

FACULDADES IBMEC

MÉTODOS QUANTITATIVOS

PROF. Cleuton Sampaio de Melo Júnior

Introdução.....	3
Seleção de Variáveis.....	4
Explicação das variáveis.....	4
Ajuste do primeiro modelo.....	6
Seleção Stepwise.....	7
Primeiro Modelo.....	8
Análise de Multicolinearidade.....	9
Segundo Modelo.....	10
Teste de significância de F:.....	10
Coeficiente Ajustado de Determinação (R_a^2).....	11
Análise dos Resíduos.....	12
Teste de Heterocedasticidade.....	12
Normalidade da distribuição de probabilidade dos Erros.....	13
Independência de erros de observações diferentes – Autocorrelação dos resíduos.....	14
Conclusão.....	17
Anexo I – Séries Originais.....	18
Anexo II – Séries Padronizadas.....	20
Anexo III – Testes de Multicolinearidade.....	22
Anexo IV Estatística F.....	23
Anexo V – Teste de Heterocedasticidade.....	24

Introdução

Nesta pesquisa, buscamos estabelecer um relacionamento entre o Desemprego e diversas variáveis econômicas, utilizando a Análise de Regressão como ferramenta.

Existe muita especulação a respeito das causas reais do desemprego no país e, embora seja muito difícil estabelecer uma relação de causa-efeito, é possível saber, dentro de um período de tempo, quais foram os fatores que mais diretamente influenciaram o desemprego.

Devemos salientar que o próprio conceito de desemprego é bastante discutível, uma vez que existem várias formas paralelas de contratação de mão-de-obra, como: mercado informal (camelôs, empresas não legalizadas), atividades ilegais (tráfico de drogas, contrabando etc), cooperativas e micro-empresas. Logo, para termos uma base inicial, optamos pelo índice de desemprego nas regiões metropolitanas, considerado pelo IPEA.

Todas as séries empregadas têm periodicidade mensal, sendo considerado o período entre Janeiro de 1995 e Junho de 2001.

Foram utilizados os softwares: SPSS e Excel nos estudos relativos a este modelo.

Seleção de Variáveis

Todas as variáveis foram obtidas do Web Site do IPEA (www.ipea.gov.br), sendo selecionadas:

- Tendência comum da inflação (%) – Mensal
- Rendimento médio do trabalho principal - RMs – Mensal
- Taxa de juros - TR (% a.m.) – Mensal
- PIB - preços de mercado - índice real - base móvel
- Taxa de câmbio - R\$ / US\$ - comercial - venda - média – Mensal
- Taxa de desemprego - aberto (%) - referência: 30 dias - RMs – Mensal
- Salário mínimo - Mensal

Explicação das variáveis

Tendência comum da inflação (%) – Mensal

Fonte: Instituto de Pesquisa Econômica Aplicada (IPEA)

Unidade: -

Comentário: Esta medida supõe que as taxas de variação do IPCA, Índice de Preços por Atacado (IPA) e INCC são compostas por uma tendência comum e por componentes transitórios específicos a cada um desses índices. A inflação latente é medida pela tendência comum, estimada por modelos formato espaço de estados. Para maiores detalhes, ver: Boletim Conjuntural 48, jan/2000.

Atualizado em: 08 de fevereiro de 2002

Rendimento médio do trabalho principal - RMs – Mensal

Fonte: Instituto Brasileiro de Geografia e Estatística, Pesquisa Mensal de Emprego (IBGE/PME)

Unidade: R\$

Comentário: Regiões metropolitanas (RMs) - Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo e Porto Alegre. Refere-se a pessoas com 15 anos ou mais de idade.

Atualizado em: 25 de janeiro de 2002

Taxa de juros - TR (% a.m.) – Mensal

Fonte: Banco Central do Brasil, Boletim, Seção Mercado financeiro e de capitais (BCB Boletim/M.Finan.)

Unidade: -

Comentário: Quadro: Taxas de juros. Obs.: Taxa Referencial de juros (TR). Criada para ser usada como referência nas transações financeiras realizadas no Brasil. Calculada pelo Banco Central com base em uma amostra dos juros pagos pelos Certificados de Depósito Bancário (CDB) das 30 maiores instituições financeiras do país.

Atualizado em: 21 de janeiro de 2002

PIB - preços de mercado - índice real - base móvel

PIB - preços de mercado - índice real - base móvel (média 1990 = 100) - Mensal

Fonte: Instituto de Pesquisa Econômica Aplicada (IPEA)

Unidade: -

Comentário: Para 1980-1990: Sistema de Contas Nacionais. De 1991 em diante: novo Sistema de Contas Nacionais. Obs.: Produto Interno Bruto (PIB). Elaboração IPEA. Série encadeada do índice de base móvel, ajustada à série anual de variação real do PIB a preços de mercado (IBGE).

