

CONTRIBUTOS À COMPREENSÃO DA DISTRIBUIÇÃO DE VERBAS DESTINADAS À EDUCAÇÃO NOS MUNICÍPIOS PORTUGUESES EM 1870-1880

Gilda Soromenho

Instituto da Educação da Universidade de Lisboa
gspereira@ie.ul.pt

Resumo

No âmbito do projeto Atlas-Repertório dos Municípios na Educação e na Cultura em Portugal (1820-1986) foram observados diversos indicadores referentes aos anos 1870-1880, em 287 municípios dos 21 distritos do Continente e Ilhas. O objetivo era determinar de que fatores dependia a distribuição de verbas para a Educação a um município nessa época. Nesta perspetiva realizou-se numa primeira análise um estudo exploratório sobre as correlações dos vários indicadores e posteriormente uma análise de regressão múltipla cuja variável dependente é o orçamento geral atribuído ao município. Constatou-se que a distribuição das verbas por município dependia essencialmente de duas variáveis latentes, uma que é o compósito de um conjunto de indicadores sobre a dimensão da população, número de escolas, número de habitantes que só sabem ler ou que sabem ler e escrever e outra que está relacionada com a taxa de alfabetização e com a taxa real de escolarização. Verificou-se também que nesta época já existia uma certa politização na Educação, uma vez que distribuição de verbas dependia da localização do município.

Palavras-Chave: Análise Fatorial em Componentes Principais; Análise de Regressão Múltipla; ANOVA.

Abstract

Under the project Atlas-Index of Municipalities regarding Education and Culture in Portugal (1820-1986) were observed several indicators for the years 1870-1880 in 287 municipalities of 21 districts of mainland and islands Madeira and Azores. The objective was to determine the main factors from which depended the budget assigned to Education on a municipality in the years 1870-1880. An exploratory correlation study



of the various indicators was held, followed by a multiple regression analysis with dependent variable the overall budget allocated to the municipality. It was found that the budgets assigned to the municipality depended essentially of two latent variables, which are the composite of a set of the previous indicators, one related with the size of population county, number of schools, number of people who only know how to read, or read and write and the other related to the literacy rate and the actual rate of schooling. It was also found that there was already certain politicization in education budget, since the distribution of funds depended on the allocation of the municipality.

Keywords: Principal Component Analysis; Multiple Regression; ANOVA.

Introdução

No âmbito do projeto Atlas-Repertório dos Municípios na Educação e na Cultura em Portugal (1820-1986) foram observados os seguintes indicadores em 287 municípios dos 21 distritos do Continente e Ilhas: dimensão da população (P); número de habitantes entre os seis e os dez anos, e entre os onze e os quinze anos; número de habitantes que só sabem ler (L); número de habitantes que sabem ler e escrever (LE); número total de escolas (E) e orçamento geral destinado à Educação.

A partir dos indicadores registados foram ainda criados novos indicadores nomeadamente: dimensão da população em idade escolar que resulta da soma do número de habitantes entre seis e dez anos com os que têm entre onze e quinze anos (IE); taxa de alfabetização (TA) que resulta do quociente entre o número de habitantes que só sabe ler com os que sabem ler e escrever pela dimensão da população ($TA = \frac{L+LE}{P} \times 100\%$); taxa real de escolarização que resulta do quociente entre a dimensão da população em idade escolar pela dimensão da população ($TRE = \frac{IE}{P} \times 100\%$). Para a realização das análises estatísticas efetuadas foi utilizado o software SPSS 20.0.

Inicialmente procedeu-se ao estudo das correlações entre os indicadores anteriormente referidos através de uma Análise em Componentes Principais, com rotação Varimax e posteriormente à estimação de um modelo de Regressão Linear Múltipla que explicasse a verba destinada à Educação em cada município. Verificou-se que essa verba depende de duas variáveis latentes, originadas pela Análise em Componentes Principais, e da localização do município.



Metodologia

Numa primeira abordagem procedeu-se a uma análise descritiva dos indicadores considerados nos 287 municípios que constituíram a unidade estatística. Devido à existência de grandes disparidades entre estes (Figura 1), padronizou-se cada um dos indicadores, isto é, subtraiu-se o seu valor pela média e dividiu-se o resultado pelo desvio-padrão. Retiraram-se do estudo os municípios que, em algum dos indicadores, os valores figuravam fora do intervalo (-3,3) por serem considerados outliers, o que interferia fortemente nos pressupostos de normalidade das análises posteriores. Dos 287 municípios iniciais restaram 269, que já apresentavam valores de assimetria e curtose dentro dos limites considerados usuais (Figura 2), Maroco (2010).

