

LEANDRO BARBOSA

**UTILIZAÇÃO DE TÉCNICAS DE ANÁLISE MULTIVARIADA NA
AVALIAÇÃO DE CARACTERÍSTICAS QUANTITATIVAS DE UMA
POPULAÇÃO F2 DE SUÍNOS**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Zootecnia, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2003

“Dele, por ele e para ele são todas as coisas.
A Deus a glória por toda a eternidade!
Amém”

Rom 11, 36

Aos meus pais José Barbosa e Edna, com admiração.

Ao meu irmão Oseander, com carinho.

À Ana Andréa, minha namorada, com muito amor.

AGRADECIMENTO

Primeiramente, a Deus, por estar sempre presente em todos os momentos da minha vida.

À Universidade Federal de Viçosa e ao Departamento de Zootecnia, pela oportunidade de realização do curso.

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), pela concessão da bolsa de estudos.

Ao professor Paulo Sávio Lopes, pela orientação, pela amizade, pela atenção, pela paciência, pela confiança e pelo apoio, minha eterna admiração e gratidão.

Ao professor Adair José Regazzi, pelos ensinamentos, pelos aconselhamentos, pelos incentivos, pela amizade, pela disposição e colaboração na realização deste trabalho, minha profunda admiração e meu muito obrigado.

À professora Simone Eliza Facioni Guimarães, pelo aconselhamento, pela atenção, pelo apoio e pela amizade.

Ao professor Robledo de Almeida Torres, pelos ensinamentos, pela amizade, pelas sugestões, pela confiança e pelo apoio.

Ao professor Ricardo Frederico Euclides, pela amizade, pelas sugestões, pelo incentivo e pela orientação durante a iniciação científica.

Ao professor Lúcio Alberto de Miranda Gomide, pelas sugestões, pela atenção e pelo apoio.

À professora Carmen Silva Pereira, pelos ensinamentos, pela amizade, pela ajuda, pela confiança e pelo apoio.

Ao meu amor, Ana Andréa, pelo carinho, pela paciência, pela compreensão, pelo apoio e pela confiança a mim dedicados.

A toda a minha família, pela formação do meu caráter.

Aos amigos da pós-graduação, Lindemberg, Fernanda, Marcelinho, Peloso, Kecya, Guilherme, Aldrin, Claudinho, Jaime, Ricardinho, Policarpo, Urbano, Vicente, Gustavo, Raquel, Mirela, Marcus Vinícius, Elizangela, José Pereira, Gisele, Adriana, Gelson, Kênia, Alex, Carlos, Danielle, Fausto, Fred, Jane, Kleibe e Samuel, pela amizade, pelos trabalhos que realizamos e pelo agradável convívio.

Ao amigo Rodolphinho, pela amizade, pelo agradável convívio, pelo companheirismo, pela atenção e ajuda imprescindível na iniciação científica.

Aos amigos de longa data, Adhemar, Geléia, Gutinha, Barney, Bilica e Rufião, pela amizade e pelo agradável convívio.

Aos demais professores, funcionários e alunos do Departamento de Zootecnia (DZO), pelo companheirismo, pelo incentivo, pela amizade e por fazer do DZO um ambiente agradabilíssimo de trabalho.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

BIOGRAFIA

LEANDRO BARBOSA, filho de José das Virgens Barbosa e Edna Maria Mota Barbosa, nasceu em São Paulo, Estado de São Paulo, em 13 de janeiro de 1978.

Em março de 1996, iniciou o curso de Zootecnia pelo Departamento de Zootecnia da Universidade Federal de Viçosa, em Viçosa, MG, onde foi bolsista de iniciação científica por alguns anos, acompanhando e conduzindo pesquisas no Programa de Melhoramento Genético de Aves do Departamento de Zootecnia da UFV.

Em agosto de 2001, graduou-se em Zootecnia pela Universidade Federal de Viçosa.

Em agosto de 2001, iniciou o curso de Mestrado em Zootecnia na Universidade Federal de Viçosa, na área de Melhoramento Genético Animal.

Em 18 de julho de 2003, submeteu-se ao exame final de defesa de tese para obtenção do título de *Magister Scientiae* em Zootecnia, na Universidade Federal de Viçosa.

ÍNDICE

RESUMO.....	ix
ABSTRACT.....	xi
1 - INTRODUÇÃO	1
2 - REVISÃO DE LITERATURA.....	3
2.1. Componentes Principais	4
2.2. Correlações Canônicas.....	12
3. REFERÊNCIAS BIBLIOGRÁFICAS	16

CAPÍTULO I

Avaliação de Características Quantitativas de Suínos por meio de Componentes Principais	
RESUMO.....	19
ABSTRACT.....	21
INTRODUÇÃO	23
MATERIAL E MÉTODOS	24
Características Avaliadas.....	24
Características de desempenho	25
Características de carcaça.....	25
Características de qualidade da carne.....	26
Metodologia Estatística	26
Análise de Componentes Principais.....	27
Análise dos Dados.....	30
RESULTADOS E DISCUSSÃO	31
Características de desempenho	31
Características de carcaça.....	36
Características de qualidade	45

CONCLUSÕES	51
LITERATURA CITADA	52

CAPÍTULO II

Estudo da Associação entre Características Quantitativas de Suínos por meio de Correlação Canônica

RESUMO.....	54
ABSTRACT.....	56
INTRODUÇÃO	58
MATERIAL E MÉTODOS	59
Características Avaliadas.....	59
Características de desempenho	60
Características de carcaça.....	60
Características de qualidade da carne.....	60
Metodologia Estatística	61
Análise de Correlação Canônicas	61
Teste de significância	63
Análise dos Dados.....	64
RESULTADOS E DISCUSSÃO	65
Características de desempenho e carcaça	65
Características de desempenho e qualidade da carne	69
Características de carcaça e qualidade da carne.....	70
CONCLUSÕES	76
LITERATURA CITADA	77
4. RESUMO E CONCLUSÕES	78

RESUMO

BARBOSA, Leandro, M.S., Universidade Federal de Viçosa, julho de 2003. **Utilização de técnicas de análise multivariada na avaliação de características quantitativas de uma população F2 de suínos.** Orientador: Paulo Sávio Lopes. Conselheiros: Adair José Regazzi e Simone Eliza Facioni Guimarães.

Este estudo objetivou reduzir a dimensionalidade dos dados de 844 animais de uma população F2 de suínos por meio de componentes principais, bem como avaliar a associação entre as características de desempenho, de carcaça e de qualidade da carne, utilizando correlação canônica. Na redução da dimensionalidade, observou-se que, para as características de desempenho, de carcaça e de qualidade da carne, respectivamente 6, 17 e 4 componentes principais, obtidos a partir da matriz de correlação, apresentaram variância inferior a 0,7 (autovalor inferior a 0,7), sugerindo respectivamente 6, 17 e 4 variáveis para descarte, por apresentarem maiores coeficientes, em valor absoluto, a partir do último componente principal. A razão para isso é que variáveis altamente correlacionadas com os componentes de menor variância representam variação praticamente insignificante. Com base nesses resultados, recomendam-se as seguintes variáveis para serem mantidas em experimentos futuros: tamanho da leitegada ao nascimento, número de tetos, peso aos 77 dias de idade, consumo de ração e conversão alimentar dos 77 aos 105 dias de idade (desempenho); idade ao abate, peso da banda direita, comprimento da carcaça pelo Método Brasileiro de Classificação de Carcaça, maior espessura de toucinho na região da copa, na linha dorso-lombar, profundidade de lombo, pesos do baço, do coração, da banda direita resfriada, do pernil, da copa, da paleta, do carré, das costelas, da papada, do filezinho e do rim (carcaça); e pH medido aos 45 minutos e 24 horas *post-mortem*, luminosidade, gordura intramuscular, perda por

cozimento e índice de vermelho (qualidade da carne). No estudo da associação entre os grupos de variáveis observou-se que aqueles com características de desempenho e de carcaça não são considerados independentes, pois foram fortemente correlacionados, visto que a correlação do primeiro par canônico foi de 0,78. Os dois primeiros pares canônicos foram estatisticamente relevantes ($P < 0,01$ e $P < 0,05$, respectivamente). No exame das variáveis que compõem o primeiro par canônico, para as características de desempenho, houve predomínio absoluto da variável peso aos 77 dias de idade; em se tratando das características de carcaça, a variável idade ao abate teve predomínio absoluto. No segundo par canônico, para as características de desempenho, houve predomínio absoluto da variável número de tetos; para as características de carcaça, a variável peso das costelas teve a maior contribuição, juntamente com a variável peso da banda direita. Para os grupos características de desempenho e características de qualidade da carne, observou-se que existe independência, pois a correlação do primeiro par canônico foi de 0,23, sendo estatisticamente irrelevante ($P > 0,05$). Para os grupos características de carcaça e características de qualidade da carne, observou-se que não existe independência, sendo os três primeiros pares canônicos estatisticamente relevantes ($P < 0,01$). No exame das variáveis que compõem o primeiro par canônico, para as características de carcaça, houve predomínio da variável peso do filezinho, seguida, principalmente, da variável comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça; no caso das características de qualidade da carne, houve predomínio absoluto da variável luminosidade. No que diz respeito às variáveis que compõem o segundo par canônico, para as características de carcaça, houve maior predomínio da variável peso do carré, seguida principalmente das variáveis comprimento da carcaça pelo Método Brasileiro de Classificação de Carcaça, maior espessura de toucinho na região da copa, na linha dorso-lombar, e pesos da paleta e da banda direita resfriada; para as características de qualidade da carne, a variável gordura intramuscular teve predomínio absoluto. No exame das variáveis que compõem o terceiro par canônico, para as características de carcaça, houve maior predomínio da variável peso da banda direita resfriada; para as características de qualidade da carne, a variável perda por cozimento teve maior predomínio, seguida, principalmente, da variável índice de vermelho.

ABSTRACT

BARBOSA, Leandro, M.S., Universidade Federal de Viçosa, July 2003. **Use of Multivariate Analysis to Evaluate Quantitative Traits of a F2 Swine Population.** Adviser: Paulo Sávio Lopes. Committee Members: Adair José Regazzi e Simone Eliza Facioni Guimarães.

This study aimed to reduce the data dimension of a F2 swine population of 844 animals using principal components, as well as evaluating the association among the performance, carcass and meat quality traits, using canonical correlation. In performance, carcass e meat quality traits it was observed that, 6, 17 and 4 principal components respectively, showed variation lower than 0.7; suggesting that, respectively, 6, 17 and 4 variables could be discarded. The reason is that when variable are highly correlated with the principal components of smaller variance, their variation is practically insignificant. Based in those results it is recommended the use of the following variables in future experiments: litter size at birth, teat number, weight at 77 days of age, feed intake and feed:gain ratio from 77 to 105 days of age (performance traits); slaughter age, right half carcass weight, carcass length by the Brazilian Carcass Classification Method, higher backfat thickness at boston shoulder area (in the midline), loin depth, spleen weight, heart weight, cold right half carcass weight, ham weight, boston shoulder weight, picnic shoulder weight, loin (bone-in) weight, spareribs weight, jowl weight, sirloin weight and kidney weight (carcass traits); pH measured 45 minutes and 24 hours *post-mortem*, lightness, intramuscular fat, cooking loss and redness (meat quality traits). In the study of the association between sets of variables, it was observed that performance and carcass traits, are not considered independent, because they were strongly correlated, since the first canonical correlation (r) was 0.78. The first two canonical variate pairs were estatistically significant ($P < 0.01$ and $P < 0.05$,

respectively). In the performance traits, there was an absolute predominance of the weight at 77 days of age variable, in the first canonical variate pair; and in the carcass traits, slaughter age had absolute predominance. In the second canonical variate pair, there was an absolute predominance of the teat number variable (performance traits) and spareribs weight followed by the right half carcass weight variable (carcass traits). It was observed independence in the sets of performance and meat quality traits, because the first canonical correlation (r) was 0.23 ($P > 0.05$). There was no independence between the sets of carcass and meat quality traits, which the first three canonical variate pairs are statistically significant ($P < 0.01$). In the carcass traits, there was a predominance of the sirloin weight variable, followed mainly, by the carcass length by to the Brazilian Carcass Classification Method, in the first canonical variate pair; and in the meat quality traits, lightness had absolute predominance. In the second canonical variate pair, in the carcass traits, there was larger predominance of the loin (bone-in) weight variable, followed mainly, by the carcass length by the Brazilian Carcass Classification Method, higher backfat thickness at boston shoulder area (in the midline), picnic shoulder weight and cold right half carcass weight (carcass traits); intramuscular fat had absolute predominance in the meat quality traits. In the third canonical variate pair, there was larger predominance of cold right half carcass weight variable, in the carcass traits; and cooking loss had larger predominance, followed mainly, by the redness variable in the meat quality traits.

1 - INTRODUÇÃO

A eficiência da produção e a qualidade dos suínos determinam a viabilidade e a competitividade do setor suinícola. Portanto, diante de um mercado cada vez mais exigente e da competição de outros produtores de carne, o suinocultor terá que ter meios de produzir suínos a baixo custo e com a melhor qualidade possível. A utilização das melhores raças disponíveis, em conjunto com o melhoramento genético através de seleção e de sistemas de cruzamento, são os meios que os produtores dispõem para melhorar a eficiência de produção e a qualidade de seus animais (FREITAS, 1996).

A grande maioria das características de importância econômica submetidas à seleção é de natureza quantitativa. A seleção de indivíduos com base nessas características é relativamente difícil, em razão da complexidade de sua base genética e do elevado grau de influência do ambiente a que estão sujeitas. Além disso, essas características são relacionadas em magnitude e sentido variáveis, de forma que a seleção em uma provoca mudanças em outras, as quais, se não forem consideradas, podem levar a erros na avaliação dos indivíduos.

Segundo COSTA (1999), as correlações entre ganho de peso diário e espessura de toucinho (Large White = 0,31; Landrace = 0,33; e Duroc = 0,02) indicam que a associação entre essas duas características pode atrasar o progresso genético quando estas são selecionadas separadamente em programas de melhoramento genético, razão por que deverão ser selecionadas pelo uso de metodologias ou procedimentos multivariados.

Na maioria das pesquisas são avaliadas várias características, com o objetivo de estudar a influência de tratamentos sobre determinadas amostras ou populações. De maneira geral, realizam-se análises estatísticas por características, tirando, posteriormente, conclusões conjuntas comuns a todas as características, sem levar em consideração o nível de significância conjunto. Dessa forma, as correlações entre as características não são aproveitadas (SAKAGUTI, 1994).

Dentre os métodos de seleção para mais de uma característica, os que utilizam avaliação e interpretação simultânea de um conjunto de características têm apresentado, na maioria das vezes, os melhores resultados. A estatística que trata de duas ou mais variáveis simultaneamente é conhecida como multivariada.

Para muitos tipos de dados biológicos, a correlação entre as variáveis e as informações providas por análise univariada podem ser incompletas em se tratando de

um conjunto de variáveis. Assim, as técnicas de análise multivariada combinam, simultaneamente, múltiplas informações que não são possíveis de ser obtidas com o uso da análise univariada.

No entanto, para Finny (1956), citado por ABREU (1997), nem sempre um experimento planejado para ser avaliado por métodos univariados poderá ser analisado por técnicas multivariadas, pois estas também têm limitações e certas pressuposições que precisam ser respeitadas. A avaliação de diversas variáveis em um experimento não significa que a análise multivariada, se for aplicável, deva incluir todas, daí a importância de se conhecer, profundamente, o fenômeno que se está estudando.

As técnicas de análise multivariada são apropriadas para o estudo do conjunto de variáveis que contém intercorrelações; em virtude de as variáveis serem consideradas simultaneamente, são obtidas interpretações que não seriam possíveis com o uso de estatísticas univariadas (JAMES e McCULLOCH, 1990).

Entretanto, apesar de as técnicas multivariadas terem eficiência comprovada e proporcionarem enriquecimento das informações extraídas de dados experimentais, é necessária, para seu uso, a disponibilidade de recursos computacionais. No entanto, com o aumento dos recursos da informática nos últimos anos, a técnica atraiu a atenção dos pesquisadores das diversas áreas, tornando o seu emprego potencialmente grande e, conseqüentemente, o seu conhecimento indispensável (LIBERATO, 1995).

A oportunidade de resumir grandes grupos de dados, especialmente em estádios exploratórios de uma investigação, tem contribuído para o crescente interesse em métodos multivariados (LIBERATO, 1995).

Atualmente, contando com a maior disponibilidade dos recursos na área de informática, as pesquisas podem explorar melhor essas técnicas de análise. Assim, o presente trabalho teve por objetivo avaliar a utilização de algumas técnicas de análise multivariada, a fim de analisar a redução da dimensionalidade do espaço multivariado e a associação entre grupos de variáveis.

2 - REVISÃO DE LITERATURA

Rao (1952), citado por SAKAGUTI (1994), fez uma breve revisão dos trabalhos com análise multivariada. Segundo ele, o ponto de partida das pesquisas que tentavam generalizar as análises univariadas de variância para o caso de variáveis múltiplas foi dado por Wishart em 1928, estudando a distribuição amostral simultânea de variâncias e covariâncias em amostras de uma população normal multivariada. Posteriormente, Hotelling (1931) verificou a distribuição T^2 , que é uma extensão natural da distribuição de t de Student, para uma população com distribuição normal multivariada. Wilks (1932), seguindo o método da razão de verossimilhança (Neyman e Pearson, 1908, 1931; Pearson e Neyman, 1930), obteve generalizações apropriadas na análise de variância, aplicáveis a diversas variáveis.

O propósito de uma análise multivariada é obter alguma compreensão sobre a estrutura da situação e as principais características dos dados. Essa estrutura pode tomar várias formas, como ser requerida para dividir os dados em grupos ou decidir se um grupo conhecido está associado às observações. Também, pode-se ajustar algum modelo matemático aos resultados, por meio da análise multivariada, ou esta pode ser requerida para associar dois tipos de observação. Essas situações necessitam de diferentes análises, e é sempre importante conhecer exatamente quais questões serão respondidas e qual ordem de análise pode respondê-las (MARRIOT, 1974).

Por outro lado, JOHNSON e WICHERN (1988) consideram a análise multivariada um *mixed bag*, sendo difícil estabelecer um esquema de classificação para as técnicas multivariadas que seja largamente aceito e, também, indique as conveniências das técnicas. Contudo, eles enfatizam que a escolha de métodos e os tipos de análise a serem empregados são determinados, principalmente, pelos objetivos da investigação.

Existem diferentes técnicas de análise multivariada, cada qual adequada a diferentes objetivos e tipos de dados. KENDALL (1980) as classificou em: técnicas de avaliação de interdependência (componentes principais, correlação canônica, análise de agrupamento, análise de fatores e escala) e técnicas de avaliação da dependência (regressão, relação funcional, múltipla contingência e análise discriminante).

As técnicas de avaliação de interdependência fornecem o modelo estrutural mais complexo, que, em conjunto com a teoria da distribuição fornecida pelas técnicas de

avaliação de dependência, permite a formulação de testes de hipóteses, bem como avaliação de significância (KENDALL, 1980).

2.1. Componentes Principais

A técnica de componentes principais foi originalmente descrita por Pearson (1901) como um meio de ajustar hiperplanos segundo o método dos mínimos quadrados. Posteriormente, foi consolidada por Hotelling (1933, 1936) para o propósito particular de analisar estruturas de correlações. Naquela época, os cálculos da análise eram extremamente desanimadores para mais de umas poucas variáveis, pois eram realizados manualmente. No entanto, de acordo com Manly (1986), seu uso só foi difundido após os computadores eletrônicos tornarem-se amplamente disponíveis (LIBERATO, 1995).

A análise de componentes principais é uma técnica multivariada, que, segundo Kendal (1950), é de avaliação da interdependência, ou seja, estuda as relações de um conjunto de variáveis entre si (XAVIER, 1996).

A técnica de componentes principais é aplicada a um conjunto de variáveis relativas a n indivíduos. O método consiste na transformação do conjunto original de variáveis em outro conjunto - os componentes principais - de dimensões equivalentes, porém com certas propriedades importantes. O objetivo é resumir a informação contida no complexo de variáveis originais, eliminando as informações redundantes existentes em decorrência da correlação entre variáveis. O que ocorre, na realidade, é uma troca de variáveis, em que as variáveis observadas são trocadas por novas variáveis abstratas, os componentes principais, esperando-se que os primeiros componentes guardem, em si, toda informação existente nas variáveis originais (Perez – Gonzales, 1992, citado por LIBERATO, 1995).

Segundo REGAZZI (2002), a técnica de componentes principais tem como principal objetivo a redução da dimensionalidade do conjunto original de variáveis com a menor perda de informação possível, além de permitir o agrupamento de indivíduos (tratamentos, genótipos, etc.) similares, mediante exames visuais em dispersões gráficas no espaço bi ou tridimensional, de fácil interpretação geométrica.

Um dos propósitos do uso da técnica de componentes principais na pesquisa animal é o de resumir um grande conjunto de características em outro menor, de sentido biológico, além de examinar as correlações entre as características estudadas, avaliar a

importância de cada característica e promover a eliminação daquelas que contribuem pouco. Ela visa ainda construir índices que possibilitem o agrupamento de indivíduos e permitir o agrupamento de indivíduos com mais alto grau de similaridade, mediante exames visuais em dispersões gráficas no espaço bi ou tridimensional.

ABEYWARDENA (1972) recomendou a técnica de componentes principais como o método de estimação de repetibilidade em animais ou em árvores. Trabalhando com dados simulados, concluiu que o método de análise de componente principal, quando comparado com métodos que utilizam a análise de variância, é mais eficiente e mais estável, pois permite realçar, medir e eliminar certos aspectos latentes de variação que a análise de variância não pode distinguir. Esse mesmo autor afirma ainda que o método de análise de componente principal também pode fornecer um critério de seleção mais eficiente, uma vez que os elementos do vetor latente (autovetor) permitem obter uma média ponderada, superior à média aritmética, a qual pode estar viesada por condições anormais que podem ter ocorrido, de tempo em tempo, durante o período das observações.

Segundo BAKER et al. (1988), a análise dos componentes principais pode revelar relações não identificadas previamente, contribuindo para melhor interpretação dos dados. Além disso, muitas vezes, representa um método de explicação da estrutura de covariância entre as várias características medidas em uma população animal.

Na maioria das situações, o número de componentes principais necessários para extrair completamente a informação contida em grupos de variáveis é igual ao número de variáveis sob consideração. Isso é verdadeiro se a matriz de covariância (ou correlação) correspondente às variáveis for não-singular. Se o *rank* (número de autovalores não-nulos) da matriz de covariância for menor que a dimensão da matriz de covariância (isto é, o número de variáveis), então o número de componentes principais é igual ao *rank* da matriz de covariância (DUNTEMAN, 1984).

