

**AVALIAÇÃO MONTE CARLO DE TESTES  
ASSINTÓTICOS E DE BOOTSTRAP PARA  
AUTOCORRELAÇÃO RESIDUAL**

**VÂNIA DE FÁTIMA LEMES DE MIRANDA**

**2004**

5832  
04

VÂNIA DE FÁTIMA LEMES DE MIRANDA

**AVALIAÇÃO MONTE CARLO DE TESTES ASSINTÓTICOS E  
DE BOOTSTRAP PARA AUTOCORRELAÇÃO RESIDUAL**

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

Orientador

Prof. Daniel Furtado Ferreira

LAVRAS  
MINAS GERAIS - BRASIL

2004

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

Miranda, Vânia da Fátima Lemes de  
Avaliação Monte Carlo de testes assintóticos e de Bootstrap para  
autocorrelação residual / Vânia de Fátima Lemes de Miranda. -- Lavras :  
UFLA, 2004.

77 p. : il.

Orientador: Daniel Furtado Ferreira.  
Dissertação (Mestrado) – UFLA.  
Bibliografia.

1. Método Monte Carlo. 2. Experimentação agropecuária. I. Universidade  
Federal de Lavras. II. Título.

CDD-519.282  
-633.2195

VÂNIA DE FÁTIMA LEMES DE MIRANDA

**AVALIAÇÃO MONTE CARLO DE TESTES ASSINTÓTICOS E DE  
BOOTSTRAP PARA AUTOCORRELAÇÃO RESIDUAL**

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

APROVADA em 06 de fevereiro de 2004

Prof. Dr. Marcelo Tavares

UFU

Prof. Dr. Lucas Monteiro Chaves

UFLA

Prof. Dr. Ruben Delly Veiga

UFLA



Prof. Dr. Daniel Furtado Ferreira  
UFLA  
(Orientador)

LAVRAS  
MINAS GERAIS - BRASIL

**Ofereço**

**A Deus pela vida e capacidade para esta realização**

**Dedico,**

**Aos meus pais, Antônio e Maria Aparecida e aos meus irmãos Clayton, Neiton, Neilton, Helton e Wesley, que sempre estiveram comigo.**

**Aos meus tios e tias, em especial a Ivonilda “tia toca” amizade e companheirismo sempre.**

## AGRADECIMENTOS

À Universidade Federal de Lavras (UFLA), ao Departamento de Ciências Exatas (DEX) e em especial ao Programa de Pós-Graduação em Agronomia/Estatística e Experimentação Agropecuária pelos conhecimentos.

À FAPEMIG, pela bolsa de estudos

Ao professor Daniel Furtado Ferreira, pela orientação, pelas críticas, pelos ensinamentos e, principalmente, pela amizade e disponibilidade durante o curso.

A todos os professores do Departamento pelos ensinamentos e amizade.

Aos meus colegas do curso, Rafaela, José Waldemar, Janser, Neia, Renata, Eliana, Leite.

As funcionárias do DEX, Maria, Maristela, Odília e Selma pela disponibilidade e colaboração durante o curso.

As minhas amigas Gisele e Rafaela pela amizade, companheirismo e paciência no convívio diário na república São Benedito, e a todos os amigos do GPP, em especial ao Breno e a Lili e os padrinhos Leonel e Cláudia.

A minha tia Ivonilda, que sempre estive disponível a me ajudar.

A todos aqueles que direta ou indiretamente colaboraram e entenderam esta difícil, mas, importante etapa de minha vida.

Obrigada!!!!!!!!!!!!!!

# SUMÁRIO

|  | <b>Página</b> |
|--|---------------|
| RESUMO.....  | i             |
| ABSTRACT.....  | ii            |
| 1 INTRODUÇÃO.....  | 1             |
| 2 REFERENCIAL TEÓRICO.....                                       | 3             |
| 2.1 Correlação e autocorrelação.....                             | 3             |
| 2.2 Teste de Durbin-Watson.....                                  | 6             |
| 2.3 Teste t.....   | 9             |
| 2.4 Teste normal para $\rho$ .....                               | 10            |
| 2.5 Teste normal para $\rho$ com correção de viés.....           | 11            |
| 2.6 Teste normal para $c$ de Young.....                          | 12            |
| 2.7 Bootstrap.....   | 15            |
| 2.7.1 Limites de confiança de bootstrap padrão.....              | 16            |
| 2.7.2 Intervalos de confiança baseados em percentis simples..... | 17            |
| 2.7.3 Bootstrap com correção de viés acelerado.....              | 20            |
| 2.8 Simulação Monte Carlo.....                                   | 23            |
| 2.9 Taxas de erros tipo I e tipo II e Poder.....                 | 24            |
| 2.10 Distribuições.....  | 25            |
| 2.10.1 Distribuição Normal.....                                  | 25            |
| 2.10.2 Distribuição Uniforme.....                                | 26            |
| 2.10.3 Distribuição exponencial.....                             | 27            |
| 3 METODOLOGIA.....   | 28            |
| 3.1 Teste de Durbin-Watson (DW).....                             | 29            |
| 3.2 Teste de DW bootstrap (DWB).....                             | 31            |
| 3.3 Teste de bootstrap $\rho$ ( $B\rho$ ).....                   | 32            |

|  |    |
|--|----|
| 3.4 Teste bootstrap de $\rho$ com correção de viés acelerado ( $B\rho Ca$ )..... | 33 |
| 3.5 Teste t ( $t\rho$ ).....   | 35 |
| 3.6 Teste t de bootstrap ( $t\rho B$ ).....                                      | 36 |
| 3.7 Teste normal para $\rho$ ( $N\rho$ ).....                                    | 36 |
| 3.8 Teste normal para $\rho$ com correção de viés ( $N\rho c$ ).....             | 37 |
| 3.9 Teste normal para C de Young (1941) (NC).....                                | 37 |
| 3.10 Teste bootstrap para C (BC).....  | 39 |
| 3.11 Viés e Erro Quadrático Médio (EQM) dos três estimadores de $\rho$ .....     | 39 |
| 4 RESULTADOS E DISCUSSÃO.....  | 40 |
| 4.1 Propriedades dos estimadores de $\rho$ .....                                 | 40 |
| 4.2 Erro tipo I.....   | 45 |
| 4.3 Poder.....   | 53 |
| 4.4 Considerações finais.....  | 65 |
| 5 CONCLUSÕES.....  | 67 |
| REFERÊNCIAS BIBLIOGRÁFICAS.....  | 68 |
| ANEXOS.....  | 70 |

## RESUMO

MIRANDA, Vânia de Fátima Lemes. **Avaliação Monte Carlo de testes para autocorrelação residual assintóticos e de bootstrap.** 2004. 77p. dissertação (Mestrado em Estatística e Experimentação Agropecuária) Universidade Federal de Lavras, Lavras MG\*

Muitas são as situações experimentais em que o pesquisador se depara com a necessidade de ajustar modelos, sejam eles lineares ou não lineares. A presença de autocorrelação nos resíduos afeta as inferências nos parâmetros de regressão. A diagnose dessa correlação é feita formalmente por meio do teste de Durbin-Watson (DW), entre outros. Este trabalho objetivou a avaliar a abordagem bootstrap, utilizando simulação Monte Carlo aplicada ao teste de Durbin-Watson, teste de bootstrap para o parâmetro de autocorrelação  $\rho$  com e sem correção acelerada de viés, teste  $t$ , teste  $t$  de bootstrap, teste normal para  $\rho$ , teste normal para  $\rho$  com correção de viés, teste normal para  $C$  de Young e teste bootstrap para  $C$ . Adicionalmente, a qualidade dos três estimadores de  $\rho$  foi estudada avaliando-se o viés e o erro quadrático médio. Simulou-se, para isso, uma estrutura de autocorrelação de primeira ordem. Como critério de avaliação, as taxas de erro tipo I e o poder destes testes foram comparados com aqueles do teste de DW. As principais conclusões obtidas são: o viés dos estimadores  $\hat{r}$ ,  $r_{1+}$  e  $C$  aumentam com o aumento de  $k$  e de  $\rho$ ; a variância dos estimadores  $\hat{r}$ ,  $r_{1+}$  e  $C$  não são afetadas pelo aumento de  $k$  e de  $\rho$ ; os testes DW,  $N\rho$ ,  $t\rho$ ,  $B\rho$  e  $B\rho Ca$  são rigorosos e menos poderosos do que seus competidores; o teste DW é o mais rigoroso de todos e apresentou menor poder; os testes DWB e BC são equivalentes; os testes DWB, BC,  $t\rho B$ ,  $N\rho c$  e  $NC$  são considerados os melhores por apresentarem tamanho de teste idênticos ao valor nominal, maior poder e por serem robustos; o teste  $NC$  foi considerado rigoroso para pequenos valores de  $\alpha$  e de  $n$ ; Além disso o teste  $N\rho c$  é recomendado nesse trabalho devido à simplicidade e facilidade para ser aplicado.

---

\*Comitê Orientador: Daniel Furtado Ferreira (Orientador)

## ABSTRACT

MIRANDA, Vânia de Fátima Lemes. **Monte Carlo evaluation of asymptotic and bootstrap tests for residual autocorrelation.** 2004. 77 p. Dissertation (Master in Statistics and Agricultural Experimentation) – Universidade Federal de Lavras, Lavras MG.\*

There are many experimental situations that researcher needs to fit linear or non linear models. The residual autocorrelation affects the inferences on the regression parameters. The diagnosis of that autocorrelation is made formally through Durbin-Watson's test (DW) among others. This work aimed to evaluate through Monte Carlo simulation the Durbin-Watson's bootstrap approach, direct bootstrap test for autocorrelation parameter ( $\rho$ ) with and without accelerated bias correction, t test, bootstrap t test, normal  $\rho$  test, normal  $\rho$  with bias correction, normal test for Young's C statistics, bootstrap C test approach. Additionally, the quality of the three estimators of  $\rho$  was studied evaluating the bias and mean square error. It was simulated an autocorrelation structure of first order. As evaluation criterion the type I error rate and the power of these tests were compared with the test of DW. The main conclusions are: the bias of the  $\hat{r}$ ,  $r_{1+}$  and C estimators increase with the increase of k and  $\rho$ ; the variance of the three estimators are not affected for the increase of k and  $\rho$ ; the DW,  $N\rho$ ,  $t\rho$ ,  $B\rho$  and  $B\rho Ca$  tests are rigorous and less powerful than their competitors; the DW test is the most rigorous of all and presented the smallest power; the DWB and BC tests are equivalents; the DWB, BC,  $t\rho B$ ,  $N\rho c$  and NC tests are considered the best because they have identical size to the nominal value, larger power and because they are robust; the NC test was considered rigorous for small values of  $\alpha$  and of n; the  $N\rho c$  test is recommended here due to its simplicity and easiness to be applied.

---

\*Guidance Committee: Daniel Furtado Ferreira (Adviser)

# 1 INTRODUÇÃO

Muitas são as situações experimentais em que o pesquisador se depara com a necessidade de ajustar modelos, sejam eles lineares ou não lineares. As metodologias de análises usuais pressupõem que os resíduos, parte aleatória desses modelos, são aditivos, homocedásticos, com distribuição normal e independentes. Dentre as técnicas utilizadas para avaliar o atendimento a essas pressuposições destacam-se as análises gráficas dos resíduos. Nesse contexto, pode-se checar a independência entre os resíduos por intermédio de um gráfico dos resíduos versus as observações ordenadas em ordem cronológica. A correlação residual afeta a precisão das estimativas e a qualidade da inferência. A dependência entre os resíduos sucessivos é denominada de autocorrelação, ou ainda, de correlação serial.

A diagnose dessa correlação é feita formalmente por meio do teste de Durbin-Watson (DW). No entanto, segundo Jeong e Chung (2001) este teste apresenta algumas limitações. A primeira delas e a mais crítica reside no fato de a distribuição da estatística do teste não ser matematicamente tratável e, ainda, ser dependente da matriz de delineamento. Devido a estas duas características, vários autores, tais como Hannan (1957) e Hannan & Terrel (1968), apresentaram aproximações numéricas fornecendo valores críticos, considerando a matriz de delineamento, porém estas aproximações apresentam baixo poder.

Recentemente, técnicas computacionais têm sido aplicadas para realizar testes de hipóteses e obter estimativas de parâmetros. Entre essas técnicas o método de bootstrap tem se destacado. Jeong e Chung (2001) apresentaram a generalização do teste de bootstrap para o teste de DW. No entanto, algumas situações restritas foram avaliadas, tais como a escolha da distribuição residual e os tamanhos de amostras utilizados. Por outro lado, os testes para autocorrelação revisados e estudados por De Carlo e Tryon (1993) são simples e relativamente

eficientes. Todavia, nenhum relato de comparação desses procedimentos com os métodos de DW e de suas versões de bootstrap foi encontrado na literatura. Também foram encontrados, na literatura, relatos de versões de bootstrap desses testes apresentados por De Carlo e Tryon (1993).

Em virtude do que foi mencionado, neste trabalho objetivou-se avaliar a abordagem bootstrap, utilizando simulação Monte Carlo aplicada ao teste de Durbin-Watson, teste de bootstrap para o parâmetro de autocorrelação  $\rho$  com e sem correção acelerada de viés, teste t, teste t de bootstrap, teste normal para  $\rho$ , teste normal para  $\rho$  com correção de viés, teste normal para C de Young e teste bootstrap para C. Adicionalmente, a qualidade dos três estimadores de  $\rho$  foi estudada avaliando-se o viés e o erro quadrático médio. Simulou-se, para isso, uma estrutura de autocorrelação de primeira ordem. Como critério de avaliação, as taxas de erro tipo I e o poder destes testes foram comparados com aqueles do teste de DW.

## 2 REVISÃO DE LITERATURA

### 2.1 Correlação e autocorrelação

A correlação mede o grau de relacionamento entre duas variáveis aleatórias, não sendo necessário distinguir se as variáveis são dependentes ou independentes, como ocorre no caso da análise de regressão. A correlação não é afetada pela escala da variável, como ocorre normalmente com a covariância. A covariância possui a mesma dimensão dos dados e a correlação é a versão padronizada da covariância, não possuindo dimensão. A vantagem de se utilizar a correlação em substituição à covariância é imediata por eliminar qualquer influência da unidade de medida e por facilitar a interpretação. O coeficiente de correlação possui a particularidade de variar entre -1 e 1. Se a correlação for nula indica que as variáveis não possuem uma relação linear; se for maior que zero, indica que existe uma relação linear positiva entre as variáveis, ou seja, elas tendem a variar no mesmo sentido, e se a correlação for menor que zero, indica que existe uma relação linear negativa, ou seja, elas tendem a variar em sentidos opostos, pequenos valores de uma variável tendem a estar relacionados com grandes valores da outra variável (Hoffmann e Vieira, 1987).

O coeficiente de correlação populacional é definido por:

$$\rho_{XY} = \frac{\text{Cov}[X, Y]}{\sigma_X \sigma_Y}, \quad (2.1)$$

em que X e Y são as variáveis aleatórias,  $\text{Cov}[X, Y]$ ,  $\sigma_X$  e  $\sigma_Y$  representam a covariância populacional entre X e Y, desvio padrão populacional de X e desvio

padrão populacional de Y, respectivamente, sendo que  $\sigma_X > 0$  e  $\sigma_Y > 0$  (Mood, Graybill e Boes, 1996).

A covariância populacional  $Cov[X, Y]$  entre as variáveis aleatórias X e Y é definida por:

$$\begin{aligned} Cov[X, Y] &= E\{[X - E(X)][Y - E(Y)]\} \\ &= E(X, Y) - E(X)E(Y) \end{aligned} \quad (2.2)$$

A covariância mede o grau de associação entre duas variáveis. Se a covariância for positiva, indica que o aumento do valor de uma variável tende a ser acompanhado com o aumento do valor da outra. Se for negativa, indica que as variáveis apresentam tendências de variação opostas, e se for nula, indica que não é possível prever o que acontece com uma variável (Bearzoti, 1998).

O coeficiente de correlação linear amostral é chamado de coeficiente de correlação momento-produto de Pearson em homenagem a seu precursor Karl Pearson (1857-1936). Considerando uma amostra bivariada de tamanho n, dada por  $X_1, X_2, \dots, X_n$  e  $Y_1, Y_2, \dots, Y_n$ , define-se o coeficiente de correlação por:

$$r = \frac{n \sum_{i=1}^n x_i y_i - \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)}{\sqrt{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \sqrt{n \sum_{i=1}^n y_i^2 - \left( \sum_{i=1}^n y_i \right)^2}} \quad (2.3)$$

Testes são utilizados para verificar se a correlação populacional  $\rho$  é nula. Para isso, pode-se utilizar testes apropriados utilizando tabelas específicas que dependem do valor nominal da significância do teste  $\alpha$  e do tamanho da amostra. Para efetuar o teste t, verifica-se se o valor encontrado para r é

significativamente diferente de zero, observando se o seu módulo excede o valor tabelado; caso contrário, conclui-se que há evidências para rejeitar a hipótese nula de ausência de correlação linear. A segunda maneira de testar essa hipótese é por meio da distribuição t de Student. A estatística do teste, sob a pressuposição de que a amostra bivariada (X e Y) é proveniente de uma distribuição normal e sob a hipótese nula  $H_0 : \rho = 0$ , é dada da seguinte forma, de acordo com Triola (1999):

$$t = \frac{r}{s_r}, \quad (2.4)$$

em que:

$$s_r = \frac{1 - r}{\sqrt{n - 2}}. \quad (2.5)$$

O coeficiente de correlação amostral é representado por r, como já anteriormente apresentado. O seu quadrado é denominado de coeficiente de determinação e representa a porcentagem da variação de Y que é explicada pela regressão, sendo que seu valor varia entre 0 e 1 (Neter e Wasserman, 1974).

Para o caso da autocorrelação, sua definição pode ser considerada como a correlação dos valores de uma variável no instante t com os valores, dessa mesma variável, defasados no tempo. Em geral, a autocorrelação possui interesse maior quando aplicada aos resíduos de um modelo linear ou não linear. Os resíduos quando são correlacionados no tempo ou no espaço são ditos autocorrelacionados ou correlacionados serialmente. Se for considerado um modelo de regressão, pressupõe-se que os resíduos, parte aleatória deste modelo, são não correlacionados, com distribuição normal e homocedásticos. Para

observações ordenadas no tempo, a pressuposição de independência é freqüentemente violada. Especificamente, pode-se verificar que os resíduos são geralmente correlacionados positivamente quando se trabalha com observações ordenadas no tempo (Neter, Wasserman e Dutner, 1974). Se a autocorrelação for maior que zero, diz-se que os erros estão positivamente autocorrelacionados, e se for menor que zero, há autocorrelação negativa (Hoffmann e Vieira, 1987). Uma alternativa muito empregada para detectar se os resíduos da regressão são ou não autocorrelacionados é realizar análises gráficas dos resíduos versus as observações ordenadas em ordem cronológica (Draper e Smith, 1998).

