503	78.21167
Desvio-Padrão	1.410394	5.829266	10.00103
Assimetria	2.78456	0.6846554	0.4029397
Curtose	11.36261	0.6990607	0.2503199
1º Quartil	0.1012789	12.79817	42.93491
2º Quartil	0.454777	16.34436	49.33172
3º Quartil	1.32381	20.48216	56.31924
Dist. Interquart.	1.222531	7.683998	13.38432

Fonte: elaborada pelo autor.



Quadro 42 – Histogramas de 10000 valores gerados da Distribuição Qui-Quadrado pelo algoritmo Logaritmo Produto Uniformes

Fonte: Elaborado pelo autor.

9. GERADORES DE NÚMEROS ALEATÓRIOS DA DISTRIBUIÇÃO F

Neste capítulo, mostraremos dois algoritmos para geração de números aleatórios da Distribuição F. Assim como os algoritmos da Distribuição Qui-Quadrado, esses dois algoritmos também são dependentes de geradores de outras distribuições de probabilidade. No caso da Distribuição F, seus algoritmos são dependentes de bons geradores da Distribuição Qui-Quadrado ou da Distribuição Beta.

9.1 Gerador baseado na Distribuição Qui-Quadrado

Este é o gerador disponível por padrão no software R para geração de ocorrências da Distribuição F.

Gerar $Y_1 \sim \chi^2(\nu_1)$, $Y_2 \sim \chi^2(\nu_2)$.

Sair $X = \frac{Y_1/\nu_1}{Y_2/\nu_2}$.

Os números aleatórios da Distribuição Qui-Quadrado em questão podem ser gerados através de um algoritmo específico. Em nossos estudos, utilizaremos o gerador da Qui-Quadrado disponível no software R, o qual, como já mencionamos anteriormente, é baseado na Distribuição Gama.

9.2 Gerador baseado na Distribuição Beta

Este algoritmo é baseado em uma transformação da Distribuição Beta.

Segundo Gentle (1998), este algoritmo é melhor que o citado anteriormente.

Gerar $Y \sim Beta(\nu_1/2, \nu_2/2)$.

Sair $X = \frac{\nu_2 Y}{\nu_1 (1-Y)}$.

O número aleatório proveniente da Distribuição Beta pode ser gerado através do algoritmo abaixo, o qual é dependente da Distribuição Gama.

Gerar $X_1 \sim Gama(\alpha, 1)$, $X_2 \sim Gama(\beta, 1)$.

Sair $X = X_1 / (X_1 + X_2)$.

10. ESTIMATIVAS DAS AMOSTRAS DA DISTRIBUIÇÃO F

Nos capítulos subsequentes, apresentaremos as estimativas com amostras de tamanho $n = 10000$ da Distribuição F com graus de liberdade m e n iguais, respectivamente, a 1, 5; 5, 1; e 5, 5 para as seguintes medidas: esperança, máximo e mínimo, amplitude, desvio-padrão, quartis, distância interquartilica e valor-p para o teste de aderência. Os valores “exatos” das medidas para a Distribuição F com os graus de liberdade citados acima são, respectivamente: esperança = 1.666667, não definida, 1.666667; máximo = 320.2978, 5764049558, 195.3586 e mínimo = $1.734891e-10$, 0.003122094, 0.005118791; amplitude = 320.2978, 5764049558, 195.3535; desvio-padrão = 4.714045, não definido, 2.981424; 1º quartil = 0.1133813, 0.590853, 0.5277992; 2º quartil = 0.5280738, 1.893675, 1 e 3º quartil = 1.692468, 8.819793, 1.89466; distância interquartilica = 1.579087, 8.22894, 1.366861. Considera-se como máximo o valor x tal que a probabilidade $P(X \leq x) = 0.99999$, e como mínimo o valor x tal que a probabilidade $P(X \leq x) = 0.00001$.

10.1 Estimativas com amostras de tamanho $n = 10000$

Nas tabelas abaixo, apresentamos a média e o erro padrão dos valores estimados pelos geradores para cada uma das medidas analisadas. Nos gráficos, ilustramos o comportamento das estimativas dos geradores ao longo das 100 amostras.

10.1.1 Tempo de Execução

Tabela 82 – Média e erro padrão dos tempos de execução, em segundos, medidos em 100 amostras de tamanho 10000 da Distribuição F

Gerador	F(1,5)		F(5,1)		F(5,5)	
	Média	E.P.	Média	E.P.	Média	E.P.
F pela Beta	5.5949	0.09471376	6.0283	0.1654382	6.7995	0.04710562
F pela Qui-Quadrado	6.0362	0.1441575	6.3813	0.1575613	7.1162	0.05329127

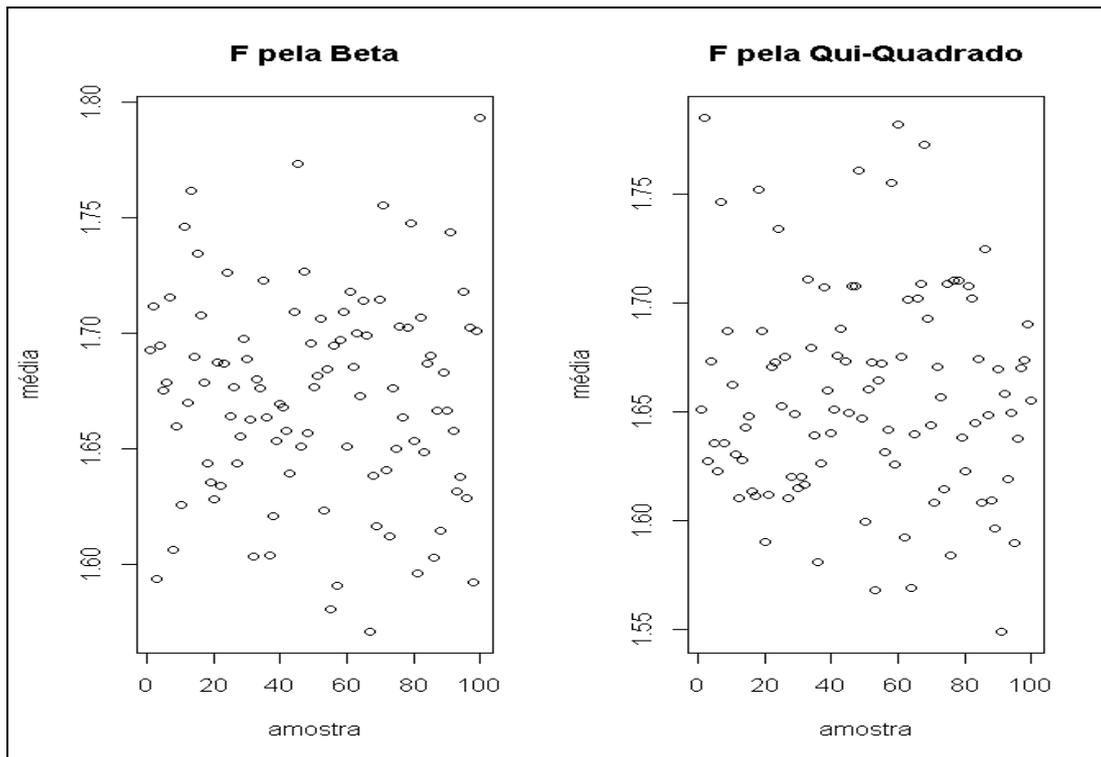
Fonte: elaborada pelo autor.

10.1.2 Esperança

Tabela 83 - Média e erro padrão dos valores estimados para a esperança da distribuição F em 100 amostras de tamanho 10000

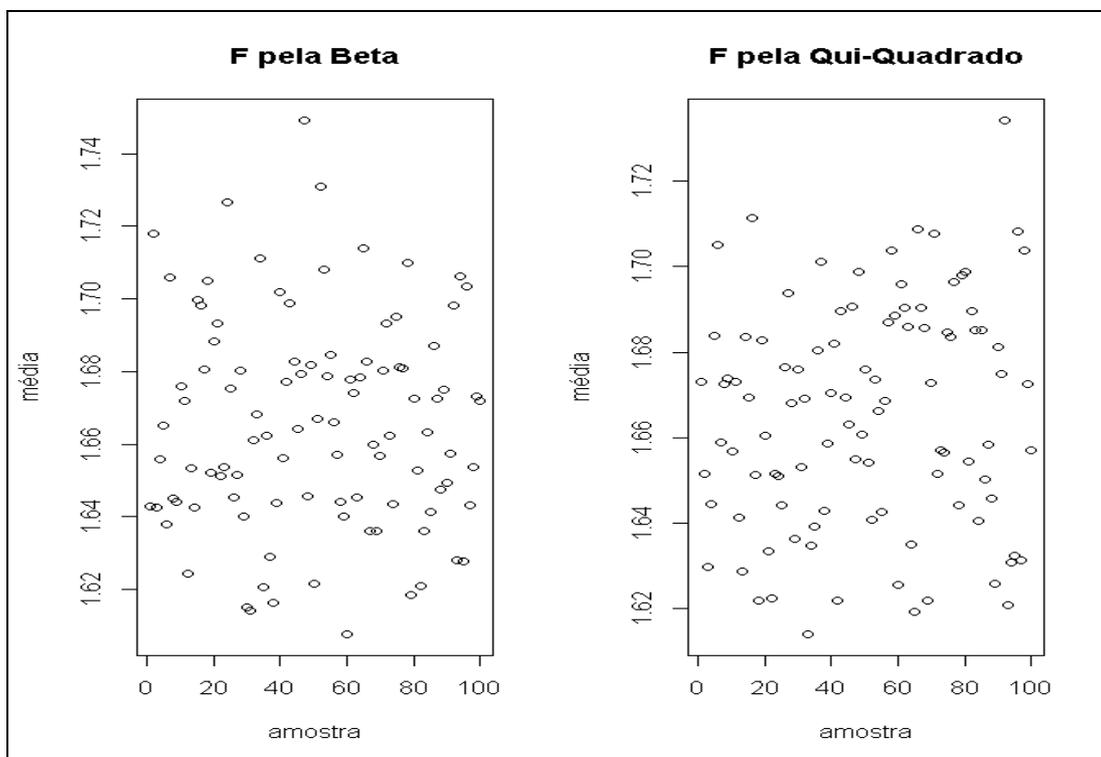
Gerador	F(1,5)		F(5,1)		F(5,5)	
	Média	E.P.	Média	E.P.	Média	E.P.
F pela Beta	1.672302	0.04438049	N.D.	N.D.	1.665627	0.02899394
F pela Qui-Quadrado	1.657483	0.04817021	N.D.	N.D.	1.665134	0.02617753

Fonte: elaborada pelo autor.



Quadro 43 - Gráficos dos valores estimados para a esperança da distribuição $F(1,5)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 44 - Gráficos dos valores estimados para a esperança da distribuição $F(5,5)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

10.1.3 Máximo e Mínimo

Tabela 84 - Média e erro padrão dos valores estimados para o máximo da distribuição F em 100 amostras de tamanho 10000

Gerador	F(1,5)		F(5,1)		F(5,5)	
	Média	E.P.	Média	E.P.	Média	E.P.
F pela Beta	197.2603	132.9748	25150953071	201002694257	121.4064	71.96039
F pela Qui-Quadrado	174.0411	123.6288	17495947392	99385108881	105.4764	66.48341

Fonte: elaborada pelo autor.

Tabela 85 - Média e erro padrão dos valores estimados para o mínimo da distribuição F em 100 amostras de tamanho 10000

Gerador	F(1,5)		F(5,1)		F(5,5)	
	Média	E.P.	Média	E.P.	Média	E.P.
F pela Beta	3.961582e-08	1.630644e-07	0.006844281	0.003312341	0.01122935	0.004708949
F pela Qui-Quadrado	3.796628e-08	8.386135e-08	0.006977915	0.003302552	0.01129847	0.005106866

Fonte: elaborada pelo autor.

10.1.4 Amplitude

Tabela 86 - Média e erro padrão dos valores estimados para a amplitude da distribuição F em 100 amostras de tamanho 10000

Gerador	F(1,5)		F(5,1)		F(5,5)	
	Média	E.P.	Média	E.P.	Média	E.P.
F pela Beta	197.2603	132.9748	25150953071	201002694257	121.3951	71.96015
F pela Qui-Quadrado	174.0411	123.6288	17495947392	99385108881	105.4651	66.48325

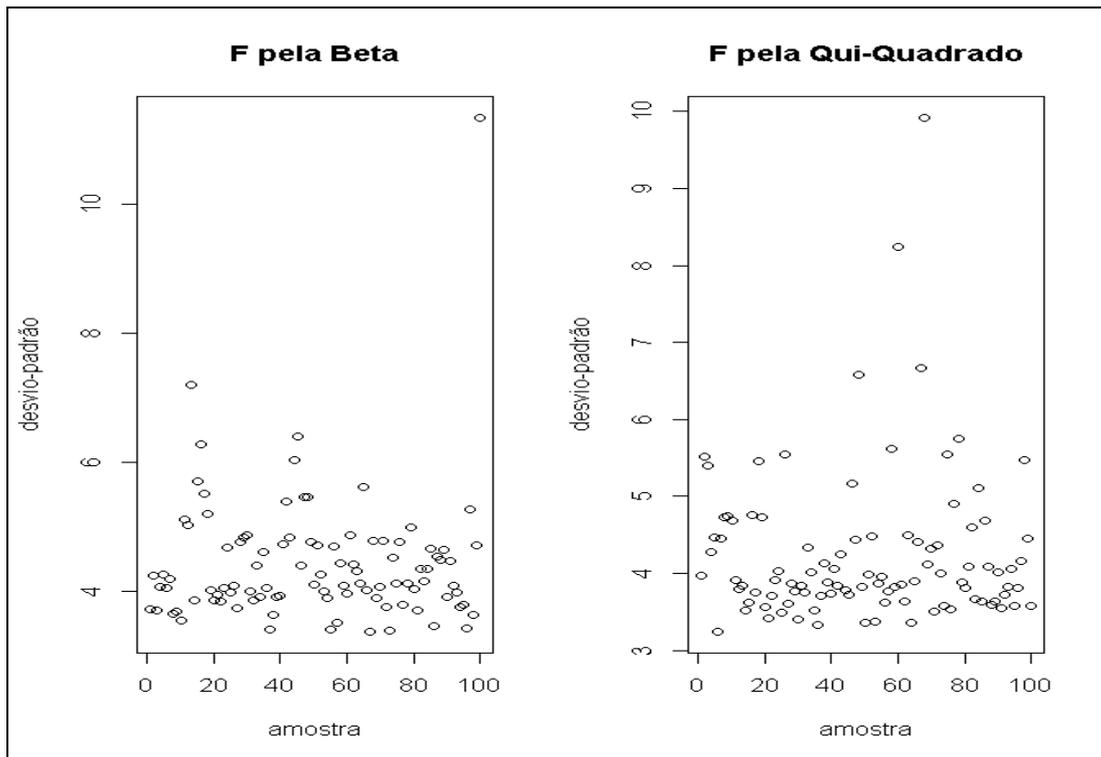
Fonte: elaborada pelo autor.

