

NORMALIZAÇÃO E PADRONIZAÇÃO POR MEIO DA TRANSFORMAÇÃO LOGARÍTMICA EM ESTUDOS ARQUEOMÉTRICOS DE CERÂMICAS

J.O. Santos¹, C.S. Munita², C. Vergne³, P.M.S. Oliveira⁴

1 – INTRODUÇÃO

Atualmente, a análise material dos artefatos arqueológicos tem sido fundamental para construir uma representação de uma cultura antiga, uma vez que estas análises permitem realizar a sua caracterização, conhecer as tecnologias de manufatura e usos (Vandiver, 2001).

Para estudos de proveniência, por meio da composição química elementar das cerâmicas, têm sido utilizadas técnicas tais como PIXE (Roumié et al., 2006), PIGE (Felicíssimo et al., 2005), etc. Entretanto, a AANI tem sido a principal técnica para estudo de proveniência de cerâmicas arqueológicas, uma vez que é capaz de produzir análise multielementar com limites de detecção ao nível de elementos traços, com exatidão e precisão adequadas a este tipo de estudo (Glascock, 2000).

A formação de grupos estatisticamente significantes das amostras analisadas e a avaliação de amostras de origens desconhecidas, para alocação em um dos grupos estabelecidos, são objetivos básicos dos estudos composicionais da cerâmica. Para estes fins têm sido aplicadas técnicas multivariadas de agrupamento, análise fatorial, análise discriminante e análise por componentes principais (Mommensen et al., 1988).

Com relação à normalidade, especificamente, faz-se necessário a sua verificação, devida, principalmente, ao uso de um grande número de variáveis, o qual torna as distorções potenciais e vieses mais significativos quando as suposições são violadas. Entretanto, em quase todos os métodos multivariados, as estimativas e predições realizadas de acordo com um determinado modelo multivariado podem ser obtidas, mesmo se a normalidade multivariada for violada.

Tendo em vista estas proposições, neste trabalho estudou-se a normalidade univariada dos elementos analisados e da distribuição conjunta destes elementos.

¹ Centro Federal de Educação Tecnológica de Sergipe, CEFET/SE-UNED LAGARTO. CEP 49.055-260, Sergipe, SE, Brasil. E-mail: josantos@ipen.br

² Pesquisador, Instituto de Pesquisas Energéticas e Nucleares – IPEN-CNEN/SP, C. P. 11049 CEP.: 05422-970, São Paulo/SP, e-mail: camunita@net.ipen.br Tel. 0 XX (11) 3816.9289.

³ Museu Arqueológico de Xingó, MAX/UFS, Av Marechal Rondon s/n, São Cristóvão, SE, Brasil

⁴ Aluno do programa de doutorado em estatística. Depto de Estatística, Instituto de Matemática e Estatística – USP, C. P.: 66281, CEP. 05315-970, São Paulo/SP, e-mail: poliveir@ime.usp.br, Tel. 0 XX (11) 3106.7760.

Inicialmente, a normalidade univariada dos elementos foi verificada por meio da construção dos histogramas dos elementos para cada sítio estudado e, em seguida, realizou-se um teste formal de normalidade por meio do teste de Shapiro-Wilks (Royston, 1982). Para verificação da normalidade multivariada foi realizado um teste formal de normalidade multivariada, o qual corresponde a uma generalização do teste de Shapiro-Wilks para o caso multivariado, e que está implementado no programa estatístico R (Jarek, 2005).

Para este fim, foi realizada uma análise de uma base de dados composta por 46 amostras cerâmicas arqueológicas e 11 variáveis, determinadas por ativação neutrônica. As amostras foram coletadas em escavações arqueológicas realizadas no sítio Justino, localizado no estado de Sergipe.

2 - Material e Métodos

2.1 – Material analisado

O sítio arqueológico Justino, localizado no Noroeste do estado de Sergipe, Brasil, fora analisado e interpretado, com objetivo de estudar a influência da normalização dos dados.

2.2 - Normalização e padronização dos dados

A normalidade multivariada dos dados é requerida, em muitos casos, visto que muito dos métodos estatísticos multivariados, incluindo análise de variância, análise fatorial com extração por função de verossimilhança, correlação canônica entre outros métodos, assumem que os dados seguem este tipo de distribuição (Baxter, 1999).

