

# SOBRE UM PROCEDIMENTO EM AMOSTRAGEM PARA CORRIGIR A PRODUÇÃO DE LEITE BOVINO

ANTONIO FRANCISCO IEMMA\*, HUMBERTO DE CAMPOS\*\* E DÉCIO BARBIN\*\*\*

## RESUMO

*Técnica das estimativas por regressão na amostragem por conglomerado em estágio duplo, para corrigir a produção de leite bovino. O procedimento em questão mostrou vantagens sobre os procedimentos convencionais, mormente quanto à determinação dos intervalos de confiança para as produções corrigidas.*

## 1 – INTRODUÇÃO

De modo geral, a variação de caracteres quantitativos é devida, segundo LOBO<sup>(9)</sup>, a componentes genéticos, ambientais e à interação entre eles. Com o propósito de seleção e melhoramento genético animal, torna-se necessário “eliminar” os vieses, oriundos da parcela ambiental para que a contribuição genética seja aquilatada com maior precisão.

As causas ambientais citadas mais amiúde como responsáveis pela variação da produção leiteira são: idade da vaca, mês e ano da parição, número de ordenhas diárias, duração da lactação, intervalos entre partos, binômio preço da ração-preço do leite, dentre outras.

A eliminação do erro inerente às causas ambientais, tem sido processada desde GOWEN<sup>(4)</sup>, com base na Teoria dos Fatores de Correção. Esses fatores, que geralmente o Zootecnista encontra sob formas de tabelas, são de dois tipos:

– Fatores Multiplicativos, como os de GOWEN<sup>(4)</sup>, JOHANSSON<sup>(8)</sup>, MILLER<sup>(10)</sup>, IEMMA<sup>(5)</sup>, entre outros.

– Fatores Aditivos, como os de SANDERS<sup>(14)</sup>, SEARLE<sup>(15)</sup>, IEMMA<sup>(6)</sup>, etc.

Apesar de universalmente adotados e de sua reconhecida eficiência, os fatores de correção apresentam o inconveniente de fornecerem apenas uma estimativa por ponto; da produção corrigida, sem possibilitar o conheci-

mento de suas variâncias e intervalos de confiança. Este fato, que aos leigos pode parecer irrelevante, causa sérios transtornos aos zootecnistas no momento das tomadas de decisão, dada a grande variabilidade comprovadamente existente na produção leiteira.

Assim, este estudo teve por objetivo a obtenção de um processo que possibilitasse a correção das produções, segundo um número razoável de causas ambientais, de modo que além das estimativas por ponto fossem também obtidas suas variâncias e intervalos de confiança.

## 2 – MATERIAL

Para a realização deste trabalho, contou-se com os registros de produções de vacas holandesas malhadas de preto, de rebanhos paulistas, retiradas dos arquivos da Associação Brasileira de Criadores de Bovinos da Raça Holandesa, sediada na cidade de São Paulo.

Foram anotados para cada animal: número de registro na ABCBRH, produção total de leite e gordura, idade em anos e meses, ordem da lactação, número de ordenhas diárias, mês e ano da parição, duração da lactação em dias.

O processamento dos dados foi executado no IBM 1130 da ESALQ/USP.

## 3 – MÉTODOS

### 3.1 – Causas Ambientais

A opção pelas causas ambientais

aqui abordadas, foi feita por adoção das citações de SEARLE & HENDERSON<sup>(16)</sup> que atuaram em rebanhos holandeses dos Estados Unidos, Canadá e Nova Zelândia; JOHANSSON<sup>(8)</sup> que atuou em rebanhos da Holanda e países baixos; e NAUFEL<sup>(11)</sup> que realizou profundos estudos em rebanhos holandeses do Brasil.

Assim, o elenco das causas ambientais ficou constituído de: idade da parição, duração da lactação, número de ordenhas diárias, mês e ano da parição.

### 3.2 – Procedimento Amostral

Dada a impossibilidade de se obter uma listagem (rol) de todas as produções leiteiras existentes, de rebanhos paulistas, na ABCBRH, optou-se pela Amostragem por Conglomerados, onde os rebanhos foram tomados como unidades primárias e os animais como unidades secundárias.