Alguns problemas ocorrem se os resíduos são autocorrelacionados, quando os procedimentos de mínimos quadrados ordinários são utilizados. Todavia, caso os coeficientes de mínimos quadrados ordinários sejam não viesados, eles não têm as propriedades de variância mínima. Além disso, o método de mínimos quadrados pode não estimar adequadamente a variância do erro. Intervalo de confiança e testes de hipóteses que utilizem as distribuições t e F não são estritamente aplicáveis (Neter, Wasserman e Dutner, 1985).

O estudo gráfico da correlação serial é uma técnica usada em análises de séries temporais, simples e de fácil interpretação. Mas a forma mais apropriada para a checagem formal da correlação serial é feita por meio do teste de Durbin-Watson (Draper e Smith, 1998).

## 2.2 Teste de Durbin-Watson

O trabalho de Durbin-Watson (1950) trata da teoria na qual o teste foi baseado. Esse trabalho considera, a princípio, o seguinte modelo de regressão:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i, \quad (2.6)$$

em que  $y_i$  é a  $i$ -ésima observação amostral da variável dependente  $Y$ ;  $\beta_0$  é o intercepto,  $\beta_l$  é o coeficiente de regressão associado à  $l$ -ésima variável independente  $X_l$  e  $u_i$  é o erro aleatório suposto normal e homocedástico;  $i = 1, 2, \dots, n$  e  $l = 1, 2, \dots, k$ .

O teste de DW é o mais usado para verificar a existência da autocorrelação nos resíduos da regressão. Sob a hipótese de existência de autocorrelação, assume-se, para os resíduos, um modelo autorregressivo de primeira ordem, dado por:

$$u_t = \rho u_{t-1} + e_t, \quad (2.7)$$

em que  $\rho$  é o parâmetro de autocorrelação tal que  $|\rho| < 1$  e  $e_t$  é o ruído branco, para o qual se assume distribuição normal com média zero e variância  $s^2$  e independência.

Considerando os valores das variáveis independentes conhecidos e fixos, o teste de DW verifica se é plausível que o parâmetro de autocorrelação seja nulo. A estatística do teste é obtida de acordo com a equação (2.8). Inicialmente, é aplicado o método de mínimos quadrados ordinários para estimar os parâmetros e o resíduo  $u_t$  no modelo (2.6) da seguinte forma  $\hat{u}_t = Y_t - \hat{Y}_t$ . Então a estatística do teste de Durbin-Watson é baseada no seguinte valor:

$$d = \frac{\sum_{t=2}^n (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^n \hat{u}_t^2}. \quad (2.8)$$

Para  $n$  grande,  $d$  é aproximadamente igual a  $2(1 - r)$ , sendo  $r$  o coeficiente de correlação entre  $\hat{u}_t$  e  $\hat{u}_{t-1}$ . O valor de  $d$  varia entre zero e quatro, quando  $r = 1$  ou  $r = -1$ . O valor de  $d$  próximo de zero indica a existência de autocorrelação positiva nos erros e um valor de  $d$  próximo de 4 indica que os erros estão negativamente autocorrelacionados (Hoffmann e Vieira, 1987).

O teste de DW faz uso da estatística  $d$ . Uma desvantagem deste teste é que a sua distribuição depende do tamanho da amostra, do número de parâmetros estimados e da matriz  $X$  de delineamento, sendo, desta forma, impossível construir um único valor crítico. A distribuição de  $d$  possui domínio entre 0 e 4 e é simétrica em relação a 2. Pela dificuldade de obter um único valor crítico de  $d$ , Durbin-Watson (1951) encontraram limites inferior e superior ( $d_L$  e  $d_U$ ) para realizar o teste. No entanto, o poder do teste tornou-se baixo. Vários estudiosos tentaram melhorar o poder propondo, por exemplo, uma aproximação para a distribuição de  $d$  pela distribuição beta. Durbin e Watson (1951) também utilizaram a aproximação de Jacobi para obter os quantis de  $d_L$  e  $d_U$ . O teste é usualmente aplicado usando tabelas apropriadas contendo estes limites. Uma maneira de efetuar o teste é comparar o valor de  $d$  observado com  $d_L$  e  $d_U$  obtidos em tabelas construídas para vários valores para  $n$ ,  $k$  (número de variáveis independentes) e níveis de significância de 1 %, 2,5% e 5% (Hoffmann e Vieira, 1987).

Para testar a hipótese  $H_0: \rho = 0$  versus  $H_1: \rho > 0$ , o valor de  $d$  é comparado com  $d_L$  e  $d_U$ , conforme já comentado anteriormente. Se  $d < d_L$  a hipótese nula é rejeitada, ou seja, conclui-se que a correlação é positiva; se  $d > d_U$ , a hipótese nula não deve ser rejeitada naquele nível nominal de significância adotado; e se  $d_L \leq d \leq d_U$  o resultado é não conclusivo. Quando não se conhece o sinal da correlação serial populacional, é feito o teste bilateral. A hipótese  $H_0: \rho = 0$  versus  $H_1: \rho \neq 0$  é testada comparando o valor de  $d$  ou de  $4 - d$  com  $d_L$  e  $d_U$ . Se  $d < d_L$  ou  $4 - d < d_L$ , a hipótese nula deve ser rejeitada, ou seja, os resíduos são

correlacionados serialmente; se  $d > d_U$  e  $4 - d < d_U$  a hipótese nula não deve ser rejeitada, ou seja, os resíduos não estão correlacionados; e se  $d_L \leq d \leq d_U$  ou  $d_L \leq 4 - d \leq d_U$ , o resultado é não conclusivo. Finalmente, para testar as hipóteses  $H_0: \rho = 0$  versus  $H_1: \rho < 0$  o valor de  $d$  é comparado com  $4 - d_L$  e  $4 - d_U$ , ou também pode-se comparar  $4 - d$  com  $d_L$  e  $d_U$ . Assim, se  $d > 4 - d_L$ , a hipótese nula é rejeitada, ou seja, a correlação é negativa; se  $4 - d_U < d < 4 - d_L$ , o resultado é não conclusivo; e se  $d > 4 - d_U$ , a hipótese nula não deve ser rejeitada (Draper e Smith, 1998).

O teste é válido, no entanto, se os erros tiverem distribuição normal com média zero e variância constante, e se as variáveis independentes não forem aleatórias (Hoffmann e Viera, 1998).

### 2.3 Teste t

Para testar a hipótese de ausência de correlação é possível aplicar diretamente um teste para o parâmetro  $\rho_{XY}$ . Assim, utilizando o estimador da correlação populacional ( $\rho_{XY}$ ) de mínimos quadrados obtido na equação (2.3), é possível testar a hipótese  $H_0: \rho_{XY} = 0$  utilizando o teste apresentado na equação (2.4), o qual tem distribuição t de Student com  $n - 2$  graus de liberdade sob  $H_0$ . Esta distribuição não pode ser usada para obter intervalos de confiança para  $\rho_{XY}$ , pois a distribuição de  $r$  não é tratável quando a hipótese nula é falsa (Moretin e Bussab, 2003).

Esse teste pode ser adaptado para o parâmetro  $\rho$  de autocorrelação, utilizando o estimador  $\hat{\rho}$  apresentado na equação (2.12) da seguinte forma:

$$t_c = \frac{\hat{\rho}\sqrt{n-3}}{\sqrt{1-\hat{\rho}^2}}. \quad (2.9)$$

Essa estatística sob a hipótese  $H_0 : \rho = 0$  possui distribuição aproximadamente t de Student com  $n - 3$  graus de liberdade. A razão para utilizar  $n - 3$  graus de liberdade é dada pelo fato de que é feito um pareamento entre os valores da variável de um instante  $t$  com o instante imediatamente anterior  $t - 1$ . Assim, são formados  $n - 1$  pares, sendo os graus de liberdade dados por  $n - 3$ .

#### 2.4 Teste normal para $\rho$

Como não existe uma forma universal para se obter a distribuição de  $r$ , pois essa depende da distribuição da população bivariada que está sendo amostrada, então não é possível apresentar um único teste para todos os possíveis casos. No entanto, se a amostra é tipicamente grande ( $n > 500$ ), a distribuição converge para a distribuição normal com média zero e desvio padrão  $1/\sqrt{n}$ . Da mesma forma, não existe uma forma universal da distribuição do estimador ( $\hat{\rho}$ ) de autocorrelação. Portanto, aproximações normais assintóticas são apresentadas na literatura.

Assim, o estimador do erro padrão de  $\hat{\rho}$  apresentado por De Carlo e Tryon (1993), o qual foi proposto por Moran (1948), é dado por:

$$S_{(\hat{\rho})} = \frac{n-2}{n\sqrt{n-1}}; \quad (2.10)$$

e o teste normal padrão sob a hipótese nula  $H_0: \rho=0$  é dado por:

$$Z_c = \frac{\hat{\rho}n\sqrt{n-1}}{n-2}. \quad (2.11)$$

Para  $n$  grande, as estatísticas  $t$  e  $Z_c$  se tornam assintoticamente as mesmas.

## 2.5 Teste normal para $\rho$ com correção de viés

Um estimador de primeira ordem para  $\rho$  é dado por:

$$\hat{\rho} = \frac{\sum_{t=1}^{n-1} (\hat{u}_t - \bar{u})(\hat{u}_{t+1} - \bar{u})}{\sum_{t=1}^n (\hat{u}_t - \bar{u})^2}, \quad (2.12)$$

em que  $\hat{u}_t$  e  $\hat{u}_{t+1}$  são os resíduos da regressão nos tempos  $t$  e  $t-1$ , respectivamente. Pode-se notar que  $\hat{\rho}$  é similar ao coeficiente de correlação  $r$  de Pearson, sendo que a diferença existente refere-se apenas ao somatório dos numeradores de (2.12) que incluem  $n-1$  termos ao invés de  $n$  termos. Uma outra importante diferença entre  $\hat{\rho}$  e  $r$  é que o viés do  $\hat{\rho}$  é maior para amostras pequenas. Huitema e Mckean (1991) sugerem o erro do seguinte estimador modificado:

$$r_{1+} = \hat{\rho} + \frac{1}{n}. \quad (2.13)$$

O viés de  $r_{1+}$  para autocorrelação positiva é menor que o viés de  $\hat{\rho}$ , particularmente para amostras menores ( $n < 50$ ) (De Carlo e Tryon, 1993). O erro padrão de  $r_{1+}$  é o mesmo de  $\hat{\rho}$ , e sugere-se usar o mesmo teste normal padrão apresentado na equação (2.14) substituindo  $\hat{\rho}$  por  $r_{1+}$ .

## 2.6 Teste normal para C de Young

A estatística C de Young (1941) também é um estimador de autocorrelação de primeira ordem, o qual é baseado em somas de diferenças entre sucessivos resíduos, sendo dado por:

$$C = 1 - \frac{\sum_{i=1}^{n-1} (\hat{u}_i - \hat{u}_{i+1})^2}{2 \sum_{i=1}^n (\hat{u}_i - \bar{u}_.)^2} \quad (2.15)$$

Por intermédio de simples álgebra pode-se descrever a relação da estatística C com  $\hat{\rho}$ , como é mostrado a seguir. Os termos do numerador podem ser reescritos da seguinte forma:

$$(\hat{u}_i - \hat{u}_{i+1})^2 = [(\hat{u}_i - \bar{u}_.) - (\hat{u}_{i+1} - \bar{u}_.)]^2.$$

Reescrevendo a equação, tem-se:

$$C = 1 - \frac{\sum_{i=1}^{n-1} (\hat{u}_i - \bar{u}_.)^2 + \sum_{i=1}^{n-1} (\hat{u}_{i+1} - \bar{u}_.)^2 - 2 \sum_{i=1}^{n-1} (\hat{u}_i - \bar{u}_.) (\hat{u}_{i+1} - \bar{u}_.)}{2 \sum_{i=1}^n (\hat{u}_i - \bar{u}_.)^2}.$$

Observa-se que o ultimo termo da equação é igual a  $\hat{\rho}$ , assim:

$$C = \hat{\rho} + 1 - \frac{\sum_{i=1}^{n-1} (\hat{u}_i - \bar{u}_.)^2 + \sum_{i=1}^{n-1} (\hat{u}_{i+1} - \bar{u}_.)^2}{2 \sum_{i=1}^n (\hat{u}_i - \bar{u}_.)^2}.$$

Os termos do numerador podem ser reescritos usando as seguintes identidades:

$$\sum_{i=1}^{n-1} (\hat{u}_i - \bar{u}_.)^2 = \sum_{i=1}^n (\hat{u}_i - \bar{u}_.)^2 - (\hat{u}_n - \bar{u}_.)^2;$$

$$\sum_{i=1}^{n-1} (\hat{u}_{i+1} - \bar{u}_.)^2 = \sum_{i=1}^n (\hat{u}_i - \bar{u}_.)^2 - (\hat{u}_1 - \bar{u}_.)^2.$$

Logo, tem-se:

$$C = \hat{\rho} + 1 - \frac{2 \sum_{i=1}^{n-1} (\hat{u}_i - \bar{u}_.)^2 - (\hat{u}_n - \bar{u}_.)^2 - (\hat{u}_1 - \bar{u}_.)^2}{2 \sum_{i=1}^n (\hat{u}_i - \bar{u}_.)^2}.$$

Substituindo

$$\frac{2 \sum_{i=1}^{n-1} (\hat{u}_i - \bar{u}_.)^2}{2 \sum_{i=1}^n (\hat{u}_i - \bar{u}_.)^2}$$

por 1 e cancelando os termos, tem-se:

$$C = \hat{\rho} + \frac{(\hat{u}_n - \bar{u}_.)^2 + (\hat{u}_1 - \bar{u}_.)^2}{2 \sum_{i=1}^n (\hat{u}_i - \bar{u}_.)^2}. \quad (2.16)$$

A análise desta equação mostra que a estatística C é similar ao estimador modificado  $r_{1+}$ , e converge para  $\hat{\rho}$  se a amostra for grande.

A estatística d de Durbin-Watson é baseada na diferença de sucessivos resíduos e é similar à estatística C de Young devido ao fato de  $\bar{u}_.$ , por definição, ser zero (média dos resíduos). Assim, o denominador da equação (2.16) é igual ao denominador da estatística d de Durbin-Watson, logo, tem-se:

$$C = 1 - \frac{1}{2}d. \quad (2.17)$$

Os testes para estas duas estatísticas são diferentes, uma vez que a estatística d é comparada com os valores críticos tabelados (limite inferior e limite superior), enquanto a estatística C é comparada com os valores críticos da distribuição normal padrão.

O erro padrão da estatística C é dado por:

$$S_{(c)} = \sqrt{\frac{n-2}{n^2-1}} \quad (2.18)$$

e o valor de  $Z$  calculado é dado por:

$$Z_c = \frac{C\sqrt{n^2-1}}{\sqrt{n-2}}. \quad (2.19)$$

O teste estatístico, neste caso, é comparado com os valores críticos da distribuição normal (De Carlo e Tryon, 1993).

## 2.7 Bootstrap

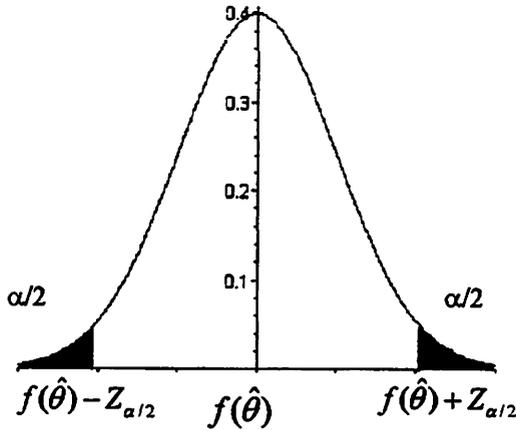
A técnica do bootstrap foi considerada sistematicamente por Efron (1979), apesar de genericamente o método ter sido usado em circunstâncias particulares antes disso. A essência do bootstrap consiste na idéia de que na ausência de qualquer outro conhecimento da população, a distribuição dos valores encontrados em uma amostra aleatória de tamanho  $n$  da população é o melhor guia da distribuição da população. Portanto, para aproximar o que acontece quando a população é reamostrada, basta reamostrar a amostra. A amostragem é feita com reposição, o que é a única diferença entre bootstrap e aleatorização em muitas aplicações (Manly, 1998).

A partir de uma amostra aleatória de tamanho  $n$  de uma população ( $X_1, X_2, \dots, X_n$ ) é obtida uma nova amostra de tamanho  $n$  com reposição dessa amostra. Se o parâmetro populacional de interesse for  $\theta$  e o seu estimador for  $\hat{\theta}$ , em cada amostra original e de bootstrap (amostra com reposição da amostra original) são calculados com os dados amostrais. O processo de bootstrap é

ou

$$P[f(\hat{\theta}) - Z_{\alpha/2} < f(\theta) < f(\hat{\theta}) + Z_{\alpha/2}] = 1 - \alpha. \quad (2.23)$$

Se a função é conhecida, basta aplicar a função inversa para obter o IC para  $\theta$ . Geralmente, a função não é conhecida. Suponha que a distribuição do estimador transformado seja gerada por reamostragem (bootstrap) dos dados originais, i.e.,  $f(\hat{\theta}_j)$ . A distribuição dessa função é apresentada na Figura 2.1.



**Figura 2.1.** Distribuição de bootstrap da transformação de  $\hat{q}_j$  dada por  $f(\hat{\theta}_j)$ , considerada normal com média  $f(\hat{\theta})$  e desvio padrão 1.

É interessante verificar que os limites do IC são:

$$L = f(\hat{\theta}) - Z_{\alpha/2} \text{ e } U = f(\hat{\theta}) + Z_{\alpha/2}. \quad (2.24)$$

Então, uma forma de gerar o IC para  $f(\theta)$  é reamostrar a amostra original e encontrar os limites que excedem as frações  $\alpha/2$  (L) e  $1-\alpha/2$  (U) dos valores gerados. No entanto, essa é a parte crucial devido à natureza monotônica da transformação considerada. A ordem das estimativas de bootstrap transformadas consiste na mesma ordem das estimativas de bootstrap não transformadas  $(\hat{\theta}_j)$ . Portanto, basta apenas realizar o bootstrap do estimador não transformado  $(\hat{\theta}_j)$  e tomar, respectivamente, os percentis  $\alpha/2$  (L) e  $1-\alpha/2$  (U) dessa distribuição.

