

Uso da Análise de Correspondência Múltipla na identificação de fatores associados ao retorno econômico na atividade leiteira no Estado de Minas Gerais, Brasil

Waldomiro Barioni Júnior^{1†}, Claudia De Mori², Artur C. Camargo³, André Luiz M. Novo⁴, Marcela M. B. Vinholis⁵

¹Embrapa Pecuária Sudeste.

²Embrapa Pecuária Sudeste. E-mail: claudia.de-mori@embrapa.br.

³Embrapa Pecuária Sudeste. E-mail: artur.camargo@embrapa.br.

⁴Embrapa Pecuária Sudeste. E-mail: andre.novo@embrapa.br.

⁵Embrapa Pecuária Sudeste. E-mail: marcela.vinholis@embrapa.br.

Resumo: O estudo aplicou a Técnica Multivariada de Análise de Correspondência Múltipla para avaliar a existência de associação entre o retorno econômico da atividade leiteira e os indicadores de produtividade, composição de rebanho, custo de produção, receita bruta e assistência técnica. Os dados são provenientes de um questionário estruturado aplicado em 374 propriedades leiteiras de Minas Gerais, Brasil. Concluiu-se que o retorno econômico superior na produção de leite, expresso pela margem bruta por área de terra, está associada a maior produção de leite por vaca, maior número de vacas por área de terra e maior percentual de vacas no rebanho, que por sua vez está associado ao maior tempo de assistência técnica recebida pelo produtor rural. Ainda que o custo de produção operacional por animal tenha sido superior quando houve aumento de produtividade, a receita bruta por animal suplantou o custo, resultando em margem bruta superior. A assistência técnica do produtor rural é um mecanismo importante para capacitar o produtor rural e levar conhecimento sobre novas tecnologias, o que resulta em ganhos de produtividade no campo. A técnica de Análise de Correspondência Múltipla mostrou-se adequada para aplicação e tomada de decisão em análises do setor agropecuário.

Palavras-chave: Análise multivariada; Produção de leite; Retorno financeiro.

Abstract: The study applied the Multiple Correspondence Analysis Technique in order to evaluate the association among the economic return of dairy production, yield, herd composition, production cost, gross income and rural technical assistance. A survey questionnaire was used to get information on 374 dairy farms in Minas Gerais, Brazil, which comprise the cross section dataset. It is concluded that the greater economic return on dairy production, expressed by the gross margin per hectare, is associated to greater milk production per cow greater number of cows per hectare and to higher percentage of cows in the herd, which in turn is associated with longer rural assistance. Although the greater production cost per animal when the yield increases, the gross revenue per animal surpass the cost, resulting in a greater gross margin. The rural assistance is an important mechanism to capacity the farmer and to deliver knowledge about new technologies, which results in productivity gains. The Multiple Correspondence Analysis technique is adequate for application and decision-making in analyzes of the agricultural sector.

Keywords: Multivariate analysis; Dairy production; Economic return.

Introdução

Com um faturamento de R\$ 68,7 bilhões, segundo dados da Associação Brasileira da Indústria de Alimentos (2019), a indústria láctea ocupou a segunda posição no ranking de faturamento dos principais setores alimentícios do país em 2018. A produção brasileira foi de 8,90 bilhões de litros de leite em 2017, é a quinta maior produção mundial envolvendo 1,1 milhão de propriedades, (ANUÁRIO LEITE 2018, 2018).

[†]Autor correspondente: waldomiro.barioni@embrapa.br.

Dentre os estados brasileiros, Minas Gerais destaca-se na produção de leite do país, com maior volume de produção e captação. Em 2017, o Estado foi responsável por 25,5%⁶ da produção nacional e 26,3%⁶ da captação anual dos laticínios registrados (6,5 bilhões de litros), contando com um contingente de 223 mil produtores de leite em atividade (20,3% do total do Brasil⁶) e um rebanho de 5,8 milhões de vacas (31,2% do rebanho nacional de vacas[‡]). O segmento de recebimento e processamento do Estado possui 771 laticínios com SIF e um total de 1,2 milhão de trabalhadores envolvidos com a atividade leiteira gerando um valor bruto de produção de leite em 2017 de R\$10,1 bilhões (ANUÁRIO LEITE 2018, 2018).

