

ESTUDO DOS ERROS TIPOS I E II PARA TESTES DE ADERÊNCIA UTILIZANDO SÉRIES SINTÉTICAS

Danilo Nogueira de Souza¹; Ticiane Marinho de Carvalho Studart³; José Nilson B. Campos³; Marcelo Leão Nogueira²; Renata Mendes Luna³

RESUMO --- O artigo se propõe a estudar a sensibilidade de dois testes de aderência – Kolmogorov-Smirnov e Qui-Quadrado – no ajuste de uma série de dados a uma determinada distribuição de probabilidade. A sensibilidade dos testes de ajustamento foi medida através dos erros tipos I e II. Foram utilizadas várias séries sintéticas com diversas extensões, as quais seguiam uma distribuição de probabilidade conhecida. Estas séries foram submetidas aos dois testes, os quais avaliaram o ajustamento a quatro distribuições de probabilidade: Gama, Gumbel, Lognormal e Normal. A sensibilidade também considerou a variabilidade dos dados, representada no coeficiente de variação. Concluiu-se que o teste de Kolmogorov-Smirnov apresenta erros do tipo II maiores do que o teste do Qui-Quadrado, para todos os coeficientes de variação utilizados. Porém, o teste de Kolmogorov-Smirnov apresenta erros do tipo I nulos. Ou seja, para todos os coeficientes de variação analisados, este teste sempre “reconhece” a distribuição real da série. Foi observado que a extensão das séries tem grande influência na sensibilidade dos testes de aderência. Dependendo do coeficiente de variação, a rejeição de uma distribuição falsa, para ambos os testes, só ocorre com séries muito extensas. Em outros casos, nenhum dos dois testes apresenta sensibilidade para rejeitar uma distribuição falsa.

ABSTRACT --- This paper intends to study the sensibility of two goodness of fit tests – Kolmogorov-Smirnov and Chi-Square – when a set of data is compared with a certain probability distribution. The sensibility for these tests was analyzed using Type I and Type II errors. Several synthetic series with a large number of values, generated by a known distribution (Gamma), were used. These series were submitted to the tests, which assessed the fitting of the data with four distributions: Gamma, Gumbel, Lognormal and Normal. Also, the sensibility analysis considered the variability of the data, represented by the coefficient of variation. In the end, it was found that Kolmogorov-Smirnov test shows higher Type II errors (acceptance of the false distribution) than the Chi-Square test for all the coefficients of variation. But, Kolmogorov-Smirnov test shows zero Type I errors (rejection of the true distribution). It means that this test always “recognizes” the set’s real distribution. It was observed that the series’ number of values greatly influences on the tests’ sensibility. Depending on the coefficient of variation, the rejection of a false distribution, for both tests, occurs only for very large series. In other cases, none of the tests is able to reject the false distribution.

Palavras-chave: testes de aderência, distribuição de probabilidades, séries sintéticas

1 - INTRODUÇÃO

Um problema importante e recorrente em estudos hidrológicos é a escolha da distribuição de probabilidades que se ajusta à variável hidrológica, em especial, a eventos extremos. Encontra-se na literatura vários estudos que abordam o tema, principalmente no que se refere a cheias, entre os quais se pode citar Benson (1968), Wallis (1988) e Vogel *et al.* (1993), que estudaram diversas distribuições de probabilidades para as cheias nos Estados Unidos; Rossi *et. al.* (1984), que estudaram as séries de cheias italianas e, principalmente, Cunnane (1989), em um estudo detalhado para a World Meteorological Organization (WMO).

Na prática, baseando-se na experiência, seleciona-se uma distribuição de probabilidades que se sabe que usualmente se ajusta a uma dada variável hidrológica, estima-se seus parâmetros pelos dados amostrais, e testa-se o ajustamento por meio de um teste de aderência. Se satisfatório, assume-se que esta mesma função se ajusta à população desta mesma variável hidrológica. Muitas vezes não apenas uma, mas duas ou três funções distribuições de probabilidade passam no teste de aderência e cabe ao hidrólogo a missão de escolher uma delas.

No processo de escolha da distribuição de probabilidades, através do teste de hipóteses, o tomador de decisão pode vir a cometer um dos dois tipos de erro:

Erro tipo I: quando o teste rejeita uma hipótese verdadeira. Ou seja, os valores empíricos têm o seu comportamento descrito, de fato, pelo modelo teórico testado, sendo este, contudo, considerado inadequado pelo teste.

Erro tipo II: quando o teste não rejeita uma hipótese falsa. Ou seja, os valores empíricos, na verdade, não são descritos pelo modelo testado, sendo este, entretanto, considerado adequado pelo teste.

Na realidade, uma vez que nunca se saberá a distribuição de probabilidades *verdadeira*, não há como saber qual erro se está cometendo.

Neste trabalho gerou-se aleatoriamente uma amostra, seguindo uma distribuição de Probabilidades Gama II e analisou-se o desempenho de dois testes clássicos de aderência - Kolmogorov-Smirnov (K-S) e Qui-Quadrado (Q-Q) em reconhecer a distribuição verdadeira. Testou-se a aderência das distribuições de probabilidades Gama II, Gumbel, Normal e Lognormal aos dados gerados.

Para se compreender a sensibilidade dos testes ao tamanho da amostra e a variabilidade natural da série dos dados, gerou-se séries sintéticas com coeficientes de variação iguais a 0,1, 0,5,

1,0 e 1,5 (e mesma média). Quanto à extensão das séries, analisou-se, para cada CV, 30 séries de 50, 100, 200, 300, 400, 500, 600, 700, 800, 900 e 1000 dados.

2 - MATERIAIS E MÉTODOS

A metodologia utilizada foi a que se segue.

2.1 - Distribuições de probabilidade utilizadas

As distribuições de probabilidade testadas neste trabalho foram: Gama, Gumbel, Lognormal e Normal.

2.1.1 - Distribuição Gama

A distribuição Gama tem a seguinte função densidade:

$$f(x) = \frac{x^{(\alpha-1)} (e)^{\left(-\frac{x}{\beta}\right)}}{\beta^\alpha \Gamma(\alpha)} \quad \text{para } x > 0, \alpha > 0, \beta > 0; \text{ e } f(x) = 0 \text{ para outros valores.} \quad (\text{Eq.1})$$

Sendo $\Gamma(\alpha)$ o valor da distribuição Gama, a qual é definida por:

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

Os parâmetros α (valor inteiro positivo) e β são estimados da amostra dos dados, segundo Yevjevich, (1972). A média e a variância da distribuição Gama são dadas, respectivamente, por:

$$\mu = \alpha\beta \quad (\text{Eq. 3})$$

e

$$\sigma^2 = \alpha\beta^2 \quad (\text{Eq. 4})$$

Em hidrologia, a distribuição Gama é usada para o ajuste de distribuições de frequências de precipitações anuais e deflúvios anuais.

2.1.2 - Distribuição de Gumbel

Segundo Yevjevich (1972), as distribuições de valores extremos têm um interesse particular e são de grande importância em hidrologia, sendo frequentemente associadas com cheias e secas, suprimentos ou déficits.

A distribuição de Gumbel é uma distribuição de extremos do tipo exponencial, a qual é dada pela equação 5:

$$f(x) = e^{(-e^{(-\alpha(x-\beta))})} \quad (\text{Eq. 5})$$

Para a condição: $\alpha > 0$.

Onde: α é um parâmetro de posição e β é um parâmetro de escala.

2.1.3 -Distribuição Normal

A distribuição Normal é considerada a distribuição de probabilidade mais importante, uma vez abrange uma grande quantidade de fenômenos.

$$f(x) = \frac{(e)^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}}{\sigma\sqrt{2\pi}} \quad (\text{Eq. 6})$$

Para todo $X \in Z$.

Onde: μ = média da distribuição;

σ = desvio-padrão da distribuição.