Hall (1992) sugere um segundo método baseado no mesmo argumento anterior, porém tomando-se a distribuição dos erros. Para a reamostragem da população o erro é definido por  $\varepsilon = \hat{\theta} - \theta$  e o erro de bootstrap por  $\varepsilon_j = \hat{\theta}_j - \hat{\theta}$ . Com base nessa definição é possível gerar dois erros  $\varepsilon_L$  e  $\varepsilon_H$ , em que:

$$P(\varepsilon_L < \hat{\theta}_B - \hat{\theta} < \varepsilon_H) = 1 - \alpha, \quad (2.25)$$

sendo,

$$P(\varepsilon < \varepsilon_L) = P(\varepsilon > \varepsilon_H) = \alpha / 2. \quad (2.26)$$

Com base na premissa de que os erros de bootstrap fornecem boas estimativas do erro verdadeiro, então:

$$P(\varepsilon_L < \hat{\theta} - \theta < \varepsilon_H) = P(\hat{\theta} - \varepsilon_H < \theta < \hat{\theta} - \varepsilon_L) = 1 - \alpha, \quad (2.27)$$

que é a segunda forma de gerar IC de bootstrap. Usando ainda a distribuição dos erros, a expressão anterior pode ser reescrita por:

$$P(2\hat{\theta} - \hat{\theta}_H < \theta < 2\hat{\theta} - \hat{\theta}_L) = 1 - \alpha, \quad (2.28)$$

em que

$$\hat{\theta}_L = \varepsilon_L + \hat{\theta} \text{ e } \hat{\theta}_H = \varepsilon_H + \hat{\theta}. \quad (2.29)$$

Se a distribuição de bootstrap é assimétrica os intervalos de confiança (IC) pelos dois metodos não deverão apresentar resultados semelhantes.

### 2.7.3 Bootstrap com correção de viés acelerado.

O teste de bootstrap com correção de viés acelerado foi introduzido por Efron e Tibshirani (1986), o qual é baseado na existência da transformação  $f(\hat{\theta})$  de  $\hat{\theta}$ , tal que  $f(\hat{\theta})$  terá distribuição normal com média  $f(\hat{\theta}) - Z_0 \{1 + af(\theta)\}$  e desvio padrão  $1 + af(\theta)$ , em que  $a$  e  $Z_0$  são constantes reais. Assumindo que, para o momento,  $a$ ,  $Z_0$  e  $f(\theta)$  são conhecidos, então:

$$P \left\{ -Z_{\alpha/2} < \frac{f(\hat{\theta}) - f(\theta) + Z_0 [1 + af(\theta)]}{1 + af(\theta)} < Z_{\alpha/2} \right\} = 1 - \alpha. \quad (2.30)$$

Rearranjando as desigualdades:

$$P \left\{ \frac{f(\hat{\theta}) + Z_0 - Z_{\alpha/2}}{1 - a(Z_0 - Z_{\alpha/2})} < f(\theta) < \frac{f(\hat{\theta}) + Z_0 + Z_{\alpha/2}}{1 - a(Z_0 + Z_{\alpha/2})} \right\} = 1 - \alpha \quad (2.31)$$

ou

$$P\{L < f(\theta) < U\} = 1 - \alpha. \quad (2.32)$$

Para calcular o limite L por bootstrap, os seguintes resultados podem ser usados:

$$P\{f(\hat{\theta}_j) < L\} = P \left\{ f(\hat{\theta}_j) < \frac{f(\hat{\theta}) + Z_0 - Z_{\alpha/2}}{1 - a(Z_0 - Z_{\alpha/2})} \right\}. \quad (2.33)$$

Logo,

$$\begin{aligned} P\{f(\hat{\theta}_j) < L\} &= P \left\{ \frac{f(\hat{\theta}_j) - f(\hat{\theta})}{1 + af(\hat{\theta})} + Z_0 < \frac{Z_0 - Z_{\alpha/2}}{1 - a(Z_0 - Z_{\alpha/2})} + Z_0 \right\}; \\ &= P \left\{ Z < \frac{Z_0 - Z_{\alpha/2}}{1 - a(Z_0 - Z_{\alpha/2})} + Z_0 \right\} = P_L \end{aligned} \quad (2.35)$$

então,

$$Z_L = \frac{Z_0 - Z_{\alpha/2}}{1 - a(Z_0 - Z_{\alpha/2})} + Z_0 \quad (2.34)$$



e

$$P_L = F(Z_L). \quad (2.35)$$

Não é possível trabalhar com a função de transformação, pois ela não é conhecida. Devido à transformação ser crescente e monotônica, esse é o limite inferior para  $q$ . Dessa forma, o limite inferior é o valor da distribuição de bootstrap que exerce a quantidade  $P_L = F(Z_L)$ . Analogamente, o limite superior é o limite que supera  $P_U = F(Z_U)$ , em que

$$Z_U = \frac{Z_0 + Z_{a/2}}{1 - a(Z_0 + Z_{a/2}) + Z_0}. \quad (2.36)$$

O valor de  $Z_0$  pode ser obtido como no caso anterior, sendo dado por:

$$P(\hat{q}_j > \hat{q}) = P(Z > Z_0) = P; \quad (2.37)$$

então,

$$Z_0 = Z_P, \quad (2.38)$$

em que  $P$  é a proporção de estimativas de bootstrap que superam  $\hat{q}$ .

A constante  $a$  não pode ser derivada de uma forma simples. Uma aproximação é apresentada por Efron e Tibshirani (1993), baseada em Jackknife.

$$a \approx \frac{\sum_{i=1}^n (\hat{\theta}_\bullet - \hat{\theta}_{-i})^3}{6 \left[ \sum_{i=1}^n (\hat{\theta}_\bullet - \hat{\theta}_{-i})^2 \right]^{3/2}}, \quad (2.39)$$

em que

$$\hat{\theta}_\bullet = \frac{\sum_{i=1}^n \hat{\theta}_{-i}}{n} \quad (2.40)$$

e  $\hat{\theta}_{-i}$  é o estimador de  $\theta$  eliminando a  $i$ -ésima observação.

## 2.8 Simulação Monte Carlo

A simulação é usada para servir como uma primeira avaliação de um sistema para gerar novas estratégias de ação e regras de decisões antes de se correr o risco de experimentá-las no sistema real. Este procedimento já era usado desde os primeiros tempos pelo homem (Naylor et al. 1988).

O nome Monte Carlo está relacionado com a cidade de mesmo nome, no Principado de Mônaco. O nome é originado principalmente devido aos jogos de azar, principalmente devido à roleta, que é um mecanismo simples para gerar números aleatórios. Esses números eram gerados manualmente ou mecanicamente. Modernamente são usados computadores para gerá-los. Esses números na verdade são números pseudo-aleatórios (Morettin e Bussab, 2003). Assim, Monte Carlo é usado para o uso de técnicas computacionais em que geram amostras de acordo com determinadas distribuições teóricas conhecidas,

visando estudar o comportamento de diferentes técnicas estatísticas que poderiam ser empregadas num dado problema (Dachs, 1988)

## 2.9 Taxas de erros tipo I e tipo II e Poder

Quando se faz um teste de hipóteses, é impossível evitar completamente cometer os erros tipo I e tipo II. O ideal é procurar manter uma probabilidade relativamente pequena de cometê-los (Mayer, 1984).

O erro tipo I é o erro que se comete ao rejeitar a hipótese nula, quando esta é verdadeira. A probabilidade de se cometer o erro tipo I, representado por  $\alpha$ , é identificada como o nível de significância do teste. O pesquisador controla a taxa de erro tipo I decidindo o nível de risco  $\alpha$  que ele está disposta a tolerar, em termos de rejeitar a hipótese nula quando ela é efetivamente verdadeira. Quando o nível de significância é especificado antes de o teste de hipótese ser realizado, o risco de cometer em erro tipo I está diretamente sob controle.

O erro tipo II, por outro lado, é o erro que se comete ao aceitar a hipótese nula, quando na verdade esta é falsa, e a probabilidade de cometê-lo é igual a  $\beta$ . Este erro não é possível ser controlado diretamente (Mood, Graybill e Boes, 1974). Somente com uma amostra suficientemente grande pode reduzir a probabilidade de cometer o erro tipo II ( $\beta$ ), fixando um baixo risco para o erro tipo I ( $\alpha$ ) (Guerra e Donaire, 1982).

| Decisão Estatística | Situação efetiva         |                          |
|---------------------|--------------------------|--------------------------|
|                     | $H_0$ Verdadeiro         | $H_0$ Falso              |
| Não rejeitar $H_0$  | Confiança ( $1-\alpha$ ) | Erro tipo II ( $\beta$ ) |
| Rejeitar $H_0$      | Erro tipo I ( $\alpha$ ) | Eficácia ( $1-\beta$ )   |

O poder de um teste é definido como a probabilidade de rejeitar a hipótese nula  $H_0$ , quando ela é falsa, sendo dado por  $(1 - \beta)$  (Mood, Graybill e Boes, 1974). O poder depende do valor de  $\alpha$ , da hipótese alternativa e do tamanho da amostra (Tonhasca Jr, 1991). Em um teste decisório deve-se considerar os erros tipo I e tipo II. Preferencialmente, deve-se determinar uma amostra que seja suficientemente grande para poder reduzir a probabilidade de se cometer o erro tipo II ( $\beta$ ), fixando um baixo risco  $\alpha$  para a probabilidade de se cometer o erro tipo I (Guerra e Donaire, 1982).

## 2.10 Distribuições

### 2.10.1 Distribuição Normal

A distribuição normal é uma das distribuições mais utilizadas e a mais importante de uma variável aleatória contínua. Variáveis como peso, altura, volume, dentre inúmeras outras, são em geral pressupostas como obedecendo a uma distribuição normal (Bearzoti, 1998).

Uma variável aleatória  $X$  tem distribuição normal se sua função densidade é dada por:

$$f_x(x) = f_x(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (2.41)$$

em que  $\mu$  e  $\sigma$  satisfazem  $-\infty < \mu < \infty$  e  $\sigma > 0$ .

Se uma variável aleatória  $X$  tem distribuição normal média  $\mu$  e variância  $\sigma^2$ , denota-se por:  $X \sim N(\mu, \sigma^2)$  (Drapper e Smith, 1998). Se uma variável aleatória

Z tem distribuição normal média 0 e variância 1, ela possui distribuição denominada normal padrão e sua função densidade é dada por:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}. \quad (2.42)$$

### 2.10.2 Distribuição Uniforme

A distribuição uniforme é utilizada quando uma variável aleatória está contida em um intervalo ao longo do qual qualquer valor é igualmente provável. Portanto, uma variável aleatória X tem distribuição uniforme se sua função densidade é dada por:

$$f_x(x) = f_x(x; \alpha, \beta) = \frac{1}{\beta - \alpha} I_{[\alpha, \beta]}(x), \quad (2.43)$$

em  $I_{[\alpha, \beta]}(x)$  é uma função indicadora de X que retorna o valor 1 se  $x \in [\alpha; \beta]$ , e os parâmetros  $\alpha$  e  $\beta$  satisfazem  $-\infty < \alpha < \beta < \infty$ .

Assim, X tem distribuição uniforme no intervalo  $[\alpha, \beta]$  (Drapper e Smith, 1998). A distribuição uniforme é geralmente aplicada aos procedimentos de amostragem em estatística de qualquer densidade. Quando se diz que um elemento da população foi tomado ao acaso, isso significa que todo elemento apresenta a mesma probabilidade de ser tomado. Nesse caso, o procedimento de coleta está associado à distribuição uniforme (Bearzoti, 1998).

### 2.10.3 Distribuição exponencial

A distribuição exponencial é utilizada no estudo de certas variáveis cujas realizações vão se tornando menos prováveis à medida que aumentam (Bearzoti, 1998). Uma variável aleatória  $X$  tem distribuição exponencial se sua função densidade é dada por:

$$f_x(x) = f_x(x; \lambda) = \lambda e^{-\lambda x} I_{[0, \infty)}(x), \quad (2.44)$$

em que  $\lambda > 0$  (Draper e Smith, 1998).

Essa distribuição é amplamente utilizada em linhas de espera ou na teoria de filas para medir o tempo decorrido entre as chegadas, no âmbito da prestação de serviços de pedágios de pontes, nos caixas automáticos ou numa sala de emergência de um hospital (Levine, Berenson e Stephan, 2000)

### 3 METODOLOGIA

Para avaliar a performance de vários testes para autocorrelação residual de primeira ordem, foi realizado esse trabalho. Os testes avaliados consistem daqueles rotineiramente empregados nos testes de hipótese sobre autocorrelação. As suas respectivas adaptações para bootstrap foram implementadas. Para realizar as avaliações utilizou-se simulação Monte Carlo. Para isso foi adaptado o modelo linear apresentado:

$$y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + b_k X_{k,i} + u_i, \quad (3.1)$$

em que  $y_i$  é a variável dependente relativa à  $i$ -ésima observação simulada,  $b_j$  é o parâmetro associado à  $j$ -ésima variável para  $j = 1, 2, \dots, (k-1)$  ou o intercepto para  $j = 0$ ;  $u_i$  segue um processo de autocorrelação de ordem 1, AR(1), dado pelo seguinte modelo:

$$u_i = ru_{i-1} + e_i \quad |r| < 1, \quad (3.2)$$

em que  $r$  é o parâmetro de autocorrelação e  $e_i$  é supostamente independente de  $u_{i-1}$ , para todo  $i, i = 1, 2, \dots, n$ .

A hipótese a ser testada é dada por:

$$H_0 : r = 0 \text{ vs } H_1 : r > 0. \quad (3.3)$$

Os valores de  $b_j$  foram fixados em 1, sem perda de generalidade, e os valores de  $X$  foram gerados aleatoriamente a partir de uma distribuição uniforme

(0,1), conforme procedimentos descrito por Jeong e Chung (2001). Para o ruído branco consideraram-se diferentes distribuições, quais sejam: normal, uniforme e exponencial. Todas as distribuições consideradas possuíam média 0 e variância 1. No caso da exponencial e da uniforme, considerou-se um modelo centrado em zero. Com a distribuição normal, procurou-se propiciar um ambiente favorável ao teste original de Durbin-Watson (1950); com as demais distribuições, ambientes desfavoráveis foram proporcionados.

Os valores paramétricos de  $\rho$  foram dados por 0,0; 0,1; 0,3; 0,5; 0,7 e 0,9 e os tamanhos amostrais ( $n$ ) considerados foram 10; 20; 50; 100 e 200. Os valores de  $k$  considerados foram 3; 5; 10 e 20, sendo que suas utilizações dependeram do tamanho da amostra considerado. Para cada configuração foram simuladas 2000 amostras de Monte Carlo; no caso dos testes que envolveram bootstrap, 2000 amostras com reposição foram retirados de cada amostra de Monte Carlo. Em cada configuração foram aplicados 8 testes e o número de resultados significativos (erro tipo I e poder) nas 2000 amostras de Monte Carlo, considerando significância com mais de 1% e 5%, foi computado. Os testes aplicados estão discutidos nas seções subseqüentes.

### 3.1) Teste de Durbin-Watson (DW)

Reapresentado o modelo (3.1) matricialmente, obtém-se:

$$\underline{y} = \underline{x}\underline{\beta} + \underline{u}. \quad (3.4)$$

Assim, inicialmente foi obtido o estimador de quadrados mínimos para  $\underline{\beta}$  (Drapper e Smith, 1998). Esse estimador é dado por:

$$\hat{\beta} = (x'x)^{-1}x'y. \quad (3.5)$$

O vetor de erros  $\underline{u}$  foi estimado por:

$$\hat{u} = y - x\hat{\beta}. \quad (3.6)$$

Assim, o estimador da estatística de Durbin-Watson foi obtido por:

$$d = \frac{\hat{u}'A\hat{u}}{\hat{u}'\hat{u}}, \quad (3.7)$$

em quem a matriz  $A_{(n \times n)}$  tem a seguinte estrutura:

$$A \equiv \begin{bmatrix} 1 & -1 & 0 & . & . & . & . & . & 0 \\ -1 & 2 & -1 & . & . & . & . & . & 0 \\ 0 & -1 & 2 & & & & & & 0 \\ . & & & . & & & & & . \\ . & & & & . & & & & . \\ . & & & & & . & & & . \\ . & & & & & & 2 & -1 & \\ 0 & 0 & 0 & . & . & . & -1 & 1 & \end{bmatrix}. \quad (3.8)$$

Os valores de  $d$  foram confrontados com os limites críticos da tabela de Durbin-Watson (1977). Assim,  $H_0$  deve se rejeitada se  $d < d_1$ .

### 3.2 Teste de DW bootstrap (DWB)

Para cada amostra simulada foi aplicado o teste de DW na sua versão de bootstrap. Utilizou-se para isso o procedimento de bootstrap recursivo apresentado por Jeong e Chung (2001) e descrito a seguir.

a) obter uma estimativa de  $\beta$  por quadrados mínimos ordinários utilizando a equação (3.5). Obter a estimativa de  $\hat{u}$  pela equação (3.6);

b) estimar  $\rho$  a partir de  $\hat{u}$  e calcular  $\hat{\epsilon}$ . O estimador de quadrados mínimos de  $\rho$  (Drapper e Smith, 1998) é dado por:

$$\hat{\rho} = \frac{\sum_{i=1}^n (\hat{u}_i \hat{u}_{i-1})}{\sum_{i=2}^n \hat{u}_i^2}; \quad (3.9)$$

c) reamostrar  $\hat{\epsilon}$  para construir um vetor de bootstrap residual  $\hat{\epsilon}^*$ ;

d) recursivamente construir o vetor de resíduos  $\hat{u}^*$  pela equação (3.2).

Neste estágio,  $H_0: \rho = 0$  deve ser imposta artificialmente para a construção da distribuição nula; portanto,  $\hat{u}^*$  é igual a  $\hat{\epsilon}^*$ ;

e) usando  $x$  e  $\hat{u}^*$ , criar os pseudo-valores  $\hat{y}^*$  pela equação (3.1);

f) usando  $x$  e  $\hat{y}^*$ , calcular a estatística de DWB ( $d^*$ ) por:

$$d^* = \frac{\hat{u}^{*'} A \hat{u}^*}{\hat{u}^{*'} \hat{u}^*}, \quad (3.10)$$

em que A está descrita na equação (8);

g) repetir 2000 vezes os passos de (c) a (f) para construir a distribuição de bootstrap de  $d^*$  ( $F_d^*$ );

h) testar  $H_0$  comparando os valores críticos de  $F_d^*$  com o valor de  $d$  estimado a partir dos dados originais.

