

Análise Estatística para a produção de milho entre os anos de 1962 e 2010

Rafaela da Silva Gomes¹, Vitória Régia Lopes dos Santos², Luciene Resende Gonçalves³

1. Introdução

O milho é produzido em diferentes regiões brasileiras, com condições ambientais e socioeconômicas diversas. Ele é cultivado em quase todo o território nacional. Em torno de 77% da área plantada e 92% da produção encontra-se nas regiões Sul, Sudeste e Centro-Oeste, sendo que a região Sul participou com 42% da área e 53% da produção; Sudeste com 19% da área e 19% da produção e Centro-Oeste com 15% da área e 19% da produção (DUARTE; MATTOSO; GARCIA, 2003)

A produção brasileira de milho tem apresentado tendência de elevação desde o fim da década de 80 em função do desenvolvimento técnico, mas também de diversos fatores micros e macroeconômicos que propiciaram o crescimento da produção. Assim, gostaríamos de testar a influência de algumas variáveis econômicas (preço do produto, PIB, inflação) e área cultivada no nível de produção do milho, através da técnica de regressão linear múltipla.

2. Referencial teórico

De grande importância econômica, o grão é o principal macro ingrediente para a produção de rações de suínos e aves. Desse modo, além do mercado agrícola, o milho também tem um grande peso na competitividade do mercado de carnes brasileiro (CRUZ, 2008). Segundo Pereira filho e Borgi (2017):

A semente de milho é, dentre todas as tecnologias agrícolas hoje empregadas no cultivo, a que mais se desenvolveu nos últimos tempos. Com os avanços da biotecnologia, houve grande impacto para a produção de milho no Brasil o que permite a sua flexibilidade de cultivo nas mais diferentes regiões e condições de clima e solo. Graças aos avanços genéticos, temos disponíveis cultivares mais responsivas, a incorporação de características como resistência a doenças, insetos, assim como às moléculas de herbicidas como glifosato e glifosinato para o controle eficiente de plantas daninhas.

Ainda o milho vem sendo utilizado na indústria de alta tecnologia, como por exemplo na fabricação de biocombustíveis como o etanol. A produção de milho-etanol é eficiente em termos energéticos, na medida em que produz 34% mais energia do que a necessária para o produzir, inclusive cultivar o milho, colhê-lo, transportá-lo e destilá-lo em etanol. O crescimento na produção de etanol proporcionou um estímulo econômico para a agricultura, porque a maior parte do etanol é feito de milho (SHAPOURI; DUFFIELD; WANG, 2002).

Com tudo isso, o milho tem alta capacidade de movimentar o mercado de trabalho e influenciar o crescimento do PIB brasileiro. De acordo com censo agropecuário do IBGE (2017), 1.6 milhões de estabelecimentos agropecuários cultivam o milho em grão, sendo produzidas 91 milhões de toneladas. Em 2017 o PIB brasileiro cresceu em 1,0% puxado pelo crescimento da participação

¹ UNIFAL, ICSA, Discente. email: rafadasilva43@yahoo.com.br.

² UNIFAL, ICSA, Discente. email: vitoria.regia83@yahoo.com.br.

³ UNIFAL, ICSA, Docente. email: luliresende.lr@gmail.com.

agropecuária, principalmente em vista do crescimento da produção de milho cresceu em 55,2%. Quanto a geração de empregos, o setor agropecuário corresponde a cerca de 20% do emprego nacional, sendo que o milho figura entre as culturas que apresentam maior impacto no número de empregos no agronegócio (COSTA; GUILHOTO; IMORI, 2013).

3. Materiais e Métodos

Na pesquisa foi aplicado uma análise, composta pela regressão linear múltipla, essa técnica possibilita a avaliação da influência simultânea das variáveis independentes e permite analisar as hipóteses definidas para o estudo. A regressão linear múltipla é uma técnica estatística que tenta resolver o problema análogo à regressão simples que no caso seria analisar a relação entre uma variável dependente e mais de uma variável independente (HAIR ET AL., 2005). Neste trabalho utilizamos variáveis independentes com medidas em unidades diferentes tornando difícil mostrar a sua importância na pesquisa. Dessa forma foi necessário aplicar a base logarítmica para se ter uma interpretação das estimativas. Uma medida utilizada para saber o melhor ajustamento é o R^2 .