Figura 1 – Análise descritiva dos indicadores P, IE, LE, L, E, TA, TRE e orçamento geral nos 287 municípios

| Descriptive Statistics | | | | | | | | | |
|---|-----------|-----------|-----------|-----------|----------------|-----------|------------|-----------|------------|
| | N | Minimum | Maximum | Mean | Std. Deviation | Skewness | | Kurtosis | |
| | Statistic | Statistic | Statistic | Statistic | Statistic | Statistic | Std. Error | Statistic | Std. Error |
| dimensão da população (P) | 287 | 880 | 246311 | 15830,10 | 17683,962 | 8,449 | ,144 | 102,843 | ,287 |
| dimensão da população em idade escolar (IE) | 287 | 214 | 34839 | 3167,05 | 2969,656 | 5,373 | ,144 | 48,127 | ,287 |
| número de habitantes que sabe ler e escrever (LE) | 287 | 50 | 96044 | 2272,06 | 6134,664 | 13,202 | ,144 | 195,342 | ,287 |
| número de habitantes que só sabe ler (L) | 287 | 0 | 14323 | 511,17 | 996,303 | 10,243 | ,144 | 133,258 | ,287 |
| número total de escolas (E) | 287 | 0 | 161 | 12,09 | 15,253 | 5,740 | ,144 | 45,055 | ,287 |
| orçamento geral | 285 | 16,000 | 11705,000 | 383,40089 | 738,806660 | 12,994 | ,144 | 195,978 | ,288 |
| taxa de alfabetização (TA) | 287 | 3,66 | 44,81 | 14,7887 | 5,54354 | 1,155 | ,144 | 3,526 | ,287 |
| taxa real de escolarização (TRE) | 287 | 14,14 | 52,84 | 20,3521 | 2,61993 | 6,828 | ,144 | 82,342 | ,287 |
| Valid N (listwise) | 285 | | | | | | | | |

Figura 2 – Análise descritiva dos indicadores P, IE, LE, L, E, TA, TRE e orçamento geral sem os outliers

| Descriptive Statistics | | | | | | | | | |
|---|-----------|-----------|-----------|-----------|----------------|-----------|------------|-----------|------------|
| | N | Minimum | Maximum | Mean | Std. Deviation | Skewness | | Kurtosis | |
| | Statistic | Statistic | Statistic | Statistic | Statistic | Statistic | Std. Error | Statistic | Std. Error |
| dimensão da população (P) | 269 | 880 | 52779 | 13987,65 | 9245,770 | 1,526 | ,149 | 2,755 | ,296 |
| dimensão da população em idade escolar (IE) | 269 | 214 | 10995 | 2819,81 | 1856,711 | 1,498 | ,149 | 2,647 | ,296 |
| número de habitantes que sabe ler e escrever (LE) | 269 | 50 | 10142 | 1711,35 | 1479,175 | 2,100 | ,149 | 2,256 | ,296 |
| número de habitantes que só sabe ler (L) | 269 | 0 | 2351 | 392,43 | 359,176 | 1,904 | ,149 | 2,974 | ,296 |
| número total de escolas (E) | 269 | 1 | 48 | 10,25 | 7,956 | 1,977 | ,149 | 2,642 | ,296 |
| taxa de alfabetização (TA) | 269 | 3,66 | 28,88 | 14,4393 | 4,96199 | ,538 | ,149 | -,015 | ,296 |
| taxa real de escolarização (TRE) | 269 | 14,33 | 27,39 | 20,2435 | 1,75637 | ,727 | ,149 | 1,582 | ,296 |
| orçamento geral | 269 | 16,000 | 1135,000 | 302,78188 | 201,829792 | 1,519 | ,149 | 2,487 | ,296 |
| Valid N (listwise) | 269 | | | | | | | | |



