

Estudo de características de produção de matrizes de corte por meio da análise de componentes principais

[Study of meat-type chickens production traits by principal components analysis]

M. Yamaki^{1,4}, G.R.O. Menezes^{2,4}, A.L.C. Paiva^{1,4}, L. Barbosa^{1,5}, R.F. Silva^{1,5},
R.B. Teixeira^{2,4*}, R.A. Torres^{3,4}, P.S. Lopes³

¹Aluno de pós-graduação - UFV – Viçosa, MG

²Aluno de graduação - UFV – Viçosa, MG

³Departamento de Zootecnia - UFV – Viçosa, MG

⁴Bolsista do CNPq

⁵Bolsista da CAPES

RESUMO

Avaliou-se o descarte de variáveis de produção, em análises de componentes principais, de três linhagens de matrizes de corte do Programa de Melhoramento Genético da Universidade Federal de Viçosa, utilizando informações de 270 aves, sendo 90 de cada linhagem. As características analisadas foram dias para postura do primeiro ovo (DPPO), taxa de postura da 22^a a 56^a semana (TP), peso médio individual na 32^a (PMI1), na 40^a (PMI2), na 48^a (PMI3), na 56^a (PMI4) e na 64^a semana (PMI5) e peso médio do ovo, obtido pela média da pesagem de três ovos na 32^a (PMO1), na 40^a (PMO2), na 48^a (PMO3), na 56^a (PMO4) e na 64^a semana (PMO5). Dos 12 componentes principais, sete apresentaram variância menor do que 0,7 (autovalor menor do que 0,7), sugerindo-se sete variáveis para descarte. As variáveis descartadas foram aquelas que apresentaram maiores coeficientes, em valor absoluto, a partir do último componente principal. Observou-se correlação linear simples e significativa entre as variáveis descartadas e as não descartadas, que indica redundância de variáveis, razão do descarte. Recomendam-se as variáveis: DPPO, TP, PMI4, PMO1 e PMO4 para o estudo de características da produção de matrizes de frango de corte por meio da análise de componentes principais.

Palavras-chave: frango de corte, análise multivariada, descarte de variável, correlação

ABSTRACT

Records of 270 meat-type chickens from three lines, 90 of each one, were used to discard variables in a principal component analysis. Data were obtained from meat-type chicken lines of the genetic breeding program of the Universidade Federal de Viçosa. The following traits were evaluated: days at first egg (DFE), egg production rate (EPR) from 22nd to 56th week, body weights at 32nd (BW1), 40th (BW2), 48th (BW3), 56th (BW4), and 64th weeks of age (BW5), and average of three egg weights, at 32nd (EW1), 40th (EW2), 48th (EW3), 56th (EW4) and at 64th weeks (EW5). From the 12 principal components, seven showed variance lower than 0,7 (eigenvalue lower than 0,7), suggesting seven variables to be discarded. Variables which showed the highest coefficients, in absolute value, in the last principal component were discarded. Highly correlated variables with the smaller principal components variance explain a small part of the whole variation. In addition, discarded variables in function of the significant simple linear correlation with the nondiscarded variable, were considered redundants. The variables DFE, EPR, BW4, EW1, and EW4 are recommended for principal component analysis of broiler matrix.

Keywords: broiler, discard of variable, multivariate analysis, correlation

Recebido em 23 de julho de 2008

Aceito em 27 de janeiro de 2009

*Autor para correspondência (corresponding author)

E-mail: rafaelzootecnia@yahoo.com.br

INTRODUÇÃO

Em muitas situações, os pesquisadores tendem a avaliar o maior número possível de características e geram, com isso, acréscimo considerável de trabalho. Quando o número de características mensuradas é grande, há possibilidade de algumas delas serem de pouca contribuição para a discriminação dos indivíduos avaliados. Neste caso, despende-se muito tempo para mensurar grande número de variáveis sem melhoria na precisão de avaliação das aves, além de tornar mais complexa a análise e a interpretação dos resultados. Assim, podem-se eliminar as características redundantes e de difícil mensuração, o que reduziria o tempo e o custo dos experimentos.