10.1.5 Desvio-Padrão

Tabela 87 - Média e erro padrão dos valores estimados para o desvio-padrão da distribuição F em 100 amostras de tamanho 10000

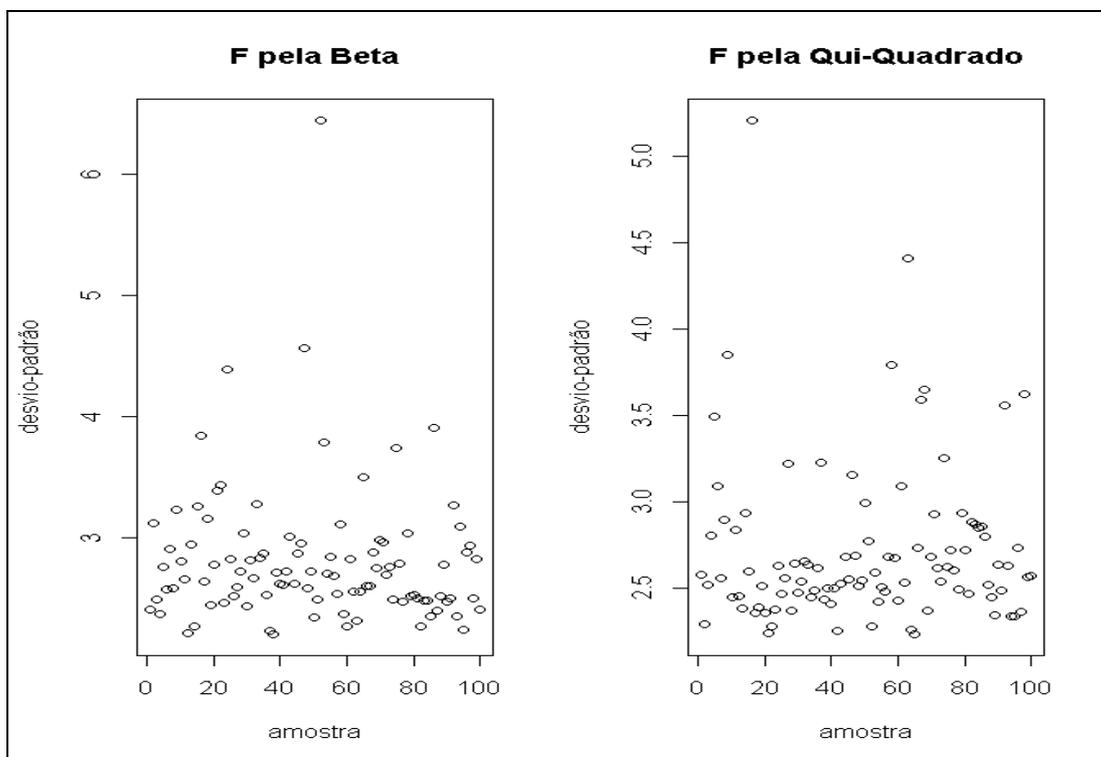
Gerador	F(1,5)		F(5,1)		F(5,5)	
	Média	E.P.	Média	E.P.	Média	E.P.
F pela Beta	4.418839	0.9934298	N.D.	N.D.	2.812776	0.5698404
F pela Qui-Quadrado	4.240205	0.9862142	N.D.	N.D.	2.719484	0.4649891

Fonte: elaborada pelo autor.



Quadro 45 - Gráficos dos valores estimados para o desvio-padrão da distribuição $F(1,5)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 46 - Gráficos dos valores estimados para o desvio-padrão da distribuição $F(5,5)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

10.1.6 Quartis

Tabela 88 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição F em 100 amostras de tamanho 10000

Gerador	F(1,5)		F(5,1)		F(5,5)	
	Média	E.P.	Média	E.P.	Média	E.P.
F pela Beta	0.1132289	0.004189741	0.5901057	0.01286061	0.5274519	0.007155641
F pela Qui-Quadrado	0.1133491	0.004092506	0.5913281	0.01269458	0.5276451	0.007807439

Fonte: elaborada pelo autor.

Tabela 89 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição F em 100 amostras de tamanho 10000

Gerador	F(1,5)		F(5,1)		F(5,5)	
	Média	E.P.	Média	E.P.	Média	E.P.
F pela Beta	0.529043	0.01191873	1.894693	0.04607719	1.000535	0.01215471
F pela Qui-Quadrado	0.5278787	0.01308743	1.889677	0.04872984	0.9994687	0.01286403

Fonte: elaborada pelo autor.

Tabela 90 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição F em 100 amostras de tamanho 10000

Gerador	F(1,5)		F(5,1)		F(5,5)	
	Média	E.P.	Média	E.P.	Média	E.P.
F pela Beta	1.696818	0.03184986	8.811734	0.3188664	1.893191	0.02697753
F pela Qui-Quadrado	1.693924	0.03577599	8.830767	0.3079152	1.897278	0.02554698

Fonte: elaborada pelo autor.

10.1.7 Distância Interquartilica

Tabela 91 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição F em 100 amostras de tamanho 10000

Gerador	F(1,5)		F(5,1)		F(5,5)	
	Média	E.P.	Média	E.P.	Média	E.P.
F pela Beta	1.583589	0.03074358	8.221629	0.3139174	1.365739	0.02462552
F pela Qui-Quadrado	1.580574	0.03493599	8.239439	0.3033141	1.369633	0.0218095

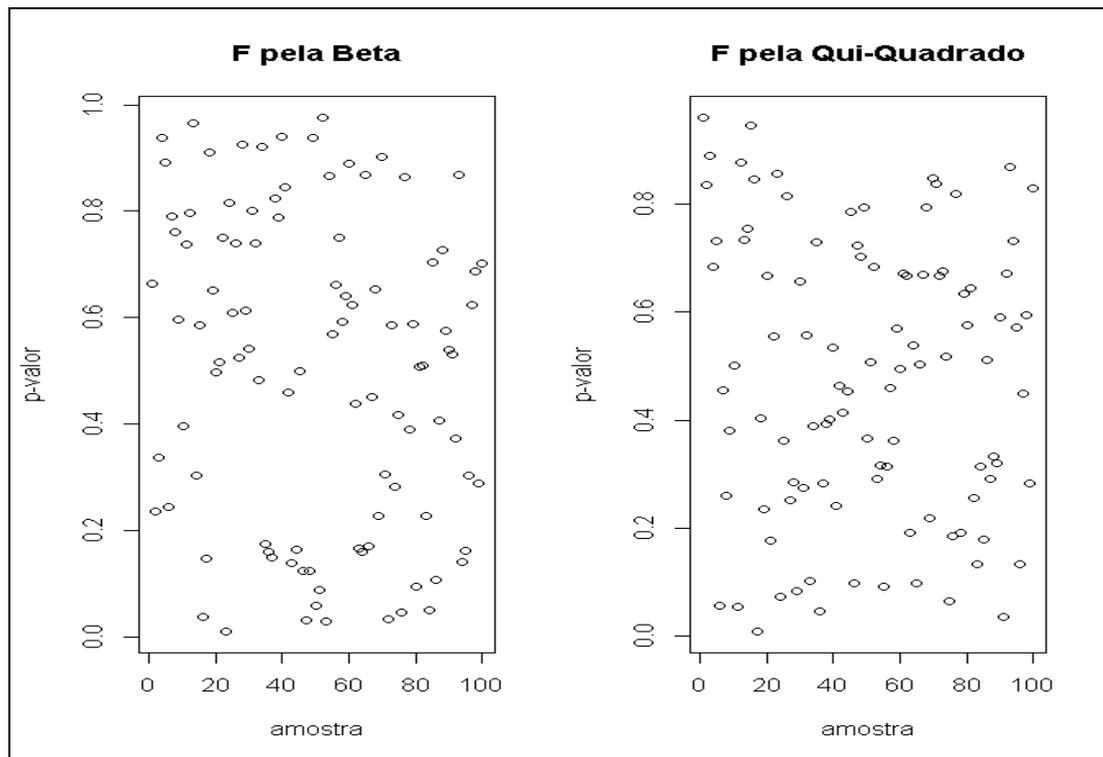
Fonte: elaborada pelo autor.

10.1.8 Teste de Aderência

Tabela 92 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição F em 100 amostras de tamanho 10000

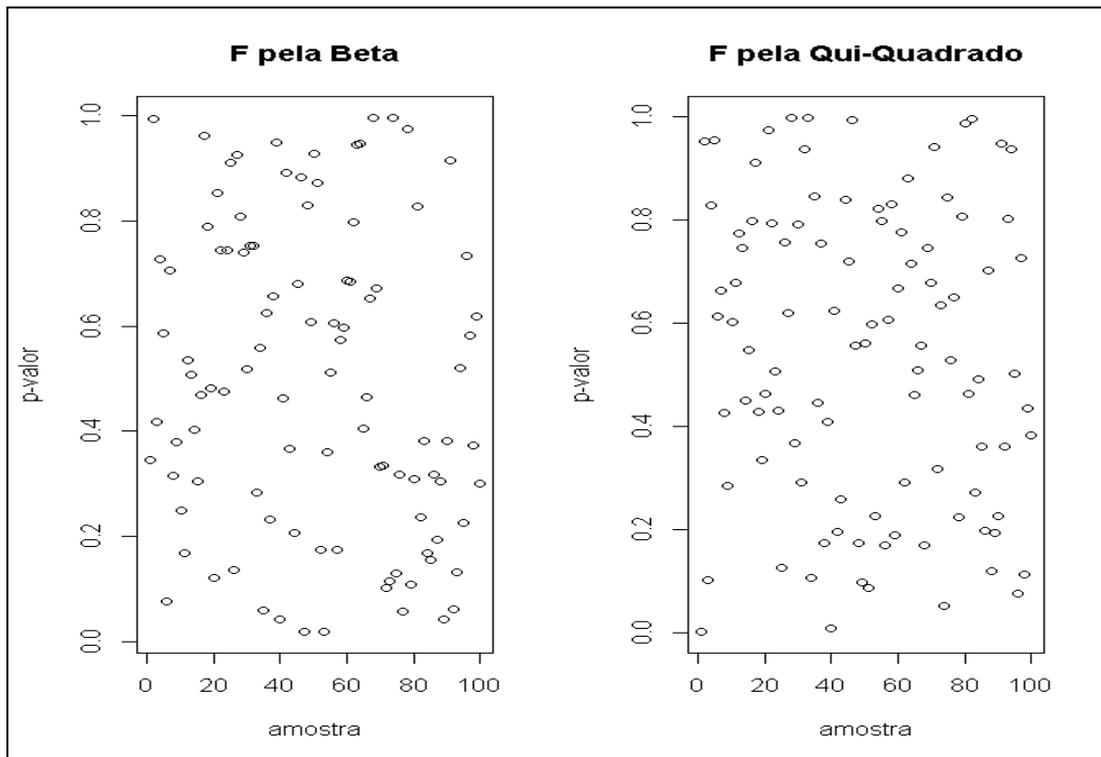
Gerador	F(1,5)		F(5,1)		F(5,5)	
	Média	E.P.	Média	E.P.	Média	E.P.
F pela Beta	0.5027308	0.2903874	0.4959093	0.2914188	0.5124999	0.318903
F pela Qui-Quadrado	0.473582	0.2593207	0.5397403	0.2882822	0.4834621	0.3132357

Fonte: elaborada pelo autor.



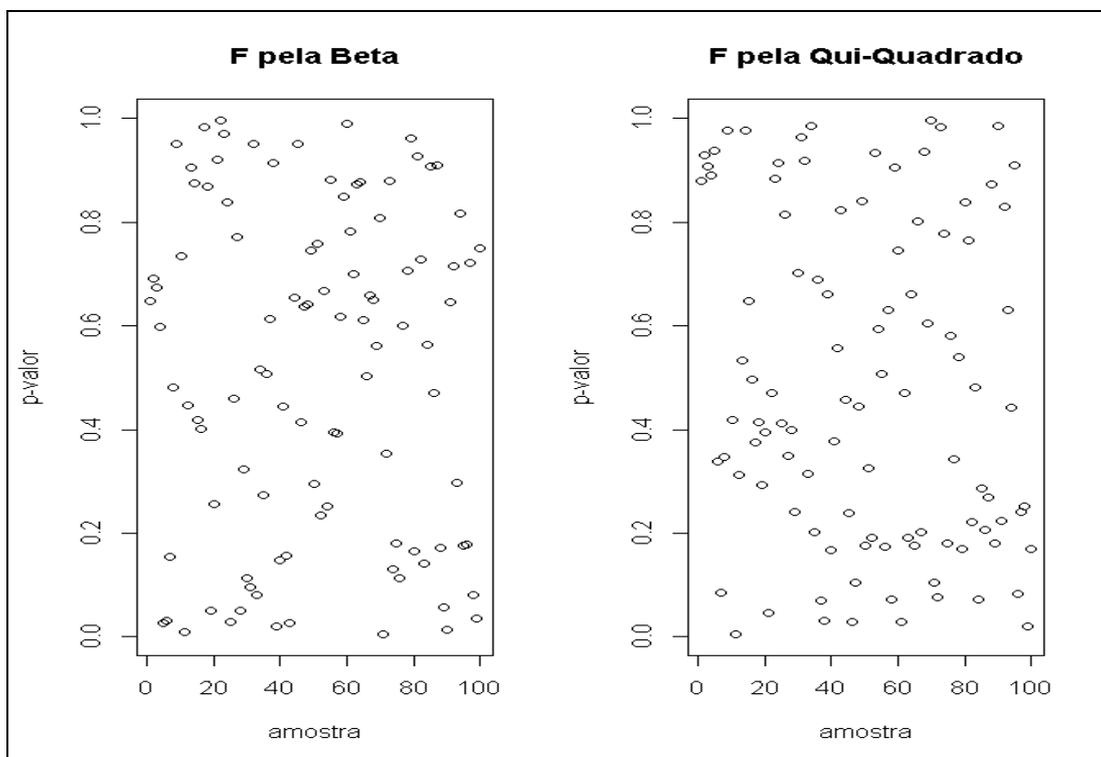
Quadro 47 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição F(1,5) em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 48 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $F(5,1)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 49 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $F(5,5)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

10.1.9 Conclusões

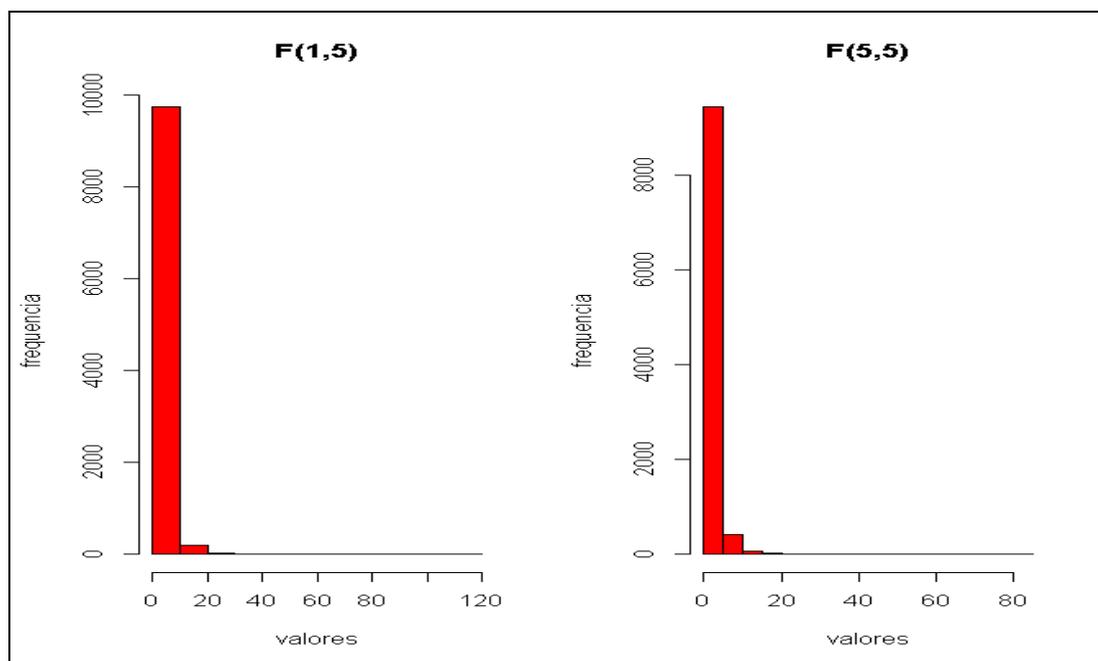
Ambos os algoritmos passaram sem problemas pelo Teste Qui-Quadrado de Aderência, independentemente se o nível de significância escolhido foi 5% ou 1%.

