

TIPOLOXÍA DOS FOGARES GALEGOS A TRAVÉS DA APLICACIÓN DE TÉCNICAS FACTORIAIS DE ANÁLISE DE TÁBOAS MÚLTIPLES

López Vizcaíno, M^a Esther ¹, Sánchez Fernández, Patricio ².

¹ Instituto Galego de Estatística

² Universidade de Vigo

RESUMO

O obxectivo deste estudio será establecer unha tipoloxía dos fogares galegos segundo características económicas e xeográficas relacionadas co nivel de ingresos, así como a presenza ou ausencia de certas comodidades e mailo seu emprazamento xeográfico. Con esta finalidade empregaranse dúas técnicas factoriais: a análise de correspondencias múltiple (*ACM*) e a análise factorial múltiple (*AFM*)

Palabras chave: análise de correspondencias múltiple, análise factorial múltiple, tipoloxía

1. INTRODUCCIÓN

A vinculación entre a estatística e as ciencias sociais e, en especial, a economía é un feito incuestionable e cada vez máis presente no desenvolvemento social dun territorio. Este traballo pode servir como boa mostra deste feito, ó combinar un estudio de carácter marcadamente socioeconómico con técnicas estatísticas relativamente recentes no mundo académico.

Así, o obxectivo principal que se pretende acadar é a obtención dunha tipoloxía que permita clasificar ós distintos fogares galegos en función de determinadas características. A obtención de tipoloxías proporciona unha ferramenta extremadamente útil ós investigadores sociais ou económicos debido, fundamentalmente, a que permite a simplificación dunha determinada realidade.

Esta simplificación proporciona un cadro de referencia que permite entre outras accións, definir correctamente un ámbito de medidas ou decisións, identificar con precisión ós destinatarios de determinadas políticas, transferir determinadas políticas entre unidades e, en última instancia, realizar unha correcta avaliación dos programas levados a cabo.

Polo que respecta ás técnicas estatísticas, neste estudio empregárase unha técnica clásica e habitual na análise de datos de enquisas como é a Análise de Correspondencias Múltiple (*ACM*) e unha técnica que se podería clasificar dentro da análise multivariante moderna, denominada Análise Factorial Múltiple (*AFM*) que axuda a resolver o problema que se presenta cando nos atopamos con táboas de tres índices, un índice para identifica-los individuos que son obxecto de estudio, un segundo índice para as variables que se mediron sobre os individuos, e un terceiro índice para as diversas situacións en que se realizaron as medicións.

2. MATERIAL E MÉTODOS

Os datos elixidos son os da *Enquisa de condicións de vida das familias* (IGE, 2001) correspondentes ó ano 1999. O deseño mostral da enquisa está concibido para dar resultados por áreas (agrupacións de comarcas) que son as seguintes:

- **Área 1: Coruña Oriental:** Arzúa, Betanzos, Eume, Melide, Ordes e Ortegal.
- **Área 2: Coruña Occidental:** A Barcala, Barbanza, Bergantiños, Soneira, Muros, Fisterra, Sar, Noia e Xallas.
- **Área 3: Comarca de Santiago.**
- **Área 4: Comarca da Coruña.**
- **Área 5: Comarca de Ferrol.**
- **Área 6: Lugo Sur:** A Ulloa, Fonsagrada, Ancares, Terra de Lemos, Chantada, Quiroga e Sarria.
- **Área 7: Comarca de Lugo.**
- **Área 8: Lugo Norte:** A Terra Chá, A Mariña Central, Oriental e Occidental e Meira.
- **Área 9: Ourense Occidental e Oriental:** Allariz, Maceda, Baixa Limia, Terra de Caldelas, Carballiño, Terra de Celanova, Ribeiro, Terra de Trives, Valdeorras, Verín, Viana e A Limia.
- **Área 10: Comarca de Ourense.**
- **Área 11: Pontevedra Occidental:** O Morrazo, O Baixo Miño, Caldas e O Salnés.
- **Área 12: Pontevedra Oriental:** Taboas-Terra de Montes, O Condado, Deza e A Paradanta.
- **Área 13: Comarca de Pontevedra.**
- **Área 14: Comarca de Vigo.**