2.1.4 -Distribuição Lognormal

Segundo Yevjevich (1972), “quando os logaritmos naturais, $\ln x$, de uma variável x são normalmente distribuídos, então diz-se que a distribuição de x segue uma distribuição de probabilidade Lognormal. A distribuição de probabilidade Lognormal para a variável $\ln x$ é dada pela equação 7:

Para $x \geq 0$:

$$f(x) = \frac{(e)^{\left(-1/2 \frac{(\ln x - \mu_n)^2}{\sigma_n^2}\right)}}{x \sigma_n \sqrt{2\pi}} \quad (\text{Eq. 7})$$

Para $x < 0$: $f(x) = 0$.

O índice n em μ_n e σ_n significa que estes parâmetros são determinados dos logaritmos naturais da variável aleatória. As relações entre as médias e as variâncias da variável aleatória e dos seus logaritmos naturais são dadas pelas equações 8 e 9:

$$\mu = (\mathbf{e})^{\left(\mu_n + 1/2 \sigma_n^2\right)} \quad (\text{Eq. 8})$$

$$\sigma^2 = \mu^2 \left((\mathbf{e})^{\left(\sigma_n^2\right)} - 1 \right) \quad (\text{Eq. 9})$$

Onde μ e σ^2 são, respectivamente, a média e o desvio-padrão da variável aleatória.

Assim como a distribuição Gama, a Lognormal também é usada para ajustar precipitações e deflúvios anuais.

2.2 - Testes de hipóteses

Para a realização dos testes de hipóteses - e conseqüente estudo dos erros - fez-se uso de dois testes clássicos de aderência, a saber: Kolmogorov-Smirnov e Qui-Quadrado. Em ambos os testes foi adotado um nível de significância igual a 5%.

2.2.1 - Qui-Quadrado

Segundo Johnson (2000), “para testar se as discrepâncias entre as frequências observadas e as esperadas podem ser atribuídas ao acaso, faz-se uso da estatística” da pela equação 10:

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i} \quad (\text{Eq. 10})$$

Onde o_i e e_i são, respectivamente, as frequências observadas e esperadas. A distribuição da amostra desta estatística segue aproximadamente uma distribuição Qui-Quadrado com $k - m$ graus de liberdade, onde k é o número de termos na fórmula para o χ^2 e m é o número de valores, obtidos dos dados observados, que são necessários para o cálculo das frequências esperadas.

2.2.2 - Kolmogorov-Smirnov

O teste de Kolmogorov-Smirnov é um teste não paramétrico baseado na máxima diferença absoluta $\mathbf{D}$ entre os valores de uma distribuição empírica de uma amostra aleatória de tamanho n e os valores de uma distribuição teórica específica.

Este teste é geralmente mais eficiente para a adequação do ajustamento do que o teste do Qui-Quadrado, para pequenas amostras, onde o teste do Qui-Quadrado não se aplica. Contudo, o teste de Kolmogorov-Smirnov não pode ser usado para distribuições discretas (Johnson, 2000).

2.3 - Recursos computacionais utilizados

Para a geração das séries testadas, utilizou-se o programa Simres versão 1.2, desenvolvido pelo Departamento de Engenharia Hidráulica e Ambiental (DEHA) da Universidade Federal do Ceará (UFC). Já as análises das séries foram realizadas com o uso do *software* estatístico Statgraphics versão 16.1.11.

3 - ANÁLISE DOS RESULTADOS

O erro do tipo I se refere somente ao teste de aderência da série sintética com a distribuição Gama (distribuição verdadeira); já o erro do tipo II foi analisado para os testes de ajustamento nas outras três distribuições testadas, que são falsas.

3.1 – Erros do Tipo I

Nas figura 1 a 4 são apresentados os resultados das análises dos erros do Tipo I, considerando o CV = 0,1; 0,5; 1,0 e 1,5; para o teste do Qui-Quadrado. O teste Kolmogorov-Smirnov não cometeu Erro do Tipo I, ou seja, para qualquer extensão da série, não rejeitou a hipótese verdadeira da distribuição ser Gama.

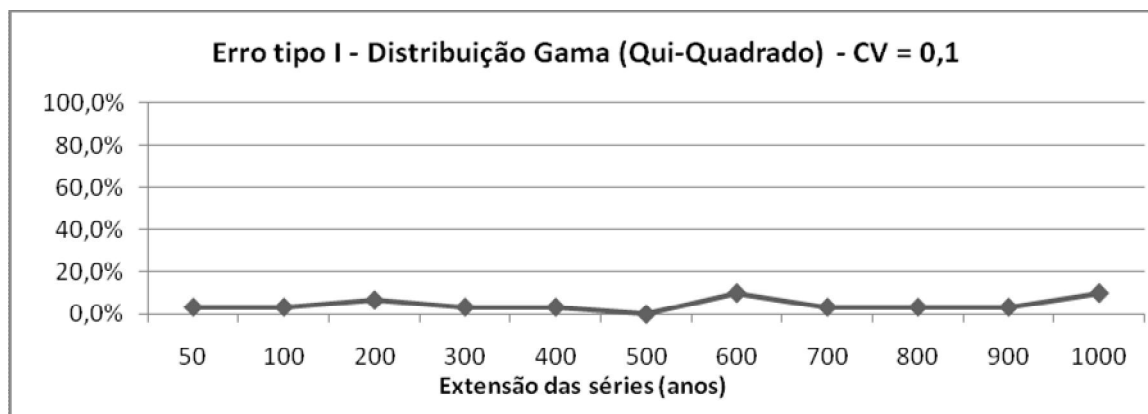


Figura 1 - Erro tipo I para o teste do Qui-Quadrado (distribuição Gama - CV = 0,1)

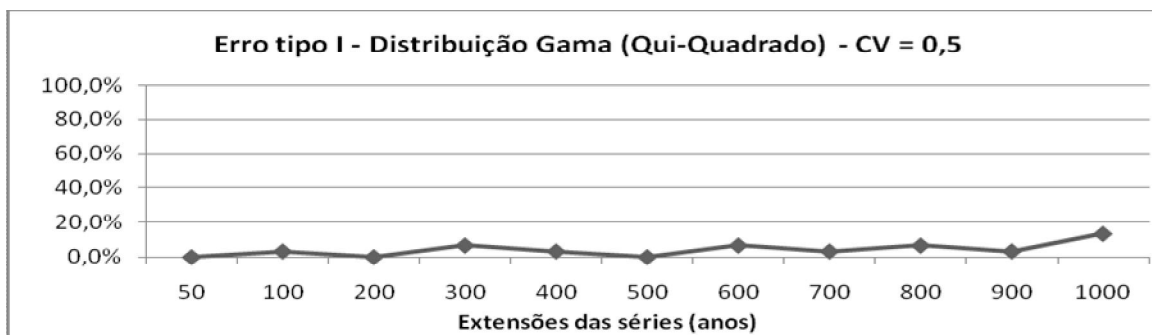


Figura 2 - Erro tipo I para o teste do Qui-Quadrado, considerando distribuição Gama (CV = 0,5)