As estimativas do gerador baseado na Distribuição Beta mostraram-se mais próximas aos valores exatos se comparadas às estimativas do gerador baseado na Distribuição Qui-Quadrado. Assim, recomendamos o algoritmo baseado na Distribuição Beta para gerar amostras da Distribuição F com tamanho $n = 10000$.

Tabela 93 – Estimativas médias do algoritmo F pela Beta em 100 amostras de tamanho 10000 da Distribuição F

Medidas	F pela Beta		
	F(1,5)	F(5,1)	F(5,5)
Tempo de Execução	5.5949 s	6.0283 s	6.7995 s
Esperança	1.672302	N.D.	1.665627
Máximo	197.2603	25150953071	121.4064
Mínimo	3.961582e-08	0.006844281	0.01122935
Amplitude	197.2603	25150953071	121.3951
Desvio-Padrão	4.418839	N.D.	2.812776
1º Quartil	0.1132289	0.5901057	0.5274519
2º Quartil	0.529043	1.894693	1.000535
3º Quartil	1.696818	8.811734	1.893191
Dist. Interquart.	1.583589	8.221629	1.365739

Fonte: elaborada pelo autor.



Quadro 50 – Histogramas de 10000 valores gerados da Distribuição F pelo algoritmo F pela Beta

Fonte: Elaborado pelo autor.

11. GERADORES DE NÚMEROS ALEATÓRIOS DA DISTRIBUIÇÃO t DE STUDENT

Neste capítulo, apresentaremos os algoritmos geradores da Distribuição t de Student. Assim como os algoritmos anteriores da Distribuição Qui-Quadrado e da Distribuição F, estes algoritmos também podem ser dependentes de geradores de outras distribuições de probabilidade, como veremos a seguir. Lembrando que ν são os graus de liberdade da Distribuição t de Student.

11.1 Método de Rejeição Polar

Este algoritmo foi proposto por Bailey (1994). Sua vantagem é não depender do gerador de qualquer outra distribuição de probabilidade além da Distribuição Uniforme.

1. Gerar $v_1, v_2 \sim U[-1;1]$.

2. Fazer $r^2 = v_1^2 + v_2^2$.

3. Se $r^2 \geq 1$ então

Volte para o passo 1.

senão

$$\text{Sair } X = v_1 \sqrt{\frac{\nu(r^{-4/\nu} - 1)}{r}}.$$

11.2 Rejeição da densidade t(3)

Este algoritmo foi proposto por Best (1978), e está baseado na rejeição da função densidade de probabilidade da Distribuição t com 3 graus de liberdade.

Repita

Repita

Gerar $U_1, U_2 \sim U[0;1]$.

Fazer $U_2 = U_2 - \frac{1}{2}$.

Até que $U_1^2 + U_2^2 \leq U_1$

Fazer $X = \sqrt{3} \frac{U_2}{U_1}$.

Gerar $U_3 \sim U[0;1]$.

Fazer $Z = X^2$, $W = 1 + \frac{Z}{3}$.

Fazer $Y = 2 \log \left(\frac{\frac{9}{16} W^2}{U} \right)$.

Fazer $Aceitar = [Y \geq 1 - Z]$.

Se não $Aceitar$ então

$$Aceitar = \left[Y \geq (\nu + 1) \log \left(\frac{\nu + 1}{\nu + Z} \right) \right].$$

Até que $Aceitar = Verdadeiro$

Sair X .

11.3 Gerador baseado na Distribuição Gama

Gerar $S \sim U[-1;1]$.

Gerar $G_1 \sim Gama(1/2, 1)$, $G_2 \sim Gama(\nu/2, 1)$.

Se $S < 0$ então

$$\text{Sair } X = -\sqrt{\nu} \sqrt{\frac{G_1}{G_2}}.$$

senão

$$\text{Sair } X = \sqrt{\nu} \sqrt{\frac{G_1}{G_2}}.$$

A Distribuição Gama pode ser gerada pelos algoritmos já mencionados anteriormente. Em nossas análises, utilizaremos o algoritmo disponível no software R.

11.4 Gerador baseado na Distribuição Normal e na Distribuição Qui-Quadrado

Este algoritmo está baseado na geração de uma ocorrência da Distribuição Normal padrão e de uma ocorrência da Distribuição Qui-Quadrado.

Gerar $Z \sim N(0,1)$.

Gerar $Y \sim \chi^2(\nu)$.

$$\text{Sair } X = \frac{Z}{\sqrt{Y/\nu}}.$$

Este é o algoritmo disponível no software R para geração de números aleatórios da Distribuição t de Student.

12. ESTIMATIVAS DAS AMOSTRAS DA DISTRIBUIÇÃO t DE STUDENT

Nos capítulos subsequentes, apresentaremos as estimativas com amostras de tamanho $n = 100$ e $n = 10000$ da Distribuição t de Student com graus de liberdade ν iguais a 3, 15 e 30 para as seguintes medidas: esperança, máximo e mínimo, amplitude, desvio-padrão, assimetria, curtose, quartis, distância interquartílica e valor-p para o teste de aderência. Os valores “exatos” das medidas para a Distribuição t de Student com graus de liberdade ν iguais a 3, 15 e 30 são, respectivamente: esperança = 0, 0, 0; máximo = 47.92773, 6.108868, 5.054032 e mínimo = -47.92773, -6.108868, -5.054032; amplitude = 95.85546, 12.21774, 10.10806; desvio-padrão = 1.732051, 1.074172, 1.035098; assimetria = não definida, 0, 0; curtose = infinito, 0.5454545, 0.2307692; 1º quartil = -0.7648923, -0.6911969, -0.6827557; 2º quartil = 0, 0, 0 e 3º quartil = 0.7648923, 0.6911969, 0.6827557; distância interquartílica = 1.221773, 1.382394, 1.365511. Considera-se como máximo o valor x tal que a probabilidade $P(X \leq x) = 0.99999$, e como mínimo o valor x tal que a probabilidade $P(X \leq x) = 0.00001$.

12.1 Estimativas com amostras de tamanho $n = 100$

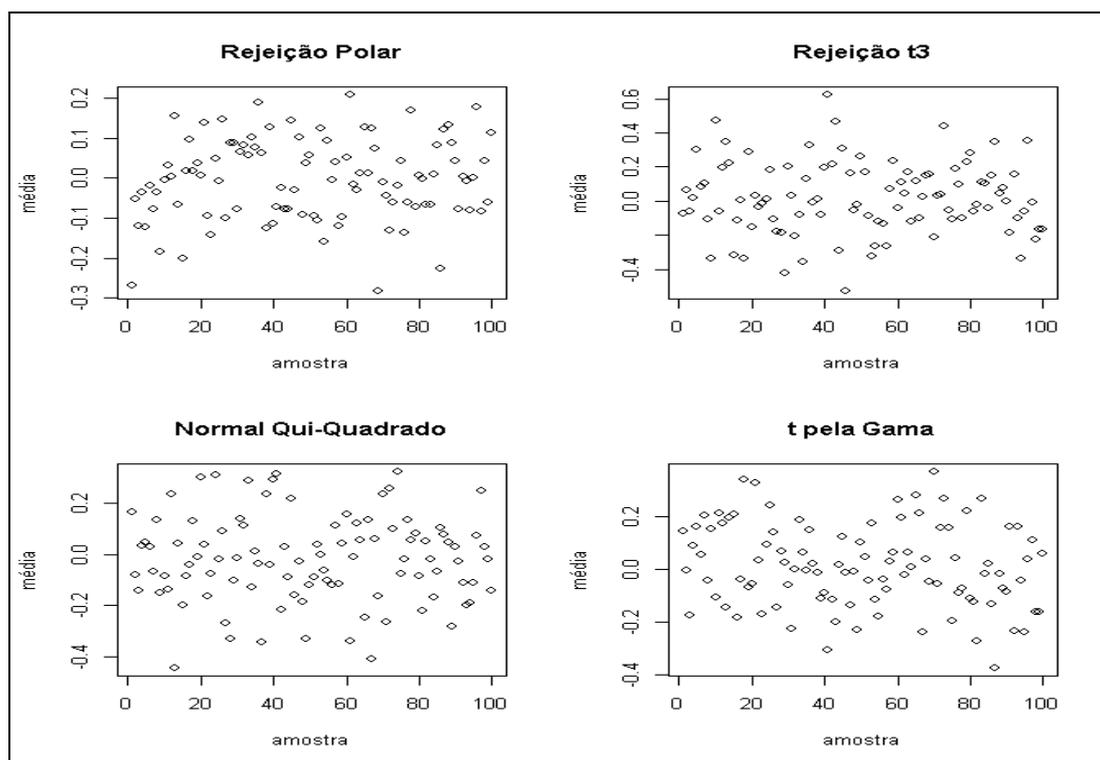
Nas tabelas abaixo, apresentamos a média e o erro padrão dos valores estimados pelos geradores para cada uma das medidas analisadas. Nos gráficos, ilustramos o comportamento das estimativas dos geradores ao longo das 100 amostras.

12.1.1 Esperança

Tabela 94 – Média e erro padrão dos valores estimados para a esperança da distribuição t de Student em 100 amostras de tamanho 100

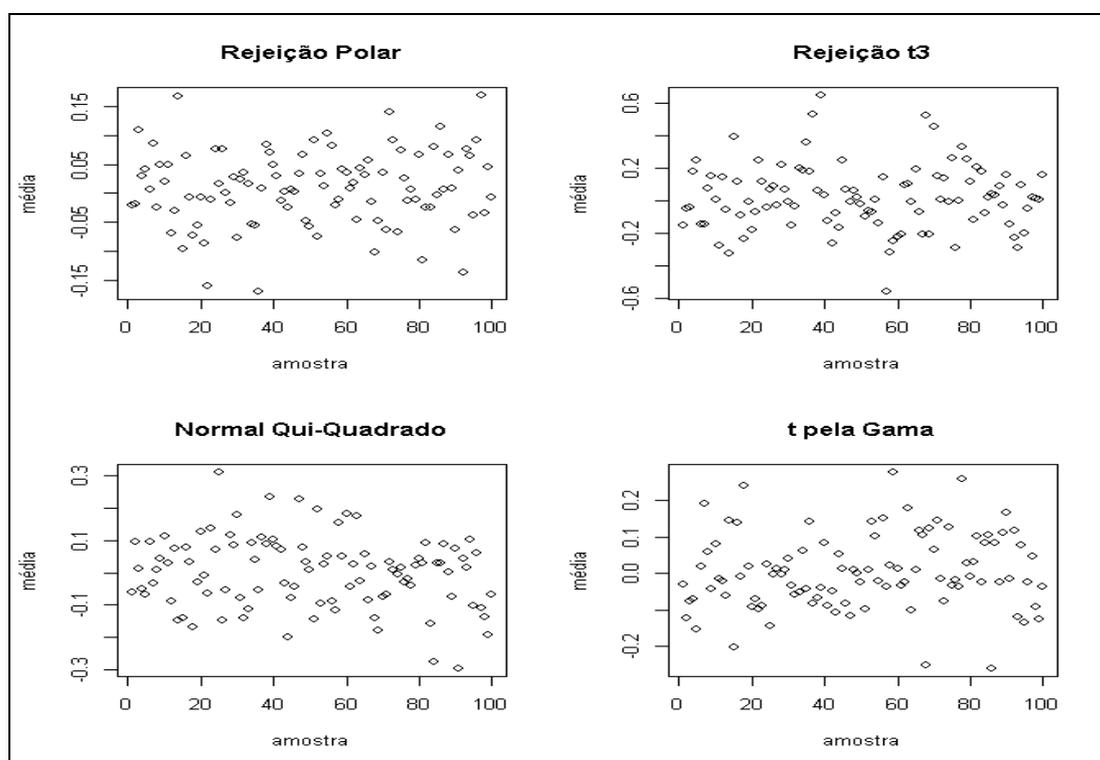
Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	-0.004598665	0.1019409	0.009325166	0.06447905	0.01287283	0.07085583
Rejeição t(3)	0.01732537	0.2116898	0.01938277	0.1997123	-0.008298479	0.1865934
Normal/Qui-Quadrado	-0.01986648	0.1692288	0.0004833724	0.1098877	0.002908149	0.1114964
t pela Gama	0.00971455	0.1572465	0.004024177	0.1019915	0.006322134	0.1047522

Fonte: elaborada pelo autor.