Atualizado em: 31 de outubro de 2001

Taxa de câmbio - R\$ / US\$ - comercial - venda - média – Mensal

Fonte: Banco Central do Brasil, Boletim, Seção Balanço de Pagamentos (BCB Boletim/BP)

Unidade: R\$

Comentário: Quadro: Taxas de câmbio do real. Obs.: Cotações para contabilidade. Os valores dos dois últimos meses são obtidos na Gazeta Mercantil.

Atualizado em: 01 de fevereiro de 2002

Taxa de desemprego - aberto (%) - referência: 30 dias - RMs – Mensal

Fonte: Instituto Brasileiro de Geografia e Estatística, Pesquisa Mensal de Emprego (IBGE/PME)

Unidade: -

Comentário: Regiões metropolitanas (RMs) - Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo e Porto Alegre. Refere-se a pessoas com 15 anos ou mais de idade.

Atualizado em: 25 de janeiro de 2002

Salário mínimo – Mensal

Fonte: Gazeta Mercantil

Unidade: R\$

Comentário: Valor nominal.

Atualizado em: 14 de fevereiro de 2002

Estas séries estão no CD-Rom anexo à pesquisa, dentro do arquivo “PesquisaCleuton.XLS”, do Excel, e dentro do arquivo “cleuton2.sav”, do SPSS.

Estão reproduzidas no **ANEXO I**.

Como algumas variáveis apresentavam escala diferente das outras (muito maiores), utilizamos o seu valor “z” (número de desvios-padrões) para transformá-las. Elas estão no **ANEXO II**.

Ajuste do primeiro modelo

Segundo Mendenhall (1996), item 4.9, existem três tipos básicos de Modelos de Regressão:

- Linear de primeira ordem
- Interação
- De segunda ordem

Modelo Linear de Primeira Ordem:

Composto pela Interseção, ou β_0 , e cada variável independente multiplicada pelo seu respectivo coeficiente:

$$E(y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

Modelo de Interação:

Este modelo assume que existe uma interação ou um produto entre duas variáveis, que é acrescentado ao modelo:

$$E(y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$$

Modelo de Segunda Ordem:

Possui termos quadrados, em acréscimo ao modelo de Interação:

$$E(y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \beta_4 X_1^2 + \beta_5 X_2^2$$

Como não temos nenhuma razão para supor que o relacionamento entre estas variáveis seja Linear, Mendenhall (1996) item 4.9, recomenda que se utilize primeiramente o modelo de interação ou de segunda ordem.

Neste caso, construímos vários modelos iniciais, efetuando regressões no SPSS. Começamos pelo Linear de Primeira Ordem, até o Linear de Segunda Ordem, passando pelo uso somente dos quadrados das variáveis.

O modelo que melhor se enquadrou, segundo o Rquadrado Ajustado, foi o linear de segunda ordem:

X1	Tendência da Inflação
X2	Salário mínimo, normalizado
X3	Taxa de Referência – TR mensal
X4	PIB - preços de mercado - índice real - base móvel, normalizado
X5	Rendimento, normalizado
X6	Taxa de Câmbio, normalizada
X1X2	X1 * X2
X1X3	X1 * X3
X1X4	X1 * X4
X1X5	X1 * X5
X1X6	X1 * X6

X2X3	X2 * X3
X2X4	X2 * X4
X2X5	X2 * X5
X2X6	X2 * X6
X3X4	X3 * X4
X3X5	X3 * X5
X3X6	X3 * X6
X4X5	X4 * X5
X4X6	X4 * X6
X5X6	X5 * X6
X1Q	X1^2
X2Q	X2^2
X3Q	X3^2
X4Q	X4^2
X5Q	X5^2
X6Q	X6^2

Seleção Stepwise

A seleção Stepwise se baseia em modelos individuais de cada candidata a variável independente, com a variável dependente, sendo a sua forma:

$$E(y) = \beta_0 + \beta_1 X_1$$

Primeira Etapa:

Conduzir testes de hipótese com cada candidata:

$$H_0: \beta_1 = 0$$

$$H_a: \beta_1 \neq 0$$

Observar e anotar os valores da Estatística T de cada resultado. Deve-se verificar também o “valor-p” (valor de significação de cada Valor T associado. Deve ser menor que 0,05 ou 5%) associado a cada variável. Se igual ou maior que 0,05, está fora da área de rejeição da hipótese nula, devendo ser descartada do Modelo.

A variável candidata com maior valor absoluto de Estatística T será considerada a primeira variável para a segunda etapa.

Segunda Etapa:

Faz-se regressão entre a variável Dependente, a Primeira Variável, selecionada na primeira etapa, e qualquer uma das outras, segundo o Modelo:

$$E(y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

Verifica-se o “valor p” da variável X2, o qual deve estar menor que 0,05 para que ela faça parte do Modelo. Deve-se, também, observar o “valor p” da variável incluída anteriormente, neste caso: X1, de modo que esteja abaixo de 0,10, o que provocaria a sua retirada do Modelo.