Assim, adotou-se por população o contingente de 40 rebanhos que mais efetivamente movimentaram os arquivos da ABCBRH por um período de 10 anos: 1967 a 1976. Com base nessa população procedeu-se ao primeiro estágio da amostragem, sorteando-se uma amostra de 4 desses rebanhos, equivalendo a 10% das U.P. da população.

De cada Unidade Primária (Rebanhos), foram retiradas 180 Unidades Secundárias (Animais), distribuídos por 5 estratos (idades) com um número fixo de 36 U.S. por estrato, como segue

\*Professor Assistente Doutor no Instituto Básico de Biologia Médica e Agrícola da UNESP, Campus de Botucatu.

\*\*Professor Titular do Departamento de Matemática e Estatística da ESALQ - USP, Piracicaba.

\*\*\*Professor Adjunto do Departamento de Matemática e Estatística da ESALQ - USP, Piracicaba.

1o. Estágio		2o. Estágio					m <sub>i</sub> constante	Fator de Sub Amostragem
UP	Mi	U.S. Por Idade (Anos)						
		2	3	4	5	6		
1	1031	36	36	36	36	36	180	17,46
2	951	36	36	36	36	36	180	18,93
3	666	36	36	36	36	36	180	27,03
4	1025	36	36	36	36	36	180	17,56
M = 3673		144	144	144	144	144	m = 720	

Fixou-se o número de 36 unidades secundárias por estrato, para que as estimativas por regressão linear fossem sempre obtidas através do mesmo número de pares, como descrito a seguir:

3.3 – Estimativas das Produções Adultas, a Nível de Estratos

Para a obtenção das produções corrigidas (estimativas das produções em idade adulta) cada grupo de 180 animais, dos respectivos rebanhos, foi estratificado em 5 idades de parição: 2, 3, 4, 5 e 6 anos. As quatro primeiras representaram as idades jovens e a idade de 6 anos, a de produção máxima ou idade adulta. Conforme citado em (3.2), à cada idade foram atribuídas as produções de 36 animais.

Dada a grande variabilidade existente na produção leiteira, mesmo dentro dos estratos Rebanho-idade, houve-se por bem não tomar os 36 animais inteiramente ao acaso, mas sim distribuí-los durante os 12 meses do ano.

Com raciocínio análogo considerou-se que essa variação seria ainda muito grande mesmo para a intersecção rebanho-idade-mês da parição e desse modo como a cada mês corresponderam 3 animais, eles foram sorteados do período de 10 anos citado anteriormente. Aqui sim, a casualização se fez presente.

Desse modo, cada estrato (idade) ficou constituído de 12 parcelas, sendo cada uma a média das produções dos 3 animais obtidas ao acaso, dos anos de 1967 a 1976.

As estimativas por regressão linear, das produções em idade adulta sobre as produções em idades jovens foram obtidas em cada estrato, com base nas médias mensais, de acordo com os procedimentos adaptados de SAMPFORD(13) e COCHRAN(3).

$$\hat{Y}(RL)_{ih} = \bar{y}_i + \hat{b}_{ih}(\bar{X}_{ih} - \bar{x}_{ih})$$

i = 1, 2, 3, 4 rebanhos  
h = 1, 2, 3, 4 idades jovens

onde

$\hat{Y}(RL)_{ih}$  é a produção média em idade adulta, estimada por regressão linear para animais (U.S.) do rebanho i, (u.P.) que pariram na idade (estrato) L.

$\bar{y}_i$  é a estimativa da produção média em idade adulta, obtida de animais da unidade primária i, aos 6 anos.

$\hat{b}_{ih}$  é a estimativa do coeficiente de regressão linear entre as produções médias mensais da idade h e da idade adulta, no rebanho i.

$\bar{X}_{ih}$  é a produção média populacional da unidade primária i, na idade h.

$\bar{x}_{ih}$  é a produção média estimada da unidade secundária em idade h, obtida da sub amostra da unidade primária i.