Segundo Mecklin & Mundfrom (2004), a identificação da normalidade multivariada pode ser obtida por meio de gráficos e coeficientes de correlação. Dentro deste ponto de vista o gráfico dos percentis (Q-Q plot) tem sido utilizado com bastante frequência (Hair et al., 2005). Apesar da praticidade do método Q-Q plot, o mesmo não constitui um teste formal para verificar desvio de normalidade multivariada. Para contornar esta limitação, diversos autores têm proposto métodos formais para verificação de normalidade por meio de estatísticas que utilizam o vetor de média, coeficiente de assimetria e curtose (Mardia, 1974). Estas estatísticas são calculadas e comparadas com seus respectivos valores críticos para verificação de desvio da distribuição normal multivariada (Mardia & Kent, 1991). Outro teste para normalidade multivariada dos dados, o que é uma extensão do teste de Shapiro-Wilks para o caso

univariado, implementado no programa estatístico R, tem sido utilizado com bastante frequência (Royston, 1995).

No estudo da composição química da cerâmica tem sido observado que os elementos analisados se distribuem log – normalmente (Glascok, 1992). Isto é preferível por duas razões básicas: primeiro, porque tem sido observado que diversas concentrações elementares são normalizadas quando tomado os seus valores logarítmicos. A segunda razão é que a transformação dos dados de concentração em uma escala logarítmica compensa as diferenças de magnitudes das variáveis medidas (padronização).

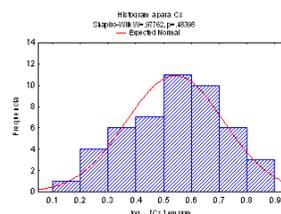
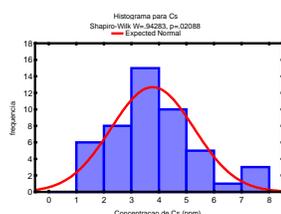
Visando a padronização e normalização dos dados, neste trabalho a interpretação dos dados, por meio da análise discriminante, foi realizada após transformação. A verificação da normalidade univariada dos dados foi verificada por meio do teste de Shapiro-Wilks (Hair et al., 2005) e para teste de normalidade multivariada foi aplicado uma extensão desse teste, o qual está implementado no programa R.

A medida da qualidade da distinção entre os grupos composicionais referentes ao Cemitério B e Cemitério C, do sítio Justino, após cada tipo de transformação, foi efetuada por meio da distância Mahalanobis, a qual mede a distância entre os centróides dos grupos, ponderada pela dispersão dos grupos.

3 – Resultados e discussão

Os testes de normalidade foram realizados para os dados brutos e para os dados transformados de acordo com a transformação logarítmica.

A observação dos histogramas para os dados brutos não apresentou assimetrias visivelmente acentuadas para nenhum dos elementos determinados com CV menor 10% (Figura 1). Entretanto, após testar a normalidade univariada por meio do teste de Shapiro-Wilks, verificou-se que os elementos Cr, Cs, Sc, Eu e Hf não apresentaram uma distribuição significativamente normal ($p < 0,05$). Por outra parte, por meio da transformação logarítmica dos dados, observou-se que as concentrações elementares das amostras, para cada sítio estudado, ajustaram-se satisfatoriamente à distribuição normal (Figura 1) e os testes de normalidade indicaram que a hipótese de normalidade univariada dos dados após transformação logarítmica é significante ($p > 0,1$).



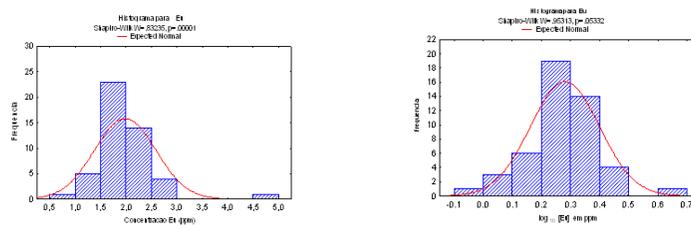


Figura 1 – Histogramas para as distribuições de Cs e Eu nas amostras do sítio Porto Belo, para os dados brutos (esquerda) e para \log_{10} da concentração do elemento (direita).

Assim, apesar de ter sido verificado um desvio da normalidade multivariada dos dados, a interpretação dos grupos composicionais será obtida após transformação logarítmica das concentrações, visto que esta transformação promove a padronização das concentrações dos diversos elementos e a normalização individual das variáveis utilizadas para formação dos grupos. Como tem sido apontado por alguns autores, o desvio de normalidade não é crítica em estudos composicionais, visto que são utilizadas técnicas de agrupamento robustas a estes desvios, tais como Análise das Componentes Principais, Análise Discriminante, Análise de Conglomerados, entre outras (Hair et al., 2005; Beier & Mommsen, 1994; Barroso & Artes, 2003).

