

ALEXANDRA MARIA RIOS  
CABRAL\*  
ISIS DIDIER LINS\*\*

aplicação de bootstrap para teste de  
durbin-watson - uma contribuição para  
cenários na economia

\* Professora da  
Universidade Federal  
de Alagoas e aluna  
de doutorado do Pro-  
grama de Pós-Gradua-  
ção em Engenharia  
de Produção da  
Universidade Federal  
de Pernambuco.

\*\* Aluna de douto-  
rado do Programa  
de Pós-Graduação  
em Engenharia  
de Produção da  
Universidade Federal  
de Pernambuco.

## RESUMO

Em economia, trabalhar com modelos de regressão linear clássicos com o intuito de investigar comportamentos sejam macroeconômicos ou microeconômicos é algo bastante comum de se encontrar. Entretanto, dada as várias as suposições que devem ser atendidas ao se utilizar este método, e dada à natureza de algumas variáveis econômicas, é comum se violar uma ou mais hipóteses. Para diagnosticar a presença de erros autocorrelacionados, o teste de Durbin-Watson é o mais usado, porém apresenta a limitação de não fornecer valores críticos exatos para a distribuição de probabilidade de sua estatística de teste, tornando-se inconclusivo em determinados casos. Com o intuito de fornecer uma solução para esta limitação, propõem-se dois algoritmos bootstrap. Um deles pode ser aplicado em situações onde se pode afirmar que os erros “bem comportados” seguem uma distribuição normal com média zero e variância (bootstrap paramétrico). O outro, por sua vez, não requer qualquer suposição sobre a sua distribuição (bootstrap não-paramétrico). Os algoritmos propostos são aplicados a dois exemplos práticos e são obtidos, para cada um deles, os valores críticos exatos da distribuição da estatística de Durbin-Watson para os níveis de significância de 1%, 5% e 10%, bem como o formato resultante da densidade de probabilidade empírica.

## PALAVRAS-CHAVE

teste de Durbin-Watson, bootstrap paramétrico, bootstrap não-paramétrico.

## ABSTRACT

In economics, working with classical linear regression models in order to investigate macroeconomic or microeconomic behaviors is very common. However, given the various assumptions that should be satisfied when using this method, and given the nature of some economic variables, it is usual to violate one or more assumptions. To diagnose the presence of autocorrelated errors, the Durbin-Watson is the most used. However, it shows the limitation of not providing exact critical values for the probability distribution of its test statistic, becoming, in certain cases, inconclusive. In order to provide a solution to this limitation, two bootstrap schemes are proposed. One of them can be used when the “well behaved” errors follow a normal distribution with zero mean and variance (parametric bootstrap). The other one does not require any suppositions about their probability distribution (non-parametric bootstrap). The proposed algorithms are applied to two examples. For both of them, the exact critical values of the Durbin-Watson distribution for the significance levels 1%, 5% and 10% are obtained as well as the shape of the resulting empirical probability density.

## KEY WORDS

Durbin-Watson test, parametric bootstrap, non-parametric bootstrap.

## INTRODUÇÃO

Em economia, trabalhar com modelos de regressão linear clássicos com o intuito de investigar comportamentos sejam macroeconômicos ou microeconômicos é algo bastante comum de se encontrar. Entretanto, são várias as suposições que devem ser atendidas antes de se escolher um modelo econométrico que seja eficiente tanto sob a ótica econômica quanto sob os padrões estatísticos necessários ao se usar, por exemplo, o método dos mínimos quadrados ordinários.

Dentre as várias hipóteses a serem atendidas, tem-se uma que é a de que os erros gerados por uma regressão deveriam ser não autocorrelacionados. Quando isso não acontece, torna-se inviável manter o modelo econométrico inicialmente adotado já que os estimadores passam a ser não eficientes, o que obrigaria o pesquisador a realizar alguns ajustes para obter resultados estatisticamente confiáveis.

Existem várias causas que podem gerar erros autocorrelacionados, na sua grande maioria encontrados apenas em séries temporais. Uma delas seria a influência prolongada de choques na economia onde os seus efeitos se prolongariam por mais de um período de tempo. Greves, catástrofes ambientais ou até mesmo guerras, provavelmente afetariam a economia de uma localidade em períodos seguintes o que contribuiria para o aparecimento de erros autocorrelacionados.

O fator inercial é outra importante causa. Comportamentos anteriores podem continuar a exercer grandes influências na economia podendo-se citar a inflação brasileira antes da implantação do plano de estabilização econômica do Real.

Ou ainda, choques aleatórios que afetem a atividade econômica de uma região que podem fazer que com que a atividade econômica de uma ou mais localidades próximas sofram impactos devidos aos vínculos econômicos entre elas. Este raciocínio pode ser estendido a pólos industriais que possuem um grande efeito multiplicador e terminam por abranger mais de um município. O pólo fabril de Pernambuco pode ser citado como exemplo.

A não inclusão de variáveis importantes no modelo, ou uma má especificação do modelo a ser trabalhado são outras fortes causas do aparecimento de erros autocorrelacionados. De forma alguma poderia

se esquecer de incluir a taxa de desemprego num modelo que tentasse explicar a oferta de mão-de-obra de uma região; ou tentar explicar como os investimentos se comportaram ao longo de um período de tempo sem considerar o efeito do PIB e da taxa de juros na mesma época. Evitar a todo custo interpolações ou suavizações nos dados também é muito importante.