Os valores de  $\hat{b}_{ih}$  foram obtidos com o usualmente, por

$$\hat{b}_{ih} = \frac{\sum_{p=1}^{12} x_{ihp} y_{ip} - \frac{(\sum_{p=1}^{12} x_{ihp})(\sum_{p=1}^{12} y_{ip})}{p}}{\sum_{p=1}^{12} x_{ihp}^2 - \frac{(\sum_{p=1}^{12} x_{ihp})^2}{p}}$$

onde p = 1, ..., 12 meses

$x_{ihp}$  é a produção no p-ésimo mês, da idade h, do rebanho i.

$y_{ip}$  é a produção adulta no p-ésimo mês, do rebanho i.

Na determinação dos intervalos de confiança das produções corrigidas, a nível de estratos, procedeu-se como em SAMPFORD(13)

$$\left\{ \begin{aligned} &\hat{Y}(RL)_{ih} - ts [ \hat{Y}(RL)_{ih} ] ; \\ &\hat{Y}(RL)_{ih} + ts [ \hat{Y}(RL)_{ih} ] \end{aligned} \right\}$$

onde

$$\hat{V} [ \hat{Y}(RL)_{ih} ] = \frac{N_{ih} - n}{N_{ih}} \cdot \frac{QM RES_{ih}}{n}$$

$$e s = \sqrt{\hat{V} [ \hat{Y}(RL)_{ih} ]}$$

e  $N_{ih}$  é o tamanho do estrato h na unidade primária i  
n é o tamanho dos estratos nas unidades secundárias.

e ainda:

$$QM RES_{ih} = \frac{SQ RES_{ih}}{n - 2}$$

$$SQ RES_{ih} = SQ TOTAL_{ih} - SQ (RL)_{ih}$$

3.4 – Estimativas a Nível de Conglomerados

A partir das produções corrigidas a nível de estratos, isto é, eliminados os efeitos de idade, mês e ano da parição, a obtenção das estimativas dos parâmetros a nível de conglomerados foi feita segundo SAMPFORD(13) para Amostragem por Conglomerados de Diferentes Tamanhos em Estágio Duplo.

Assim, para A = 40 Unidades Primárias na população e a = 4 Unidades Primárias na amostra,  $M_i$  = número de unidades secundárias por unidade na população e  $m_i$  = 180 unidades secundárias por unidade primária na sub amostra, adotou-se:

$$\bar{y} = \frac{\sum_{i=1}^a y_i}{\sum_{i=1}^a M_i} \text{ onde } y_i = M_i \bar{y}_i$$

$$e \bar{y}_i = \frac{\sum_{h=1}^H y_{ih}}{H}$$

ademais:

$y_{ih} = \hat{Y}(RL)_{ih}$  é a produção média corrigida de animais do rebanho i que pariram na idade h.  
 $\bar{y}$  é a média por unidade secundária, que aqui poderia ser

#### 4.2 – Estimativas das Produções Adultas a Nível de Estratos

O confronto entre as estimativas dos coeficientes de regressão linear das produções em idade adulta sobre as produções em idades jovens, para cada rebanho, revelou a existência de diferenças significativas ( $\alpha = 0,05$ ) entre eles. Este fato, de acordo com COCHRAN<sup>(3)</sup>, motivou o uso dos  $\hat{b}_{ih}$  a nível de estratos em lugar dos coeficientes combinados  $\hat{b}_{ic}$  que eventualmente seriam calculados a nível de rebanhos.

Assim, segundo a metodologia citada em (3.3), foram obtidas as estimati-

QUADRO 2 – Distribuição de  $\hat{b}_{ih}$

Rebanhos	IDADE (anos)			
	2	3	4	5
1	2,27	-0,17	0,49	1,18
2	-0,29	0,04	-0,25	0,40
3	0,06	-0,29	0,15	0,11
4	0,73	0,65	0,82	0,56

vas das produções em idade adulta, por ponto e por intervalo, os valores das semi amplitudes desses intervalos ( $d_{ih}$ ) em termos de porcentagem da respectiva média e os coeficientes de variação ( $CV_{ih}$ ) das produções corrigidas.

QUADRO 3 – Distribuição das Produções Estimadas em Idade Adulta, em kg; de  $d_{ih}$  e de  $CV_{ih}$ %.