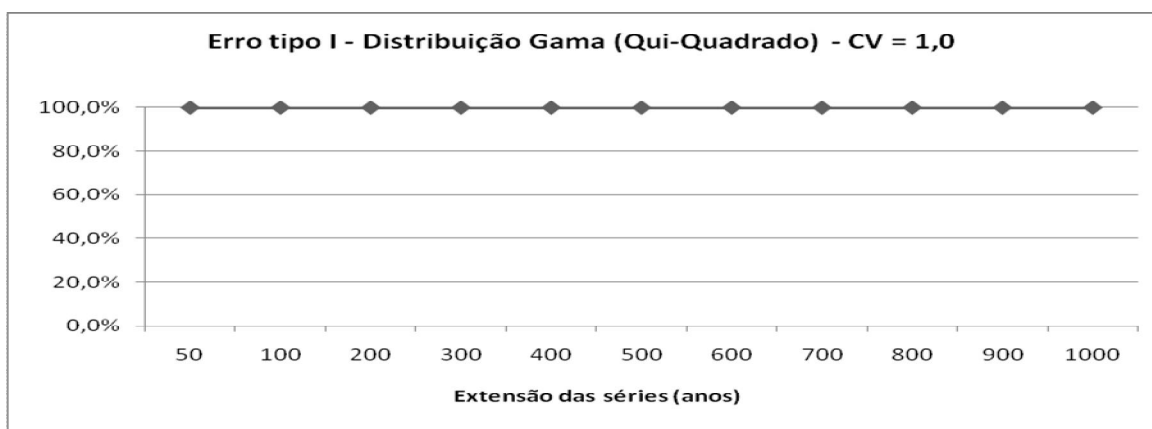


Figura 3 - Erro tipo I para o teste do Qui-Quadrado, considerando distribuição Gama (CV = 1,0)

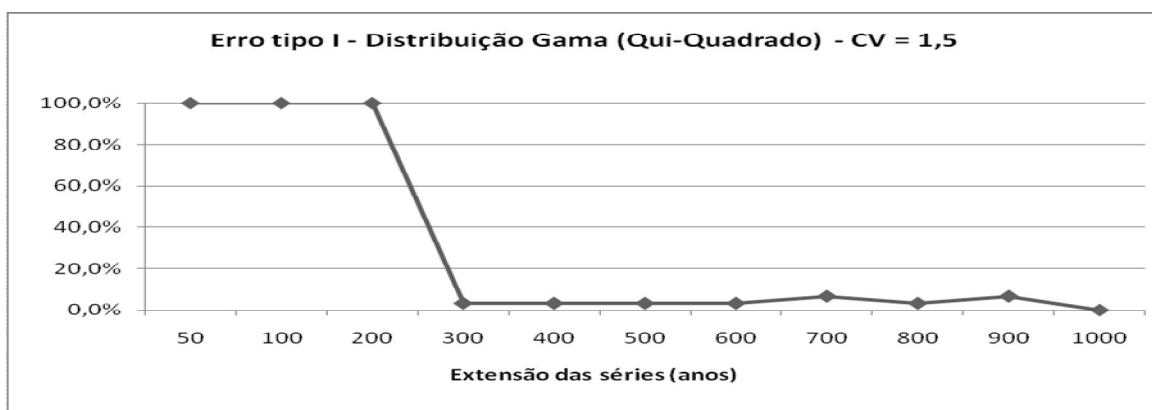


Figura 4 - Erro tipo I para o teste do Qui-Quadrado, considerando distribuição Gama (CV = 1,5)

Observa-se que o teste do Qui-Quadrado apresenta um comportamento bastante irregular no ajustamento da distribuição verdadeira (Gama) para baixos Cvs. Já no caso do CV igual a 1,0; comete erro do Tipo I em todas as 30 séries e finalmente, no CV igual a 1,5; começa a ter sensibilidade (ou seja, não comete mais erro do Tipo I, em todas as séries, a partir de da extensão da série (n) igual a 300.

3.2 – Erros do Tipo II – análise comparativa dos cvs

3.2.1 - Distribuição de Gumbel

As figuras 5 a 8 apresentam os resultados para as 30 séries com extensões de 50 a 1000 anos do testes Kolmogorov-Smirnov e Qui-Quadrado. Observa-se na Figura 5 que ambos os testes apresentaram valores elevados para o erro do Tipo II e que estes erros vão diminuindo com o tamanho da série, zerando a partir de n igual a 800 anos.

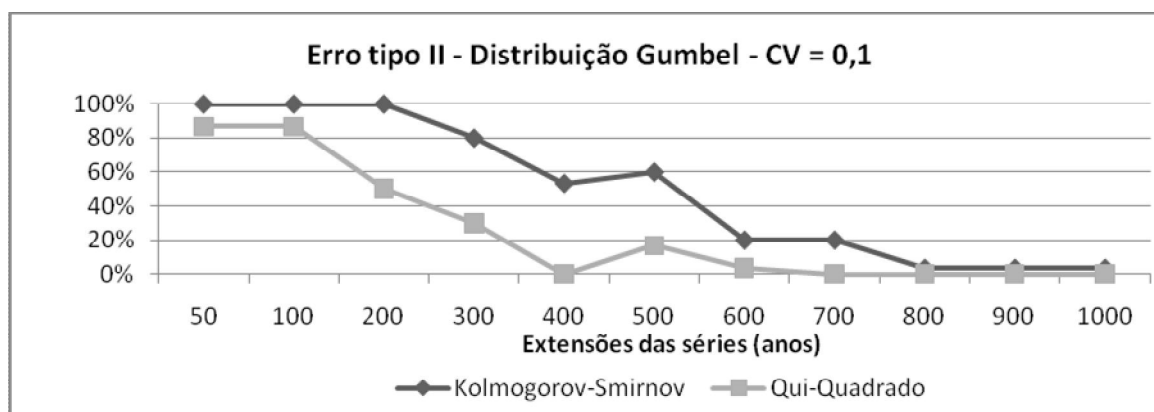


Figura 5- Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando distribuição Gumbel ($CV = 0,1$)

Observa-se na Figura 6, que, para praticamente todas as séries, em todas as extensões, ambos os testes não rejeitaram a hipótese falsa, cometendo o erro do Tipo II.

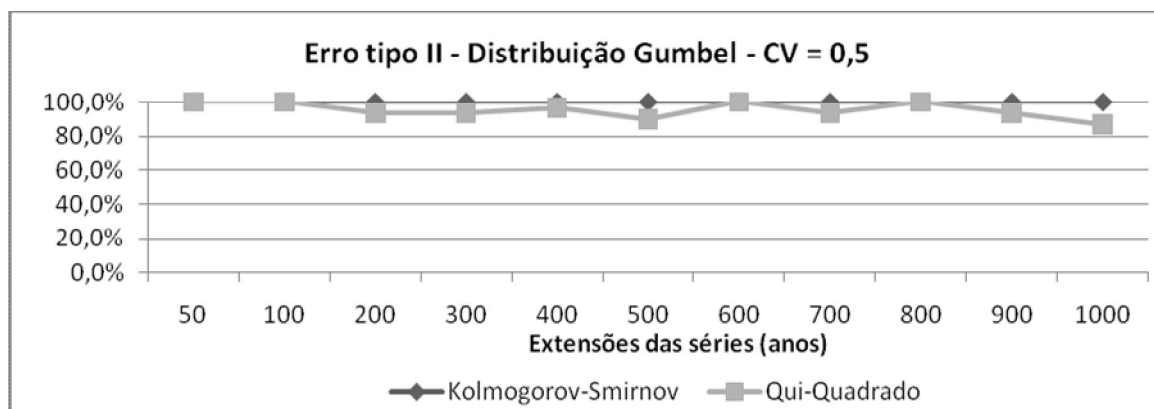


Figura 6- Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando uma distribuição Gumbel ($CV = 0,5$)

Observa-se, tanto na Figura 7 como na 8, um comportamento semelhante observado na Figura 5, ou seja, a diminuição do erro do Tipo II, em ambos os testes, com o aumento da extensão da séries e também com o aumento do CV. Este comportamento só não foi observado para CV igual a 0,5; que precisa ser melhor avaliado e pode ser fruto de algum problema na geração dos números aleatórios com tal variabilidade. Observa-se ainda que, para altos CVs, o teste do Qui-Quadrado

apresenta melhor desempenho, ou seja, não apresenta para nenhuma as séries, a partir de n igual a 50, erros do Tipo II.