Nesste contexto, o teste mais famoso para identificar a presença de erros autocorrelacionados, segundo Gujarati (2006), é o de Durbin-Watson. Este teste é amplamente utilizado em modelos econômicos mas nem sempre fornece uma resposta por possuir uma região chamada de “inconclusiva” onde não se pode afirmar se existe ou não uma autocorrelação serial.

Assim sendo, este trabalho tem como objetivo propor uma solução para essa limitação, o que acontece quando o valor calculado para a estatística de teste pertence ao intervalo formado pelos valores limites inferiores e superiores desta distribuição (zona de inconclusão).

Para tanto, propõem-se dois algoritmos *bootstrap*. Um deles pode ser aplicado em situações onde se pode afirmar que os erros “bem comportados” ( $\varepsilon_t = e_t - \rho e_{t-1}$ ) seguem de fato uma distribuição de probabilidade normal com média zero e variância  $\sigma^2$  (*bootstrap* paramétrico). O outro, por sua vez, não requer hipótese sobre a distribuição de probabilidade de  $\varepsilon_t$  (*bootstrap* não-paramétrico).

O uso de algoritmos *bootstrap* vem se tornando uma importante ferramenta bastante utilizada por pesquisadores pois é uma alternativa eficiente que em muitas vezes leva a resultados muito mais precisos do que as abordagens tradicionais são capazes de fazê-lo. Isto acontece porque em muitos casos, a inferência sobre as distribuições sugeridas pela teoria assintótica só valem para amostras de tamanho infinitamente grande (o que nem sempre é o que se tem à disposição). Com a técnica *bootstrap*, as estimativas de parâmetros e estatísticas de teste são calculados de forma bastante convencional, mas os valores dos níveis de significância, valor-*p* e intervalos de confiança são calculados com base em distribuições obtidas por simulação, e não mais baseados na teoria assintótica, o que poderia ser de pouca credibilidade para pequenas amostras.

Mackinnon (2002) cita como um exemplo disto o teste J, comumente utilizado em modelos de regressão não aninhados, que rejeita a hipótese nula com demasiada frequência. Em casos extremos, até mesmo para amostras de tamanhos tão grandes quanto 50 observações, um teste

assintótico J ao nível de 0,05 poderia rejeitar uma hipótese nula verdadeira mais de 80% das vezes. Sendo assim, uma das razões para a utilização de testes *bootstrap* ao invés de testes assintóticos é que se espera cometer pequenos erros na probabilidade de rejeição de uma estatística ao fazê-lo.

Este texto está dividido em 5 seções incluindo esta introdução. A seção 2 contém uma descrição sobre o teste de Durbin-Watson; a seção 3 descreve os passos utilizados para a montagem dos algoritmos a que se propõe, bem como as rotinas preparadas na plataforma computacional do *software* R (R Development Core Team, 2010); a seção 4 traz dois exemplos econômicos em que os algoritmos propostos foram utilizados. Por fim, a seção 5 fornece alguns comentários finais.

## TESTE DE DURBIN-WATSON

Para a introdução do teste de Durbin-Watson, considere o seguinte modelo de regressão linear:

$$y_t = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + e_t, \quad t = 1, 2, \dots, T \quad (1)$$

em que  $T$  é o número de observações,  $K$  é o número de parâmetros no modelo e  $e_t$  são os erros gerados segundo um modelo autorregressivo de primeira ordem (AR(1)):

$$(\varepsilon_t = e_t - \rho e_{t-1}) \quad (2)$$

em que  $\rho$  é a correlação entre erros vizinhos no tempo ( $-1 < \rho < 1$ ) e  $e_t$  são os erros “bem comportados”, ou seja, tratam-se de erros aleatórios, independentes, com distribuição normal com média 0 e variância  $\sigma^2$ . A suposição de normalidade é importante para a dedução das distribuições limites inferior e superior da estatística de teste de Durbin-Watson (HILL, 2003).

Enfatiza-se a importância do processo AR(1) para geração dos termos  $e_t$ . Se outro tipo de processo autorregressivo, que não o de primeira ordem, gera os erros  $e_t$ , não se pode aplicar o teste em questão.

Além disso, o modelo de regressão não pode apresentar como Regressores valores defasados da variável resposta, uma vez que é necessária a hipótese de variáveis explicativas não-estocásticas para a realização do teste

de Durbin-Watson. Quando existem Regressores correspondentes a valores defasados da variável resposta ou para qualquer caso em que o erro esteja correlacionado a um Regressão  $r$  o teste de Durbin-Watson não é confiável. Como alternativa pode-se utilizar o teste  $h$  de Durbin (KENNEDY, 2009).

Outro pressuposto importante é a presença de observações para todos os períodos considerados. Por exemplo, se os dados são bimestrais para um período de 3 anos, a base de dados deve apresentar um total de 18 observações, uma para cada bimestre. Caso haja a omissão de algum dado, o teste não é capaz de detectar autocorrelação serial de primeira ordem de maneira confiável.

As hipóteses envolvidas no teste de Durbin-Watson são:

$$H_0 : \rho = 0 \quad \text{vs.} \quad H_1 : \rho > 0 \quad (3)$$

quando a hipótese alternativa refere-se à autocorrelação positiva e

$$H_0 : \rho = 0 \quad \text{vs.} \quad H_1 : \rho < 0 \quad (4)$$

quando se quer avaliar autorrelações negativas. Em ambos os casos, a hipótese nula está relacionada à ausência de autocorrelação serial dos erros.

