

Projecção de lactações incompletas

Projection of incomplete lactations

C. Roquete^{1,2} & J. Castro^{1,2}

RESUMO

A Análise Factorial não sendo uma metodologia recente tem, na sua essência, justificação teórica para a questão da projecção de lactações incompletas, que cada vez mais se justifica devido ao desenvolvimento quase exponencial da capacidade dos computadores em termos de memória, velocidade e armazenamento de dados.

Neste contexto foi aplicado e ensaiado um desenvolvimento computacional, com base em uma população de bovinos leiteiros da região Sul de Portugal, tendo-se demonstrado ser perfeitamente exequível um tratamento matemático-estatístico, tendo por base a Análise Factorial, para um elevado número de registo do Contraste Oficial Leiteiro.

Não sendo a única, é no entanto, mais uma metodologia para apoio da selecção e melhoramento da população bovina de leite, com a grande vantagem de não fornecer estimativas tendenciosas ou enviesadas.

Em resultado final, pela utilização da técnica, houve um incremento em cerca de 10% do número de lactações envolvidas na avaliação genética da população em estudo.

ABSTRACT

The Factorial Analysis not being a recent methodology, it has in its essence theoretical justification, for the problem of the projection of lactations, that each time more is justified due to the almost exponential development of the capacity of the computers in memory terms, speed and storage of data.

In this context it was applied and assayed a computational development, on the basis of a dairy cattle population of South region of Portugal, that it demonstrated as is perfectly justified a mathematician-statistician treatment, having for base the Factorial Analysis, for an enormous number of register of the Official Milking Test.

Being not the only one it is, however, another methodology for support of the selection and improvement of the dairy cattle population, with the great advantage not to supply unbiased estimates.

As a final result, by the use of the technique, there was an increment in about 10% of the total number of involved lactations in the genetic evaluation of the population in study.

¹ Departamento de Zootecnia, Universidade de Évora, Apartado 94, 7002-554 Évora; ² Instituto de Ciências Agrárias Mediterrâneas (ICAM), Universidade de Évora e-mail: croquete@uevora.pt

INTRODUÇÃO

A produção na lactação é normalmente calculada a partir das produções de apenas alguns dias, dando origem a uma lactação estandardizada de duração de 305 dias (Wiggans & Grossman, 1980).

Os métodos de cálculo podem ser diversos, sendo o mais utilizado o denominado de Fleishmen (método 39 da FAO - 1951, dirigido e fiscalizado pelo Comité Europeu de Contraste Leiteiro e Manteigueiro e controlado em Portugal pela D.G.V.).

Nas análises de caracterização da população Frísia é usual entrar apenas em linha de conta com as lactações terminadas normalmente e com duração superior a 180 dias e que, portanto, passarão a ser equivalentes a lactações estandardizadas de 305 dias.

Podemos focar a importância das lactações parciais e/ou incompletas, se considerarmos o valor do progresso genético associado a um esquema de selecção em bovinos leiteiros, cuja sub-população controlada não seja suficientemente grande. Assim, um programa óptimo de selecção requer um compromisso entre intensidade de selecção aplicada e a precisão do teste de descendência (Robertson, 1957).

Um dos pressupostos incide no facto que cada descendente tem igual importância para os propósitos do teste e de que o ponto crítico para o teste de descendência é o tamanho dos grupos das filhas dos touros em teste (Hinks, 1974). Na prática, os grupos são bastante desequilibrados e, quanto mais exigimos que a variância residual se encurte, mais estamos a incrementar o desequilíbrio e a tornar menos efectiva a fórmula base, de que cada filha deve ter a mesma importância.

Portugal é um país em que podemos considerar a população bovina leiteira pequena, em termos efectivos de testagem (uso limitado da inseminação artificial e serviços de

testagem). Como é evidente, o factor económico é uma componente importante de um esquema de selecção, e o incremento de estruturas não é, na maior parte dos casos a solução. Esta pode passar por caminhos que, teoricamente menos precisos, levem à resolução de pelo menos parte do problema. Deste modo, a utilização de lactações parciais e/ou incompletas pode ser uma alternativa.