O objetivo para determinar de que fatores dependia a atribuição de verbas a um município nos anos 1870-1880 foi atingido estimando um modelo de regressão linear múltipla, em que a variável dependente era o orçamento geral e que foi explicada em função dos restantes indicadores (variáveis explicativas). Contudo, para estimar este modelo, era necessário que as variáveis explicativas apresentassem correlações fracas entre si, o que intuitivamente se percebe não ser o caso, porque existia uma correlação elevada entre, por exemplo, a dimensão da população e a dimensão da população em idade escolar. Assim, foi realizada uma Análise Fatorial em Componentes Principais, com rotação Varimax, para identificar os grupos de indicadores mais relacionados entre si, Pallant (2010). Obteve-se um KMO (Kaiser Meyer Olkin Measure of Sampling Adequacy) de 0,707 e dois fatores (variáveis latentes) que atingiram uma percentagem de explicação de 81,36%. O primeiro fator explicou cerca de 58,00% da variação total e o segundo 23,36%. Através da Figura 3 pode observar-se que o primeiro fator, variável latente designada por características populacionais do município, é formado pelos indicadores: dimensão da população; dimensão da população em idade escolar; número de habitantes que sabem ler e escrever; número de habitantes que só sabem ler; número total de escolas, porque são estes os indicadores mais relacionados com o primeiro fator. O segundo fator, variável latente designada por escolaridade no município, é composto pelos indicadores taxa de alfabetização e taxa real de escolarização, por serem estes os indicadores mais relacionados com o segundo fator.

Analisou-se ainda se estas duas variáveis latentes eram discriminantes dos 21 distritos, tomando como amostra os municípios de cada distrito. Pode observar-se pela Figura 4 que de facto existem diferenças muito significativas entre os distritos, $\text{sig}=0.000 < \alpha=0.05$, pelo que se rejeita a hipótese de igualdade dos valores médios das variáveis latentes por distrito. Tal indica que as variáveis latentes encontradas podem ser utilizadas para explicar a atribuição de verbas. Por outro lado, como provêm de uma Análise Fatorial em Componentes Principais, têm correlação nula entre si, o que é ideal para a sua utilização como variáveis explicativas num modelo de regressão linear múltipla.

Para estimar o modelo de regressão que irá analisar os indicadores dos quais depende o orçamento geral atribuído a um município foram utilizadas:

- As duas variáveis latentes (caraterísticas populacionais do município e escolaridade no município)



- Uma variável dummy designada por Litoral, que assume o valor 1 se o município pertencer ao litoral e zero se pertencer ao interior
- Uma variável dummy designada por Norte e que assume o valor 1 se o município for do Norte e zero se pertencer ao Centro ou ao Sul
- Uma variável dummy designada por Centro, que assume o valor 1 se o município for do Centro e zero se pertencer ao Norte ou ao Sul
- Uma variável dummy designada por Ilhas e que assume o valor 1 se o município for das ilhas (Madeira ou Açores) e zero se pertencer ao Continente.

Figura 3 – Resultados da Análise em Componentes Principais com rotação Varimax

| Total Variance Explained | | | | | | | | | |
|--------------------------|---------------------|---------------|--------------|-------------------------------------|---------------|--------------|-----------------------------------|---------------|--------------|
| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 4,313 | 61,621 | 61,621 | 4,313 | 61,621 | 61,621 | 4,060 | 57,997 | 57,997 |
| 2 | 1,382 | 19,738 | 81,359 | 1,382 | 19,738 | 81,359 | 1,635 | 23,362 | 81,359 |
| 3 | ,599 | 8,551 | 89,910 | | | | | | |
| 4 | ,418 | 5,978 | 95,888 | | | | | | |
| 5 | ,236 | 3,367 | 99,255 | | | | | | |
| 6 | ,050 | ,721 | 99,976 | | | | | | |
| 7 | ,002 | ,024 | 100,000 | | | | | | |

Extraction Method: Principal Component Analysis.

| Rotated Component Matrix ^a | | |
|---|-----------|-------|
| | Component | |
| | 1 | 2 |
| dimensão da população (P) | ,970 | ,000 |
| dimensão da população em idade escolar (IE) | ,971 | -,102 |
| número de habitantes que sabe ler e escrever (LE) | ,893 | ,331 |
| número de habitantes que só sabe ler (L) | ,863 | ,221 |
| número total de escolas (E) | ,745 | ,311 |
| taxa de alfabetização (TA) | ,279 | ,815 |
| taxa real de escolarização (TRE) | ,026 | -,840 |

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.