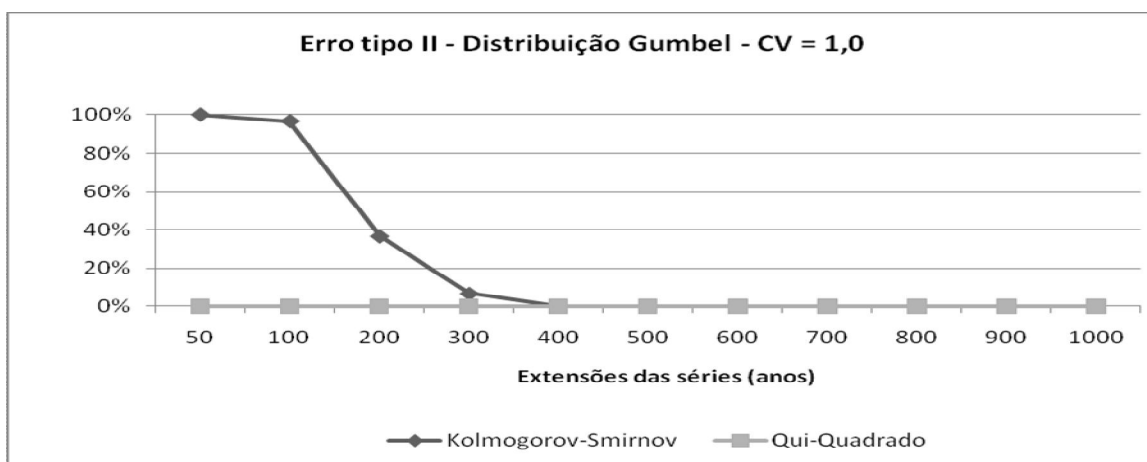


Figura 7 - Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando uma distribuição Gumbel ($CV = 1,0$)

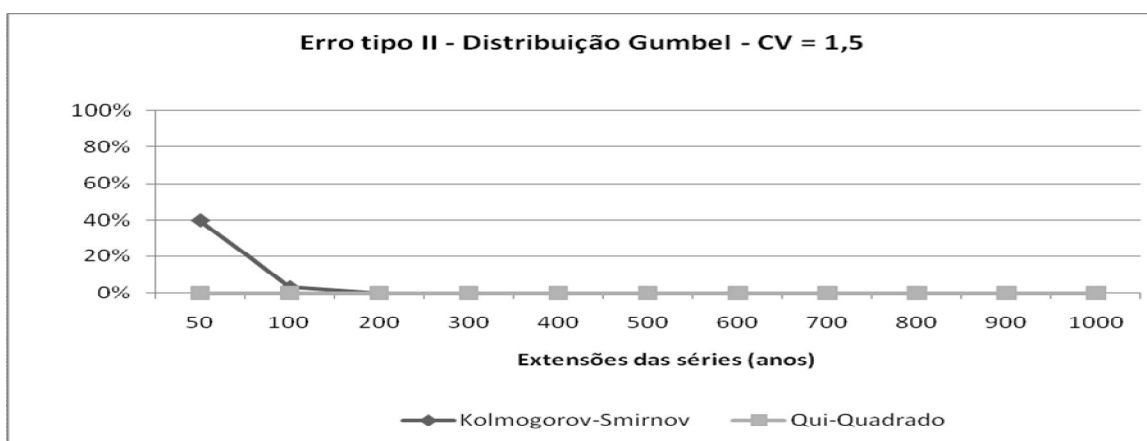


Figura 8 - Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando uma distribuição Gumbel ($CV = 1,5$)

3.2.2 - Distribuição Lognormal

As figuras 9 a 12 apresentam os resultados para as 30 séries com extensões de 50 a 1000 anos do testes Kolmogorov-Smirnov e Qui-Quadrado. Observa-se um comportamento semelhante ao da distribuição de Gumbel, ou seja, ambos os testes se tornam mais sensíveis em detectar a distribuição de probabilidade falsa (cometem menos o erro do Tipo II), a medida em que n e CV aumentam seus valores. Sendo, dentre os dois testes, o do Qui-Quadrado a ter melhor desempenho.

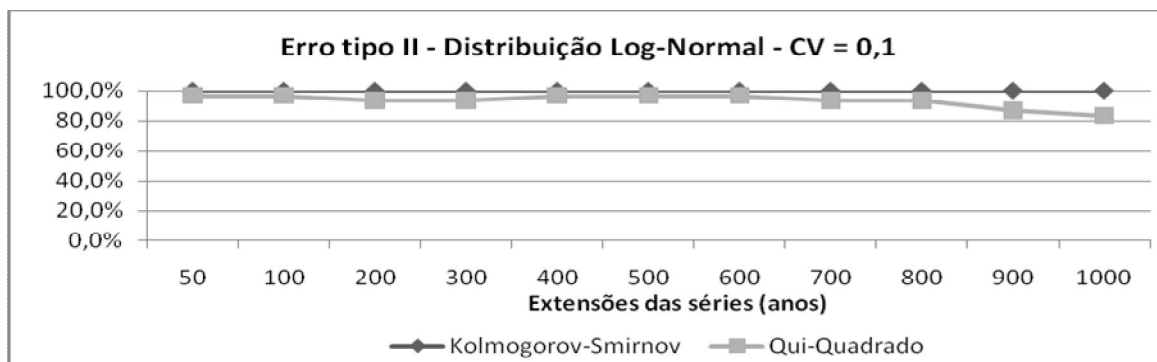


Figura 9 - Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando distribuição Lognormal (CV = 0,1)

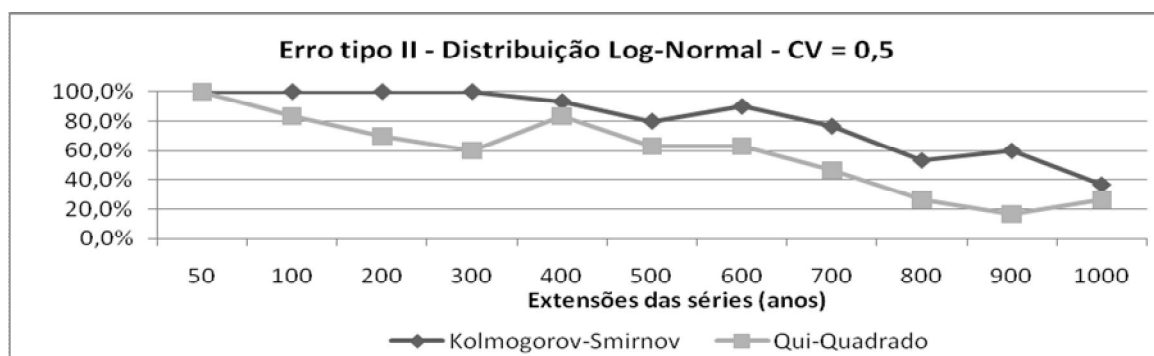


Figura 10 - Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando uma distribuição Lognormal (CV = 0,5)

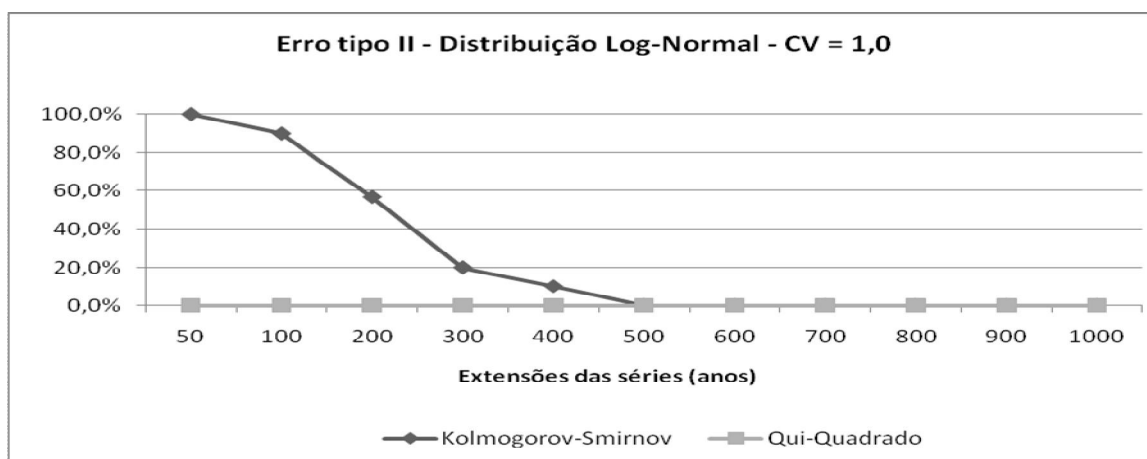


Figura 11 - Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando uma distribuição Lognormal (CV = 1,0)