Os dados utilizados na pesquisa foram retirados do Ipea data (2019) que é uma base de dados macroeconômicos, financeiros e regionais do Brasil mantida pelo Ipea. Com acesso gratuito, oferece também catálogo de séries e fontes, dicionário de conceitos econômicos, histórico das alterações da moeda nacional e dicas sobre métodos e fontes utilizadas. As séries são mostradas desde o nível municipal e, em alguns casos, podem ser agregadas para Áreas Mínimas Comparáveis (AMC), recurso que permite realizar comparações intertemporais consistentes para municípios que mudaram de área ou contorno ao longo do tempo.

A variável dependente utilizada foi a produção de milho por tonelada anual no período entre 1962 e 2010, as variáveis independentes utilizadas nesse trabalho foram área colhida de milho por hectares, preços anuais do milho atualizados conforme o ano de 2010 usando o deflator, índice de inflação anual o IGP-DI e por fim o PIB do Brasil anual em dólar na qual esses dados foram retirados do The World Bank. Os anos escolhidos foram entre 1962 e 2010, em razão da disponibilidade pelo Ipea Data que só apresentava esses anos não sendo possível obter dados atualizados. Houve a possibilidade de encontrar dados mais atualizados em outros sites, no entanto as fontes eram diferentes das utilizadas pelo Ipea data. O objetivo do trabalho consiste em descobrir a influência das variáveis independentes na produção do milho.

As variáveis presentes neste estudo são as seguintes:

- Produção do milho: A produção brasileira de milho é a terceira maior do mundo, com incrementos de produtividade nas últimas 10 safras de 6,8% ao ano. A área de milho cultivado em 2^a safra representa 62% da área e 65% da produção nacional. Essa evolução é possível devido à introdução de tecnologias de produção em clima tropical. O milho é a base energética da alimentação de aves, suínos e bovinos, além de ser destinado à alimentação humana. O milho tem papel fundamental para a nossa rotação de culturas, pois produz uma grande quantidade de palha que auxilia na proteção do solo, na reciclagem de nutrientes e no incremento de matéria orgânica no solo;
- Área colhida do milho por hectares: é a capacidade da área em relação a colheita medida em hectares;
- Preços anuais do milho: é uma estimativa dos preços da produção do milho medidas anualmente em reais e atualizadas conforme o ano de 2010, ou seja, foi utilizado o deflator para que os preços tenham médias mais aproximadas.

- IGP-DI anual: foi instituído em 1944 com a finalidade de medir o comportamento de preços em geral da economia brasileira.
- PIB do Brasil: O produto interno bruto do Brasil é o principal indicador do tamanho da economia do país. Corresponde à soma (em valores monetários) de todos os bens e serviços finais produzidos internamente, em determinado período (normalmente 1 ano). Nesse caso usamos o PIB em dólar para não usar o deflator, pois em dólar a inflação já foi corrigida.

Nesse trabalho foi retirado a área colhida de milho por hectares, pois não apresentou significância suficiente e teve problemas de colinearidade o que não auxiliou na construção do modelo.

A partir das variáveis usadas no modelo foram levantadas as seguintes hipótese:

H1: Quando os preços anuais do milho aumentam, a produção aumenta, pois quando algo afeta o preço e a quantidade de equilíbrio utilizamos a teoria da demanda e oferta. Então quanto maior os preços mais os produtores tendem a ofertar o seu produto, o que faz com que a produção aumente em vista do aumento da demanda. É possível que a produção seja influenciada diretamente pelos preços, por causa da lei da oferta e da demanda.

H2: Quando a inflação estiver alta isto fará com que a produção diminua, produzir ficará mais caro, e com a alta dos preços isso influenciará toda a produtividade do milho. A inflação influencia diretamente os preços e isso pode atrapalhar os produtores, pois mesmo que as pessoas estejam consumindo mais, a produção do milho poderá diminuir, porque para produzir precisa-se de insumos e se eles estiverem caros a produção diminuirá, os produtores não acharão propício aumentar a produção nesse ambiente econômico.

