



ISSN: 1696-8352 - BRASIL – FEBRERO 2016

ANÁLISE ECONOMETRICA DO IMPACTO DO PREÇO E DA ÁREA COLHIDA NA OFERTA DE GOIABA NO ESTADO DO PARÁ

Brenda de Farias Oliveira Cardoso

UEPA

Gabriel Rodrigues de Oliveira Gadelha

UEPA

Heriberto Wagner Amanajás Pena

UEPA; FACI/DEVRY

Resumo

Neste trabalho determinou-se a análise da quantidade ofertada de goiaba em relação ao preço do fruto por tonelada e sua área colhida com o auxílio dos softwares Excel 2007 e EvIEWS 3.0, e a partir do Banco de Dados Agregados disponibilizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) foi possível coletar os dados de quantidade produzida de goiaba ao longo de 21 (vinte e um) anos no estado do Pará. Com estes dados e para estimação do modelo foi possível realizar os testes de correlação, regressão linear, desenvolver relações estatísticas e encontrar às variáveis componentes para propor uma equação de Oferta de goiaba em função do tempo, através do método de regressão múltipla. Foram feitos os testes necessários e analisaram-se os resultados obtidos a partir do software EvIEWS 3.0.

Palavras-chave: Oferta; Banco de Dados; Modelagem Matemática; Estatística Aplicada.

ECONOMETRIC ANALYSIS OF THE IMPACT OF PRICE AND CROPPED AREA IN OFFER GOIABA IN PARA STATE.

Abstract

In this work we determined the analysis of the quantity supplied of guava on the price of fruit per ton and area harvested with the help of Excel 2007 software and EvIEWS 3.0, and from the Aggregates Database provided by the Brazilian Institute of Geography and Statistics (IBGE) it was possible to collect the amount of data produced guava over 21 (twenty-one) years of Pará's state. With this data and model estimation was possible to realize the correlation tests, linear regression develop statistical relationships and find the variable components to propose a guava Offer equation a function of time, through multiple regression method. The necessary tests were made and analyzed the results obtained from the EvIEWS 3.0 software.

Keywords: Offering; Database; Mathematical Modeling; Applied statistics.

1. Introdução

A goiaba tem suas origens nos trópicos americanos, e é distribuída hoje ao longo das áreas tropicais e subtropicais do mundo (Kwee e Chong, 1990; Gonzaga Neto e Soares, 1994; Medina, 1988). Não há uma resposta concreta com relação ao seu local de origem, deduzindo-se que a mesma originou-se de uma área que se estende do México meridional pela ou através da América Central (Morton, 1987). Apesar de todo esse questionamento, a goiaba hoje é unanimidade em diversos países, podendo ser localizada em diferentes regiões devido a fácil disseminação e capacidade de manter-se em diversos climas.

No cenário agrícola das grandes potências econômicas mundiais (Estados Unidos e União Europeia), a goiaba ainda não possui grande espaço entre os produtos de hortifrúti. Mesmo esses países/continente tendo o mercado que mais consome este tipo de produto, a fruta é tida como exótica e é comercializada com altos preços e em baixa escala. Apesar de o Brasil ser um dos maiores produtores da fruta, existem fatores que desestimulam a participação do mesmo no mercado internacional. São eles: o elevado grau de perecibilidade do fruto (o que dificulta e exige um cuidado maior para realizar a exportação); e a falta de conhecimento dos consumidores deste mercado mais rentável.

Esse cenário crítico muda ao analisar a cultura da goiaba no Brasil, que por se tratar de um clima tropical e favorável à fruticultura ajuda e estimula o crescimento da produção da goiaba. Além disso, está é uma fruta que se adapta a quase todo o tipo clima, fazendo com que se produza nas cinco regiões do país. Na região sudeste, destaca-se os estados de São Paulo, Minas Gerais e Rio de Janeiro como os maiores produtores; Bahia, Pernambuco e Paraíba, na região nordeste; Goiás, no centro-oeste e Rio Grande do Sul e Paraná na região sul. Mesmo a região norte não estando entre os maiores produtores, o estado possui uma gama de municípios que produzem o fruto e tem elevado potencial para o cultivo, como a cidade de Dom Eliseu, localizado no nordeste do Pará. De 1999 a 2014 o município manteve a produção de 1.000 a 20.000 toneladas por ano (IBGE).

Além desses fatores, vale ressaltar a importância da fruticultura para elevar a economia local ajudando assim o desenvolvimento do país. Os efeitos desse desenvolvimento é o aumento de emprego, que gera a independência financeira da população rural, levando ao crescimento dessas regiões e consequente distribuição de renda regional mais justa. A partir dessa análise, selecionou-se a cultura da goiaba como objeto de estudo, com finalidade de avaliar o impacto do preço e da área colhida na oferta no estado do Pará através da proposição de uma equação econométrica da oferta do fruto na região, assim como auxiliar possíveis previsões de oferta elaboradas futuramente baseadas neste trabalho. Desse modo, o artigo em questão está dividido segundo a seguinte ordem: Introdução, Referencial Teórico, Metodologia e Análise Econométrica, Discussão dos Resultados e Considerações Finais.