A importância de um componente se avalia por meio da porcentagem de variância que ele retém. Assim, o primeiro componente é o mais importante, pois retém a maior parte da variação total encontrada nos dados originais. Independência ou ortogonalidade (ausência de correlação) entre os componentes é propriedade útil, porque significa que esses componentes estão medindo diferentes “dimensões” dos dados (MANLY, 1986; MARDIA et al., 1997). Não há, portanto, perda de informação na transformação. Assim, se um dos componentes principais possui grande variância em

relação à variação total, então este componente principal captura a maioria da informação existente no grupo de variáveis originais.

A viabilidade de utilização dos componentes principais dependerá da possibilidade de resumir o conjunto de variáveis originais em apenas poucos componentes, o que significará ter uma aproximação do comportamento dos indivíduos oriundos de um espaço multidimensional em um espaço bi ou tridimensional. A variância, antes dispersa entre as variáveis correlacionadas, é agora atribuída, em sua quase totalidade, a um número menor de componentes. Desse modo, a análise proporcionará simplificação considerável nos cálculos estatísticos e na interpretação dos resultados, principalmente quando o número dos indivíduos avaliados for relativamente grande (CRUZ, 1990). Cada componente principal é uma combinação linear das variáveis originais independentes entre si, sendo estimado com o propósito de reter, em ordem de estimação, o máximo de informação em termos de variação total contida nos dados iniciais (CRUZ e REGAZZI, 1997).

Para DUNTEMAN (1984), a decisão de quantos componentes reter é um problema complexo e, na maioria dos casos, tem sido mesmo baseada em métodos simples e práticos, como: “plotar” os autovalores e determinar onde cessam os grandes e iniciam os pequenos (*scree plot*), ou excluir componentes cujo autovalor seja menor que a média. Em geral, os primeiros componentes principais, que acumulam pelo menos 70 a 80% da variação total, têm sido utilizados (CRUZ e REGAZZI, 1997). Entretanto, segundo MARDIA et al. (1997), os componentes principais que deveriam ser inclusos são aqueles responsáveis pelo acúmulo de pelo menos 90% da variação total.

Segundo Kendal (1980), Mardia et al. (1979) e Duterman (1984), citados por CRUZ (1990), um novo conjunto de variáveis, originado de uma análise multivariada com base em componentes principais ou em variáveis canônicas que, em ordem de estimação, retêm o máximo de informação em termos de variação total, explicará tanto melhor a variabilidade manifestada entre os indivíduos avaliados, quanto menor for o número de componentes que acumulem pelo menos 80% da variação total.

ALCALDE e NEGUERUELA (2001), utilizando técnicas de componentes principais para avaliar a influência de condições finais de armazenamento na coloração de carcaça de carneiros, verificaram que o primeiro componente principal explicava 87,64% da variação total. No entanto, os autores decidiram utilizar os dois primeiros componentes principais como um método para melhorar os resultados e fazer a representação gráfica.

SCHAFER et al. (2002), ao estudarem a relação de variáveis pós-morte envolvidas na perda por gotejamento na carne de suíno, observaram que 89% da variação total foi explicada pelos dois primeiros componentes principais, indicando uma eficiente redução da dimensionalidade.

Entretanto, DESTEFANIS et al. (2000), ao trabalharem com componentes principais nas análises químicas, físicas e sensoriais da carne de bovinos jovens (qualidade da carne), num total de 18 variáveis, constataram que 34% da variação total era explicada pelo primeiro componente principal, 54,4% pelos dois primeiros e 62,6% pelos três primeiros componentes principais. Pelo fato de os três primeiros componentes terem explicado em torno de 63% da variação total, esses autores concluíram que a técnica de componentes principais demonstra ser um procedimento muito efetivo para resumir o julgamento da qualidade da carne.

ABREU et al. (1999), ao avaliarem a capacidade de combinação de linhagens de matrizes de frango de corte, usando a técnica de componentes principais em 11 características produtivas, observaram que foram necessários os seis primeiros componentes principais para explicar mais de 80% da variação total das médias dos cruzamentos, em que o primeiro componente foi responsável por 21%, o segundo por 17%, o terceiro por 16%, o quarto por 11%, o quinto por 8% e o sexto por 8%.

Os melhores resultados são obtidos quando as variáveis originais são fortemente correlacionadas, positiva ou negativamente. Segundo MANLY (1986), nesse caso, é completamente concebível que 20 ou 30 variáveis originais possam ser adequadamente representadas por dois ou três componentes principais. Se um ou poucos componentes principais explicam a maior parte da variação em um grande grupo de variáveis, então se tem, efetivamente, redução da dimensionalidade do grupo de dados.

A análise de componentes principais é influenciada pela escala das variáveis, sendo os coeficientes destes componentes altamente dependentes das unidades de medida associadas com cada variável - esse aspecto é discutido por DUNTEMAN (1984) e MARDIA et al. (1997). É aconselhável que os caracteres estudados possuam a mesma unidade e dimensões não muito discrepantes (ou seja, variância com semelhantes ordens de magnitude). Como na maioria das vezes isso não ocorre, recomenda-se utilizar as variáveis reduzidas, ou seja, variáveis padronizadas com média igual a zero e variância igual à unidade (ou variáveis padronizadas apenas com variância igual à unidade). Isso significa que a matriz de covariância é transformada em uma matriz de correlações antes que a análise de componentes principais seja

conduzida. Caso contrário, a variável com maior variância possuirá maior peso na estimação dos componentes, sendo esse peso mais em função das unidades de medida do que da estrutura básica das observações. Assim, as combinações lineares das quantidades originais poderiam ter pouco significado, e as variáveis padronizadas e a matriz de correlações deveriam ser utilizadas (MORRISON, 1976). A padronização das variáveis assegura que todas elas possuam, inicialmente, igual peso na análise (MANLY, 1986). De acordo com IEZZONI e PRITTS (1991), se os caracteres avaliados têm diferentes dimensões, então, geralmente, os dados não possuem distribuição normal multivariada. Assim, a extrapolação de resultados para uma população maior não deverá ser realizada. Uma solução é a padronização das variáveis; entretanto, segundo DUNTERMAN (1984) e MORRISON (1976), a distribuição normal multivariada não é necessária para que a análise de componentes principais seja aplicada.

Embora a análise, formalmente, não requeira a distribuição normal multivariada, ela é mais apropriada para variáveis quantitativas contínuas. Quando a distribuição de razões e proporções é razoavelmente próxima à normal, a análise pode ser útil, mas pode não captar relações não-lineares (LIBERATO, 1995). Para dados constituídos de contagens, razões, proporções ou percentagens, pode-se, previamente à análise de componentes principais, realizar transformações para tornar sua distribuição mais apropriada, ou então utilizar um método não-linear (JAMES e McCULLOCH, 1990). STAUFFER et al. (1985) submeteram dados de percentagem à transformação arco seno da raiz quadrada da percentagem e PIMENTEL GOMES (1984) submeteu dados de contagem à transformação de raiz quadrada da percentagem, antes de realizarem a análise de componentes principais.

Para caracteres quantitativos com distribuição normal multidimensional, sua informação é proporcional à variância. Se uma variável mensurada possuir maior variação que outras, então ela poderá dominar o cálculo dos componentes principais. Na prática, é desejável que todas as variáveis tenham a mesma influência no cálculo dos componentes principais. Assim, se as variâncias são igualadas, por padronização, então cada caráter contribui com igual informação. Embora a padronização equalize o tamanho de diferentes caracteres, ela não remove a diferença de tamanho entre indivíduos (SNEATH e SOKAL, 1973). Segundo LIBERATO (1995), não é necessária a padronização de variáveis binárias, porém é recomendada para variáveis ordinais e quantitativas, quando se procede à análise de componentes principais, ou quando se

estima a distância euclidiana. Não é necessária a padronização de variáveis na estimativa da distância de Mahalanobis e na análise canônica.

Uma vantagem adicional da técnica dos componentes principais é a de possibilitar uma avaliação da importância de cada caráter estudado sobre a variação total disponível entre os indivíduos avaliados. O interesse nesta avaliação reside na possibilidade de se descartarem caracteres que contribuem pouco para a discriminação no material avaliado, reduzindo, dessa forma, mão-de-obra, tempo e custo despendidos.

Diferentes métodos estatísticos, que procuram fazer a seleção de subconjuntos de variáveis, foram propostos e utilizados por vários pesquisadores, com eficiências variadas. Contudo, a eleição de determinada técnica, dentre as disponíveis, vai depender, basicamente, da estrutura dos dados e dos objetivos da investigação.

JOLLIFFE (1972, 1973) avaliou, em dados simulados e reais, oito métodos de descarte de variáveis, em que dois estavam relacionados com regressão múltipla, quatro com componentes principais e dois com técnicas de conglomeração. Concluiu que os métodos possuíam eficiência equivalente em rejeitar variáveis. Essas técnicas multivariadas são sensíveis em discriminar as variáveis menos importantes que se caracterizam pela redundância, por já estarem fortemente correlacionadas com outras variáveis presentes na análise ou pela sua invariância.

A redução no número de características, com a eliminação daquelas que menos contribuem para a variação total, deve facilitar as interpretações sem causar perda considerável de informação.

Diversas proposições sobre o descarte de variáveis têm sido apresentadas por BEALE et al. (1967) e JOLLIFFE (1972, 1973). Na análise de componentes principais, quando se utiliza matriz de correlação, tem-se adotado o critério de avaliação da importância dos caracteres a partir dos coeficientes de ponderação (autovetores) das variáveis associados aos últimos componentes, que, por estimação, retêm proporção mínima da variação total (MARDIA et al., 1997). Com base no princípio de que a importância, ou variância, dos componentes principais decresce do primeiro para o último, tem-se que os últimos componentes são responsáveis pela explicação de uma fração muito pequena da variância total. Assim, a variável que domina (aquela que possui o maior coeficiente) no componente principal de menor autovalor (menor variância) deve ser a menos importante para explicar a variância total e, portanto, passível de descarte (PEREIRA, 1989; REGAZZI, 2002). A razão é que variáveis

altamente correlacionadas com os menores componentes principais representam variação praticamente insignificante (DUNTEMAN, 1984).

Quando, em um componente de menor variância, o maior coeficiente de ponderação está associado a um caráter previamente descartado, tem-se optado por não fazer nenhum outro descarte com base nos coeficientes daquele componente, mas prosseguir a identificação da importância relativa dos caracteres no outro componente de variância imediatamente superior (REGAZZI, 2002).

Uma variação do método de descarte de variáveis consiste em refazer a análise de componentes principais com as variáveis remanescentes, a cada estágio de descarte da variável associada com o componente de menor autovalor. Esse processo é conduzido até que todos os componentes principais tenham autovalores altos (MARDIA et al., 1997).

CURY (1993) realizou o descarte das variáveis por meio da técnica dos componentes principais, descartando a variável com maior coeficiente do componente com menor autovalor em sucessivas análises, isto é, a cada descarte repetiu a análise considerando o novo subconjunto de variáveis.

Segundo ABREU (1997), as técnicas padrões da análise multivariada envolvem manipulações de matrizes de ordem igual ao número de variáveis, e um modo de evitar problemas é fazer, preliminarmente, uma análise dos componentes principais e descartar as variáveis de maior importância nos últimos componentes. A análise é, então, realizada com as variáveis restantes, o que evita os problemas com as matrizes mal condicionadas.

JOHNSON e WICHERN (1998) recomendam cautela na interpretação dos resultados de qualquer método de seleção de variáveis. Para eles, independentemente do critério de seleção adotado, não há garantia de que um subconjunto selecionado seja realmente o melhor; eles atentam para o fato de que um subconjunto selecionado em uma amostra, com base em algum dos critérios existentes, pode ser insuficiente para análise em amostras futuras. Além disso, os problemas associados com procedimentos de seleção de variáveis são aumentados com a presença de grandes correlações entre as variáveis ou entre as combinações lineares de variáveis.

As variáveis selecionadas e descartadas devem apresentar correlações significativas entre si, ou seja, as variáveis descartadas devem ser redundantes (ser responsáveis pelo mesmo tipo de correlação já contida nas variáveis selecionadas). Por outro lado, as variáveis selecionadas devem ter baixas correlações entre si. Desse modo,

cada variável será responsável por um tipo de informação biológica exclusiva, e a ação conjunta delas será complementar para a descrição geral dos indivíduos ou populações estudados.

Uma crítica que se pode fazer ao uso dos componentes principais é o fato de ser uma combinação linear, cujos coeficientes de ponderação não são estabelecidos na ordem direta da importância dos caracteres (ABREU, 1997). Nem sempre os componentes principais correlacionam-se com as características de maior importância na avaliação das linhagens (ABREU et al., 1999).

Por imposição, os vetores característicos são ortogonais, implicando que a soma das variâncias de cada componente corresponde ao somatório das variâncias de cada variável, denominada variância total, e que a covariância entre os componentes é nula (PEREIRA, 1989).

O coeficiente de uma variável individual representa a contribuição desta variável para o componente principal. Esse valor depende das outras variáveis que são incluídas no estudo. Se diferentes conjuntos de variáveis são incluídos, os coeficientes serão diferentes. Os coeficientes dos componentes principais são também denominados elementos de autovetores, e a variância associada a cada componente principal é denominada autovalor. A correlação entre a variável original padronizada e o componente principal é denominada *loading* e indica o quanto uma variável poderia substituir o componente principal (JAMES e McCULLOCH, 1990).

Outra característica da técnica de componentes principais é a de trabalhar com médias amostrais, ou poder ser usada nas situações em que não há repetições de dados. Os componentes principais dependem somente da matriz de covariâncias entre as variáveis, podendo ser muito sensíveis a observações atípicas ou indicar associação quando, de fato, existe pouca. De acordo com JOHNSON e WICHERN (1998), observações atípicas podem ter considerável efeito nas correlações e, conseqüentemente, nos componentes principais. Em razão de os coeficientes dos componentes serem altamente sujeitos à variabilidade amostral, JAMES e McCULLOCH (1990) relataram que não se deveria dar muita ênfase aos valores exatos dos coeficientes.

O estabelecimento de grupos de similaridade com base na simples inspeção visual da dispersão é questionável e pode, às vezes, não ser tão claro na visualização do grau de similaridade dos indivíduos estudados. Há casos em que a análise não consegue resumir o complexo de informações das variáveis originais, ou seja, os dois ou três

primeiros componentes principais contêm porcentagem da variância total insuficiente para avaliação segura da similaridade. Nessas situações, algumas vezes é utilizado o recurso da análise de agrupamento, conjugada à análise de componentes principais. Com os indivíduos simulando pontos em um espaço multidimensional, como determinado por seus escores em relação aos componentes principais, a distância geométrica entre quaisquer dois indivíduos pode ser calculada.

Segundo SOUZA (1989), a grande importância do conhecimento da técnica dos componentes principais reside no fato de ela constituir um procedimento básico, do qual derivam vários outros métodos de análise de dados multivariados.

2.2. Correlações Canônicas

Correlação canônica é uma técnica utilizada para analisar as relações entre dois conjuntos de variáveis, na qual se estima a máxima correlação entre as combinações lineares das variáveis. O número de correlações canônicas é igual ao menor número de variáveis que constitui um dos conjuntos, e sua magnitude sempre decresce com a ordem em que são estimadas. Entretanto, o primeiro coeficiente é sempre maior ou igual, em valor absoluto, a qualquer coeficiente de correlação, simples ou múltipla, entre as variáveis do primeiro e do segundo grupo (REGAZZI, 2002).

A correlação canônica é uma generalização da correlação e regressão, que é aplicável quando os atributos de um único grupo de objetos podem ser divididos naturalmente em dois conjuntos (JAMES e McCULLOCH, 1990). De um ponto de vista, a análise de correlação canônica é uma extensão da regressão múltipla. No caso da regressão múltipla, existe somente uma variável dependente. Por outro lado, a situação da correlação canônica envolve, no mínimo, duas variáveis do “lado dependente” e, no mínimo, duas variáveis do “lado independente”. Embora alguns pesquisadores possam considerar um grupo de variáveis independentes e outro como variáveis dependentes, o método não faz essas distinções entre os dois grupos (LIBERATO, 1995).

Segundo JOHNSON e WICHERN (1998), a análise de correlação canônica busca identificar e quantificar a associação entre dois conjuntos de variáveis.

A correlação canônica mede a força da associação entre dois conjuntos de variáveis. O método consiste em encontrar um vetor de coeficientes para cada um dos dois grupos de variáveis, tal que a correlação entre as duas combinações lineares seja máxima. Determina-se o primeiro par de combinações lineares que possuem a maior

correlação. A seguir, determina-se outro par de combinações que possuem a segunda maior correlação entre todos os pares não correlacionados com o par selecionado inicialmente, e assim sucessivamente. As combinações lineares são denominadas variáveis canônicas, sendo suas correlações chamadas de correlações canônicas.

A análise de correlação canônica talvez seja facilmente entendida e interpretada se as relações das variáveis em cada um dos dois conjuntos puderem ser transformadas em um menor número de variáveis poucos correlacionadas (KHATTREE e NAIK, 2000).

A análise de correlação canônica possui certas propriedades similares às da análise de componentes principais; contudo, esta última considera as interações dentro de um grupo de variáveis, enquanto a análise de correlação canônica considera a relação entre dois grupos (MARDIA et al., 1997).

Se as variáveis originais são padronizadas com média 0 e variância 1, então os coeficientes de correlação não têm unidade de medida, e as correlações canônicas são interpretadas em termos de variáveis padronizadas (JOHNSON e WICHERN, 1998).

Segundo KHATTREE e NAIK (2000), freqüentemente se está interessado na execução de análise canônica em conjunto de variáveis padronizadas que possua média 0 e desvio-padrão 1. Isso porque muitas variáveis são medidas em diferentes escalas e, conseqüentemente, são trazidas para igual fundamento; neste caso, a matriz de variâncias e covariâncias populacionais para os dois conjuntos de dados será igual à matriz de correlação. A análise de correlação canônica executada a partir dos dados padronizados resultará em diferentes coeficientes, os quais são chamados de coeficientes de correlação canônica padronizados.

É de interesse determinar a correlação para cada variável canônica do conjunto com a variável individual dentro do conjunto. Essas correlações indicam a contribuição de cada variável original para compor a variável canônica, ajudando na interpretação desta (TIMM, 1975).

Além disso, usando a correlação dentro de um conjunto, para melhor entender as variáveis canônicas, pode-se, também, examinar a relação entre a variável canônica num conjunto e uma variável individual no outro conjunto (TIMM, 1975).

Segundo JOHNSON e WICHERN (1998), se as variáveis canônicas são “bons” resumos de seus respectivos conjuntos de variáveis, as associações entre estas podem ser descritas em termos de variação canônica e suas correlações. Isso é útil para se ter um resumo da extensão das medidas para a qual a variação canônica informa a variação

de seus respectivos conjuntos de dados; é de utilidade também na ocasião de calcular a proporção da variância em um conjunto de variáveis explicado pelas variáveis canônicas do outro conjunto.

Entretanto, correlações entre variáveis canônicas e variáveis originais são interpretadas com precaução. Elas fornecem somente informação univariada, não indicando como as variáveis contribuem conjuntamente com a análise canônica. Por isso, muitos pesquisadores preferem avaliar a contribuição das variáveis originais diretamente nos coeficientes padronizados (JOHNSON e WICHERN, 1998).

Se o tamanho da amostra for grande, então é desejável que ela seja dividida ao meio. A primeira metade dela pode ser usada para construir e avaliar a variável canônica e as correlações canônicas de toda a amostra. Os resultados podem ser “validados” com o restante das observações. As mudanças, caso existam, na natureza da análise canônica poderão prover uma indicação das variabilidades amostrais e da estabilidade das conclusões (JOHNSON e WICHERN, 1998).

Os testes de hipóteses, e algum teste de significância, executados em análise de correlação canônica devem ser vistos com cuidado. Em pequenas amostras, talvez muitos testes assintóticos sejam inapropriados em determinadas situações (KHATTREE e NAIK, 2000). Observação semelhante também foi feita por JOHNSON e WICHERN (1998).

Segundo DUNTERMAN (1984), o método é usualmente utilizado com variáveis quantitativas contínuas, sendo necessário assumir a existência de normalidade multivariada quando testes de significância estatística são utilizados. A análise também pode ser utilizada quando há mistura de variáveis contínuas e qualitativas, ou se todas as variáveis são qualitativas (JOHNSON e WICHERN, 1998). No entanto, segundo JAMES e McCULLOCH (1990), o procedimento não é ideal para variáveis qualitativas.

DUFOUR et al. (2000), ao estudarem o estado de maturação de queijo, aplicando espectro de fluorescência e infravermelho, observaram maior discriminação, para quatro grupos de maturação, quando se utilizou a análise de correlação canônica do que quando se empregou a análise com componentes principais. Esses mesmos autores observaram que, na prática, a correlação canônica não poderia ser diretamente aplicada nos dados de fluorescência espectral, devido à alta colinearidade entre as variáveis. Como consequência, foram feitas análises de componentes principais para cada conjunto de dados, e a análise de correlação foi feita na matriz de escores dos componentes principais. Nesse mesmo trabalho, dois conjuntos de variáveis

(infravermelho e fluorescência espectral) em queijos foram estudados; observou-se que os três primeiros pares canônicos apresentaram correlação canônica ao quadrado (r^2) igual a 0,89, 0,64 e 0,35, que foram significativos a 1% de probabilidade pelo teste F. Essas correlações indicaram que as variáveis canônicas demonstram descrição comum para ambas as amostras: fluorescência e infravermelho.

DUFOUR et al. (2001), ao utilizarem análise de correlação canônica em queijo por meio do método de fluorescência espectral para medir o vínculo entre dois grupos de variáveis medidas numa mesma amostra, observaram alta correlação, sendo os coeficientes de correlação canônica ao quadrado (autovalores) 1 e 2 iguais a 0,93 e 0,80, respectivamente.

MASOOD et al. (2001) constataram alta correlação entre características de qualidade final do trigo (variáveis dependentes) e outras características físico-químicas e bioquímicas (variáveis independentes) em 30 cultivares de trigo, quando utilizaram análise de correlação canônica.

VADIVELLOO (1995), ao estudar os fatores que contribuem para o valor nutritivo da palha de arroz em diferentes variedades, para utilização na alimentação de bovinos, observou que as correlações canônicas entre composição e digestibilidade em matéria seca da palha inteira com morfologia e características agronômicas foram pequenas e não-significativas ($p > 0,05$).