A variável que apresentar o maior valor absoluto de Estatística T, deverá entrar como a segunda variável, ao lado de X1.

Próximas etapas:

Repete-se os testes da Etapa 2, agora tendo como Modelo:

$$E(y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_n X_i$$

Igualmente, deve-se tomar o cuidado de verificar se o “valor p” de alguma das variáveis anteriores ultrapassou 0,10.

Mendenhall (1996), página 242

Primeiro Modelo

Submetemos o primeiro modelo completo ao processo de Seleção Stepwise, através do SPSS, obtendo o seguinte resultado:

$$E(y) = 17,505 - 0,980X_1 + 0,315X_6 - 0,469X_5X_6 + 1,734X_2 - 2,112X_1X_2 \\ - 0,722X_2X_5 + 0,163X_3X_4$$

Análise de Multicolinearidade

A Multicolinearidade é o relacionamento (Correlação) forte entre duas variáveis independentes. Quando duas ou mais variáveis independentes, utilizadas no Modelo, são correlacionadas, ocorre o fenômeno da Multicolinearidade.

A presença deste fenômeno indica que é problemática a interpretação de cada β individual, com o propósito de buscar causa e efeito. Uma das alternativas para sua correção seria manter as variáveis problemáticas, evitando inferências de causa e efeito e restringindo estimativa sobre valores futuros a uma região de experimentação.

Para detectar a multicolinearidade, devemos analisar o indicador de Variance Inflation Factor (VIF) de cada parâmetro β . Caso algum deles seja maior que 10, existe a multicolinearidade.

Segundo alguns autores, um valor VIF maior ou igual a 10 seria problemático, já, para outros, um valor superior a 7 também seria. A abordagem mais simples seria retirar todas as variáveis cujo VIF fosse superior a 10. Neste caso, estaríamos correndo o risco de retirar variáveis de contribuição significativa para o Modelo.

Neste caso temos o seguinte quadro:

Variável	VIF
TIX1	5,854
CBX6	6,748
X5X6	10,501
SMZX2	61,989
X1X2	102,288
X2X5	14,372
X3X4	3,543

Fizemos então as seguintes análises e tentativas:

1) Analisamos a correlação entre as variáveis, usando o quadro “Coefficient Correlations”, do SPSS, que está no **ANEXO III**. Descobrimos as seguintes correlações:

Variável	Correlação 1	Correlação 2	Correlação 3	Correlação 4
X2	-0,60 X6	-0,905 X1X2	-0,788 X2X5	
X1X2	0,407 X6	-0,449 X5X6	-0,905 X2	-0,907 X2X5
X2X5	-0,460 X5X6	-0,788 X2	0,907 X1X2	

Decidimos retirar do Modelo as variáveis X1X2 e X2X5, porque são derivadas das originais.

2) Na tentativa seguinte, ainda tivemos multicolinearidade:

Variável	VIF
TIX1	5,197
CBX6	218,407
X5X6	8,263
X6Q	217,531
SMZX2	10,997

Analizamos novamente o quadro de “Coefficients Correlations”, do SPSS e decidimos retirar do Modelo a variável X6Q.

3) Na terceira tentativa, ficamos com o seguinte quadro de VIF:

Variável	VIF
TIX1	55,601
CBX6	5,183
X5X6	53,241
X1Q	97,44
RZX5	74,991
PIZX4	2,163
TRX3	8,073

Mais uma vez, analisando o “Coefficient Correlations”, do SPSS, decidimos retirar as variáveis X1q e X5X6.

4) No último modelo, ficamos com o seguinte quadro:

Variável	VIF
TIX1	5,313
CBX6	1,96
RZX5	6,963

Tentamos outras experiências, inserido outras variáveis, mas este, realmente ficou sendo o Modelo pós Multicolinearidade.

Segundo Modelo

$$E(y) = 17,085 - 1,131X1 + 0,455X6 - 0,614X5$$

Teste de significância de F:

$$H_0: \beta_1 = \beta_1 = \beta_1 = \dots = \beta_1 = 0$$

H_a : Pelo menos um dos β 's é diferente de Zero.

No **ANEXO IV** temos o resultado do Segundo Modelo, com o valor não significativo de F, logo rejeitando a hipótese nula e não rejeitando a hipótese alternativa.

Coeficiente Ajustado de Determinação (R_a^2)

Este coeficiente é derivado do Coeficiente Múltiplo de Determinação ou Rquadrado. O Rquadrado é calculado pelo:

$$R^2 = 1 - SSE / SS_{yy}$$

Onde $SSE = \sum (y_i - y_{pi})^2$

E $SS_{yy} = \sum (y_i - y_{mi})^2$

y_{pi} = Valor previsto de y_i e y_{mi} = valor médio de y_i .