2. Referencial Teórico

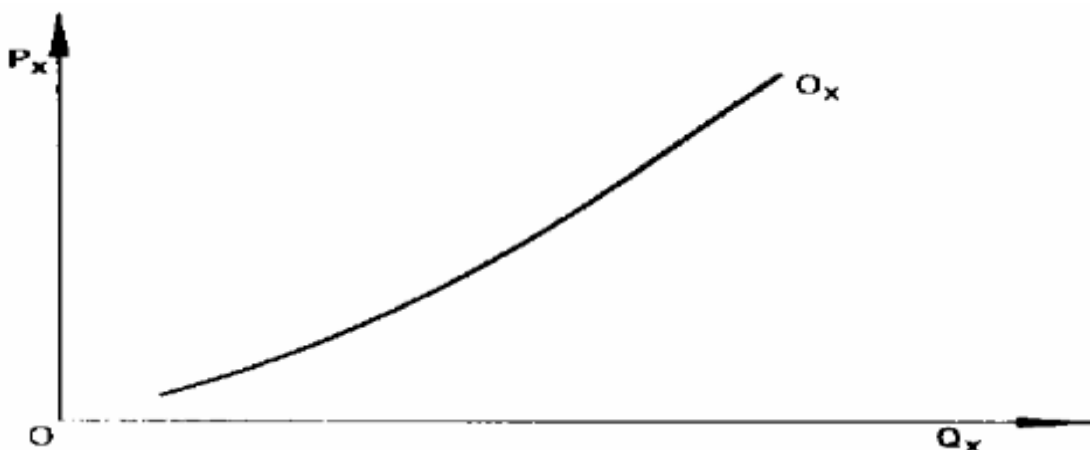
2.1. Teoria elementar da Oferta

Estipula-se oferta como a quantidade a ser disponibilizada de um bem ou serviço, no qual se deseja vender em determinada margem de tempo. Ou seja, a oferta é tratada pelos produtores como uma eventual quantidade no qual se acredita que há potencial total de venda. Já a demanda são os interessados neste bem ou serviço, e que varia de acordo com o tempo, gerando um fluxo. Existem diversos fatores que tornam possível realizar uma análise

minuciosa acerca da oferta, que são argumentos levantados por Pinho e Vasconcellos (2004) como:

- A oferta de um bem depende de seu próprio preço, admitindo a hipótese “*ceteris paribus*”, quanto maior for o preço de um bem, mais interessante se torna produzi-lo e, portanto, a oferta é maior. Relacionando a quantidade ofertada de um bem com seu preço obtemos a curva de oferta (Figura 1).

Figura 1 – Curva da Oferta



Fonte: adaptado de Pinho e Vasconcellos (2004)

- A oferta do bem x depende dos preços dos fatores de produção. De fato, o preço dos fatores, juntamente com a tecnologia empregada, determina o custo de produção. Havendo aumento do preço de fator, aumentaria o custo de produção. Os bens em cuja produção se emprega grandes quantidades deste fator sofrerão aumentos de custo significativos, enquanto aqueles que o empregam pouco sofrerão menos. Por exemplo: aumentando o preço da terra, teremos um grande aumento no custo de produção de café, enquanto em outros setores, que utilizam em menor intensidade o fator terra, teremos aumentos menores de custos. Desta forma, a mudança no preço de um fator acarretará alterações na lucratividade, relativa das produções, e isto ocasionará deslocamentos nas curvas de ofertas das diferentes mercadorias (o mesmo raciocínio se pode fazer em relação à mudança na tecnologia, de produção).
- A oferta de um bem pode ser alterada por mudança nos preços dos demais bens produzidos. Se os preços dos demais bens subirem e o preço do bem x permanecer idêntico, sua produção torna-se menos atraente em relação à produção dos outros bens, consequentemente diminuindo sua oferta. Neste caso, temos um deslocamento da curva de oferta para a esquerda. Podemos sintetizar estas relações matematicamente:

$$O_x = f(P_x, P_1 \dots P_{n-1}, \pi_1, \pi_2 \dots \pi_m, T)$$

Onde:

O_x = quantidade ofertada do bem x

$P_x = P$, o preço do bem x

P_i = o preço do bem i, $i = 1, 2 \dots n - 1$

$j \pi$ = o preço dos fatores de produção, $j = 1, 2 \dots m$

T = tecnologia

2.2. Econometria da Oferta

De acordo com Pinho e Vasconcellos (2004), O estudo do método estatístico é feito, didaticamente, por meio de duas categorias: a Estatística Matemática e a Estatística Aplicada. A primeira possui a importante função de fornecer uma justificativa formal dos processos de trabalho do método estatístico, ou seja, sob o aspecto da matemática pura, com resultados que sirvam de embasamento para a análise. Por sua vez, a Estatística Aplicada utiliza processos de análise estatística ao tratamento de dados de um determinado campo específico, analisando os mesmos para que a partir disso seja possível avaliar e tirar conclusões a respeito dos resultados.

Ainda segundo os autores, em cada campo de aplicação da Estatística, os fatos a ela pertinentes estão sujeitos a determinadas estruturas características e que são responsáveis pelos mesmos. Essas estruturas poderão ou não, teoricamente, ser ignoradas. Ignorando-as, elas se tornam observações mais vulneráveis, já que haverá a necessidade de manutenções da regularidade da mesma sob o aspecto da experiência cotidiana.

Com base nisso, para a realização do método estatístico sobre a análise dos dados numéricos de observação pode ser feita seguindo duas orientações: medida sem teoria; medida com teoria. Pinho e Vasconcellos (2004) afirma que o procedimento metodológico da “medida sem teoria” é orientada da seguinte maneira:

- a) Observação da realidade e constatação de possíveis estados de regularidade;
- b) Adoção da hipótese de que o estado de regularidade, empiricamente constatado na amostra, existe na população das observações possíveis;
- c) Descrição do estado de regularidade observada
- d) Adoção dessa descrição como válida para a população dos resultados possíveis.

Vale ressaltar que essa dedução é realizada por meio de raciocínio dos prováveis estados de regularidade que existem na população, e mesmo que isso possa trazer respostas que encaminhem para uma conclusão sobre a estrutura responsável pelas observações, é importante saber que essa análise é realizada sob o ponto de vista empírico e que muitas vezes não descreverá a realidade. Considera-se a condição de *ceteris paribus* (tudo o mais constante) para que se tenha a real dimensão de uma condição (como por exemplo: quantidade demandada e preço), mas que permanece frágil por ser empírica, podendo ser admitida apenas por um momento, até que se disponha de teorias explicativas e concretas.

Paralelamente, o este autor diz que, por sua vez, a “medida com teoria” baseia-se (como o próprio nome já sugeriu) na possibilidade de assegurar informações sob o ponto de vista explicativo, que comprove com resultados algo teórico, gerando um mecanismo no qual cada dado referido tenha um argumento, realizada pelos seguintes pontos:

- a) Considerar os dados estatísticos que se pretende analisar;
- b) Considerar uma teoria (conjunto de hipóteses) que se proponha a explicar o mecanismo responsável pelos dados estatísticos observados; e
- c) Verificar, por meio de processos estatísticos adequados, a diferença entre o comportamento apresentado pelos dados observados e o comportamento que esses mesmos dados deveriam apresentar se a teoria fosse adequada para descrevê-los.