Por outra banda, as 22 variables empregadas na análise son: ingresos netos mensuais, canalización pública de augas residuais, recollida de lixo, auga quente, calefacción, ascensor, lavalouza, microondas, nº de televisores, antena parabólica, canle de pago, vídeo, cámara de vídeo, cadea de música, ordenador, internet, coche, moto, garaxe, segunda vivenda, humidade, deficiencias de illamento térmico e compra de prensa por algún membro do fogar.

Referente á selección dos fogares como unidade de análise, a pretensión última é achegarse á unidade básica de actuación e decisión na sociedade. Finalmente, en canto o criterio xeográfico empregado, cómpre sinalar que a unidade de referencia é a comarca.

En canto ás técnicas estatísticas, a *ACM* permite estudar un conxunto de *I* individuos descritos por un grupo de *J* variables cualitativas. Os datos poderanse presentar nunha táboa do tipo “individuos x variables”, os primeiros en filas e os segundos en columnas e na intersección da fila e a columna aparece un valor que non posúe propiedade numérica. Pola súa banda, a *AFM* é un método que permite o tratamento simultáneo, como activas, de diferentes táboas de datos referidas ó mesmo conxunto de individuos, podendo ter diferente número de variables e ser de distinta natureza (cualitativa ou cuantitativa), aínda que dentro de cada unha das táboas deben ser da mesma natureza. Ademais, pondera as diferentes táboas equilibrando a influencia que sobre a análise global poida exercer cada unha delas. O procedemento consiste en compara-las nubes dos individuos a través das matrices de produtos escalares (matriz de covarianzas ou de correlación, segundo a transformación máis apropiada dos datos), equilibrando a súa influencia.

Os métodos tradicionais empregados para o estudo de enquisas só permiten estudia-las relacións entre as variables da mesma natureza, todas cuantitativas ou todas cualitativas, o que limita a súa utilización. Isto pode superarse mediante o emprego da técnica da *AFM*, que permite estrutura-la información en grupos diferentes e proporciona unha visión máis rica do fenómeno.

3. RESULTADOS

No tocante á **Análise de Correspondencias Múltiple** os elementos principais son as 49 modalidades das 22 variables correspondentes ás características dos fogares. Centrandonos especificamente no plano dos dous primeiros factores, hai que sinala-los seguintes aspectos:

En primeiro lugar, a proxección das características dos fogares sobre o primeiro plano factorial (Gráfico 1) amosa como as dúas modalidades da mesma variable (tenza ou ausencia) están sempre aliñadas coa orixe dos eixes. Máis aínda, certas modalidades, como a tenza e ausencia de cadea de música, teñen posicións case simétricas debido a que os fogares de cada modalidade son case iguais (57% e 43%). Cando isto non ocorre, é dicir os fogares non están totalmente equilibrados, a modalidade de maior peso está cerca da orixe e a modalidade de menor peso esta descentrada, como ocorre neste caso coa tenza e ausencia de internet nos fogares.

Asemade, o gráfico revela unha estrutura particular que convén resaltar: o segundo factor produce unha separación entre os dous tipos de modalidades, de tal xeito que tódalas modalidades relativas á ausencia están a esquerda do eixo mentres que as relativas á presenza están á dereita. Isto, parece indicar que globalmente hai un certo parecido entre o conxunto das ausencias e, por outro lado, entre o conxunto das modalidades presentes.