Se o Rquadrado for próximo de 1, significa que o modelo está explicando bem a variação da variável independente.

O problema com o Rquadrado é que ele não é ajustado para o número de observações e o número de parâmetros do Modelo. Logo, ele pode ser forçado a assumir um valor alto quando temos muitos parâmetros ou observações. Uma medida mais sensata é o RquadradoAjustado, cujo cálculo leva em consideração estes números, sendo menos suscetível ao “mascaramento”. Sua fórmula é:

$$R_a^2 = 1 - ((n - 1) / (n - (k + 1))) * (1 - R^2)$$

Onde n = número de observações e k = número de parâmetros.

O RquadradoAjustado do Modelo é: 0,795, significando que 79,5% dos erros estão explicados pelo modelo.

Análise dos Resíduos

Para que uma regressão seja considerada válida, ela tem que atender as quatro condições básicas:

- 1) A média da distribuição de probabilidades do erro aleatório ε deve ser zero.
- 2) A variância da distribuição de probabilidades de ε deve ser constante.
- 3) A distribuição de probabilidades de ε deve ser normal
- 4) Os erros associados a qualquer duas observações diferentes devem ser independentes.

A primeira condição pode ser atendida com um número maior de observações, sendo que a média dos erros deverá tender a zero.

Teste de Heterocedasticidade

Uma das condições necessárias para a validade das inferências de uma regressão é que o termo do erro aleatório, ε , tenha uma variância constante para todos os níveis de variáveis independentes.

Quando esta condição é satisfeita, o modelo é dito Homocedástico. Quando são observadas variações desiguais para diferentes conjuntos de variáveis independentes, o modelo é dito Heterocedástico.

A razão mais comum de heterocedasticidade é que a variância do y previsto é uma função de sua média ($E(y)$). Neste caso, pode-se recalculer o modelo aplicando-se uma das funções *conhecidas*: Poisson, Binomial ou Multiplicativa.

Mendenhall (1996), páginas 395 a 397

Para testar a condição de Homocedasticidade, vamos seguir a sugestão de Mendenhall (1996), pg 404, nota de rodapé:

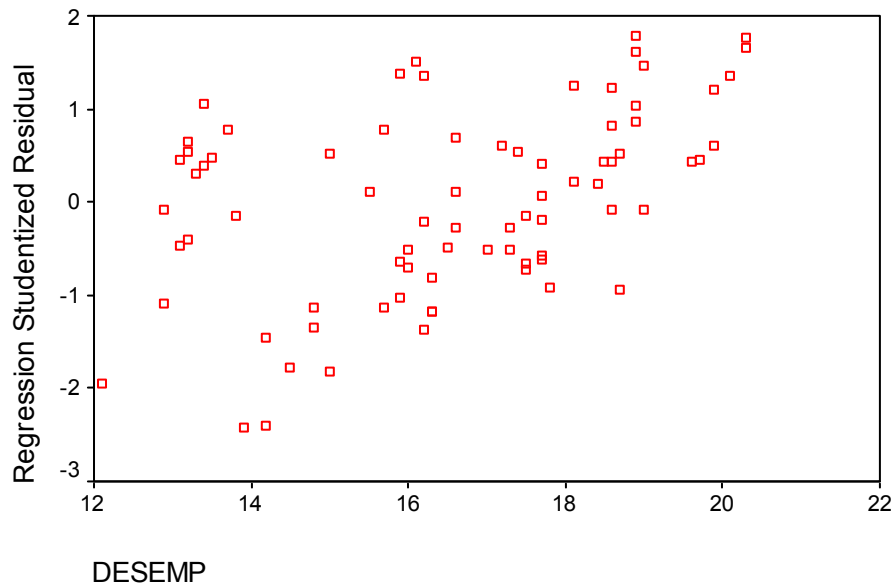
“Considere testar o valor absoluto dos resíduos em função das variáveis independentes no Modelo, por exemplo: teste o modelo $E(|\varepsilon|) = \beta_0 + \beta_1 X_1 \dots$, um valor não significativo de F implica que a condição de Homocedasticidade está satisfeita.”

Jogamos o modelo no Excel (**ANEXO V**) e fizemos a regressão, obtendo a listagem dos resíduos. Após criar uma coluna com o Seu valor absoluto, a substituímos no modelo e refizemos a regressão, obtendo o valor de significância de $F = 0,02172349$, logo, nosso modelo é **Homocedástico**.