De forma sucinta, o método econométrico pode ser dito como um método que analisa quantitativamente a economia, aliando a Teoria econômica com a Matemática, com o intuito de gerar uma linguagem interpretativa tanto para quantificar, tanto para modelar estatisticamente, permitindo assim a realização de uma análise com um olhar teórico, alinhando isso junto ao comportamento da realidade.

2.3. Regressão linear

De acordo com Sell (2005), a regressão linear é uma técnica de métodos quantitativos, utilizada para melhorar a capacidade de análise do comportamento dos custos e o aumento da objetividade e confiabilidade das informações contábeis, através da análise individual das variáveis que possam estar relacionadas com a composição do custo. Stevenson (1986), por sua vez, ressalta ser este método uma equação que descreve o relacionamento entre fatores em termos matemáticos. Ao mesmo tempo, para ele a regressão traduz a análise de dados amostrais, verificando se e como duas ou mais variáveis estão relacionadas numa determinada população, gerando como resultado uma equação matemática com o relacionamento descrito. Desse modo, Com o conhecimento de valores de outras variáveis, esta equação pode servir de auxílio para realizar estimativas de valores futuros. Nesse sentido, Sell (2005) ao abordar o método de regressão, evidencia as possibilidades existentes deste modelo, a saber:

- - A regressão Linear Simples: estabelece uma equação matemática linear que descreve o relacionamento entre duas variáveis, uma dependente e outra independente, com a finalidade de estimar valores para uma variável, com base em valores conhecidos da outra.
- - A regressão Linear Múltipla: envolve três ou mais variáveis, com uma variável dependente e duas ou mais variáveis independentes, com a finalidade de melhorar a capacidade de predição em confronto com a regressão linear simples.

Nesse sentido, percebe-se que a diferença entre uma e outra é somente a quantidade de variáveis a ser levado em consideração. A regressão linear tem como resultado a equação de uma reta, enquanto a regressão linear múltipla (tendo três variáveis) resultará num plano, porém se tiver “n” variáveis, resultará num hiperplano. Com a regressão é possível realizar diversas operações, que vão desde a estimativa de valores de uma variável, explicar estes valores, e a partir disso é possível estimar futuros valores de uma variável. Paralelamente, Sell (2005) acrescenta uma gama de informações a respeito do significado de cada estimativa gerada, tais como:

- Correlação - Literalmente significa co-relacionamento, evidencia até que ponto os valores de uma variável estão relacionados com os de outra.
- Regressão - Através de estimativas dos parâmetros, mostra o efeito da variável explicada X sobre a dependente Y, enquanto a correlação mede o grau de associação entre Y e a variável independente X.

De acordo com Matos (2000), Multicolinearidade - Refere-se à correlação entre duas variáveis explicativas ou entre uma delas e as demais, incluídas na equação de um modelo. Isso significa que esse fenômeno ocorre quando duas variáveis X1 e X2 tem aproximadamente a mesma medida, tornando assim a correlação entre eles próximo à perfeição.

- R² ajustado - Exibe o coeficiente de determinação múltipla, que é uma medida do grau de ajustamento da equação de regressão múltipla aos dados amostrais. Um ajuste perfeito resulta em R² = 1, um ajuste muito bom acarreta um valor próximo de 1 e um

ajuste fraco ocasiona um valor de R^2 próximo de zero. O coeficiente múltiplo de determinação de R^2 é uma medida de aderência da equação de regressão aos dados amostrais. O coeficiente de determinação ajustado é o coeficiente múltiplo de determinação R^2 modificado de modo a levar em conta o número de variáveis e o tamanho da amostra. O coeficiente de determinação ou de explicação R^2 mede a parcela da variação de Y explicada pela variação dos X.

- A variância residual, S^2 - Mede o grau de dispersão entre os valores observados e os estimados de Y, sendo que a raiz quadrada de S^2 é denominada erro-padrão da estimativa.
- Erro padrão - Mede a dispersão da estimativa do parâmetro, indicada pelos dados amostrais, podendo ser obtido pela raiz quadrada da variância residual. É uma espécie de desvio padrão que mede a dispersão em torno da reta de regressão.
- Estatística F - Utilizada para testar o efeito conjunto das variáveis independentes sobre a dependente, ou seja, serve para verificar se, pelo menos, um dos X explica a variação do Y. Desse modo, a hipótese nula (H_0) indicará que nenhum dos X afeta Y, enquanto a hipótese alternativa (H_1) assegura que, pelo menos uma das variáveis independentes influenciará a variável dependente Y, dentro dos graus de liberdade exigidos e dado um grau de significância. Considerando um nível de significância igual a 0,05, se F de significação for $< 0,05$, a regressão é significativa, mas se for $\geq 0,05$, a regressão não é significativa.
- Estatística T (para os coeficientes das variáveis independentes ou valor da prova) - Tem a finalidade de testar se o efeito de cada uma das variáveis independentes sobre a dependente é ou não estatisticamente significativo.
- Intervalos de confiança - É o intervalo dentro dos quais o valor verdadeiro do parâmetro populacional cairá, respeitando um determinado nível de confiança. Se a hipótese nula for rejeitada e o valor estimado de “b” for positivo, zero não pertencerá ao intervalo.
- O Valor P - Segundo Lapponi, (2000), é o maior valor de nível de significância para o qual o teste é significativo. Ou é o maior valor de nível de significância que rejeita a hipótese nula, H_0 . Valor P é o nível de significância observado. Para julgamento compara-se Valor P com o nível de significância ou erro tolerado que julgar mais adequado. O critério de decisão para o Valor P será: escolher o nível de significância α ; Se Valor P $< \alpha$, então, rejeitar a H_0 . É uma medida de significância global da equação de regressão múltipla e uma boa medida de aderência da equação aos dados amostrais.