JUSTIFICAÇÃO TEÓRICA

Confrontemos a situação de trabalharmos apenas com lactações completas e a possibilidade de utilização de lactações parciais ou incompletas. Assim, se optarmos por lactações incompletas e tendo presente Hinks (1974) em que:

- r = precisão do teste de descendência;
- h^2 = heritabilidade de um carácter;
- n_1 = número de descendentes do reprodutor em teste;
- w = número efectivo de descendentes.

$$r = \left(\frac{0,25 \cdot n_1 \cdot h^2}{1 - 0,25 \cdot h^2 (n_1 - 1)} \right)^{1/2}$$

Teoricamente, (r) é dado pela fórmula anterior, mas na prática há que ter em atenção o número efectivo de descendente dos reprodutores (t , factor de ponderação) em cada grupo contemporâneo (rebanho, ano, época de parto, etc), o que irá reduzir drasticamente esse factor. Então, (r') será:

$$w = \sum \frac{n_{11} \cdot n_{12}}{n_{11} + n_{12}} + \dots + \sum \frac{n_{1p} \cdot n_{2p}}{n_{1p} + n_{2p}}$$

$$= \sum w_1 + \dots + \sum w_p$$

$$t = \frac{w(1 + 0,25 \cdot h^2 \cdot (n_1 - 1))}{n_1(1 + 0,25 \cdot h^2 \cdot (w - 1) + c^2)}$$

$$r' = r \cdot t$$

Como (w), por definição não pode exceder (n₁), e c² igual a 0 ou 0,1, teremos que (t) será sempre menor que a unidade e (r) excederá sempre (r'). Portanto na prática nunca se atingirá o valor teórico, uma vez que ele está estreitamente ligado ao número efectivo de contemporâneas, para cada nível do factor ambiental. Deste modo, será mais fácil obter um número efectivo elevado para descendentes com lactações parciais e/ou incompletas do que apenas com lactações totais.

Quando utilizamos apenas registos de lactações completas, para propósitos de melhoramento genético de bovinos leiteiros, parte da componente hereditária é perdida (Syrstad, 1965). Por este motivo, é importante encontrar métodos apropriados para avaliar registos leiteiros que incluam lactações incompletas, com a ideia subjacente de que pretendemos na realidade fazer com que registos de duração inferior aos 305 dias sejam comparáveis. Nunca com o fito de prognosticar a produção da lactação completa. Ainda dentro do princípio de utilização de lactações incompletas, Danell (1982), aponta para a possibilidade de se obter a mesma confiança nas estimativas genéticas embora obtidas de 4 a 7 meses mais cedo.

Van Vleck & Henderson (1961) atestam da importância da utilização dos registos leiteiros incompletos e investigam o problema da combinação da informação parcial e completa, segunda a óptica dos índices de selecção.

Consideremos (i) grupos de registos com médias independentes e um índice para touros baseado nas médias das respectivas filhas (\bar{X}_i), sendo

$$I = \sum b_i (\bar{X}_i - \mu_i)$$

em que (μ_i) é a média geral do grupo i-ésimo e (b_i) o factor de ponderação. Desenvolvendo teremos:

$$I = b_1(\bar{X}_1 - \mu_1) + b_2(\bar{X}_2 - \mu_2)$$

com os factores de ponderação calculados pelo sistema de equações:

$$\{B_1 + B_2 \cdot r_{\bar{X}_1\bar{X}_2} = r_{\bar{X}_1} G_1$$

$$\{B_1 \cdot r_{\bar{X}_1\bar{X}_2} + B_2 = r_{\bar{X}_2} G_2$$

sendo (r) o coeficiente de correlação entre as variáveis indicadas e (G_i) é o genótipo com respeito ao tipo de lactação (1 - completa; 2 - incompleta) e (B_i) os coeficientes de regressão estandardizados correspondentes a (b_i):