A técnica de componentes principais foi originalmente descrita por Pearson (1901) e, posteriormente, aplicada por Hotelling (1933), citado por Jolliffe (1973), em diversas áreas da ciência. De acordo com Manly (1988), o uso da técnica de componentes principais acentuou-se quando houve maior disponibilidade de recursos na área computacional. O método consiste na transformação do conjunto original de variáveis em outro, os componentes principais, de dimensões equivalentes, porém com a propriedade de que cada componente retém porcentagem da variância original e de que as variâncias são decrescentes do primeiro ao último componente principal (Morrison, 1976).

Componentes principais são combinações lineares de variáveis aleatórias ou estatísticas, que possuem propriedades especiais, em termos de variâncias. Entre essas propriedades especiais, destaca-se o fato de que o primeiro componente principal possui a máxima variância, tendo o segundo componente principal a segunda maior variância e, assim, sucessivamente, de forma que o máximo de informação, em termos de variação total, está contido nos primeiros componentes. Além disso, os componentes principais representam vetores linearmente independentes. (Anderson, 1958).

Segundo Baker et al. (1988), a análise dos componentes principais pode revelar relações não identificadas previamente e contribuir para melhor interpretação dos dados. Além disso, muitas vezes representa método para explicar a

estrutura de covariância entre as várias características medidas em uma população.

Decisões acerca da seleção de animais em função de determinada característica podem ser tomadas de forma intuitiva, negligenciando seus componentes e as ligações existentes entre eles (Roso e Fries, 1995). Nesse contexto, a análise de componentes principais deve contribuir na interpretação das relações existentes entre as variáveis e, conseqüentemente, na tomada de decisões (Baker et al., 1988; Roso e Fries, 1995).

O trabalho teve o objetivo de, ao utilizar análises de componentes principais, avaliar as variáveis mensuradas em experimentos genéticos e estabelecer critérios para eliminação de variáveis redundantes, sem que houvesse perda apreciável na explicação da variação total do experimento.

MATERIAL E MÉTODOS

Os dados utilizados são provenientes de três linhagens de aves de corte do Programa de Melhoramento Genético da Universidade Federal de Viçosa e totalizam 270 aves avaliadas, sendo 90 de cada linha. As aves foram alojadas em gaiolas individuais na 22ª semana de idade.

Utilizando-se a matriz de correlação entre todas as variáveis mensuradas, fez-se avaliação da multicolinearidade entre as características para avaliar a dependência linear entre as variáveis, que pode redundar em matrizes singulares ou mal condicionadas.

As características analisadas foram dias para postura do primeiro ovo (DPPO); taxa de postura da 22ª a 56ª semana (TP); peso médio individual na 32ª (PMI1), na 40ª (PMI2), na 48ª (PMI3), na 56ª (PMI4) e na 64ª semana (PMI5); e peso médio do ovo obtido pela média da pesagem de três ovos na 32ª (PMO1), na 40ª (PMO2), na 48ª (PMO3), na 56ª (PMO4) e na 64ª semana (PMO5).

A análise de componentes principais consiste em transformar, a partir da matriz de correlação, um conjunto de variáveis Z_1, Z_2, \dots, Z_p em um novo conjunto de p variáveis (Y_1 (CP1), Y_2 (CP2), Y_p (CPp)) não-correlacionadas entre si e que explicam a variância total observada entre os dados (Regazzi, 2002). A idéia principal do procedimento consiste em utilizar os primeiros

componentes principais, que contenham a maior variabilidade dos dados originais, descartando-se racionalmente os demais componentes e reduzindo a dimensionalidade e o número de variáveis mensuradas.

O descarte de variáveis foi feito eliminando-se a variável com maior correlação no componente principal, referente ao menor autovalor (menor variância), variável esta julgada menos importante para explicar a variação total dos dados (Regazzi, 2002), porque as variáveis altamente correlacionadas com os componentes principais de menor variância, de acordo com Mardia et al. (1997), representam variação praticamente insignificante. O critério do número de variáveis descartadas foi, conforme recomendações de Jolliffe (1973), eliminar tantas variáveis quanto for o número de componentes cuja variância (autovalor) fosse menor que 0,7. Todas as análises foram feitas utilizando o programa SAS System, versão 8.0/1999/.