A estatística de teste de Durbin-Watson é:

$$d = \frac{\sum_{t=2}^T (\hat{e}_t - \hat{e}_{t-1})^2}{\sum_{t=1}^T \hat{e}_t^2} \quad (5)$$

ou em notação matricial

$$d = \frac{\hat{e}' A \hat{e}}{\hat{e} \hat{e}} \quad (6)$$

em que

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix} \quad (7)$$

A distribuição de  $d$  para amostras finitas depende dos Regressores. Dessa forma, é comum utilizar tabelas que fornecem apenas *intervalos* para os valores críticos determinados pelas distribuições limites inferior e superior  $d_L$  e  $d_U$ , respectivamente. Quando o valor calculado de  $d$  pertencer aos intervalos  $[d_{L_c}, d_{U_c}]$  ou  $[4 - d_{U_c}, 4 - d_{L_c}]$ , torna-se impossível decidir sobre a rejeição ou não da hipótese nula de não-autocorrelação dos erros (zona de indecisão).

A autocorrelação amostral  $\hat{\rho}$  é definida por:

$$\hat{\rho} = \frac{\sum_{t=2}^T \hat{e}_t \hat{e}_{t-1}}{\sum_{t=1}^T \hat{e}_t^2} \quad (8)$$

Como  $\sum_{t=1}^T \hat{e}_t^2$  e  $\sum_{t=2}^T \hat{e}_{t-1}^2$  diferem em apenas uma observação, os somatórios são aproximadamente iguais (GUJARATI, 2006). Portanto, pode-se reescrever a estatística  $d$  (Equação (5)) em função de  $\hat{\rho}$  (Equação (8)):

$$d \approx 2(1 - \hat{\rho}) \quad (9)$$

Da Equação (9), se  $\hat{\rho} = -1$  (autocorrelação negativa perfeita), então  $d = 4$ ; se  $\hat{\rho} = 1$  (autocorrelação positiva perfeita), então  $d = 0$ ; se  $\hat{\rho} = 0$  (ausência de autocorrelação), então  $d = 2$ . Dessa maneira, as decisões a seguir são possíveis para os testes de autocorrelação positiva e negativa, respectivamente:

$H_0 : \rho = 0$  vs.  $H_1 : \rho > 0$ ; calcular  $d$  e

$$\left\{ \begin{array}{l} \text{rejeitar } H_0, \text{ se } d < d_{L_c}; \\ \text{n\~{a}o rejeitar } H_0, \text{ se } d > d_{U_c}; \\ \text{inconclusivo, se } d_{L_c} < d < d_{U_c}. \end{array} \right\} \quad (10)$$

$H_0 : \rho = 0$  vs.  $H_1 : \rho < 0$ ; calcular  $d$  e

$$\left\{ \begin{array}{l} \text{rejeitar } H_0, \text{ se } d > 4 - d_{L_c}; \\ \text{n\~{a}o rejeitar } H_0, \text{ se } d < 4 - d_{U_c}; \\ \text{inconclusivo, se } 4 - d_{U_c} < d < 4 - d_{L_c}. \end{array} \right\} \quad (11)$$

#### BOOTSTRAP PARA TESTE DE DURBIN-WATSON

Foram propostos dois algoritmos *bootstrap* para realizao do teste de Durbin-Watson, um parametrico e outro no-parametrico. A descrio de cada uma das versoes bem como as respectivas rotinas implementadas no R so apresentadas nas proximas duas subseoes.

#### ALGORITMO BOOTSTRAP PARAMETRICO PARA TESTE DE DURBIN-WATSON

Na verso parametrica do algoritmo *bootstrap*, e necessaria a suposio de normalidade dos erros "bem comportados"  $\epsilon_t$ . O modelo estimado com os dados originais fornece valores ajustados para a variavel resposta, a variancia estimada dos erros  $\hat{\sigma}^2$  e a estatistica de Durbin-Watson  $d_0$ . Ento, a partir dos valores ajustados e de erros gerados segundo uma distribuio normal com media 0 e variancia  $\hat{\sigma}^2$ , constroem-se  $B$  vetores de valores simulados para  $y$ , em que  $B$  e o numero de iteraoes *bootstrap*. Em seguida, realizam-se  $B$  regressoes dos  $y$ 's simulados sobre as variaveis independentes e, para cada uma delas, obtem-se o valor da estatistica de



Durbin-Watson  $d_i$ ,  $i=1, 2, \dots, B$ . Assim,  $d_0, d_1, d_2, \dots, d_B$  permitem construir uma densidade de probabilidade empírica (sob  $H_0$ ) da estatística de Durbin-Watson para o caso analisado. Portanto, torna-se fácil a obtenção de valores críticos exatos, podendo-se proceder à rejeição ou não da hipótese nula. É importante enfatizar que, ao se adotar um algoritmo *bootstrap* para o teste de Durbin-Watson, não há o risco de inconclusão. A rotina do algoritmo *bootstrap* paramétrico implementado no R é mostrada a seguir.

```
#####
#### PROGRAMA: Algoritmo bootstrap paramétrico para teste de Durbin-Watson ####
#### para autocorrelação de primeira ordem dos erros #####
#### AUTORAS: Alexandra Maria Rios Cabral e Isis Didier Lins #####
###
## DATA: 21/12/2010 #####
#####
```