3 - REFERÊNCIAS BIBLIOGRÁFICAS

- ABEYWARDEMA, V. An application of principal component analysis in genetics. **Jornal of Genetics**, v.61, p.27-51, 1972.
- ABREU, V.M.N. **Estudo da capacidade de combinação de linhagens de matrizes de frango de corte, utilizando técnicas uni e multivariada**. Viçosa, MG: UFV, 1997, 112p. Dissertação (Doutorado em Genética e Melhoramento) – Universidade Federal de Viçosa, 1997.
- ABREU, V.M.N.; SILVA, M.A.; CRUZ, C.D.; FIGUEIREDO, E.A.P.; ABREU, P.G. Capacidade de combinação de características produtivas de linhagens de matrizes de frangos de corte, usando a técnica de componentes principais. **Revista Brasileira de Zootecnia**, v.28, p.250-257, 1999.
- ALCADE, M.J.; NEGUERUELA, A.I. The influence of final conditions on meat colour in light lamb carcasses. **Meat Science**, v.57, p.117-123, 2001.
- BAKER, J.F.; STEWART, T.S.; LONG, C.R.; CARTWRIGHT, T.C. Multiple regression and principal components analysis of puberty and growth in cattle. **Journal Animal Science**, v.66, p.2147-2158, 1988.
- BEALE, E.M.L.; KENDALL, M.G.; MANN, D.W. The discarding of variables in multivariate analysis. **Biometrika**, v.54, p.357-365, 1967.
- COSTA, A.R. **Avaliação genética e valores econômicos de características de desempenho em suínos**. Viçosa, MG: UFV, 1999, 50p. Dissertação (Mestrado em Zootecnia) – Universidade Federal de Viçosa, 1999.
- CRUZ, C.D. **Aplicação de algumas técnicas multivariada no melhoramento de plantas**. Piracicaba, SP: ESALQ/USP, 1990, 188p. Dissertação (Doutorado em Genética e Melhoramento) – Escola Superior de Agricultura ‘Luiz de Queiroz’, 1990.
- CRUZ, C.D.; REGAZZI, A.J. **Modelos biométricos aplicado ao melhoramento genético**. 2. ed., Viçosa, MG: UFV, Impr. Univ., 1997. 390p.
- CURY, R. **Dinâmica evolutiva e caracterização de germoplasma de mandioca (Manihot esculenta Crantz) na agricultura autóctone do sul do Estado de São Paulo**. Piracicaba, SP: ESALQ/USP, 1993. 103p. Dissertação (Tese MS) - Escola Superior de Agricultura ‘Luiz de Queiroz’, 1993.
- DESTEFANIS, G.; BARGE, M.T.; BRUGIAPAGLIA, A.; TASSONE, S. The use of principal component analysis (PCA) to characterize beef. **Meat Science**, v.56, p.255-259, 2000.
- DUFOUR, E.; MAZEROLLES, G.; DEVAUX, M.F.; DUBOZ, G.; DUPLOYER, M.H.; MOUHOUS RIOU, N. Phase transition of triglycerides during semi-hard cheese ripening. **International Dairy Journal**, v.10, p.81-93, 2000.

- DUFOUR, E.; DEVAUX, M.F.; FORTIER, P.; HERBERT, S. Delineation of the structure of soft cheeses at the molecular level by fluorescence spectroscopy - relationship with texture. **International Dairy Journal**, v.11, p.465-473, 2001.
- DUNTEMAN, G.H. **Introduction to multivariate analysis**. Beverly Hills, Sage: 1984. 237p.
- FREITAS, R.T.F. **Estudo de divergência genética de suínos em cruzamentos, utilizando-se técnicas de análise multivariada**. Viçosa, MG: UFV, 1996, 152p. Dissertação (Doutorado em Genética e Melhoramento) – Universidade Federal de Viçosa, 1996.
- IEZZONI, A.; PRITTS, M.P. Applications of principal component analysis to horticultural research. **HortScience**, v.26, p.334-338, 1991.
- JAMES, F.C.; McCULLOCH, C.E. 1990. Multivariate analysis in ecology and systematics: Panacea or pandora's box? **Annual Review Ecology Systematic**, v.21, p.129-166, 1990
- JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. 2. ed., Englewood Cliffs, Prentice Hall, 1988. 607p.
- JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. 4. ed., Englewood Cliffs, Prentice Hall, 1998. 816p.
- JOLLIFFE, I.T. Discarding variables in a principal component analysis. I. Artificial data. **Applied Statistics**, v.21, p.160-173, 1972.
- JOLLIFFE, I.T. Discarding variables in a principal component analysis. II. Real data. **Applied Statistics**, v.22, p.21-31, 1973.
- KENDAL, M. **Multivariate analysis**. Charles Griffin an Co, Ltd., 1980. 210p.
- KHATTREE, R.; NAIK, D.N. **Multivariate data reduction and discrimination with SAS software**. SAS Institute Inc., Cary, NC, USA, 2000. 558p.
- LIBERATO, J.R. **Aplicação de técnica de análise multivariada em fitopatologia**. Viçosa, MG: UFV, 1995, 144p. Dissertação (Mestrado em Fitopatologia) – Universidade Federal de Viçosa, 1995.
- MANLY, B.F.J. **Multivariate statistical methods: A primer**. London, Chapman and Hall, 1986. 159p.
- MARDIA, K.V.; KENT, J.T.; BIBBY, J.M. **Multivariate analysis**. 6. ed., London, Academic press, 1997. 518p.
- MORRISON, D.F. **Multivariate statistical methods**. 2. ed., New York, McGraw Hill, 1976. 415p.

- MASOOD, S.B.; FAQIR, M.A.; DICK, J.V.Z.; MUMTAZ, S. Development of predictive models for end-use quality of spring wheats through canonical analysis. **International Journal of Food Science and Technology**, v.36, p.433-440, 2001
- PIMENTEL GOMES, F. **A estatística moderna na pesquisa agropecuária**. Piracicaba, Potafos, 1984. 160p.
- PEREIRA, A.V. **Utilização de análise multivariada na caracterização de germoplasma de mandioca. (Manihot esculenta Crantz)**. Piracicaba: ESALQ, 1989, 180p. Dissertação (Tese DS) – Escola Superior de Agricultura “Luiz de Queiroz”, 1989.
- REGAZZI, A.J. **Análise multivariada**. Viçosa: Universidade Federal de Viçosa, 2002. (INF-766) (notas de aula).
- SAKAGUTI, E.S. **Utilização de técnicas de análise multivariada na avaliação de cruzamento dialélicos em coelhos**. Viçosa, MG: UFV, 1994, 170p. Dissertação (Mestrado em Genética e Melhoramento) – Universidade Federal de Viçosa, 1994.
- SCHAFER, A.; ROSENVOLD, K.; PURSLOW, P.P.; ANDERSEN, H.J.; HENCKEL, P. Physiological and structural events post mortem of importance for drip loss in pork. **Meat Science**, v.61, p.355-366, 2002.
- SNEATH, P.H.A.; SOKAL, R.R. **Numerical taxonomy. the principles and practice of numerical classification**. San Francisco, W.H. Freeman, 1973. 573p.
- SOUZA, A.L. **Análise multivariada para manejo de florestas naturais: alternativas de produção sustentada de madeira para serraria**. Curitiba, PR: UFPR, 1989, 255p. Dissertação (Doutorado em Manejo Florestal) – Universidade Federal do Paraná, 1989.
- STAUFFER, D.F.; GARTON, E.O.; STEINHORST, R.K. A comparison of principal components from real and random data. **Ecology**, v.66, p.1693-1698, 1985.
- TIMM, H.N. **Multivariate analysis with application in education and psychology**. Monterey, University of Pittsburg, 1975. 689p.
- VADIVELLOO, J. Factors contributing to varietal differences in the nutritive value of rice straw. **Animal Feed Science and Technology**, v.54, p.45-53, 1995.
- XAVIER, A. **Aplicação da análise multivariada da divergência genética no melhoramento de eucalyptus ssp**. Viçosa: UFV, 1996, 128p. Dissertação (Doutorado em Engenharia Florestal) – Universidade Federal de Viçosa, 1996.

CAPÍTULO I

Avaliação de Características Quantitativas de Suínos por meio de Componentes Principais

Resumo: Dados de 844 animais de uma população F2 de suínos foram utilizados para avaliar a redução da dimensionalidade do espaço multivariado, por meio de componentes principais. Foram consideradas as seguintes características de desempenho: tamanho da leitegada ao nascimento (TLN), tamanho da leitegada à desmama (TLD), número de tetos (NT), pesos ao nascer (PN), aos 21 (P21), aos 42 (P42), aos 63 (P63) e aos 77 dias de idade (P77), ganho de peso médio diário (GPD), consumo de ração (CR) e conversão alimentar (CA) dos 77 aos 105 dias de idade; de carcaça: idade ao abate (IDA), peso da banda direita (PBDIR), rendimento de carcaça com pés e cabeça (RCARC), comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça (MBCC), comprimento de carcaça pelo Método Americano (MLC), maior espessura de toucinho na região da copa, na linha dorso-lombar (SH), espessura de toucinho imediatamente após a última costela, na linha dorso-lombar (UC), espessura de toucinho entre a última e a penúltima vértebra lombar, na linha dorso-lombar (UL), espessura de toucinho medida imediatamente após a última costela, a 6,5 cm da linha dorso-lombar (P2), área de olho de lombo (AOL), profundidade de lombo (PROFLOMB), espessura do bacon (EBACON), pesos do pulmão (PULMAO), do coração (COR), do fígado (FIG), do baço (BACO) e do rim (RIM), comprimento total do intestino delgado (INTEST), pesos da banda direita resfriada (PBDIRRES), do pernil (PP), do pernil sem pele e sem gordura (PPL), da copa (PCOPA), da copa sem pele e sem gordura (PCOPAL), da paleta (PPA), da paleta sem pele e sem gordura (PPAL), do carré (PC), do lombo (PL), do bacon (PB), das costelas (PCOS), da papada (PAPADA), da cabeça (CABEÇA), do filezinho (PF) e da banha rama (PBR); e de qualidade da carne: pH medido 45 minutos e 24 horas *post-mortem* (pH₄₅, pH_u, respectivamente), perda de peso por gotejamento (GOTEJ), perda de peso por cozimento (COZ), perda de peso total (PTOT), gordura intramuscular (GORINT), maciez objetiva (MACIEZ), luminosidade (L), índice de vermelho (a) e índice de amarelo (b). Para as características de desempenho, de carcaça e de qualidade da carne, respectivamente 6, 17 e 4 componentes principais, obtidos a partir da matriz de correlação, apresentaram variância inferior a 0,7 (autovalor inferior a 0,7). Isso sugere, respectivamente, 6, 17 e 4 variáveis

para descarte, por apresentarem maiores coeficientes, em valor absoluto, a partir do último componente principal. A razão para isso é que variáveis altamente correlacionadas com os componentes de menor variância representam variação praticamente insignificante. Com base nesses resultados, recomendam-se as seguintes variáveis para serem mantidas em experimentos futuros: TLN, NT, P77, CR e CA (desempenho); IDA, PBDIR, MBBC, SH, PROFLOMB, BACO, COR, PBDIRres, PP, PCOPA, PPA, PC, PCOS, PAPADA, PF e RIM (carcaça); e pH₄₅, pH_u, L, a, GORINT e COZ (qualidade da carne).

Palavras-chave: análise multivariada, carcaça, correlação, descarte de variáveis, desempenho, qualidade da carne

Evaluation of Swine Quantitative Traits Using Principal Components

Abstract: Records of 844 animals of a swine population were used to discard variables in a principal components analysis. The following traits were used: 1 - performance: litter size at birth (BL), litter size at weaning (WL), teat number (TN), birth weight (BW), weight at 21 days of age (W21), at 42 (W42), at 63 (W63), at 77 days of age (W77), average daily gain (ADG), feed intake (FI) e feed:gain ratio (FGR) from 77 to 105 days of age; 2 - carcass: slaughter age (SA), right half carcass weight (HCW), carcass yield with feet and head (CY), carcass length by the Brazilian Carcass Classification Method (BLC), carcass length by the American Carcass Classification Method (ALC), higher backfat thickness at boston shoulder area, in the midline (SH), backfat thickness after last rib, in the midline (UC), backfat thickness between last 1st-2nd lombar vertebrae, in the midline (UL), backfat thickness after last rib, at 6.5 cm from the midline (P2), loin eye area (LEA), loin depth (LD), bacon depth (BD), lungs weight (LW), heart weight (HW), liver weight (LIW), spleen weight (SPLW), kidney weight (KW), intestine length (IL), cold right half carcass weight (CCW), ham weight (HAW), skinless and fatless ham weight (SFHW), boston shoulder weight (BSW), skinless and fatless boston shoulder weight (SFBSW), picnic shoulder weight (PSW), skinless and fatless picnic shoulder weight (SFPSW), loin (bone-in) weight (BLW), loin weight (LW), bacon weight (BAW), spareribs weight (SPW), jowl weight (JW), head weight (HEW), sirloin weight (SLW) and belly fat weight (BFW); and 3 - meat quality: pH evaluated at 45 minutes after slaughter (pH₄₅), pH evaluated at 24 hours after slaughter (pH_u), drip loss (DL), cooking loss (CL), total loss (TL), intramuscular fat (IMF), objective tenderness (OT), lightness (L), redness (a), yellowness (b). In performance, carcass and meat quality traits it was observed that, 6, 17 and 4 principal components respectively, showed variation lower than 0.7; suggesting that, respectively, 6, 17 and 4 variables could be discarded. The reason is that when variable are highly correlated with the principal components of smaller variance, their variation is practically insignificant. Based in those results it is recommended the use of the following variables in future experiments: BL, TN, W77, FI e FGR (performance traits); SA, HCW, BLC, SH, LD, SPLW, HW, CCW, HAW, BSW, PSW, BLW, SPW, JW, SLW, and KW (carcass traits); pH₄₅, pH_u, L, a, IMF and CL (meat quality traits).

Key Words: carcass, correlation, discarding variables, meat quality, multivariate analysis, performance

INTRODUÇÃO

A análise multivariada permite avaliar a contribuição da inclusão de algumas características quando outras estão presentes na análise. Na prática, não há limite para o número de características que podem ser obtidas nos experimentos; portanto, é relevante examinar se algumas poderiam ser desconsideradas na coleta das observações ou nas análises estatísticas.

Em muitas situações os pesquisadores tendem a avaliar maior número de características, gerando, com isso, acréscimo considerável de trabalho. No entanto, após análise inicial, pode ser possível identificar somente aquelas mais importantes, aumentando a eficiência e reduzindo o custo de experimentos futuros. Quando o número de características torna-se elevado, é possível que muitas delas contribuam pouco para a discriminação dos indivíduos avaliados, por serem relativamente invariantes entre estes, por serem redundantes em virtude de serem altamente correlacionadas ou, ainda, pelo fato de uma característica ser combinação linear de outras. Essa situação apresenta, como consequência, aumento no trabalho de caracterização, sem melhoria na precisão, além de tornar mais complexa a análise e interpretação dos dados. Assim, podem-se eliminar aquelas características redundantes e de difícil mensuração, o que reduziria tempo e custos de experimentos.

Segundo JOLLIFFE (1972), a redução no número de variáveis seria útil quando algumas que estivessem presentes dificultassem a análise dos dados e fornecessem pouca ou nenhuma informação adicional.

O descarte das variáveis pode ser feito por meio da análise de componentes principais, a partir dos coeficientes de ponderação das variáveis associadas aos últimos componentes, que retêm proporções mínimas da variação total. De acordo com JOLLIFFE (1973), os resultados serão mais satisfatórios se o número de variáveis rejeitadas for igual ao número de componentes cuja variância é inferior a 0,7. Esse critério é estabelecido para os casos em que os dados são padronizados; assim, estimam-se os autovalores a partir da matriz de correlações, e sua soma corresponde ao número total de variáveis analisadas.

Objetivou-se no presente trabalho reduzir a dimensionalidade dos dados, eliminando as informações redundantes, e recomendar que estas não sejam avaliadas em experimentos futuros.

MATERIAL E MÉTODOS

Os dados foram obtidos de um experimento conduzido na Granja de Melhoramento de Suínos do Departamento de Zootecnia da Universidade Federal de Viçosa, em Viçosa – MG, no período de novembro de 1998 a julho de 2001.

Foram formadas inicialmente duas famílias, provenientes do cruzamento de dois varrões da raça nativa Brasileira (Piau) com 18 fêmeas originadas de linhagem desenvolvida na UFV, pelo acasalamento de animais das raças Landrace, Large White e Pietran. A geração F1 nasceu nos meses de março a maio de 1999 e era composta por 106 fêmeas e 134 machos. Dentre os machos F1, foram selecionados 11 varrões, provenientes de diferentes leitegadas, os quais foram acasalados (monta natural) com 54 fêmeas, evitando-se o acasalamento de parentes. Esses animais foram acasalados, nos meses de fevereiro a outubro de 2000, para produção da geração F2, a qual nasceu entre junho de 2000 e fevereiro de 2001. Assim, foram obtidos 844 animais F2, divididos em cinco lotes:

- lote 1 = animais nascidos entre 20/06/00 e 03/07/00;
- lote 2 = animais nascidos entre 03/08/00 e 23/08/00;
- lote 3 = animais nascidos entre 16/09/00 e 01/11/00;
- lote 4 = animais nascidos entre 30/11/00 e 25/12/00; e
- lote 5 = animais nascidos entre 19/01/01 e 12/02/01.

Os três primeiros lotes foram obtidos de porcas do primeiro parto; os dois últimos já eram porcas de segunda parição.

Foram avaliados 844 animais da geração F2 para características de desempenho, dos quais 550 foram abatidos. Os animais experimentais – fêmeas e machos castrados – foram abatidos com média de 64,79 kg e desvio-padrão de 5,06 kg.

Informações detalhadas sobre o manejo dos animais foram descritas por PIRES (2003).

Características Avaliadas

Preliminarmente, foi realizado um teste para diagnóstico do efeito da multicolinearidade ou dependência linear entre as variáveis, que pode levar à formação de matrizes singulares ou mal condicionadas, sendo este um problema na utilização de certas técnicas de análise multivariada que necessita de inversão de matrizes. Após essa

análise, foram identificadas e descartadas as variáveis que provocaram forte multicolinearidade nas características de desempenho, carcaça e qualidade da carne.

Para as características de desempenho, a que provocou forte multicolinearidade quando presente foi o peso aos 105 dias de idade.

Em se tratando das características de carcaça, as que provocaram forte multicolinearidade quando presentes foram peso da banda esquerda; peso da carcaça; peso ao abate; menor espessura de toucinho na região acima da última vértebra lombar, na linha dorso-lombar; e espessura de toucinho a 6,5 cm da linha dorso-lombar, equivalente a P2.

Para as características de qualidade da carne, as que provocaram forte multicolinearidade quando presentes foram saturação e tonalidade.

Características de desempenho

Foram avaliadas as seguintes características de desempenho: tamanho da leitegada ao nascimento (TLN); tamanho da leitegada à desmama – 21 dias (TLD); número de tetos (NT); pesos ao nascer (PN), aos 21 (P21), aos 42 (P42), aos 63 (P63) e aos 77 dias de idade (P77); ganho de peso médio diário (GPD); consumo de ração (CR); e conversão alimentar (CA) dos 77 aos 105 dias de idade.

Características de carcaça

Foram avaliadas as seguintes características de carcaça: idade ao abate (IDA); peso da banda direita (PBDIR); rendimento de carcaça com pés e cabeça (RCARC); comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça (MBCC); comprimento de carcaça pelo Método Americano (MLC); maior espessura de toucinho na região da copa, na linha dorso-lombar (SH); espessura de toucinho imediatamente após a última costela, na linha dorso-lombar (UC); espessura de toucinho entre a última e a penúltima vértebra lombar, na linha dorso-lombar (UL); espessura de toucinho medida imediatamente após a última costela, a 6,5 cm da linha dorso-lombar (P2); área de olho de lombo (AOL); profundidade de lombo (PROFLOMB); espessura do bacon (EBACON); pesos do pulmão (PULMAO), do coração (COR), do fígado (FIG), do baço (BACO) e do rim (RIM); e comprimento total do intestino delgado (INTEST).

A meia carcaça direita foi resfriada a 4°C por 24 horas e, então, se procedeu à sua dissecação, para obter os cortes da carcaça: pesos da banda direita resfriada (PBDIRres), do pernil (PP), do pernil sem pele e sem gordura (PPL), da copa (PCOPA), da copa sem pele e sem gordura (PCOPAL), da paleta (PPA), da paleta sem pele e sem gordura (PPAL), do carré (PC), do lombo (PL), do bacon (PB), das costelas (PCOS), da papada (PAPADA), da cabeça (CABEÇA), do filezinho (PF) e da banha rama (PBR).

Características de qualidade da carne

Foram avaliadas as seguintes características de qualidade da carne suína: pH medido aos 45 minutos e 24 horas *post-mortem* (pH₄₅, pH_u, respectivamente); maciez objetiva (força de cisalhamento - MACIEZ); gordura intramuscular (GORINT); capacidade de retenção de água [perda por gotejamento (GOTEJ), perda por cozimento (COZ) e perda de peso total (PTOT)]; e coloração. A coloração foi determinada pelo sistema HUNTER LAB, sendo medidos a luminosidade (L), o índice de vermelho (a) e o índice de amarelo (b), em espectrofotômetro COLORQUEST II.

As metodologias empregadas na avaliação das características de qualidade da carne foram descritas por BENEVENUTO JUNIOR (2001).

Metodologia Estatística

Os dados submetidos à análise de componentes principais foram previamente ajustados para efeitos fixos e covariáveis. O modelo estatístico adotado foi:

$$y_{ijk} = \mu + S_i + L_j + b(C_{ijk} - \bar{C}) + e_{ijk}$$

em que:

y_{ijk} = característica observada no animal k , no lote j , no sexo i ;

μ = constante inerente a toda observação;

S_i = efeito fixo do sexo i , $i = 1, 2$ (1 = macho castrado e 2 = fêmea);

L_j = efeito fixo do lote j , $j = 1, 2, 3, 4, 5$;

C_{ijk} = valor observado da covariável, no animal k , no lote j , no sexo i ;

\bar{C} = média da covariável;

b = coeficiente de regressão linear da característica, em função da covariável;

e_{ijk} = erro aleatório associado a cada observação.

Com base no trabalho de PIRES (2003), para ajuste dos dados, foram utilizadas as seguintes características e respectivas covariáveis:

- Característica peso ao nascer (PN) – tamanho de leitegada ao nascimento (TLN);
- Características P21, P42, P63 e P77 – tamanho de leitegada a desmama (TLD);
- Características CR, GPD e CA – peso aos 77 dias de idade (P77);
- Características de carcaça – peso da carcaça (PCARC);
- Características de cortes de carcaça – peso da banda direita resfriada (PBDIRres); e
- Características de qualidade da carne – idade ao abate (IDA).

Análise de Componentes Principais

O método de análise de componentes principais, a partir da matriz de correlação, consiste em transformar um conjunto de variáveis Z_1, Z_2, \dots, Z_p em um novo conjunto de variáveis $Y_1 (CP_1), Y_2 (CP_2), \dots, Y_p (CP_p)$, com as seguintes propriedades:

- I. Cada componente principal (CP_i) é uma combinação linear das características padronizadas (Z_j s), ou seja:

$$CP_i = a_{i1}Z_1 + a_{i2}Z_2 + \dots + a_{ip}Z_p = \sum_{j=1}^p a_{ij}Z_j$$

com: $i e j = 1, 2, \dots, p$

- II. A soma dos quadrados dos coeficientes a_{ij} , com $j = 1, 2, \dots, p$, é igual à unidade, isto é:

$$a_{i1}^2 + a_{i2}^2 + \dots + a_{ip}^2 = 1, \text{ ou seja, } \sum_{j=1}^p a_{ij}^2 = 1 \text{ para todo } i$$

III. De todas essas combinações lineares possíveis, CP_1 é a que tem a maior variância, e assim sucessivamente, até completar o conjunto dos p CP_i 's, indo de CP_1 até CP_p , ou seja:

$$Var(CP_1) \geq Var(CP_2) \geq K \geq Var(CP_p)$$

IV. As combinações lineares formadas são não-correlacionadas, isto é:

$$Cov(CP_1, CP_2) = Cov(CP_1, CP_3) = K = Cov(CP_{p-1}, CP_p) = 0$$

ou

$$Cov(CP_i, CP_j) = 0 \quad \text{para } i \neq j$$

Dessa forma, um novo conjunto de p variáveis não-correlacionadas entre si e arrançadas numa ordem decrescente de variâncias é definido. A idéia principal por trás desse procedimento é de que poucos, dentre os primeiros componentes principais, contenham a maior variabilidade dos dados originais; contudo, pode-se racionalmente descartar os demais componentes, reduzindo o número de variáveis.