RESULTADOS E DISCUSSÃO

Os resultados obtidos para os componentes principais, seus respectivos autovalores e porcentagens da variância total explicada por componente são apresentados na Tab. 1.

Dos 12 componentes principais, sete (58,3%) apresentaram variância menor que 0,7 (autovalor menor que 0,7). Assim os cinco primeiros componentes principais que explicam 77,19% da variação total foram selecionados.

Os seis primeiros componentes principais explicaram 82,8% da variação total, confirmando os resultados de Abreu et al. (1999a) que, ao analisarem a combinação de características produtivas de linhagens de matrizes de frangos de corte, também relataram que os seis primeiros componentes principais explicam mais de 80% da variação total dos dados.

Strapasson et. al. (2000), ao selecionarem variáveis na caracterização de germoplasma de *Paspalum* sp. por meio de componentes principais, com o intuito de descartar as variáveis consideradas redundantes ou não-discriminantes, conseguiram selecionar oito variáveis, consideradas mais importantes na explicação da

variabilidade presente na coleção de acessos do germoplasma estudado.

Tabela 1. Componentes principais (CP), autovalores (λ_i) e porcentagem da variância, explicada pelos componentes (% VCP), das características de produção de matrizes de frango de corte

Componentes principais	(λ_i)	% VCP	% VCP (acumulada)
CP ₁	3,46	28,83	28,83
CP ₂	2,85	23,76	52,59
CP ₃	1,13	9,42	62,01
CP ₄	1,00	8,33	70,34
CP ₅	0,82	6,85	77,19
CP ₆	0,67	5,62	82,81
CP ₇	0,52	4,36	87,16
CP ₈	0,46	3,87	91,03
CP ₉	0,39	3,26	94,29
CP ₁₀	0,29	2,39	96,68
CP ₁₁	0,22	1,84	98,52
CP ₁₂	0,18	1,48	100

As sete variáveis que apresentaram maiores coeficientes, em valor absoluto, a partir do último componente principal, foram descartadas, conforme apresentado na Tab. 2.

Neste estudo, as variáveis que podem ser descartadas e em ordem de menor importância para explicar a variação total dos dados são: PMI2, PMI3, PMO5, PMO2, PMI5, PMO3, PMI1.

Os resultados, apesar de se referirem à espécie diferente, assemelham-se aos encontrados por Barbosa (2003), que avaliaram 13 características de desempenho em suínos e concluíram que oito delas (61,5%) foram redundantes, portanto podiam ser descartadas.

Abreu et al. (1999b), ao trabalharem com produção de ovos de matrizes de frango de corte, obtiveram melhores resultados com os dois primeiros componentes principais que explicaram 98% da variação total dos dados. Quanto menor o número de variáveis que expliquem maior fração da variação total, maior a economia de tempo, o que é altamente desejável.

Tabela 2. Coeficientes de ponderação das características de matrizes de frango de corte com os componentes principais descartados em ordem de menor importância

Característica	Coeficiente						
	CP ₆	CP ₇	CP ₈	CP ₉	CP ₁₀	CP ₁₁	CP ₁₂
DDPO	0,389212	0,194931	0,184014	0,251019	0,129184	0,058381	0,021451
TP	0,017217	-0,132851	-0,126335	0,016228	0,011622	0,076659	0,041735
PMO1	-0,493404	0,334938	0,212531	0,155646	0,029069	-0,032784	-0,068397
PMO2	0,393811	-0,389982	0,181420	-0,589363	-0,038857	-0,072479	-0,033397
PMO3	-0,144868	-0,599195	0,076763	0,526729	0,227642	0,136229	-0,022728
PMO4	-0,008738	0,421402	-0,306516	-0,262613	0,567590	0,005214	-0,018962
PMO5	0,008628	0,240940	0,015986	0,051518	-0,747787	0,010998	0,150468
PMI1	0,606293	0,255948	0,186679	0,396236	0,079342	0,012415	-0,031142
PMI2	-0,006520	-0,089304	-0,448175	0,034034	-0,012665	0,260454	0,638423
PMI3	-0,035890	-0,109223	-0,286224	0,124422	-0,068783	-0,721871	-0,312600
PMI4	-0,096984	-0,001084	0,033306	-0,123458	-0,098542	0,594679	-0,581379
PMI5	-0,224094	-0,001199	0,673240	-0,164935	0,163583	-0,151489	0,352440