Isto pode ser comprovado no gráfico de plotagem dos resíduos:

Scatterplot

Dependent Variable: DESEMP

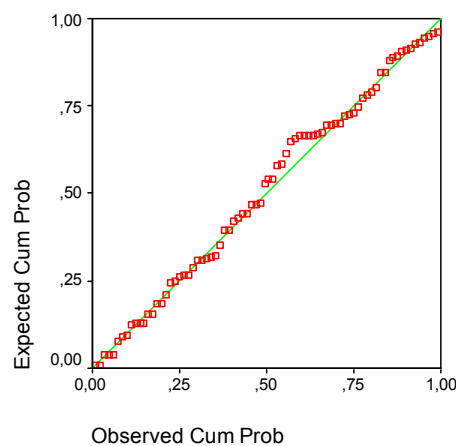


Normalidade da distribuição de probabilidade dos Erros

Existem alguns testes para saber se a distribuição de probabilidade dos erros será normal. Um deles é observar o gráfico de distribuição de probabilidade normal (Normal Probability Plot):

Normal P-P Plot of Regression Stand:

Dependent Variable: DESEMP



Este gráfico deve estar aproximadamente como uma linha reta.

Outro teste é o de que no máximo 5% dos resíduos padronizados esteja fora do intervalo -2 e +2. Fizemos este teste no Excel e obtivemos o valor de: 0,025641 ou **2,5%**. Logo, consideramos satisfeita esta condição. Isto pode ser observado no **ANEXO V**.

Independência de erros de observações diferentes – Autocorrelação dos resíduos

O teste de Autocorrelação dos resíduos é feito através da estatística Durbin-watson:

$$d = \left(\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2 \right) / \left(\sum_{t=1}^n \varepsilon_t^2 \right)$$

Onde ε = Erro previsto (com “^”).

O intervalo da estatística “d” deve ser entre 0 e 4, e sua interpretação é:

- $d \approx 2$, não há autocorrelação entre os resíduos.
- $d \approx 0$, há forte autocorrelação positiva.
- $d \approx 4$, há forte autocorrelação negativa.

Mendenhall (1996), página 430.

Segundo o **ANEXO IV** temos um índice de 0,354, representando forte autocorrelação positiva entre os resíduos. Isto significa que o erro de hoje está influenciando o de amanhã.

O efeito causado pela autocorrelação no modelo linear geral depende do padrão de autocorrelação. Um dos padrões mais comuns é que a autocorrelação entre observações consecutivas no tempo é positiva. Quando o resíduo na observação em t é positivo (y observado maior que o esperado), logo o valor da observação seguinte ($t+1$) tende a ser positivo também. Isto demonstra a existência de autocorrelação positiva entre resíduos.

Outra propriedade comum é que a autocorrelação entre dois pontos diminui rapidamente quando se aumenta a distância (t) entre eles. Modelos que apresentam este comportamento são conhecidos como **Modelos Autorregressivos de Primeira Ordem**. E as séries que dependem apenas da distância entre os pontos são conhecidas como **Séries Temporais Estacionárias**.

Quais são as consequências da Autocorrelação? Segundo Mendenhall (1996) pg 534, “*Embora os estimadores dos Betas continuem não-enviezados, os Erros Padrões, dados pela teoria dos mínimos quadrados, são normalmente menores que os verdadeiros erros padrões quando os resíduos estão positivamente autocorrelacionados.*”. Isto aumentaria as chances de erros do Tipo I, incluindo variáveis não significativas para o modelo.

Existem duas explicações para Autocorrelação dos resíduos:

- 1) Problemas de séries temporais (Tendência Secular, Efeito Cíclico ou Sazonalidade).
- 2) Segundo explicação do Prof Gérson, poderia ser a *Falta de variáveis independentes*.

Se o problema for um dos efeitos de séries temporais, poderíamos utilizar algumas das técnicas de construção de séries temporais, citadas em Mendenhall (1996) pg 528, para criar novas séries temporais, como: Lagging ou outras técnicas de suavização como: média móvel, exponencial ou Holt-Winters.

Ainda Mendenhall (1996) pgs 533 a 536, ensina a ajustar o Modelo de Mínimos Quadrados através do seu produto pelo Modelo de Autoregressão, criando um modelo transformado apropriado.

Como utilizamos o SPSS, fizemos vários testes criando novas Séries Temporais, utilizando as diversas funções disponíveis, como:

Difference. Nonseasonal difference between successive values in the series. The order is the number of previous values used to calculate the difference. Because one observation is lost for each order of difference, system-missing values appear at the beginning of the series. For example, if the difference order is 2, the first two cases will have the system-missing value for the new variable.

Seasonal difference. Difference between series values a constant span apart. The span is based on the currently defined periodicity. To compute seasonal differences, you must have defined date variables (Data menu, Define Dates) that include a periodic component (such as months of the year). The order is the number of seasonal periods used to compute the difference. The number of cases with the system-missing value at the beginning of the series is equal to the periodicity multiplied by the order. For example, if the current periodicity is 12 and the order is 2, the first 24 cases will have the system-missing value for the new variable.