Cabe aqui lembrar que a equação da reta, na regressão linear, é representada por: “ $y=a+b.x$ ”, onde “y”, representa a variável dependente, “a” o intercepto ou a interseção, “b” o coeficiente angular e o “x”, representa a variável independente. Ao mesmo tempo, é fundamental ressaltar o conceito de Teste de Hipótese. Ele é um processo capaz de afirmar, com base em dados amostrais, se uma hipótese sob prova é correta ou não. É uma afirmação que admite se certo efeito está presente ou não. Por hipótese, entende-se certa afirmação condicionada acerca de uma população, e classificam-se em dois tipos: a) hipótese nula ($H=0$), quando se admite não haver diferença entre a informação fornecida pela realidade e a afirmação da hipótese; b) hipótese alternativa ($H\neq 0$), quando se admite haver diferença entre a informação fornecida pela realidade e a afirmação da hipótese. Portanto, o processo de teste consiste em aceitar ou rejeitar a hipótese n

ula ($H=0$), com base na diferença entre o valor hipotético e seu estimado.

2.4. Histograma: teste de Jaque-Bera

Este teste serve para analisar se a série está distribuída normalmente ou não. Verifica-se a diferença entre a assimetria e a curtose da série junto a distribuição normal, baseado na fórmula a seguir:

$$Jarque - Bera = \frac{N - k}{6} \left(S^2 + \frac{(K - 3)^2}{4} \right)$$

Onde:

S = assimetria;

k = curtose;

K = número de coeficientes estimados para se criar a série.

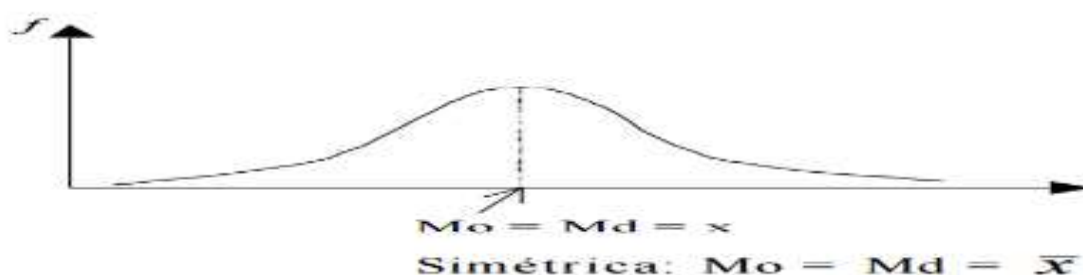
2.4.1. Assimetria

As medidas de assimetria indicam o grau de assimetria de uma distribuição de frequências unimodal em relação a uma linha vertical que passa por seu ponto mais elevado.

2.4.2. Distribuição Simétrica

Graficamente, uma distribuição simétrica tem associada a si uma curva de frequências unimodal apresentando duas "caudas" simétricas em relação a uma linha vertical que passa por seu ponto mais alto (eixo de simetria).

Figura 2 – Exemplo de Distribuição Simétrica

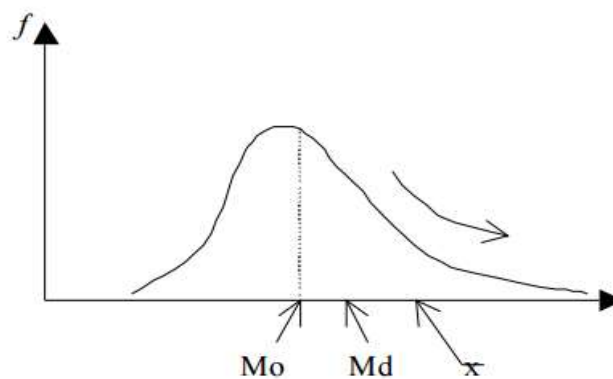


Fonte: Lociks (2001)

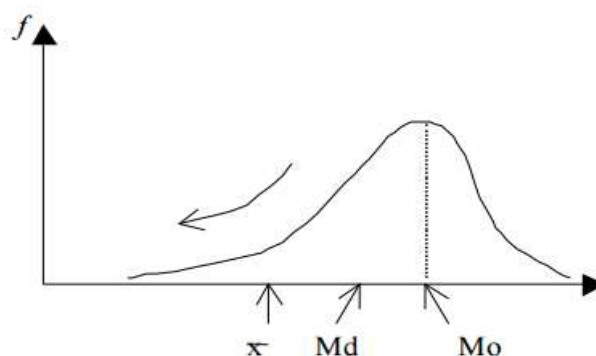
2.4.3. Distribuições Assimétricas

Uma distribuição assimétrica tem associada a si uma curva de frequências unimodal que apresenta, a partir do seu ponto mais alto, uma "cauda" mais longa para a direita (assimetria positiva) ou para a esquerda (assimetria negativa). Nas distribuições assimétricas os valores da moda, da mediana e da média divergem sendo que a média sempre estará do mesmo lado que a cauda mais longa.

Figura 3 – Exemplo de Distribuições Assimétricas



Assimétrica à direita (Positiva): \bar{x} à **direita** da Mo ($Mo < Md < \bar{x}$)



Assimétrica à esquerda (Negativa): \bar{x} à **esquerda** da Mo ($\bar{x} < Md < Mo$)

Fonte: Lociks (2001)

Com a assimetria e curtose elevadas, a curva em forma de sino que é característica peculiar de uma distribuição normal torna-se muito irregular, fazendo com que a expressão de Jarque-Bera tenha um valor maior que 6,0. Parâmetro esse fundamental para considerar, ou não, a distribuição normal de determinado conjunto de dados. Neste caso, se o valor for maior que 6,0, a hipótese de normalidade pode ser desconsiderada, já que o valor estará tendendo a não normalidade.

3. Metodologia e Análise Econométrica

3.1. Coleta de dados

O presente trabalho foi realizado através do auxílio do Banco de Dados Agregados disponibilizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Através deste meio, foi possível coletar os dados de quantidade produzida de goiaba ao longo de 21 (vinte e um) anos, assim como sobre o valor de receita com esta produção e a respeito da área colhida durante essa série histórica. Durante a realização do trabalho, optou-se por averiguar somente a região produtora de goiaba presente apenas no estado do Pará. Durante este estudo, foi constatada a hegemonia de apenas uma cidade como principal produtora do fruto em todo o estado, representando mais de 93% do total produzido na série histórica analisada, a cidade de Dom Eliseu (Figura 4).