Suposições: (1) Regressores não-estocásticos.  
 # (2) Termos de erro (e(t)) gerados unicamente por processo AR(1).  
 # (3) Erros “bem comportados” (epsilon(t)) normalmente distribuídos  
 # com média 0 e variância constante.  
 # (4) Conjunto de dados completo, não há períodos sem observação.  
 # (5) Estimação da distribuição de probabilidade empírica da  
 # estatística de teste de Durbin-Watson sob  $H_0$  ( $\rho = 0$ ).

```
# Bibliotecas
require(lmtest)
require(stats)
```

```
# Dados
dados = read.table("c:\dados.txt", header = TRUE) # Leitura de arquivo
attach(dados)
T = length(dados[,1]) # Número de observações
J = length(dados[1,]) # Número de colunas de dados
x = as.matrix(cbind(dados[,2:J])) # Matriz de Regressores
```

```

X = as.matrix(cbind(1,x)) # Matriz X com coluna referente ao intercepto
y = as.matrix(dados[,1]) # Variável resposta
T = length(x[,1]) # Número de observações

# Regressão linear
ajuste = lm(y~x) # Estimação do modelo por MQO
b = coef(ajuste) # Estimativas dos parâmetros do ajuste por MQO
K = length(b) # Número de parâmetros do modelo
e = as.matrix(residuals(ajuste)) # Resíduos do ajuste

# Gráfico de dispersão dos resíduos
plot(e, col = "red", pch = 21, bg = "red", xlab = "Observação", ylab
= "Resíduo")
abline(0,0)
lines(e, lty = 2, lwd = 1.5, col = "green")

# Gráficos ACF
par(mfrow = c(1,2), pty = "s")
acf(e, main = "Função de autocorrelação")
acf(diff(e), main = "Função de autocorrelação (primeira diferença)")

# Montagem da matriz A
A = matrix(0, nrow = T, ncol = T)
A[1,1] = 1;
A[1,2] = -1;
A[T,T-1] = -1;
A[T,T] = 1;
for (t in 3:T-1)
{ A[t,t-1] = -1
  A[t,t] = 2
  A[t,t+1] = -1 }

B = 9999 # Número de simulações bootstrap
d0 = (t(e) %*% A %*% e)/(t(e) %*% e) # Estatística de Durbin-Watson
para ajuste
D = rbind(d0) # Vetor D que conterà B+1 estatísticas de Durbin-Watson

```

```

# Bootstrap paramétrico
# Gerar erros segundo uma distribuição normal com média = 0 e
variância = sigma2
# para o cálculo do y simulado; sob H0, rho = 0, logo e(t) = epsilon(t)
y.ajustado = X %*% b # Valores ajustados
sigma2 = t(e) %*% e / (T - K) # Estimativa da variância dos erros
y.simulado = y.ajustado[,1] + matrix(rnorm(T * B, mean = 0, sd =
sqrt(sigma2)), T, B) # Valores simulados
ajustes = lm(y.simulado~x) # Estimativas de B modelos de Regressão

# Matriz de resíduos (T x B) dos B modelos estimados
residuos = as.matrix(residuals(ajustes))
for (j in 1:B)
{ # Estatística de Durbin-Watson referente ao j-ésimo modelo estimado
ds = (t(residuos[,j]) %*% A %*% residuos[,j]) /
(t(residuos[,j]) %*% residuos[,j])
D = rbind(D,ds) }

# Saída
# Estatística de Durbin-Watson para modelo ajustado com dados
originais e
# valores críticos da distribuição de probabilidade empírica sob
H0 (rho = 0)
d0
quantile(D, c(0.01, 0.05, 0.1, 0.9, 0.95, 0.99))
# Visualização do formato da densidade de probabilidade
hist(D, col = "green", xlim=c(0,4), freq = FALSE, main = "", xlab = "d",
ylab = "Densidade de probabilidade empírica")
lines(density(D), col = "red", lwd = 2)

```

#### ALGORITMO BOOTSTRAP NÃO-PARAMÉTRICO PARA TESTE DE DURBIN-WATSON

O procedimento não-paramétrico tem várias etapas idênticas à versão paramétrica. A diferença consiste na etapa de geração dos erros para obtenção dos valores simulados da variável resposta, uma vez que não se

faz a hipótese de normalidade dos erros “bem comportados” (a suposição (3) na rotina do algoritmo paramétrico não é mais necessária).

Em vez de gerar erros a partir de uma distribuição específica, é utilizada a reamostragem com reposição dos resíduos. e as amostras resultantes são escalonadas com conforme a expressão a seguir:

$$\tilde{e}_t = \left( \frac{T}{T-K} \right)^{1/2} \hat{e}_t \quad (12)$$

Segundo Mackinnon (2002), os resíduos obtidos pelo modelo de regressão linear utilizando o método dos mínimos quadrados (MQO) possuem uma variância menor do que os erros gerados por reamostragens. Assim o autor propõe o escalonamento dos termos dos erros gerados pela Equação (12) antes de realizar as  $B$  regressões do algoritmo não-paramétrico.

Os erros reamostrados e escalonados em conjunto com os  $y$ 's ajustados a partir do modelo estimado com os dados originais fornecem os valores simulados de  $y$  e o algoritmo prossegue como no caso paramétrico.

Dessa forma, o algoritmo *bootstrap* não-paramétrico para o teste de Durbin-Watson é semelhante ao anterior bastando substituir o trecho referente a “# Bootstrap paramétrico” da rotina anterior por:

```
# Bootstrap não-paramétrico
# Gerar erros a partir de reamostragem com reposição dos resíduos
do modelo
# estimado com dados originais; sob H0, rho = 0, logo e(t) = epsilon(t)
es = matrix(0, nrow = T, ncol = B)
pos = matrix(round(runif(T * B, min=0.51, max=T)), T, B) # Sorteio
das amostras
for (i in 1:B) { es[,i] = e[pos[,i],1] }
es = sqrt(T / (T-K)) * es # Escalonamento dos resíduos reamostrados
y.ajustado = X %*% b # Valores ajustados
y.simulado = y.ajustado[,1] + es # Valores simulados
ajustes = lm(y.simulado~x) # Estimativas de B modelos de Regressão
```

## EXEMPLOS DE APLICAÇÃO

Para ilustrar a aplicação do teste de Durbin-Watson, foram considerados os seguintes exemplos práticos apresentados por Hill, Griffiths e Judge (2003):

- Função investimento;
- Relação entre oferta de emprego e taxa de desemprego.