Vale salientar que a obtenção dos componentes principais pode ser a partir de uma matriz de covariâncias (**S**) ou de uma matriz de correlações entre as variáveis originais (**R**) de X_1, X_2, \dots, X_p . Assim, a técnica de componentes principais caracteriza-se por trabalhar com a média amostral ou ser usada nas situações em que não haja repetições de dados. Entretanto, as estimativas dos componentes principais, quando se usa a matriz **S**, podem ser muito diferentes daquelas encontradas quando se utiliza a matriz **R**. Por essa razão, é recomendado o uso de matriz **S** somente naqueles casos em que as unidades originais não são fixadas arbitrariamente, mas sim sugeridas por razões objetivas, ou apenas nos casos em que os caracteres apresentam uma mesma unidade e uma dimensão não muito discrepante.

Tendo em vista o grande número de variáveis medidas em unidades diferentes, foi necessária a padronização destas variáveis X_j ($j = 1, 2, \dots, p$); nesse caso, a estrutura de dependência de X_j foi dada pela matriz de correlação **R**.

Dessa forma, a diferença das escalas de medidas das variáveis foi retirada pelo uso de variáveis reduzidas (variáveis padronizadas), conforme apresentado a seguir:

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{S(X_j)}$$

em que:

Z_{ij} = variável padronizada no indivíduo i , na característica j ;

X_{ij} = valor observado no indivíduo i , na característica j ;

\bar{X}_j = média estimada da característica j ; e

$S(X_j)$ = desvio-padrão dos dados da característica j .

Assim, a matriz de correlação das variáveis padronizadas é:

$$R = \begin{bmatrix} 1 & r_{12} & \Lambda & r_{1p} \\ r_{21} & 1 & \Lambda & r_{2p} \\ \text{M} & \text{MO} & \text{M} & \\ r_{p1} & r_{p2} & \Lambda & 1 \end{bmatrix}$$

sendo:

$$r_{j'j''} = r(X_j, X_{j'}) = \text{Cov}(Z_j, Z_{j'}) = \frac{\text{Cov}(X_j, X_{j'})}{\sqrt{\hat{\text{Var}}(X_j)\hat{\text{Var}}(X_{j'})}$$

Os componentes principais foram obtidos a partir da matriz de correlação com base nas seguintes expressões:

$$|R - \lambda I| = 0 \quad \Rightarrow \text{fornece os autovalores } \lambda_1, \lambda_2, \dots, \lambda_p$$

$$[R - \lambda_i I] \tilde{a}_i = \tilde{\phi} \quad \Rightarrow \text{fornece os autovetores } \tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_p$$

em que:

R = matriz de correlação entre as características avaliadas;

λ_i = autovalores da matriz R ;

\tilde{a}_i = autovetor associado ao autovalor λ_i ;

I = matriz identidade de ordem p (p = número de características); e

$\tilde{\phi}$ = vetor nulo, de dimensão $p \times 1$.

Os autovetores \tilde{a}_i foram normalizados de modo a obter \tilde{a}_i^* tal que $\tilde{a}_i^* \tilde{a}_i^* = 1$ para $i=1,2,\dots,p$ e $\tilde{a}_i^* \tilde{a}_j^* = 0$ para $i \neq j$.

Assim, os autovalores de \mathbf{R} corresponderam às variâncias de cada componente, e os autovetores normalizados corresponderam aos coeficientes de ponderação das características padronizadas.

A importância relativa de um componente principal foi avaliada pela percentagem de variância total que ele explica, ou seja, a percentagem de seu autovalor em relação ao total dos autovalores de todos os componentes, ou a percentagem de seu autovalor em relação ao traço da matriz \mathbf{R} , que é dado por:

$$\text{importância de } CP_j = \frac{\hat{Var}(CP_j)}{\sum_{j=1}^p \hat{Var}(CP_j)} \cdot 100 = \frac{\lambda_j}{\sum_{j=1}^p \lambda_j} \cdot 100 = \frac{\lambda_j}{\text{traço}(\mathbf{R})} \cdot 100$$

Para descarte de variáveis, a variável que domina (aquela que possui maior correlação) no componente principal de menor autovalor (menor variância) deve ser menos importante para explicar a variância total e, portanto, passível de descarte (REGAZZI, 2002).

O critério do número de variáveis descartadas foi, conforme recomendações de JOLLIFFE (1972, 1973), baseado em dados simulados e reais; quando a análise de componentes principais foi feita a partir da matriz de correlação, estabeleceu-se o critério de que o número de variáveis descartadas deve ser igual ao número de componentes cuja variância (autovalor) é inferior a 0,7.

Análise dos Dados

Todas as análises foram feitas utilizando o programa SAS System for Windows^{NT}, versão 8.0, licenciado pela UNIVERSIDADE FEDERAL DE VIÇOSA (SAS Institute Inc., Cary, NC, USA, 1999).

RESULTADOS E DISCUSSÃO

Características de desempenho

O número de observações, as médias corrigidas e os desvios-padrão das características de desempenho são apresentados na Tabela 1. Após serem feitas as consistências nos dados, 435 observações foram utilizadas na análise.

Tabela 1 – Número de observações, médias corrigidas e desvios-padrão das características de desempenho

Característica	Número de observações	Média corrigida	Desvio-padrão *
TLN (nº)	798	10,42052	2,13024
TLD (nº)	713	9,27431	1,89787
NT (nº)	804	13,19901	1,28136
PN (kg)	798	1,18584	0,26711
P21 (kg)	584	4,94395	1,10331
P42 (kg)	595	8,38423	1,81944
P63 (kg)	578	16,54801	3,14553
P77 (kg)	535	21,55254	3,96670
CR (kg)	595	40,29999	6,52833
GPD (kg)	585	0,53587	0,12155
CA (kg/kg)	580	2,80183	0,62664

TLN - tamanho da leitegada ao nascimento; TLD - tamanho da leitegada ao desmame; NT - número de tetos; PN - peso ao nascer, P21 - aos 21, P42 - aos 42, P63 - aos 63 e P77 - aos 77 dias de idade; GPD - ganho de peso médio diário; CR - consumo de ração; e CA - conversão alimentar dos 77 aos 105 dias de idade.

* Desvio-padrão dos dados corrigidos para efeitos fixos e covariáveis.

Os resultados obtidos para os componentes principais, os autovalores e as percentagens da variância explicada por cada componente são apresentados na Tabela 2.

Tabela 2 – Componentes principais (CP), autovalores (λ_i) e percentagem da variância explicada pelos componentes (% var. CP), das características de desempenho

Componentes principais	λ_i	% var. CP	% var. CP (acumulada)
CP ₁	3,66645313	33,33	33,33
CP ₂	2,03806818	18,53	51,86
CP ₃	1,76553181	16,05	67,91
CP ₄	0,98916384	8,99	76,90
CP ₅	0,90767612	8,25	85,15
CP ₆	0,65728174	5,98	91,13
CP ₇	0,38070829	3,46	94,59
CP ₈	0,23790821	2,16	96,75
CP ₉	0,19937425	1,81	98,57
CP ₁₀	0,10214737	0,93	99,49
CP ₁₁	0,05568705	0,51	100

Os três primeiros componentes principais explicaram 67,91% da variação total.

Conforme pode ser observado na Tabela 2, seis variáveis (54,54% das 11 características avaliadas) podem ser descartadas de acordo com o critério de JOLLIFFE (1972, 1973), em razão de o número de componentes que apresentaram autovalores menores que 0,7 ser seis.

STRAPASSON et al. (2000), ao utilizarem a análise de componentes principais em forrageiras do gênero *Paspalum*, observaram uma redução de 53, 68 e 43% dos descritores reprodutivos, vegetativos e agrônômicos, respectivamente, do conjunto inicialmente considerado.

As seis variáveis que apresentaram maiores coeficientes, em valor absoluto, a partir do último componente principal são passíveis de descarte, conforme apresentado na Tabela 3. A razão para isso é que variáveis altamente correlacionadas com os componentes principais de menor variância representam variação praticamente insignificante.

As variáveis passíveis de descarte são, respectivamente, em ordem de menor importância para explicar a variação total: ganho de peso médio diário, peso aos 63 dias de idade, peso aos 42 dias de idade, tamanho da leitegada à desmama, peso aos 21 dias de idade e peso ao nascer.

Com base nesses resultados, recomendam-se as seguintes variáveis para serem mantidas em experimentos futuros: tamanho da leitegada ao nascimento, número de tetos, peso aos 77 dias de idade, consumo de ração e conversão alimentar dos 77 aos 105 dias de idade.

Do ponto de vista econômico, a conversão alimentar deve receber atenção especial, pois a alimentação representa considerável parte do custo total de produção de suínos. Uma restrição à avaliação dessa característica, para fins de melhoramento genético, é o custo elevado de sua mensuração, uma vez que os animais têm de ser criados em baias individuais e o fornecimento de ração deve ser anotado. Mesmo assim, as referências encontradas justificam o custo de mensuração, visto que o ganho genético obtido nessa característica é compensador (TORRES FILHO, 2001).

PIRES (2003) encontrou QTL significativo para consumo de ração. De acordo com este autor, um mesmo gene ou grupo gênico, localizado em torno de 100 cM do cromossomo 6, pode estar atuando sobre as características consumo de ração e ganho de peso diário. Isso sugere uma correlação significativa entre essas duas variáveis. Neste

trabalho, observou-se correlação de 0,62, conforme demonstrado na Tabela 4, explicando, em parte, o fato de a variável ganho de peso diário ter sido sugerida para descarte.

Apesar de terem sido avaliadas cinco pesagens (PN, P21, P42, P63 e P77), apenas uma foi selecionada, sendo as demais redundantes. Além disso, a redução no número de pesagens causaria menos estresse ao animal na avaliação de característica.

Como os tamanhos de leitegada ao nascimento e à desmama são fortemente correlacionados, conforme demonstrado na Tabela 4, a análise de componentes principais sugeriu para descarte o tamanho da leitegada ao nascimento, por esta ser redundante.

Adicionalmente, pôde-se observar que as variáveis sugeridas para descarte apresentaram correlação linear simples significativa com as demais, ou seja, são redundantes; por outro lado, as variáveis selecionadas apresentaram menor correlação entre si (Tabela 4). Desse modo, cada variável 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 ponderação das características e suas correlações (em porcentagem) com os seis últimos componentes principais menos importantes para explicar a variação total das 11 características de desempenho

Característica	Coeficientes (Correlações - %)					
	CP ₁₁	CP ₁₀	CP ₉	CP ₈	CP ₇	CP ₆
TLN	-0,063016 (-1,487)	-0,003536 (-0,113)	0,170324 (7,605)	0,656094 (32,002)	0,133733 (8,252)	0,120510 (9,770)
TLD	0,080231 (1,893)	0,032339 (1,034)	-0,170240 (-7,601)	-0,674098 (-32,880)	-0,115558 (-7,130)	0,003342 (0,271)
NT	0,008138 (0,192)	0,006693 (0,214)	0,012571 (0,561)	-0,059993 (-2,926)	-0,017670 (-1,090)	0,167317 (13,565)
PN	-0,028369 (-0,669)	-0,014371 (-0,459)	0,068005 (3,037)	-0,001533 (-0,075)	0,311037 (19,191)	-0,869807 (-70,518)
P21	0,067130 (1,584)	0,002410 (0,077)	-0,375472 (-16,765)	0,264599 (12,906)	-0,754422 (-46,549)	-0,127693 (-10,352)
P42	-0,113538 (-2,679)	0,240499 (7,686)	0,775471 (34,626)	-0,165141 (-8,055)	-0,199369 (-12,301)	0,160086 (12,979)
P63	0,030643 (0,723)	-0,782567 (25,011)	-0,063809 (-2,849)	-0,059501 (-2,902)	0,251109 (15,494)	0,273581 (22,180)
P77	0,047853 (1,129)	0,572141 (18,286)	-0,405022 (-18,085)	0,031015 (1,513)	0,444946 (27,454)	0,284868 (23,095)
CR	-0,429773 (-10,142)	-0,027883 (-0,891)	-0,064364 (-2,874)	0,032115 (1,566)	-0,027363 (-1,688)	0,036083 (2,925)
GPD	0,710125 (16,758)	0,016526 (0,528)	0,119256 (5,325)	0,041187 (2,009)	-0,003996 (-0,247)	0,025803 (2,092)
CA	0,528311 (12,467)	-0,006158 (0,197)	0,091722 (4,096)	0,083764 (4,086)	0,031736 (1,958)	-0,031788 (2,577)

TLN - tamanho da leitegada ao nascimento; TLD - tamanho da leitegada ao desmame; NT - número de tetos; PN - peso ao nascer, P21 - aos 21, P42 - aos 42, P63 - aos 63 e P77 - aos 77 dias de idade; GPD - ganho de peso médio diário; CR - consumo de ração; e CA - conversão alimentar dos 77 aos 105 dias de idade.

Tabela 4 – Coeficientes de correlação simples entre as características de desempenho

	TLN	TLD	NT	PN	P21	P42	P63	P77	CR	GPD	CA
TLN	1,0000										
TLD	0,7520	1,0000									
NT	0,0099	-0,0367	1,0000								
PN	0,0407	0,1010	0,1366	1,0000							
P21	0,0006	0,0693	0,1222	0,5381	1,0000						
P42	-0,0281	0,0343	0,0893	0,4764	0,7496	1,0000					
P63	-0,0026	0,0265	0,1083	0,4688	0,6945	0,8308	1,0000				
P77	-0,0257	-0,0154	0,1057	0,4544	0,6397	0,7649	0,8848	1,0000			
CR	0,0565	0,1169	0,0271	0,0504	-0,0094	-0,0101	-0,0080	0,0128	1,0000		
GPD	0,0273	-0,0007	0,0191	0,0636	-0,0345	-0,0193	-0,0144	-0,0135	0,6239	1,0000	
CA	0,0061	0,0468	-0,0162	-0,0407	-0,0018	0,0236	-0,0126	-0,0032	-0,0808	-0,7530	1,0000

TLN - tamanho da leitegada ao nascimento; TLD - tamanho da leitegada ao desmame; NT - número de tetos; PN - peso ao nascer, P21 - aos 21, P42 - aos 42, P63 - aos 63 e P77 - aos 77 dias de idade; GPD - ganho de peso médio diário; CR - consumo de ração; e CA - conversão alimentar dos 77 aos 105 dias de idade.

Características de carcaça

O número de observações, as médias corrigidas e os desvios-padrão das características de carcaça são apresentados nas Tabela 5.

Tabela 5 – Número de observações, médias corrigidas e desvios-padrão das características de carcaça

Característica	Número de Observações	Média corrigida	Desvio-padrão *
IDA (dias)	521	147,49089	9,02621
PBDIR (kg)	526	26,93527	0,44033
RCARC (%)	510	82,03599	1,75048
MBCC (cm)	527	86,12846	3,30273
MLC (cm)	521	71,81434	2,51823
SH (mm)	526	40,34599	4,78134
UC (mm)	528	19,83661	4,28438
UL (mm)	526	28,45176	5,31851
P2 (mm)	526	16,82945	3,27629
PROFLOMB (mm)	456	44,00011	4,08628
AOL (cm ²)	461	26,54827	3,17069
PULMÃO (kg)	520	0,45527	0,07511
COR (kg)	526	0,23587	0,02657
FIG (kg)	524	1,26710	0,14541
BACO (kg)	526	0,09341	0,01798
INTEST (m)	523	18,43255	1,70151
PBDIRres (kg)	537	26,30870	2,65343
PP (kg)	531	7,29198	0,38003
PPL (kg)	531	5,01165	0,40448
PCOPA (kg)	537	2,34227	0,23007
PCOPAL (kg)	532	1,69171	0,19906
PPA (kg)	533	4,87903	0,34134
PPAL (kg)	536	2,70860	0,30019
PC (kg)	528	3,48642	0,31579
PL (kg)	530	1,02937	0,13702
PB (kg)	529	2,67993	0,29944
EBACON	527	24,80329	5,67158
PCOS (kg)	536	1,52593	0,19546
PAPADA (kg)	534	0,71201	0,19139
CABEÇA (kg)	533	1,51383	0,16300
PF (kg)	534	0,22255	0,03073
PBR (kg)	532	0,45402	0,13047
RIM (kg)	525	0,12740	0,01646

IDA - idade ao abate; PBdir - peso da banda direita; Rcarc - rendimento de carcaça com pés e cabeça; MBCC - comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça; MLC - comprimento de carcaça pelo Método Americano; SH - maior espessura de toucinho na região da copa, na linha dorso-lombar; UC - espessura de toucinho imediatamente após a última costela, na linha dorso-lombar; UL - espessura de toucinho entre a última e a penúltima vértebra lombar, na linha dorso-lombar; P2 - espessura de toucinho medida imediatamente após a última costela, a 6,5 cm da linha dorso-lombar; PROFLOMB - profundidade de lombo; AOL - área de olho de lombo; PULMÃO - peso do pulmão; COR - peso do coração; FIG - peso do fígado; BACO - peso de baço; INTEST - comprimento total do intestino delgado; PBDIRres - peso da banda direita resfriada; PP - peso do pernil; PPL - peso do pernil limpo; PCOPA - peso da copa; PCOPAL - peso da copa limpa; PPA - peso da paleta; PPAL - peso da paleta limpa; PC - peso do carré; PL - peso do lombo; PB - peso do bacon; EBACON - espessura do bacon; PCOS - peso das costelas; PAPADA - peso da papada; CABEÇA - peso da cabeça; PF - peso do filezinho; PBR - peso da banha rama; e RIM - peso do rim.

* Desvio-padrão dos dados corrigidos para efeitos fixos e covariáveis.

Após serem feitas as consistências nos dados, 367 observações foram utilizadas na análise.

Os resultados obtidos para os componentes principais, seus respectivos autovalores e percentagens da variância explicada por cada um são apresentados na Tabela 6.

Tabela 6 – Componentes principais (CP), autovalores (λ_i) e percentagem da variância explicada pelos componentes (% var. CP), das características de carcaça

Componentes principais	λ_i	% var. CP	% var. CP (acumulada)
CP ₁	7,11502642	21,56	21,56
CP ₂	3,19399983	9,68	31,24
CP ₃	2,18870988	6,63	37,87
CP ₄	1,73999864	5,27	43,14
CP ₅	1,63646299	4,96	48,10
CP ₆	1,31877385	4,00	52,10
CP ₇	1,23500965	3,74	55,84
CP ₈	1,14304141	3,46	59,31
CP ₉	1,10809914	3,36	62,66
CP ₁₀	1,03843389	3,15	65,81
CP ₁₁	0,99010615	3,00	68,81
CP ₁₂	0,91540061	2,77	71,59
CP ₁₃	0,85822729	2,60	74,19
CP ₁₄	0,82194009	2,49	76,68
CP ₁₅	0,80542306	2,44	79,12
CP ₁₆	0,71385057	2,16	81,28
CP ₁₇	0,62873197	1,91	83,19
CP ₁₈	0,57485497	1,74	84,93
CP ₁₉	0,54872679	1,66	86,59
CP ₂₀	0,49459181	1,50	88,09
CP ₂₁	0,48374203	1,47	89,56
CP ₂₂	0,44892429	1,36	90,92
CP ₂₃	0,41991465	1,27	92,19
CP ₂₄	0,37466169	1,14	93,32
CP ₂₅	0,35257200	1,07	94,39
CP ₂₆	0,33327336	1,01	95,40
CP ₂₇	0,30013388	0,91	96,31
CP ₂₈	0,28127010	0,85	97,16
CP ₂₉	0,25297852	0,77	97,93
CP ₃₀	0,24386258	0,74	98,67
CP ₃₁	0,16312866	0,49	99,16
CP ₃₂	0,14761302	0,45	99,61
CP ₃₃	0,12851621	0,39	100

Os três primeiros componentes principais explicaram apenas 37,87% da variação total observada. Entretanto, apesar de a técnica de componentes principais ter sido ineficiente em reduzir o espaço multidimensional em um espaço bi ou tridimensional; a técnica tem como vantagem adicional avaliar a importância de cada característica

estudada sobre a variação total, possibilitando o descarte das variáveis menos discriminantes (redundantes), por já estarem correlacionadas com outras variáveis, pela sua invariância ou por serem combinação linear de outras características.

Dos 33 componentes principais, 17 apresentaram variância inferior a 0,7 (autovalor inferior a 0,7). Assim, as 17 variáveis que apresentaram maiores coeficientes, em valor absoluto, a partir do último componente principal, são passíveis de descarte, conforme apresentado na Tabela 7. A razão para isso é que variáveis altamente correlacionadas com os componentes principais de menor variância representam variação praticamente insignificante.

DAHER et al. (1997), utilizando a técnica de componentes principais para descarte dos caracteres que menos contribuíram para o estudo da divergência genética entre acessos de capim-elefante, observaram que, de um total de 22 caracteres avaliados consistentemente em três anos, selecionaram-se oito como os mais importantes para a determinação de divergência genética.

As variáveis sugeridas para descarte são, respectivamente, em ordem de menor importância para explicar a variação total: comprimento de carcaça pelo Método Americano de Classificação de Carcaça; peso do pernil sem pele e sem gordura; peso da copa sem pele e sem gordura; espessura de toucinho medida imediatamente após a última costela, a 6,5 cm da linha dorso-lombar; espessura do bacon; área de olho de lombo; espessura de toucinho entre a última e a penúltima vértebra lombar, na linha dorso-lombar; peso do lombo; peso da banha rama; peso da paleta sem pele e sem gordura; espessura de toucinho imediatamente após a última costela, na linha dorso-lombar; peso do fígado; comprimento total do intestino delgado; peso da cabeça; peso do bacon; rendimento de carcaça com pés e cabeça; e peso do pulmão.

Com base nesses resultados, recomendam-se as seguintes variáveis para serem mantidas em experimentos futuros: idade ao abate; peso da banda direita; comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça; maior espessura de toucinho na região da copa, na linha dorso-lombar; profundidade de lombo; peso do baço; peso do coração; peso da banda direita resfriada; peso do pernil; peso da copa; peso da paleta; peso do carré; peso das costelas; peso da papada; peso do filezinho; e peso do rim.