Segundo Lopes e Carvalho (2000), a análise econômica da atividade leiteira é importante, pois possibilita detectar os pontos críticos do sistema produtivo, buscando soluções para aumentar os lucros ou diminuir os custos de produção. A margem bruta por unidade de área, diferença entre a receita obtida pela venda do produto e os custos operacionais despendidos para sua produção, é uma medida importante para verificar o retorno econômico da atividade no curto prazo assim como avaliar a competitividade da atividade leiteira com as demais opções do uso da terra. É condicionada por diferentes fatores econômicos e zootécnicos, tais como a receita bruta por vaca, o custeio por vaca, o percentual de vacas no rebanho, a produção média de leite das vacas em lactação, dentre outros indicadores. Identificar os fatores diretamente associados à margem bruta pode auxiliar na orientação dos produtores na melhoria da gestão de seus sistemas produtivos.

Este estudo teve como objetivo a aplicação da Técnica Multivariada de Análise de Correspondência Múltipla para avaliar a existência de associações entre o retorno econômico por unidade de área (margem bruta/ha) da atividade leiteira e os indicadores zootécnico-econômicos. Os dados são provenientes de um questionário estruturado aplicado em propriedades leiteiras de Minas Gerais que integram o Programa Balde Cheio da Embrapa.[§]

Material e Métodos

O conjunto de dados utilizados na análise refere-se a 374 propriedades mineiras, distribuídas em 139 municípios do estado, acompanhadas pelo Programa Balde Cheio⁷, compreendendo o período de janeiro a dezembro de 2017. Os dados são gerados por uma planilha de acompanhamento mensal de indicadores econômico-zootécnico e compreende variáveis de natureza quantitativa relativas à área, rebanho, produção leiteira, qualidade de leite, produtividade, despesas com custeio e com investimentos, receita total, margem bruta, dentre outros, totalizando 37 variáveis.

A área das propriedades variou de 1,1 a 770,6 ha e quase $\frac{3}{4}$ das propriedades (72,2%) são caracterizadas como sendo de pequeno porte por possuir menos de quatro módulos rurais. Em termos de rebanho, as propriedades possuíam de 6,0 a 390,0 vacas, com média de 38,9 vacas/ propriedade. A produção de leite diária média foi de 394,7 litros/propriedade/dia, oscilando entre 27,2 a 4.091,0 litros/propriedade/dia. A observação destes valores ratifica a aplicação da ACM, na medida em que permite contemplar valores extremos, realmente observados, na execução da análise. Considerando a amplitude e valores extremos observados nas coletas de campo, o uso de Técnica de Análise de Correspondência Múltipla (ACM), tratando os dados de forma categórica, se mostra oportuna para ampliar o nível de significância dos resultados, ainda que haja alguma redução na precisão. A ACM

[‡] Percentuais calculados pelos autores com base nos valores nominais apresentados por FAEMG, 2018.

[§] O Programa Balde Cheio, criado pela Embrapa Pecuária Sudeste em 1998, tem por objetivo promover transferência de tecnologias de produção e gerenciais aos técnicos e produtores de leite através da capacitação e do acompanhamento das propriedades. Em 2017, o Programa atuou em 229 municípios de 10 estados brasileiros, com atendimento a 911 propriedades rurais, envolvendo 97 técnicos e 82 instituições parceiras.

é uma técnica estatística exploratória utilizada para verificar associações ou similaridades entre variáveis qualitativas ou contínuas categorizadas. Isto é, esta técnica tem a função de estudar as correspondências entre as categorias das variáveis. As posições das categorias de cada variável podem ser interpretadas como associações, no plano multidimensional, por representação gráfica (BENZÉCRI, 1992; BARIONI, 1995; GREENACRE, 2007).

Neste trabalho, utilizou-se uma matriz de dez variáveis de natureza sócioeconômica e zootécnica. O retorno econômico, variável resposta, foi definida pelo indicador Margem Bruta por área (**MB_AREA**), expresso em R\$/ha. Os dados zootécnicos e econômicos, variáveis explicativas, usados na ACM foram: percentual de vacas no rebanho (**VR**, expresso em %); percentual de vacas em lactação no rebanho (**VLR**, expresso em %); produção diária média de leite das vacas em lactação (**M_VL**, expressa em L/vaca/dia); produção diária média de leite das vacas do rebanho (**M_VR**, expressa em L/vaca/dia); número de vacas em lactação por área (**VL_AREA**, expresso em unidade/ha); receita bruta por vaca por ano (**REC_VAC**, expresso em R\$/vaca); custeio por vaca por ano (**COE_VAC**, expresso em R\$/vaca/ano); margem bruta por vaca por ano (**MB_VAC**, expresso em R\$/vaca/ano); e tempo de assistência técnica recebida por meio da participação no Programa Balde Cheio (**TBC**, expresso em meses).