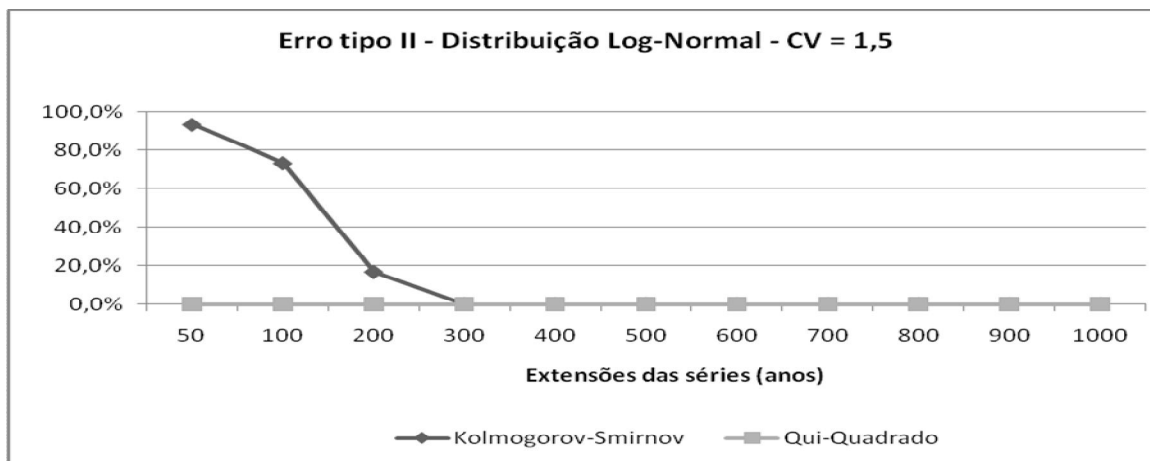


Figura 12 - Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando uma distribuição Lognormal (CV = 1,5)

3.2.3 - Distribuição Normal

As figuras 13 a 16 apresentam os resultados para as 30 séries com extensões de 50 a 1000 anos do testes Kolmogorov-Smirnov e Qui-Quadrado. Observa-se um comportamento semelhante aos das distribuições de Gumbel e Lognormal, entretanto com uma sensibilidade em detectar a distribuição errada (rejeitar a hipótese falsa) maior. Ou seja, para altos CVs (1,0 e 1,5) nenhum erro do Tipo II é cometido no teste Qui-Quadrado e, o mesmo para Kolmogorov Smirnov, a partir de n igual a 100.

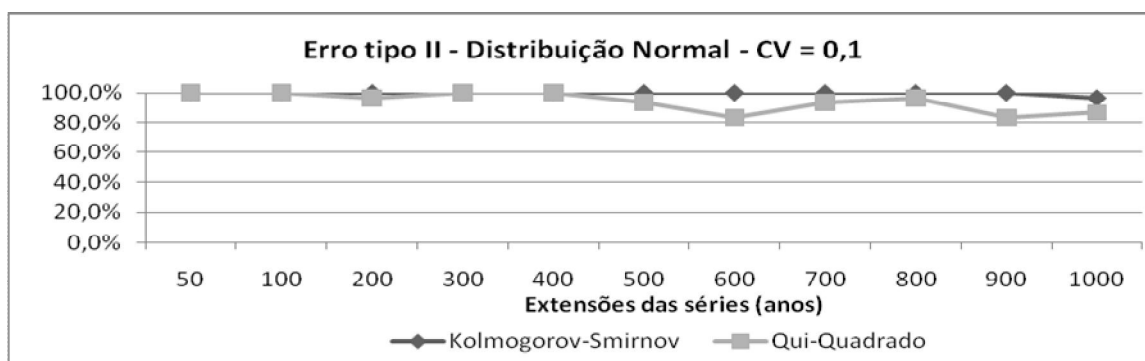


Figura 13- Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando distribuição Normal (CV = 0,1)

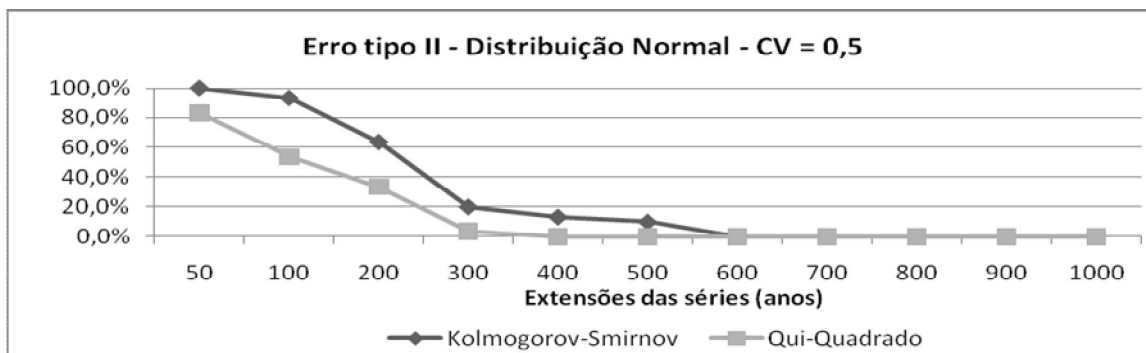


Figura 14 - Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando uma distribuição Normal (CV = 0,5)

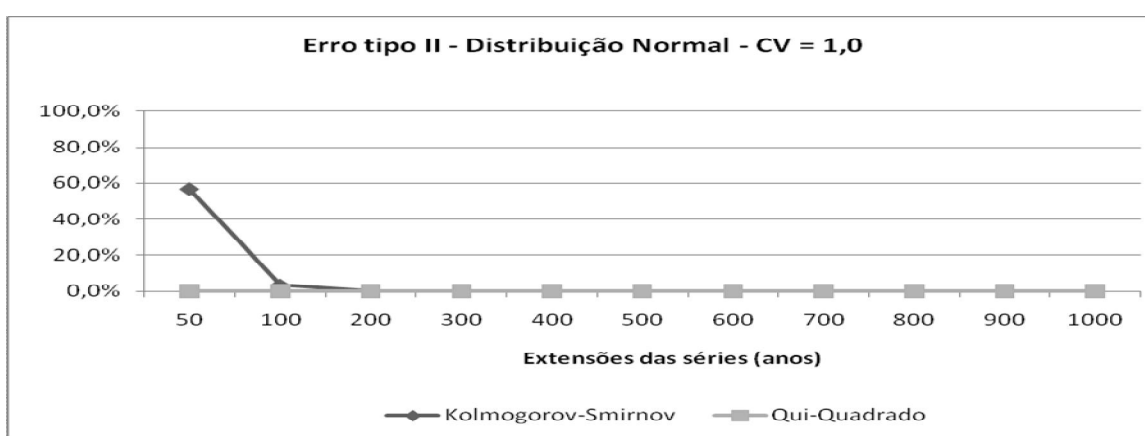


Figura 15 - Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando uma distribuição Normal (CV = 1,0)