Figura 4 – Principal município produtor de goiaba do Pará



Fonte: IBGE

Paralelamente, após uma análise preliminar dos dados, houve a necessidade de ajustá-los através do uso do logaritmo natural para que os resultados se tornassem mais confiáveis, tanto em relação ao gráfico de distribuição desses valores, como em relação às análises de correlação e regressão linear realizadas neste estudo. Posteriormente, através dos softwares: Excel 2007 e Eviews 3.0, foi possível realizar os testes de correlação, regressão linear, desenvolver relações estatísticas e encontrar as variáveis componentes para propor uma equação de Oferta de goiaba em função do tempo, através do método de regressão múltipla.

3.2. Modelo Teórico

O trabalho em questão buscou analisar a quantidade ofertada de goiaba em relação ao preço do fruto por tonelada e a área colhida de goiaba. Para isso, 21 (vinte e uma) amostras foram coletadas da série histórica de produção de goiaba, onde a variável dependente figura como a Quantidade Produzida de goiaba (QP) e as variáveis independentes são: a Área Colhida de goiaba (AC) e o Preço por Tonelada (PT). Desse modo, segundo esse direcionamento temos o seguinte modelo teórico:

$$QP = f(PT; AC)$$

(1)

Onde,

QP = Quantidade Produzida de goiaba

PT = Preço da goiaba por Tonelada

AC = Área Colhida de goiaba

2.3. Modelo Econométrico

Paralelamente, baseado no modelo teórico supracitado, um modelo econométrico foi proposto objetivando apresentar uma representação simplificação da situação real em análise. Para isso, se fez uso de um modelo de regressão linear múltiplo executado pelo software Excel 2007, pois o modelo em questão apresenta mais de uma variável independente. Ao mesmo tempo, para este estudo se optou pelo uso do Nível de Significância de 10% (dez por cento), assim como, segundo evidenciado durante a realização deste trabalho, houve a necessidade de adequar o modelo econométrico em questão para um modelo logarítmico, segundo a equação (2):

$$\ln QP_t = \alpha + \beta_1 \ln PT_t + \beta_2 \ln AC_t + \mu_t$$

(2)

Onde,

- $\ln QP_t$ = representa a Variável Dependente em log natural (Quantidade Produzida de goiaba) em função do tempo;
- α = Representa o Intercepto;
- $\beta_1 \ln PT_t$ = representa a Variável independente em log natural (Preço por Tonelada) em função do tempo;
- $\beta_2 \ln AC_t$ = representa a Variável independente em log natural (Área Colhida de goiaba) em função do tempo;
- μ_t = Termo de erro.

3.4. Hipóteses Estabelecidas

Para a realização do presente estudo, algumas hipóteses foram estabelecidas a fim de analisar a precisão do modelo proposto, a saber:

H0: $\alpha = 0$ (hipótese nula), de que o intercepto é zero, ou seja, quando as variáveis preço e área destinada à colheita forem iguais à zero, não haverá quantidade produzida de goiaba;

Ha: $\alpha > 0$ (hipótese alternativa), independente do preço e área destinada à colheita haverá oferta de goiaba;

H0: $\beta_1 = 0$ (hipótese nula), de que a variável preço não influencia na oferta de goiaba.

Ha: $\beta_1 > 0$ (hipótese alternativa) de que a variável preço influencia na oferta de goiaba.

H0: $\beta_2 = 0$ (hipótese nula), de que a variável área destinada à colheita não influencia na oferta de goiaba.

Ha: $\beta_2 > 0$ (hipótese alternativa) de que a variável área destinada à colheita influencia na oferta de goiaba.

4. Discussão dos resultados

O presente trabalho foi baseado em dados provenientes do Banco de Dados Agregados do IBGE. Nele, foram coletados os dados referentes à produção por tonelada de goiaba no estado do Pará (Quadro 1), assim como sobre a área colhida destinada a produção do fruto no estado. Paralelamente, através dos valores de receita disponibilizados pelo Instituto citado, foi possível estimar o preço por tonelada de goiaba através da razão entre a quantidade produzida e a receita. Desse modo, contando com um total de 21 (vinte e uma) observações, os dados abrangem desde o período de estabelecimento do Real como moeda corrente na federação (1994) até o dado mais recente a respeito dessa produção, ou seja, 2014.

Quadro 1 – Série histórica de produção de Goiaba no estado do Pará

Ano	Quant. Prod Goiaba (ton)	Preço/Tonelada	Área colhida (Hect)
1994	0	R\$ 0,00	0
1995	160	R\$ 50,00	2
1996	160	R\$ 62,50	2
1997	618	R\$ 17,80	13

1998	0	R\$	0,00	0
1999	8756	R\$	82,34	107
2000	8520	R\$	157,16	230
2001	1050	R\$	118,10	95
2002	2725	R\$	173,58	109
2003	2500	R\$	450,00	100
2004	1377	R\$	310,09	270
2005	3640	R\$	409,89	325
2006	4153	R\$	508,31	376
2007	5092	R\$	476,04	376
2008	18672	R\$	537,86	1263
2009	13844	R\$	524,70	944
2010	20182	R\$	621,59	1381
2011	5552	R\$	791,25	499
2012	6462	R\$	679,20	365
2013	2944	R\$	880,10	173
2014	2862	R\$	919,64	177

Fonte: IBGE

Todavia, após um estudo preliminar dos dados através de um estudo de correlação e regressão linear, os parâmetros selecionados apresentaram resultados insatisfatórios para uma análise acurada. Por esse motivo, se optou por adequá-los através do uso do Log Natural em todas as variáveis conforme exposto no Quadro 2.