$$B_i = \frac{\sigma_{\bar{X}_i}}{\sigma_{G_i}} \cdot b_i \Rightarrow b_i = B_i \cdot \frac{\sigma_{G_i}}{\sigma_{\bar{X}_i}} = B_i \cdot \frac{\sigma_{G_i}}{\sigma_{\bar{X}_i}} \cdot \frac{\sigma_{G_i}}{\sigma_{G_i}}$$

e segundo Syrstad (1965) por razões práticas é plausível considerar σ_{G2} igual a σ_{G1} . Então

$$b_i = B_i \cdot h_{\bar{X}_i} \cdot \frac{\sigma_{G_i}}{\sigma_{G_i}} = B_i \cdot h_{\bar{X}_i}$$

onde ($h_{\bar{X}_i}$) representa a estimativa da raiz quadrada da heritabilidade média baseada nos meios irmãos, isto é:

$$h_{\bar{X}}^2 = \frac{0,25 \cdot h^2 \cdot n}{1 + (n-1) \cdot 0,25 \cdot h^2}$$

em que (h^2) representa a heritabilidade do registo individual, (n) o número de registos incluídos na média e como o quadrado do coeficiente de correlação entre o índice (I) e o genótipo (B) é no geral igual a:

$$R_{IG_1}^2 = \sum B_i \cdot r_{\bar{X}_i G_1}$$

teremos a confiança ou fiabilidade das estimativas função dos conjuntos de grupos de informação que por sua vez é função dos número de elementos dentro de cada grupo de meios-irmãos (lactações completas ou incompletas), como indicado pelos autores citados.

O valor da correlação genética entre registos de lactações incompletas e os completos,

indicada por diversos autores, é elevada, na ordem dos 0,8 para registos com duração inferior a 100 dias e de cerca de 0,9 quando controlados entre os 100 e 200 dias. Quanto à correlação entre registos dos controles leiteiros consecutivos, como exemplo, antes e após os 100 dias é da ordem de 0,85 (Madden, 1955; Rendel *et al.*, 1957; Pirchrter, 1961; Searle, 1961; Van Vleck e Henderson, 1961; Decking, 1964; citado por Syrstad, 1965).

No entanto, é apontado por VanRaden *et al.*, (1991) que existe uma correlação genética entre as lactações incompletas projectadas e as reais superior a 0,92, o que indica que as lactações incompletas são valiosas se modeladas adequadamente. Ainda indica que os ganhos genéticos esperados com a sua utilização serão na ordem de 0,27 a 0,30 vezes, mas permitirão avaliações genéticas mais estáveis, principalmente para as primeiras lactações.

Um dos pontos críticos que a informação de lactações incompletas, quando projectadas para 305 dias de duração de lactação possui, é o de assumir como igual, a variância fenotípica e a genética como se de lactações completas se tratassem (Weller, 1988; Wiggans *et al.*, 1988).

Portanto, se a correlação fenotípica e genotípica entre registos de produção de lactações completas e incompletas são elevadas, razão existirá para a utilização destas quando da estimativa dos valores genéticos (Auran, 1976; Danell, 1982).

Numa revisão bibliográfica realizada por Auran & Mocquot (1974) são apontadas algumas das razões para que se utilizem lactações incompletas virtualmente prolongadas:

a) os valores genéticos das vacas podem ser calculados em qualquer data sem a atendermos à data ou época de parto que é o mesmo que dizer que a ordenação das vacas

pode ser realizado independente da fase de lactação;

b) todas as lactações incompletas devem ser incluídas em testes de descendência, uma vez que aumentam a precisão ou que pelo menos mantêm a mesma precisão devido ao maior número de grupos de descendentes;

c) a inclusão de apenas lactações completas nas análises genéticas inerentes aos testes de descendência implica uma escolha prévia dos registos o que pode tornar tendenciosa a estimativa dos valores genéticos e, deste modo, é de esperar que touros cujas filhas tenham sido sujeitas a maiores taxas de refugo acabe por ter o valor genético sobrestimado.