Desta forma, utilizou-se uma matriz de dados reais dada por X_{ij} ($i=374$ propriedades de leite e $j=10$ variáveis quantitativas). As variáveis foram categorizadas em três classes de intensidade: 1-baixa, 2-média e 3-alta. Exceto a variável tempo no Balde Cheio contendo 5 classes. Foram geradas tabelas de contingência simples para cada variável; Teste de Qui-quadrado e Coeficiente de Contingência de Pearson entre a variável resposta **MB_AREA** e as demais variáveis explicativas; e finalmente a realização da Análise de Correspondência Múltipla (ACM) com representação dos resultados em forma de mapa bidimensional (DIM1 e DIM2), envolvendo todas as variáveis. Os pacotes estatísticos usados foram o SAS (2012) e o STATISTICA (2004).

O conceito básico da ACM esta na estatística de Qui-quadrado (χ^2), expressa pela formula a seguir:

$$\chi^2 = \sum_{ij} \frac{(n_{ij} - e_{ij})^2}{e_{ij}} = \sum_{ij} \frac{(n_{ij} - n_{i.}n_{.j}/n)^2}{n_{i.}n_{.j}/n}, \text{ com } (I-1)(J-1) \text{ graus de liberdade,}$$

Onde: $e_{ij} = n_{i.}n_{.j}/n$ é o valor "esperado" na cela (i,j) da matriz baseada sobre as marginais linhas e colunas, $n_{i.}$ e $n_{.j}$ de uma tabela de contingencia, respectivamente. O Nível de significância adotado foi de 10%, por considerar um estudo observacional.

Resultados e Discussão

Na Tabela 1 estão descritas as 10 variáveis com as suas respectivas classes de intensidade (1:baixa, 2:média e 3:alta) e as respectivas frequências absolutas (N) e relativas (%), provenientes dos 374 produtores entrevistados, bem como as probabilidades geradas pelo Teste de Hipótese de Qui-quadrado (χ^2) e Coeficiente de Contingência (CC) entre o cruzamento da variável resposta (**MB_AREA**) e as variáveis explicativas.

O Teste Qui-quadrado mostrou que há dependência ($p \leq 10\%$) entre a variável resposta Margem Bruta por área (**MB_AREA**) e todas as variáveis explicativas. O CC mostrou forte associação de **MB_AREA** com **VL_AREA** e **MB_VAC**; moderada associação com **REC_VAC**, **VLR**, **MVL** e **M_VR**, e; fraca associação com **VR**, **COE_VAC** e **TBC**.

Tabela 1. Descrição das variáveis usadas na Análise de Correspondência Múltipla (ACM), com as respectivas categorias, frequências absolutas (N) e relativas (%) de um total de 374 produtores, as probabilidades do Teste de Qui-quadrado (χ^2) e o Coeficiente de Contingência (CC).