A proporção de simulações Monte Carlo em que  $H_0: \rho = 0$  foi rejeitada é armazenada. Um total de  $4 \times 10^6$  análises foi realizado para cada configuração. É conveniente salientar que esse teste está sendo aplicado sob a hipótese nula de  $\rho = 0$ , pois não existe um valor nulo da distribuição de Durbin-Watson.

### 3.3 Teste de bootstrap $\rho$ ( $B\rho$ )

O teste de bootstrap é aplicado diretamente no estimador de quadrados mínimos da autocorrelação de primeira ordem. Para isso os passos de (a) a (c) da seção 3.2 são repetidos. Os passos necessários são dados por:

(a) a (c) mesmo do DWB;

d) a partir do resíduo  $e_i^*$ , fez-se  $\hat{u}_i^* = e_i^*$ ; os demais valores do vetor  $\hat{u}^*$  são recriados recursivamente por:

$$\hat{u}_i^* = \hat{\rho} \hat{u}_{i-1}^* + \hat{e}_i^*, \quad i = 2, 3, \dots, n; \quad (3.11)$$

e) uma estimativa de  $\rho^*$  é obtida por quadrados mínimos ordinários:

$$\hat{\rho}^* = \frac{\sum_{i=1}^n (\hat{u}_i^* \hat{u}_{i-1}^*)}{\sum_{i=2}^n \hat{u}_i^{2*}}; \quad (3.12)$$

f) repetir os passos de (c) a (e) para construir a distribuição de bootstrap de  $\hat{\rho}^*$  ( $F_{\hat{\rho}^*}$ );

g) testar  $H_0 : \rho = 0$  comparando os valores críticos de  $F_{\hat{\rho}^*}$  com o valor nulo de  $\rho$  (zero).

Novamente foi computada a proporção de rejeições de  $H_0$  nas 2000 amostras de Monte Carlo geradas por cada configuração.

### 3.4 Teste bootstrap de $\rho$ com correção de viés acelerado (BpCa)

No item 3.3 foi apresentado o estimador de  $\rho$  ( $\hat{\rho}$ ) na amostra original e os estimadores de bootstrap ( $\hat{\rho}_b$ ). Em cada reamostragem verificou se  $\hat{\rho}_b > \hat{\rho}$ . A proporção de vezes em que essa desigualdade foi verificada nas 2000 reamostragens de bootstrap foi computada. Assim, o estimador dessa proporção é definido por:

$$\hat{P}(\hat{\rho}_b > \hat{\rho}) = P. \quad (3.13)$$

O próximo passo é estimar  $Z_0$  por :

$$Z_0 = \Phi^{-1}(1 - P), \quad (3.14)$$

em que  $\Phi^{-1}(\cdot)$  é a inversa da função de distribuição da normal padrão. Da mesma forma foi obtido  $Z_\alpha$  (quantil  $\alpha$  superior da normal padrão), uma vez que o teste é unilateral, por:

$$Z_\alpha = \Phi^{-1}(1-\alpha). \quad (3.15)$$

O valor da constante “a” dever ser estimado por:

$$a \approx \frac{\sum_{i=1}^n (\hat{\rho}_\cdot - \hat{\rho}_{-i})^3}{6 \left[ \sum_{i=1}^n (\hat{\rho}_\cdot - \hat{\rho}_{-i})^2 \right]^{3/2}}, \quad (3.16)$$

em que  $\hat{\rho}_{-i}$  é estimador de  $\rho$  da amostra original simulada eliminando-se a  $i$ -ésima observação obtida, utilizando a expressão (3.9); e  $\hat{\rho}_\cdot$  é dado por:

$$\hat{\rho}_\cdot = \frac{\sum_{i=1}^n \hat{\rho}_{-i}}{n}. \quad (3.17)$$

Com esses valores determinam-se, em seguida, as seguintes quantidades:

$$Z_1 = \frac{Z_0 - Z_\alpha}{[1 - a(Z_0 - Z_\alpha)] + Z_0} \quad (3.18)$$

e

$$\rho_1 = \Phi(Z_1). \quad (3.19)$$

Com os valores de  $\rho_I$  e a amostra ordenada obteve-se o quantil  $100\rho_I\alpha 100\%$  da distribuição de bootstrap de  $\rho$ . Esse limite foi denominado de  $\hat{\rho}_I$ . Assim, a significância empírica foi estimada computando-se a proporção de vezes em que o valor hipotético de  $\rho$  ( $H_0 : \rho = 0$ ) foi inferior a  $\hat{\rho}_I$  nas 2000 simulações de Monte Carlo para cada configuração.

### 3.5 Teste t ( $t_\rho$ )

Para cada simulação realizada aplica-se o teste clássico de t para  $\rho = 0$  de uma distribuição normal bivariada. Esse teste adaptado para autocorrelação é dado por:

$$t_c = \frac{\hat{\rho}\sqrt{n-3}}{\sqrt{1-\hat{\rho}^2}}. \quad (3.20)$$

Sob  $H_0 : \rho = 0$ , a distribuição de  $t_c$  é aproximadamente uma t de Student com  $v = n - 3$  graus de liberdade. Computou-se a proporção de vezes em que o teste apresentou-se significativo nos níveis nominais adotados, da mesma forma como foi realizado anteriormente.

### 3.6 Teste t de bootstrap ( $t_\rho B$ )

Em cada simulação realizada o procedimento de bootstrap da seção 3.2 foi implementado e as estimativas de  $\hat{\rho}_b$  foram obtidas. A essas 2000 estimativas o teste de equação (3.20) foi aplicado, substituindo  $\hat{\rho}$  por  $\hat{\rho}_b$ , obtendo  $t_b$ . Computou-se a proporção de vezes em que o  $t_b$  superou  $t_c$  da amostra original para obter a significância de bootstrap. Nas 2000 amostras de Monte Carlo, computou-se a proporção de vezes em que essa significância foi inferior aos níveis nominais adotados (1% e 5%). Essas foram as estimativas do erro tipo I ou do poder para esse teste.

### 3.7 Teste normal para $\rho$ ( $N\rho$ )

Inicialmente, o erro padrão de  $\hat{\rho}$  de (De Carlo, Tryon, 1993) foi estimado por:

$$S_{(\hat{\rho})} = \frac{n-2}{n\sqrt{n-1}}. \quad (3.21)$$

Posteriormente, aplicou-se um teste normal padrão sob  $H_0: \rho = 0$  da seguinte forma:

$$Z_c = \frac{\hat{\rho}n\sqrt{n-1}}{n-2}. \quad (3.22)$$

Rejeitou-se  $H_0: \rho = 0$  se  $P(Z > Z_c) \leq \alpha$ , em que  $\alpha$  é o valor nominal de significância. Computou-se a proporção de vezes em que a hipótese foi rejeitada nas 2000 simulações.

### 3.8 Teste normal para $\rho$ com correção de viés (N $\rho$ c)

O estimador de  $\hat{\rho}$  é viesado (De Carlo e Tryon, 1993), então aplica-se a correção de viés dada por:

$$r_{1+} = \hat{\rho} + \frac{1}{n} \quad (3.23)$$

Da mesma forma como apresentado na seção (3.7), aplicou-se o teste normal padrão para  $H_0: \rho = 0$ . Para isso, substitui-se  $\hat{\rho}$  por  $r_{1+}$  na equação (3.22). Os demais foram exatamente os mesmos descritos em 3.7.

### 3.9 Teste normal para C de Young (1941) - (NC)

Os estimadores  $\hat{\rho}$  e  $r_{1+}$  são baseados em somas de produtos. A estatística C proposta por Young (1941) é um estimador de  $\rho$  baseado em somas de diferenças. Esse estimador foi utilizado e é dado por:

$$C = 1 - \frac{\sum_{i=1}^{n-1} (\bar{u}_i - \bar{u}_{i+1})^2}{2 \sum_{i=1}^n (\bar{u}_i - \bar{u})^2}, \quad (3.24)$$

em que  $\bar{u} = \frac{\sum_{i=1}^n \hat{u}_i}{n} = 0$  (por definição).

Esse estimador pode ser obtido utilizando a sua relação com  $\hat{\rho}$  da eq. (3.9) ou da relação d da equação (3.7) por:

$$C = \hat{\rho} + \frac{(\bar{u}_n - \bar{u})^2 + (\bar{u}_i - \bar{u})^2}{2 \sum_{i=1}^n (\bar{u}_i - \bar{u})^2} = 1 - \frac{d}{2}. \quad (3.25)$$

O estimador de erro padrão de C é dado por:

$$S_{(C)} = \sqrt{\frac{n-2}{n^2-1}}. \quad (3.26)$$

Para o teste de hipótese  $H_0 : \rho = 0$ , aplicou-se o teste:

$$Z_c = \frac{C\sqrt{n^2-1}}{\sqrt{n-2}}, \quad (3.27)$$

o qual segue a normal padrão assintoticamente. Se  $Z_c > Z_\alpha$  ou  $P(Z_c > Z_\alpha) \leq \alpha$ , rejeita-se a hipótese nula. Novamente, computou-se a proporção de vezes em que a hipótese nula foi rejeitada nas 2000 simulações.

### 3.10 Teste bootstrap para C (BC)

Em cada simulação aplicou-se o procedimento de bootstrap da seção 3.2 e obteve-se a estimativa C da equação (3.25), denominada  $C_b$ . O seu valor padronizado foi obtido utilizando a equação (3.27), substituindo C por  $C_b$  e

denominando de  $Z_{Cb}$ . Computou-se a proporção de vezes em que  $Z_{Cb}$  superou  $Z_c$  da amostra original de Monte Carlo. Nas 2000 amostras de Monte Carlo calculou-se a proporção de vezes em que essa significância de bootstrap foi inferior aos níveis nominais (1% e 5%).

### 3.11 Viés e Erro Quadrático Médio (EQM) dos três estimadores de $\rho$

O viés e o erro quadrático médio dos três estimadores de  $\rho$  ( $\hat{\rho}$ ,  $r_{1+}$  e  $C$ ) foram computados nas 2000 simulações, conforme descrição apresentada a seguir. Seja  $\hat{\Theta}$  um estimador de  $\rho$  que representa  $\hat{\rho}$ ,  $r_{1+}$  ou  $C$ ; então, o viés foi dado por:

$$\text{Viés}(\hat{\Theta}) = \frac{\sum_{m=1}^{2000} \hat{\Theta}_m}{2000} - \rho. \quad (3.28)$$

E o erro quadrático médio (EQM) de  $\hat{\Theta}$  foi dado por:

$$\text{EQM}(\hat{\Theta}) = \frac{\sum_{m=1}^{2000} (\hat{\Theta}_m - \rho)^2}{2000}. \quad (3.29)$$

Os estimadores de  $\rho$  foram comparados em relação às suas propriedades estatísticas.

## 4 RESTADOS E DISCUSSÃO

### 4.1 Propriedades dos estimadores de $\rho$

Na Tabela 4.1.1 estão apresentados o viés dos estimadores de  $\hat{\rho}$ ,  $r_{1+}$  e C do parâmetro de autocorrelação  $\rho$  em função do tamanho da amostra. Para a situação de  $\rho=0$ , o viés de  $\hat{\rho}$  é negativo e em módulo superior aos vieses de  $r_{1+}$  e C. Assim, o  $\hat{\rho}$  subestima  $\rho$  e, em geral, os estimadores de  $r_{1+}$  e C superestimam o valor paramétrico. Com o aumento do valor de  $n$ , como é apontado por De Carlo e Tryon (1993), os estimadores de  $r_{1+}$  e C aproximaram  $\hat{\rho}$ . Assim, os vieses tendem a ser semelhantes. Para pequenas amostras com  $\rho=0$ , os vieses de  $r_{1+}$  e C são menores do que o viés de  $\hat{\rho}$ . Sendo as situações de pequenas amostras as situações mais críticas, os estimadores  $r_{1+}$  e C, por apresentarem menores vieses, são recomendados.

Com o aumento do valor paramétrico de  $\rho$  observou-se, para os três estimadores, um aumento do viés em modulo. Os vieses para essas situações dos três estimadores são negativos. Assim, por exemplo, o viés de  $\hat{\rho}$  para  $\rho = 0$  é de -0,095 e passa a ser de -0,650 para  $\rho$  igual a 0,9 com  $k = 3$ . Os vieses de C e  $r_{1+}$  são bastante similares para todos os valores de  $\rho$  e  $n$  para  $k = 3$ . O aumento do valor de  $n$  diminui consideravelmente os vieses observados, embora se verifique que com o aumento de  $\rho$ , amostras maiores são necessárias para se obter o mesmo viés para um  $n$  fixo e  $\rho$  inferior. Assim, por exemplo, para  $k = 3$ ,  $n = 100$  e  $\rho = 0,1$ , o viés de C foi -0,009. Nessa mesma situação, exceto pela  $\rho = 0,9$ , o viés passou a ser de -0,060. Para se ter o mesmo viés da situação de  $\rho = 0,1$ , seria necessário trabalhar com  $n$  bem maior do que 100.

**Tabela 4.1.1** Viés dos estimadores de  $\hat{f}$ ,  $r_{1+}$  e C do parâmetro do autocorrelação  $\rho$ ,  $k = 3$  em função do tamanho da amostra.

| $\rho$ | n         |          |         |           |          |         |
|--------|-----------|----------|---------|-----------|----------|---------|
|        | 10        |          |         | 20        |          |         |
|        | $\hat{f}$ | $r_{1+}$ | C       | $\hat{f}$ | $r_{1+}$ | C       |
| 0,0    | -0,0954   | 0,0046   | 0,00365 | -0,0478   | 0,0022   | 0,0021  |
| 0,1    | -0,1526   | -0,0526  | -0,0497 | -0,0729   | -0,0230  | -0,0230 |
| 0,5    | -0,3791   | -0,2791  | -0,2750 | -0,1935   | -0,1435  | -0,1428 |
| 0,9    | -0,6504   | -0,5504  | -0,5222 | -0,3582   | -0,3082  | -0,2942 |
| $\rho$ | n         |          |         |           |          |         |
|        | 50        |          |         | 100       |          |         |
|        | $\hat{f}$ | $r_{1+}$ | C       | $\hat{f}$ | $r_{1+}$ | C       |
| 0,0    | -0,0132   | 0,0068   | 0,0070  | -0,0121   | -0,0021  | -0,0022 |
| 0,1    | -0,0306   | -0,0105  | -0,0104 | -0,0193   | -0,0093  | -0,0088 |
| 0,5    | -0,0833   | -0,0633  | -0,0644 | -0,0431   | -0,0331  | -0,0336 |
| 0,9    | -0,1454   | -0,1254  | -0,1250 | -0,0690   | -0,0590  | -0,0603 |

Na Tabela 4.1.2 estão apresentados os vieses de  $\hat{f}$ ,  $r_{1+}$  e C para  $k = 10$  devido ao fato de se ter observado comportamento similar para valores de  $k$  intermediários entre 3 e 10; somente os extremos foram apresentados. Para os estimadores C e  $r_{1+}$  com  $\rho = 0$ , nota-se que deve haver um ponto de equilíbrio considerando  $n$  e  $k$  em que o viés será nulo. À medida que  $\rho$  afasta de zero, o viés aumenta, para esses estimadores considerando  $n$  fixado, com o aumento de  $k$ .

**Tabela 4.1.2** Viés dos estimadores de  $\hat{f}$ ,  $r_{1+}$  e  $C$  do parâmetro do autocorrelação  $\rho$ , para  $k = 10$  em função de  $n$  e  $\rho$ .

| $\rho$ | $n$       |          |         |           |          |         |
|--------|-----------|----------|---------|-----------|----------|---------|
|        | 10        |          |         | 20        |          |         |
|        | $\hat{f}$ | $r_{1+}$ | $C$     | $\hat{f}$ | $r_{1+}$ | $C$     |
| 0,0    | -         | -        | -       | -0,0518   | -0,0018  | -0,0011 |
| 0,1    | -         | -        | -       | -0,1051   | -0,0550  | -0,0540 |
| 0,5    | -         | -        | -       | -0,3526   | -0,3026  | -0,3025 |
| 0,9    | -         | -        | -       | -0,6177   | -0,5676  | -0,5604 |
| $\rho$ | $n$       |          |         |           |          |         |
|        | 50        |          |         | 100       |          |         |
|        | $\hat{f}$ | $r_{1+}$ | $C$     | $\hat{f}$ | $r_{1+}$ | $C$     |
| 0,0    | -0,0230   | -0,0030  | -0,0040 | -0,0117   | -0,0017  | -0,0016 |
| 0,1    | -0,0412   | -0,0212  | -0,0210 | -0,0225   | -0,0125  | -0,0124 |
| 0,5    | -0,1488   | -0,1288  | -0,1305 | -0,0714   | -0,0614  | -0,0622 |
| 0,9    | -0,2655   | -0,2455  | -0,2448 | -0,1304   | -0,1204  | -0,1217 |

Na Tabela 4.1.3 está apresentado EQM dos estimadores de  $\hat{f}$ ,  $r_{1+}$  e  $C$  de  $\rho$  para  $k = 3$  em função de  $n$  e de  $\rho$ . Pode-se observar, considerando  $\rho = 0$ , que os EQM de  $\hat{f}$ ,  $r_{1+}$  e  $C$  são bastante similares, embora o  $r_{1+}$  apresente valores menores de EQM que os seus competidores. Devido ao fato de o EQM ser função do viés ao quadrado e da variância do estimador, é natural que seu valor diminua com o aumento de  $n$ . Como os vieses dos três estimadores tomados ao quadrado para  $\rho = 0$  são praticamente quantidades inexpressivas, o EQM se torna praticamente função da variância dos estimadores. Como os EQM's são parecidos para  $\rho = 0$ , infere-se que as variâncias dos três estimadores sejam similares. Conseqüentemente, o aumento de  $\rho$  ira propiciar um aumento do EQM, pois já foi verificado anteriormente que há um aumento no viés. Da mesma forma, espera-se que os EQM's de  $r_{1+}$  e  $C$  sejam ligeiramente menores do que o de  $\hat{f}$ . Esses resultados podem ser comprovados com a consulta da Tabela 4.1.3.