Como pode ser observado na Tabela 8, as espessuras de toucinho apresentaram alta correlação entre si, justificando em parte o fato de que, das quatro medidas de

espessura de toucinho incluídas na análise, apenas uma foi selecionada, sendo as demais redundantes.

Área de olho de lombo foi sugerida para descarte por apresentar forte correlação com outras características, principalmente profundidade de lombo e peso do lombo (Tabela 8). Ressalta-se ainda que a área de olho de lombo é uma característica de difícil mensuração, em razão da necessidade de uso de planímetro em decalque vegetal tomado da seção transversal do músculo *Longissimus dorsi*.

Apesar de a área de olho de lombo (AOL) ter sido sugerida para descarte, segundo CROSS et al. (1975), ela é um melhor preditor (maiores coeficientes de correlação) do rendimento em cortes do que a profundidade do lombo (PROFLOMB). No entanto, deve-se ressaltar que, pela metodologia utilizada, PROFLOMB não foi passível de descarte, provavelmente por não apresentar correlação significativa com grande número de características, diferentemente do que ocorreu com AOL.

Os cortes limpos (sem pele e sem gordura) foram, todos, sugeridos para descarte. A não-utilização de cortes limpos em análises futuras evitará dispêndio de tempo e dinheiro, sendo necessária somente a avaliação nos cortes inteiros.

Na Tabela 8 estão apresentados os coeficientes de correlação simples entre as características de carcaça.

Tabela 7 – Coeficientes de ponderação das características com os 18 últimos componentes principais menos importantes para explicar a variação total das 33 características de carcaça

Característica	Coeficientes					
	CP ₃₃	CP ₃₂	CP ₃₁	CP ₃₀	CP ₂₉	CP ₂₈
IDA	0,034076	-0,014725	0,049423	0,147713	0,029384	-0,061004
PBdir	0,023598	-0,008791	0,067036	0,018927	0,092763	-0,221761
Rcarc	0,003865	-0,013645	0,016430	0,049960	-0,000818	0,004667
MBCC	-0,562351	-0,064659	0,305719	-0,016153	0,107005	0,057809
MLC	0,607403	-0,010966	-0,346560	-0,057115	0,163715	-0,128305
SH	-0,032519	-0,068465	-0,058161	-0,033953	-0,076541	0,001647
UC	0,018350	0,030381	-0,026938	-0,151752	-0,135123	-0,047333
UL	0,017841	0,082836	0,073912	-0,437512	0,155390	-0,137251
P2	0,064436	0,153829	0,172025	0,705993	0,251902	-0,155659
PROFLOMB	0,022746	-0,009601	0,006522	0,148441	-0,244550	-0,315839
AOL	0,009931	-0,033434	-0,035354	0,095145	0,303093	0,522813
PULMÃO	-0,012015	0,041615	-0,004583	-0,023043	-0,031224	-0,084776
COR	0,006911	0,032057	0,008188	-0,003247	0,032139	0,050190
FIG	-0,067764	-0,001122	0,010880	0,095134	-0,139765	0,046548
BACO	0,099163	-0,060940	-0,011879	0,008137	0,035865	-0,054770
INTEST	-0,019303	0,029756	0,102072	0,201243	0,061232	0,036268
PBDIRres	-0,021968	-0,012475	0,028895	-0,034598	-0,066351	0,018561
PP	-0,097527	-0,352363	-0,262139	0,173321	-0,024765	-0,024559
PPL	0,277175	0,523103	0,476312	-0,134191	0,096790	0,029226
PCOPA	0,242412	-0,412436	0,400079	-0,016584	-0,051679	0,040120
PCOPAL	-0,264683	0,479080	-0,478205	0,021148	0,035035	-0,047845
PPA	0,008358	0,244096	0,050019	0,128390	-0,170560	0,151203
PPAL	0,065131	-0,189521	-0,019876	-0,057885	0,053559	0,097381
PC	0,120237	0,060893	-0,019484	0,147782	-0,364070	0,338335
PL	-0,210548	-0,070857	0,051137	-0,077010	0,187601	-0,415553
PB	0,034628	0,042502	-0,044803	0,027294	0,042945	0,122192
EBACON	-0,025842	-0,006036	-0,058056	-0,128329	0,567103	0,210740
PCOS	0,018282	-0,145560	-0,036330	0,081740	0,050380	0,079323
PAPADA	0,012204	0,126431	-0,001376	0,105521	-0,110796	0,039189
CABEÇA	-0,083662	-0,013926	-0,027555	-0,061388	-0,011590	0,273835
PF	-0,034204	-0,049838	-0,145373	-0,008499	-0,042133	0,085831
PBR	-0,037780	0,077333	0,090541	-0,199443	-0,322729	0,113114
RIM	0,030539	-0,032188	-0,012409	0,017753	-0,001643	0,071898

IDA - idade ao abate; PBdir - peso da banda direita; Rcarc - rendimento de carcaça com pés e cabeça; MBCC - comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça; MLC - comprimento de carcaça pelo Método Americano; SH - maior espessura de toucinho na região da copa, na linha dorso-lombar; UC - espessura de toucinho imediatamente após a última costela, na linha dorso-lombar; UL - espessura de toucinho entre a última e a penúltima vértebra lombar, na linha dorso-lombar; P2 - espessura de toucinho medida imediatamente após a última costela, a 6,5 cm da linha dorso-lombar; PROFLOMB - profundidade de lombo; AOL - área de olho de lombo; PULMÃO - peso do pulmão; COR - peso do coração; FIG - peso do fígado; BACO - peso de baço; INTEST - comprimento total do intestino delgado; PBDIRres - peso da banda direita resfriada; PP - peso do pernil; PPL - peso do pernil limpo; PCOPA - peso da copa; PCOPAL - peso da copa limpa; PPA - peso da paleta; PPAL - peso da paleta limpa; PC - peso do carré; PL - peso do lombo; PB - peso do bacon; EBACON - espessura do bacon; PCOS - peso das costelas; PAPADA - peso da papada; CABEÇA - peso da cabeça; PF - peso do filezinho; PBR - peso da banha rama; e RIM - peso do rim.

Tabela 7 – Coeficientes de ponderação das características com os 18 últimos componentes principais menos importantes para explicar a variação total das 33 características de carcaça (CONTINUAÇÃO)

Característica	Coeficientes					
	CP ₂₇	CP ₂₆	CP ₂₅	CP ₂₄	CP ₂₃	CP ₂₂ *
IDA	0,062619	0,073209	0,132425	0,005942	-0,160181	0,194452
PBdir	0,017121	-0,049212	0,297064	0,417925	0,171722	-0,038645
Rcarc	-0,027145	-0,010937	0,085333	0,099880	0,040897	0,142403
MBCC	-0,003504	-0,120701	0,051712	0,029951	0,023518	-0,077665
MLC	-0,016024	0,138667	-0,086444	0,027416	0,140889	-0,011397
SH	-0,090717	-0,074445	0,090754	0,268157	0,302815	-0,280113
UC	-0,402275	0,294509	0,189120	-0,019561	-0,476102	-0,286089
UL	0,633706	-0,142804	-0,003953	-0,203015	-0,082095	0,072474
P2	0,101803	0,003621	-0,360042	-0,032237	-0,028980	-0,122799
PROFLOMB	-0,117479	-0,354750	0,058839	-0,000452	0,119334	0,161904
AOL	0,185211	0,150950	0,071947	0,046355	0,143916	-0,368274
PULMÃO	0,142120	-0,066343	-0,065111	0,190287	0,060641	-0,013158
COR	-0,061127	-0,006043	-0,033597	-0,101159	0,000631	0,136056
FIG	-0,085099	-0,004847	0,169978	-0,100349	0,218225	0,105866
BACO	0,062868	-0,032965	0,078345	-0,046614	0,049994	-0,075357
INTEST	-0,059130	0,101773	0,076635	0,084379	-0,161337	0,186261
PBDIRres	-0,047410	0,017867	-0,087707	-0,043439	-0,049650	0,005912
PP	0,030413	0,053024	0,101679	-0,037087	0,119744	-0,003233
PPL	-0,208469	-0,127558	0,017431	-0,009086	0,073740	-0,077399
PCOPA	0,006810	0,094025	0,017374	-0,208388	-0,029881	0,053324
PCOPAL	-0,015305	-0,089607	-0,069495	-0,093940	-0,081546	-0,028940
PPA	0,272149	0,412129	0,319077	0,066898	0,101529	0,176708
PPAL	0,134303	-0,255406	-0,237267	0,512600	-0,422472	-0,048596
PC	0,111068	-0,113054	-0,025997	-0,080103	-0,146543	0,272599
PL	0,055335	0,529715	-0,177419	-0,034795	0,011710	0,094468
PB	0,068015	-0,072727	0,225864	-0,168418	0,038175	-0,245631
EBACON	-0,309511	-0,040577	0,100861	0,148747	0,061944	0,523893
PCOS	0,033888	-0,131230	-0,008374	-0,082942	0,118329	0,061793
PAPADA	0,234965	0,160327	0,160946	0,333067	-0,189940	0,129244
CABEÇA	-0,112037	0,088197	-0,376620	-0,148145	0,026288	0,058831
PF	0,000633	-0,036562	-0,039339	-0,067815	-0,002984	0,176146
PBR	-0,026914	0,234768	-0,444426	0,313084	0,364326	0,065654
RIM	-0,039729	0,047445	-0,080235	0,088609	-0,230334	-0,029163

* No 22º componente principal não houve sugestão para descarte; haja vista que o maior coeficiente foi de uma características passível de descarte num componente anterior.

Tabela 7 – Coeficientes de ponderação das características com os 18 últimos componentes principais menos importante para explicar a variação total das 33 características de carcaça (CONTINUAÇÃO)

Característica	Coeficientes					
	CP ₂₁	CP ₂₀	CP ₁₉	CP ₁₈	CP ₁₇	CP ₁₆
IDA	-0,240764	-0,085629	0,015963	-0,103855	0,152264	0,139649
PBdir	-0,016264	-0,045353	0,152934	-0,070204	-0,282538	-0,227326
Rcarc	-0,225101	-0,374538	0,225075	0,055016	0,510284	-0,327475
MBCC	0,211903	-0,165318	0,195917	0,307987	0,073080	0,083085
MLC	0,092870	-0,124030	0,156996	0,205718	0,115079	0,122532
SH	0,230266	0,080110	-0,134154	-0,296317	0,491087	0,260937
UC	-0,007173	0,101823	0,261583	0,204731	0,069077	0,015011
UL	-0,040348	-0,073911	0,170654	0,084311	0,123015	0,144282
P2	0,105371	0,06703	0,066542	0,148335	0,080607	-0,029624
PROFLOMB	0,128708	0,092803	0,180488	0,012669	-0,034771	0,313628
AOL	-0,177836	-0,004094	-0,117414	0,076558	-0,034071	0,099515
PULMÃO	0,043082	0,164000	-0,083872	-0,172663	0,144209	-0,437974
COR	0,223050	-0,123864	-0,130193	0,198863	0,046490	-0,123631
FIG	-0,412075	0,318135	-0,022258	0,310405	0,331127	-0,105134
BACO	-0,032315	0,056563	0,110284	0,303696	0,056675	-0,075831
INTEST	-0,296580	-0,383472	0,028953	-0,296901	0,005245	0,368553
PBDIRres	-0,011087	0,030810	-0,097996	-0,003570	0,232364	0,152732
PP	0,053455	-0,176961	0,276407	0,059270	-0,150935	0,069656
PPL	0,014823	-0,042588	0,131912	-0,059979	0,034722	0,032936
PCOPA	0,147490	0,112420	-0,112086	-0,015350	0,108817	0,058128
PCOPAL	-0,040691	0,029700	-0,022721	-0,003835	0,079748	0,071685
PPA	0,195114	0,083105	0,072564	-0,042194	0,012500	0,123079
PPAL	-0,141073	0,200415	0,094430	-0,014480	0,108277	0,101528
PC	0,001423	0,077461	0,057667	-0,126854	-0,074014	-0,226409
PL	-0,028220	0,186889	0,052151	-0,252025	0,100247	-0,108201
PB	0,150016	0,017387	0,376769	-0,355494	0,044098	-0,152622
EBACON	0,125515	0,197206	0,021249	0,026884	-0,021784	-0,000073
PCOS	-0,083723	0,137474	0,389255	-0,062684	0,002752	-0,006281
PAPADA	0,283328	0,210981	0,067887	0,236710	0,056813	0,108491
CABEÇA	0,102744	0,190012	0,421200	-0,199508	0,075826	0,060090
PF	0,321698	-0,193253	-0,155774	0,019461	0,233214	0,028382
PBR	-0,112206	-0,180169	0,148036	0,172503	-0,126847	0,061749
RIM	0,274953	-0,366529	0,003706	-0,031432	0,082237	-0,268536

Tabela 8 – Coeficientes de correlação simples entre as características de carcaça

	IDA	PBdir	Rcarc	MBCC	MLC	SH	UC	UL	P2	proflomb	AOL	pulmão	COR	FIG	BACO	INTEST
IDA	1,000															
PBdir	-0,019	1,000														
Rcarc	-0,015	0,001	1,000													
MBCC	-0,061	0,049	-0,251	1,000												
MLC	-0,120	0,066	-0,276	0,824	1,000											
SH	0,123	-0,105	0,101	-0,347	-0,355	1,000										
UC	0,116	-0,015	0,170	-0,448	-0,496	0,429	1,000									
UL	0,055	-0,016	0,125	-0,508	-0,497	0,525	0,609	1,000								
P2	0,057	-0,046	0,183	-0,510	-0,520	0,522	0,620	0,674	1,0000							
proflomb	0,055	-0,026	0,068	-0,077	-0,095	-0,119	0,071	-0,001	-0,156	1,000						
AOL	0,067	0,111	0,107	-0,048	-0,078	-0,143	0,040	-0,101	-0,235	0,596	1,000					
pulmão	-0,065	-0,077	-0,240	0,285	0,301	-0,255	-0,183	-0,272	-0,300	0,006	0,014	1,000				
COR	0,062	0,071	-0,161	0,186	0,187	-0,168	-0,166	-0,225	-0,294	0,143	0,168	0,214	1,000			
FIG	-0,160	0,035	-0,369	0,252	0,361	-0,167	-0,247	-0,107	-0,211	-0,109	-0,122	0,253	0,151	1,000		
BACO	0,040	-0,028	-0,060	0,054	-0,054	-0,099	-0,088	-0,177	-0,175	0,134	0,154	0,094	0,123	-0,048	1,000	
INTEST	-0,291	0,003	-0,349	0,239	0,337	-0,097	-0,207	-0,093	-0,244	-0,093	-0,145	0,266	0,111	0,344	0,049	1,000
Pbdirres	-0,065	0,112	0,070	-0,037	-0,006	-0,076	-0,055	-0,012	0,016	-0,016	0,010	-0,069	-0,105	-0,057	0,024	-0,034
PP	-0,003	-0,135	-0,001	-0,065	-0,116	-0,040	-0,013	0,003	-0,080	0,108	0,162	0,072	0,082	-0,032	0,105	-0,108
PPL	-0,026	-0,065	-0,096	0,297	0,240	-0,332	-0,394	-0,485	-0,581	0,217	0,356	0,205	0,235	0,112	0,227	0,001
Pcopa	-0,032	0,407	0,066	0,126	0,105	-0,093	-0,101	-0,124	-0,136	0,025	0,114	0,009	0,070	-0,006	0,019	0,040
Pcopal	0,013	0,419	0,017	0,232	0,209	-0,184	-0,210	-0,261	-0,296	0,083	0,212	0,035	0,158	0,017	0,144	0,061
PPA	-0,048	0,013	-0,129	0,302	0,291	-0,238	-0,232	-0,299	-0,271	0,022	-0,105	0,217	0,180	0,102	0,080	0,034
PPAL	-0,022	-0,004	-0,153	0,387	0,349	-0,385	-0,380	-0,494	-0,491	0,143	0,141	0,209	0,328	0,199	0,186	0,079
PC	0,041	0,085	0,141	-0,071	-0,164	0,279	0,274	0,246	0,194	0,114	0,212	-0,183	-0,080	-0,233	0,045	-0,187
PL	-0,029	-0,030	-0,001	0,148	0,165	-0,215	-0,236	-0,346	-0,430	0,384	0,498	0,069	0,227	0,018	0,233	-0,033
PB	0,071	-0,181	0,010	-0,207	-0,203	0,136	0,152	0,196	0,286	-0,101	-0,238	-0,097	-0,109	-0,002	-0,045	0,010
Ebacon	0,095	-0,147	0,171	-0,525	-0,548	0,466	0,536	0,598	0,645	-0,001	-0,176	-0,255	-0,301	-0,234	-0,089	-0,220
PCOS	-0,035	0,038	-0,108	0,115	0,109	-0,105	-0,063	-0,158	-0,175	-0,055	-0,000	0,083	0,088	0,032	-0,008	0,103
Papada	-0,118	-0,107	0,083	-0,092	-0,048	0,020	-0,061	0,024	-0,004	-0,085	-0,041	-0,138	-0,151	0,003	-0,096	0,079
Cabeça	-0,055	0,295	-0,100	0,294	0,381	-0,270	-0,349	-0,276	-0,333	-0,030	-0,068	0,197	0,125	0,234	0,006	0,214
PF	0,065	0,078	-0,005	-0,019	-0,059	-0,181	-0,026	-0,147	-0,152	0,211	0,303	0,044	0,068	-0,013	0,177	-0,050
PBR	0,203	-0,062	0,115	-0,482	-0,493	0,380	0,500	0,523	0,594	0,003	-0,063	-0,275	-0,192	-0,228	-0,028	-0,178
RIM	-0,066	0,072	-0,281	0,078	0,112	-0,077	-0,121	-0,008	-0,129	0,087	0,021	0,168	0,078	0,376	0,030	0,186

Tabela 8 – Coeficientes de correlação simples entre as características de carcaça (CONTINUAÇÃO)

	Pbdirres	PP	PPL	Pcopa	Pcopal	PPA	PPAL	PC	PL	PB	Ebacon	PCOS	papada	Cabeça	PF	PBR	RIM
Pbdirres	1,000																
PP	0,022	1,000															
PPL	-0,006	0,569	1,000														
Pcopa	-0,030	-0,066	0,034	1,000													
Pcopal	-0,013	0,031	0,238	0,793	1,000												
PPA	-0,016	0,047	0,143	0,107	0,140	1,000											
PPAL	-0,037	0,194	0,523	0,112	0,302	0,475	1,000										
PC	-0,014	0,142	0,005	-0,026	0,003	-0,324	-0,179	1,000									
PL	-0,018	0,241	0,576	0,018	0,176	-0,018	0,359	0,299	1,000								
PB	0,053	-0,140	-0,317	-0,183	-0,287	-0,242	-0,204	-0,031	-0,154	1,000							
Ebacon	0,022	-0,018	-0,479	-0,146	-0,306	-0,254	-0,473	0,277	-0,309	0,309	1,000						
PCOS	-0,034	-0,179	0,117	0,039	0,134	0,099	0,060	-0,139	0,080	-0,148	-0,207	1,000					
Papada	0,071	-0,053	-0,057	-0,059	-0,141	-0,496	-0,292	-0,041	-0,013	0,154	0,050	0,034	1,000				
Cabeça	0,001	-0,084	0,130	0,179	0,191	0,244	0,203	-0,305	-0,041	-0,227	-0,417	-0,024	-0,002	1,000			
PF	0,015	0,097	0,354	0,016	0,058	0,010	0,196	-0,009	0,312	-0,094	-0,1774	0,122	-0,035	-0,043	1,000		
PBR	-0,034	-0,049	-0,449	-0,060	-0,172	-0,325	-0,416	0,161	-0,272	0,384	0,599	-0,083	0,042	-0,447	-0,071	1,000	
RIM	-0,050	-0,051	0,049	0,031	0,076	0,094	0,053	-0,133	0,044	-0,041	-0,103	0,111	-0,074	0,110	-0,059	-0,094	1,000

IDA - idade ao abate; PBdir - peso da banda direita; Rcarc - rendimento de carcaça com pés e cabeça; MBCC - comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça; MLC - comprimento de carcaça pelo Método Americano; SH - maior espessura de toucinho na região da copa, na linha dorso-lombar; UC - espessura de toucinho imediatamente após a última costela, na linha dorso-lombar; UL - espessura de toucinho entre a última e a penúltima vértebra lombar, na linha dorso-lombar; P2 - espessura de toucinho medida imediatamente após a última costela, a 6,5 cm da linha dorso-lombar; proflomb - profundidade de lombo; AOL - área de olho de lombo; Pulmão - peso do pulmão; COR - peso do coração; FIG - peso do fígado; BACO - peso de baço; INTEST - comprimento total do intestino delgado; PBDIRres - peso da banda direita resfriada; PP - peso do pernil; PPL - peso do pernil limpo; PCOPA - peso da copa; PCOPAL - peso da copa limpa; PPA - peso da paleta; PPAL - peso da paleta limpa; PC - peso do carré; PL - peso do lombo; PB - peso do bacon; EBACON - espessura do bacon; PCOS - peso das costelas; PAPADA - peso da papada; CABEÇA - peso da cabeça; PF - peso do filezinho; PBR - peso da banha rama; e RIM - peso do rim.

Características de qualidade da carne

O número de observações, as médias corrigidas e os desvios-padrão das características de qualidade da carne são apresentados na Tabela 9. Após serem feitas as consistências dos dados, 326 observações foram utilizadas na análise.

Tabela 9 – Número de observações, médias e desvios-padrão das características de qualidade da carne

Característica	Número de observações	Média corrigida	Desvio-padrão *
pH ₄₅	510	6,49099	0,25714
pH _u	514	5,70665	0,14535
L	459	44,96616	1,97374
a	449	0,68580	0,59029
b	455	6,59036	0,51791
GORINT (%)	466	1,54064	0,56964
GOTEJ (%)	515	3,23606	1,66173
COZ (%)	508	32,63745	2,35889
MACIEZ (g/1,2 cm)	397	5550	841,61781
PTOT (%)	404	34,26022	2,64539

pH₄₅ - pH medido 45 minutos após o abate; pH_u - pH medido 24 horas após o abate; L – luminosidade; a - índice de vermelho; b - índice de amarelo; GORINT - gordura intramuscular; GOTEJ - perda por gotejamento; COZ - perda por cozimento; MACIEZ – maciez objetiva; e PTOT - perda total.

* Desvio-padrão dos dados corrigidos para efeitos fixos e covariáveis.

Os resultados obtidos para componentes principais, autovalores e percentagens da variância explicada por cada componente encontram-se na Tabela 10.