$R_i$	Parâmetros	IDADE (Anos)			
		2	3	4	5
1	$\hat{Y}(RL)_{1h}$	4981,72	4669,05	4534,06	4545,24
	$d_{1h}$	4,35	7,88	7,92	6,89
	$CV_{1h}$	9,78	12,89	12,85	11,18
2	$\hat{Y}(RL)_{2h}$	4968,43	5070,86	4991,51	5059,30
	$d_{2h}$	4,24	4,21	4,15	3,69
	$CV_{2h}$	6,92	6,95	6,86	6,16
3	$\hat{Y}(RL)_{3h}$	5125,67	5057,44	5091,09	5120,53
	$d_{3h}$	2,91	3,09	3,23	3,20
	$CV_{3h}$	5,24	5,23	5,25	5,21
4	$\hat{Y}(RL)_{4h}$	5002,45	5020,23	4885,72	4765,99
	$d_{4h}$	3,15	3,68	3,88	3,65
	$CV_{4h}$	5,16	5,97	6,28	5,95

Os valores de  $d_{ih}$ , observados no quadro 3, mostraram-se bastante razoáveis, considerando-se a alta variabilidade inerente à produção leiteira. Sua variação de 2,91 a 7,92% e que em apenas 3 estratos do rebanho 1 foi superior a 5% parece evidenciar a validade de se haver tomado um número fixo de 12 unidades por estrato (36 animais).

Apenas a título de ilustração, calculou-se para o rebanho 2, a composição da amostra estratificada necessária para se obter a precisão média aqui alcançada ( $\bar{d} = 4,07\% \bar{y}$  est), se não houvessem sido aplicadas as técnicas dos conglomerados e da regressão. A amostra baseada na Partilha de Neymann resultou:

Idade (Anos)	2	3	4	5	6	Total
Animais	124	133	141	141	125	664

que representa quase que a totalidade dos 720 animais exigidos para os 4 rebanhos.

Por sua vez, o coeficiente de variação das produções corrigidas, pode ser interpretado, segundo SEARLE & HENDERSON<sup>(16)</sup>, como uma boa medida da validade do processo. Para esses autores, coeficientes de variação desejáveis estariam em torno de 20%. Sob esse aspecto, os resultados aqui obtidos foram satisfatórios, pois conforme consta do quadro 3, o maior coeficiente de variação encontrado, foi de 12,85%.

Ademais, para GOWEN<sup>(4)</sup>, SEARLE & HENDERSON<sup>(16)</sup>, NAUFEL<sup>(11)</sup>, se o processo de correção é adequado então os coeficientes de variação são aproximadamente constantes dentro dos rebanhos. Este fato pode ser também observado no quadro 3, com exceção do rebanho 1, que se mostrou discrepante aos 2 anos.

#### 4.3 – Estimativas das Produções Adultas, a Nível de Unidade Secundárias

No ítem anterior foram determinadas as estimativas por ponto e por intervalo, das produções médias corrigidas de uma determinada idade jovem para a idade adulta.

Aqui, a nível de unidades secundárias, determinaram-se as estimativas das produções médias, por ponto e por intervalo, para animais dos rebanhos paulistas, em idade adulta.

De acordo com a metodologia citada em (3.4) foram obtidos:

$$\bar{\bar{y}} = 4911,89 \text{ kg}$$

$$\hat{V}(\bar{\bar{y}}) = 7264,8460 \text{ kg}^2$$

$$\therefore [C(\bar{\bar{Y}}) : 4741,42 \leq \bar{\bar{Y}} \leq 5082,36 \text{ kg} \\ (\alpha = 0,05)$$

Que fornece uma precisão, em termos de porcentagem de  $\bar{\bar{y}}$ , de 3,47%.

#### 4.4 – Comparação com um Método Tradicional

Determinaram-se os fatores multiplicativos pelo método de GOWEN<sup>(4)</sup> e as produções médias mensais, a nível de estratos, foram através deles corrigidas. As médias das produções por idade, assim corrigidas constam do quadro 4.

Seguindo os procedimentos descritos em CAMPOS<sup>(2)</sup> o Teste das Ordens Assinaladas foi aplicado a cada idade isoladamente, obtendo-se os níveis mínimos de significância de 0,375; 0,625; 0,563 e 0,375 para 2, 3, 4 e 5 anos respectivamente. Desse modo, não foram detectadas diferenças significativas entre as produções corrigidas através do processo aqui descrito e através do método de Gowen.