Quadro 2 – Ajuste dos dados através do Log Natural				
Ano	Quant. Prod Goiaba (ton)	preço/Tonelada	Área colhida (Hect)	
1994	0,00	R\$	0,00	0,00
1995	5,08	R\$	3,91	0,69
1996	5,08	R\$	4,14	0,69
1997	6,43	R\$	2,88	2,56

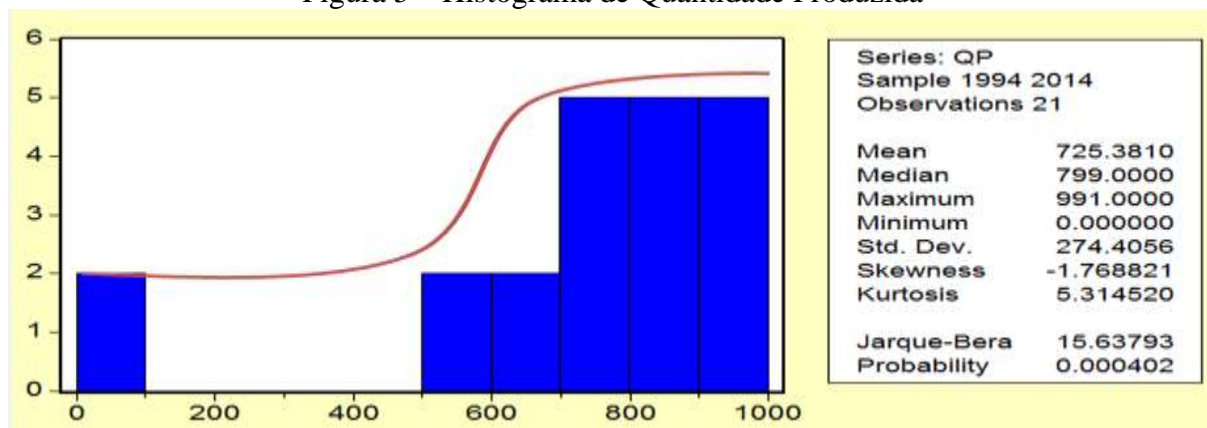
1998	0,00	R\$	0,00	0,00
1999	9,08	R\$	4,41	4,67
2000	9,05	R\$	5,06	5,44
2001	6,96	R\$	4,77	4,55
2002	7,91	R\$	5,16	4,69
2003	7,82	R\$	6,11	4,61
2004	7,23	R\$	5,74	5,60
2005	8,20	R\$	6,02	5,78
2006	8,33	R\$	6,23	5,93
2007	8,54	R\$	6,17	5,93
2008	9,83	R\$	6,29	7,14
2009	9,54	R\$	6,26	6,85
2010	9,91	R\$	6,43	7,23
2011	8,62	R\$	6,67	6,21
2012	8,77	R\$	6,52	5,90
2013	7,99	R\$	6,78	5,15
2014	7,96	R\$	6,82	5,18

Fonte: Autores (2015)

4.1. Análise dos histogramas dos dados e análise de sua normalidade

Objetivando verificar se os dados analisados apresentam um comportamento probabilístico semelhante a uma distribuição normal, houve a necessidade de plotar os valores referentes à: Quantidade Produzida de goiaba (QP), Preço por Tonelada (PT) e Área Colhida (AC). Ao mesmo tempo, a análise do método de *Jarque-Bera* foi realizada com a mesma finalidade.

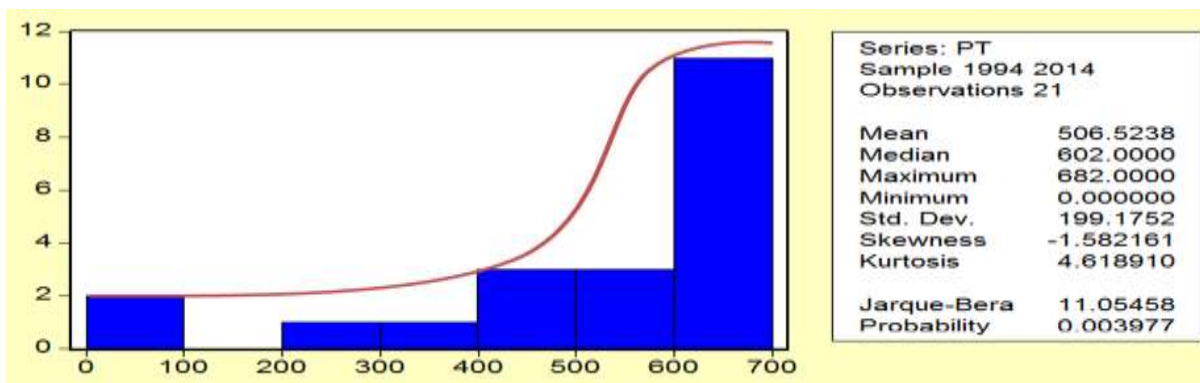
Figura 5 – Histograma de Quantidade Produzida



Fonte: Software Eviews 3.0

Conforme observado na Figura 5, o gráfico da Quantidade Produzida apresenta assimetria, com um Coeficiente de Simetria (*Shewness*) de -1.768821, um Coeficiente de Achatamento (índice de *Kurtosis*) de 5.314520 e valor de probabilidade de *Jarque-Bera* maior do que 6, indicando, desse modo, ser possível rejeitar a hipótese de distribuição normal.

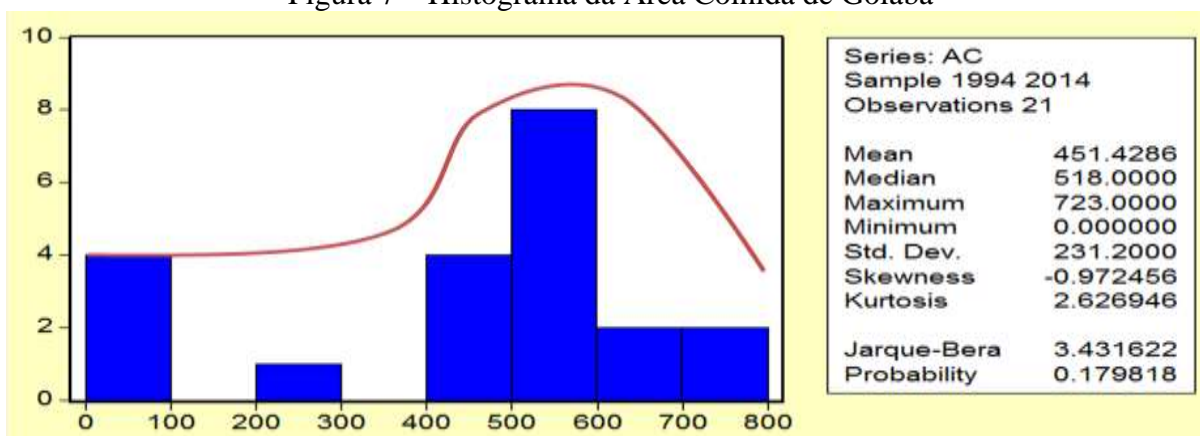
Figura 6 – Histograma de Preço por Tonelada