Tabela 10 – Componentes principais (CP), autovalores (λ_i) e percentagem da variância explicada pelos componentes (% var. CP), nas características de qualidade da carne

Componentes principais	λ_i	% var. CP	% var. CP (acumulada)
CP ₁	2,75699821	27,57	27,57
CP ₂	1,93548805	19,35	46,92
CP ₃	1,37236409	13,72	60,65
CP ₄	1,11265755	11,13	71,78
CP ₅	0,83768128	8,38	80,15
CP ₆	0,82546364	8,25	88,41
CP ₇	0,54030368	5,40	93,81
CP ₈	0,30719654	3,07	96,88
CP ₉	0,25340500	2,53	99,42
CP ₁₀	0,05844197	0,58	100

Os três primeiros componentes principais explicaram 60,65% da variação total, demonstrando que a técnica de componente principal foi efetiva para resumir o julgamento das características de qualidade da carne. Resultado semelhante foi

encontrado por DESTEFANIS et al. (2000), que, trabalhando com componentes principais nas análises químicas, físicas e sensoriais da carne de bovinos jovens, verificaram que, pelo fato de os três primeiros componentes terem explicado em torno de 63% da variação total, pode-se concluir que a técnica de componentes principais demonstra ser um procedimento muito efetivo para resumir o julgamento da qualidade da carne.

Conforme demonstrado na Tabela 10, quatro variáveis podem ser descartadas em razão do número de componentes principais, que apresentaram autovalores inferiores a 0,7, segundo o critério de JOLLIFFE (1972, 1973).

As quatro variáveis que apresentaram maiores coeficientes, em valor absoluto, nos últimos componentes principais são passíveis de descarte, conforme pode ser verificado na Tabela 11.

As variáveis passíveis de descarte são, respectivamente, em ordem de menor importância para explicar a variação total: perda de peso total, índice de amarelo, perda por gotejamento e maciez objetiva.

Com base nesses resultados, recomendam-se as seguintes variáveis para serem mantidas em experimentos futuros: pH medido aos 45 minutos após abate, pH medido 24 horas após abate, luminosidade, índice de vermelho, gordura intramuscular e perda por cozimento.

A variável perda de peso total foi a menos importante para explicar a variação total, por ser uma combinação linear da perda por gotejamento e perda por cozimento ($PTOT = GOTEJ + COZ$); em razão disso, foi a primeira variável a ser sugerida para descarte.

Segundo HUFF-LONERGAN et al. (2002), medidas de pH são comumente usadas para prever várias características de qualidade da carne. O pH_{45} e pH_u , por apresentarem baixa correlação entre si, foram sugeridos para serem mantidos em experimentos futuros. CASTEELS et al. (1995) verificaram correlação não-significativa entre estas duas características, o que corrobora os resultados encontrados por BENEVENUTO JUNIOR (2001) e aqueles verificados neste experimento (Tabela 12). Tanto o pH_{45} quanto o pH_u são importantes na determinação da qualidade da carne suína.

As características de índice de cor são muito importantes, porque na carne fresca, em geral, exercem influência na aceitabilidade do consumidor (JEREMIAH, 1994; BREWER e Mc KEITH, 1999). Neste trabalho, dois atributos de índice de cor

foram sugeridos para serem mantidos em experimentos futuros (luminosidade e índice de vermelho), sendo redundantes as demais características de índice de cor.

NORMAN et al. (2003), ao avaliarem a relação de índice de cor com a aceitabilidade do consumidor, observaram que a luminosidade apresentou correlação ($p < 0,01$), com índice de amarelo $r = 0,62$, resultado semelhante ao do presente trabalho ($r = 0,70$), justificando, em parte, o fato de a característica índice de amarelo ter sido passível de descarte.

Estudos têm demonstrado que maciez é um dos mais importantes atributos da qualidade da carne (VAN OECKEL et al., 1999a), visto que determina a aceitabilidade e a satisfação do consumidor. Embora a maciez seja uma importante característica de qualidade da carne, seu descarte se justifica em parte, por apresentar correlação significativa ($p < 0,01$) com índice de cor (DAVIS et al., 1975; HODGSON et al., 1991; NORMAN et al., 2003) e com pH_{45} , pH_u e índice de cor (BENEVENUTO JUNIOR, 2001).

A gordura intramuscular, não considerada passível de descarte no presente trabalho, tem influência positiva na suculência, na maciez, no sabor e na conservação da carne (GARCÍA-MACÍAS, 1996; BARROS, 2001). Além disso, pode ser de grande aplicabilidade no melhoramento genético de suínos, visto que o gene da IMF (Gordura Intramuscular) tem efeito no conteúdo de gordura intramuscular. Animais homozigotos para o alelo IMF apresentaram 3,9% de gordura intramuscular, comparado com 1,8% em animais não-portadores ou portadores de apenas um alelo (De Vries et al., 2000, citado por BENEVENUTO JUNIOR, 2001).

As características de qualidade da carne que não apresentaram correlação significativa com o teor de gordura intramuscular (GORINT) foram o pH (tanto aos 45 minutos quanto 24 horas pós-abate) e a luminosidade (L), motivo pelo qual se justifica, em parte, o fato de não serem consideradas passíveis de descarte. Correlação não-significativa da GORINT com medidas de pH_{45} e pH_u também foi verificada por BREDAHL et al. (1998).

Van OECKEL et al. (1999b) encontraram correlação significativa ($p < 0,001$) entre pH_{45} e perda por gotejamento, com um valor de $-0,34$, inferior ao encontrado no presente trabalho ($-0,66$), o que pode ser devido aos valores médios de pH_{45} de 5,86 naquele trabalho e de 6,49 neste, uma vez que, segundo WARRISS e BROWN (1987) e Van LAACK et al. (1994), a correlação entre pH_{45} e perda por gotejamento é melhor representada para valores de pH_{45} superiores a 6,1. PIRES (2003) encontrou QTLs

significativos para as características pH_{45} e perda por gotejamento, indicando que mesmos grupos gênicos podem estar atuando sobre ambas as características. A existência de correlação entre pH_{45} e perda por gotejamento explica, em parte, o fato de a perda por gotejamento ter sido sugerida para descarte.

A única característica relacionada com a capacidade de retenção de água sugerida para ser mantida em experimentos futuros foi a perda por cozimento, sendo as demais redundantes.

Na Tabela 12 estão apresentados os coeficientes de correlação simples entre as 10 características de qualidade da carne.

Tabela 11 – Coeficientes de ponderação das características e suas correlações (em porcentagem) com os últimos componentes principais menos importantes para explicar a variação total das 10 características de qualidade da carne

Característica	Coeficientes (Correlação - %)			
	CP ₁₀	CP ₉	CP ₈	CP ₇
pH ₄₅	0,026838 (0,649)	0,003597 (0,181)	0,651813 (36,127)	-0,272676 (-20,043)
pH _u	0,001842 (0,045)	0,097319 (4,899)	-0,001329 (-0,074)	0,098048 (7,207)
L	-0,022363 (-0,541)	-0,596230 (-30,014)	0,088130 (4,885)	0,486951 (35,793)
a	0,003241 (0,078)	0,017016 (0,857)	0,052993 (2,937)	0,546995 (40,207)
b	0,010190 (0,246)	0,716457 (36,066)	-0,176730 (-9,795)	0,126306 (9,284)
GORINT	-0,013946 (-0,337)	-0,050987 (-2,567)	-0,002452 (-0,136)	0,003815 (0,280)
GOTEJ	0,363777 (8,794)	0,175011 (8,810)	0,660508 (36,609)	0,092438 (6,795)
COZ	0,564991 (13,659)	-0,171275 (-8,622)	-0,273296 (-15,148)	-0,161015 (-11,835)
MACIEZ	0,015175 (0,367)	0,242389 (12,202)	0,061661 (3,418)	0,567306 (41,700)
PTOT	-0,739384 (-17,874)	-0,010480 (-0,528)	0,136233 (7,551)	-0,086230 (-6,338)

pH₄₅ - pH medido 45 minutos após o abate; pH_u - pH medido 24 horas após o abate; L – luminosidade; a - índice de vermelho; b - índice de amarelo; GORINT - gordura intramuscular; GOTEJ - perda por gotejamento; COZ - perda por cozimento; MACIEZ – maciez objetiva; e PTOT - perda total.

Tabela 12 – Coeficientes de correlação simples entre as características de qualidade da carne

	pH45	pH24	L	a	b	GORINT	GOTEJ	COZ	MACIEZ	PTOT
pH ₄₅	1,0000									
pH _u	0,0906	1,0000								
L	0,0221	-0,1822	1,0000							
a	-0,0618	0,0512	-0,3377	1,0000						
b	0,0387	-0,2632	0,6987	-0,2638	1,0000					
GORINT	-0,0998	0,0331	0,0004	-0,1051	0,0213	1,0000				
GOTEJ	-0,6680	-0,1274	0,1255	-0,0016	0,1396	0,1168	1,0000			
COZ	-0,1258	0,0610	0,0709	0,0371	0,1764	0,0534	0,2703	1,0000		
MACIEZ	0,2528	0,1668	-0,1687	-0,1163	-0,3280	0,0473	-0,3389	0,0081	1,0000	
PTOT	-0,3837	-0,0055	0,0924	0,0357	0,1877	0,0778	0,6342	0,8477	-0,1324	1,0000

pH₄₅ - pH medido 45 minutos após o abate; pH_u - pH medido 24 horas após o abate; L – luminosidade; a - índice de vermelho; b - índice de amarelo; GORINT - gordura intramuscular; GOTEJ - perda por gotejamento; COZ - perda por cozimento; MACIEZ – maciez objetiva; e PTOT - perda total.

CONCLUSÕES

Tendo em vista os resultados obtidos, podem ser descartadas 54,54%, 51,51% e 40,00% das características avaliadas de desempenho, de carcaça e de qualidade da carne, respectivamente.

Em razão do grande número de variáveis redundantes que podem ser descartadas, espera-se relativa economia de tempo e custo em experimentos futuros, sem causar perda considerável de informação.

LITERATURA CITADA

- BARROS, L.B. **Efeito de níveis de lisina da dieta sobre a qualidade da carne de fêmeas suínas abatidas em diferentes pesos.** Viçosa, MG: UFV, 2001. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos) - Universidade Federal De Viçosa, 2001.
- BENEVENUTO JUNIOR, A.A. **Avaliação de rendimento de carcaça e de qualidade da carne de suínos comerciais, de raça nativa e cruzados.** Viçosa, MG: UFV, 2001, 93p. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos) – Universidade Federal de Viçosa, 2001.
- BREDAHL, L.; GRUNERT, K.G.; FERTIN, C. Relating consumer perceptions of pork quality to physical product characteristics. **Food Quality and Preference**, v.9, p.273-281, 1998.
- BREWER, M.S.; McKEITH, F.K. Consumer-rated quality characteristics as related to purchase intent of fresh pork. **Journal of Food Science**, v.64, p.171–174, 1999.
- CASTEELS, M.; OECKEL, M.J.V.; BOSCHAERTS, L.; SPINCEMAILLE, G.; BOUEQUÉ, C.V. The relationship between carcass, meat and eating quality of three pig genotypes. **Meat Science**, v. 40, p. 253-269, 1995.
- CROSS, H.R.; SMITH, G.C.; CARPENTER, Z.L.; KOTULA, A.W. Relationship of carcass scores and measurements to five points for lean cut yields in barrow and gilt carcass. **Journal of Animal Science**, v. 41, p.1318-1326, 1975.
- DAHER, R. F.; MORAES, C. F.; CRUZ, C. D.; PEREIRA, A. V.; XAVIER, D. F. Seleção de caracteres morfológicos discriminantes em capim-elefante (*pennisetum purpureum schum.*). **Revista Brasileira de Zootecnia**, v.26, p.247-254, 1997.
- DAVIS, G.W.; SMITH, G.C.; CARPENTER, Z.L.; CROSS, H.R. Relationships of quality indicators to palatability attributes of pork loins. **Journal of Animal Science**, v.41, p.1305–1313, 1975.
- DESTEFANIS, G.; BARGE, M.T.; BRUGIAPAGLIA, A.; TASSONE, S. The use of principal component analysis (pca) to characterize beef. **Meat Science**, v.56, p.255-259, 2000.
- GARCÍA-MACÍAS, J.A.; GISPert, M.; OLIVER, M.A.; DIESTRE, A.; ALONSO, P.; MUÑOZ-LUNA, A.; SIGGENS, K.; CUTHBERT-HEAVENS, D. The effects of cross, slaughter weight and halothane genotype on leanness and meat and fat quality in pig carcasses. **Journal of Animal Science**, v.63, p.487-496, 1996.
- HODGSON, R.R.; DAVIS, G.W.; SMITH, G.C.; SAVELL, J.W.; CROSS, H.R. Relationships between pork loin palatability traits and physical characteristics of cooked chops. **Journal of Animal Science**, v.69, p.4856–4858, 1991.

- HUFF-LONERGAN, E.; BAAS, T.J.; MALEK, M.; DEKKERS, J.C.M.; PRUSA, K.; ROTHSCHILD M.F. Correlations among selected pork quality traits. **Jornal of Animal Science**. v.80, p.617–627, 2002.
- JEREMIAH, L.E. Consumer responses to pork loin chops with diferent degrees of muscle quality in two western Canadian. **Canadian Journal of Animal Science**, v.74, p.425–432, 1994.
- JOLLIFFE, I.T. Discarding variables in a principal component analysis. I. Artificial data. **Applied Statistics**, v.21, p.160-173, 1972.
- JOLLIFFE, I.T. Discarding variables in a principal component analysis. II. Real data. **Applied Statistics**, v.22, p.21-31, 1973.
- NORMAN, J.L.; BERG, E.P.; HEYMAN H.; LORENZEN, C.L. Pork loin color relative to sensory and instrumental tenderness and consumer acceptance. **Meat Science**. v.65, p.927–933, 2003.
- PIRES, A.V. **Mapeamento de locos de características quantitativas, no cromossomo seis suíno**. Viçosa, MG: UFV, 2003, 86p. Dissertação (Doutorado em Genética e Melhoramento) – Universidade Federal de Viçosa, 2003.
- REGAZZI, A. J. **Análise multivariada**. Viçosa: Universidade Federal de Viçosa, 2002. (INF-766) (notas de aula).
- SAS, 1999. Software: versão 8.0. SAS Institute Inc., Cary, NC, USA.
- STRAPASSON, E.; VENCOVSKY, R.; BATISTA, L.A.R. Seleção de descritores na caracterização de germoplasma de paspalum sp. por meio de componentes principais. **Revista Brasileira de Zootecnia**, v.29, p.373-381, 2000.
- TORRES FILHO, R.A. **Avaliação genética de características de desempenho e reprodutivas em suínos**. Viçosa, MG: UFV, 2001, 79p. Dissertação (Mestrado em Genética e Melhoramento) – Universidade Federal de Viçosa, 2001.
- Van LAACK, R.L.J.M.; KAUFFMAM, R.G.; SYBESMA, W.; SMULDERS, F.J.M.; EIKELNBONN,G.; PINHEIRO J.C. Is colour brightness (L-value) a reliable indicator of water-holding capacity in porcine muscle. **Meat Science**, v.38, p.193-201, 1994.
- Van OECKEL, M.J.; WARNANTS, N.; BOUCQUÉ, Ch.V. Pork tenderness estimation by taste panel, Warner-Bratzler shear force and on-line methods. **Meat Science**. v.53 p.259-267, 1999a.
- Van OECKEL, M.J.; WARNANTS, N.; BOUEQUÉ, Ch.V. Comparison of different methods for measuring water hoding capacity and juiciness of pork versus on-lines screening methods. **Meat Science**, v.51, p.313-320, 1999b.
- WARRIS, P.D.; BROWN, S.N. The relationships between initial pH, reflectance and exudation in pigs muscle. **Meat Science**, v. 20, p.65-74, 1987.

CAPÍTULO II

Estudo da Associação entre Características Quantitativas de Suínos por meio de Correlação Canônica

Resumo: Dados de 844 animais de uma população F2 de suínos foram utilizados para avaliar a associação entre características de desempenho, carcaça e qualidade da carne, por meio de correlação canônica. Foram consideradas as seguintes características de desempenho: tamanho da leitegada ao nascimento (TLN), número de tetos (NT), peso aos 77 dias de idade (P77), consumo de ração (CR) e conversão alimentar dos 77 aos 105 dias (CA); de carcaça: idade ao abate (IDA), peso da banda direita (PBDIR), comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça (MBCC), maior espessura de toucinho na região da copa, na linha dorso-lombar (SH), profundidade de lombo (PROFLOMB), pesos do coração (COR), do baço (BACO), do rim (RIM), da banda direita resfriada (PBDIRRES), do pernil (PP), da copa (PCOPA), da paleta (PPA), do carré (PC), das costelas (PCOS), da papada (PAPADA) e do filezinho (PF); e de qualidade da carne: pH medido aos 45 minutos e 24 horas *post-mortem* (pH₄₅, pH_u, respectivamente), luminosidade (L), gordura intramuscular (GORINT), perda por cozimento (COZ) e índice de vermelho (a). Observou-se que os grupos características de desempenho e características de carcaça não são considerados independentes, pois foram fortemente correlacionados, visto que a correlação do primeiro par canônico (r) foi de 0,78. Os dois primeiros pares canônicos foram estatisticamente relevantes ($P < 0,01$ e $P < 0,05$, respectivamente). No exame das variáveis que compõem o primeiro par canônico, para as características de desempenho, houve predomínio absoluto da variável P77; para as características de carcaça, a variável IDA teve predomínio absoluto. No segundo par canônico, para as características de desempenho, houve predomínio absoluto da variável NT; para as características de carcaça, a variável PCOS teve a maior contribuição, juntamente com a variável PBDIR. Para os grupos características de desempenho e características de qualidade da carne, observou-se que existe independência, pois a correlação do primeiro par canônico (r) foi de 0,23, sendo estatisticamente irrelevante ($P > 0,05$). Para os grupos características de carcaça e características de qualidade da carne, observou-se que não existe independência, sendo os três primeiros pares canônicos estatisticamente relevantes ($P < 0,01$). No exame das variáveis que compõem o primeiro par canônico, para as características de carcaça, houve predomínio da variável PF, seguida,

principalmente, pela variável MBCC; para as características de qualidade da carne, a variável L teve predomínio absoluto. No exame das variáveis que compõem o segundo par canônico, para as características de carcaça, houve maior predomínio da variável PC, seguida, principalmente, das variáveis MBCC, SH, PPA e PBDIRres; para as características de qualidade da carne, a variável GORINT teve predomínio absoluto. Em se tratando das variáveis que compõem o terceiro par canônico, para as características de carcaça, houve maior predomínio da variável PBDIRres; para as características de qualidade da carne, a variável COZ teve maior predomínio, seguida, principalmente, da variável “a”.

Palavras-chave: carcaça, correlação, desempenho, pares canônicos, qualidade da carne, variáveis canônicas

Study of Swine Quantitative Traits Using Canonical Correlation

Abstract: Records of 844 animals of a swine population were used to evaluate the association among the performance, carcass and meat quality traits, using canonical correlation analysis. The following traits were used: 1 - performance: litter size at birth (BL), teat number (TN), weight at 77 days of age (W77), feed intake (FI) and feed:gain ratio (FGR) from 77 to 105 days of age; 2 - carcass: slaughter age (SA), right half carcass weight (HCW), carcass length by the Brazilian Carcass Classification Method (BLC), higher backfat thickness at boston shouder area, in the midline (SH), loin depth (LD), heart weight (HW), spleen weight (SPLW), kidney weight (KW), cold right half carcass weight (CCW), ham weight (HAW), boston shoulder weight (BSW), picnic shoulder weight (PSW), loin (bone-in) weight (BLW), spareribs weight (SPW), jowl weight (JW) and sirloin weight (SLW); and 3 - meat quality: pH avaluated at 45 minutes after slaughter (pH_{45}), pH avaluated at 24 hours after slaughter (pH_u), lightness (L), intramuscular fat (IMF), cooking loss (CL) and redness (a). It was observed that performance and carcass traits, are not considered independent, because they were strongly correlated, since the first canonical correlation (r) was 0.78. The first two canonical variate pairs were estatistically significant ($P < 0.01$ and $P < 0.05$, respectively). In the performance traits, there was an absolute predominance of the W77 variable, in the first canonical variate pair; and in the carcass traits, SA had absolute predominance. In the second canonical variate pair, there was an absolute predominance for the TN variable (performance traits) and SPW had the largest contribution followed by the HCW variable, in the carcass traits. It was observed independence in the sets of performance and meat quality traits, because the first canonical correlation (r) was 0.23 ($P > 0.05$). There was no independence between the sets of carcass and meat quality traits, which the first three canonical variate pairs are estatistically significant ($P < 0.01$). In the carcass traits, there was a predominance of the SLW variable, followed mainly, by the BLC, in the first canonical variate pair; in the meat quality traits, L had absolute predominance. In the second canonical variate pair, in the carcass traits, there was larger predominance of the BLW variable, followed mainly, by the BLC, SH, PSW e CCW; and IMF had absolute predominance in the meat quality traits. In the third canonical variate pair, there was larger predominance of the CCW, in the carcass traits; and CL had larger predominance, followed mainly, by the ‘a’ variable in the meat quality traits.

Key Words: canonical variates, canonical variate pair, carcass, correlation, meat quality, performance

INTRODUÇÃO

O melhoramento animal ao longo de várias gerações altera a composição gênica dos indivíduos, buscando atender a demanda dos produtores e consumidores. O objetivo do melhoramento genético de suínos é a melhoria no desempenho, no rendimento de carcaça e na qualidade da carne.

A seleção de suínos para melhoria na taxa de crescimento, na eficiência alimentar e nas características de rendimento de carcaça tem causado piora na qualidade da carne. Essas respostas correlacionadas negativas para piora na qualidade da carne ocorreram em razão da correlação genética existente entre as características.

Identificar e quantificar a associação entre as características de desempenho, de carcaça e de qualidade da carne são de fundamental importância para promover seleção mais eficiente, de forma a atender os objetivos do melhoramento genético de suínos. Um método que possibilita o estudo dessa associação entre dois conjuntos de variáveis é a análise de correlação canônica.

A análise de correlação canônica busca identificar e quantificar a associação entre dois conjuntos de variáveis (JOHNSON e WICHERN, 1998). Segundo REGAZZI (2002), esta análise se baseia na determinação de variáveis canônicas ortogonais em cada conjunto de variáveis; por isso, as variáveis originais em cada conjunto devem ser linearmente independentes. Se essa condição não for satisfeita de início, será necessário descartar as variáveis que são combinações lineares das demais (variáveis redundantes).

O coeficiente de correlação canônica, para cada par canônico, mede a intensidade da correlação, enquanto a composição das variáveis canônicas exprime a natureza da associação.

Os objetivos do presente trabalho foram identificar e quantificar a associação entre as características de desempenho e carcaça, entre características de desempenho e qualidade da carne e entre características de carcaça e qualidade da carne.

MATERIAL E MÉTODOS

Os dados foram obtidos de um experimento conduzido na Granja de Melhoramento de Suínos do Departamento de Zootecnia da Universidade Federal de Viçosa, em Viçosa – MG, no período de novembro de 1998 a julho de 2001.