H3: Quando o PIB anual do Brasil estiver em alta, haverá um aumento da produção do milho, em razão do PIB ser um dos indicadores mais utilizados na macroeconomia, e tem o objetivo principal de mensurar a atividade econômica de uma região. Na contagem do PIB, considera-se apenas bens e serviços finais, excluindo da conta todos os bens de consumo intermediários, ou seja, se o PIB estiver em alta as pessoas conseguirão ter mais poder de compra, a situação econômica do país vai estar melhor, este cenário pode influenciar na produção de milho do país, pois os produtores poderão ver este cenário otimista e querer produzir mais.

Para corroborar com a pesquisa e mostrar a necessidade de transformação dos dados para a base *log* será apresentado alguns gráficos de dispersão:

Figura 1- Gráfico de dispersão Prod X Preço

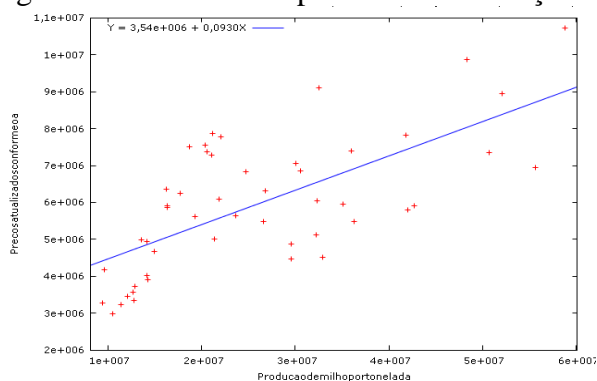


FIGURA 1- Gráfico de dispersão Prod X Preço
Fonte: Elaboração Própria do autor

Figura 2 - Gráfico de dispersão Prod X Área

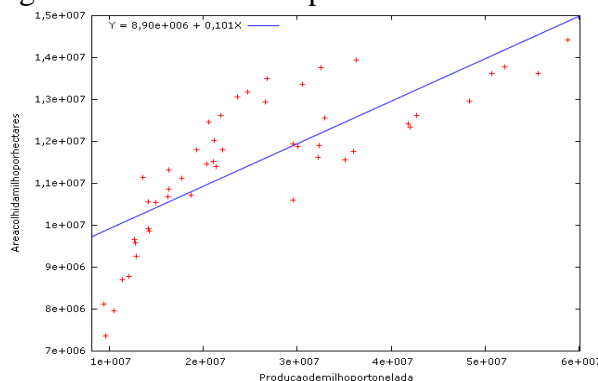


FIGURA 2- Gráfico de dispersão Prod X Área
Fonte: Elaboração Própria do autor

Figura 3 - Gráfico de dispersão Prod X Infla

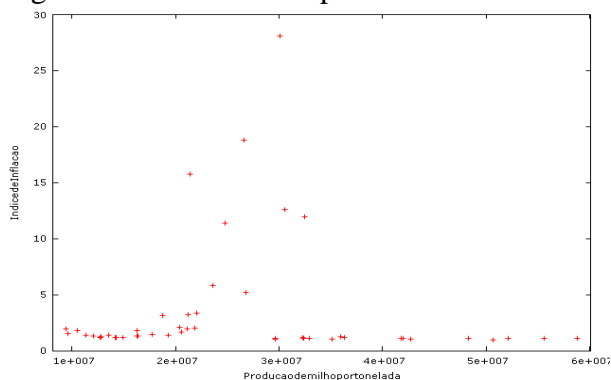
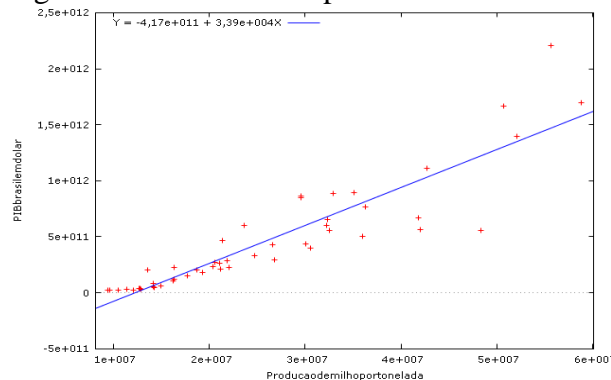
FIGURA 3- Gráfico de dispersão Prod X Infla
Fonte: Elaboração Própria do autor

Figura 4 - Gráfico de dispersão Prod X PIB

FIGURA 4- Gráfico de dispersão Prod X PIB
Fonte: Elaboração Própria do autor

As Figuras 1, 2 e 4 na qual são apresentadas o gráfico de dispersão, mostra uma relação com a variável dependente, sendo possível traçar uma reta. Já na Figura 3 é possível notar que o gráfico mostra uma dispersão que não é possível traçar uma reta, isto ocorre, pois, as escalas numéricas são diferentes, pois a inflação está em índice e a produção em toneladas, fora a grande discrepância entre estas duas variáveis por isso é necessário a logaritimização.