Variável (unidade)	Sigla	Categoria	N	%	p-value ¹ (CC) ²
Porcentagem de vacas no rebanho (%)	VR1	0 <= PVR <= 45	107	28,61	0,001 (0,272)
	VR2	45 < VR <= 60	185	49,47	
	VR3	VR > 60	82	21,93	
Porcentagem de vacas em lactação no rebanho (%)	VLR1	0 <= VLR <= 33	105	28,07	0,001 (0,314)
	VLR2	33 < VLR <= 45	167	44,65	
	VLR3	VLR > 45	102	27,27	
Produção média diária de leite das vacas em lactação (L/vaca/dia)	MVL1	0 <= MVL <= 10	109	29,14	0,001 (0,357)
	MVL2	10 < MVL <= 15	188	50,27	
	MVL3	MVL > 15	77	20,59	
Produção média diária de leite das vacas do rebanho (L/vaca/dia)	M_VR1	0 <= M_VR <= 8	133	35,56	0,001 (0,366)
	M_VR2	8 < M_VR <= 11	132	35,29	
	M_VR3	M_VR > 11	109	29,14	
Número de vacas em lactação por área (unidade/ha)	VL_AREA1	0 <= VL_AREA <= 0,5	115	30,75	0,001 (0,589)
	VL_AREA2	0,5 < VL_AREA <= 1,0	141	37,70	
	VL_AREA3	VL_AREA > 1,0	118	31,55	
Receita por vaca por ano (R\$/vaca/ano)	REC_VAC1	0 <= REC_VAC <= 10	131	35,03	0,001 (0,377)
	REC_VAC2	10 < REC_VAC <= 15	155	41,44	
	REC_VAC3	REC_VAC > 15	88	23,53	
Custeio por vaca por ano (R\$/vaca/ano)	COE_VAC1	0 <= COE_VAC <= 6	105	28,07	0,0165 (0,177)
	COE_VAC2	6 < COE_VAC <= 10	157	41,98	
	COE_VAC3	COE_VAC > 10	112	29,95	
Margem bruta por vaca por ano (R\$/vaca/ano)	MB_VAC1	MB_VAC <= 2,5	118	31,55	0,001 (0,605)
	MB_VAC2	2,5 < MB_VAC <= 5,0	151	40,37	
	MB_VAC3	MB_VAC > 5,0	105	28,07	
Tempo que participa do Balde Cheio (meses)	TBC1	0 < TBC <= 12	31	8,38	0,0634 (0,196)
	TBC2	12 < TBC <= 24	98	26,49	
	TBC3	24 < TBC <= 36	92	24,86	
	TBC4	36 < TBC <= 48	48	12,97	
	TBC5	TBC > 48	101	27,30	
Margem Bruta por área (R\$/ha) - Variável resposta	MB_AREA1	MB_AREA <= 700	119	31,82	-
	MB_AREA2	700 < MB_AREA <= 1600	11	31,55	
	MB_AREA3	MB_AREA > 1600	137	36,63	

¹P-value do teste de hipótese de Qui-quadrado (p=0,10).

²(CC: Coeficiente de Contingência de Pearson): indica a força da dependência entre duas variáveis qualitativas e é dado por $CC = \sqrt{(\chi^2)/(\chi^2+n)}$.

Critério de interpretação: CC < 0,1: associação muito fraca; CC = 0,1 a < 0,3: associação fraca CC = 0,3 a 0,5: associação moderada e CC > 0,5: associação forte. Fonte: <http://www.acastat.com/statbook/chisqassoc.htm>

Considerando todas as variáveis da Tabela 1 como “input” para a ACM e analisando a visualização gráfica representada pelo plano bidimensional (DIM1 e DIM2) gerado pela ACM com explicação de 32,56% da inércia total dos dados (Figura 1), observa-se que há uma associação direta,

Sigmae, Alfenas, v.8, n.2, p. 636-641, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18º Simpósio de Estatística Aplicada à Experimentação Agrônoma (SEAGRO).

seguindo uma trajetória crescente da esquerda para a direita do mapa (seta verde), entre as categorias de baixa (1), média (2) e alta intensidade (3), marcadas em cinza (Zona1, Zona2 e Zona3). A margem bruta baixa (MB_AREA1, \leq R\$700,00/ha), na esquerda do gráfico, Zona1, está associada às categorias de baixa intensidade das variáveis explicativas. No centro do mapa, Zona2, está localizada a margem bruta intermediária (MB_AREA2, entre R\$700,00/ha e R\$1600,00/ha), também associada às categorias de média intensidade para as demais variáveis. No lado direito do mapa, Zona3, encontra-se a margem bruta superior (MB_AREA3, $>$ R\$1600,00/ha), associada às categorias de maior intensidade de todas as nove variáveis explicativas (VR3, VLR3, ..., MB_VA3). Isto é, o rendimento econômico, representado pela Margem Bruta (MB_AREA), cresce da esquerda para direita (DIM1). Ainda que o custo de produção operacional por vaca tenha sido superior na Zona 3 do mapa (COE_VAC3), a maior produção média diária de leite por vaca (MVL3 e M_VR3) e a alta receita bruta por vaca (REC_VAC3), resultaram em uma maior margem bruta (MB_AREA3). Ressalta-se que o produtor rural que recebe assistência técnica há algum tempo no Programa Balde Cheio (TBC1 a TBC4) já parte de um retorno econômico intermediário (MB_AREA2, R\$700,00 a R\$1600,00/ha) e produtores com mais tempo no programa (TBC5, $>$ 48 meses) estão associados a maiores ganhos de produtividade (MVL3; VL_AREA3) e retorno econômico superior (MB_AREA3, $>$ R\$1600,00/ha). A variável TBC foi considerada como suplementar na ACM.