**Tabela 4.1.3 EQM dos estimadores de  $\hat{f}$ ,  $r_{1+}$  e C de  $\rho$  para  $k = 3$  em função de  $n$  e de  $\rho$ .**

| n      |           |          |        |           |          |        |
|--------|-----------|----------|--------|-----------|----------|--------|
| 10     |           |          | 20     |           |          |        |
| $\rho$ | $\hat{f}$ | $r_{1+}$ | C      | $\hat{f}$ | $r_{1+}$ | C      |
| 0,0    | 0,0806    | 0,0715   | 0,0804 | 0,0432242 | 0,0409   | 0,0429 |
| 0,1    | 0,0976    | 0,0779   | 0,0851 | 0,0482994 | 0,0435   | 0,0458 |
| 0,5    | 0,2188    | 0,1529   | 0,1546 | 0,0776732 | 0,0608   | 0,0601 |
| 0,9    | 0,4947    | 0,3746   | 0,3507 | 0,1606625 | 0,1273   | 0,1171 |
| 50     |           |          | 100    |           |          |        |
| $\rho$ | $\hat{f}$ | $r_{1+}$ | C      | $\hat{f}$ | $r_{1+}$ | C      |
| 0,0    | 0,0192    | 0,0191   | 0,0195 | 0,0099    | 0,0097   | 0,0099 |
| 0,1    | 0,0202    | 0,0194   | 0,0199 | 0,0106    | 0,0103   | 0,0137 |
| 0,5    | 0,0241    | 0,0212   | 0,0210 | 0,0098    | 0,0090   | 0,0090 |
| 0,9    | 0,0309    | 0,0255   | 0,0248 | 0,0083    | 0,0070   | 0,0069 |

Na Tabela 4.1.4 estão apresentados os EQM's para  $k = 10$  dos estimadores  $\hat{f}$ ,  $r_{1+}$  e C. Para  $\rho = 0$ , poucas mudanças no EQM foram observadas, pois o viés nessa situação é bem pequeno. Como o viés foi afetado pelo aumento de  $k$ , ou seja, o viés aumentou com o aumento de  $k$ , observa-se na Tabela 4.1.4 que o EQM também aumentou. Assim, embora os comentários anteriormente realizados para  $k = 3$  continuem valendo, em geral, para  $k = 10$ , a alteração ocorreu somente nos níveis de EQM que foram maiores na mesma situação de  $k = 10$  (tabelas 4.1.3 e 4.1.4). As exceções ocorrem geralmente para grandes valores de  $n$ . Esses resultados diferem daqueles apresentados por De Carlo e Tryon (1993) devido a esses autores não considerarem o estudo desse problema em situações de regressão. Os trabalhos de simulação deles envolveram apenas o uso de uma variável autocorrelacionada. O presente trabalho tem o mérito de destacar a influência do número de variáveis regressoras no viés e no EQM, por meio do viés, nos estimadores de  $\rho$ .

Tabela 4.1.4 EQM dos estimadores de  $\hat{\rho}$ ,  $r_{1+}$  e C de  $\rho$  para  $k = 10$  em função de  $n$  e de  $\rho$ .

| $\rho$ | n            |          |        |              |          |        |
|--------|--------------|----------|--------|--------------|----------|--------|
|        | 10           |          |        | 20           |          |        |
|        | $\hat{\rho}$ | $r_{1+}$ | C      | $\hat{\rho}$ | $r_{1+}$ | C      |
| 0,0    | -            | -        | -      | 0,0464       | 0,0437   | 0,0459 |
| 0,1    | -            | -        | -      | 0,0541       | 0,0461   | 0,0477 |
| 0,5    | -            | -        | -      | 0,1674       | 0,1346   | 0,1355 |
| 0,9    | -            | -        | -      | 0,4256       | 0,3663   | 0,3588 |
| $\rho$ | n            |          |        |              |          |        |
|        | 50           |          |        | 100          |          |        |
|        | $\hat{\rho}$ | $r_{1+}$ | C      | $\hat{\rho}$ | $r_{1+}$ | C      |
| 0,0    | 0,0193       | 0,0188   | 0,0191 | 0,0097       | 0,0096   | 0,0096 |
| 0,1    | 0,0205       | 0,0192   | 0,0195 | 0,0101       | 0,0098   | 0,0099 |
| 0,5    | 0,0386       | 0,0330   | 0,0334 | 0,0134       | 0,0121   | 0,0121 |
| 0,9    | 0,0831       | 0,0729   | 0,0721 | 0,0214       | 0,0189   | 0,0191 |

## 4.2 Erro tipo I

Nas Tabelas 4.2.1, 4.2.2 e 4.2.3 estão apresentadas as taxas de erro tipo I para os dez testes estudados neste trabalho, considerando valores nominais de 5% de significância e distribuição normal para os resíduos. A distribuição normal é considerada a situação ideal para o teste DW, uma vez que a teoria de seu desenvolvimento foi baseada em normalidade dos resíduos. O teste de DW apresentou taxas de erro tipo I significativamente inferiores ao valor nominal de  $\alpha = 5\%$  para todos os valores de  $n$  e  $k$  simulados. A hipótese de igualdade das taxas de erro observadas para cada teste ao valor nominal  $\alpha$  (5% ou 1%) foi testada por meio do intervalo de 99% de confiança para proporções (Leemis e Trivedi, 1996). Esse intervalo foi de 3,83% e 6,39%. Assim, as taxas observadas foram confrontadas com esses limites para o teste de hipótese de igualdade ao valor nominal de 5%. Essa hipótese foi testada para eliminar os

erros de Monte Carlo presentes na amostra de tamanho 2000 para cada configuração.

Dessa forma, o teste DW foi considerado rigoroso. Com o aumento do valor de  $k$ , para um mesmo valor de  $n$ , verifica-se que a taxa de erro tipo I do teste DW diminui consideravelmente (Tabelas 4.2.1, 4.2.2 e 4.2.3). Da mesma forma, fixando o valor  $k$  e aumentando o valor de  $n$ , verifica-se que a taxa de erro tipo I empírica deste teste (DW) aumenta, embora com  $n \leq 200$  ainda continue sendo significativamente ( $P < 0,01$ ) inferior ao valor nominal. É conveniente salientar que um teste se torna menos rigoroso à medida que o valor real de sua significância diminui. Assim, como as taxas de erro tipo I do teste DW são inferiores a 5%, espera-se que o seu poder seja maior.

**Tabela 4.2.1** Taxas de erro tipo I para os dez testes de autocorrelação em função do número de variáveis  $k = 3$  e do tamanho da amostra para distribuição normal considerando  $\alpha = 5\%$ .

| Testes | n       |         |          |         |         |
|--------|---------|---------|----------|---------|---------|
|        | 10      | 20      | 50       | 100     | 200     |
| DW     | 0,0015* | 0,0065* | 0,0225*  | 0,0275* | 0,0265* |
| DWB    | 0,0515  | 0,0440  | 0,0555   | 0,0510  | -       |
| Bp     | 0,0015* | 0,0135* | 0,03050* | 0,0420  | 0,0325* |
| BpCa   | 0,0310* | 0,0415  | 0,0450   | 0,0450  | 0,0355* |
| tp     | 0,0025* | 0,0140* | 0,0360   | 0,0430  | -       |
| tpB    | 0,0520  | 0,0450  | 0,0585   | 0,0525  | -       |
| Np     | 0,0240* | 0,0270* | 0,0455   | 0,0425  | -       |
| Npc    | 0,0590  | 0,0470  | 0,0590   | 0,0500  | -       |
| NC     | 0,0510  | 0,0480  | 0,0575   | 0,0500  | -       |
| BC     | 0,0515  | 0,0440  | 0,0555   | 0,0510  | -       |

**Tabela 4.2.2** Taxas de erro tipo I para os dez testes de autocorrelação em função do número de variáveis  $k = 10$  e do tamanho da amostra para distribuição normal considerando  $\alpha = 5\%$ .

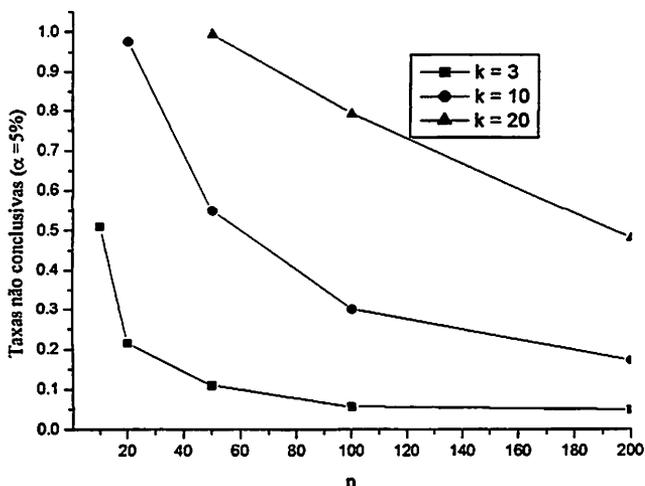
| Testes | N  |         |         |         |         |
|--------|----|---------|---------|---------|---------|
|        | 10 | 20      | 50      | 100     | 200     |
| DW     | -  | 0,0000* | 0,0005* | 0,0010* | 0,0065* |
| DWB    | -  | 0,0490  | 0,0515  | 0,0485  | -       |
| Bp     | -  | 0,0125* | 0,0300* | 0,0420  | 0,0365* |
| BpCa   | -  | 0,0395  | 0,0405  | 0,0465  | 0,0380* |
| tp     | -  | 0,0165* | 0,0265* | 0,0385  | -       |
| tpB    | -  | 0,0490  | 0,0560  | 0,0495  | -       |
| Np     | -  | 0,0290* | 0,0365* | 0,0410  | -       |
| Npc    | -  | 0,0500  | 0,0535  | 0,0500  | -       |
| NC     | -  | 0,0490  | 0,0500  | 0,0500  | -       |
| BC     | -  | 0,0490  | 0,0515  | 0,0485  | -       |

**Tabela 4.2.3.** Taxas de erro tipo I para os dez testes de autocorrelação em função do número de variáveis  $k = 20$  e do tamanho da amostra para distribuição normal considerando  $\alpha = 5\%$ .

| Testes | n  |    |         |         |         |
|--------|----|----|---------|---------|---------|
|        | 10 | 20 | 50      | 100     | 200     |
| DW     | -  | -  | 0,0000* | 0,0000* | 0,0005* |
| DWB    | -  | -  | 0,0420  | 0,0490  | -       |
| Bp     | -  | -  | 0,0260* | 0,0370* | 0,0340* |
| BpCa   | -  | -  | 0,0375* | 0,0395  | 0,0330* |
| tp     | -  | -  | 0,0250* | 0,0400  | -       |
| tpB    | -  | -  | 0,0430  | 0,0480  | -       |
| Np     | -  | -  | 0,0305* | 0,0420  | -       |
| Npc    | -  | -  | 0,0445  | 0,0490  | -       |
| NC     | -  | -  | 0,0435  | 0,0515  | -       |
| BC     | -  | -  | 0,0420  | 0,0490  | -       |

O teste DW possui essa deficiência, possivelmente, devido ao fato de não apresentar resultados conclusivos quanto à rejeição ou não da hipótese de nulidade para a correlação serial em algumas regiões dos valores da estatística  $d$  observada. Nessa simulação foi mensurada a taxa de resultados não conclusivos do teste que podem ser observados na Figura 4.2.1. Observa-se que o aumento do valor de  $n$  proporciona uma redução nas taxas não conclusivas do teste DW.

Para um mesmo valor de  $n$ , observa-se que à medida que o valor de  $k$  aumenta, o valor das taxas não conclusivas também aumenta. Esse resultado é preocupante, uma vez que com 20 variáveis, mesmo para  $n = 200$ , a taxa não conclusiva de DW foi de 48,05%, que é muito alta. Assim, o teste DW é possivelmente muito afetado.



**Figura 4.2.1** Taxas não conclusivas do teste DW para o valor nominal de significância de 5%.

O teste DW na versão bootstrap foi nomeado DWB e os resultados encontrados para as taxas de erro tipo I permitem inferir que uma sensível melhora foi obtida. Isso ocorreu devido ao fato de que todas as taxas de erro tipo I observadas para  $n$  variando de 10 a 200 e  $k$  variando de 3 a 20 foram iguais ao valor nominal de 5%, exceto pelos erros de Monte Carlo presentes nas estimativas dessas taxas. Assim, conclui-se que a versão bootstrap desse teste foi eficiente para o controle da taxa de erro tipo I e que o teste DWB apresentou tamanho do teste igual ao valor nominal  $\alpha = 5\%$ .

Os testes DW e DWB são testes indiretos para a autocorrelação serial. Os testes baseados diretamente na autocorrelação serial estudados nesse trabalho são  $t_p$ ,  $N_p$ ,  $N_{pc}$  e  $NC$ . As taxas de erro tipo I desses testes e de suas respectivas versões de bootstrap se encontram nas Tabelas 4.2.1, 4.2.2 e 4.2.3. Os testes baseados nos estimadores viesados de  $\rho$  são  $t_p$  e  $N_p$ . Para esses dois testes, as taxas de erro tipo I tenderam a ser significativamente diferentes e menores do que o valor nominal  $\alpha = 5\%$  para amostras pequenas. À medida que o tamanho da amostra aumenta esses testes a partir de  $n \geq 100$  apresentam valores semelhantes ao nominal (5%), principalmente com  $k \geq 20$ . O teste normal para  $\rho$  apresentou taxas que convergiam mais rapidamente para o valor nominal com o aumento de  $n$  do que o teste  $t_p$ . Quanto maior for o valor  $k$ , em geral é necessário um maior valor de  $n$  para que as taxas de erro tipo I dos testes  $t_p$  e  $N_p$  sejam semelhantes ao valor nominal de significância.

O teste  $N_{pc}$ , o qual é baseado em um estimador não viesado ( $r_{1+}$ ), apresentou taxas de erro tipo I iguais ao valor nominal em todas as situações estudadas referentes a  $k$  e a  $n$ . A simples correção de viés melhorou consideravelmente o resultado do teste  $N_p$ . O mesmo resultado foi obtido para o teste  $NC$ , em que  $C$  é também um estimador de  $\rho$  com menor viés do que  $\hat{f}$  (De Carlo e Tryon, 1993). Assim, os testes  $N_{pc}$  e  $NC$  possuem tamanhos iguais ao valor nominal de 5% para todas as situações estudadas, sob distribuição normal.

Dentre os testes de bootstrap para  $\rho$ , o teste  $B_p$  foi o que apresentou pior resultado. Para  $k = 3$ , por exemplo, somente com  $n = 100$  a significância encontrada foi semelhante ao valor nominal  $\alpha = 5\%$  ( $P > 0,01$ ). Nos demais valores de  $n$ , inclusive para  $n = 200$ , o valor da taxa de erro tipo I foi inferior ao valor nominal de 5% ( $P < 0,01$ ). Pouca alteração foi observada passando de  $k = 3$  para  $k = 10$ . Para  $k = 20$ , todas as taxas de erro tipo I foram significativamente inferiores ao valor nominal de 5% ( $P < 0,01$ ). Esse teste

tendeu a ser rigoroso e deve apresentar menor poder (Mood, Graybill e Boes, 1974). O teste de bootstrap para  $\rho$  com correção de viés (BpCa) se apresentou melhor no controle da taxa de erro tipo I, em relação Bp. Esse teste tendeu a ser mais rigoroso com pequenas amostras em relação ao k usado. Também foi mais rigoroso, com taxas de erro tipo I inferiores ao valor nominal, à medida que o valor k aumentou (tabelas, 4.2.2, 4.2.2 e 4.2.3). Com o aumento de n para 200 eram esperadas taxas de erro tipo I semelhantes ao valor nominal. Surpreendentemente isso não ocorreu, sendo que o teste BpCa apresentou-se rigoroso para os três valores de k (3, 10 e 20).

Os outros dois testes de bootstrap baseados em quantidades pivôs apresentaram, para todas as situações de k e n, resultados idênticos ao valor nominal, exceto pelo erro de Monte Carlo. Assim, esses dois testes, na situação ideal de normalidade, apresentam tamanho igual ao valor nominal de 5%. É importante enfatizar que nenhum dos 10 testes apresentou taxas de erro tipo I superiores ao valor nominal de 5%.

Nas Tabelas 4.2.4, 4.2.5 e 4.2.6 estão apresentadas as taxas de erro tipo I para os testes considerando a distribuição uniforme e valor nominal de significância de 5%. Os resultados, em geral, concordam com aqueles obtidos para a distribuição normal. As maiores alterações foram observadas para o teste BpCa. Em algumas situações de k e n houve uma piora (teste rigoroso) em relação à situação de normalidade, em outras, os valores das taxas de erro tipo I se igualaram ( $P > 0,01$ ) ao valor nominal de  $\alpha = 5\%$ . Isso pode ser visualizado confrontando as Tabelas 4.2.1, 4.2.2 e 4.2.3 com as Tabelas 4.2.4, 4.2.5 e 4.2.6. Os resultados não conclusivos para o teste de DW foram similares aos da distribuição normal e isso não foi apresentado. Assim, por não alterarem a sua performance com a mudança de distribuição, os testes foram considerados robustos.