Foram formadas inicialmente duas famílias, provenientes do cruzamento de dois varrões da raça Brasileira (Piau) com 18 fêmeas originadas de linhagem desenvolvida na UFV, pelo acasalamento de animais das raças Landrace, Large White e Pietran. A geração F1 nasceu nos meses de março a maio de 1999 e era composta por 106 fêmeas e 134 machos. Dentre os machos F1, foram selecionados 11 varrões, provenientes de diferentes leitegadas, os quais foram acasalados (monta natural) com 54 fêmeas, evitando-se o acasalamento de parentes. Esses animais foram acasalados, nos meses de fevereiro a outubro de 2000, para produção da geração F2, a qual nasceu entre junho de 2000 e fevereiro de 2001. Assim, foram obtidos 844 animais F2, divididos em cinco lotes:

- lote 1 = animais nascidos entre 20/06/00 e 03/07/00;
- lote 2 = animais nascidos entre 03/08/00 e 23/08/00;
- lote 3 = animais nascidos entre 16/09/00 e 01/11/00;
- lote 4 = animais nascidos entre 30/11/00 e 25/12/00; e
- lote 5 = animais nascidos entre 19/01/01 e 12/02/01.

Os três primeiros lotes foram obtidos de porcas do primeiro parto; os dois últimos já eram porcas de segunda parição.

Foram avaliados 844 animais da geração F2 para características de desempenho, dos quais 550 foram abatidos. Os animais experimentais – fêmeas e machos castrados – foram abatidos com média de 64,79 kg e desvio-padrão de 5,06 kg.

Informações detalhadas sobre o manejo dos animais foram descritas por PIRES (2003).

Características Avaliadas

As características avaliadas foram previamente submetidas à análise de componentes principais para eliminação das variáveis redundantes (conforme descrito no capítulo I), uma vez que, na análise de correlações canônicas, cada conjunto de

variáveis deve ser linearmente independente (KHATTREE e NAIK, 2000; REGAZZI, 2002).

VAINIONPÄÄ et al. (2000) e DUFOUR et al. (2001), para obter variáveis linearmente independentes em seus conjuntos de dados, utilizaram a análise de componentes principais para posteriormente empregar a análise de correlação canônica.

Cabe ressaltar que antes de efetuar a análise de componentes principais foi realizado um teste para efeito de multicolinearidade (conforme descrito no capítulo I), em que foram eliminadas de início variáveis linearmente dependentes.

Características de desempenho

Foram avaliadas as seguintes características de desempenho: tamanho da leitegada ao nascimento (TLN), número de tetos (NT), peso aos 77 dias de idade (P77), consumo de ração (CR) e conversão alimentar dos 77 aos 105 dias (CA).

Características de carcaça

Avaliaram-se as seguintes características de carcaça: idade ao abate (IDA); peso da banda direita (PBDIR); comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça (MBCC); maior espessura de toucinho na região da copa, na linha dorso-lombar (SH); profundidade de lombo (PROFLOMB); e pesos do coração (COR), do baço (BACO), do rim (RIM), da banda direita resfriada (PBDIRres), do pernil (PP), da copa (PCOPA), da paleta (PPA), do carré (PC), das costelas (PCOS), da papada (PAPADA) e do filezinho (PF).

Características de qualidade da carne

Foram avaliadas as seguintes características de qualidade da carne suína: pH medido aos 45 minutos e 24 horas *post-mortem* (pH₄₅ e pH_u, respectivamente), gordura intramuscular (GORINT), perda por cozimento (COZ), índice de vermelho (a) e luminosidade (L).

As metodologias empregadas na avaliação das características de qualidade da carne foram descritas por BENEVENUTO JUNIOR (2001).

Metodologia Estatística

Os dados submetidos à análise de correlação canônica foram previamente ajustados para efeitos fixos e covariáveis. O modelo estatístico adotado foi:

$$y_{ijk} = \mu + S_i + L_j + b(C_{ijk} - \bar{C}) + e_{ijk}$$

em que:

y_{ijk} = característica observada no animal k , no lote j , no sexo i ;

μ = constante inerente a toda observação;

S_i = efeito fixo do sexo i , $i = 1, 2$ (1 = macho castrado e 2 = fêmea);

L_j = efeito fixo do lote j , $j = 1, 2, 3, 4, 5$;

C_{ijk} = valor observado da covariável, no animal k , no lote j , no sexo i ;

\bar{C} = média da covariável;

b = coeficiente de regressão linear da característica, em função da covariável; e

e_{ijk} = erro aleatório associado a cada observação.

Com base no trabalho de PIRES (2003), para ajuste dos dados, foram utilizadas as seguintes características e respectivas covariáveis:

- Característica P77 – tamanho de leitegada a desmama (TLD);
- Características CR e CA – peso aos 77 dias de idade (P77);
- Características de carcaça – peso da carcaça (PCARC);
- Características de cortes de carcaça – peso da banda direita resfriada (PBDIRres); e
- Características de qualidade da carne – idade ao abate (IDA).

Análise de Correlações Canônicas

A técnica de correlação canônica foi utilizada para analisar as relações entre dois conjuntos de variáveis, em que foi estimada a máxima correlação entre combinações lineares das variáveis.

Foram calculadas, inicialmente, duas combinações lineares (variáveis canônicas), uma sobre cada conjunto de variáveis, de forma que os coeficientes de correlação de Pearson entre elas fossem máximos. Em seguida, foram calculadas duas outras variáveis canônicas, uma sobre cada conjunto de variáveis, formando o segundo par canônico, com a condição de que as variáveis canônicas dentro de cada conjunto fossem ortogonais, e assim sucessivamente.

As variáveis canônicas têm as seguintes propriedades:

$$Var(U_k) = Var(V_k) = 1$$

$$Cov(U_k, U_l) = Corr(U_k, U_l) = 0 \quad k \neq l$$

$$Cov(V_k, V_l) = Corr(V_k, V_l) = 0 \quad k \neq l$$

$$Cov(U_k, V_l) = Corr(U_k, V_l) = 0 \quad k \neq l$$

para $k, l = 1, 2, \dots, p$.

Os coeficientes de correlação canônica entre as variáveis dos pares sucessivos são obviamente decrescentes.

A análise de correlação canônica visa encontrar combinações lineares $U = \underline{a}' \underline{X}$ e $V = \underline{b}' \underline{Y}$ de tal modo que $r^2(U, V)$ seja máxima.

Por definição:

$$r^2(\underline{a}' \underline{X}, \underline{b}' \underline{Y}) = \frac{(\underline{a}' S_{12} \underline{b})^2}{(\underline{a}' S_{11} \underline{a})(\underline{b}' S_{22} \underline{b})},$$

e o problema da correlação canônica pode ser resolvido encontrando-se a solução para:

$\max_{\underline{a}, \underline{b}} (\underline{a}' S_{12} \underline{b})^2$, sujeito às restrições $\underline{a}' S_{11} \underline{a} = 1$ e $\underline{b}' S_{22} \underline{b} = 1$. As duas restrições

referem-se ao fato de que a correlação é invariante à mudança de escala. Essa é uma propriedade importante: em lugar das matrizes de covariâncias S_{11} , S_{22} e S_{12} , podem-se utilizar as matrizes de correlações R_{11} , R_{22} e R_{12} (KHATTREE e NAIK, 2000).

No entanto, as variáveis foram padronizadas conforme descrito para os componentes principais (capítulo I). Nesse caso, têm-se $S_{11} = R_{11}$, $S_{22} = R_{22}$ e $S_{12} = R_{12}$, em que \mathbf{R} representa uma matriz de correlação amostral.

A primeira correlação canônica (r_1) entre as combinações lineares $U_1 = \underline{a}' \underline{X}$ e $V_1 = \underline{b}' \underline{Y}$ é dada por: $r_1 = \sqrt{\lambda_1}$, sendo λ_j (autovalores) determinado

pela equação característica: $|\mathbf{R}_{22}^{-1} \mathbf{R}_{21} \mathbf{R}_{11}^{-1} \mathbf{R}_{12} - \lambda \mathbf{I}| = 0$.

Os coeficientes (autovetores) das variáveis canônicas (\underline{a} e \underline{b}) são determinados resolvendo a equação: $\left[R_{22}^{-1} R_{21} R_{11}^{-1} R_{12} - \lambda I \right] \underline{b} = 0$, e o vetor \underline{a} é obtido de:

$$\underline{a} = \frac{1}{\sqrt{\lambda}} R_{11}^{-1} R_{12} \underline{b}$$

Os autovetores \underline{a} e \underline{b} foram normalizados por $\underline{a}' R_{11} \underline{a} = 1$ e $\underline{b}' R_{22} \underline{b} = 1$, respectivamente.

Os coeficientes são escolhidos de tal modo que:

$$r_1^2(U_1, V_1) > r_2^2(U_2, V_2) > \dots > r_j^2(U_j, V_j)$$

isto é, $\lambda_1 > \lambda_2 > \dots > \lambda_j$, que são os j autovalores não-nulos de $R_{22}^{-1} R_{21} R_{11}^{-1} R_{12}$, para $j = \min(p, q)$.

Fica evidenciado, então, que o coeficiente de correlação canônica de ordem j é a raiz quadrada do autovalor λ_j , ou seja:

$$r_j = \frac{\underline{a}'_j R_{12} \underline{b}_j}{\left[\left(\underline{a}'_j R_{11} \underline{a}_j \right) \left(\underline{b}'_j R_{22} \underline{b}_j \right) \right]^{1/2}} = \underline{a}'_j R_{12} \underline{b}_j = \sqrt{\lambda_j} \text{ pois } \underline{a}'_j R_{11} \underline{a}_j \text{ e } \underline{b}'_j R_{22} \underline{b}_j$$

são iguais a um.

Usando as variáveis padronizadas (Z), a correlação entre as variáveis originais e as variáveis canônicas é dada por:

$$\text{corr}(Z_y, U_j) = R_{11} \underline{a}_j$$

$$\text{corr}(Z_x, V_j) = R_{22} \underline{b}_j$$

para $j = 1, 2, \dots, p$

Teste de significância

Foi utilizado o teste da razão de máxima verossimilhança, cuja estatística é a Λ de Wilks, obtido pelo procedimento PROC CANCORR do SAS (1999); cujas expressões são apresentadas por KHATTREE e NAIK (2000).

Para testar:

H_0 : A s -ésima correlação canônica e todas que se seguem são zero

vs H_a : não H_0

A estatística:

$$\Lambda_s = \prod_{i=s+1}^k (1-r_i^2)$$

é transformada na estatística F , dada por:

$$F = \left(\frac{df_2}{df_1} \right) \left(\frac{1-\Lambda_s^{1/m_s}}{\Lambda_s^{1/m_s}} \right),$$

$$m_s = \left\{ \frac{(p-s)^2 (q-s)^2 - 4}{(p-s)^2 + (q-s)^2 - 5} \right\}^{1/2}, \text{ quando } (p-s)(q-s) = 2, \text{ então } m_s = 1$$

e:

$$df_1 = (p-s)(q-s)$$

$$df_2 = m_s [n - (p+q+3)/2] - [(p-s)(q-s) - 2]/2$$

$F_{\text{tabelado}} = F_{\alpha}(df_1, df_2)$, sendo α o nível de significância.

A regra de decisão é a seguinte:

se $F_{\text{calc}} \geq F_{\text{tab}}$ rejeita-se H_0 (estatisticamente relevante, isto é, a associação existe); caso contrário, não se rejeita H_0 .

Análise dos Dados

Todas as análises foram feitas utilizando o programa SAS System for Windows^{NT}, versão 8.0, licenciado pela UNIVERSIDADE FEDERAL DE VIÇOSA (SAS Institute Inc., Cary, NC, USA, 1999).

RESULTADOS E DISCUSSÃO

Características de desempenho e carcaça

Na Tabela 1 são apresentados os coeficientes padronizados, a correlação canônica, a correlação canônica ao quadrado e o teste estatístico para os pares canônicos entre características de desempenho e carcaça.

Tabela 1 – Coeficientes padronizados, correlação canônica (r), correlação canônica ao quadrado (r^2) e teste estatístico para os pares canônicos entre características de desempenho e carcaça

Variáveis	Pares canônicos				
	1º	2º	3º	4º	5º
TLN	0,0574	-0,0978	-0,1984	0,4378	0,8726
NT	-0,0222	0,9258	0,0417	0,3888	-0,0793
P77	0,8058	-0,3030	-0,1270	0,4172	-0,2907
CR	0,4321	0,1032	0,7705	-0,3790	0,2789
CA	-0,3714	-0,2827	0,6670	0,5753	-0,1152
IDA	-0,8665	0,2012	0,0527	-0,1387	0,0004
PBDIR	-0,0906	-0,4946	0,0714	0,1187	0,0135
MBCC	-0,0489	0,2699	-0,4772	0,1523	-0,2339
SH	0,0327	-0,3301	0,2454	-0,1361	0,1821
PROFLOMB	0,0078	0,2229	-0,0231	-0,2363	0,1318
COR	-0,1276	-0,1409	0,1968	-0,0897	-0,4514
BACO	0,0306	-0,0956	-0,1663	0,0221	-0,1309
PBDIRres	0,4794	0,2853	0,1828	-0,5670	-0,2341
PP	0,0605	-0,1999	-0,0204	0,1151	-0,2085
PCOPA	0,0944	0,0624	-0,2023	0,0600	0,4894
PPA	0,0167	-0,3201	-0,2455	0,2547	-0,1842
PC	0,0540	0,3778	0,2739	0,7713	-0,1946
PCOS	0,0608	0,5194	-0,0917	-0,0863	0,2531
PAPADA	-0,0249	-0,1178	0,0193	0,2908	0,2735
PF	0,0267	-0,2605	-0,2691	-0,0691	0,2978
RIM	0,0585	0,1802	0,2044	0,1900	0,1262
r	0,780449	0,337017	0,260733	0,193052	0,124935
F	5,45 **	1,37 *	0,98 ns	0,68 ns	0,43 ns
r^2	0,609101	0,113581	0,092419	0,049492	0,030255

TLN - tamanho da leitegada ao nascimento; NT - número de tetos; P77 - peso aos 77 dias de idade; CR - consumo de ração; CA - conversão alimentar dos 77 aos 105 dias de idade; IDA - idade ao abate; Pbdir - peso da banda direita; MBCC - comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça; SH - maior espessura de toucinho na região da copa, na linha dorso-lombar; PROFLOMB - profundidade de lombo; COR - peso do coração; BACO - peso de baço; PBDIRres - peso da banda direita resfriada; PP - peso do pernil; PCOPA - peso da copa; PPA - peso da paleta; PC - peso do carré; PCOS - peso das costelas; PAPADA - peso da papada; PF - peso do filezinho; RIM - peso do rim.

** Significativo a 1% ($P < 0,01$).

* Significativo a 5% ($P < 0,05$).

ns - Não-significativo a 5% ($P > 0,05$).

Conforme apresentado na Tabela 1, observa-se que os grupos características de desempenho e características de carcaça não são considerados independentes. Os dois primeiros pares canônicos são estatisticamente relevantes ($P < 0,01$ e $P < 0,05$, respectivamente). Isso mostra que as características de desempenho e carcaça avaliadas são fortemente correlacionadas, pois a correlação do primeiro par canônico (r) foi de 0,78.

Primeiro par canônico

Em razão de existir mais de um par canônico estatisticamente relevante, a discussão foi feita para cada par canônico separadamente.

A proporção de variância comum para as duas variáveis canônicas no primeiro par canônico, que é dado pelo coeficiente de correlação canônica ao quadrado (r^2), foi de 60,91%.

O exame das variáveis que compõem o primeiro par canônico apresenta, do lado das características de desempenho, predomínio absoluto da variável peso aos 77 dias de idade (P77). Do lado das características de carcaça, a variável idade ao abate (IDA) teve predomínio absoluto.

Com o objetivo de auxiliar no procedimento de interpretação, alguns artifícios têm sido desenvolvidos, a fim de ajudar o pesquisador a identificar que tipo de relações está sendo representado por um coeficiente de correlação canônica; um desses artifícios é representado pela correlação entre as variáveis originais e as variáveis canônicas (ABREU e VETTER, 1978).

No exame de correlação entre a variável original e sua correlação canônica, confirma-se que **P77** é a mais importante na variável canônica U_1 , e a variável **IDA** é a mais importante na variável canônica V_1 , conforme pode ser observado na Tabela 2. Observa-se que maior peso aos 77 dias de idade (característica de desempenho) proporciona uma menor idade ao abate (característica de carcaça). Cabe ressaltar que, em muitos casos, a interpretação com base nos coeficientes padronizados e nas correlações entre as variáveis originais e as variáveis canônicas podem ser muito diferentes.

Tabela 2 – Correlação entre as características de desempenho e suas variáveis canônicas (U) e entre as características de carcaça e suas variáveis canônicas (V)

	U ₁	U ₂	U ₃	U ₄	U ₅
TLN	0,0717	-0,0749	-0,1676	0,3844	0,9019
NT	0,0986	0,8981	0,0629	0,4124	-0,0984
P77	0,8027	-0,1955	-0,0838	0,4548	-0,3219
CR	0,4804	0,1613	0,7076	-0,3650	0,3305
CA	-0,3870	-0,3019	0,6120	0,5968	-0,1686
	V ₁	V ₂	V ₃	V ₄	V ₅
IDA	-0,8614	0,1591	0,1260	-0,2493	-0,0370
PBDIR	0,0674	-0,3174	0,0191	0,1371	0,1922
MBCC	-0,0072	0,2852	-0,6674	0,2210	-0,2952
SH	-0,0740	-0,2383	0,5781	0,0054	0,1571
PROFLOMB	-0,0306	0,2243	-0,0822	-0,1786	0,0671
COR	-0,2023	-0,0995	-0,0577	-0,0794	-0,4714
BACO	-0,0037	-0,0602	-0,1782	0,0013	-0,2480
PBDIRres	0,4621	0,1996	0,2002	-0,5298	-0,2298
PP	0,0512	-0,2616	0,0258	0,1401	-0,3652
PCOPA	0,0589	-0,0666	-0,3380	0,1476	0,3547
PPA	0,0281	-0,1642	-0,5794	-0,0346	-0,2788

TLN - tamanho da leitegada ao nascimento; NT - número de tetos; P77 – peso aos 77 dias de idade; CR - consumo de ração; CA - conversão alimentar dos 77 aos 105 dias de idade; IDA - idade ao abate; Pbdir - peso da banda direita; MBCC - comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça; SH - maior espessura de toucinho na região da copa, na linha dorso-lombar; PROFLOMB - profundidade de lombo; COR - peso do coração; BACO - peso de baço; PBDIRres - peso da banda direita resfriada; PP - peso do pernil; PCOPA - peso da copa; PPA - peso da paleta; PC - peso do carré; PCOS - peso das costelas; PAPADA - peso da papada; PF - peso do filezinho; RIM - peso do rim.

Com base na Tabela 2, tem-se que:

$$U_{y_1}^2 = \frac{(0,0717)^2 + (0,0986)^2 + (0,8027)^2 + (0,4804)^2 + (-0,3870)^2}{5} = 0,2079$$

e

$$V_{x_1}^2 = \frac{(-0,8614)^2 + (0,0674)^2 + (-0,0072)^2 + \Lambda + (-0,0146)^2 + (0,0375)^2}{16} = 0,0643$$

Assim, 20,79% da variância das características de desempenho é explicada pela variável canônica U_1 , e 6,43% da variância das características de carcaça, pela variável canônica V_1 .

Segundo par canônico

A proporção de variância comum para as duas variáveis canônicas no segundo par canônico, que é dada pelo coeficiente de correlação canônica ao quadrado (r^2), foi de 11,36%.

O exame das variáveis que compõem o segundo par canônico apresenta, do lado das características de desempenho, predomínio absoluto da variável número de tetos (NT). Do lado das características de carcaça, a variável peso das costelas (PCOS) teve a maior contribuição, seguida pela variável peso da banda direita, conforme apresentado na Tabela 1.

No exame de correlação entre a variável original e sua correlação canônica, confirma-se que a variável **NT** é a mais importante na variável canônica U_2 , e a variável **PCOS** é a mais importante na variável canônica V_2 ; a variável peso da papada teve contribuição desprezível, conforme pode ser constatado na Tabela 2. Observa-se que, quanto maior o número de tetos (características de desempenho), maior é o peso das costelas (características de carcaça).

Com base na Tabela 2, tem-se que:

$$U_{y_2}^2 = \frac{(-0,0749)^2 + (0,8981)^2 + (-0,1955)^2 + (0,1613)^2 + (-0,3019)^2}{5} = 0,1935$$

e

$$V_{x_2}^2 = \frac{(0,1591)^2 + (-0,3174)^2 + (0,2852)^2 + \Lambda + (-0,2158)^2 + (0,1919)^2}{16} = 0,05196$$

Assim, 19,35% da variância das características de desempenho é explicada pela variável canônica U_2 , e 5,20% da variância das características de carcaça, pela variável canônica V_2 .

Características de desempenho e qualidade da carne

Na Tabela 3 são apresentados os coeficientes padronizados, a correlação canônica, a correlação canônica ao quadrado e o teste estatístico para os pares canônicos entre características de desempenho e qualidade da carne.

Tabela 3 – Coeficientes padronizados, correlação canônica (r), correlação canônica ao quadrado (r^2) e teste estatístico para os pares canônicos entre características de desempenho e qualidade da carne

Variáveis	Pares canônicos				
	1º	2º	3º	4º	5º
TLN	-0,3418	-0,2248	0,6472	-0,4730	-0,4507
NT	0,6850	0,1821	0,5219	0,4367	-0,2468
P77	-0,3108	0,8751	-0,1805	-0,1952	-0,3028
CR	0,2296	0,3072	0,3132	-0,4510	0,7470
CA	0,5599	-0,1395	-0,3432	-0,6732	0,3254
pH ₄₅	0,0281	-0,1590	0,9283	-0,2358	-0,2890
pH _u	-0,3209	-0,3911	0,0327	0,1370	0,8904
L	-0,2255	-0,8303	-0,0712	0,6015	-0,1597
a	0,9031	-0,4008	-0,0258	0,3397	0,1054
GORINT	0,1654	0,3079	0,4168	0,2914	0,2152
COZ	-0,0813	0,4821	0,2732	0,6179	-0,0672
r	0,229934	0,170860	0,137675	0,054359	0,033925
F	1,07 ns	0,79 ns	0,58 ns	0,21 ns	0,17 ns
r^2	0,052870	0,029193	0,018954	0,002955	0,001151

TLN - tamanho da leitegada ao nascimento; NT - número de tetos; P77 - peso aos 77 dias de idade; CR - consumo de ração; CA - conversão alimentar dos 77 aos 105 dias de idade; pH₄₅ - pH medido 45 minutos após o abate; pH_u - pH medido 24 horas após o abate; L - luminosidade; a - índice de vermelho; GORINT - gordura intramuscular; e COZ - perda por cozimento.

ns - Não-significativo a 5% ($P > 0,05$).

Conforme apresentado na Tabela 3, observa-se que os grupos características de desempenho e características de qualidade da carne são considerados independentes. Isso mostra que as características de desempenho e qualidade avaliadas são não-correlacionadas, pois a correlação do primeiro par canônico (r) foi de 0,23, sendo estatisticamente não-significativa ($P > 0,05$).