Centered moving average. Average of a span of series values surrounding and including the current value. The span is the number of series values used to compute the average. If the span is even, the moving average is computed by averaging each pair of uncentered means. The number of cases with the system-missing value at the beginning and at the end of the series for a span of n is equal to $n/2$ for even span values and for odd span values. For example, if the span is 5, the number of cases with the system-missing value at the beginning and at the end of the series is 2.

Prior moving average. Average of the span of series values preceding the current value. The span is the number of preceding series values used to compute the average.

The number of cases with the system-missing value at the beginning of the series is equal to the span value.

Running median. Median of a span of series values surrounding and including the current value. The span is the number of series values used to compute the median. If the span is even, the median is computed by averaging each pair of uncentered medians. The number of cases with the system-missing value at the beginning and at the end of the series for a span of n is equal to $n/2$ for even span values and for odd span values. For example, if the span is 5, the number of cases with the system-missing value at the beginning and at the end of the series is 2.

Cumulative sum. Cumulative sum of series values up to and including the current value.

Lag. Value of a previous case, based on the specified lag order. The order is the number of cases prior to the current case from which the value is obtained. The number of cases with the system-missing value at the beginning of the series is equal to the order value.

Lead. Value of a subsequent case, based on the specified lead order. The order is the number of cases after the current case from which the value is obtained. The number of cases with the system-missing value at the end of the series is equal to the order value.

Smoothing. New series values based on a compound data smoother. The smoother starts with a running median of 4, which is centered by a running median of 2. It then resmooths these values by applying a running median of 5, a running median of 3, and hanning (running weighted averages). Residuals are computed by subtracting the smoothed series from the original series. This whole process is then repeated on the computed residuals. Finally, the smoothed residuals are computed by subtracting the smoothed values obtained the first time through the process. This is sometimes referred to as T4253H smoothing.

Como não obtivemos melhora significativa da estatística Durbin-Watson, consideramos que o problema esteja mais relacionado à falta de variáveis independentes, o que poderia ser resolvido com mais pesquisas. Portanto vamos aceitar, dentro da nossa proposição, o Modelo Encontrado até agora, lembrando os problemas potenciais causados pela Autocorrelação positiva.

Conclusão

O último modelo encontrado é este:

$$E(y) = 17,085 - 1,131X1 + 0,455X6 - 0,614X5$$

Onde X1 é a Tendência comum da Inflação, X6 é a taxa de Câmbio e X5 é o Rendimento Médio do Trabalho Principal.

Observando o Modelo, concluímos que existe Correlação positiva entre o Desemprego e o Câmbio, porém negativa entre o Desemprego e a Tendência da Inflação e o Rendimento Médio.

Anexo I – Séries Originais

Data	Tendência comum da inflação (%)	Sal Mínim o	TR Mensal	Dese mpreg o Taxa(%)	PIB	Rendi mento médio	Câmbio
1995 01	1,776771	85	2,1013	12,1	112,8426	420,34	0,847136
1995 02	1,715498	70	1,8531	12,9	109,5376	435,78	0,84075
1995 03	1,753528	70	2,2998	13,2	123,4769	450,71	0,88937
1995 04	1,808986	70	3,4667	13,5	115,0755	467,6	0,907529
1995 05	1,870854	100	3,2471	13,4	116,4449	487,02	0,897364
1995 06	1,800053	100	2,8863	13,2	118,3114	499,44	0,914048
1995 07	1,760191	100	2,9905	13,1	117,7354	509,48	0,928762
1995 08	1,664566	100	2,6045	12,9	118,3815	521,96	0,941957
1995 09	1,63611	100	1,9393	13,3	116,5566	530,43	0,95275
1995 10	1,612452	100	1,654	13,4	118,7657	537,86	0,959681
1995 11	1,549462	100	1,4387	13,7	117,3866	561,62	0,963375
1995 12	1,478955	100	1,34	13,2	111,2698	600,62	0,968285
1996 01	1,391922	100	1,2526	13,1	111,8908	576,38	0,974477
1996 02	1,281407	100	0,9625	13,8	110,376	587,91	0,981079
1996 03	1,17828	100	0,8139	15	118,3495	587,37	0,986119
1996 04	1,143993	100	0,6597	15,9	117,4556	594,71	0,990114
1996 05	1,081579	112	0,5888	16,1	120,6217	609,63	0,995282
1996 06	1,003887	112	0,6099	16,2	118,5793	619,36	1,001332
1996 07	0,930541	112	0,5851	15,7	125,7916	639,63	1,006939
1996 08	0,829414	112	0,6275	15,5	123,329	644,18	1,013414
1996 09	0,75656	112	0,662	14,8	122,1427	636,43	1,01929
1996 10	0,7152	112	0,7419	14,8	124,9253	636,95	1,025114
1996 11	0,681157	112	0,8146	14,5	121,4805	641,44	1,030385
1996 12	0,663106	112	0,8717	14,2	117,9507	686,66	1,037121
1997 01	0,679717	112	0,744	13,9	116,7187	641,75	1,042868
1997 02	0,628506	112	0,6616	14,2	114,5389	642,73	1,0493
1997 03	0,613271	112	0,6316	15	123,1649	634,11	1,056653
1997 04	0,609388	112	0,6211	15,9	124,6455	649,94	1,0609
1997 05	0,566259	120	0,6354	16	122,9218	666,7	1,06827
1997 06	0,540342	120	0,6535	16	124,8489	664,5	1,074595
1997 07	0,482535	120	0,658	15,7	128,108	675,23	1,080652
1997 08	0,439228	120	0,627	15,9	125,2824	684,17	1,087867
1997 09	0,414068	120	0,6474	16,3	127,5396	681,18	1,093591
1997 10	0,412896	120	0,6553	16,5	129,655	689,63	1,1001
1997 11	0,418309	120	1,5334	16,6	122,8352	695,49	1,1073
1997 12	0,444552	120	1,3085	16,6	119,5091	744,11	1,1134
1998 01	0,478004	120	1,1459	16,6	116,6696	700,7	1,119933
1998 02	0,473647	120	0,4461	17,2	114,4369	696,29	1,127122
1998 03	0,460395	120	0,8995	18,1	127,0712	685	1,1337
1998 04	0,425159	120	0,472	18,9	124,9461	679,52	1,141195
1998 05	0,41091	130	0,4543	18,9	126,0961	675,01	1,14813
1998 06	0,339897	130	0,4913	19	127,1187	682,89	1,154571
1998 07	0,292046	130	0,5503	18,9	130,1293	678,72	1,161491
1998 08	0,242962	130	0,3749	18,9	125,9988	685,74	1,171724