**Tabela 4.2.4** Taxas de erro tipo I para os dez testes de autocorrelação em função do número de variáveis  $k=3$  e do tamanho da amostra para distribuição uniforme considerando  $\alpha = 5\%$ .

| Testes | n       |         |         |         |        |
|--------|---------|---------|---------|---------|--------|
|        | 10      | 20      | 50      | 100     | 200    |
| DW     | 0,0000* | 0,0075* | 0,0215* | 0,0040* | -      |
| DWB    | 0,0470  | 0,0585  | 0,0550  | 0,0565  | -      |
| Bp     | 0,0040* | 0,0130* | 0,0325* | 0,0355* | 0,0420 |
| BpCa   | 0,0400  | 0,0345* | 0,0400  | 0,0365* | 0,0420 |
| tp     | 0,0035* | 0,0225* | 0,0360* | 0,0430  | -      |
| tpB    | 0,0485  | 0,0590  | 0,0555  | 0,0580  | -      |
| Np     | 0,0180* | 0,0405  | 0,0415  | 0,0475  | -      |
| Npc    | 0,0505  | 0,0660  | 0,0580  | 0,0595  | -      |
| NC     | 0,0465  | 0,0605  | 0,0545  | 0,0575  | -      |
| BC     | 0,0470  | 0,0585  | 0,0550  | 0,0565  | -      |

**Tabela 4.2.5** Taxas de erro tipo I para os dez testes de autocorrelação em função do número de variáveis  $k=10$  e do tamanho da amostra para distribuição uniforme considerando  $\alpha = 5\%$ .

| Testes | n  |         |         |         |        |
|--------|----|---------|---------|---------|--------|
|        | 10 | 20      | 50      | 100     | 200    |
| DW     | -  | 0,0000* | 0,0010* | 0,0030* | -      |
| DWB    | -  | 0,0435  | 0,0575  | 0,0455  | -      |
| Bp     | -  | 0,0130* | 0,0310* | 0,0355* | 0,0420 |
| BpCa   | -  | 0,0405  | 0,0380  | 0,0365* | 0,0420 |
| tp     | -  | 0,0155* | 0,0385  | 0,0375* | -      |
| tpB    | -  | 0,0455  | 0,0560  | 0,0470  | -      |
| Np     | -  | 0,0265* | 0,0455  | 0,0380* | -      |
| Npc    | -  | 0,0475  | 0,0575  | 0,0475  | -      |
| NC     | -  | 0,0415  | 0,0555  | 0,0445  | -      |
| BC     | -  | 0,0435  | 0,0575  | 0,0455  | -      |

**Tabela 4.2.6** Taxas de erro tipo I para os dez testes de autocorrelação em função do número de variáveis  $k = 20$  e do tamanho da amostra para distribuição uniforme considerando  $\alpha = 5\%$ .

| Testes | n  |    |         |         |         |
|--------|----|----|---------|---------|---------|
|        | 10 | 20 | 50      | 100     | 200     |
| DW     | -  | -  | 0,0000* | 0,0000* | -       |
| DWB    | -  | -  | 0,0485  | 0,052   | -       |
| Bp     | -  | -  | 0,0330* | 0,0315* | 0,0370* |
| BpCa   | -  | -  | 0,0440  | 0,0350* | 0,0370* |
| tp     | -  | -  | 0,0325* | 0,0365* | -       |
| tpB    | -  | -  | 0,0525  | 0,0490  | -       |
| Np     | -  | -  | 0,0360* | 0,0430  | -       |
| Npc    | -  | -  | 0,0520  | 0,0510  | -       |
| NC     | -  | -  | 0,0510  | 0,0535  | -       |
| BC     | -  | -  | 0,0485  | 0,0520  | -       |

Na Tabela 4.2.7 estão apresentados os resultados das taxas de erro tipo I para os testes de autocorrelação considerando  $k = 3$ ,  $\alpha = 5\%$  e distribuição exponencial para gerar uma condição mais inapropriada para distribuição dos resíduos, o qual, diferentemente da distribuição uniforme, é uma distribuição assimétrica à direita. Embora essa condição atípica tenha sido simulada, os testes apresentaram praticamente os mesmos resultados obtidos para as distribuições normal e uniforme. Assim, os testes DWB, tpB, NpC, NC e Bc apresentaram tamanho igual ao nominal de 5%. Os demais testes, em geral, para a maioria das situações, foram considerados rigorosos, ou seja, apresentaram taxas de erro tipo I inferiores ( $P < 0,001$ ) ao valor nominal de 5%. Nesse momento, é oportuno observar que o teste DWB e o teste BC apresentam os mesmos resultados para a taxa de erro tipo I (tabelas 4.2.1 a 4.2.7). Portanto, esse fato fornece um forte indício de que esses testes equivalentes.

**Tabela 4.2.7** Taxas de erro tipo I para os dez testes de autocorrelação em função do número de variáveis  $k=3$  e do tamanho da amostra para distribuição exponencial considerando  $\alpha = 5\%$ .

| Testes | n       |         |         |         |         |
|--------|---------|---------|---------|---------|---------|
|        | 10      | 20      | 50      | 100     | 200     |
| DW     | 0,0025* | 0,0080* | 0,0170* | 0,0270* | 0,0285* |
| DWB    | 0,0505  | 0,0525  | 0,0500  | 0,0485  | -       |
| Bp     | 0,0025* | 0,0140* | 0,0275* | 0,0385  | 0,0365* |
| BpCa   | 0,0335* | 0,0365* | 0,0350* | 0,0400  | 0,0380* |
| tp     | 0,0040* | 0,0195* | 0,0340* | 0,0360* | -       |
| tpB    | 0,0495  | 0,0545  | 0,0535  | 0,0480  | -       |
| Np     | 0,0230* | 0,0340* | 0,0415  | 0,0405  | -       |
| Npc    | 0,0540  | 0,0570  | 0,0545  | 0,0485  | -       |
| NC     | 0,0535  | 0,0510  | 0,0530  | 0,0495  | -       |
| BC     | 0,0505  | 0,0525  | 0,0500  | 0,0485  | -       |

Na Tabela 4.2.8 estão apresentadas as taxas de erro tipo I para os testes em função de  $n$ , para  $k=3$ , distribuição normal e valor nominal de significância de 1%. Para esse caso, o intervalo de 99% de confiança para a taxa de erro tipo I é de 0,52 e 1,73. Assim, testes com taxas empíricas superiores a 1,73% são considerados não rigorosos e inferiores a 0,52%, rigorosos. Os testes que possuem tamanho igual ao valor nominal de 1% terão taxas empíricas entre 0,52% e 1,73%. As oscilações encontradas são devidas ao erro de Monte Carlo. Nesse caso de  $\alpha = 1\%$ , optou-se por apresentar apenas a situação de  $k=3$  distribuição normal, haja vista que comportamentos similares encontrados para  $\alpha = 5\%$  foram observados.

Os testes considerados rigorosos para  $\alpha = 5\%$  continuaram rigorosos para  $\alpha = 1\%$ . O teste DW da mesma forma como ocorreu para 5%, apresentou, para todos os valores de  $n$  estudados, taxas empíricas inferiores ( $P < 0,01$ ) ao valor nominal de 1%. Os melhores testes, ou seja, aqueles que para todos os  $n$  apresentaram tamanhos idênticos ao nominal, foram também os mesmos. A exceção ocorreu para o teste de NC, o qual com  $n \leq 20$ , mostrou-se rigoroso. Assim, o único teste não baseado em bootstrap que se apresentou com controle

do erro tipo I foi o Npc. Isso de certa forma mostra a superioridade dos testes baseados em bootstrap, mesmo para amostras pequenas, como  $n = 10$ .

**Tabela 4.2.8.** Taxas de erro tipo I para os dez testes de autocorrelação em função do número de variáveis  $k = 3$  e do tamanho da amostra para distribuição normal considerando  $\alpha = 1\%$ .

| Testes | n       |         |         |         |         |
|--------|---------|---------|---------|---------|---------|
|        | 10      | 20      | 50      | 100     | 200     |
| DW     | 0,0050* | 0,0005* | 0,0005* | 0,0045* | 0,0050* |
| DWB    | 0,0085  | 0,0070  | 0,0075  | 0,0085  | 0,0090  |
| Bp     | 0,0000* | 0,0050* | 0,0020* | 0,0060  | 0,0060  |
| BpCa   | 0,0095  | 0,0060  | 0,0050* | 0,0070  | 0,0060  |
| tp     | 0,0000* | 0,0020* | 0,0025* | 0,0040* | 0,0070  |
| tpB    | 0,0085  | 0,0085  | 0,0100  | 0,0090  | 0,0095  |
| Np     | 0,0020* | 0,0040* | 0,0045* | 0,0040* | 0,0070  |
| Npc    | 0,0065  | 0,0095  | 0,0070  | 0,0085  | 0,0085  |
| NC     | 0,0045* | 0,0050* | 0,0085  | 0,0070  | 0,0090  |
| BC     | 0,0085  | 0,0070  | 0,0075  | 0,0085  | 0,0090  |

É conveniente salientar que o comportamento dos 5 testes que se destacaram nesse trabalho em relação ao controle do erro tipo I para as distribuições uniforme e exponencial foi similar àquele da distribuição normal. Esse resultado evidencia uma performance robusta desses testes.

### 4.3 Poder

Na Tabela 4.3.1 está apresentado o poder dos 10 testes para correlação serial obtidos mediante simulação Monte Carlo sob  $H_1$ , considerando  $\rho$  igual 0,1, 0,5 e 0,9 em função de  $n$ . De maneira geral, pode-se verificar que o aumento de  $n$  provoca aumento do poder, o que é esperado pela teoria dos testes de hipóteses (Mood, Graybill e Boes, 1974). Também verifica-se que o aumento de  $\rho$  de 0,1 para 0,9 provoca grande aumento no poder dos testes, como é esperado pela teoria. O poder dos testes para  $\rho = 0,1$ , mesmo quando  $n = 20$ , pode ser

considerado baixo (inferior a 4%). Poder da ordem de 80% ou mais para  $\rho = 0,5$  é encontrado com  $n$  entre 20 e 50 e para  $\rho = 0,9$  com  $n$  entre 10 e 20. Para grandes valores de  $n$  e  $\rho$ , os poderes dos testes tendem a se igualar e aproxima-se de 100%.

O teste de DW apresentou o menor poder dentre os 10 testes estudados em todas as configurações apresentadas nas Tabelas 4.3.1, 4.3.2 e 4.3.3. A diferença em relação aos três testes mais poderosos se destaca principalmente para pequenos valores de  $\rho$  e de  $n$ . Em algumas situações, o poder do teste é significativamente ( $P < 0,01$ ) inferior ao valor nominal de significância de 5%. Os resultados para este teste não são similares aos encontrados por Jeong e Chung (2001). O teste  $B_p$  foi inferior a todos os outros, superando apenas em poder o teste DW original. Da mesma forma que o teste DW,  $B_p$  apresentou poder inferior ao valor nominal de 5% para pequenos valores de  $n$  e de  $\rho$ . O mesmo teste com correção de viés,  $B_{pCa}$ , por outro lado, apresentou poder bem maior que o teste sem correção,  $B_p$ . Grandes diferenças foram observados, independentemente do valor de  $\rho$ , para pequenas amostras, o que se torna uma grande vantagem desse teste sobre o teste sem correção de viés. Esse resultado está de acordo com os resultados observados para o viés dos estimadores de  $\rho$  em pequenas amostras nesse trabalho e em De Carlo e Tryon (1993). Os resultados desse trabalho estão, no entanto, bem aquém daqueles obtidos por Jeong Chung (2001). Esses autores encontraram para  $n = 10$ ,  $k = 3$  e  $\rho > 0,7$  poder da ordem de 99% para esse teste, enquanto no presente trabalho o poder não ultrapassou 36%. Não foi encontrada nenhuma explicação plausível para as diferenças encontradas.

**Tabela 4.3.1** Poder dos dez testes de autocorrelação para  $k = 3$ ,  $\rho = 0,1$  em função do tamanho da amostra para distribuição normal considerando  $\alpha = 5\%$ .

| Testes | n      |        |        |        |        |
|--------|--------|--------|--------|--------|--------|
|        | 10     | 20     | 50     | 100    | 200    |
| DW     | 0,0040 | 0,0170 | 0,0745 | 0,1600 | 0,3200 |
| DWB    | 0,0670 | 0,0955 | 0,1645 | 0,2425 | -      |
| Bp     | 0,0015 | 0,0030 | 0,0955 | 0,1975 | 0,3640 |
| BpCa   | 0,0485 | 0,0425 | 0,1275 | 0,2115 | 0,3695 |
| tp     | 0,0055 | 0,0365 | 0,1210 | 0,2020 | -      |
| tpB    | 0,0725 | 0,0995 | 0,1715 | 0,2450 | -      |
| Np     | 0,0355 | 0,0700 | 0,1415 | 0,2130 | -      |
| Npc    | 0,0820 | 0,1040 | 0,1720 | 0,2420 | -      |
| NC     | 0,0740 | 0,1025 | 0,1660 | 0,2410 | -      |
| BC     | 0,0670 | 0,0955 | 0,1645 | 0,2425 | -      |

**Tabela 4.3.2** Poder dos dez testes de autocorrelação em função do número de variáveis  $k = 3$ ,  $\rho = 0,5$  e do tamanho da amostra para distribuição normal considerando  $\alpha = 5\%$ .

| Testes | n      |        |        |        |        |
|--------|--------|--------|--------|--------|--------|
|        | 10     | 20     | 50     | 100    | 200    |
| DW     | 0.0215 | 0,2525 | 0,8680 | 0,9995 | 1,0000 |
| DWB    | 0.2110 | 0,5455 | 0,9365 | 0,9975 | -      |
| Bp     | 0,0210 | 0,3335 | 0,9230 | 0,9985 | 1,0000 |
| BpCa   | 0,1685 | 0,4910 | 0,9375 | 0,9985 | 1,0000 |
| tp     | 0,0425 | 0,3715 | 0,9115 | 0,9970 | -      |
| tpB    | 0,2185 | 0,5530 | 0,9365 | 0,9975 | -      |
| Np     | 0,1285 | 0,4645 | 0,9205 | 0,9970 | -      |
| Npc    | 0,2265 | 0,5595 | 0,9365 | 0,9970 | -      |
| NC     | 0,2090 | 0,5490 | 0,9355 | 0,9975 | -      |
| BC     | 0,2110 | 0,5455 | 0,9365 | 0,9975 | -      |

**Tabela 4.3.3** Poder dos dez testes de autocorrelação em função do número de variáveis  $k = 3$ ,  $\rho = 0,9$  e do tamanho da amostra para distribuição normal considerando  $\alpha = 5\%$ .

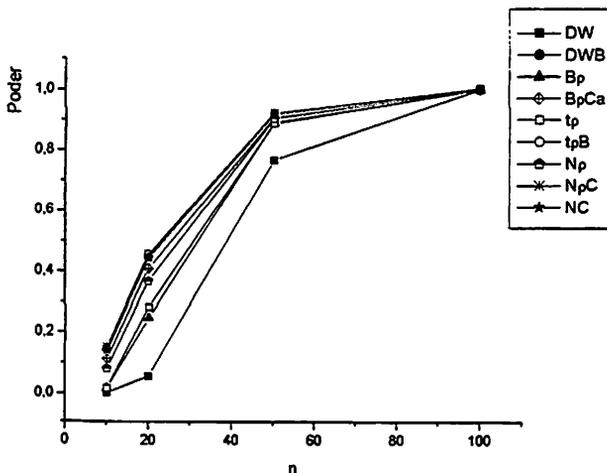
| Testes | n      |        |        |        |        |
|--------|--------|--------|--------|--------|--------|
|        | 10     | 20     | 50     | 100    | 200    |
| DW     | 0,0755 | 0,7585 | 0,9990 | 1,0000 | 1,0000 |
| DWB    | 0,4380 | 0,9145 | 0,9995 | 1,0000 | -      |
| Bp     | 0,0635 | 0,7710 | 1,0000 | 1,0000 | 1,0000 |
| BpCa   | 0,3530 | 0,8615 | 1,0000 | 1,0000 | 1,0000 |
| tp     | 0,0945 | 0,8085 | 0,9995 | 1,0000 | -      |
| tpB    | 0,4040 | 0,9070 | 0,9995 | 1,0000 | -      |
| Np     | 0,2635 | 0,8710 | 0,9995 | 1,0000 | -      |
| Npc    | 0,4110 | 0,9075 | 0,9995 | 1,0000 | -      |
| NC     | 0,4295 | 0,9135 | 1,0000 | 1,0000 | -      |
| BC     | 0,4380 | 0,9145 | 0,9995 | 1,0000 | -      |

O teste tp foi, em ordem crescente de poder, o terceiro colocado, superando os testes DW e Bp. Teve pequeno poder como esses outros dois, para pequenos valores de  $\rho$  e de n. Já o teste Np apresentou performance imediatamente superior ao tp, inclusive para valor de n pequeno e grandes valores de  $\rho$ , tendo poder muito superior ao tp. Por exemplo, se for considerado  $\rho = 0,9$  e  $n = 10$ , o poder do teste Np é de 21,35% e do tp é de 9,45%. Esse teste foi inferior ao teste BpCa e aos demais não mencionados (DWB, TpB, Bc, Npc e NC). Todos esses testes, em geral, apresentaram taxas de erro tipo I inferiores ao valor nominal de 5%. Assim, era esperado que o poder desses testes fosse menor do que o daqueles que apresentaram controle do erro tipo I. Isso de fato aconteceu (Tabela 4.3.1)

Os testes que apresentaram tamanhos iguais a  $\alpha = 5\%$  foram Npc e NC e os testes de bootstrap DWB, tpB e BC. O poder desses testes é discutido a seguir, simultaneamente. Todos esses testes tiveram poderes semelhantes entre n, sendo as diferenças devidas ao erro de Monte Carlo. O poder destes testes foi superior aos dos demais testes anteriormente mencionados (DW, Bp, tp, Np, BpCa). Para amostras pequenas, a diferença entre o valor do

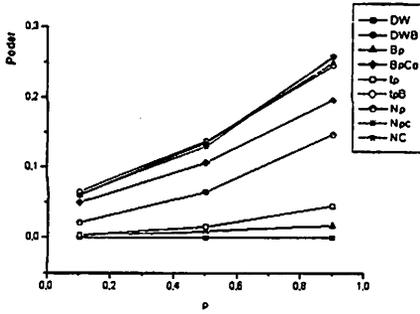
poder desses testes e o do melhor teste, que se mostrou rigoroso no controle do erro tipo I ( $N\rho$ ), foi de 16,55 pontos percentuais. É conveniente salientar que se pode observar que os poderes dos testes DWB e BC foram exatamente os mesmos. As taxas de erro tipo I também são iguais para esses dois testes. Depreende-se, assim, que esses dois testes de bootstrap na verdade são equivalentes. Isso se deve ao fato de que existe uma relação linear perfeita monótona decrescente entre  $C$  e  $d$ , ou seja,  $C = 1 - 0,5d$ . Assim, efetivamente têm-se 9 testes diferentes nesse trabalho.

Na Figura 4.3.1, os poderes dos testes em função de  $n$  para  $K = 5$  e  $\rho = 0,5$  estão apresentados para distribuição normal. O mesmo comportamento ocorre em relação à performance dos testes quanto ao poder, sendo o pior teste o DW e os melhores, os testes NC,  $N\rho$  DWB,  $\tau\rho$ B e BC, que é equivalente a DWB e não está apresentado na Figura 4.3.1. Destaca-se, para um mesmo valor de  $n$ , que os resultados obtidos para o poder são inferiores àqueles apresentados na Tabela 4.3.2 para  $k = 3$ . Assim, o efeito do aumento do valor de  $k$  afeta a magnitude do poder dos testes, diminuindo-o à medida que  $k$  aumenta. Esse efeito é mais marcante para pequenos valores de  $n$ .