Como não se conhece a distribuição populacional dos erros “bem comportados” *a priori* e as amostras têm poucas observações, o algoritmo *bootstrap* não-paramétrico é o mais indicado. De qualquer forma, os quantis da densidade empírica obtidos a partir do algoritmo paramétrico para os mesmos dados foram bem semelhantes. Assim, optou-se por apresentar apenas os resultados relativos à abordagem não-paramétrica.

## FUNÇÃO INVESTIMENTO

Da teoria econômica sabe-se que o investimento é vulnerável a crises provocadas por guerras, depressões econômicas, políticas econômicas governamentais dentre outros. Também se sabe que a demanda por investimentos é reduzida por aumentos na taxa de juros provocadas por uma política monetária restritiva e/ou uma política fiscal expansionista. Não obstante, o investimento também sofre decréscimos em períodos onde o PIB se encontra em fase de desaceleração por razões diversas.

Por conta disso, é de esperar que os erros estejam autocorrelacionados num modelo econométrico em que o impacto sobre o investimento devido a choques internos e/ou exógenos não podem estar explícitos como uma ou mais variável independente, seja por ausência de informações a respeito de tais variáveis ou por um erro na forma funcional do modelo escolhido.

A título de exemplo, foram consideradas 30 observações para investimento no ano  $t$  ( $I_t$ ) como variável dependente e os Regressores PIB no ano  $t$  ( $Y_t$ ) e taxa de juros no ano  $t$  ( $R_t$ ) (HILL et al, 2003, p.321). O seguinte modelo foi ajustado:

$$\hat{I}_t = 6.2249 + 0.7699 Y_t - 0.1842 R_t \quad R^2 = 0.8163$$

Teste t (2.479)\* (10.724)\*\*\* (-1.457)

Nota: Códigos de significância 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 '' 1

As estimativas  $\hat{\beta}_2$  e  $\hat{\beta}_3$  têm os sinais esperados, ou seja, para um aumento na renda, espera-se um efeito positivo sobre o investimento, enquanto que um crescimento na taxa de juros provoca uma redução sobre  $I$ . Os coeficientes  $\beta_1$  e  $\beta_2$  são estatisticamente significativos pelo teste  $t$ . Já o parâmetro associado à taxa de juros não foi significativamente diferente de 0 o que se leva a questionar sobre a inclusão ou não de  $R_t$  no modelo de regressão. Entretanto, a teoria econômica sugere que a taxa de juros exerce uma forte influência sobre o nível de investimento.

Uma análise de resíduos preliminar por meio de um gráfico de dispersão dos resíduos de mínimos quadrados do modelo acima (Figura 1) revela várias sequências de resíduos negativos seguidos de resíduos positivos, sugerindo a existência de autocorrelação dos erros. O gráfico da função de autocorrelação dos resíduos (ACF) também indica o mesmo, como pode ser visto no gráfico esquerdo da Figura 2, em que algumas das linhas verticais ultrapassam a linha crítica. Construindo-se o gráfico ACF da primeira diferença dos resíduos ( $\Delta e_t = e_t - e_{t-1}$ ), nota-se que a autocorrelação presente provavelmente é de primeira ordem (gráfico direito da Figura 2). Logo, pode-se aplicar o teste de Durbin-Watson para avaliar a existência de autocorrelação serial de primeira ordem, já que as demais hipóteses são satisfeitas.

Figura 1: Gráfico de dispersão dos resíduos para exemplo 1.