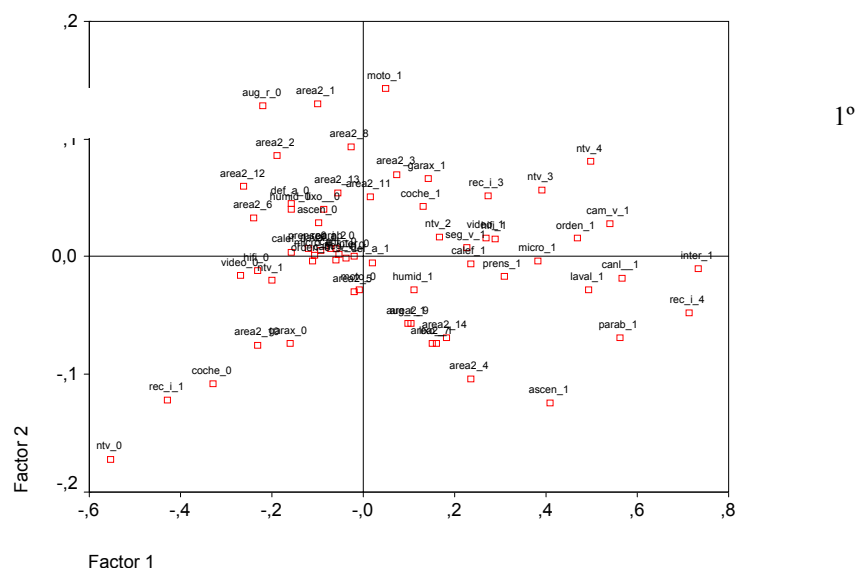
Afondando nestes resultados pode comprobarse como algunhas agrupacións son visibles sobre o plano dos dous primeiros factores. En particular, parece que a ausencia de determinadas características, así como os niveis de renda máis baixos están asociados ás áreas de Pontevedra Oriental, Coruña Oriental, Coruña Occidental e Lugo Sur. Pola contra semella que a presenza de características tales como, canle de pago, internet e un nivel de ingresos máis elevado vencéllase coas comarcas de Lugo, Vigo e A Coruña.

Igualmente, outra característica que se pode observar é que os fogares con maiores ingresos están asociados cos fogares que teñen máis presenza das características relacionadas coas novas tecnoloxías (canle de pago, internet, ordenador, ...)

O gráfico da proxección das modalidades, permite deduci-la existencia de varios grupos de fogares que manifestan un comportamento diferente respecto das características que posúen:

- Primeiro grupo (situado no 4º cuadrante): comprende os fogares que posúen un nivel de ingresos moi elevado e que posúen no fogar aparatos relacionados coas novas tecnoloxías: internet, ordenador, canle de pago, parabólica, etc
- Segundo grupo (situado no 1º cuadrante): trátase dos fogares que posúen un nivel de ingresos alto e que teñen certas comodidades tales como: vídeo, cadea de música, máis de dúas televisións, segunda vivenda, etc.
- Terceiro grupo (situado no 2º e 3º cuadrante): vén dado polos fogares con niveis de ingresos baixos e moi baixos e que están moi asociados a ausencia das modalidades correspondentes ás variables empregadas

Gráfico 1.- Plano dos dous primeiros factores (ACM)



Na tipoloxía de fogares obtida por esta análise, a caracterización das áreas xeográficas e o perfil dos fogares segundo diversas características interveñen conxuntamente, dous fogares próximos se se parecen porque teñen características. Pero esta igualdade aparece tamén no 4º cuadrante.

ocultar un desequilibrio importante, xa que nada impide que unha das áreas predomine sobre a outra. Por este motivo, o obxectivo que se persigue cando se consideran varios grupos heteroxéneos é que os dous interveñan do mesmo xeito na tipoloxía.

Con respecto á *Análise Factorial Múltiple*, hai que sinalar que estudiáronse as 22 variables para as 14 áreas xeográficas, obtendo como resultado que os factores tiñan unhas interpretacións semellantes en tódolos casos. Así, o primeiro factor está asociado a tenza ou ausencia de determinadas características, así como ós ingresos, mentres que o segundo sempre aparece asociado a algunha característica que é moi pouco común entre os fogares da enquisa, por exemplo, a ausencia de televisor no fogar. O terceiro factor xa é menos interpretable.