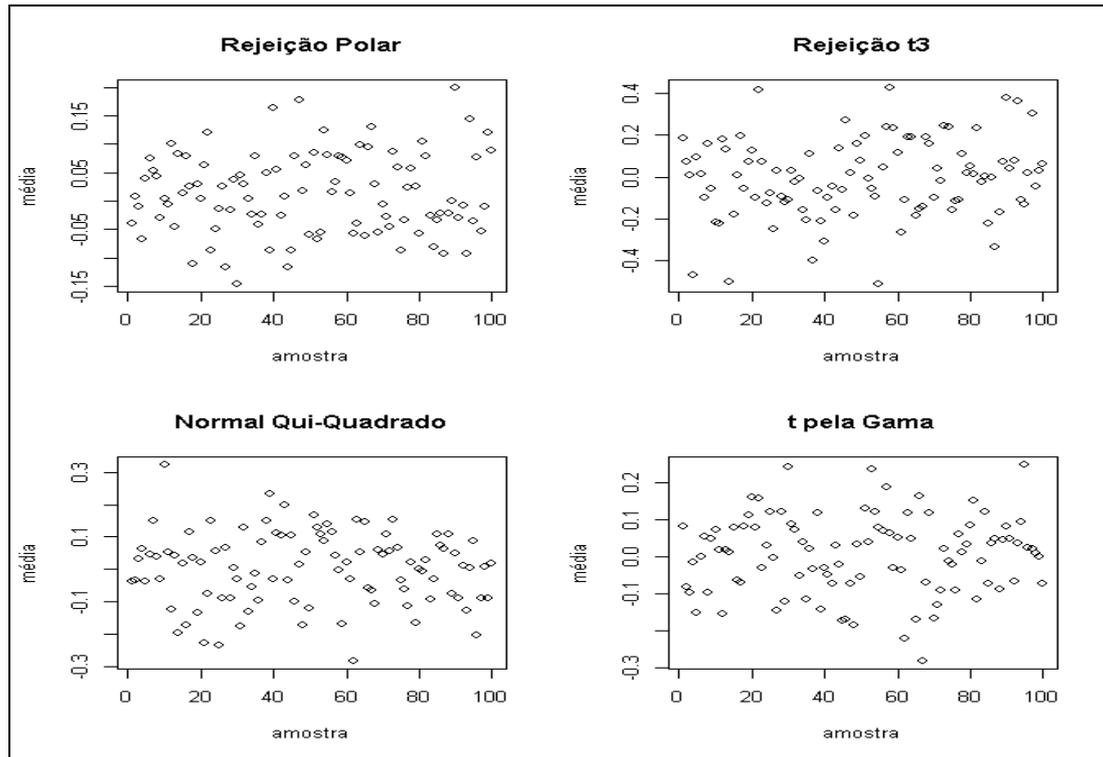
Quadro 51 – Gráficos dos valores estimados para a esperança da distribuição $t(3)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 52 - Gráficos dos valores estimados para a esperança da distribuição $t(15)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 53 – Gráficos dos valores estimados para a esperança da distribuição $t(30)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.

12.1.2 Máximo e Mínimo

Tabela 95 – Média e erro padrão dos valores estimados para o máximo da distribuição t de Student em 100 amostras de tamanho 100

Gerador	$t(3)$		$t(15)$		$t(30)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	2.096557	0.3737161	1.266886	0.03657693	1.223106	0.03310312
Rejeição $t(3)$	7.418707	5.471068	7.777956	4.732135	7.43231	5.139396
Normal/Qui-Quadrado	6.395364	3.293031	2.790624	0.5617903	2.714362	0.504266
t pela Gama	6.623744	5.112158	2.902177	0.6029363	2.708596	0.4784058

Fonte: elaborada pelo autor.

Tabela 96 – Média e erro padrão dos valores estimados para o mínimo da distribuição t de Student em 100 amostras de tamanho 100

Gerador	$t(3)$		$t(15)$		$t(30)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	-2.119795	0.3521025	-1.26584	0.03738448	-1.222045	0.03151934
Rejeição $t(3)$	-6.834417	3.452183	-7.74489	5.388736	-7.192102	4.030778
Normal/Qui-Quadrado	-7.221168	5.612062	-2.839191	0.6040959	-2.760361	0.6058763
t pela Gama	-6.306513	4.427112	-2.853139	0.5646511	-2.763708	0.5148828

Fonte: elaborada pelo autor.

12.1.3 Amplitude

Tabela 97- Média e erro padrão dos valores estimados para a amplitude da distribuição t de Student em 100 amostras de tamanho 100

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	4.216352	0.5257759	2.532726	0.05486362	2.44515	0.04354627
Rejeição t(3)	14.25312	5.974151	15.52285	7.122265	14.62441	6.317338
Normal/Qui-Quadrado	13.61653	6.564588	5.629815	0.7653944	5.474723	0.7786894
t pela Gama	12.93026	6.411619	5.755317	0.8445249	5.472303	0.7028371

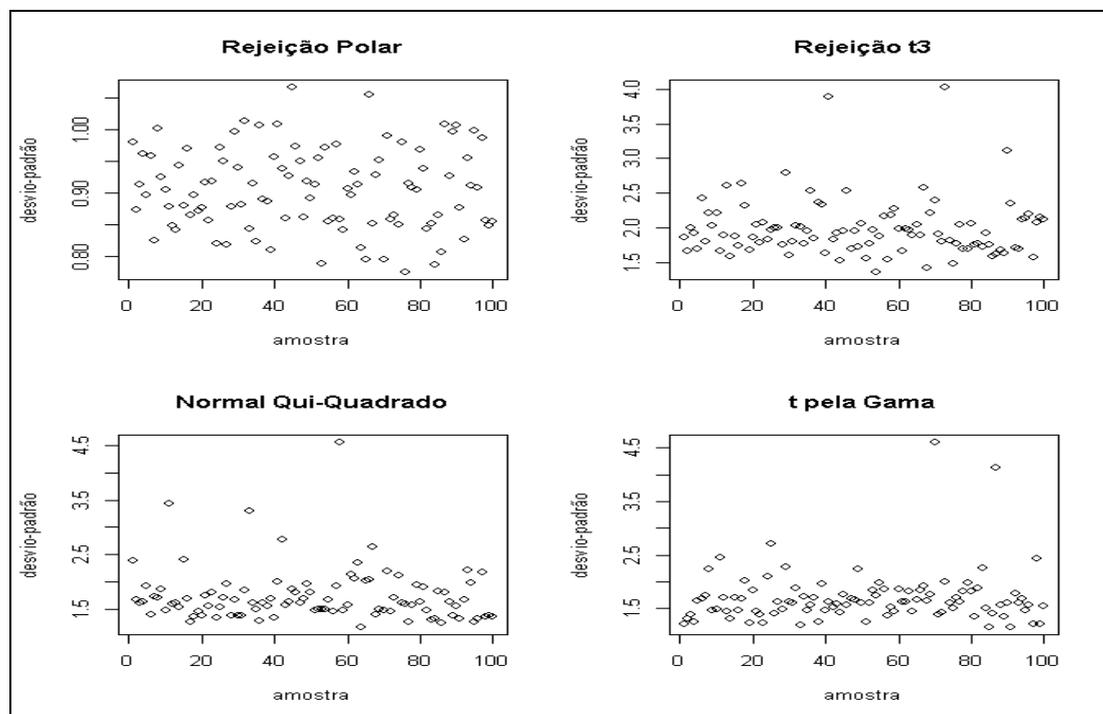
Fonte: elaborada pelo autor.

12.1.4 Desvio-Padrão

Tabela 98 – Média e erro padrão dos valores estimados para o desvio-padrão da distribuição t de Student em 100 amostras de tamanho 100

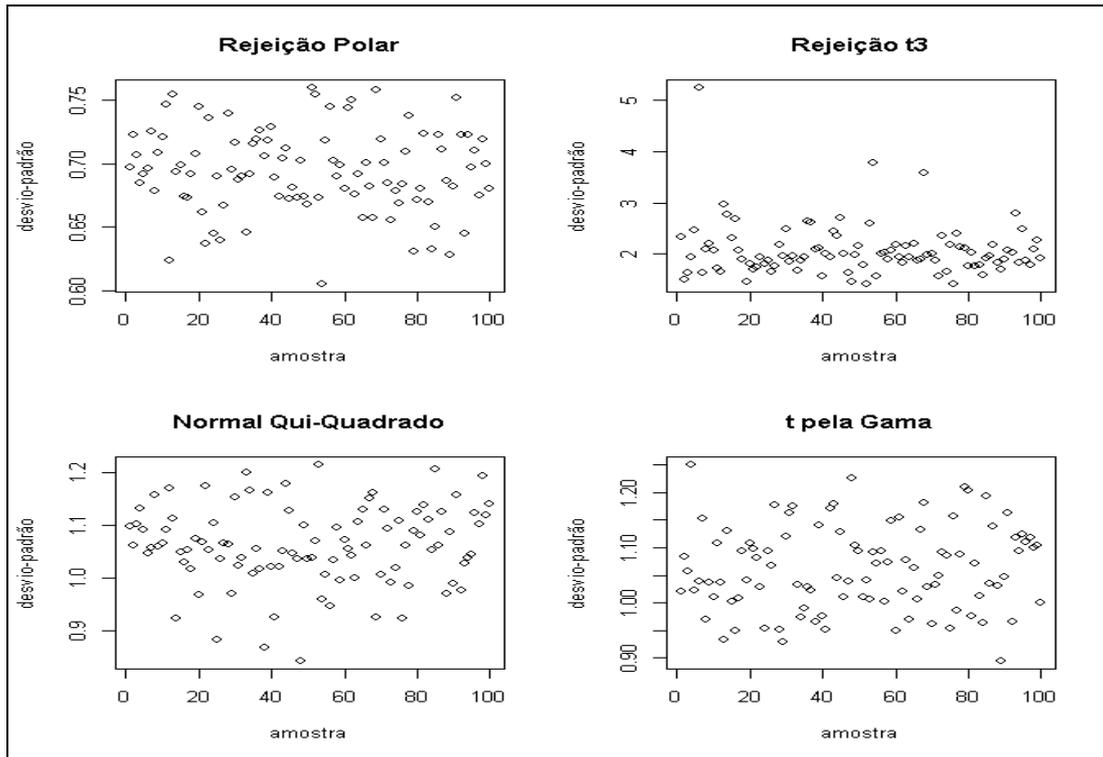
Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	0.905292	0.06460766	0.6943889	0.0328755	0.683983	0.03898796
Rejeição t(3)	1.9887	0.417468	2.068609	0.5124419	2.017714	0.424908
Normal/Qui-Quadrado	1.728108	0.4868518	1.061975	0.0747561	1.04709	0.07644573
t pela Gama	1.689665	0.4872077	1.061645	0.07613666	1.042441	0.08258455

Fonte: elaborada pelo autor.



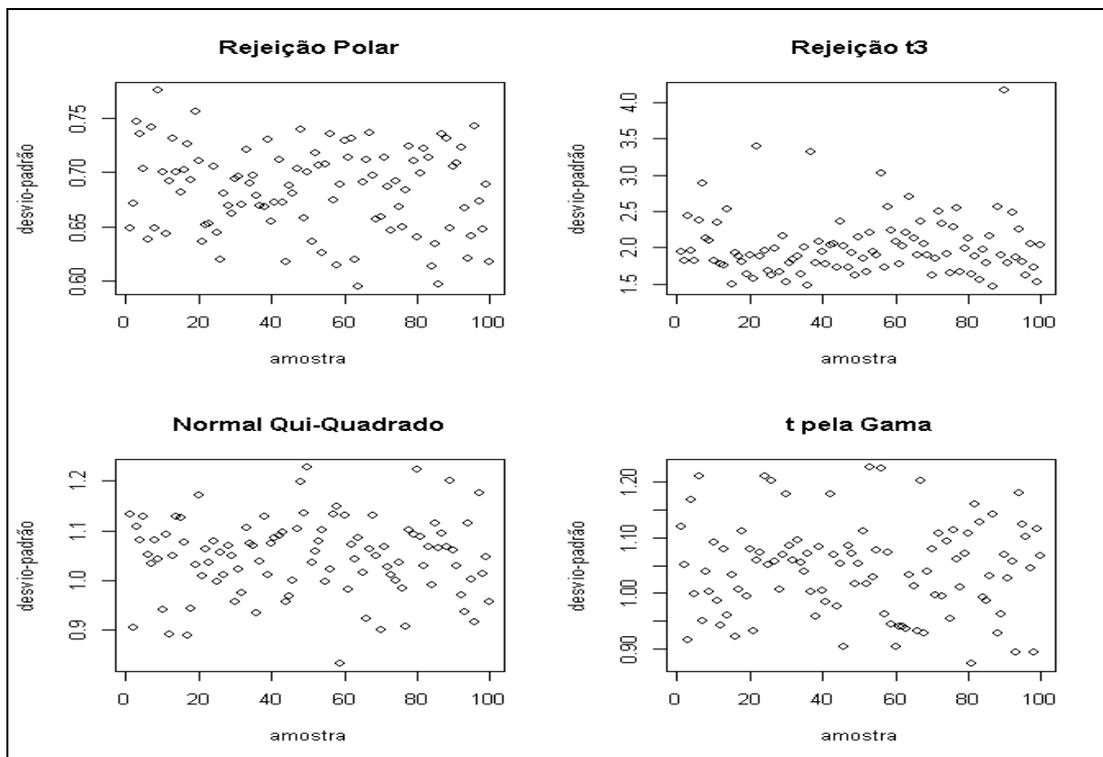
Quadro 54 – Gráficos dos valores estimados para o desvio-padrão da distribuição t(3) em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 55 – Gráficos dos valores estimados para o desvio-padrão da distribuição $t(15)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 56 – Gráficos dos valores estimados para o desvio-padrão da distribuição $t(30)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.

12.1.5 Assimetria

Tabela 99 – Média e erro padrão dos valores estimados para a assimetria da distribuição t de Student em 100 amostras de tamanho 100

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	N.D.	N.D.	-0.0009921229	0.1275464	-0.02662405	0.1420623
Rejeição t(3)	N.D.	N.D.	0.04547138	1.736186	0.04276195	1.768826
Normal/Qui-Quadrado	N.D.	N.D.	-0.01360928	0.3344405	-0.02468971	0.3028857
t pela Gama	N.D.	N.D.	0.02242046	0.2992206	-0.01294353	0.2411218

Fonte: elaborada pelo autor.

12.1.6 Curtose

Tabela 100 – Média e erro padrão dos valores estimados para a curtose da distribuição t de Student em 100 amostras de tamanho 100

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	-0.3750947	0.4074466	-1.019636	0.1517819	-1.053629	0.154776
Rejeição t(3)	7.000221	11.9084	7.281677	10.97333	6.987864	11.31793
Normal/Qui-Quadrado	9.153857	12.19695	0.386272	0.7310924	0.2540617	0.6432579
t pela Gama	8.815519	13.86484	0.424218	0.7793396	0.2427952	0.6451201

Fonte: elaborada pelo autor.

12.1.7 Quartis

Tabela 101 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição t de Student em 100 amostras de tamanho 100

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	-0.6582047	0.1421557	-0.5432306	0.09188405	-0.5437369	0.1104535
Rejeição t(3)	-1.098231	0.1704976	-1.107763	0.1650684	-1.118499	0.1627833
Normal/Qui-Quadrado	-0.7781199	0.1681507	-0.6701203	0.1401034	-0.6670346	0.1340985
t pela Gama	-0.753171	0.1476401	-0.6785246	0.1364987	-0.6756497	0.1320089

Fonte: elaborada pelo autor.