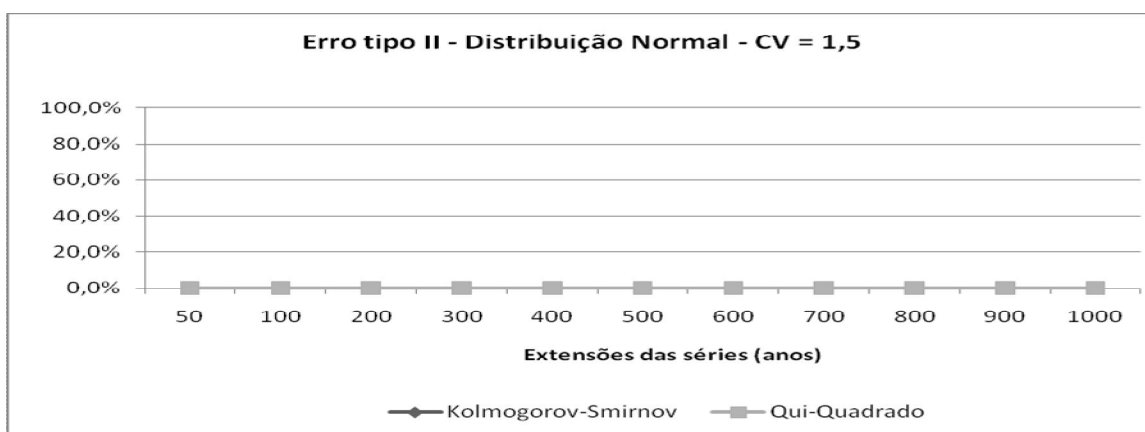


Figura 16 - Erro tipo II para os testes de Kolmogorov-Smirnov e Qui-Quadrado, considerando uma distribuição Normal (CV = 1,5)

3.3 – Erros do Tipo II – análise comparativa dos Testes

3.3.1 - Teste Kolmogorov - Smirnov

As figuras 17 a 20 apresentam os resultados para as 30 séries com extensões de 50 a 1000 anos do testes Kolmogorov-Smirnov para CV variando de 0,1 a 1,5. Observa-se também que os erros do Tipo II decrescem com o aumento do n e CV e que, entre as três distribuições, a Lognormal é a que mais apresenta erros nos testes de aderência. Comportamento anômalo, mais uma vez, é observado para CV igual a 0,5 e deve ser melhor investigado.