Fonte: Software Eviews 3.0

O mesmo ocorre para a análise do Preço por Tonelada (PT), Figura 6, neste caso o gráfico também apresenta assimetria, com um Coeficiente de Simetria (*Shewness*) de -1.582161, um Coeficiente de Achatamento (índice de *Kurtosis*) de 4.618910 e valor de probabilidade de *Jarque-Bera* maior do que 6, indicando, desse modo, ser possível rejeitar a hipótese de distribuição normal.

Figura 7 – Histograma da Área Colhida de Goiaba



Fonte: Software Eviews 3.0

Todavia, a análise gráfica da Área Colhida (Figura 7) apresentou resultados destoantes com as demais variáveis. Com um Coeficiente de Simetria (*Shewness*) de -0.972456, assim como um Coeficiente de Achatamento (índice de *Kurtosis*) de 2.626946, o gráfico da área colhida apresentou um valor de *Jarque-Bera* dentro dos padrões aceitáveis para uma distribuição normal, desse modo à hipótese de normalidade da curva não pode ser rejeitada.

4.2. Análise de Correção entre as variáveis propostas

Após a nova adequação dos dados através do uso do logaritmo natural, houve uma nova análise de correlação entre as variáveis selecionadas, segundo evidenciado no Quadro 3.

Quadro 3 – Análise de Correlação entre as variáveis

	Q.P.	P.T.	A.C.
Q.P.	1		
P.T.	0,908561801	1	

A.C.	0,906868372	0,866708378	1
------	-------------	-------------	---

Fonte: Software Excel 2007

Conforme exposto no quadro supracitado, as Variáveis selecionadas esta a análise são:

- QP: Quantidade Produzida de Goiaba (em toneladas)
- PT: Preço por Tonelada produzida
- AC: Área Colhida referente à produção de goiaba (em Hectare)

Segundo o exposto, a Matriz de Correlação evidenciou uma forte correlação entre as variáveis: Quantidade Produzida de goiaba (QP) e a Área Cultivada (AC), pois o resultado deste estudo apresentou valor próximo de 1 (um). Desse modo, indicando a não rejeição da hipótese alternativa para essas variáveis. Ao mesmo tempo, o mesmo se observou entre as variáveis: Quantidade Produzida (QP) e Preço por Tonelada (PT); assim como para: Área Colhida (AC) e Preço por Tonelada (PT), indicando também a não rejeição das hipóteses alternativas para essas análises. Portanto, o alto grau de correlação entre as variáveis propiciou a continuidade do estudo através da análise de Regressão Linear.

4.3. Análise do Coeficiente de Determinação

Após o estudo da correlação entre as variáveis ajustadas através do logaritmo natural, o estudo em questão analisou relevância do Coeficiente de Determinação, também denominado R-quadrado, com o objetivo de identificar o quanto o modelo proposto consegue explicar os valores observados na análise. Através do uso da técnica da Regressão Linear, utilizando QP (Quantidade Produzida de goiaba) como variável dependente, assim como AC (Área Colhida) e PT (Preço por Tonelada) como variáveis independentes, foi possível obter os resultados presentes no Quadro 4. Ao mesmo tempo, foi estabelecido como Intervalo de Confiança o valor de 90% (noventa por cento) e, conseqüentemente, apresentando o valor de $\alpha=0,1$.

Quadro 4 – Análise de Regressão	
<i>Estatística de regressão</i>	
R múltiplo	0,939569606
R-Quadrado	0,882791044
R-quadrado ajustado	0,869767826
Erro padrão	0,990488249
Observações	21

Fonte: Software Excel 2007

Após a determinação desta medida, foi possível constatar que o modelo proposto neste estudo consegue explicar aproximadamente 88% (oitenta e oito por cento) dos valores observados para a Quantidade Produzida (QP) de goiaba em relação ao Preço por Tonelada (PT) e a Área Colhida (AC) sendo, portanto, considerado como um parâmetro aceitável para esta análise.

4.4. Análise do modelo e interpretação estatística

Paralelamente à análise do Coeficiente de Determinação e ainda dentro da análise de Regressão Linear, outra medida foi avaliada foi o “F de Significação”. Desse modo, conforme exposto no Referencial Teórico, é aconselhável não rejeitar a hipótese alternativa, já que o

valor de F de Significação é acentuadamente menor do que o valor proposto de alta, ou seja, $\alpha=0.1$, indicando, deste modo, que a regressão é boa.

Quadro 5 – Análise do F de Significância

ANOVA					
	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	2	133,0050957	66,50254785	67,785941	4,17491E-09
Resíduo	18	17,65920549	0,981066972		
Total	20	150,6643012			

Fonte: Software Excel 2007

Dando continuidade ao estudo, a medida Valor – P foi analisada objetivando identificar os valores correspondentes aos coeficientes angulares da equação da oferta proposta.