Métodos Quantitativos
Cleuton Sampaio

1998 09	0,243148	130	0,4512	18,5	125,3797	685,86	1,180886
1998 10	0,254086	130	0,8892	18,1	125,0415	695,24	1,188367
1998 11	0,242611	130	0,6136	17,7	122,4674	715,28	1,193775
1998 12	0,257138	130	0,7434	17,4	117,6639	758,1	1,20535
1999 01	0,263406	130	0,5163	17,8	116,6693	687,15	1,5019
1999 02	0,296543	130	0,8298	18,7	114,0136	678,78	1,9137
1999 03	0,344386	130	1,1614	19,9	127,2316	677,9	1,8968
1999 04	0,380198	130	0,6092	20,3	124,4585	676,92	1,6941
1999 05	0,396188	136	0,5761	20,3	125,0792	676,78	1,6835
1999 06	0,38407	136	0,3108	19,9	126,5484	683,35	1,7654
1999 07	0,412943	136	0,2933	20,1	127,6956	674,76	1,8003
1999 08	0,371273	136	0,2945	19,6	127,1136	676,05	1,8808
1999 09	0,34924	136	0,2715	19,7	125,796	679,52	1,8981
1999 10	0,392947	136	0,2265	19	126,4462	688,36	1,9695
1999 11	0,390427	136	0,1998	18,6	125,3454	707,15	1,92989
1999 12	0,393262	136	0,2998	17,5	124,7032	757,68	1,84283
2000 01	0,417472	136	0,2149	17,7	120,2585	707,66	1,8037
2000 02	0,41125	136	0,2328	17,7	123,3947	702	1,7753
2000 03	0,428583	136	0,2242	18,4	130,4556	698,34	1,742033
2000 04	0,442341	151	0,1301	18,6	128,2219	699,57	1,768168
2000 05	0,426773	151	0,2492	18,7	131,4457	711,64	1,827932
2000 06	0,443673	151	0,214	18,6	132,9015	727,58	1,808329
2000 07	0,529665	151	0,1547	18,6	133,5872	723,91	1,797824
2000 08	0,530236	151	0,2025	17,7	134,6978	731,5	1,809239
2000 09	0,483168	151	0,1038	17,3	129,8327	733,99	1,83919
2000 10	0,488979	151	0,1316	16,3	132,855	745,84	1,879638
2000 11	0,511484	151	0,1197	16,2	130,6971	743,99	1,94801
2000 12	0,540255	151	0,0991	16,2	128,1654	805,07	1,96325
2001 01	0,548464	151	0,1369	16,3	126,6589	738,5	1,9545
2001 02	0,542904	151	0,0368	17	123,8503	742,25	2,0019
2001 03	0,533588	151	0,1724	17,3	137,4232	740,08	2,0891
2001 04	0,540967	180	0,1546	17,7	133,6749	746,12	2,1925
2001 05	0,54651	180	0,1827	17,5	136,4109	740,4	2,2972
2001 06	0,561394	180	0,1458	17,5	133,4361	750,8	2,3758