Após estudo da normalidade, realizou-se uma análise de cluster, seguida por uma análise discriminante, tanto com a base original quanto com a base transformada por meio do logaritmo. Obteve-se que as amostras dos cemitérios B e C do Justino foram grupo composicionais distintos, conforme era esperado do ponto de vista da Arqueologia. Além disso, observou-se que esta análise não diferiu significativamente, quando os dados são originais ou transformados.

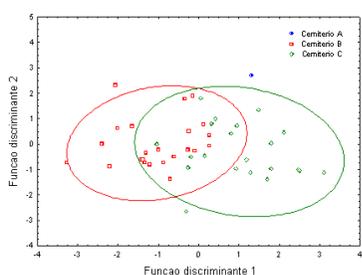


Figura 2 – Análise discriminante com base original.

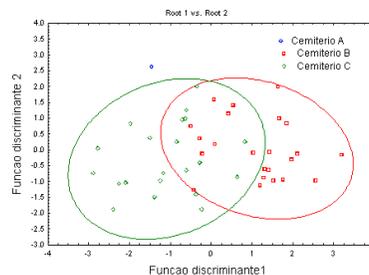


Figura 3 – Análise discriminante com transformação \log_{10} da concentração.

4 - Conclusões

Conforme os resultados da análise por ativação com nêutrons, verificou-se que após transformação logarítmica os dados apresentaram uma distribuição normal, isto é, a distribuição univariada dos dados é lognormal. Entretanto, observou-se que a distribuição conjunta dos resultados não é lognormal, conforme a extensão do teste de

Shapiro-Wilks, o que não implica em grandes prejuízos na análise composicional de cerâmicas.

5 – REFERÊNCIAS BIBLIOGRÁFICAS

- BAXTER, M.J. *Stepwise Discriminant Analysis in Archaeometry: a Critique*. Journal of Archaeological Science, 21, 659 – 666, 1994.
- FELICISSIMO, M.P., PEIXOTO, J.L., BARBOSA, M.S., DERMOTIER, G., , PIREAUX, J.J., RODRIGUES, U.P. *PIXE and PIGE analysis of Pre-colonial pottery from pantanal – MS, Brazil*. Journal of Radioanalytical and Nuclear Chemistry, 209-217 (1), 2005.
- GLASCOCK, M.D. *Characterization of ceramics at MURR by NAA and Multivariate Statistics*. In: Chemical Characterization of Ceramic Paste in Archaeology, 2000.
- HAIR JUNIOR, J.F., ANDERSON, R.E., TATHAM, R.L., BLACK, C. *Análise multivariada de dados*. Bookman, Porto Alegre, 2005.
- JAREK, S. *Mvnormtest: A Package for normality test multivariate variables*. 2001. (disponível para download em: <http://cran.r-project.org/src/contrib/Descriptions/mvnormtest.html>).
- MARDIA, K. V. *Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies*. Sankhya B, 36, 115–128, 1974.
- MECKLIN, C. J., MUNDFROM, D. J. *An appraisal and bibliography of tests for multivariate normality*. International Statistical Review, 72 (1), 123 – 138, 2004.
- MOMMSEN, H., KREUSER, A., WEBER, J. *A method for grouping pottery by chemical composition*. Archaeometry, 30, 47-57, 1988.
- ROUMIÉ, M., REYNOLDS, P., ATALLAH, C., BAKRAJI, E., ZAHRAMAN, K., NSOULI, B. *Provenance study of excavated pottery from Beirut using PIXE cluster analysis*. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 249 (1-2), 612 – 615, 2006.
- VANDIVER, P. *The role of materials research in ceramics and archaeology*. Annu. Rev. Mater. Res. 31 373 – 385, 2001.

AGRADECIMENTOS

Os autores deste trabalho gostariam de agradecer ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, 155373/2006-4) pelo apoio financeiro. Agradecimentos especial ao Museu Arqueológico de Xingó (MAX) e Petróleo Brasileiro S.A (PETROBRAS) pela pelo apoio técnico, logístico e por permitir a análise da coleção.

This document was created with Win2PDF available at <http://www.daneprairie.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.