Quanto aos métodos de avaliação das lactações totais a partir das lactações incompletas ou completas mas inferiores à duração estandardizada de 305 dias a história é longa e sinuosa. Em alguns casos existe deslocação para o lado da facilidade dos cálculos e em outros casos a opção é pela precisão. Os apoiantes da primeira opção (Mocquot & Auran, 1975) utilizaram o método das razões ao passo que os segundos (Van Vleck & Henderson, 1961; Miller & Corley, 1965; Miller *et al.*, 1972, Dordio, 1988) optam pelas regressões múltiplas.

Outras posições divergentes quanto ao problema da transformação de lactações incompletas em lactações com duração virtual de 305 dias, consistem em testar os métodos e a respectiva precisão numa amostra de registos e aplicá-los posteriormente noutros conjuntos independente de dados ou no próprio conjunto (Syrstad, 1964; Miller *et al.*, 1972; Auran & Mocquot, 1974, citados por Auran, 1976).

Os factores de projecção das lactações incompletas tem como objectivo prever a produção completa aos 305 dias, no entanto, estas têm menores variâncias genéticas e fenotípicas e, deste modo, não são óptimas

para usar nas avaliações genéticas, portanto deverão receber menos ênfase (Weller, 1988; VanRaden *et al.*, 1991).

Na aplicação do método experimental por nós proposto e ensaiado, para prever lactações de 305 dias a partir de lactações incompletas, com base na utilização da Análise Factorial, mais do que procurar uma amostra independente, foi considerado mais importante e mais correcto empregá-lo para conjuntos específicos da população em análise tendo em atenção as épocas de parto (Roquete *et al.*, 1989).

METODOLOGIA

Um dos mais divulgados métodos para projectar lactações incompletas, tem utilizado as regressões lineares simples ou múltiplas como anteriormente focado. No entanto, a metodologia que iremos apresentar tem por base a Análise Factorial que sendo recente, ressurgiu de um modo consistente no princípio do século com Sperman e, aprofundada por Thurstone mais tarde em 1947. Na sua essência, utiliza ajustamentos que se baseiam na álgebra linear com correspondente representação gráfica em que se descreve o comportamento de determinados pontos num plano em referência a um eixo (Lebart *et al.*, 1982). Com o método propomo-nos reconstituir, a partir de um número reduzido de factores, as correlações existentes entre as variáveis observadas. Teoricamente, consiste na existência de uma matriz de covariância, simétrica, definida positiva; na procura da representação da variação de cada variável em termos da contribuição que seja específica para essa variável; e as contribuições de (f) perturbações aleatórias, comuns para mais de uma variável ao mesmo tempo.

A contribuição de cada factor comum, para uma variável particular é standardiza-

da de modo a representar a correlação entre a variável em questão e o factor hipotético.

Uma das restrições para o modelo consiste em os factores não terem de ser correlacionados. Deste modo, a matriz de covariância prognosticada (Σ) será (Joreskog, 1973):

$$\Sigma = B (\Omega \Phi \Omega' + \delta^2) B' + \theta^2$$

Ω (v,f) - matriz dos factores comuns;

Φ (f,f) - matriz das correlações entre factores comuns ;

δ^2, θ^2 - matrizes diagonais contendo as variâncias específicas;

B – matriz arbitrária que permite os factores serem escalonados e combinados num número diverso de maneiras.