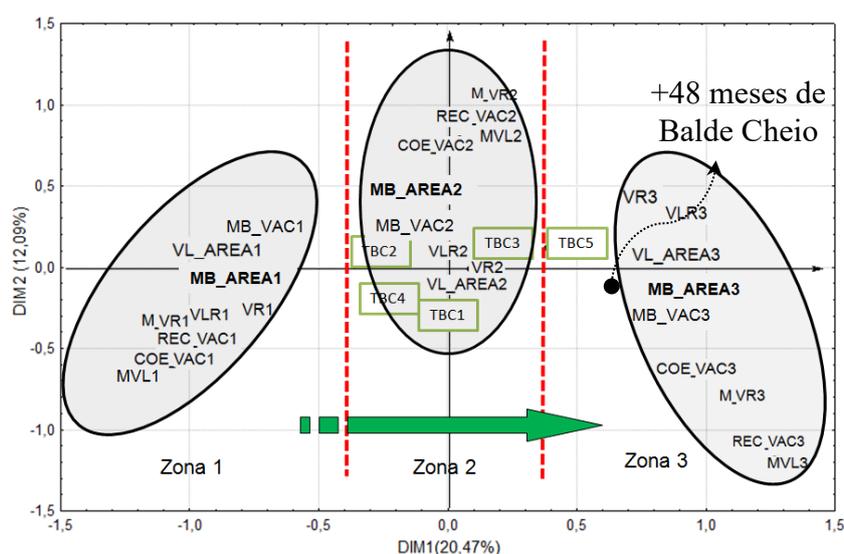


Figura 1. Mapa gerado pela ACM representando um plano bidimensional, formado pelas duas primeiras dimensões, caracterizando o perfil dos produtores em baixa, média e alta eficiência na atividade leiteira, considerando os fatores econômicos e zootécnicos em função do tempo no programa Balde Cheio (variável suplementar).

Conclusões

Os resultados da Análise de Correspondência Múltipla mostram que o indicador de retorno econômico margem bruta e as variáveis econômicas e zootécnicas selecionadas estão associados.

O retorno econômico superior na produção de leite, expresso pela margem bruta por área de terra, está associada a maior produção leiteira por vaca, maior número de vacas por área de terra e maior percentual de vacas no rebanho, que por sua vez está associado com maior tempo de capacitação e assistência técnica no programa Balde Cheio. Este é um mecanismo importante para o

fomento da adoção de novas tecnologias e ganhos de produtividade no campo, o que resulta em retornos financeiros positivos para o produtor rural.

Em termos de aplicação da técnica, a Análise de Correspondência Múltipla permitiu contemplar valores extremos que são observados nos sistemas reais demonstrando-se adequada para aplicação em casos de produção agropecuária.

Agradecimentos

Federação da Agricultura e Pecuária do Estado de Minas Gerais (FAEMG), técnicos e agricultores do Programa Balde Cheio.

Referências Bibliográficas

ASSOCIAÇÃO BRASILEIRA DA INDÚSTRIA DE ALIMENTOS. *Números do setor: Faturamento*. São Paulo: ABIA. 2019. Disponível em: https://www.abia.org.br/vsn/tmp_6.aspx?id=16 . Acesso em: 07 mar. 2019.

ANUÁRIO LEITE 2018: *Indicadores, tendências e oportunidades para quem vive no setor leiteiro*. São Paulo: Texto Comunicação Corporativa, 2018. 114 p.

BARIONI Jr., W. *Análise de correspondência na identificação dos fatores de risco associados à diarreia e à performance de leitões na fase de lactação*. Piracicaba, Escola Superior de Agricultura Luiz de Queiroz, 1995. 97p. (Tese de Mestrado).

BENZÉCRI, J. P. *Correspondence analysis handbook*. New York: Marcel Decker. 1992. 665p. FAEMG.

GREENACRE, M. *Correspondence Analysis in Practice*. Boca Raton: Chapman & Hall/CRC. 2 ed, 2007. 296p.

LOPES, M.A.; CARVALHO, F.M. *Custo de produção do leite*. Boletim Agropecuário, n.32, 2000. 42 p.

SAS INSTITUTE INC. *SAS System for Microsoft Windows*. Release 9.3. Cary: SAS Institute Inc. 2010.

STATSOFT INC. *Statistica: data analysis software system*. Version 7. Tulsa: StatSoft, Inc. 2004.. Disponível em <http://www.statsoft.com>.

Sigmae, Alfenas, v.8, n.2, p. 636-641, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).
18º Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).