DDPO: dias para o primeiro ovo; TP: taxa de postura da 22^a a 56^a semana; PMO1: peso médio do ovo na 32^a semana; PMO2: peso médio do ovo na 40^a semana; PMO3: peso médio do ovo na 48^a semana; PMO4: peso médio do ovo na 56^a semana; PMO5: peso médio do ovo na 64^a semana; PMI1: peso médio individual na 32^a semana; PMI2: peso médio individual na 40^a semana; PMI3: peso médio individual na 48^a semana; PMI4: peso médio individual na semana 56^a; PMI5: na 64^a semana.

Mascioli et al. (2000), também, ao trabalharem com espécie diferente, analisaram os componentes principais (CP) para características de crescimento na raça Canchim, observaram que os dois primeiros CP foram responsáveis por mais de 75% da variação total, tanto para os ganhos como para os pesos, e concluíram que se pode utilizar a variância fenotípica para o estudo dos componentes principais. Para os pesos, a maior parte da variância total dos dados foi justificada pelo primeiro componente principal, que explicou 74,2% da variação. Esse CP contrasta animais mais pesados com animais mais leves após o nascimento e indica que a influência do peso se torna mais importante à medida que o animal envelhece, como

observaram Barbosa e Smith (1988) na raça Charolês.

Adicionalmente, foi possível observar que as características sugeridas para descarte apresentaram correlação linear simples significativas com as não descartadas, ou seja, são redundantes. Por outro lado, as variáveis selecionadas apresentam menor correlação entre si (Tab. 3). Deste modo, cada característica selecionada deve ser responsável por um tipo de informação biológica exclusiva, e a ação conjunta destas será complementar para descrição geral dos indivíduos ou população estudados.

Tabela 3. Coeficientes de correlação simples entre as características mensuradas em matrizes de frango de corte durante o experimento

	DDPO	TP	PMO1	PMO2	PMO3	PMO4	PMO5	PMI1	PMI2	PMI3	PMI4	PMI5
DDPO	1,000											
TP	0,0789	1,0000										
PMO1	0,0778	0,0873	1,0000									
PMO2	0,1842	0,0279	0,4241	1,0000								
PMO3	0,1067	-0,0478	0,3361	0,4965	1,0000							
PMO4	0,1219	-0,0023	0,3325	0,4505	0,5248	1,0000						
PMO5	0,1948	0,0146	0,3499	0,4824	0,5980	0,6927	1,0000					
PMI1	-0,3043	-0,0040	0,0141	0,1323	-0,0572	0,0552	0,0211	1,0000				
PMI2	-0,0731	-0,0507	0,1507	0,1621	0,0108	0,0947	0,0309	0,4170	1,0000			
PMI3	0,0002	0,0454	0,1029	0,1470	0,0899	0,1723	0,1483	0,3760	0,7630	1,0000		
PMI4	-0,0828	-0,0029	0,0364	0,0606	-0,0269	0,0943	0,0702	0,4042	0,7388	0,7483	1,0000	
PMI5	-0,1273	0,0348	-0,0453	-0,0557	-0,0518	0,0240	0,0178	0,3422	0,4895	0,5956	0,7099	1,000

DDPO: dias para o primeiro ovo; TP: taxa de postura da 22^a a 56^a semana; PMO1: peso médio do ovo na 32^a semana; PMO2: peso médio do ovo na 40^a semana; PMO3: peso médio do ovo na 48^a semana; PMO4: peso médio do ovo na 56^a semana; PMO5: peso médio do ovo na 64^a semana; PMI1: peso médio individual na 32^a semana; PMI2: peso médio individual na 40^a semana; PMI3: peso médio individual na 48^a semana; PMI4: peso médio individual na semana 56^a; PMI5: na 64^a semana.