Este modelo factorial é indeterminado, uma vez que existe um número infinito de restrições para (Ω) e (Φ) que dão as mesmas soluções para (Σ). Então, teoricamente, uma vez que as restrições são arbitrárias, façamos com que os factores tenham de ser ortogonais ($\Phi = I$), permitindo a obtenção de vectores próprias que definirão combinações lineares das variáveis, não correlacionadas e com variância mínima (Lebart *et al.*, 1982).

Concretamente para a projecção das lactações incompletas, a metodologia empregue consistiu em três fases:

a) fase de preparação dos registos. Utilizámos a totalidade dos registos do contraste leiteiro: - a produção no dia do teste de cada animal que tivesse a 1ª lactação completa de 305 dias (12820 registos); e seguidamente calcularam-se as produções acumuladas por animal e para cada um dos 10 contrastes ($LACUM_{j(i)}$), teoricamente separados por intervalos de 30,5 dias. Paralelamente foi calculada a correspondente produção desconhecida ou remanescente para cada um dos contrastes ($LDESC_j$) do animal com produção standardizada ($L305_i$):

$$\text{LDESC}_{j(i)} = \text{L305}_i - \text{LACUM}_{j(i)}$$

Como a Análise Factorial utiliza variáveis estandardizadas fomos calculá-las para cada controlo, enquadrada numa classe ou grupo de manejo (k), que no nosso caso estava associado à época e idade ao primeiro parto. Danell (1982), aponta para projecções de lactações tendo em consideração os efeitos ambientais que normalmente condicionam as lactações completas:

$$Z_{jk(i)} = (\text{LDESC}_{j(i)} - \text{LDESC}_k) / \sigma_k$$

tendo este processo sido realizado como apoio de um programa informático criado e desenvolvido em Clipper (ver nota no final).

b) fase de cálculo dos factores comuns e específicos, utilizando o pacote estatístico NCSS e os dados preparados anteriormente. Este programa teve como base a equação em forma matricial:

$$Z = \Omega f + \varepsilon$$

Z - matriz das produções acumuladas em cada um dos 10 contrastes leiteiros;

ε - erros aleatórios;

Ω e f — definidos anteriormente;

em que se assumia independência entre os erros e as observações e entre os próprios factores. Então os factores ($f_{mk(i)}$) para cada animal e correspondente a cada controlo conhecido (m), são obtidos de

$$f_{mk(i)} = \Omega' R_{mk}^{-1} Z_{mk}$$

R - matriz das correlações entre as (m) variáveis (controles) conhecidos.

c) fase do cálculo das lactações projectadas. Construímos programa em Fortran 77 (anexo 2) que calcula os valores das variáveis estandardizadas após a leitura dos valores do vector dos factores comuns (f) e da matriz (Ω) estimados na alínea anterior,

$$Z_{(m+1,10)k} = \Omega_{(m+1,10)k} \cdot f_{mk}$$

depois lê cada conjunto de registos de produção no dia do controlo referentes a um

animal e através dos factores previamente estimados e armazenados, tendo em atenção a época e idade ao parto, irá fornecer a lactação projectada (L305P_{ki}) da vaca referida (i):

$$\text{L305P}_{ki} = \text{LACUM}_k + \sum_{j=m+1}^{10} Z_{jk} \cdot \sigma_{jk}$$

Posteriormente toda a informação foi utilizada na predição dos valores genéticos da população frísia do Sul de Portugal, aplicando-se o modelo animal-BLUP (Roquete, 1993), havendo a realçar a consistência das soluções e a possibilidade do incremento dos animais avaliados (mais 10%), segundo a aplicação da Análise Factorial para a projecção de lactações incompletas.

A grande vantagem deste método, que deve fornecer estimativas não tendenciosas, reside no facto que os factores serem estimados dos valores da própria amostra onde vão ser aplicados (Arendonk & Fimland, 1983), além de se ter demonstrado um processo eficaz e relativamente simples para a dimensão dos dados utilizados.