A proporção de variância comum para as duas variáveis canônicas, que é dada pelo coeficiente de correlação canônica ao quadrado (r^2), foi de apenas 5,28%.

Características de carcaça e qualidade da carne

Na Tabela 4 são apresentados os coeficientes padronizados, a correlação canônica, a correlação canônica ao quadrado e o teste estatístico para os pares canônicos entre características de carcaça e qualidade da carne.

Tabela 4 – Coeficientes padronizados, correlação canônica (r), correlação canônica ao quadrado (r^2) e teste estatístico para os pares canônicos entre características de carcaça e qualidade da carne

Variáveis	Pares canônicos					
	1º	2º	3º	4º	5º	6º
IDA	0,3015	-0,0613	0,4134	-0,1833	-0,2936	-0,0969
Pbdir	0,1635	-0,1617	-0,0399	-0,3311	-0,4008	-0,3306
MBCC	0,5109	0,4722	0,2894	0,3501	0,0141	-0,1274
SH	0,2422	-0,4416	0,1730	0,2291	0,3457	0,1283
PROFLOMB	-0,0316	-0,1181	-0,1692	0,1011	0,3617	-0,3363
COR	-0,1156	0,0687	0,2466	0,4745	-0,1111	0,0414
BACO	-0,3267	-0,1153	0,2362	0,2335	-0,1904	-0,1090
PBDIRres	-0,1055	0,4083	-0,4600	0,0747	0,1412	-0,1321
PP	-0,0582	0,2301	0,0568	-0,3975	0,2849	0,1784
PCOPA	-0,0722	-0,0619	0,1217	-0,1025	0,1095	-0,2194
PPA	-0,2244	-0,4237	-0,0836	0,1818	0,1909	0,2150
PC	-0,3861	0,5081	0,3965	-0,2539	0,1302	0,2321
PCOS	0,2249	0,2448	-0,2346	-0,0464	-0,0723	-0,1773
PAPADA	-0,2131	-0,1894	0,3285	0,0854	0,6173	-0,6017
PF	-0,5244	0,0155	0,0298	0,1743	-0,2223	0,2310
RIM	0,1521	0,1574	0,0212	0,0431	0,3363	0,4343
pH ₄₅	0,3221	0,3271	0,0975	0,2204	0,8774	-0,1828
pH _u	0,3982	-0,1583	-0,4788	0,0894	-0,0554	0,8212
L	0,9003	-0,1678	-0,0280	0,3388	-0,4583	-0,1856
a	0,1293	0,2567	0,6193	0,7192	-0,2550	0,2849
GORINT	0,2401	0,8757	0,0765	-0,3765	-0,2137	0,0496
COZ	-0,2710	0,3682	-0,6680	0,5476	0,0700	-0,2954
r	0,413829	0,348141	0,334569	0,297722	0,211018	0,175166
F	1,92 **	1,69 **	1,55 **	1,30 ns	0,95 ns	0,84 ns
r ²	0,171255	0,121202	0,111936	0,088638	0,044529	0,030683

IDA - idade ao abate; Pbdir - peso da banda direita; MBCC - comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça; SH - maior espessura de toucinho na região da copa, na linha dorso-lombar; PROFLOMB - profundidade de lombo; COR - peso do coração; BACO - peso de baço; PBDIRres - peso da banda direita resfriada; PP - peso do pernil; PCOPA - peso da copa; PPA - peso da paleta; PC - peso do carré; PCOS - peso das costelas; PAPADA - peso da papada; PF - peso do filezinho; RIM - peso do rim; pH₄₅ - pH medido 45 minutos após o abate; pH_u - pH medido 24 horas após o abate; L - luminosidade; a - índice de vermelho; GORINT - gordura intramuscular; e COZ - perda por cozimento.

** Significativo a 1% ($P < 0,01$).

ns - Não-significativo a 5% ($P > 0,05$).

Conforme apresentado na Tabela 4, observa-se que os grupos características de carcaça e características de qualidade da carne não são considerados independentes. Os

três primeiros pares canônicos são estatisticamente significativos ($P < 0,01$). Isso mostra que as características de carcaça e qualidade da carne avaliadas são correlacionadas, visto que as correlações do primeiro, segundo e terceiro pares canônicos (r) foram, respectivamente, de 0,41, 0,34 e 0,33.

Primeiro par canônico

Em razão de existir mais de um par canônico estatisticamente significativo, a discussão foi feita para cada par canônico separadamente.

A proporção de variância comum para as duas variáveis, que é dada pelo coeficiente de correlação canônica ao quadrado (r^2), foi de 17,12%.

O exame das variáveis que compõem o primeiro par canônico apresenta, do lado das características de carcaça, predomínio para a variável peso do filezinho (PF), seguida, principalmente, pela variável comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça. Do lado das características de qualidade da carne, a variável luminosidade (L) teve predomínio absoluto.

No exame de correlação entre a variável original e sua correlação canônica, confirma-se que a variável **PF** é a mais importante na variável canônica U_1 , e a variável **L** é a mais importante na variável canônica V_1 , conforme pode ser observado na Tabela 5. É importante ressaltar que, em muitos casos, as interpretações com base nos coeficientes padronizados e nas correlações entre as variáveis originais e as canônicas podem ser muito diferentes.

Conforme pode ser observado na Tabela 5, maior peso do filezinho proporcionam menores índices de luminosidade, ou seja, maiores rendimentos de carne tendem a fornecer carne de pior qualidade.

Dos aspectos relacionados com a qualidade da carne, o que tem recebido maior atenção por parte dos pesquisadores são os índices de cor, pelos reflexos que tem sobre a aceitação do produto por parte dos consumidores. A inclusão da cor da carne nos programas de seleção tem sido recomendada como medida capaz de melhorar a qualidade do produto.

Tabela 5 – Correlação entre as características de carcaça e suas variáveis canônicas (U) e entre as características de qualidade da carne e suas variáveis canônicas (V)

	U ₁	U ₂	U ₃	U ₄	U ₅	U ₆
IDA	0,2166	-0,1437	0,4366	-0,1382	-0,2917	-0,0384
Pbdir	0,0964	0,0241	-0,0546	-0,4082	-0,4015	-0,3095
MBCC	0,3625	0,4979	0,1665	0,4878	-0,0806	-0,1022
SH	0,1857	-0,3757	0,2688	0,1669	0,2877	0,2318
PROFLOMB	-0,2361	-0,0417	-0,0847	0,1192	0,2569	-0,1908
COR	-0,0440	0,1473	0,2878	0,5587	-0,2080	0,0591
BACO	-0,4404	0,0165	0,2497	0,3245	-0,2342	-0,0163
PBDIRres	-0,1158	0,3911	-0,4988	0,0253	0,0258	-0,1848
PP	-0,2471	0,2011	0,1377	-0,3373	0,1811	0,3080
PCOPA	-0,0406	-0,0639	0,1046	-0,1071	-0,1689	-0,1553
PPA	0,0184	-0,1953	-0,3006	0,3473	-0,1322	0,3038
PC	-0,2400	0,3912	0,4353	-0,4349	0,1276	0,1270
PCOS	0,1625	0,2566	-0,2673	0,1785	-0,1340	-0,2016
PAPADA	-0,1081	-0,0174	0,2122	0,0388	0,4847	-0,6475
PF	-0,5946	0,0581	-0,0284	0,1829	-0,2727	-0,0219
RIM	0,2818	0,1427	-0,0444	0,1533	0,2879	0,3554
	V ₁	V ₂	V ₃	V ₄	V ₅	V ₆
pH ₄₅	0,4224	0,1396	0,1163	0,1107	0,8788	-0,0629
pH _u	0,2177	-0,0172	-0,4284	0,0443	0,1430	0,8639
L	0,7529	-0,1940	-0,1208	0,1232	-0,3275	-0,5084
a	-0,1922	0,2258	0,5569	0,6574	-0,1617	0,3790
GORINT	0,2004	0,8378	-0,0318	0,4293	-0,2589	0,0752
COZ	-0,2811	0,3461	-0,6552	0,5395	-0,1400	-0,2472

IDA - idade ao abate; Pbdir - peso da banda direita; MBCC - comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça; SH - maior espessura de toucinho na região da copa, na linha dorso-lombar; PROFLOMB - profundidade de lombo; COR - peso do coração; BACO - peso de baço; PBDIRres - peso da banda direita resfriada; PP - peso do pernil; PCOPA - peso da copa; PPA - peso da paleta; PC - peso do carré; PCOS - peso das costelas; PAPADA - peso da papada; PF - peso do filezinho; RIM - peso do rim; pH₄₅ - pH medido 45 minutos após o abate; pH_u - pH medido 24 horas após o abate; L - luminosidade; a - índice de vermelho; GORINT - gordura intramuscular; e COZ - perda por cozimento.

Com base na Tabela 5, tem-se que:

$$U_{y_1}^2 = \frac{(0,2166)^2 + (0,0964)^2 + (0,3625)^2 + \Lambda + (-0,5946)^2 + (0,2818)^2}{16} = 0,06742$$

e

$$V_{x_1}^2 = \frac{(0,4224)^2 + (0,2177)^2 + (0,7529)^2 + \Lambda + (0,2004)^2 + (-0,2811)^2}{6} = 0,1581$$

Assim, 6,74% da variância das características de carcaça é explicada pela variável canônica U_1 , e 15,81% da variância das características de qualidade da carne, pela variável canônica V_1 .

Segundo par canônico

A proporção de variância comum para as duas variáveis canônicas no segundo par canônico, que é dada pelo coeficiente de correlação canônica ao quadrado (r^2), foi de 12,12%.

O exame das variáveis que compõem o segundo par canônico apresenta, do lado das características de carcaça, maior predomínio da variável peso do carré (PC), seguida, principalmente, das variáveis comprimento da carcaça pelo Método Brasileiro de Classificação de Carcaça (MBCC); maior espessura de toucinho na região da copa, na linha dorso-lombar (SH); peso da paleta (PPA); e peso da banda direita resfriada (PBDIRres). Do lado das características de qualidade da carne, a variável gordura intramuscular (GORINT) teve predomínio absoluto.

No exame de correlação entre a variável original e sua correlação canônica, confirma-se que a variável **GORINT** é a mais importante na variável canônica V_2 ; entretanto, diferentemente dos resultados obtidos nos coeficientes padronizados para a variável canônica U_2 , a variável **MBCC** teve maior predomínio, seguida das variáveis **PC**, **PBDIRres** e **SH**, conforme pode ser observado na Tabela 5. Contudo, apesar de existirem algumas diferenças na ordem de importância das variáveis, os resultados são semelhantes.

A análise de correlação canônica fornece uma maneira mais simples de reduzir as complexidades envolvidas em relacionar dois conjuntos de variáveis. O principal problema associado com a técnica se refere à interpretação das soluções canônicas. Isso se deve ao fato de que a análise pode identificar várias correlações que, embora estatisticamente significativas, associam as variáveis utilizadas de forma bastante heterogênea e, conseqüentemente, difícil de ser interpretada (ABREU e VETTER, 1978).

Maior comprimento da carcaça proporciona uma maior quantidade de gordura intramuscular, conforme pode ser observado na Tabela 5. O conteúdo de gordura intramuscular é uma característica muito importante para a satisfação do consumidor da carne suína. Normalmente, as melhores pontuações em painéis sensoriais para suculência de lombos são obtidas naqueles lombos que apresentam maior quantidade de gordura intramuscular; valores extremamente baixos indicam características de qualidade mais pobres (BENEVENUTO JUNIOR, 2001).

Com base na Tabela 5, tem-se:

$$U_{y_2}^2 = \frac{(-0,1437)^2 + (0,0241)^2 + (0,4979)^2 + \Lambda + (0,0581)^2 + (0,1427)^2}{16} = 0,05703$$

e

$$V_{x_2}^2 = \frac{(0,1396)^2 + (-0,0172)^2 + (-0,1940)^2 + \Lambda + (0,8378)^2 + (0,3461)^2}{6} = 0,15501$$

Dessa forma, 5,70% da variância das características de carcaça é informada pela variável canônica U_2 , e 15,50% da variância das características de qualidade da carne é explicada pela variável canônica V_2 .

Terceiro par canônico

A proporção de variância comum para as duas variáveis canônicas no terceiro par canônico, que é dada pelo coeficiente de correlação canônica ao quadrado (r^2), foi de 11,19%.

O exame das variáveis que compõem o terceiro par canônico apresenta, do lado das características de carcaça, maior predomínio da variável peso da banda direita resfriada (PBDIRres). Em se tratando das características de qualidade da carne, a variável perda por cozimento (COZ) teve maior predomínio, seguida, principalmente, da variável índice de vermelho (a).

No exame de correlação entre a variável original e sua correlação canônica, a variável **PBDIRres** confirmou a maior importância na variável canônica U_3 , no terceiro par canônico; para a variável canônica V_3 , a variável **COZ** confirmou a maior importância, como pode ser observado na Tabela 5. Observa-se que, quanto maior for o peso da banda direita resfriada (maior rendimento de carcaça), maiores serão as perdas

por cozimento (pior a qualidade da carne), que é uma característica importante para processamento industrial.

IRGANG (1997) também evidencia a relação negativa entre rendimento em carne e qualidade desta, mostrando que raças suínas com maior rendimento em carne também apresentam maior exsudação de água. Uma tendência futura das indústrias é implementar o pagamento baseado não só nos índices de rendimento de carcaça, mas também nas características de qualidade da carne (SANCEVERO e ANTUNES, 1998).

Com base na Tabela 5, tem-se que:

$$U_{Y_3}^2 = \frac{(0,4366)^2 + (-0,0546)^2 + (0,1665)^2 + \Lambda + (-0,0284)^2 + (-0,0444)^2}{16} = 0,07023$$

e

$$V_{X_3}^2 = \frac{(0,1163)^2 + (-0,4284)^2 + (-0,1208)^2 + \Lambda + (-0,0318)^2 + (-0,6552)^2}{6} = 0,15868$$

Assim, 7,02% da variância das características de carcaça é informada pela variável canônica U_3 , e 15,87% da variância das características de qualidade da carne, pela variável canônica V_3 .

CONCLUSÕES

Os conjuntos de características de desempenho e carcaça são considerados dependentes. Em se tratando das características de desempenho, a variável mais importante é o peso aos 77 dias de idade, enquanto do lado das características de carcaça a variável mais importante é idade ao abate.

Os conjuntos de características de desempenho e qualidade da carne são independentes.

Os conjuntos de características de carcaça e qualidade da carne são dependentes. Do lado das características de carcaça, a variável mais importante é peso do filezinho, seguida, principalmente, pela variável comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça; no que diz respeito às características de qualidade da carne, a variável luminosidade é a mais importante.

Conhecidas a natureza e a magnitude da associação entre os grupos de características de desempenho e carcaça, entre características de desempenho e qualidade da carne e entre características de carcaça e qualidade da carne, será possível promover a seleção de forma mais eficiente, que possa melhor atender os objetivos do melhoramento genético de suínos.

LITERATURA CITADA

- ABREU, M.A. e VETTER, D. A. Análise de relação entre conjuntos de variáveis na matriz geográfica: Correlação Canônica. In: Faissol, S. Tendências atuais na geografia urbano/regional: teorização e quantificação. Rio de Janeiro, **IBGE**. p.133-144.
- BENEVENUTO JUNIOR, A.A. **Avaliação de rendimento de carcaça e de qualidade da carne de suínos comerciais, de raça nativa e cruzados**. Viçosa, MG: UFV, 2001, 93p. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos) – Universidade Federal de Viçosa, 2001.
- DUFOUR, E.; DEVAUX, M.F.; FORTIER, P.; HERBERT, S. Delineation of the structure of soft cheeses at the molecular level by fluorescence spectroscopy - relationship with texture. **International Dairy Journal**, v.11, p.465-473, 2001.
- IRGANG, R. A Força da Genética, Suinocultura Industrial, Jun/Jul 1997.
- JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. 4. ed., Englewood Cliffs, Prentice Hall, 1998. 816p.
- KHATTREE, R.; NAIK, D.N. **Multivariate data reduction and discrimination with SAS software**. SAS Institute Inc., Cary, NC, USA, 2000. 558p.
- PIRES, A. V. **Mapeamento de locos de características quantitativas, no cromossomo seis suíno**. Viçosa, MG: UFV, 2003, 86p. Dissertação (Doutorado em Genética e Melhoramento) – Universidade Federal de Viçosa, 2003.
- REGAZZI, A.J. **Análise multivariada**. Viçosa: Universidade Federal de Viçosa, 2002. (INF-766) (notas de aula).
- SANCEVERO, A. B., ANTUNES, R.C. Qualidade de carcaça e qualidade de carne em suínos. In: 20 REUNIÃO DA SOCIEDADE BRASILEIRA DE MELHORAMENTO ANIMAL. 1998, Uberaba. **Anais ...** Uberaba: Sociedade Brasileira de Melhoramento Animal, 1998.
- SAS, 1999. Software: versão 8.0. SAS Institute Inc., Cary, NC, USA.
- VAINIONPÄÄ, J.; KERVINEN, R.; De PRADO, M.; LAURILA, E.; KARI, M.; MUSTONEN, L.; AHVENAINEN, R. Exploration of storage and process tolerance of different potato cultivars using principal component and canonical correlation analyses. **Journal of Food Engineering**, v.44, p.47-61. 2000.

4 - RESUMO E CONCLUSÕES

Dados de 844 animais foram utilizados para avaliar a redução da dimensionalidade do espaço multivariado, por meio de componentes principais; por meio de correlação canônica foram avaliadas as associações entre as características de desempenho e de carcaça, entre as de desempenho e de qualidade da carne e entre as de carcaça e de qualidade da carne.

Na redução da dimensionalidade, observou-se que, para as características de desempenho, seis componentes principais, obtidos a partir da matriz de correlação, apresentaram variância inferior a 0,7 (autovalor inferior a 0,7). Assim, as seis variáveis que apresentaram maiores coeficientes, em valor absoluto, a partir do último componente principal, são passíveis de descarte. A razão para isso é que variáveis altamente correlacionadas com os componentes de menor variância representam variação praticamente insignificante.

As variáveis passíveis de descarte são, respectivamente, em ordem de menor importância para explicar a variação total: ganho de peso médio diário, peso aos 63 dias de idade, peso aos 42 dias de idade, tamanho da leitegada à desmama, peso aos 21 dias de idade e peso ao nascer. Com base nesses resultados, recomendam-se as seguintes variáveis para serem mantidas em experimentos futuros: tamanho da leitegada ao nascimento, número de tetos, peso aos 77 dias de idade, consumo de ração e conversão alimentar dos 77 aos 105 dias de idade.

Para as características de carcaça, observou-se que 17 componentes principais, obtidos a partir da matriz de correlação, apresentaram variância inferior a 0,7 (autovalor inferior a 0,7). Dessa forma, as 17 variáveis que apresentaram maiores coeficientes, em valor absoluto, a partir do último componente principal, são passíveis de descarte.

As variáveis sugeridas para descarte são, respectivamente, em ordem de menor importância para explicar a variação total: comprimento de carcaça pelo Método Americano de Classificação de Carcaça; peso do pernil sem pele e sem gordura; peso da copa sem pele e sem gordura; espessura de toucinho medida imediatamente após a última costela, a 6,5 cm da linha dorso-lombar; espessura do bacon; área de olho de lombo; espessura de toucinho entre a última e a penúltima vértebra lombar, na linha dorso-lombar; peso do lombo; peso da banha rama; peso da paleta sem pele e sem gordura; espessura de toucinho imediatamente após a última costela, na linha dorso-lombar; peso do fígado; comprimento total do intestino delgado; peso da cabeça; peso

do bacon; rendimento de carcaça com pés e cabeça; e peso do pulmão. Com base nesses resultados, recomenda-se a manutenção das seguintes variáveis em experimentos futuros: idade ao abate; peso da banda direita; comprimento da carcaça pelo Método Brasileiro de Classificação de Carcaça; maior espessura de toucinho na região da copa, na linha dorso-lombar; profundidade de lombo; e pesos do baço, do coração, da banda direita resfriada, do pernil, da copa, da paleta, do carré, das costelas, da papada, do filezinho e do rim.

No que se refere às características de qualidade da carne, observou-se que quatro componentes principais, obtidos a partir da matriz de correlação, apresentaram variância inferior a 0,7 (autovalor inferior a 0,7). Assim, as quatro variáveis que apresentaram maiores coeficientes, em valor absoluto, a partir do último componente principal, são passíveis de descarte.

As variáveis passíveis de descartes são, respectivamente, em ordem de menor importância para explicar a variação total: perda de peso total, índice de amarelo, perda por gotejamento e maciez objetiva. Com base nesses resultados, recomendam-se as seguintes variáveis para serem mantidas em experimentos futuros: pH medido aos 45 minutos e 24 horas *post-mortem*, luminosidade, gordura intramuscular, perda por cozimento e índice de vermelho.

Em razão do grande número de variáveis redundantes que podem ser descartadas, espera-se relativa economia de tempo e custo em experimentos futuros, sem causar perda considerável de informação.

No estudo da associação entre os grupos de variáveis, observou-se que os grupos características de desempenho e características de carcaça não são considerados independentes, pois foram fortemente correlacionados, visto que a correlação do primeiro par canônico foi de 0,78. Os dois primeiros pares canônicos foram estatisticamente relevantes ($P < 0,01$ e $P < 0,05$, respectivamente).

No exame das variáveis que compõem o primeiro par canônico, para as características de desempenho, houve predomínio absoluto da variável peso aos 77 dias de idade; para as características de carcaça, a variável idade ao abate teve predomínio absoluto. No segundo par canônico, para as características de desempenho, houve predomínio absoluto da variável número de tetos; em se tratando das características de carcaça, a variável peso das costelas teve a maior contribuição, seguida da variável peso da banda direita.

No estudo da associação entre os grupos características de desempenho e características de qualidade da carne, observou-se que existe independência, pois a correlação do primeiro par canônico foi de 0,23, sendo estatisticamente irrelevante ($P > 0,05$).

No que diz respeito à associação entre os grupos características de carcaça e características de qualidade da carne, constatou-se que não existe independência; os três primeiros pares canônicos foram estatisticamente relevantes ($P < 0,01$).

No exame das variáveis que compõem o primeiro par canônico, para as características de carcaça, houve predomínio da variável peso do filezinho, seguida principalmente da variável comprimento de carcaça pelo Método Brasileiro de Classificação de Carcaça; em se tratando das características de qualidade da carne, a variável luminosidade teve predomínio absoluto. No que se refere às variáveis que compõem o segundo par canônico, para as características de carcaça, houve maior predomínio da variável peso do carré, seguida principalmente das variáveis comprimento da carcaça pelo Método Brasileiro de Classificação de Carcaça, maior espessura de toucinho na região da copa, na linha dorso-lombar, peso da paleta e peso da banda direita resfriada; para as características de qualidade da carne, houve predomínio absoluto da variável gordura intramuscular. No exame das variáveis que compõem o terceiro par canônico, para as características de carcaça, houve maior predomínio da variável peso da banda direita resfriada; em se tratando das características de qualidade da carne, a variável perda por cozimento teve maior predomínio, seguida principalmente da variável índice de vermelho.