Tabela 102 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição t de Student em 100 amostras de tamanho 100

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	-0.0006509152	0.1354208	0.004609962	0.1021405	0.01904434	0.1139502
Rejeição t(3)	0.002926382	0.2687609	0.02455318	0.2735703	-0.003234995	0.2616832
Normal/Qui-Qu.	-0.009794832	0.1302142	-0.002341417	0.1289588	0.005247962	0.1327185
t pela Gama	-0.007814776	0.1354155	-0.002381686	0.1248561	-0.005964509	0.130517

Fonte: elaborada pelo autor.

Tabela 103 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição t de Student em 100 amostras de tamanho 100

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	0.6499294	0.1427702	0.5724181	0.09618475	0.5759705	0.1037415
Rejeição t(3)	1.110021	0.1879189	1.132244	0.1697455	1.119253	0.1483883
Normal/Qui-Quadrado	0.7395151	0.1559699	0.6836792	0.1369405	0.6821984	0.1514448
t pela Gama	0.7577238	0.1413862	0.6832777	0.1493293	0.7022948	0.1526656

Fonte: elaborada pelo autor.

12.1.8 Distância Interquartilica

Tabela 104 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição t de Student em 100 amostras de tamanho 100

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	1.308134	0.1588184	1.115649	0.105916	1.119707	0.1332332
Rejeição t(3)	2.208252	0.2062949	2.240007	0.2062251	2.237752	0.1966478
Normal/Qui-Quadrado	1.517635	0.1886474	1.3538	0.1532061	1.349233	0.1594522
t pela Gama	1.510895	0.181088	1.361802	0.1775066	1.377944	0.1600078

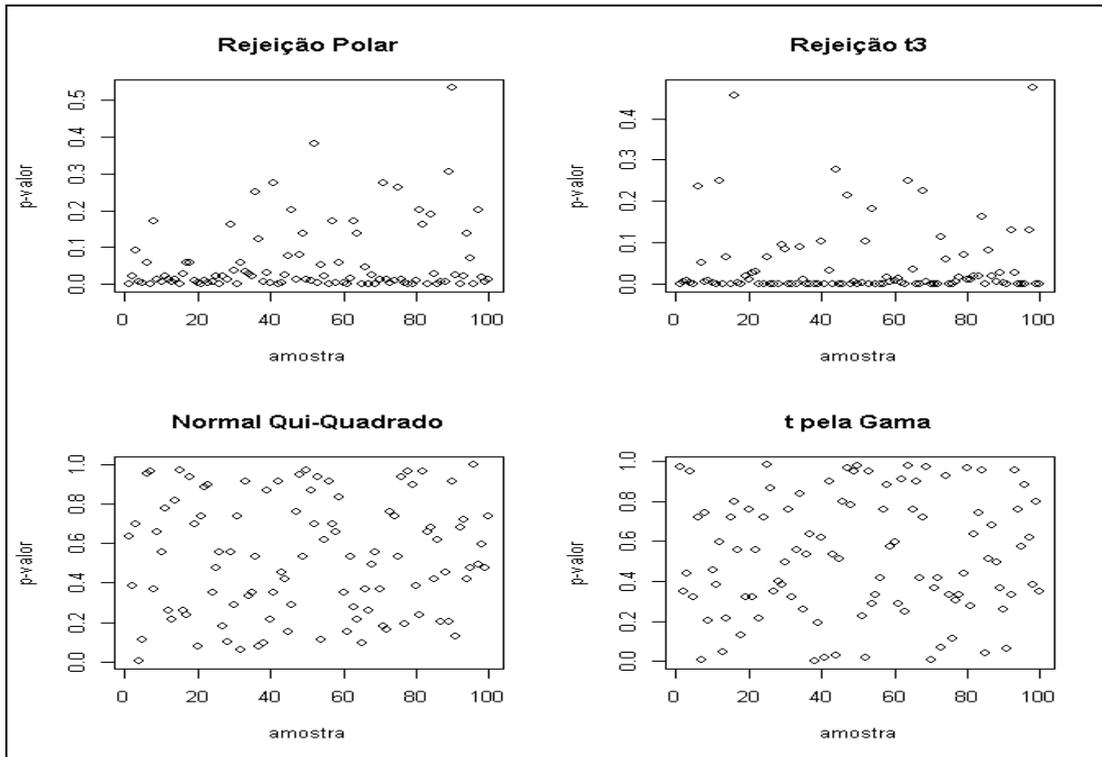
Fonte: elaborada pelo autor.

12.1.9 Teste de Aderência

Tabela 105 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição t de Student em 100 amostras de tamanho 100

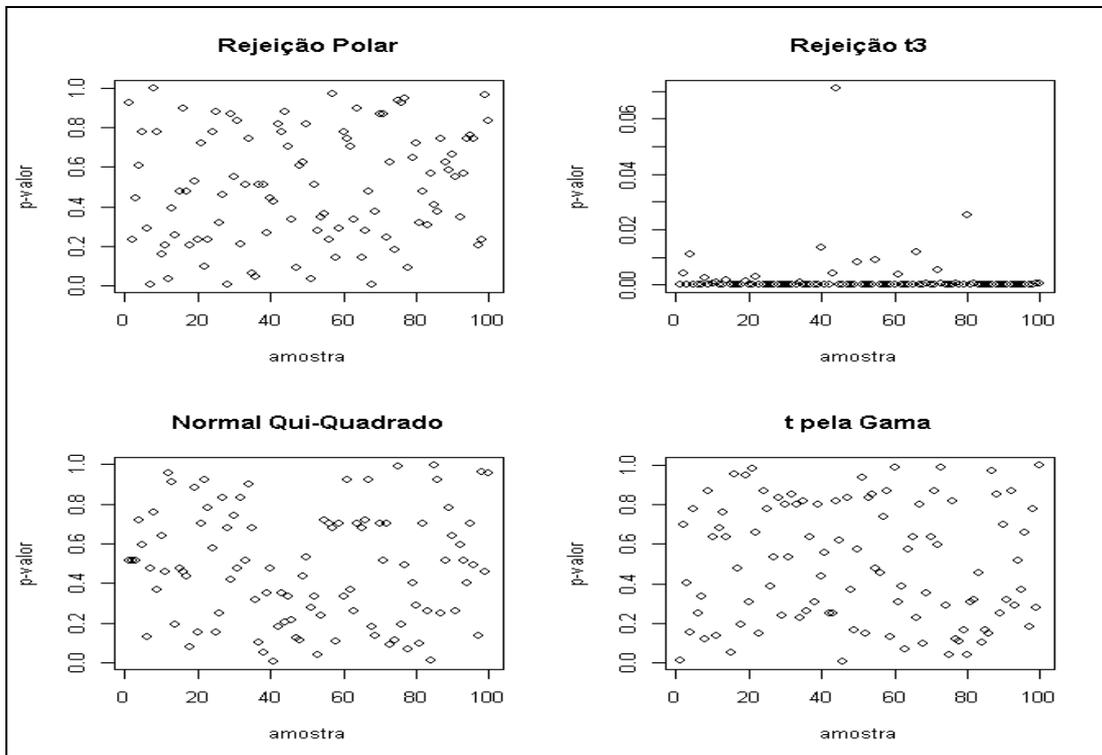
Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	0.05891308	0.09503757	0.5002167	0.2812343	0.4820837	0.2855532
Rejeição t(3)	0.0443549	0.08860526	0.001824125	0.007823174	0.001322604	0.007510335
Normal/Qui-Quadrado	0.5242399	0.2878748	0.4769547	0.2781969	0.4903221	0.2811345
t pela Gama	0.5233127	0.2940255	0.4974157	0.2990075	0.4607708	0.2826925

Fonte: elaborada pelo autor.



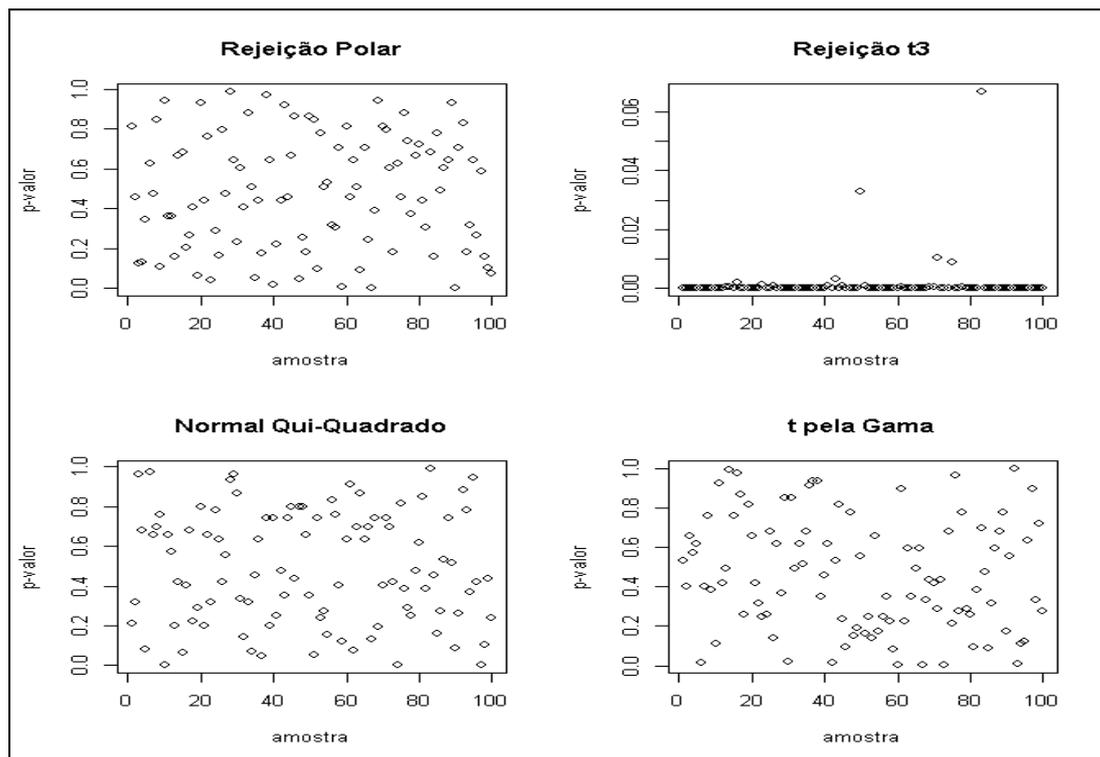
Quadro 57- Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $t(3)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 58 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $t(15)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.



Quadro 59 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $t(30)$ em 100 amostras de tamanho 100

Fonte: Elaborado pelo autor.

12.1.10 Conclusões

Para amostras de tamanho $n = 100$, o algoritmo baseado na Distribuição Normal e na Distribuição Qui-Quadrado mostrou-se um pouco mais preciso que os demais. Então, o indicamos para gerar amostras com tamanho $n = 100$ da Distribuição t de Student.

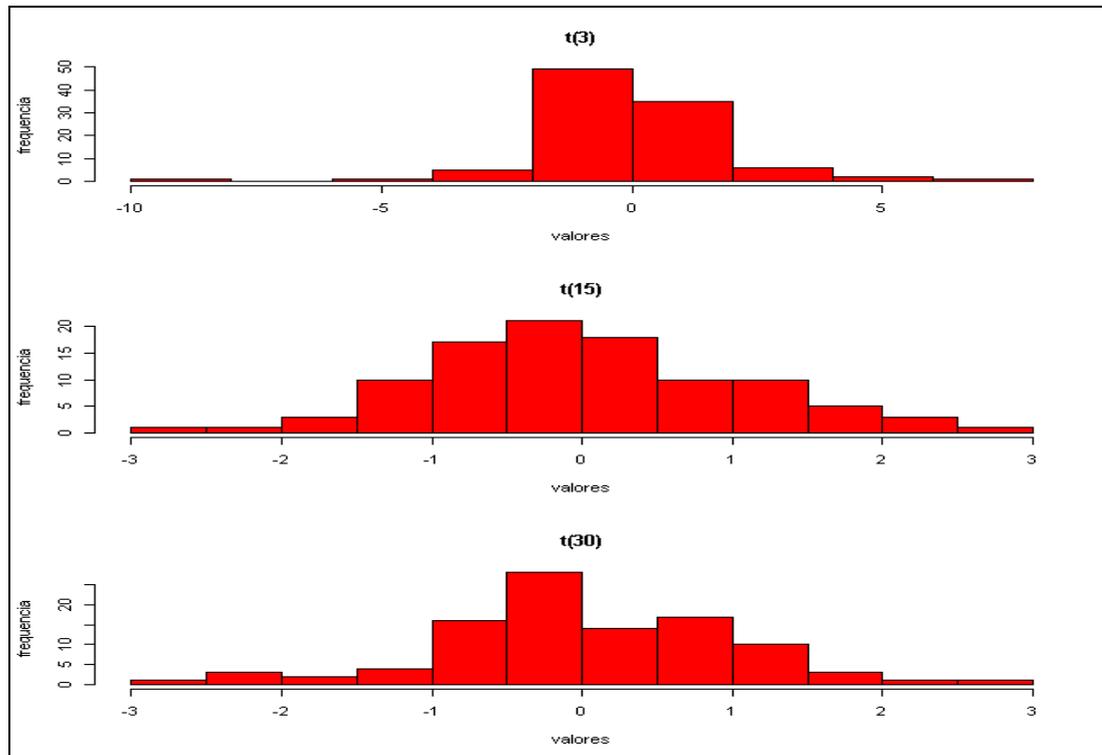
Assim como na Distribuição Qui-Quadrado, para a Distribuição t de Student também podemos observar que um algoritmo não passou no Teste Qui-Quadrado de Aderência: o algoritmo baseado na rejeição de uma Distribuição t com 3 graus de liberdade, que aqui chamamos de Rejeição t_3 . Logo, não aconselhamos o seu uso para geração de amostras com tamanho 100 da Distribuição t de Student.

Tabela 106 – Estimativas médias do algoritmo Normal/Qui-Quadrado em 100 amostras de tamanho 100 da Distribuição t de Student

Medidas	Normal/Qui-Quadrado		
	$t(3)$	$t(15)$	$t(30)$
Esperança	-0.01986648	0.0004833724	0.002908149
Máximo	6.395364	2.790624	2.714362
Mínimo	-7.221168	-2.839191	-2.760361
Amplitude	13.61653	5.629815	5.474723
Desvio-Padrão	1.728108	1.061975	1.04709

Assimetria	N.D.	-0.01360928	-0.02468971
Curtose	9.153857	0.386272	0.2540617
1º Quartil	-0.7781199	-0.6701203	-0.6670346
2º Quartil	-0.009794832	-0.002341417	0.005247962
3º Quartil	0.7395151	0.6836792	0.6821984
Dist. Interquart.	1.517635	1.3538	1.349233

Fonte: elaborada pelo autor.



Quadro 60 – Histogramas de 100 valores gerados da Distribuição t de Student pelo algoritmo Normal/Qui-Quadrado

Fonte: Elaborado pelo autor.