Para melhorar o entendimento sobre as variáveis utilizadas na pesquisa, na Tabela 1 é apresentado as estatísticas descritivas das variáveis.

Tabela – 1 Estatísticas Descritivas das Variáveis

Variáveis Estáticas Descritivas	Área colhida de Milho	Produção de milho	Preços atualizados do milho	Índice de inflação (IGP- DI)	PIB do Brasil em Dólar
Média	1,152e+007	2,584e+007	5,943e+006	3,459	4,592e+011
Desvio Padrão	1,685e+006	1,297e+007	1,791e+006	5,395	4,863e+011

Fonte: Elaboração própria do autor

A partir da Tabela 1 é possível perceber as discrepâncias de valores entre as variáveis. Com as informações das estatísticas descritivas é possível analisar que o índice de IGP-DI é o que apresenta a média menor e o desvio padrão também. E o que apresenta a maior média e desvio padrão é o PIB, pois é uma variável macroeconômica que avalia todo produto interno do bruto do país, por isso apresenta os maiores valores. Já as outras variáveis não apresentam tanta discrepância entre si, o que a demonstrando que a produção de milho é bastante controlada.

4. Resultado e discussão

Os dados após tabulados foram exportados para o *Gretl* (2019), e então aplicamos a logaritimização para que os valores resultantes não fossem discrepantes e facilitasse a interpretação das estimativas.

Utilizou-se do método de Mínimos Quadrados Ordinários. Na Tabela 2 estão dispostos os coeficientes, o erro padrão, a razão-t e o p valor associado às estimativas do modelo.

Tabela 2 - Resultados do Modelo de Regressão Linear Múltipla.

Variável dependente: log - Produção de Milho por Tonelada				
Variáveis Independentes (Todas Logaritimizadas)	Coeficiente	Erro Padrão	Razão-t	P-valor

Constante	5,37870	1,33341	4,034	0,0002	***
Preço do milho anual	0,202686	0,111403	1,819	0,0755	*
Índice de Inflação anual	-0,0523507	0,0274862	-1,905	0,0632	*
PIB do Brasil em dólar anual	0,322632	0,0256797	12,56	<0,0001	***
R-quadrado	0,898002				

*Nível significância entre 5% e 10%, ** Nível significância entre 1% e 5% e *** Nível significância entre 0% e 1%.

Fonte: Produção própria do autor.

O R^2 é uma medida descritiva da qualidade do ajuste obtido. Em geral referimo-nos ao R^2 como a quantidade de variabilidade nos dados que é explicada pelo modelo de regressão ajustado. Entretanto, o valor do coeficiente de determinação depende do número de observações (n), tendendo a crescer quando n diminui. O valor apresentado por ele foi de 0,898002, indicando que, 89,8002% das variações da “Produção de milho” são explicadas em média pelas variáveis independentes.

Para a análise dos parâmetros foi utilizado a relação *Ceteris paribus* uma predição ou uma afirmação sobre uma relação causal, empírica ou lógica indutiva entre dois estados de coisas para fazer uma análise das variáveis estudadas a influência de um fator sobre outro, sem que as demais variáveis sofram alterações. Assim mesmo as variáveis sendo transformadas para a base *log* elas não sofreram alterações e sim foi uma tentativa de deixar os dados menos discrepantes entre si. Pois como foi aplicado *log* nas variáveis dependentes e independentes, os fatores continuam constantes, ou seja, a interpretação pode ser feita.

Portanto as análises dos parâmetros observadas:

Preços anuais do milho: O coeficiente positivo apresentado mostra uma relação direta com a variável dependente, onde, por exemplo, se a variável “Preço anuais do Milho” aumentar em 100% a variável dependente aumenta em 20,27%, em média, o que está em conformidade com H1. O p-valor do teste t apresentou um valor de 0,0755, significativo a 10%.