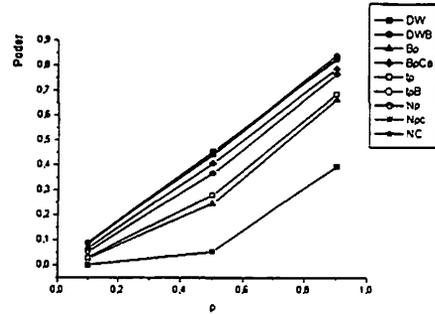


**Figura 4.3.1** Poder dos testes para distribuição normal dos resíduos em função de  $n$  com  $\rho = 0,5$ ,  $k = 5$  para  $\alpha = 0.05$

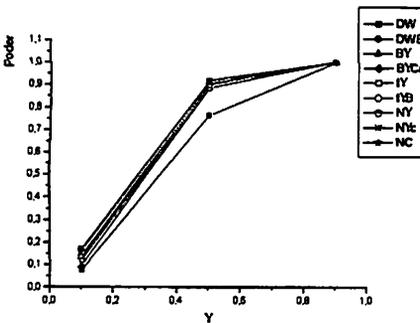
Na Figura 4.3.2 estão apresentados os poderes dos 9 testes para  $k = 5$  em função de  $\rho$  para diferentes valores de  $n$  (figura 4.3.2 a, b, c, e d). Como já foi salientado, verifica-se grande aumento do poder com o aumento do valor de  $\rho$  de 0,1 para 0,9. A comparação das 4 situações de  $n$  permite verificar grande aumento de poder. Deve ser observado que existem diferenças nas escalas das figuras. Os testes se diferenciam em maior magnitude quando  $n$  é pequeno ( $n = 10$ ) e  $\rho$  é grande ( $\rho = 0,9$ ). Se  $n$  for grande, os testes terão maior diferenciação se  $\rho$  for pequeno. Em todos os casos o teste DW original foi o que apresentou menor poder.



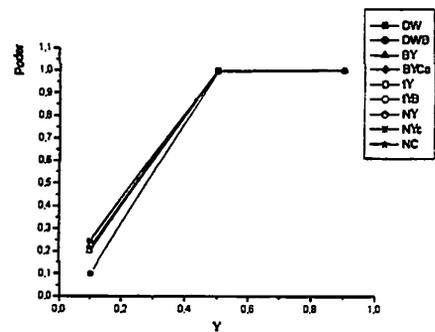
(a)



(b)



(c)



(d)

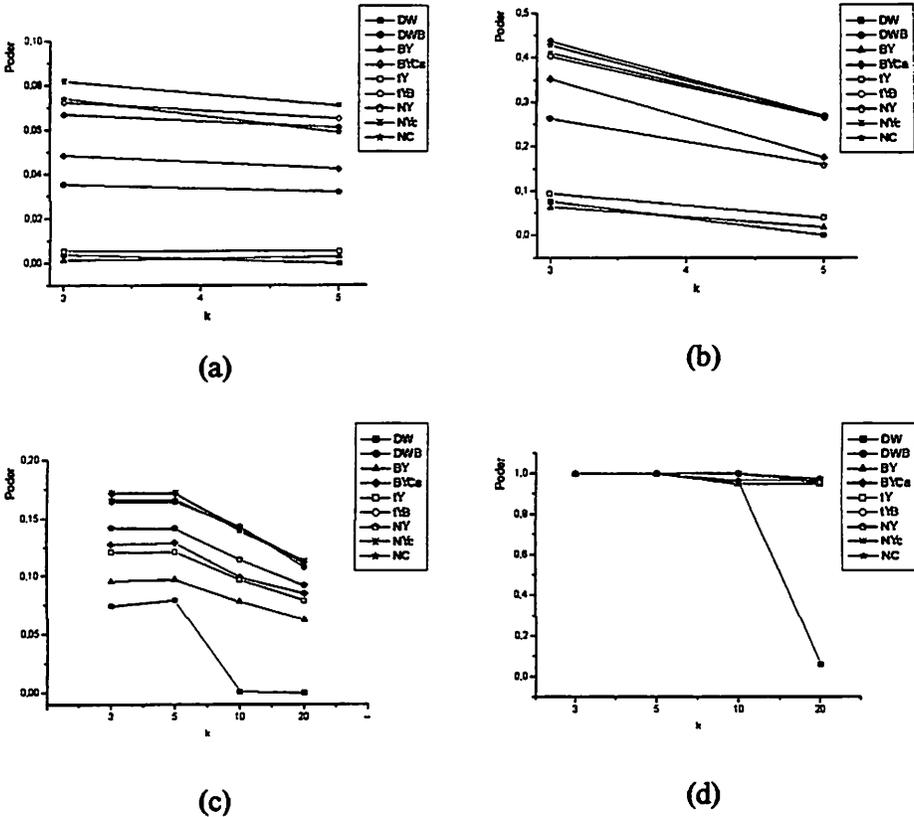
Figura 4.3.2 Poder dos testes para distribuição normal dos resíduos em função de  $\rho$  com  $k = 5$ , para  $\alpha = 0.05$  (a)  $n = 10$ , (b)  $n = 20$ , (c)  $n = 50$ , (d)  $n = 100$ .

Na Figura 4.3.3 o poder dos testes é apresentado em função de  $k$ , para algumas configurações de  $\rho$  e de  $n$  para distribuição normal. Nessa figuras pode-se observar com clareza a grande diferenciação entre os testes já destacados anteriormente.

Outro fato que chama a atenção e que ainda não foi destacado refere-se à grande diferença entre os testes  $B_p$  e  $B_{pCa}$ . A correção de viés provocou uma grande melhoria no teste, isso provavelmente devido ao viés dos estimadores

observados na Seção 4.1. O mesmo comentário pode ser estendido para  $N_p$  e  $N_{pc}$ .

O aumento de  $k$  provoca reduções no poder de todos os testes. Para amostras maiores ( $n \geq 50$ ) o teste DW, especificamente, sofre grandes perdas no poder (Figura 4.3.3 (c) e (d)). Dentre os melhores testes, destaca-se o teste  $N_{pc}$ .



**Figura 4.3.3** Poder dos testes para distribuição normal dos resíduos em função de  $k$  para  $\alpha = 0.05$  (a)  $n = 10, \rho = 0,1$  (b)  $n = 10, \rho = 0,9$  (c)  $n = 50, \rho = 0,1$  (d)  $n = 50, \rho = 0,9$

Para a distribuição uniforme, as mesmas situações apresentadas para a distribuição normal (Figuras 4.3.1, 4.3.2 e 4.3.3) foram obtidas e estão contempladas nas Figuras 4.3.4, 4.3.5 e 4.3.6. Devido ao fato de os resultados serem praticamente os mesmos, nenhum comentário adicional é feito. Apenas deve ser destacado que os testes, por esse motivo, são considerados robustos, uma vez que é violada a pressuposição de normalidade e os resultados são mantidos. O leitor pode conferir esses resultados nas Figuras 4.3.4, 4.3.5 e 4.3.6, comparando com as figuras do poder da distribuição normal equivalente.

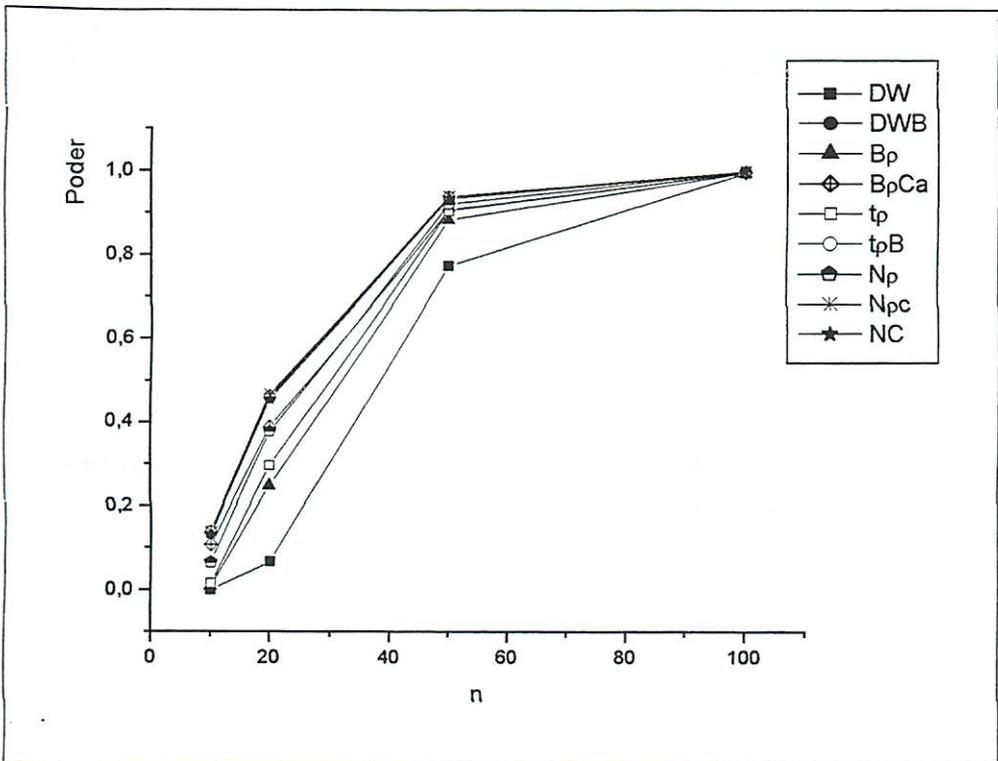
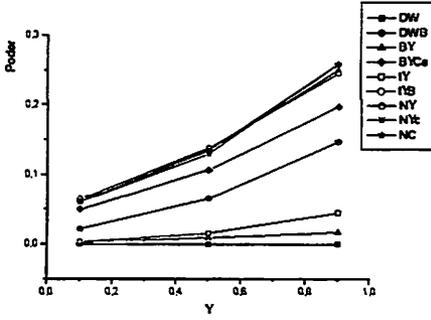
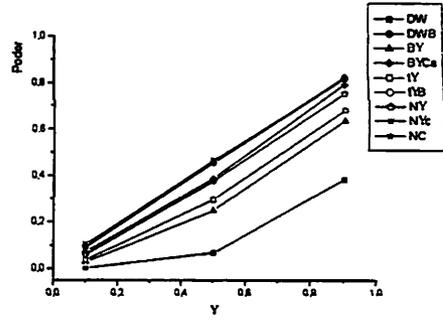


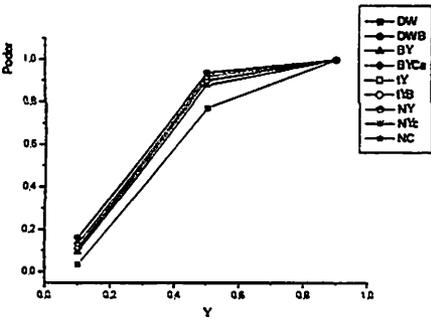
Figura 4.3.4 Poder dos testes para distribuição uniforme dos resíduos em função de n, com  $k = 5$  e  $\rho = 0$ , para  $\alpha = 0.05$



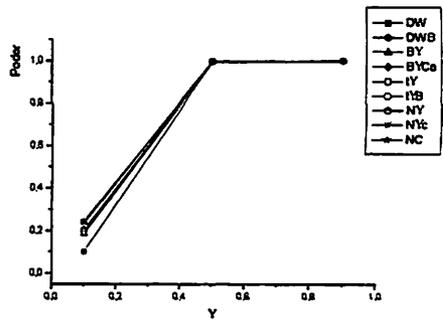
(a)



(b)

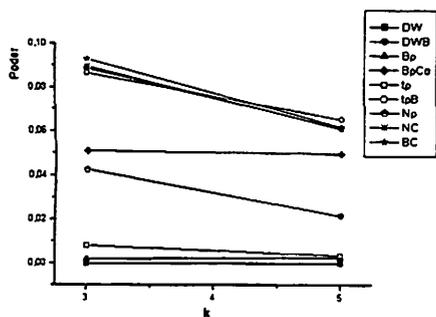


(c)

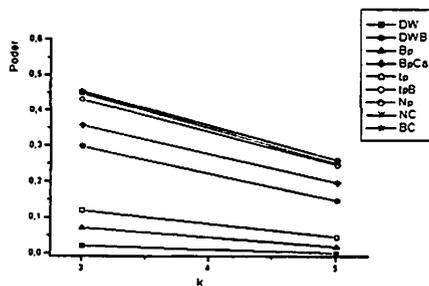


(d)

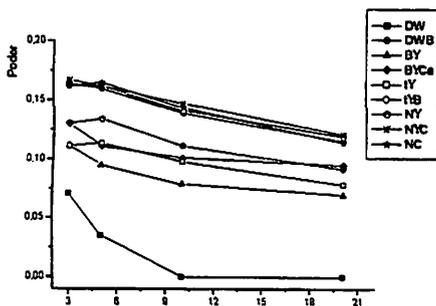
**Figura 4.3.5** Poder dos testes para distribuição uniforme dos resíduos em função de  $\rho$ , com  $k=5$  para  $\alpha=0.05$  (a)  $n=10$  (b)  $n=20$  (c)  $n=50$  (d)  $n=100$ .



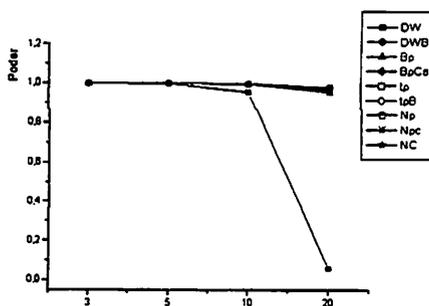
(a)



(b)



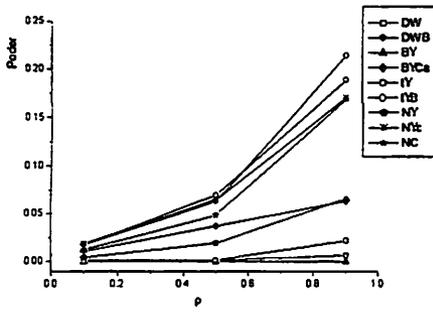
(c)



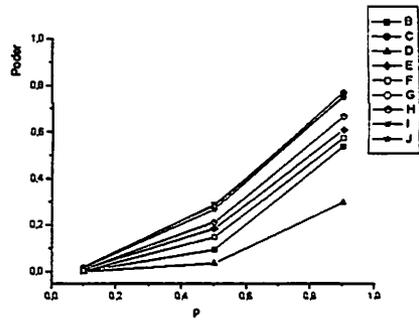
(d)

Figura 4.3.6 Poder dos testes para distribuição uniforme dos resíduos em função de  $k$  para  $\alpha = 0.05$  (a)  $n = 10, \rho = 0,1$  (b)  $n = 10, \rho = 0,9$  (c)  $n = 50, \rho = 0,1$  (d)  $n = 50, \rho = 0,9$

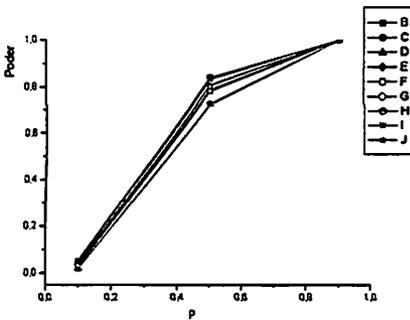
Nas Figuras 4.3.7 (a), (b), (c) e (d) estão apresentados os poderes dos testes para  $\alpha = 1\%$  e distribuição normal. Verificou-se o mesmo padrão de respostas encontrado para o valor nominal de 5%. A exceção novamente ocorre para  $n = 10$  com o teste NC, que apresentou menor poder do que os concorrentes tpB, DWB, BC e Npc. Esse resultado está de acordo com os resultados observados para o erro tipo I desse teste, o qual se apresentou rigoroso para pequenas amostras. Para valores de  $n$  maiores do que 10, essa diferença praticamente desapareceu.



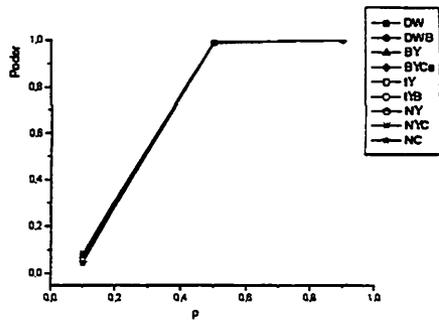
(a)



(b)



(c)



(d)

Figura 4.3.7 Poder dos testes para distribuição normal dos resíduos em função de  $\rho$  com  $k=3$  para  $\alpha=0.01$  (a)  $n=10$  (b)  $n=20$  (c)  $n=50$  (d)  $n=100$ .

#### 4.4 Considerações finais

No presente trabalho, vários testes para avaliar a presença de autocorrelação foram estudados e sua performance avaliada por meio de simulação Monte Carlo. Inicialmente, três estimadores ( $\hat{\rho}$ ,  $r_{1+}$  e C) de  $\rho$  foram comparados e observou-se que os três apresentaram viés. Os estimadores  $r_{1+}$  e C apresentaram menores vieses, principalmente com pequenas amostras, situação em que o viés é teoricamente maior. Esse fato é importante, pois pode influenciar os testes de hipóteses que se baseiam nesses estimadores. Observou-se também que as principais mudanças no EQM desses estimadores eram causadas pelo viés, permanecendo constante a variabilidade dos estimadores quando o valor paramétrico de  $\rho$  e o valor de  $k$  mudavam para um mesmo  $n$ . Esse resultado para  $k$  é importante, pois não foi comentado por De Carlo e Tryon (1993) e vem mostrar que o número de variáveis regressoras pode afetar o poder e as taxas de erro tipo I. Essa expectativa foi confirmada quando se estimou  $\alpha$  e  $\beta$  empiricamente por meio de simulação Monte Carlo.

A comparação das taxas de erro tipo I dos testes mostrou que 5 deles apresentaram tamanhos de testes iguais aos valores nominais de 5% e 1%. Esses testes foram o DWB, BC, tpB, Npc e NC. Dos cinco testes de tamanhos idênticos ao nominal, três foram testes de bootstrap. Verificou-se que DWB e BC foram equivalentes, pois além de apresentarem a mesma significância, também apresentaram o mesmo poder. A explicação obtida foi que C e d são combinações lineares um do outro. As taxas de erro tipo I para o teste BpCa em algumas configurações foram significativamente menores do que o valor nominal. Esses resultados são diferentes daqueles apresentados por Jeong e Chung (2001), os quais afirmaram que esse teste possui tamanho igual ao valor nominal. O teste original de DW mostrou taxas próximas de zero ( $\alpha = 5\%$  ou  $1\%$ ), o que o tornaria pouco poderoso. O baixo poder desse teste foi confirmado



pelos resultados das simulações efetuadas. Assim, os melhores testes envolveram de alguma forma estimadores corrigidos para viés de  $\rho$ . A exceção foi o teste  $tpB$ , que envolveu uma estatística pivô  $t$ , mas baseada em um estimador que não possui correção para viés. Manly (1998) aponta que esse tipo de teste de bootstrap tem apresentado melhores resultados do que seus competidores.