12.2. Estimativas com amostras de tamanho $n = 10000$

Nas tabelas a seguir, apresentamos a média e o erro padrão dos valores estimados pelos geradores para cada uma das medidas analisadas. Nos gráficos, ilustramos o comportamento das estimativas dos geradores ao longo das 100 amostras.

12.2.1 Tempo de Execução

Tabela 107 – Média e erro padrão dos tempos de execução, em segundos, medidos em 100 amostras de tamanho 10000 da Distribuição t de Student

Gerador	$\chi^2(1)$		$\chi^2(17)$		$\chi^2(50)$	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	0.8247	0.03975823	0.8417	0.02974827	0.8034	0.02327786
Rejeição t(3)	1.8059	0.07081196	1.8563	0.03836705	1.7489	0.02957254
Normal/Qui-Quadrado	5.5548	0.1596554	5.2568	0.08688643	4.9097	0.04003673
t pela Gama	5.7399	0.1759275	5.4802	0.06853415	5.1342	0.06108745

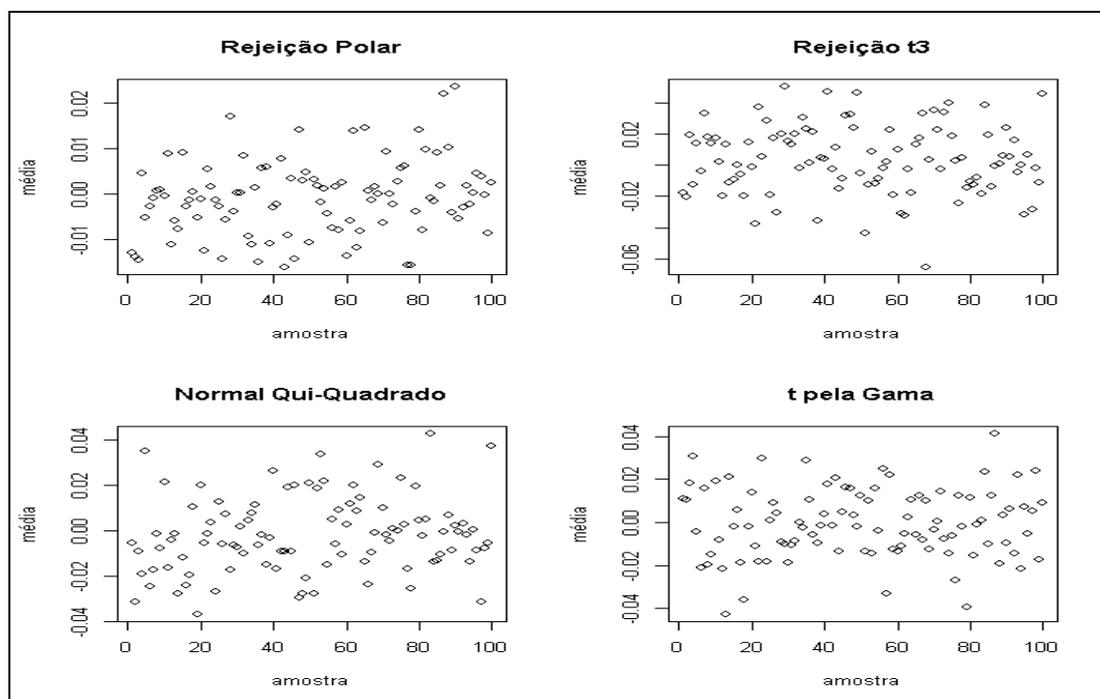
Fonte: elaborada pelo autor.

12.2.2 Esperança

Tabela 108 - Média e erro padrão dos valores estimados para a esperança da distribuição t de Student em 100 amostras de tamanho 10000

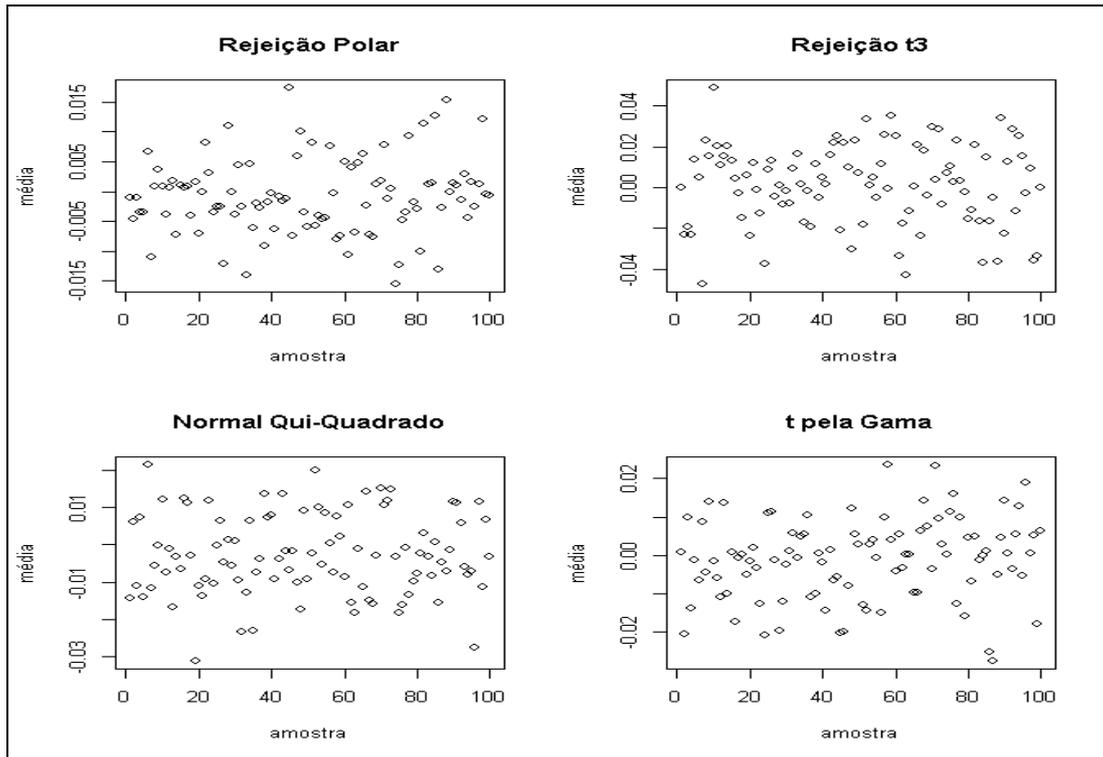
Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	-0.000994978	0.008361409	-0.0008118529	0.006314536	-0.0004204632	0.007090136
Rejeição t(3)	0.003500486	0.02200558	0.001238042	0.01963975	-0.0006904531	0.0217135
Normal/Qui-Quadrado	-0.002243976	0.01659397	-0.0027243	0.01083226	-0.0002610586	0.01029538
t pela Gama	-0.0005633687	0.01611234	-0.001068755	0.01050596	-0.001126007	0.01049263

Fonte: elaborada pelo autor.



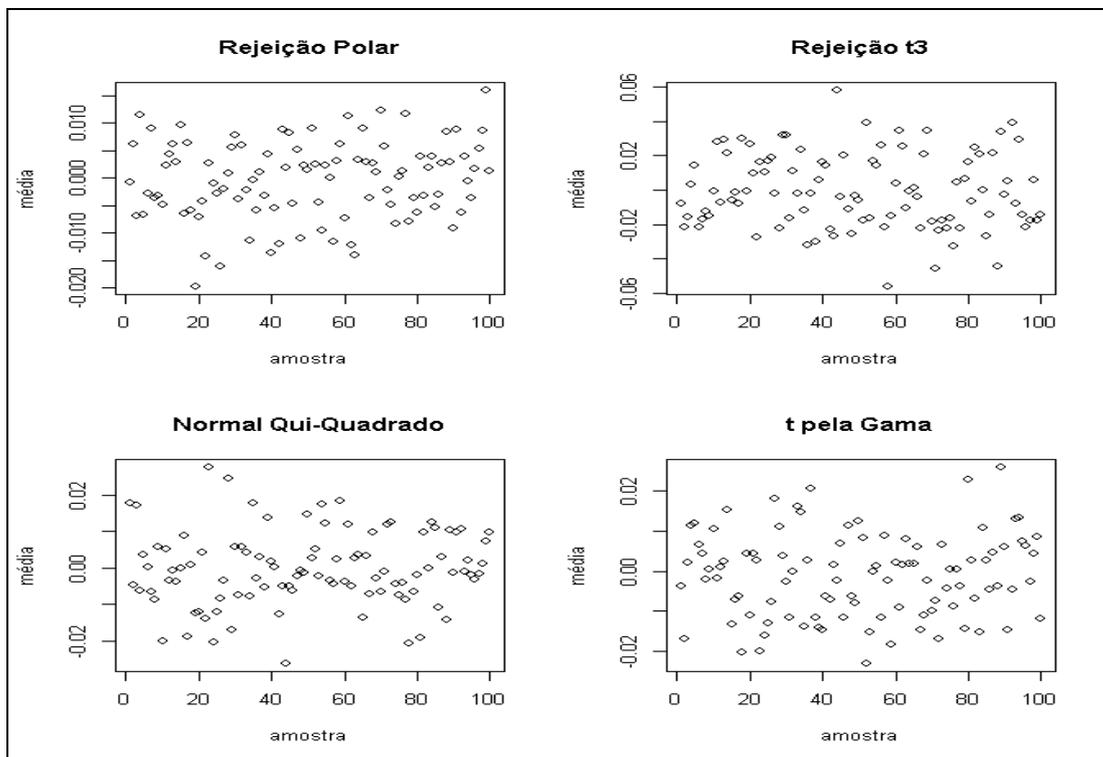
Quadro 61 - Gráficos dos valores estimados para a esperança da distribuição t(3) em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 62 - Gráficos dos valores estimados para a esperança da distribuição $t(15)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 63 - Gráficos dos valores estimados para a esperança da distribuição $t(30)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

12.2.3 Máximo e Mínimo

Tabela 109 - Média e erro padrão dos valores estimados para o máximo da distribuição t de Student em 100 amostras de tamanho 10000

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	3.253134	0.350302	1.305183	0.0004142814	1.255757	0.0005393244
Rejeição t(3)	36.23513	21.56082	32.9532	21.93775	33.07402	21.24237
Normal/Qui-Quadrado	31.99661	16.27621	5.20627	0.6005691	4.472856	0.5112167
t pela Gama	31.23282	15.83101	5.274739	0.6727141	4.41958	0.5065569

Fonte: elaborada pelo autor.

Tabela 110 - Média e erro padrão dos valores estimados para o mínimo da distribuição t de Student em 100 amostras de tamanho 10000

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	-3.231285	0.3089289	-1.305266	0.0003333101	-1.255804	0.0003652054
Rejeição t(3)	-34.21112	19.1052	-34.37653	19.46049	-31.70713	17.34393
Normal/Qui-Quadrado	-29.19614	14.97728	-5.177636	0.7496224	-4.449784	0.4296766
t pela Gama	-29.80888	14.00435	-5.138016	0.638735	-4.383875	0.4876735

Fonte: elaborada pelo autor.

12.2.4 Amplitude

Tabela 111- Média e erro padrão dos valores estimados para a amplitude da distribuição t de Student em 100 amostras de tamanho 10000

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	6.484419	0.4712651	2.610449	0.0004820758	2.511561	0.0006344208
Rejeição t(3)	70.44625	27.12852	67.32973	29.95593	64.78115	26.5967
Normal/Qui-Quadrado	61.19275	21.93637	10.38391	0.9565117	8.92264	0.6400792
t pela Gama	61.04169	20.73326	10.41275	0.9152294	8.803455	0.7591368

Fonte: elaborada pelo autor.

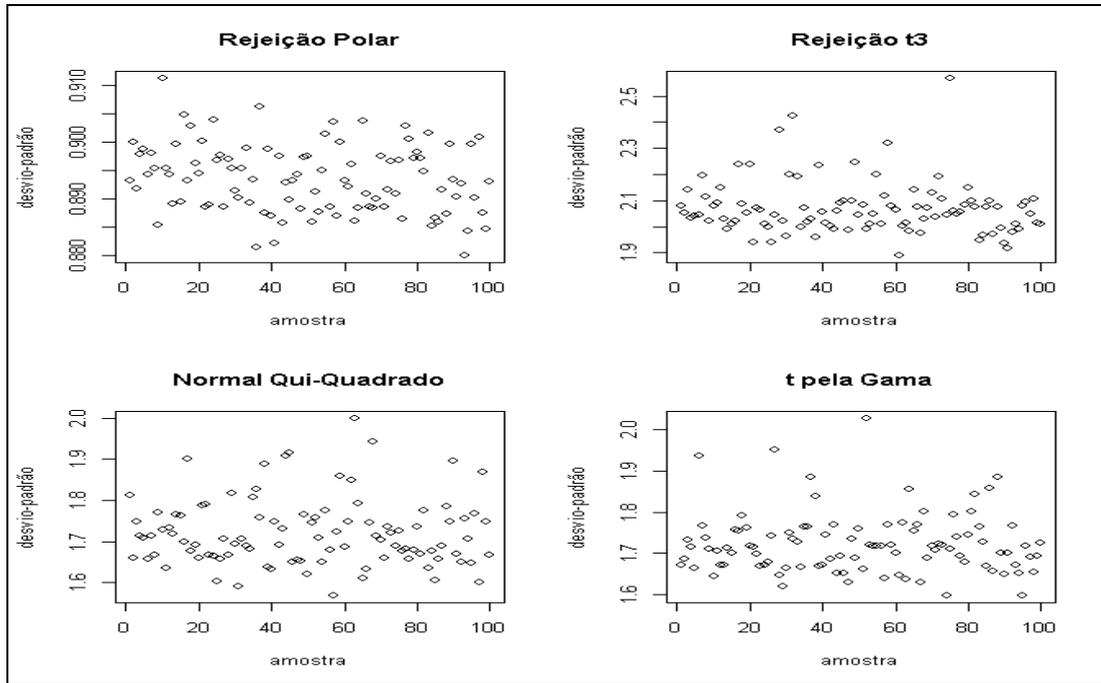
12.2.5 Desvio-Padrão

Tabela 112 - Média e erro padrão dos valores estimados para o desvio-padrão da distribuição t de Student em 100 amostras de tamanho 10000

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	0.8935672	0.006022852	0.6986402	0.003684291	0.6828484	0.003011036
Rejeição t(3)	2.070406	0.1038779	2.05683	0.113693	2.042289	0.1015588

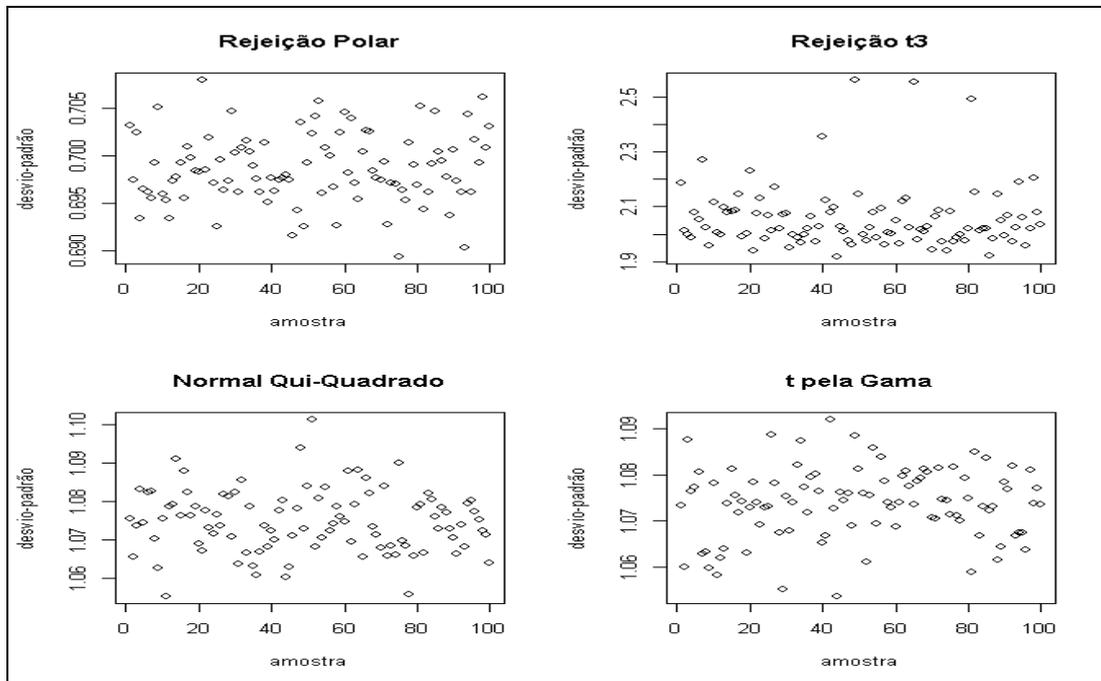
Normal/Qui-Quadrado	1.720892	0.0822379	1.074816	0.008093036	1.035618	0.007173984
t pela Gama	1.722826	0.07306908	1.0738	0.007637394	1.034485	0.007607595

Fonte: elaborada pelo autor.