Com base nesses resultados, recomendam-se as variáveis DPPO, TP, PMI4, PMO1 e PMO4 para serem mantidas nas análises de componentes principais. Das características mensuradas de peso corporal, somente PMI4 não pode ser descartada; as demais de peso corporal são redundantes. Para peso médio dos ovos, três características podem ser descartadas, com a utilização de PMO1 e PMO4 que apresentam menor correlação quando comparadas às outras características de peso médio dos ovos. As características DPPO e TP foram pouco correlacionadas com as demais variáveis, portanto não devem ser descartadas.

CONCLUSÕES

A análise de componentes principais é efetiva e permite o descarte de sete variáveis das 12 variáveis mensuradas no experimento. Quanto menor o número de variáveis necessárias para explicar a variação total observada, mais rápida e fácil se torna a análise e resulta em importante economia de tempo.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABREU, V.M.N.; SILVA, M.A.; CRUZ, C.D. et al. Capacidade de combinação de características produtivas de linhagens de matrizes de frangos de corte, usando a técnica de componentes principais. *Rev. Bras. Zootec.*, v.28, p.250-257, 1999a.
- ABREU, V.M.N.; SILVA, M.A.; CRUZ, C.D. et al. Capacidade de combinação de características de produção de ovos de linhagens de matrizes de corte usando componentes principais. *Rev. Bras. Zootec.*, v.28, p.955-959, 1999b.
- ANDERSON, T.W. *An introduction to multivariate statistical analysis*. New York: John Wiley & Sons, 1958. 374p.
- BAKER, J.F.; STEWART, T.S.; LONG, C.R. et al. Multiple regression and principal components analysis of puberty and growth in cattle. *J. Anim. Sci.*, v.66, p.2147-2158, 1988.
- BARBOSA, L. Avaliação de características de desempenho de suínos por meio de componentes principais. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 40., 2003, Santa Maria. *Anais...* Santa Maria, 2003. (resumo) CDROM.
- BARBOSA, P.F.; SMITH, W.B. Aplicação de técnicas de análise multivariada em melhoramento animal. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 25., 1988, Viçosa. *Anais...* Viçosa, 1988. p.240. (resumo).
- HOTELLING, H. Review of the triumph of mediocrity in business. *J. Am. Stat. Assoc.*, v.28, p.463-465, 1933.
- JOLLIFFE, I.T. Discarding variables in a principal component analysis, II: real data. *Appl. Stat.*, v.22, p.22-31, 1973.
- MANLY, B.F.J. *Multivariate statistical methods: a primer*. London: Chapman and Hall, 1988. 159p.
- MARDIA, K.V.; KENT, J.T.; BIBBY, J.M. et al. *Multivariate analysis*. 6.ed. London: Academic, 1997. 518p.
- MASCIOLI, A.S.; EL FARO, L.; ALENCAR, M.M. et al. Estimativas de parâmetros genéticos e fenotípicos e análise de componentes principais para características de crescimento na raça Canchim. *Rev. Bras. Zootec.*, v.29, p.1654-1660, 2000.
- MORRISON, D.F. *Multivariate statistical methods*. 2.ed. Singapore: McGraw Hill, 1976. 415p.
- PEARSON, K. On lines and planes of closest fit to system of point in space. *Philosophical Magazine*, 1901. Disponível em <<http://pbil.univ-lyon1.fr/R/pearson190.pdf/>>. Acessado em: 6 mai. 2006.
- REGAZZI, A.J. *Análise multivariada*. Viçosa: UFV, 2002. (Notas de aula - INF-766).
- ROSO, V.M.; FRIES, L.A. Componentes principais em bovinos da raça Polled Hereford à desmama e sobreano. *Rev. Soc. Bras. Zootec.*, v.24, p.728-735, 1995.
- STRAPASSON, E; VENCOSKY, R; BATISTA, L.A.R. Seleção de descritores na caracterização de germoplasma de *Paspalum* sp. por meio de componentes principais. *Rev. Bras. Zootec.*, v.29, p.373-381, 2000.