IGP-DI anual: O coeficiente negativo apresentado mostra uma relação decrescente com a variável dependente, onde, por exemplo, se a variável “IGP-DI anual” aumentar em 100% a variável dependente diminui em 5,23507%, em média, o que está em conformidade com H2. O p-valor do teste t apresentou um valor de 0,0632 com uma confiabilidade necessariamente positiva, pois está abaixo de 10%.

PIB do Brasil: O coeficiente positivo apresentado mostra uma relação direta com a variável dependente, onde, por exemplo, se a variável “PIB do Brasil” aumentar em 100% a variável dependente aumenta em 32,2632%, em média, em conformidade com H3. O p-valor do teste t apresentou um valor de <0,0001, está abaixo de 1%.

Os pressupostos utilizados na regressão linear para conferir a robustez do modelo assim como seus testes e resultados, são apresentados a seguir:

O teste de White que confere se há “Heterocedasticidade”, apresentou um p-valor de 0,476258, ou seja, a hipótese nula de que os dados não apresentam heterocedasticidade foi aceita, pois o valor foi acima de 5%.

O teste de “Normalidade dos Resíduos” apresentou um p-valor de 0,37834, ou seja, a hipótese nula de que os dados e o erro têm distribuição normal foi aceita, pois o p-valor ficou acima de 5%.

Para o teste de “Colinearidade”, os resultados dos valores normais são aceitos se menores que 10, para que eles não sejam multicolineares, neste caso, os dados obtidos através do VIF foram: 2,137 para Preço do milho anual; 1,054 para IGP-DI; 2,060 para PIB do Brasil em dólar anual.

A partir destes resultados observou-se a falta de multicolinearidade e a relação positiva com a variável dependente “Produção de milho anual”.

Após executar a regressão, notamos que todas as hipóteses formuladas foram aceitas, comprovando que as variáveis macroeconômicas escolhidas para esta pesquisa tiveram uma relação condizente com a literatura comprovado com o modelo escolhido.

5. Conclusão

Os resultados obtidos com o modelo foi que ele apresentou uma robustez bem significativa, a produção de milho é bem controlada para que não haja desperdícios e nem prejuízo aos produtores. Os resultados obtidos ficaram em conformidade com as hipóteses formuladas, ou seja, o que se obteve já condizia com que era esperado.

Este trabalho poderá ser utilizado para pesquisas futuros ou para que os produtores de milho possam observar sempre os períodos propícios para a produção, já que foi possível notar que todas as variáveis utilizadas tiveram um nível de significância abaixo de 10% PIB anual do Brasil apresentou o maior nível de significância entre as variáveis escolhidas.

Um desafio para este trabalho foi a falta de dados atualizados. Pretendemos diante da disponibilidade estimar um novo modelo e comparar para verificar se ainda as variáveis utilizadas na pesquisa tiveram a mesma influência ou semelhante.

6. Referências bibliográficas

COSTA, C. C. da; GUILHOTO, J. J. M.; IMORI, D. Importância dos setores agroindustriais na geração de renda e emprego para a economia brasileira. **Rev. Econ. Sociol. Rural**, Brasília, v. 51, n. 4, p. 787-814, 2013.

CRUZ, José Carlos et al. **A cultura do milho**. Sete Lagoas: Embrapa Milho e Sorgo, 2008.

DUARTE, J. O.; MATTOSO, M. J.; GARCIA, J. C. Importância socioeconômica. In: CRUZ, J. C. (Ed.). **Milho Sete Lagoas**: Embrapa Milho e Sorgo, 2003. (Árvore do Conhecimento)

IBGE. **Censo agropecuário 2017**. Rio de Janeiro, 2017. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/3093/agro_2017_resultados_preliminares.pdf>. Acesso em: 18 de março de 2019.

HAIR, J. F., Jr.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. **Análise multivariada de dados**. 5ª ed. Porto Alegre: Bookman, 2005.

PEREIRA FILHO, Israel Alexandre; BORGHI, Emerson. Mercado de sementes de milho no Brasil: safra 2016/2017. **Sete Lagoas: Embrapa Milho e Sorgo**, 2016.

SHAPOURI, H.; DUFFIELD, J. A.; WANG, M. Q. The energy balance of corn ethanol: an update. 2002.