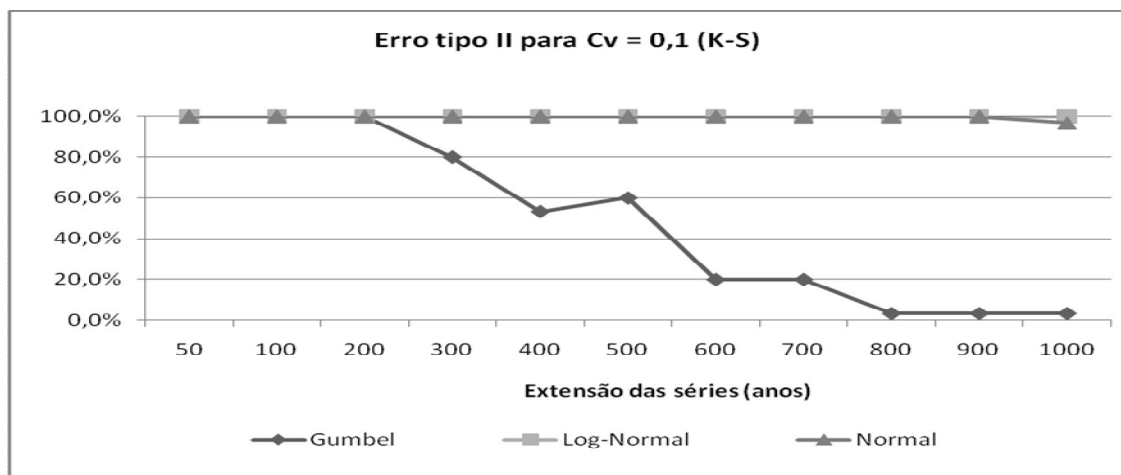


Figura 17 – Erro tipo II do teste de Kolmogorov-Smirnov para o CV = 0,1

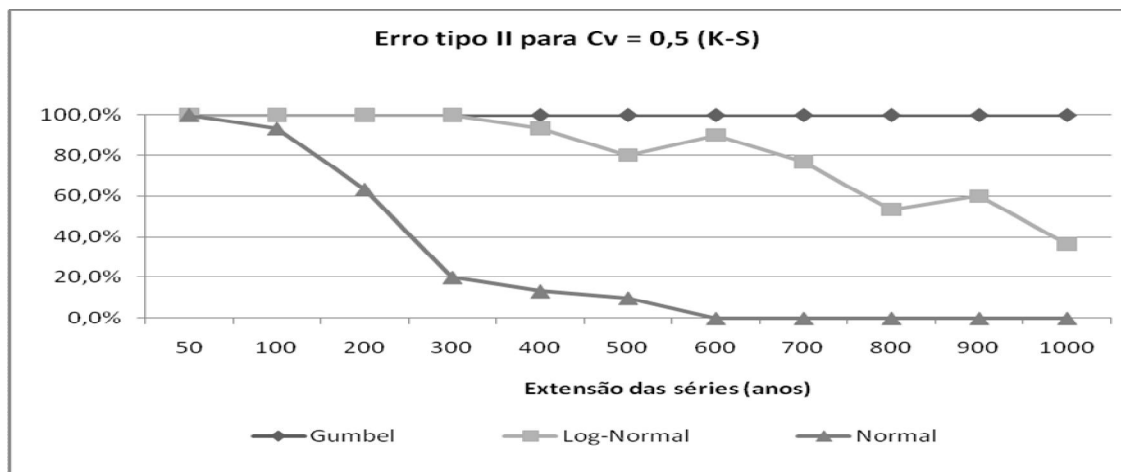


Figura 18 – Erro tipo II do teste de Kolmogorov-Smirnov para o CV = 0,5

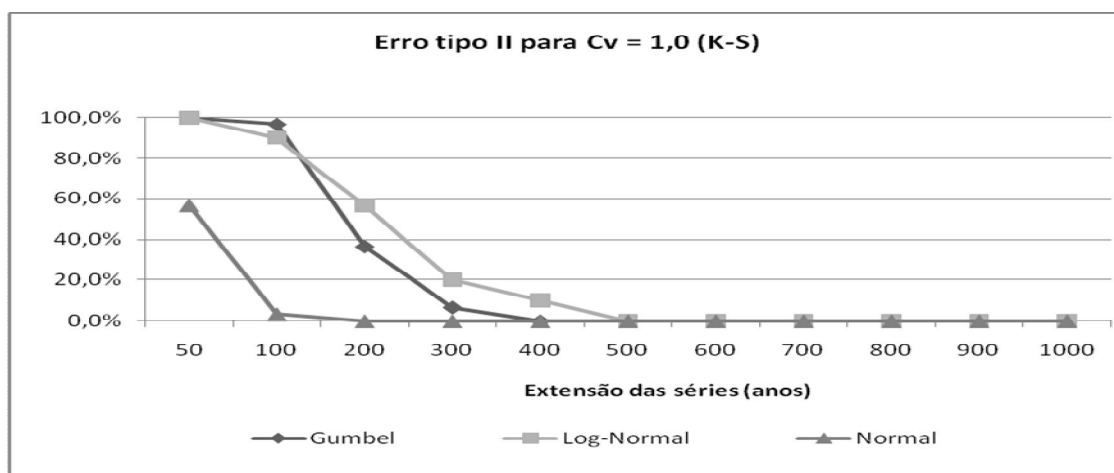


Figura 19 – Erro tipo II do teste de Kolmogorov-Smirnov para o $CV = 1,0$

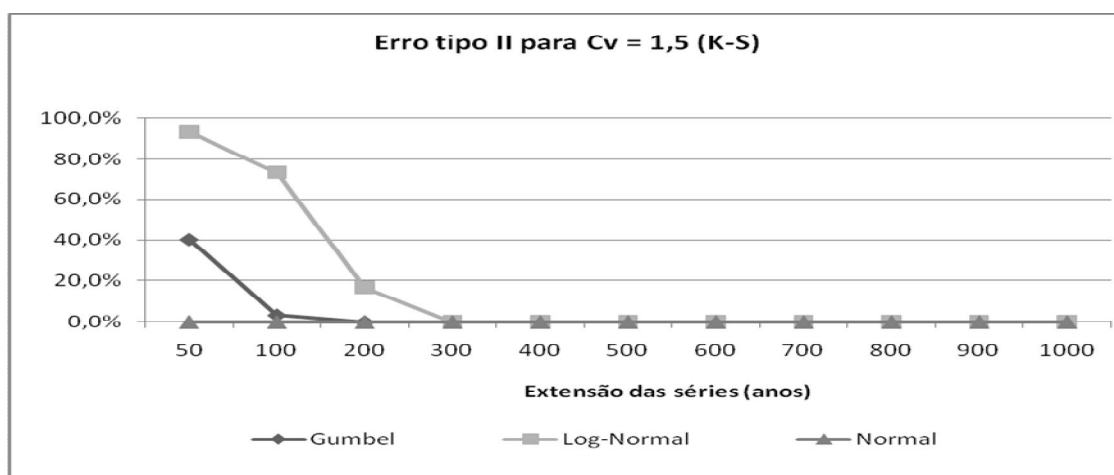


Figura 20 – Erro tipo II do teste de Kolmogorov-Smirnov para o $CV = 1,5$

3.3.2 - Teste Qui-Quadrado

As figuras 21 a 24 apresentam os resultados para as 30 séries com extensões de 50 a 1000 anos do teste Qui-Quadrado para CV variando de 0,1 a 1,5. Observa-se, mais uma vez, que os erros do Tipo II decrescem com o aumento do n e CV . Observa-se nas Figuras 23 e 24 que para altos CV s, os testes identificam facilmente a hipótese falsa, não cometendo mais erros do Tipo II.