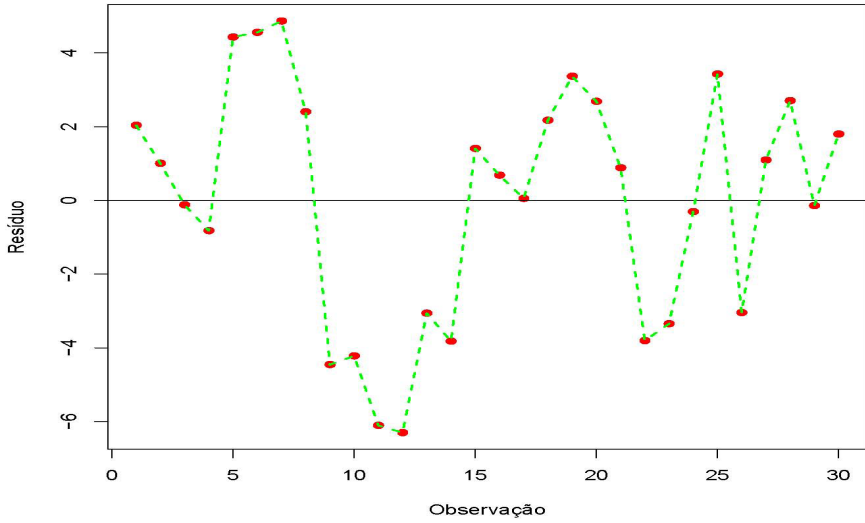
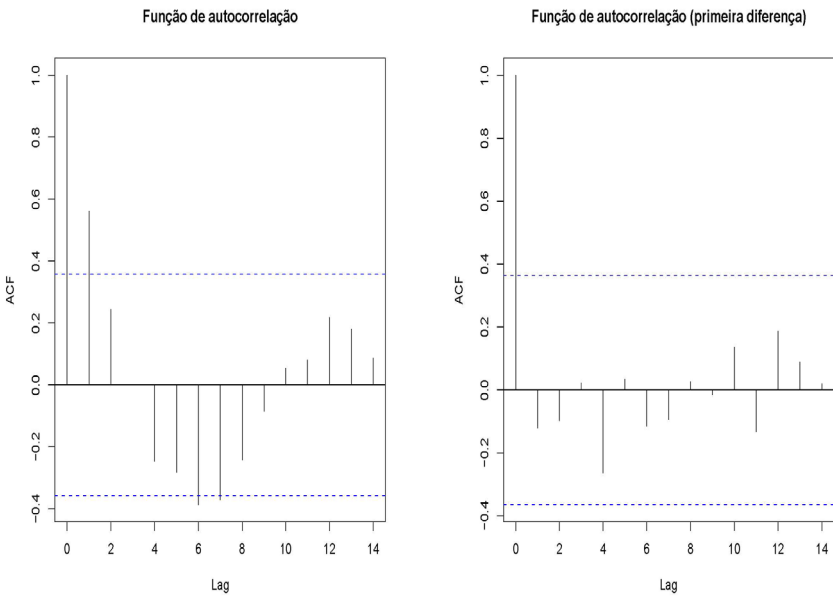


Figura 2: Gráfico da função de autocorrelação dos resíduos (esquerda) e da primeira diferença (direita)



Nesse contexto, a estatística de teste de Durbin-Watson resultou em  $d = 0.8522$ . Utilizando-se o algoritmo *bootstrap* não-paramétrico chegou-se aos quantis abaixo (saída do R) e se pode concluir que os erros apresentam autocorrelação positiva a um nível de significância inferior a 1%.

```
> quantile(D, c(0.01,0.05,0.1,0.9,0.95,0.99))
```

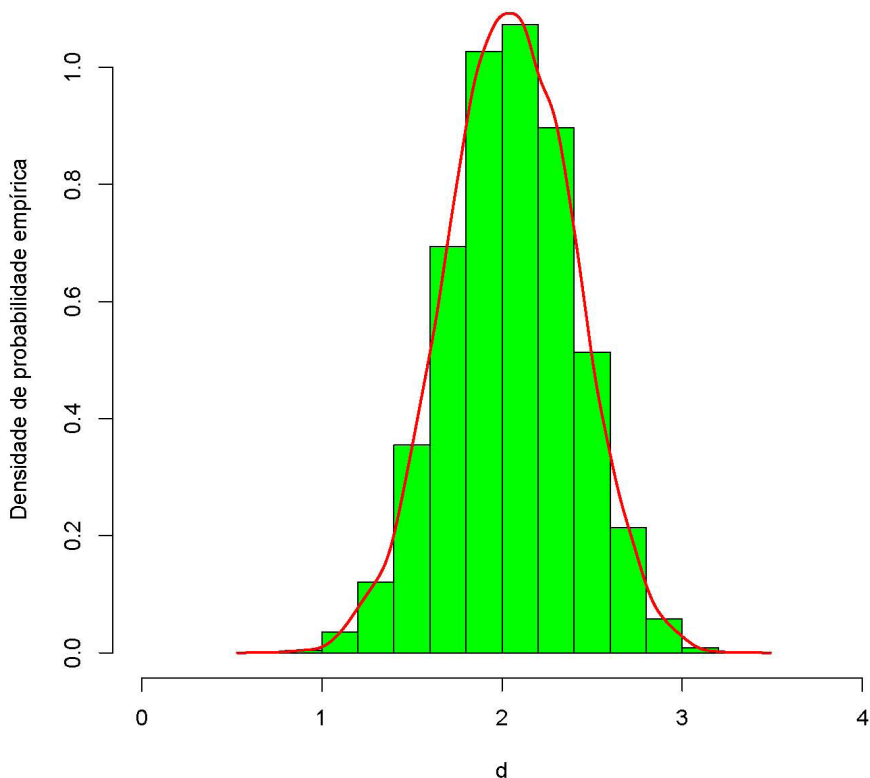
1%	5%	10%	90%	95%	99%
1.223687	1.472354	1.595466	2.494582	2.620432	2.841750

Os valores críticos tabelados  $c$  para o nível de significância 5% são respectivamente iguais a 1.284 e 1.567 para este caso ( $T=30$  e  $K=3$ ), o que leva à mesma conclusão, ou seja, existe autocorrelação serial de primeira ordem positiva. Note que o valor crítico devolvido pelo algoritmo *bootstrap*, para o mesmo nível de significância, está contido no intervalo previsto pela tabela construída a partir das distribuições limites inferior e superior associadas a  $d_L$  e a  $d_U$ , nessa ordem.

Por meio das  $B+1$  estatísticas provenientes do algoritmo *bootstrap*, construiu-se um histograma desses valores e a partir dele, pode-se ter uma ideia do formato da densidade de probabilidade da estatística de teste de Durbin-Watson para esse exemplo específico (Figura 3).



Figura 3: Densidade de probabilidade empírica para exemplo 1



#### OFERTA DE EMPREGO

A oferta de emprego é um indicador que sofre influência de várias variáveis como salário real (quanto maior o salário real, menor a oferta de emprego), variações no nível de preços no mercado (um é inversamente proporcional ao outro), e a taxa de desemprego de uma economia que quanto mais alta, menor será a oferta de emprego.