Figura 4 – Testes de igualdade dos valores médios de cada variável latente por distrito

| Tests of Equality of Group Means | | | | | |
|---|---------------|--------|-----|-----|------|
| | Wilks' Lambda | F | df1 | df2 | Sig. |
| caraterísticas populacionais do município | ,796 | 3,288 | 20 | 248 | ,000 |
| escolaridade no município | ,524 | 11,648 | 20 | 248 | ,000 |

O modelo de regressão linear múltipla explica 76,9% da variabilidade do orçamento geral atribuído a um município e como se pode observar pela Figura 5, os fatores que o influenciam mais significativamente são: as caraterísticas populacionais do município, a escolaridade no município e a localização, uma vez que $\text{sig}=0.000 < \alpha=0.005$.

Quanto maior for a dimensão das caraterísticas populacionais, nomeadamente dimensão da população em idade escolar e dimensão da população, também maior é o orçamento designado ao município, assim como quanto maior for a taxa de alfabetização também maior será a verba atribuída, porque os coeficientes destas variáveis no modelo são positivos, o que indica uma relação direta entre estes e a verba. O fato de o município estar localizado no litoral ou no interior não influencia o orçamento concedido, no entanto se este pertencer à região Norte, Centro, Sul ou Ilhas isso já introduz modificações na atribuição de verba.

Foi também observada a condição de normalidade dos resíduos pelo teste de Kolmogorov-Smirnov ($\text{sig}=0,064$) de modo a validar as conclusões anteriores, Martins (2011).

Para indagar como se processava a distribuição de verbas por localização do município foi também realizada uma Análise de Variância (ANOVA) a um fator, em que a variável dependente foi o orçamento geral e a variável independente a localização. As localizações previstas foram designadas por Norte, Centro, Sul e Ilhas. A diferença mais significativa encontra-se entre o Continente e as Ilhas (Figura 6) ($\text{sig} \leq .002$) e pode observar-se ainda uma ligeira diferença entre os orçamentos atribuídos aos municípios do Norte e do Centro, sendo que os maiores valores são atribuídos aos municípios do Norte, seguidos pelos do Sul.



Figura 5 – Resultados do modelo de regressão linear múltipla ajustado ao indicador orçamento geral

| Model Summary ^b | | | | | |
|----------------------------|-------------------|----------|-------------------|----------------------------|--|
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | |
| 1 | ,877 ^a | ,769 | ,764 | 131,105294 | |

a. Predictors: (Constant), Ilhas, características populacionais do município, Centro, litoral, escolaridade no município, Norte