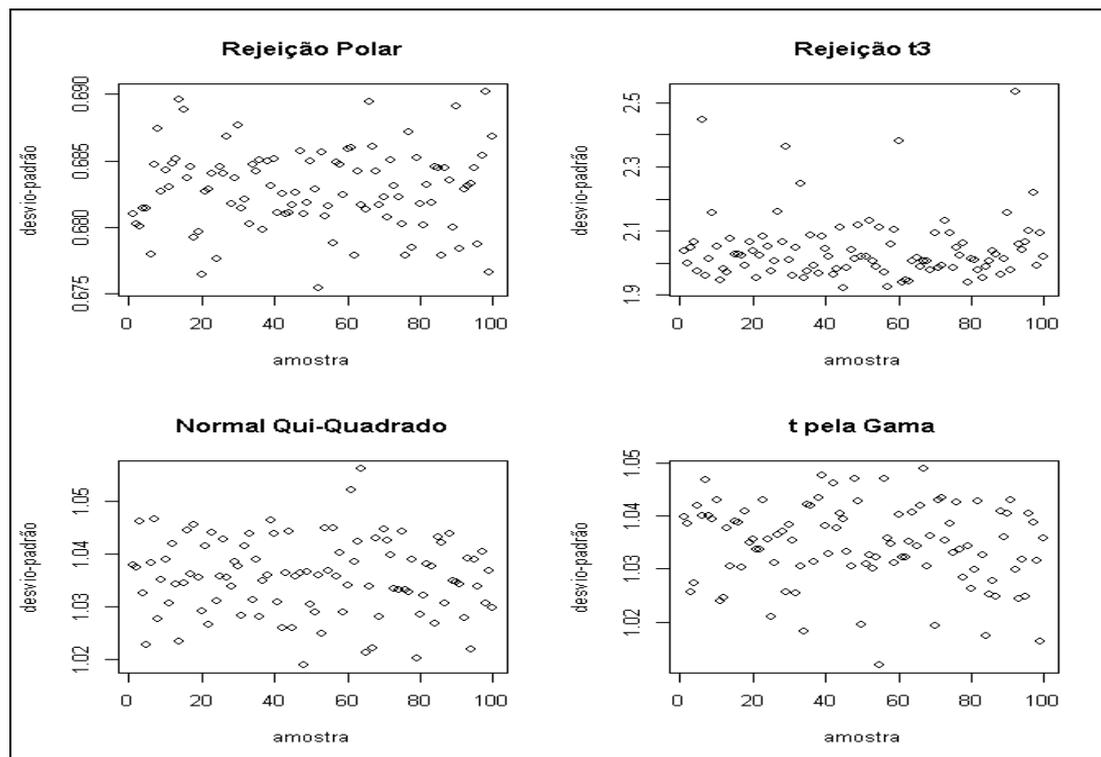
Quadro 64 - Gráficos dos valores estimados para o desvio-padrão da distribuição t(3) em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 65 - Gráficos dos valores estimados para o desvio-padrão da distribuição t(15) em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 66 - Gráficos dos valores estimados para o desvio-padrão da distribuição $t(30)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

12.2.6 Assimetria

Tabela 113 – Média e erro padrão dos valores estimados para a assimetria da distribuição t de Student em 100 amostras de tamanho 10000

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	N.D.	N.D.	0.001054805	0.01271668	0.001508181	0.01345761
Rejeição t(3)	N.D.	N.D.	-0.1156953	3.772117	0.2033567	3.779235
Normal/Qui-Quadrado	N.D.	N.D.	0.0009490708	0.03928421	-0.001829655	0.02430135
t pela Gama	N.D.	N.D.	0.003879621	0.03441255	0.002563008	0.03483463

Fonte: elaborada pelo autor.

12.2.7 Curtose

Tabela 114 – Média e erro padrão dos valores estimados para a curtose da distribuição t de Student em 100 amostras de tamanho 10000

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	-0.3682658	0.04075416	-1.052805	0.01240015	-1.080457	0.01142997
Rejeição t(3)	88.95895	242.075	80.90279	194.9372	73.8449	181.1188
Normal/Qui-Quadrado	78.50295	137.6988	0.5315311	0.1050083	0.237296	0.06950506

t pela Gama	72.95067	114.4952	0.5659147	0.1075562	0.2336864	0.07741289
-------------	----------	----------	-----------	-----------	-----------	------------

Fonte: elaborada pelo autor.

12.2.8 Quartis

Tabela 115 – Média e erro padrão dos valores estimados para o 1º quartil da distribuição t de Student em 100 amostras de tamanho 10000

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	-0.6419352	0.01278373	-0.5724046	0.009437525	-0.5647692	0.01004503
Rejeição t(3)	-1.117401	0.01702514	-1.115255	0.01696166	-1.117387	0.01812075
Normal/Qui-Quadrado	-0.7667135	0.01668055	-0.6934613	0.01511237	-0.6829064	0.01341986
t pela Gama	-0.7651948	0.01729666	-0.690652	0.01397572	-0.6828854	0.01454821

Fonte: elaborada pelo autor.

Tabela 116 – Média e erro padrão dos valores estimados para o 2º quartil da distribuição t de Student em 100 amostras de tamanho 10000

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	-0.001015602	0.01051489	-0.001316497	0.01010447	-0.0009050616	0.01059656
Rejeição t(3)	-0.002879813	0.02831074	0.004500296	0.03049669	0.001071381	0.02769272
Normal/Qui-Quadrado	-0.001869124	0.01213952	-0.002282075	0.01313385	-0.0003864402	0.0114158
t pela Gama	-0.0006331813	0.01369761	-0.001219858	0.01257827	-0.001640086	0.0125983

Fonte: elaborada pelo autor.

Tabela 117 – Média e erro padrão dos valores estimados para o 3º quartil da distribuição t de Student em 100 amostras de tamanho 10000

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	0.6399353	0.01265066	0.5705428	0.01042236	0.5631618	0.0104577
Rejeição t(3)	1.119622	0.01640077	1.117173	0.01660609	1.116966	0.01661918
Normal/Qui-Quadrado	0.7625466	0.01531475	0.6881182	0.01354994	0.682698	0.01396843
t pela Gama	0.7654078	0.0170889	0.6884739	0.01559257	0.680062	0.01322865

Fonte: elaborada pelo autor.

12.2.9 Distância Interquartilica

Tabela 118 – Média e erro padrão dos valores estimados para a distância interquartilica da distribuição t de Student em 100 amostras de tamanho 10000

Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	1.28187	0.01534214	1.142947	0.01298505	1.127931	0.01010873
Rejeição t(3)	2.237023	0.01893425	2.232428	0.01996053	2.234353	0.01992724
Normal/Qui-Quadrado	1.52926	0.01652522	1.381579	0.01570259	1.365604	0.01674402

t pela Gama	1.530603	0.02061373	1.379126	0.01717452	1.362947	0.01569743
-------------	----------	------------	----------	------------	----------	------------

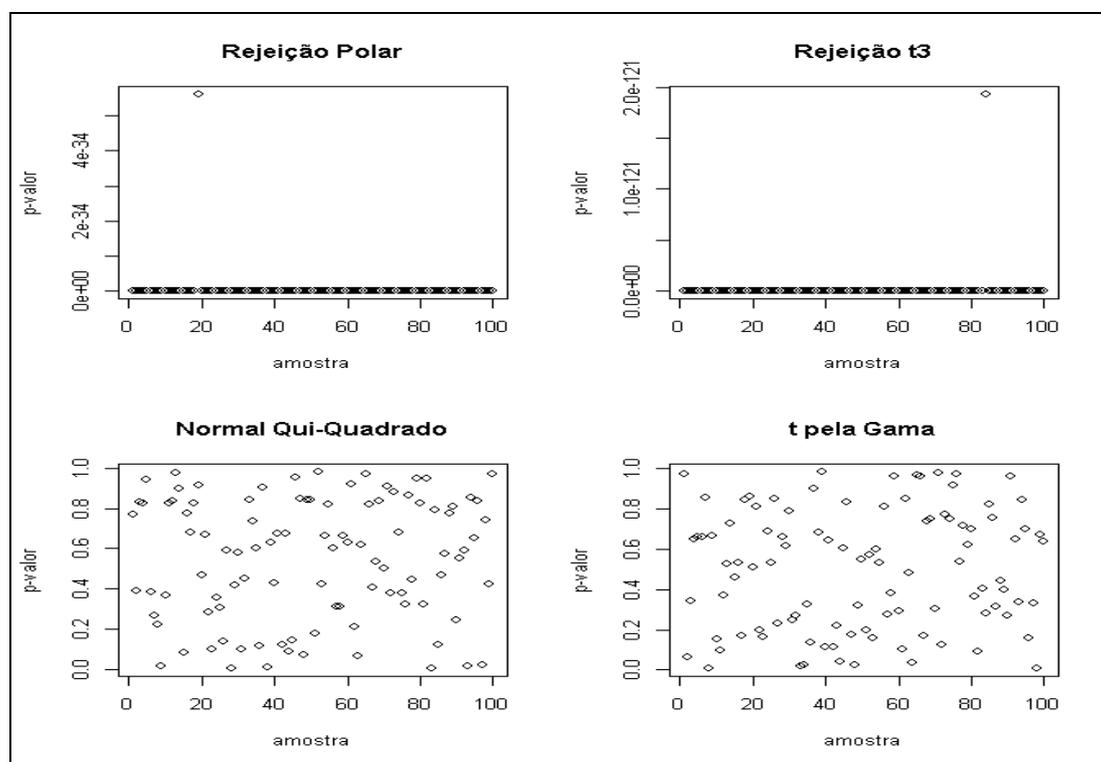
Fonte: elaborada pelo autor.

12.2.10 Teste de Aderência

Tabela 119 – Média e erro padrão dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição t de Student em 100 amostras de tamanho 10000

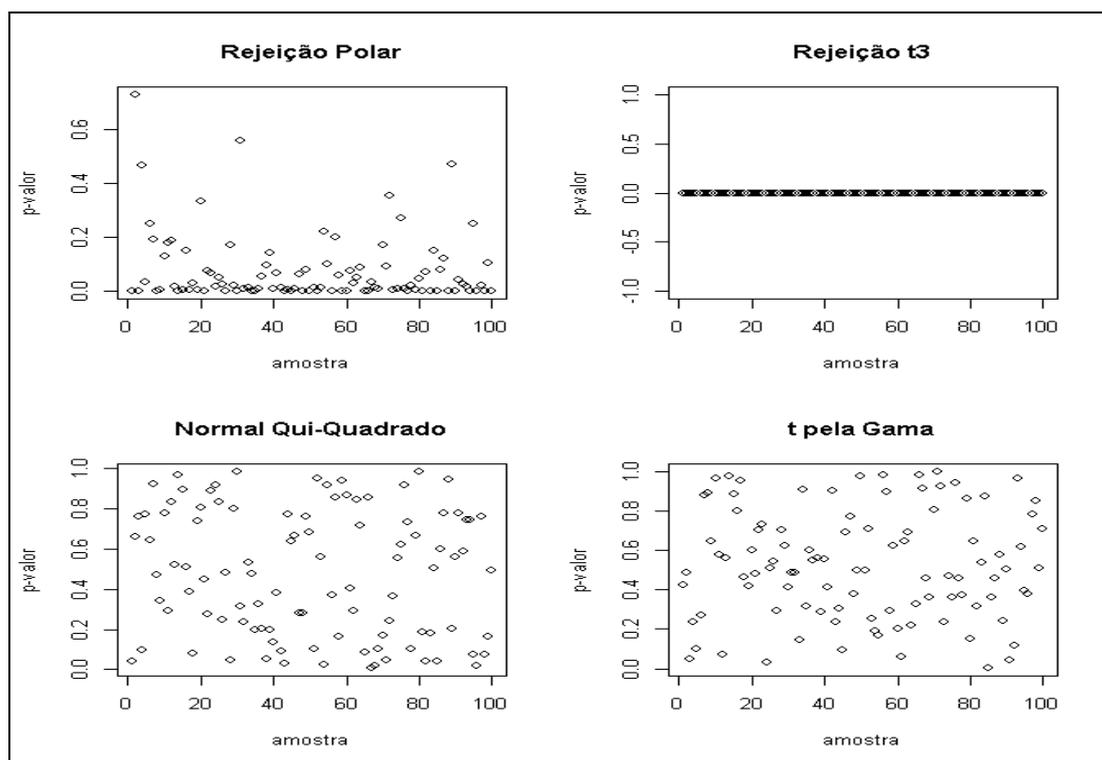
Gerador	t(3)		t(15)		t(30)	
	Média	E.P.	Média	E.P.	Média	E.P.
Rejeição Polar	5.622707e-36	5.622696e-35	0.07569451	0.1276354	0.04732809	0.1040282
Rejeição t(3)	1.933541e-123	1.931955e-122	0	0	0	0
Normal/Qui-Quadrado	0.5472772	0.3013079	0.4764712	0.3132472	0.495391	0.3019421
t pela Gama	0.4960554	0.2976483	0.5231738	0.2762151	0.4863203	0.280924

Fonte: elaborada pelo autor.