Na prática, espera-se obter um modelo econométrico que não só englobe todas essas variáveis como também espera-se que o sinal dos estimadores sejam compatíveis com a teoria econômica.

Neste segundo exemplo, optou-se por construir um modelo simples (a priori já se desconfiava que haveria autocorrelação pela ausência de variáveis explicativas importantes), com 24 observações para investigar uma relação entre número de empregos oferecidos no ano  $t$  ( $E0_t$ ) e a taxa de desemprego no mesmo período ( $D_t$ ). (HILL et al., 2003, p. 322). Utilizou-se um modelo log-log, pela conveniência de interpretação dos coeficientes em termos de elasticidades. A estimativa do modelo é:

$$\begin{array}{l} \log(E0_t) = \quad 3.5027 - 1.6116 \log(D_t) \quad R^2=0.8299 \\ \text{Teste t} \quad \quad (12.38)** \quad (-10.36)*** \end{array}$$

Nota: Códigos de significância 0 \*\*\*\*' 0.001 \*\*\*' 0.01 '\*\*' 0.05 '.' 0.1 '' 1

Os parâmetros são todos altamente significativos e as estimativas dos mesmos têm os sinais esperados. Percebe-se que um aumento de 1% na taxa de desemprego diminui a oferta de emprego em 1.6116%, o que permite concluir que a oferta é elástica em relação a  $D$ .

Da mesma forma que no exemplo anterior, procedeu-se a uma análise gráfica dos resíduos para identificação de indícios de autocorrelação serial de primeira ordem dos erros. Tanto o gráfico de dispersão dos resíduos de mínimos quadrados do modelo acima (Figura 4) quanto os gráficos ACF e ACF da primeira diferença (Figura 5) sugerem a presença de erros autocorrelacionados.

Figura 4: Gráfico de dispersão dos resíduos para exemplo 2.

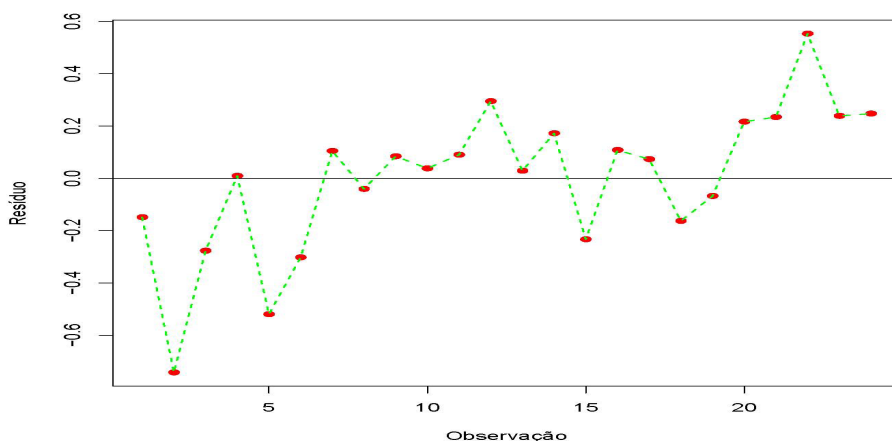
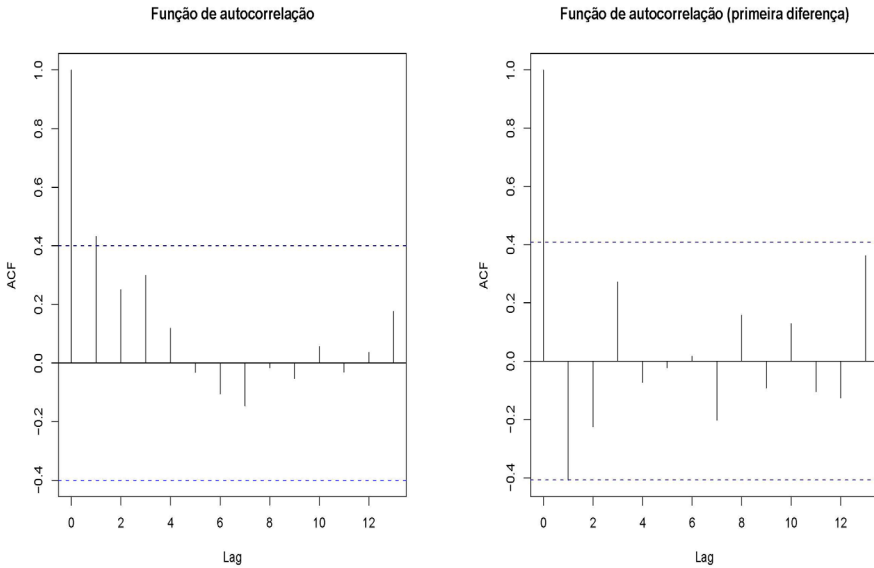


Figura 5: Gráfico da função de autocorrelação dos resíduos (esquerda) e da primeira diferença (direita)



A estatística de teste de Durbin-Watson obtida para esse exemplo foi  $d=1.0896$ . Os quantis provenientes do algoritmo *bootstrap* não-paramétrico são apresentados a seguir:

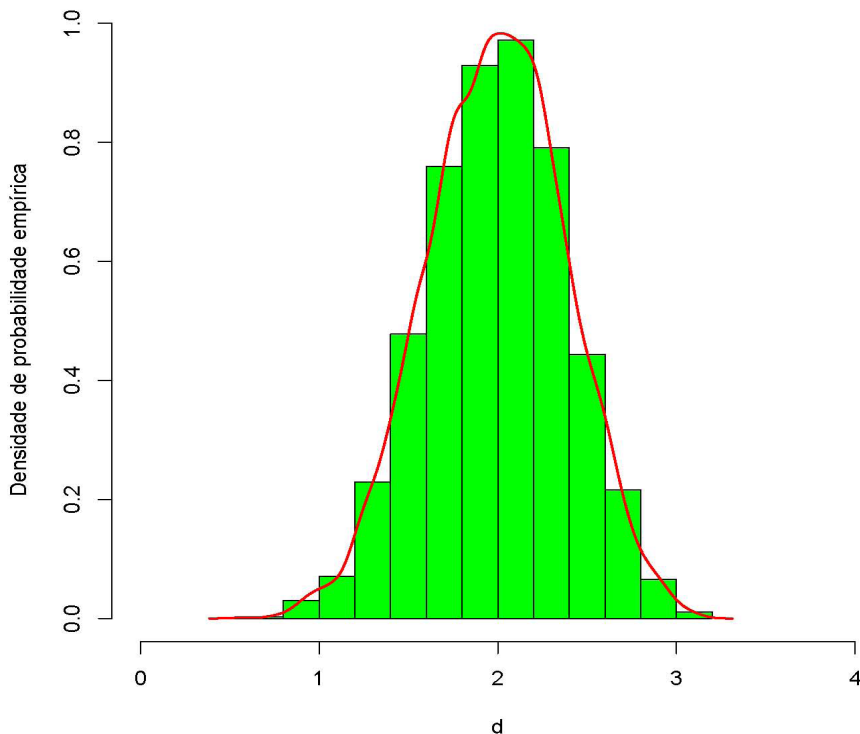
```
> quantile(D, c(0.01,0.05,0.1,0.9,0.95,0.99))
      1%      5%     10%     90%     95%     99%
1.068507 1.341022 1.487424 2.495482 2.623368 2.861997
```

Fixado o nível de significância de 1%, observa-se que a hipótese nula não é rejeitada, ou seja, não há autocorrelação serial de primeira ordem nos erros. Se forem utilizados os valores limites tabelados para essa estatística, para o mesmo nível de significância, o teste é inconclusivo. Isto porque o valor de  $d$  calculado está contido na zona de indecisão definida por  $d_{L_c} = 1.037$  e  $d_{U_c} = 1.199$ , para  $T=24$  e  $K=2$ . É fato que essa situação de dúvida não se repetiria caso um nível maior de significância fosse considerado. Porém, como o intuito neste trabalho é mostrar a vantagem do uso da técnica de bootstrap para fornecer valores críticos exatos para

essa distribuição, considera-se atendido o que se pretendia desde o início. Assim sendo, o resultado disponibilizado pelo algoritmo *bootstrap* proposto é bastante útil para resolver situações onde não se pode afirmar a existência ou não de autocorrelação serial de primeira ordem a partir de valores previamente definidos para  $d_{L_c}$  e  $d_{U_c}$ .

Por fim, a Figura 6 apresenta a forma da densidade de probabilidade da estatística de teste de Durbin-Watson para o modelo estimado entre oferta de emprego e taxa de desemprego.

Figura 6: Densidade de probabilidade empírica para exemplo 2



## CONCLUSÃO

Nesse trabalho foram propostos dois algoritmos *bootstrap* para realização do teste de Durbin-Watson utilizado para detecção de autocorrelação serial de primeira ordem dos erros associados a modelos de regressão linear. A técnica *bootstrap* permitiu que valores críticos exatos para a estatística de teste fossem obtidos, evitando assim a possibilidade de resultados inconclusivos.

Um dos algoritmos envolveu a abordagem paramétrica e outro, um procedimento não-paramétrico. No caso paramétrico, os erros são gerados de acordo com uma distribuição normal com média 0 e variância igual à estimada por meio do modelo ajustado com os dados originais. Já o algoritmo não-paramétrico não requereu suposição alguma sobre a distribuição dos erros “bem comportados” e as amostras *bootstrap* foram construídas a partir de reamostragem com reposição dos resíduos escalonados do modelo estimado inicialmente.

Para ilustrar o uso dos algoritmos, eles foram aplicados a dois exemplos econômicos práticos. O primeiro deles envolveu uma função investimento e o outro, um modelo de oferta de emprego. Para ambos os exemplos, foram obtidos os valores críticos exatos da distribuição da estatística de Durbin-Watson para os níveis de significância de 1%, 5% e 10%, bem como o formato resultante da distribuição empírica associada.

## REFERÊNCIAS

- GUJARATI, D. *Econometria básica*. Rio de Janeiro: Elsevier, 2006.  
Tradução de Maria José Cylar.
- HILL, R. C.; GRIFFITHS, W. E.; JUDGE, G. G. *Econometria*. 2. ed. São Paulo: Saraiva, 2003.
- KENNEDY, P. *Manual de Econometria*. Rio de Janeiro: Elsevier, 2009.
- MACKINNON, J. G. Bootstrapp inference in econometrics. *Canadian Journal of Economics/Revue Canadienne d'Economique*, v. 35, n. 4, p. 615-645, 2002.
- R Development Core Team. *R: A Language and Enviroment for Statistical Computing*. Vienna, Austria, 2010. ISBN 3-900051-07-0. Disponível em <<http://www.R-project.org>>.