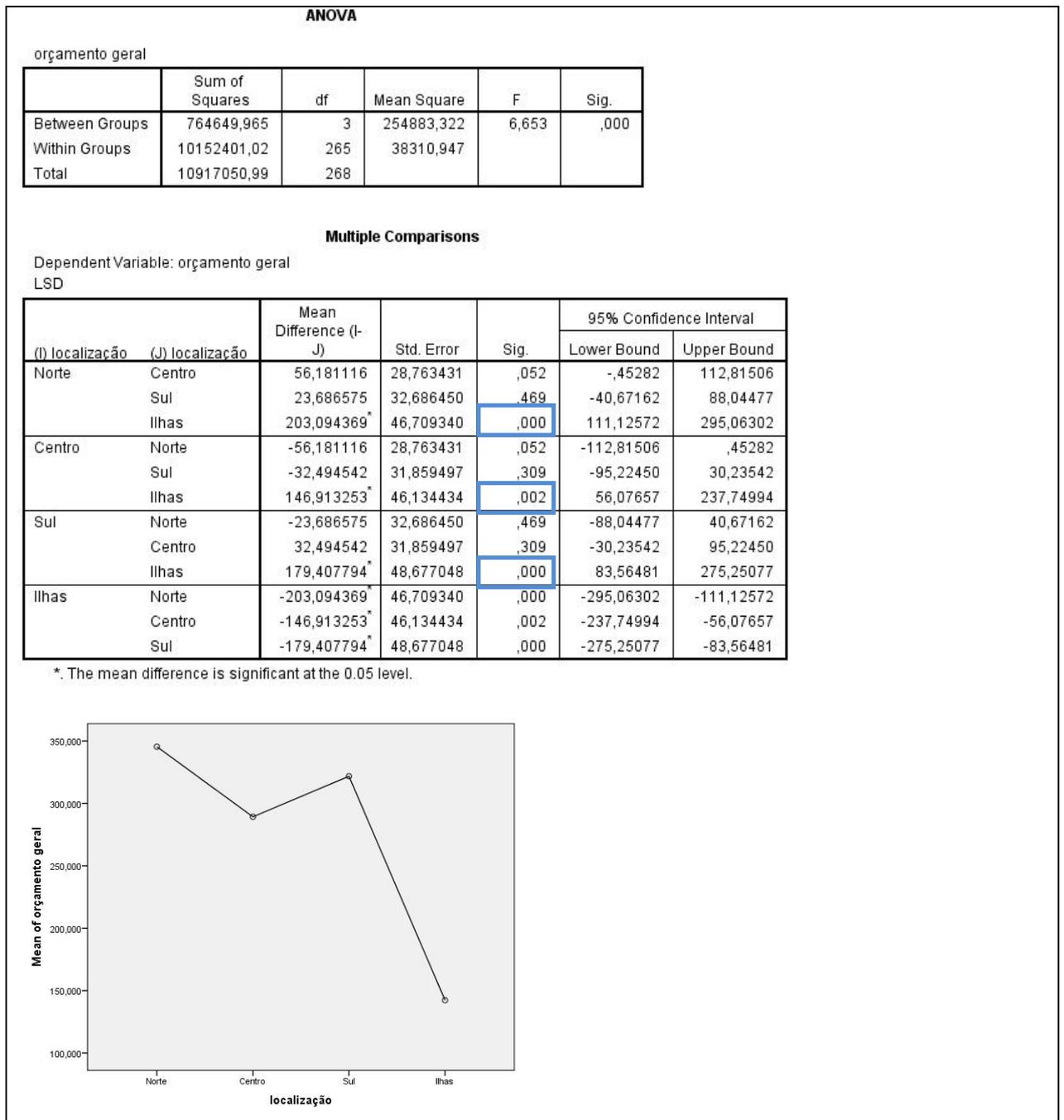
b. Dependent Variable: orçamento geral

| Coefficients ^a | | | | | | |
|---------------------------|---|-----------------------------|------------|---------------------------|--------|------|
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 407,579 | 20,165 | | 20,212 | ,000 |
| | caraterísticas populacionais do município | 163,650 | 9,119 | ,788 | 17,946 | ,000 |
| | escolaridade no município | 40,927 | 9,214 | ,200 | 4,442 | ,000 |
| | Litoral | -40,457 | 21,468 | -,099 | -1,884 | ,061 |
| | Norte | -129,880 | 23,925 | -,302 | -5,429 | ,000 |
| | Centro | -61,685 | 21,649 | -,148 | -2,849 | ,005 |
| | Ilhas | -120,400 | 34,615 | -,164 | -3,478 | ,001 |

a. Dependent Variable: orçamento geral



Figura 6 – Resultados da Anova a um fator





Conclusões

Deste estudo ressalta que a caracterização dos municípios na época 1870 a 1880 pode ser organizada recorrendo a duas características fundamentais, nomeadamente as características populacionais do município, que incluem a dimensão da população, a dimensão da população em idade escolar, o número de habitantes que só sabem ler, o número de habitantes que sabem ler e escrever, o número total de escolas, a escolaridade no município que inclui os indicadores taxa de alfabetização e taxa real de escolarização negativamente relacionadas, isto é quanto maior for a taxa de alfabetização, menor é a taxa real de escolarização.

A verba destinada à educação, em cada município, depende não só das características populacionais do município e da sua escolaridade, como também da sua localização. Os municípios da região Norte são os que têm maior atribuição de verba, seguidos dos do Sul e do Centro, sendo os municípios das Ilhas o local onde as verbas são mais reduzidas. Nesta época não se verificava ainda uma diferenciação na atribuição de verbas para a Educação entre os municípios do litoral e do interior.

Referências Bibliográficas

- Maroco, J. (2010). *Análise estatística, com utilização do SPSS* (2ª Ed.). Edições Sílabo.
- Martins, C. (2011). *Manual de Análise de Dados Quantitativos com Recurso ao IBM SPSS*. Psiquilíbrios Edições.
- Pallant, J. (2010). *SPSS Survival Manual* (5th Ed.). MacGraw Hill International.