#### 5 – CONCLUSÃO

Face aos resultados obtidos concluiu-se que o procedimento amostral adotado neste estudo pode ser considerado eficiente para corrigir a produção de rebanhos de gado holandês do Estado de São Paulo.

Ademais, a possibilidade de determinação dos intervalos de confiança das produções corrigidas, com precisão média em torno de 4% da estimativa por ponto, parece evidenciar uma vantagem deste processo sobre os processos tradicionais da Teoria dos Fatores de Correção.

QUADRO 4 – Distribuição das Médias das Produções Corrigidas Através do Método de Gowen, Segundo a Idade e o Rebanho.

Rebanho	Idade (Anos)			
	2	3	4	5
1	4681,74	4664,83	4653,05	4675,91
2	5057,19	5048,74	5053,01	5057,19
3	5110,25	5121,63	5118,10	5092,91
4	4882,15	4878,00	4872,50	4880,46

#### ABSTRACT

*In this research, the techniques for the estimation of regression in cluster sampling with double stage for correction of the milk production, was used. This procedure showed to be more advantageous than usual procedures, especially for determination of the confidence interval to the corrected productions.*

## REFERÊNCIAS BIBLIOGRÁFICAS

- 1 – BRICKET JR., R. *Melhoramento genético animal*. São Paulo, Melhoramentos, EDUSP, 1967. 269p.
- 2 – CAMPÕES, H. *Estatística experimental não-paramétrica*. 3.ed. Piracicaba, Depto. de Matemática e Estatística da ESALQ/USP, 1979. 343p.
- 3 – COCHRAN, W.G. *Técnicas de amostragem*. Fundo de Cultura, 1965. 585p.
- 4 – GOWEN, J.W. Studies in Milk Secretion. V – On the Variations and Correlations of Milk Secretion with Age. *Genetics*, 5: 11, 1920.
- 5 – IEMMA, A.F. Fatores de correção para a produção de leite segundo a idade – comparando produções ajustadas através de dois critérios. In: JORNADA CIENTÍFICA DO CAMPUS DE BOTUCATU, 7, 1977. *Anais*. Botucatu, 1977.
- 6 – \_\_\_\_\_. Fatores não genéticos para a correção da produção de leite bovino. Piracicaba, ESALQ/USP, 1980. Dissertação (Mestr.) ESALQ/USP.
- 7 – \_\_\_\_\_. Fatores de correção para a produção de leite bovino. VII – Estimativas por intervalo. In: JORNADA CIENTÍFICA DO CAMPUS DE BOTUCATU, 10, 1981. *Anais*. Botucatu, 1981.
- 8 – JOHANSSON, I. *Genetic and Environmental variations in Milk and butterfat yield - urbana*. 1961.
- 9 – LOBO, R.B. Método de correção de performance em bovinos, USP, São Paulo, 1976. 18p. (mimeografado)
- 10 – MILLER, O. A recent study of age adjustment. *J. Dairy Sci.*, 56 (7): 952, 1973.
- 11 – NAUFEL, F. Efeitos de alguns fatores ambientais e genéticos na produção de leite e gordura do rebanho experimental holandês preto e branco do departamento de produção animal de São Paulo. *Bol. Ind. Animal*, 23 (N.S): 21, 1965/66.
- 12 – PIMENTEL GOMES, F. Como varia durante o período de lactação a porcentagem de gordura do leite em Piracicaba. *Revista de Agricultura*, 19: 392, 1944.
- 13 – SAMPFORD, M.R. *An introduction to sampling theory*. 1.ed. London, Oliver and Boyd, 1962. 292p.
- 14 – SANDERS, H.G. The variations in milk yields caused by season of the year, service, age and dry period, and their elimination. *J. Agr. Sci.*, 18: 46, 1928.
- 15 – SEARLE, S.R. Simplified herd-level age – correction factors. *J. Dairy Sci.*, 43 (6): 821, 1960.
- 16 – SEARLE, S.R. & HENDERSON, C.R. Judging the effectiveness of age-correction factors. *J. Dairy Sci.*, 48: 347, 1965.