Quadro 6 – Análise do Valor - P

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>% superior</i>	<i>inferior 90,0</i>	<i>Superior 90,0%</i>
Interseção	1,242341598	0,613348503	2,025506857	0,057903	-0,04625579	2,530939	0,178756	2,305926914
preço/Tonelada	0,678750876	0,222897459	3,045126125	0,006966	0,210460691	1,147041	0,292233	1,065269247
Área colhida (Hect)	0,569853254	0,192090217	2,966591748	0,008264	0,166286682	0,97342	0,236757	0,902949909

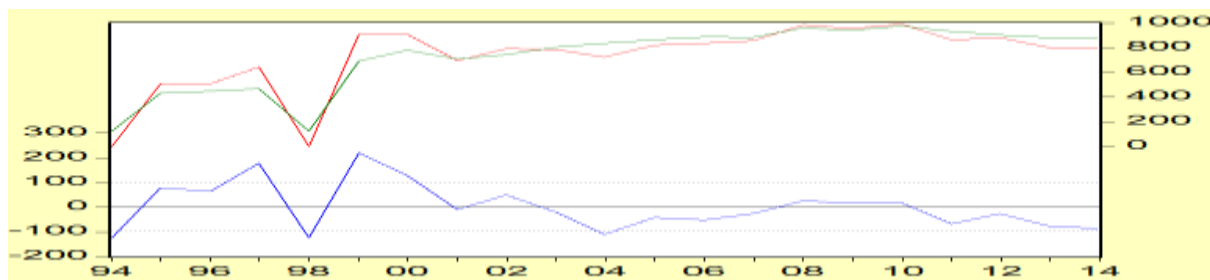
Fonte: Software Excel 2007

Desse modo, conforme evidenciado no Quadro 6, os coeficientes angulares para as variáveis PT (Preço por Tonelada) e AC (Área Colhida) são aceitáveis, pois apresentam valor inferior ao de alfa, viabilizando, deste modo, a proposição da Equação da Oferta de goiaba em relação à Área Colhida e ao Preço por Tonelada.

4.5. Análise do gráfico de Resíduos

Após a realização dos estudos de normalidade dos dados analisados, correlação entre as variáveis propostas no modelo econométrico e regressão linear tendo como variável dependente a Quantidade Produzida de goiaba (QP), assim como a Área Colhida (AC) e o Preço por Tonelada (PT) como variáveis independentes, foi gerado um gráfico de resíduo para medir a conformidade do modelo proposto. Desse modo, segundo evidenciado na Figura 8, o resultando gráfico apresentou conformidade dentro dos parâmetros máximo e mínimo estabelecidos no modelo.

Figura 8 - Gráfico de Resíduos



Fonte: Software Eviews 3.0

4.6. Proposição da Equação da Oferta para a produção de goiaba

Finalmente, após as análises de supracitadas (normalidade dos dados, correlação entre as variáveis, regressão linear e análise das medidas: Valor P e F de significância) foi possível estabelecer a equação da oferta de goiaba em função do tempo, segundo a Equação (3):

$$QP(t) = 1,242341598 + 0,678750876*PT(t) + 0,569853254*AC(t) \quad (3)$$

Onde,

- QP: Quantidade Produzida de Goiaba no tempo (t)
- PT: Preço por Tonelada no tempo (t)
- AC: Área Colhida no tempo (t)
- t: tempo

5. Considerações Finais

O trabalho em questão, após a realização dos procedimentos citados no tópico Discussão dos Resultados, culminou com a proposição de um modelo de regressão linear múltipla para a oferta de goiaba no estado do Pará. Este modelo abrange a interação entre a Quantidade Produzida de goiaba em determinado período de tempo como variável dependente, assim como a Área Colhida de goiaba e o Preço por Tonelada de goiaba como variáveis independentes. Ao mesmo tempo, ambas as variáveis desse modelo foram validadas pelo método de correlação, obtendo valores próximos do valor ideal, indicando forte correlação entre elas. Ao mesmo tempo, este modelo econométrico conta com um nível de segurança de 10% (dez por cento), estado todas as suas constantes condizentes com esta métrica adotada durante a realização da regressão linear.

Ao mesmo tempo, após a obtenção da equação de oferta de goiaba no estado do Pará, foi possível analisar as hipóteses propostas no item “Hipóteses Estabelecidas”. Segundo esta análise, a hipótese referente ao intercepto (α) deve ser considerada a não rejeição da hipótese alternativa. Ao mesmo tempo, quanto a β_1 e β_2 , a não rejeição das hipóteses alternativas também devem ser consideradas. Nesse sentido, se tem o modelo de oferta de quantidade produzida de goiaba em relação ao preço por tonelada e a área colhida.

Referências

- GONZAGA NETO, L.; Soares, J.M. (1994). **Goiaba para exportação**: aspectos técnicos da produção. Brasília: Embrapa-SPI. p. 49. (Série Publicações Técnicas FRUPEX, 5).
- IBGE. Banco de Dados Agregados. **Sistema IBGE de Recuperação Automática - SIDRA**. Disponível em: <http://www.ibge.gov.br>. Acesso em: 12 nov. 2015.
- KWEE, L.T.; Chong, K.K. (1990). **Botany and Cultivars**. In: Guava in Malaysia – Production, pests and diseases. Tropical Press: Kuala Lumpur. p. 21-51.
- LAPPONI, J. C. (2000). **Estatística Usando Excel 5 e 7**. São Paulo: Lapponi, Treinamento e Editora, p. 294.
- MATOS, O. C. (2000). **Econometria Básica**. São Paulo: Atlas, p. 124.
- MEDINA, J. C. (1988). **Goiaba**: cultura, matéria-prima, processamento e aspectos econômicos. 2.ed. rev. ampl. Campinas. p. 1-120. (ITAL. Série Frutas Tropicais, 6).
- MORTON, J. (1987). **Guava**. In: Fruits of warm climates. Julia F. Morton, Miami. pp. 356–363.
- PINHO, Diva Benevides; VASCONCELLOS, Marco Antonio S. de. **MANUAL DE ECONOMIA**. 5. ed. São Paulo: Saraiva, 2004.
- SELL, Isair. **Utilização da regressão linear como ferramenta de decisão na gestão de custos**. 2005. IX Congresso Internacional de Custos - Florianópolis, SC, Brasil - 28 a 30 de novembro de 2005. Disponível em: <<http://www.ime.usp.br/~salles/fatec/estatistica/trabalho/Utilização da regressão linear como ferramenta de decisão na gestão de custos.pdf>>. Acesso em: 16 out. 2015.
- STEVENSON, W. J. (1986). **Estatística aplicada à administração**. São Paulo: Harbra, p. 341.