Anexo II – Séries Padronizadas

SminZ	PIBZ	RendZ
-1,73546	-1,66237	-2,91362
-2,39967	-2,17943	-2,72454
-2,39967	0,001337	-2,54171
-2,39967	-1,31304	-2,33488
-1,07125	-1,09881	-2,09707
-1,07125	-0,8068	-1,94498
-1,07125	-0,89691	-1,82203
-1,07125	-0,79583	-1,66921
-1,07125	-1,08133	-1,56549
-1,07125	-0,73572	-1,4745
-1,07125	-0,95148	-1,18354
-1,07125	-1,90844	-0,70596
-1,07125	-1,81128	-1,0028
-1,07125	-2,04826	-0,8616
-1,07125	-0,80084	-0,86822
-1,07125	-0,94069	-0,77833
-0,53988	-0,44535	-0,59563
-0,53988	-0,76489	-0,47648
-0,53988	0,363461	-0,22826
-0,53988	-0,02181	-0,17254
-0,53988	-0,20739	-0,26744
-0,53988	0,227932	-0,26108
-0,53988	-0,31099	-0,20609
-0,53988	-0,86323	0,347658
-0,53988	-1,05597	-0,2023
-0,53988	-1,39698	-0,19029
-0,53988	-0,04747	-0,29585
-0,53988	0,184161	-0,102
-0,18564	-0,0855	0,103234
-0,18564	0,215979	0,076294
-0,18564	0,725864	0,20769
-0,18564	0,2838	0,317167
-0,18564	0,636935	0,280552
-0,18564	0,967877	0,384028
-0,18564	-0,09905	0,455788
-0,18564	-0,61942	1,051174
-0,18564	-1,06365	0,519588
-0,18564	-1,41295	0,465584
-0,18564	0,563657	0,327331
-0,18564	0,231192	0,260224
0,257169	0,411104	0,204996
0,257169	0,571081	0,301492
0,257169	1,042091	0,250428
0,257169	0,395878	0,336392

Métodos Quantitativos
Cleuton Sampaio

0,257169	0,299026	0,337862
0,257169	0,246116	0,452727
0,257169	-0,15659	0,69813
0,257169	-0,9081	1,222491
0,257169	-1,0637	0,353659
0,257169	-1,47917	0,251162
0,257169	0,588756	0,240386
0,257169	0,154905	0,228385
0,522853	0,252015	0,226671
0,522853	0,481869	0,307125
0,522853	0,661344	0,201935
0,522853	0,570286	0,217732
0,522853	0,364155	0,260224
0,522853	0,465881	0,368476
0,522853	0,293652	0,598573
0,522853	0,193185	1,217348
0,522853	-0,50217	0,604818
0,522853	-0,01152	0,535507
0,522853	1,093129	0,490688
1,187063	0,743677	0,50575
1,187063	1,248027	0,653556
1,187063	1,475784	0,848752
1,187063	1,58306	0,803811
1,187063	1,756817	0,896755
1,187063	0,995682	0,927247
1,187063	1,468514	1,072359
1,187063	1,130918	1,049704
1,187063	0,73484	1,797671
1,187063	0,499155	0,982475
1,187063	0,059754	1,028397
1,187063	2,183196	1,001823
2,471203	1,596782	1,075787
2,471203	2,024827	1,005742
2,471203	1,559422	1,133097

Anexo III – Testes de Multicolinearidade

Modelo	Arquivo
Inicial	reg2SegundaOrdemInicial.spo
Segundo	Res2SegundaOrdemSegundoModelo.spo
Terceiro	Reg2SegundaOrdemTerceiro.spo
Quarto	Reg2SegundaOrdemQuarto.spo
Quinto	Reg2SegundaOrdemQuintoSemMulticolinearidade.spo

Anexo IV Estatística F

Model Summary

Model	Adjusted R Square	Std. Error of the Estimate	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	,688	1,2170	170,866	1	76	,000	
2	,743	1,1056	17,089	1	75	,000	
3	,795	,9858	20,339	1	74	,000	,354

a Predictors: (Constant), TIX1

b Predictors: (Constant), TIX1, CBX6

c Predictors: (Constant), TIX1, CBX6, RZX5

d Dependent Variable: DESEMP

Anexo V – Teste de Heterocedasticidade

Modelo: $E(\varepsilon) = \beta_0 + \beta_1 X_1 \dots$

ANOVA					
	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	3	2,773068	0,924356	3,414088	0,02172349
Resíduo	74	20,03532	0,270748		
Total	77	22,80839			

No arquivo “PesquisaCleuton2.XLS”, pasta HETERO, temos a regressão completa. Na pasta “Modelo” temos, na coluna ABSRESID, o valor absoluto dos resíduos.

Teste de Normalidade da Distribuição dos Resíduos

Soma 2
Percentual 0,025641

Está no mesmo arquivo, pasta “Modelo1”.