A performance desse testes também foi avaliada sob ambientes anormais para o teste (distribuição uniforme e exponencial para o resíduo). Os resultados obtidos foram os mesmos no ambiente favorável que o da distribuição normal. Isso mostra que os teste são robustos à violação da normalidade dos resíduos. Outras distribuições devem ser contempladas em futuros trabalhos para que esses resultados possam ser generalizados de uma forma mais ampla.

Finalmente, avaliou-se o poder desses testes e aqueles que tinham sido rigorosos apresentaram menores poderes. Os testes que apresentaram tamanho idênticos ao valor nominal (5% e 1%) foram mais poderosos do que os concorrentes. Os cinco testes (DWB, BC,  $tpB$ ,  $N_{pc}$  e NC) apresentaram valores muito similares entre si, sendo as diferenças atribuídas ao erro de Monte Carlo, mas superiores aos demais testes. A mudança de distribuição não afetou o poder e não interagiu com os testes.

O número de variáveis ( $k$ ) influenciou o poder e as taxas de erro tipo I. Em geral, o aumento de  $K$  provocou reduções no poder e reduziu as taxas de erro tipo I de alguns dos testes, exceto os cinco testes considerados superiores. A maior discordância obtida foi com o testes de  $BpCa$ , cujos resultados de Jeong e Chung (2001) não puderam ser reproduzidos nesse trabalho, embora os resultados desse teste tenham tido comportamento semelhante aos dos demais, enquanto os resultados desses autores destoaram dos resultados dos outros testes obtidos por eles mesmo e neste trabalho.

## 5 CONCLUSÕES

(a) O viés dos estimadores  $\hat{\rho}$ ,  $r_{1+}$  e  $C$  aumenta com o aumento de  $k$  e de  $\rho$ ;

(b) A variância dos estimadores  $\hat{\rho}$ ,  $r_{1+}$  e  $C$  não é afetada pelo aumento de  $k$  e de  $\rho$ ;

(c) Os testes  $DW$ ,  $N\rho$ ,  $tp$ ,  $B\rho$  e  $B\rho Ca$  são rigorosos e menos poderosos do que seus competidores;

(d) O teste  $DW$  é o mais rigoroso de todos e apresentou menor poder;

(e) Os testes  $DWB$  e  $BC$  são equivalentes;

(f) Os testes  $DWB$ ,  $BC$ ,  $tpB$ ,  $N\rho c$  e  $NC$  são considerados os melhores por apresentarem tamanho de teste idênticos ao valor nominal, maior poder e por serem robustos;

(g) o teste  $NC$  foi considerado rigoroso para pequenos valores de  $\alpha$  e de  $n$ ;

(h) Além das propriedades destacadas na conclusão (f), o teste  $N\rho c$  é recomendado nesse trabalho devido à simplicidade e facilidade para ser aplicado.

## REFERÊNCIAS BIBLIOGRÁFICAS

BEARZOTI, E. **Inferência estatística**. Lavras: FAEP, 1998.

\*BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica** 5. ed. São Paulo: Atual, 2003.

DACHS, J.; NORBERTO W. **Estatística computacional: uma introdução em Turbo Pascal**. Rio de Janeiro, 1988.

DE CARLO, L. T.; TRYON, W. W. Estimating and testing autocorrelation with small samples: A comparison of the C-statistic to a modified estimator. **Behaviour Therapist**, Oxford, v. 31, n. 8, p. 781-788, Nov. 1993.

DRAPER, R. N.; SMITHE, H. **Applied regression analysis**. 3. ed. New York, 1998.

DURBIN, J.; WATSON G. S. Testing for serial correlation in least squares regression I. **Biometrika**, London, v. 37, n. 3/4, p. 409-428, Dec. 1950.

DURBIN, J.; WATSON G. S. Testing for serial correlation in least squares regression I. **Biometrika**, London, v. 38, n. 1/2, p. 159-178, June 1951.

DURBIN, J.; WATSON G. S. Testing for serial correlation in least squares regression. **Biometrika**, London, v. 58, n. 1, p. 1-19, Apr. 1971.

EFRON, B.; TIBSHIRANI, R. **An introduction to the bootstrap**. New York: Chapman & Hall, 1993.

GUERRA, M. J.; DONAIRE, D. **Estatística indutiva: teoria e exercícios**. 2. ed. São Paulo : Livraria Ciência e Tecnologia, 1982. 311 p.

HANNAN, E. J.; TERRELL, R. D. Testing for serial correlation after least squares regression. **Econometrica**, Oxford, v. 36, n. 1, 133-150, 1968.

HALL, P. **The bootstrap and edgeworth expansions**. New Your, 1992

HOFFMANN, R.; VIERA, S. **Uma Introdução à Econometria**. 2. ed. São Paulo, 1987.

BOLFAFINE;

ANDERSON, 2003

SMITH (1947)

Johnson 1987

HUITEMA, B. E.; MCKEAN, J. W. Autocorrelation estimation and inference with small samples. *Psychological Bulletin*, Washington, v. 110, n. 2, p. 291-304, Sept. 1991.

JEONG, J.; CHUNG, S. Bootstrap tests for autocorrelation. *Computational statistics & data analysis*, 2001. p. 49-69.

Young, L. C. On randomness in ordered sequences. *Annals of mathematical Statistics*, 1941 v. 12, p.296-300

LEEMIS, L.; TRIVEDI, K. S. A comparison of approximate interval estimators for the Bernoulli parameter. *The American Statistician*, Alexandria, v. 50, n. 4, p. 388-389, Nov. 1996.

LEVINE, M. D.; BERENSON, L. M.; STEPHAN, D. *Estatística: teoria e aplicações*. Rio de Janeiro, RJ: LTC, 2000. 811 p.

MANLY, B.F.J. *Randomization, bootstrap and Monte Carlo methods in biology*. 2. ed. London: Chapman-Hall, 1998. 399 p.

MAYER, P. L. *Probabilidade: aplicações à estatística*. Tradução do prof. Ruy de C.B. Lourenço Filho. 2. ed. Rio de Janeiro: LTC, 1984.

MOOD, A. M.; GRAYBILL, F. A.; BOES D. C. *Introduction to the theory of statistics*. 3. ed. New York, 1974.

MORAN, P. A. P. Some theorems on time series II. The significance of the series correlation coefficient. *Biometrika*, London, v. 35, n. 3/4, p. 255-260, Dec. 1948.

✎ NAYLOR, T. H.; BALINTFY, J. L.; BURDICK, D. S.; CHU, K. *Técnicas de simulação em computadores*. Petrópolis: Vozes, 1971. 402 p.

NETER, J.; WASSERMAN, W.; KUTNER, H. M. *Applied linear statistical models*. 2. ed. Illinois: Richard D. Irwin, 1974.

TONHASCA, Jr. A. The three "capital sins" of statistics used in biology. *Ciência e Cultura*, São Paulo, v. 43, n. 6, p. 417-422, nov./dez. 1991.

TRIOLA, M. F. *Introdução a estatística*. 7. ed. Rio de Janeiro, 1999.

## **ANEXOS**

Anexo 1 - Programa utilizado para gerar os quantis dos testes  $t_p$ ,  $t_{pB}$ ,  $N_p$ ,  $N_{pC}$ ,  $N_C$  e  $B_C$ , e o viés e EQM dos estimadores de  $\rho$ .....71

Anexo 2 – Programa utilizado para gerar os quantis dos testes DW original e inconclusivo,  $B_p$  e  $B_{pCa}$ .....74

**ANEXO 1** Programa utilizado para gerar os quantis dos testes  $t_p$ ,  $t_pB$ ,  $N_p$ ,  $N_{pc}$  NC e BC, e o viés e EQM dos estimadores de  $\rho$

```

proc iml;
n=10;k=3;alpha=0.01;
pho=0.1;
b=j(k,1,1);
/*b0=1;b1=1;b2=1;b3=1;*/
signDW=0;InccDW=0;
dl=0.340;du=1.733;
signMC=0;signbMC=0;signtMC=0;
signZpho=0;signZr1=0;SignZc=0;
signZcb=0;
nsim=2000;nboot=2000;
Biaspho=0;biasr1=0;eqmpho=0;eqmr1=0;
biasc=0;eqmc=0;
do ij=1 to nsim;
  y=j(n,1,0);x=j(n,k,0);
  ui=0;uia=0;
  do i=1 to n;
    x[i,1]=1;y[i]=0;
    do j=1 to k;
      if j>1 then x[i,j]=ranuni(0);/*gerando os valores de Xi*/
      y[i]=y[i]+b[j]*x[i,j];
    end;
    ei=rannor(0); distribuição normal;
    ei=(ranuni(0)-0.5)/(1/12)**0.5; * distribuição uniforme;
    *ei=-1-(1-ranuni(0))/1; *distribuição exponencial com k=-1 e lamb=1 - Exp(mu=0 e
var=1);
    ui=pho*uia+ei;
    y[i]=y[i]+ui;
    uia=ui;
  end;
  *print x y;
  beta=ginv(t(x)*x)*t(x)*y;
  yhat=x*beta;
  uhat=y-yhat;
  yu=uhat[2:n];
  xu=uhat[1:n-1];
  phohat=inv(t(uhat)*uhat)*t(xu)*yu;

```

```

Biaspho=biaspho+(phohat-pho)/nsim;
biasr1=biasr1+(r1-pho)/nsim;
eqmpho=eqmpho+(phohat-pho)**2/nsim;
eqmr1=eqmr1+(r1-pho)**2/nsim;
ehat=yu-xu*phohat;
ehat=uhat[1]/ehat;
d=sum((yu-xu)#(yu-xu))/(t(uhat)*uhat);
  if d<dl then SignDW=SignDW+1/nsim;
  if d>=dl & d<=du then InccDW=InccDW+1/nsim;
*print phohat d ij;
*print beta;
c=1-0.5*d;
  biasc=biasc+(c-pho)/nsim;
  eqmc=eqmc+(c-pho)**2/nsim;
  zcphohat=phohat/((n-2)/(n*(n-1)**0.5));
  zcr1=r1/((n-2)/(n*(n-1)**0.5));
zcc=c/((n-2)/(n**2-1))**0.5;
  probzphohat=1-probnorm(zcphohat);
  probzr1=1-probnorm(zcr1);
  probzc=1-probnorm(zcc);
  if probzphohat<=alpha then signZpho=signZpho+1/nsim;
  if probzr1<=alpha then signZr1=signZr1+1/nsim;
  if probzc<=alpha then signZc=signZc+1/nsim;
/* t de Monte Carlo*/
tc=phohat*(n-3)**0.5/(1-phohat**2)**0.5;
prtc=1-probt(tc,n-3);
  if prtc<=alpha then signtMC=signtMC+1/nsim;
  *print tc prtc;
/*bootstrap*/
ehatb=j(n,1,0);
uib=j(n,1,0);
yfake=j(n,1,0);
db=j(nboot,1,0);
sigdb=0;sigtcb=0;
  sigzcb=0;
do ii=1 to nboot;
do j=1 to n;
do until (iii<=n);
iii=int(ranuni(0)*n)+1;
end;
  ehatb[j]=ehat[iii];
  if j=1 then uia=0;
  else uia=uib[j-1];
*uib[j]=phohat*uia+ehatb[j];*verificar se era isso que eu queria fazer;
/*impondo artificialmente a h0: pho=0*/

```

```

u[bj]=0.*ua+ehat[bj];
yake[j]=beta[1]+u[bj];
do i=2 to k;
yake[j]=yake[j]+beta[i]*x[i];
end;

*print phat ehat ehat uhat;
betab=ginv(t(x)*t(x))*yake;
yhatb=x*betab;
uhatb=yake-yhatb;
yu=uhatb[2:n];
xu=uhatb[1:n-1];
phatb=invt(uhatb)*uhatb*(xu)*yu;
db[ii]=sum(yu-xu)/(t(uhatb)*uhatb);
if d >= db[ii] then sigdb=sigdb+1/nboot;
cb=1-0.5*db[ii];
zccb=cb/((n-2)/(n**2-1))*0.5;
tcb=phohat*(n-3)**0.5/(1-phohat**2)**0.5;
if tcb >= tc then sigtc=sigtc+1/nboot;
if zccb >= zcc then sigzcb=sigzcb+1/nboot;
end;

if sigdb <= alpha then signMC=signMC+1/nsm;
*print sigdb;
*print db;
*print sigcb;
if sigcb <= alpha then signbMC=signbMC+1/nsm;
if sigzcb <= alpha then sigzcb=sigzcb+1/nsm;
end;

print "veja com cuidado qual foi a distribuicao do erro gerada";
print "parâmetros da simulação" n alpha pho k nsm nboot;
print "Significancias dos diferentes procedimentos";
print "teste de bootstrap para d" signMC;
print "teste de DW original" sigDW InccDW;
print "teste de t MC para (auto)correlacao:" signbMC;
print "teste de t bootstrap para (auto)correlacao:" signbMC;
print "teste normal MC para (auto)correlacao:" signzpho;
print "teste normal MC para r1+:" signZr1;
print "teste normal MC para C=1-0.5*d:" signZc;
print "teste normal Bootstrap para C=1-0.5*d:" signZcb;
print "Vies e EQM de pho^r1+ e c";
print "pho^Vies:" biaspho "EQM" eqmpho;

```

```

xu=uhat[1:n-1];
phohat=inv(t(uhat)*uhat)*(xu)*yu;
ehat=yu-xu*phohat;
ehat=uhat[1]//ehat;
d=sum((yu-xu)#(yu-xu))/(t(uhat)*uhat);
if d<di then SigndW=SigndW+1/nsim;
if d>di & d<=du then InccdW=InccdW+1/nsim;
*print phohat d ij;
/*bootstrap*/
ehat=j(n,1,0);
wb=j(n,1,0);
yfake=j(n,1,0);
*db=j(nboot,1,0);
sigdb=0;
pp=0;
*print phohat;
do ii=1 to nboot;
do j=1 to n;
do until (iii<=n);
iii=int(randu(0)*n)+1;
end;
ehat[iii]=ehat[iii];
*if j=1 then wia=ehat[j]/(1-phohat**2)**0.5;
*if j=1 then wia=ehat[j];
else wia=phohat*wb[j-1]+ehat[j];
wb[j]=wia;
end;
*print phohat ehat wb uhat;
uhat=wb;
yu=uhat[2:n];
xu=uhat[1:n-1];
phohat=inv(t(uhat)*uhat)*(xu)*yu;
phohat[iii]=phohat;
if phohat>phohat then pp=pp+1/nboot;
*db[iii]=sum((yu-xu)#(yu-xu))/(t(uhat)*uhat);
if 0 <= phohat then sigdb=sigdb+1/nboot;
*print phohat sigdb;
end;
z0=probit(1-pp);
za=probit(1-alpha);
ax=j(n-1,k,0):ay=j(n-1,1,0);
theta_1=j(n,1,0);theta_m=0;
do ii=1 to n;
do jf=1 to ii-1;
ax[jf,]=X[jf,];

```

```

print "pho^: Vies: " biaspho " EQM " eqmpho;
print "r1+: Vies: " biasr1 " EQM " eqmrl;
print "c: Vies: " biasc " EQM " eqmc;
*xy=x||y;
*create saida var {x0 x1 x2 x3 y};
*append from xy;
quit;

```

## ANEXO 2 Programa utilizado para gerar os quantis dos testes DW original e inconclusivo, Bp e BpC

```

/* accelerated bias-corrected percentiles limited*/
proc iml;
n=100;k=20;alpha=0.05;
pho=0.7;
b=j(k,1,1);
/*b0=1;b1=1;b2=1;b3=1;*/
signDW=0;lnccDW=0;
dl=1.229;du=2.164;
signMCbc=0;signMC=0;
signtMC=0;signtbMC=0;
nsim=2000;nboot=2000;
phobot=j(nboot,1,0);
do ij=1 to nsim;
  y=j(n,1,0);x=j(n,k,1);
  ui=0;uia=0;
  do i=1 to n;
    x[i,1]=1;y[i]=0;
    do j=1 to k;
      if j>1 then x[i,j]=ranuni(0);/*gerando os valores de Xi*/
      y[i]=y[i]+b[j]*x[i,j];
    . end;
    *ei=rannor(0);*distribuição normal;
    ei=(ranuni(0)-0.5)/(1/12)**0.5;* distribuição uniforme;
    *ei=-1-(1-ranuni(0))/1;*distribuição exponencial com k=-1 e lamb=1 - Exp(mu=0 e
var=1);
    ui=pho*uia+ei;
    y[i]=y[i]+ui;
    uia=ui;
  end;
  beta=ginv(t(x)*x)*t(x)*y;

```

```

    ay[ji]=y[ji];
end;
do ji=ii+1 to n;
    ax[ji-1,]=X[ji,];
    ay[ji-1]=y[ji];
end;
abeta=ginv(t(ax)*ax)*t(ax)*ay;
ayhat=ax*abeta;
auhat=ay-ayhat;
ayu=auhat[2:n-1];
axu=auhat[1:n-2];
theta_i[ii]=inv(t(auhat)*auhat)*t(axu)*ayu;
    thetam=thetam+theta_i[ii]/n;
end;
*print thetam theta_i phohat;
num=0;den=0;
do ii=1 to n;
    num=num+(thetam-theta_i[ii])**3;
    den=den+(thetam-theta_i[ii])**2;
end;
aa=num/(6*den**1.5);

*print aa;
*print phobot;
b_o=phobot;
phobot[rank(phobot),]=b_o;

*zu=(z0+za)/(1-aa*(z0+za))+z0;
*zl=(z0-za)/(1-aa*(z0-za))+z0;
*alphau=probnorm(zu);
alphal=probnorm(zl);
*ii=round(alphau*nboot);
*phou=phobot[ii];
ii=round(alphal*nboot);
if ii<=0 then ii=1;
phol=phobot[ii];
*if 0.0>phou then signMCbc=signMCbc+1/nsim;
if 0.0<phol then signMCbc=signMCbc+1/nsim;
*print z0 za zu alphau ii phou;
*print phobot;
if sigdb<=alpha then signMC=signMC+1/nsim;
*print sigdb;
*print db;
end;

print "veja com cuidado qual foi a distribuicao do erro gerada";
print "parâmetros da simulação" n alpha pho k nsim nboot;

```

```
print "teste de bootstrap para rho" signMC;  
print "teste de bootstrap para rho com correcao de vies" signMCbc;  
print "teste de DW original" signDW InccDW;  
*xy=x||y;  
*create saida var {ind x2 x3 y};  
*append from xy;  
quit;
```

2007-1828

SENWA: 252525