NOTA:

Às pessoas interessadas os autores podem disponibilizar o código dos seguintes programas informáticos:

1 - Programa em Clipper para preparação dos registos de lactações incompletas para futura projecção em lactações de 305 dias.

2- Programa em Fortran 77 para projectar lactações incompletas em lactações de 305 dias.

REFERÊNCIAS BIBLIOGRÁFICAS

- Arendonk, J. M. & Fimland, E. 1983. Comparison of two methods to extend partial milk records. *Z. Tier. Zucht.*, **100**: 33-38.
- Auran, T. & Mocquot, J. C. 1974. Etudes sur la production laitière des bovins. II-

- Intérête des lactations partielles pour la sélection. *Génét. Sél. Animal*, **6(4)**:429-444.
- Auran, T. 1976 Studies on monthly and cumulative monthly milk yields records. IV – Estimating total lactation from part-lactation. *Act. Agr. Scand.* **26**:10-17.
- Danell, B. 1982. Studies on lactation yield and individual test-day yields of swedish dairy cows. IV - Extension of part-lactation records for use in sire evaluation. *Act. Agr. Scand.*, **32**: 103-114.
- Dordio, 1988. *Lactações Parciais – Análise e Utilização Prática*. Trabalho Fim de Curso de Eng.^a Zootécnica. Univ. Évora.
- Hinks, C.J. 1974. The planning and organisation of progeny testing with particular reference to numerically small populations and breeds of fairy cattle. *Z. Tier. Zuch.*, **91**:169-175.
- Joreskog 1973. *Factor Analysis*. Polico-piado.
- Lebart, L. A. Morinea & Fénelon, J.P. 1982. *Traitement des Données Statistiques- Méthodes et Programmes*. 2^a édition, Dunod. Paris.
- Miller, R. H. & Corley, E.L. 1965. DHIA factors for projecting incomplete records to 305 days. *DHIL*, Vol. 41, n^o6.
- Miller, R. H., Person & Fhorman 1972. Methods of projecting complete lactation production from part-lactation yield. *Dairy Sci.* **55**:1602-1610.
- Mocquot, J. & Auran, T. 1975. Etudes sur la production laitière des bovins II – intérêt des lactations partielles pour la sélection. *Ann. Génét. Select. Anim.*, **7(1)**:59-65.
- Robertson, A. 1957. Optimum groups sire in progeny testing and family selection. *Biometrics*, 442-453.
- Roquete, C., Feneja, P., Rodrigues, C. Carvalho, J., Hausmann, H. & Potes, N. 1989. *Análise e Incremento do Sistema de Produção Leiteiro no Perímetro de Rega do Mira*. GTZ, Resultados de projectos de investigação agrícola. Vila Real.
- Syrstad, O. 1965. Studies on dairy herd records. I-Evaluation of incomplete records. *Act. Agr. Scand.*, **14**:129-136.
- Roquete, C. 1993. *Aplicação do Modelo Animal na Caracterização Genética das Populações Frísia e Mertolenga no Alentejo*. Tese doutoramento. Universidade de Évora.
- Vanraden, P., R. Wiggans & Ernst, A. 1991. Expansion of projected lactation yield stabilize genetic variance.: *Dairy Sci.*, **74**:4344-4350.
- Van Vleck, L. & Henderson, C. 1961. Utilizing both part and complete daughter records in sire evaluation. *J. Dairy Sci.*, **44**:2068-2075.
- Weller, J. I. 1988. Inclusion of partial lactations in the genetics analysis of yield traits by differential weighting of records. *J. Dairy Sci.* **71**:1873-1881.
- Wiggans, G. R. & Grossan, M. 1980. Computing lactation records from sample-day production. *DHIL*. Vol. 56, n^o4.
- Wiggans, G. R. 1988. Animal model Evaluation of Ayrshire milk yield with all lactation, herd-sire interaction and groups based unknown parents. *J. Dairy Sci.*, **71 (Supl 2)**: 1319-1329.