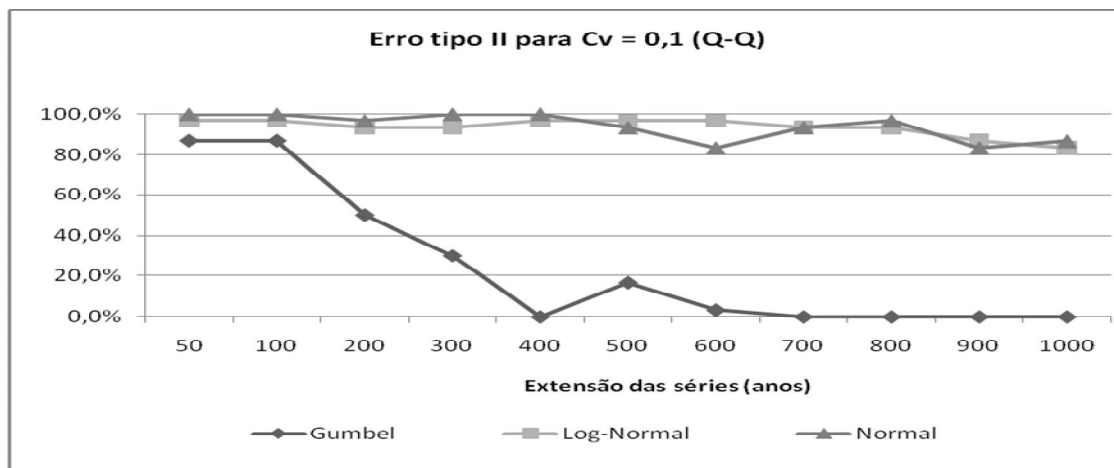


Figura 21 – Erro tipo II do teste de Qui-Quadrado para o CV = 0,1

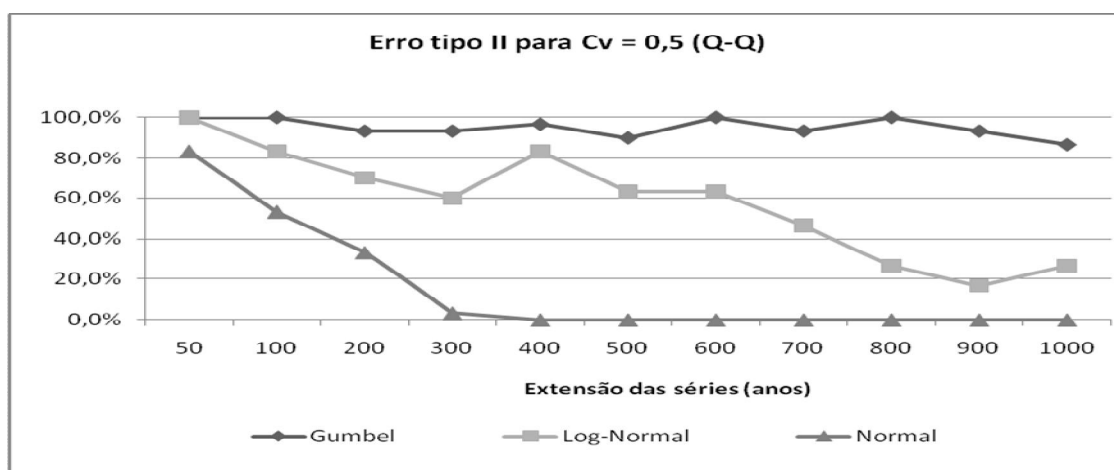


Figura 22 – Erro tipo II do teste de Qui-Quadrado para o CV = 0,5

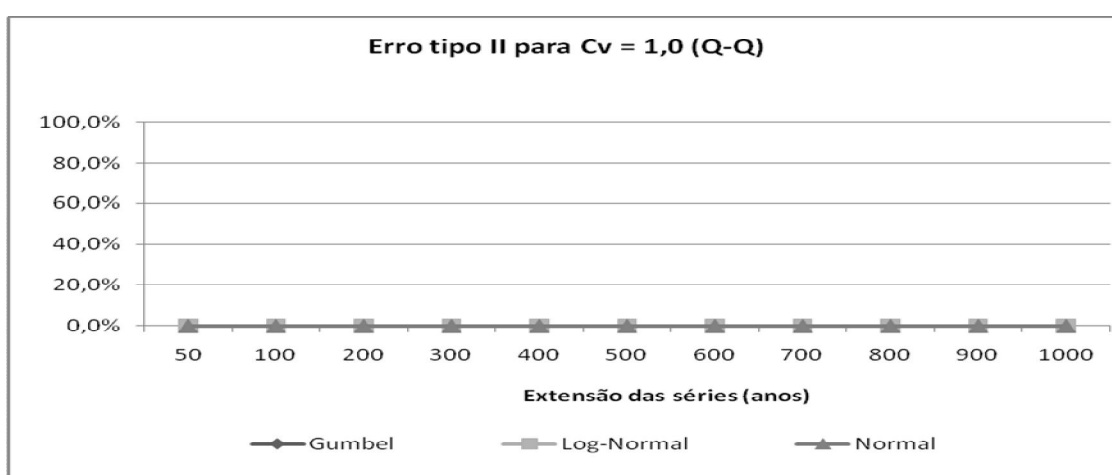


Figura 22 – Erro tipo II do teste de Qui-Quadrado para o CV = 1,0

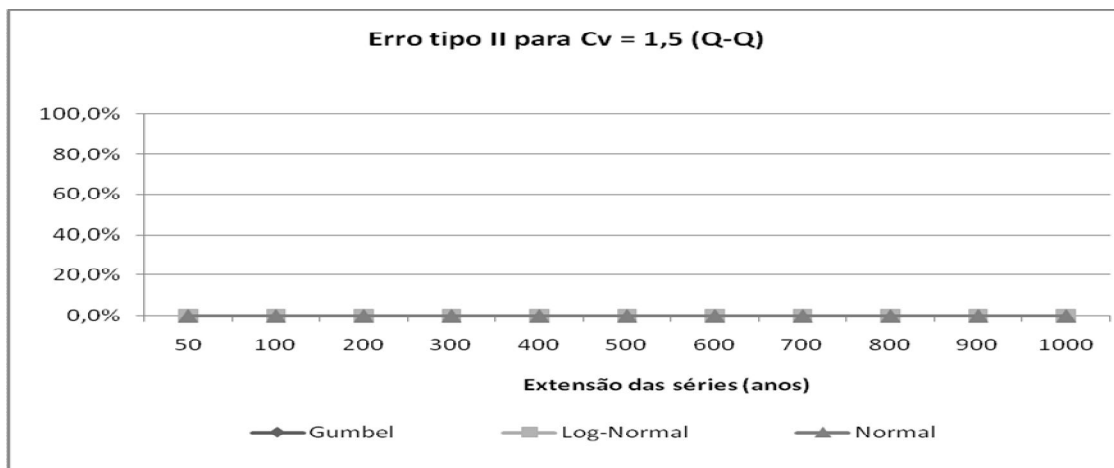


Figura 23 – Erro tipo II do teste de Qui-Quadrado para o CV = 1,5

4 - CONCLUSÕES

Para ambos os valores baixos do coeficiente de variação testados – 0,1 e 0,5 –, o teste de Kolmogorov-Smirnov apresenta erros do tipo II mais elevados do que o teste do Qui-Quadrado. Porém, o teste de Kolmogorov-Smirnov apresenta erro do tipo I nulo, ou seja, reconhece a verdadeira distribuição da série testada, em todos os casos.

Considerando um coeficiente de variação igual a 0,1; para que a distribuição Gumbel seja rejeitada pelos dois testes como falsa é necessário que as séries tenham grande extensão. Já para um coeficiente de variação igual a 0,5; nenhum dos dois testes tem sensibilidade para rejeitar a distribuição Gumbel como falsa, pois apresentam altos valores de erros do tipo II.

A distribuição Lognormal, para o coeficiente de variação de valor igual a 0,1; não é rejeitada como falsa por nenhum dos dois testes de aderência. Mas, para um coeficiente de variação igual a 0,5; o erro tipo II cometido pelos dois testes de ajustamento se reduz à medida que a extensão das séries aumenta. Pode-se inferir que a anulação do erro tipo II, para o teste da distribuição Lognormal, só se dará para uma série extremamente longa.

Dados os grandes erros do tipo II, para um CV igual a 0,1; nem o teste de Kolmogorov-Smirnov nem o teste do Qui-Quadrado têm sensibilidade para rejeitar a distribuição Normal como falsa.

Já para um CV igual a 0,5; os erros do tipo II para o teste da distribuição Normal se reduzem, até a anulação, à medida que aumenta a extensão das séries. Sendo necessária uma série menos extensa para o teste do Qui-Quadrado do que para o teste de Kolmogorov-Smirnov.

Para ambos os valores altos do coeficiente de variação testados – 1,0 e 1,5 –, assim como aconteceu para os valores baixos, o teste de Kolmogorov-Smirnov apresenta erros do tipo II mais elevados do que o teste do Qui-Quadrado. Contudo, o teste de Kolmogorov-Smirnov apresenta erro do tipo I nulo, ou seja, sempre reconhece a verdadeira distribuição da série testada em todos os casos.

Para um coeficiente de variação igual a 1,0; a distribuição Gumbel é sempre rejeitada como falsa pelo teste do Qui-Quadrado, em todos os casos. Já o teste de Kolmogorov-Smirnov só a rejeita completamente, anulando o erro do tipo II, para séries cuja extensão de dados seja maior ou igual a 400. Considerando o coeficiente de variação como sendo 1,5; novamente, o Qui-Quadrado a rejeita em todos os casos. O teste de Kolmogorov-Smirnov só a rejeita a partir de uma série com extensão maior ou igual a 100.

A distribuição Lognormal, para os dois valores altos do coeficiente de variação; é completamente rejeitada em todos os casos – erro do tipo II nulo –, pelo teste do Qui-Quadrado. O teste de Kolmogorov-Smirnov rejeita a distribuição Lognormal como falsa nos seguintes casos: coeficiente de variação igual 1,0 e séries com extensão a partir de 500, e coeficiente de variação igual a 1,5 e série com extensão a partir de 300.

A distribuição Normal é completamente rejeitada pelo teste do Qui-Quadrado, em todos os casos, para os dois valores altos do coeficiente de variação. O teste de Kolmogorov-Smirnov só a rejeita completamente, em todos os casos, para um coeficiente igual a 1,5. Para um coeficiente de valor igual a 1,0; o erro da rejeição torna-se pequeno para séries de extensão igual a 100. Já a rejeição completa da distribuição falsa, pelo teste de Kolmogorov-Smirnov, só ocorre com séries cujas extensões são maiores ou iguais a 200 valores de dados.

BIBLIOGRAFIA

ARAÚJO, E. M. *et al.* (2010). “*Análise da Aderência de Distribuições de Probabilidade aos Dados de Temperatura Máxima e Mínima do Ar em Iguatu-CE*”. Revista Caatinga, Mossoró, v. 23, n. 3, jul/set., p. 104-109.

CUNNANE, C. (1989). “*Statistical distributions for flood frequency analysis*”. Operational Hydrology Report no. 33, WMO no. 718, World Meteorological Organization, Geneva, Switzerland.

FONSECA, J. S.; MARTINS, G. A. (1996). “*Curso de Estatística*”. Editora Atlas S.A. São Paulo. 6ª edição.

- HAKTANIR, T.; HORLACHER, H. B. (1993). "*Evaluation of various distributions for flood frequency analysis*". Hydrological Sciences Journal.
- JOHNSON, R. A. (2000). "*Probability and statistics for engineers*". Prentice-Hall, Inc. United States of America. 6th edition.
- ROSSI, F., Fiorentino, M. & Versace, P. (1984). "*Two component extreme value distribution for flood frequency analysis*". Wat. Resour. Res. 20(7), p. 847-856.
- SILVINO, A. N. O. et al. (2007). "*Determinação de Vazões Extremas para Diversos Períodos de Retorno para o Rio Paraguai utilizando Métodos Estatísticos*". São Paulo, UNESP, Geociências, v. 26, n. 4, p. 369-378.
- STUDART, T. M. C.; CAMPOS, J. N. B. (2001). "*Incertezas nas estimativas da Vazão Regularizada por um Reservatório*". Revista Brasileira de Recursos Hídricos, Volume 6, n. 3, jul/set, 81-94.
- STUDART, T. M. C.; CAMPOS, J. N. B.; CARVALHO, R. M. (2005). "*Recursos Hídricos e as Incertezas Científicas: Novos Conceitos e Abordagens*" in Anais do XIV Simpósio Brasileiro de Recursos Hídricos, Porto Alegre, 2001.
- VOGEL, R. M.; THOMAS, W. O.; McMAHON, T. A. (1993). "*Flood-flow frequency model selection in the southwestern United States*". Wat. Resour. Plan. Manage. ASCE. 119(3), p. 353-366.
- WALLIS, J. R. (1988). "*Catastrophes, computing and containment: living in our restless habitat*". Speculation in Science and Technology. 11(4). p. 295-315.
- YEVJEVICH, V. (1972). "*Probability and Statistics in Hydrology*". Water Resources Publications. Fort Collins, Colorado, USA.