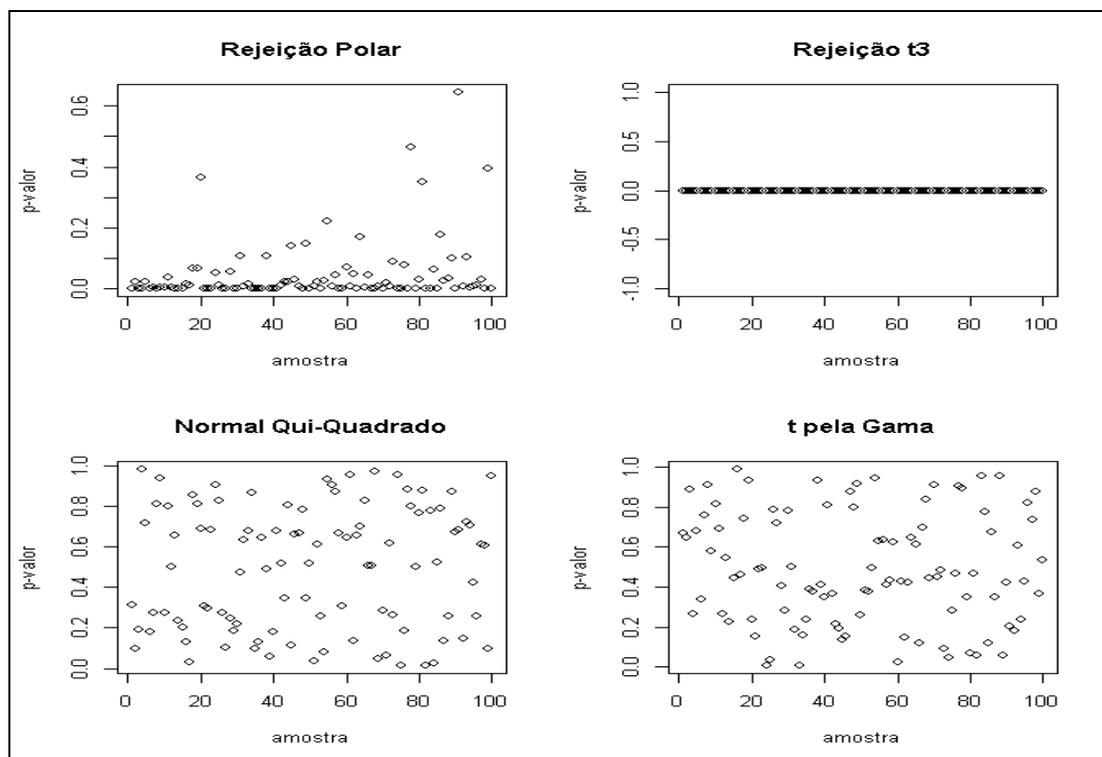
Quadro 67 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição t(3) em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 68 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $t(15)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.



Quadro 69 - Gráficos dos p-valores do Teste Qui-Quadrado de Aderência para uma distribuição $t(30)$ em 100 amostras de tamanho 10000

Fonte: Elaborado pelo autor.

12.2.11 Conclusões

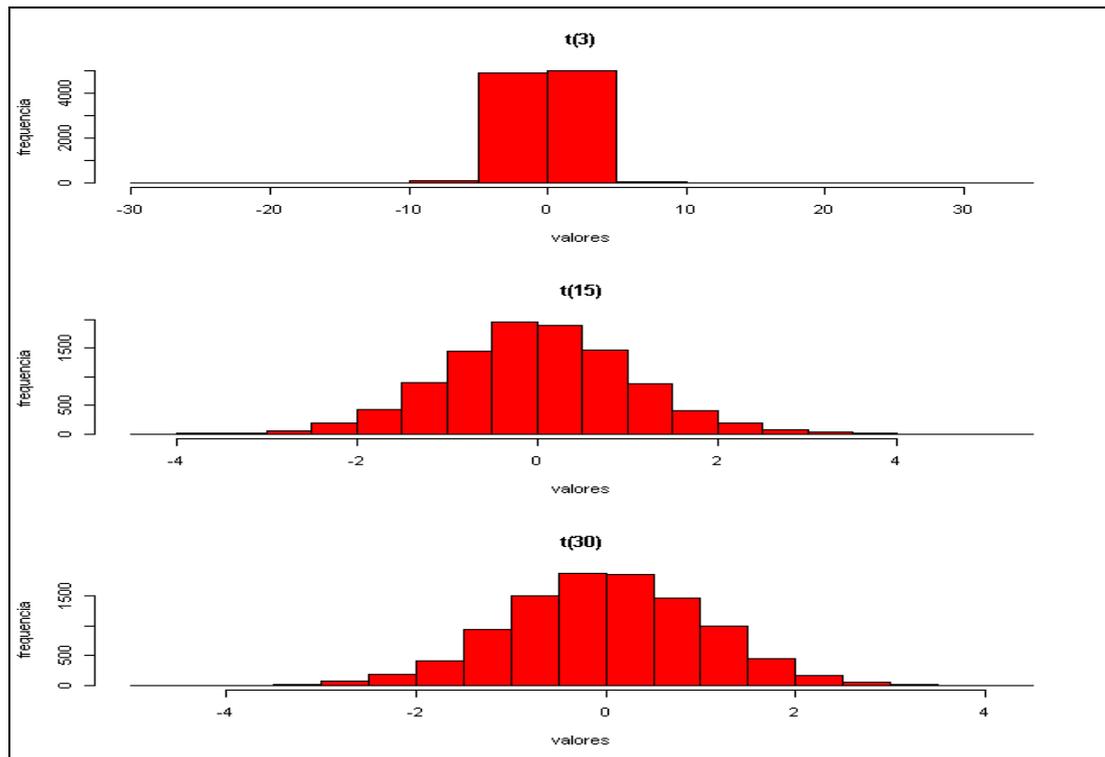
Desta vez, tivemos um resultado mais equilibrado entre os algoritmos Normal/Qui-Quadrado e t pela Gama. Qualquer um desses algoritmos poderia ser utilizado para gerar amostras com tamanho $n = 10000$ da Distribuição t sem prejuízos nos resultados. Entretanto, podemos observar uma leve vantagem para o algoritmo Normal/Qui-Quadrado, pois suas estimativas, mais uma vez, chegaram um pouco mais próximas aos valores exatos se comparadas às estimativas do algoritmo t pela Gama. Além do mais, o algoritmo Normal/Qui-Quadrado mostrou-se um pouco mais rápido que o algoritmo t pela Gama, pois seu tempo de processamento foi um pouco menor. Sendo assim, indicamos o algoritmo Normal/Qui-Quadrado como gerador de amostras com tamanho $n = 10000$ da Distribuição t de Student.

Os dois algoritmos restantes (Rejeição Polar e Rejeição t3), não passaram no Teste Qui-Quadrado de Aderência e, assim, o seu uso como gerador da Distribuição t não é recomendado.

Tabela 120 – Estimativas médias do algoritmo Normal/Qui-Quadrado em 100 amostras de tamanho 10000 da Distribuição t de Student

Medidas	Normal/Qui-Quadrado		
	t(3)	t(15)	t(30)
Tempo de Execução (s)	5.5548	5.2568	4.9097
Esperança	-0.002243976	-0.0027243	-0.0002610586
Máximo	31.99661	5.20627	4.472856
Mínimo	-29.19614	-5.177636	-4.449784
Amplitude	61.19275	10.38391	8.92264
Desvio-Padrão	1.720892	1.074816	1.035618
Assimetria	N.D.	0.0009490708	-0.001829655
Curtose	78.50295	0.5315311	0.237296
1º Quartil	-0.7667135	-0.6934613	-0.6829064
2º Quartil	-0.001869124	-0.002282075	-0.0003864402
3º Quartil	0.7625466	0.6881182	0.682698
Dist. Interquart.	1.52926	1.381579	1.365604

Fonte: elaborada pelo autor.



Quadro 70 – Histogramas de 10000 valores gerados da Distribuição t de Student pelo algoritmo Normal/Qui-Quadrado

Fonte: Elaborado pelo autor.

13. ALGORITMOS DISPONÍVEIS NO SOFTWARE R

Dentre todos os geradores apresentados neste trabalho, os algoritmos listados abaixo encontram-se disponíveis no software R, versão 2.14.2, não sendo necessária a sua implementação.

- Para a Distribuição Uniforme:
 1. Mersenne Twister (padrão).
- Para a Distribuição Normal:
 1. Inversão por Wichura (padrão).
 2. Kinderman-Ramage.
 3. Box-Muller.
 4. Ahrens-Dieter.
- Para a Distribuição Qui-Quadrado:
 1. Gama – R.
- Para a Distribuição F de Snedecor:
 1. F pela Qui-Quadrado.
- Para a Distribuição t de Student:
 1. Normal/Qui-Quadrado.

Os demais algoritmos não estão disponíveis no software R e, caso se deseje utilizá-los, deve-se proceder a sua implementação. Além do Mersenne Twister, o software também disponibiliza outros geradores para a Distribuição Uniforme, mas estes não serão tratados neste trabalho.

14. ALGORITMOS “VENCEDORES” PARA CADA DISTRIBUIÇÃO

A tabela a seguir resume os algoritmos “vencedores” para cada distribuição.

Tabela 121 – Algoritmos “vencedores” para cada distribuição

Distribuição	Tamanho de amostra (n)	Algoritmo “vencedor”
Normal	100	Kinderman-Monahan
	1000	Rejeição pela Distr. Exponencial
	10000	Box-Muller
Qui-Quadrado	100	Logaritmo Produto Uniformes
	1000	Soma Quadrados Normal
	10000	Logaritmo Produto Uniformes
F de Snedecor	10000	F pela Beta
t de Student	100	Normal/Qui-Quadrado
	10000	Normal/Qui-Quadrado

Fonte: elaborada pelo autor.

15. CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

Neste trabalho, tivemos a oportunidade de reunir, ao todo, 25 algoritmos para geração de distribuições de probabilidade: 13 para a Distribuição Normal, seis para a Distribuição Qui-Quadrado, dois para a Distribuição F e quatro para a Distribuição t de Student. A maioria desses algoritmos pode, seguramente, ser utilizada para gerar amostras da distribuição a qual se propõe. Além das distribuições Normal, Qui-Quadrado, F e t de Student não serem integráveis analiticamente, a maior parte dos algoritmos geradores desses modelos encontra-se espalhada pela literatura de Estatística Computacional e Simulação, sendo difícil encontrar uma obra que os reúna. Daí, resulta a relevância deste trabalho, pois aqui conseguimos agrupar todos esses algoritmos, o que pode se constituir em uma boa fonte de consulta para trabalhos futuros.

Para aprimorar ou estender este trabalho, pode-se, por exemplo, repetir às análises realizadas utilizando o software R para plataforma Linux, aumentar o número de replicações e utilizar tamanhos de amostra diferentes ou repetir às análises usando outro(s) gerador(es) da Distribuição Uniforme [0; 1].

REFERÊNCIAS

- BUSSAB, W.O., MORETTIN, P.A. **Estatística Básica – 5ª edição**. São Paulo: Saraiva, 2002.
- BUSTOS, O.H., FRERY, A.C. **Simulação Estocástica: Teoria e Algoritmos**. Rio de Janeiro: IMPA, 1992.
- Código-fonte do software R v.2.14.2. 2012**. Disponível em:
<<http://cran-r.c3sl.ufpr.br/src/base/R-2/>>. Acesso em: 24 abr. 2012.
- COSTA NETO, P.L.O. **Estatística – 2ª edição**. São Paulo: Blücher, 2002.
- DEVROYE, L. **Non-Uniform Random Variate Generation**. New York: Springer, 1986.
- FERREIRA, D.F. **Estatística Computacional Utilizando R**. 2010. Disponível em:
<<http://www.dex.ufla.br/~danielff/apeco.pdf>>. Acesso em: 25 mar. 2012.
- FRERY, A.C., CRIBARI-NETO, F. **Elementos de Estatística Computacional Usando Plataformas de Software Livre/Gratuito**. Rio de Janeiro: IMPA, 2005.
- GENTLE, J.E. **Random Number Generation and Monte Carlo Methods**. New York: Springer, 1998.
- INSÚA, D.R., INSÚA, S.R., JIMÉNEZ, J.M., MARTÍN, A.J. **Simulación. Métodos y aplicaciones, 2ª edición**. Ciudad del México: Alfaomega, 2009.
- JONES, O., MAILLARDET, R., ROBINSON, A. **Scientific Programming and Simulation Using R**. New York: Chapman & Hall, 2009.
- MATSUMOTO, M., NISHIMURA, T. **Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator**. 1998. Disponível em:
<<http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/ARTICLES/mt.pdf>>. Acesso em: 07 jul. 2012.
- Normas de Apresentação Tabular – 3ª edição**. Rio de Janeiro: IBGE, 1993. Disponível em:
<<http://biblioteca.ibge.gov.br/visualizacao/monografias/GEBIS%20-%20RJ/normastabular.pdf>>. Acesso em: 07 jul. 2012.
- THOMAS, D.B., LUK, W., LEONG, P.H.W., VILLASENOR, J.D. **Gaussian Random Number Generators**. 2007. Disponível em:
<http://www.cse.cuhk.edu.hk/~phwl/mt/public/archives/papers/grng_acmcs07.pdf>. Acesso em: 18 abr. 2012.
- VIALI, L., BITTENCOURT, H.R. **As distribuições de probabilidade T, F e Qui-Quadrado: teoria e prática com o uso da planilha**. In: ENCONTRO NACIONAL DE EDUCAÇÃO MATEMÁTICA. Anais... Belo Horizonte, 2007.
- VIEIRA, C.E.C., RIBEIRO, C.C., CASTRO E SOUZA, R. **Geradores de números aleatórios**. 2004. Disponível em: <ftp://ftp.inf.puc-rio.br/pub/docs/techreports/04_22_vieira.pdf>. Acesso em: 28 mar. 2012.

WIKIPEDIA. The Free Encyclopedia.

<http://en.wikipedia.org/wiki/Chi-squared_distribution>. Acesso em: 18 mar. 2012.

<http://en.wikipedia.org/wiki/Mersenne_Twister>. Acesso em: 07 jul. 2012.

<http://en.wikipedia.org/wiki/Normal_distribution>. Acesso em: 18 mar. 2012.

<http://en.wikipedia.org/wiki/Snedecor%27s_F_distribution>. Acesso em: 18 mar. 2012.

<http://en.wikipedia.org/wiki/Student%27s_t-distribution>. Acesso em: 18 mar. 2012.

<http://pt.wikipedia.org/wiki/Primo_de_Mersenne>. Acesso em: 07 jul. 2012.