Mantivéronse 3 factores para cada unha das *ACM* realizadas sobre as 14 áreas xeográficas. A matriz de correlación entre os factores dos 14 *ACM* (antes de ponderar polos inversos dos primeiros valores propios) permítenos observar que as correlacións entre os factores da mesma orde entre áreas diferentes son altas, co que se pode concluír que os factores das diferentes *ACM* teñen a mesma interpretación. Nembargantes obsérvase como esta correlación é forte para os primeiros factores, aínda que diminúe para os segundos e terceiros, o que pode ser debido a que os factores de maior orde están máis afectados polas fluctuacións da mostraxe.

Para obter unha medida de semellanza entre as diferentes áreas pódense observa-los coeficientes *RV*, que se poden interpretar como coeficientes de correlación entre as táboas. Estes coeficientes preséntanse a continuación:

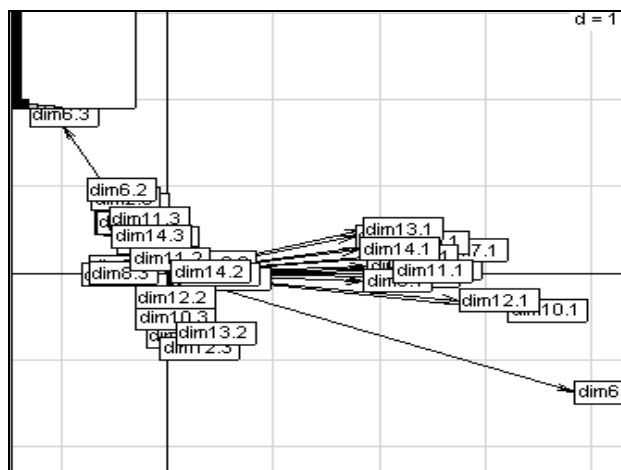
Cadro 1.- Coeficientes *RV* de Escoufier

	Cor_ori	Cor_Occ	Com_Sant	Com_Coru	Com_Ferrol	Lugo_Sur	Com_Lugo	Lugo_Norte	Ourense_OrOcc	Com_Ourense	Pontevedra_occ	Pontevedra_orien	Com_Pontevedra
Cor_ori	1,00												
Cor_Occ	0,83	1,00											
Com_Santiago	0,89	0,86	1,00										
Com_Coruña	0,79	0,76	0,87	1,00									
Com_Ferrol	0,86	0,88	0,92	0,83	1,00								
Lugo_Sur	0,77	0,79	0,71	0,54	0,78	1,00							
Com_Lugo	0,86	0,76	0,91	0,89	0,86	0,64	1,00						
Lugo_Norte	0,84	0,80	0,85	0,80	0,78	0,58	0,83	1,00					
Ourense_OrOcc	0,86	0,83	0,90	0,90	0,89	0,71	0,91	0,83	1,00				
Com_Ourense	0,91	0,81	0,82	0,69	0,81	0,73	0,81	0,85	0,78	1,00			
Pontevedra_occ	0,91	0,88	0,94	0,87	0,95	0,74	0,89	0,84	0,92	0,83	1,00		
Pontevedra_orien	0,87	0,74	0,75	0,64	0,74	0,69	0,73	0,78	0,73	0,89	0,78	1,00	
Com_Pontevedra	0,79	0,81	0,86	0,83	0,81	0,63	0,84	0,88	0,89	0,73	0,85	0,70	1,00
Com_Vigo	0,85	0,81	0,92	0,94	0,89	0,63	0,93	0,85	0,93	0,76	0,93	0,71	0,88

Nesta táboa vese como a área de Lugo Sur é a que está menos correlacionada co resto, tendo o valor máis baixo coa área de Lugo Norte (0,58). As áreas que están máis correlacionadas son a Comarca de Vigo coa comarca de A Coruña (0,94) seguidas da Comarca de Santiago con Pontevedra Occidental (0,94). Aínda así, é difícil facerse unha idea clara das relacións, debido ó número de áreas.

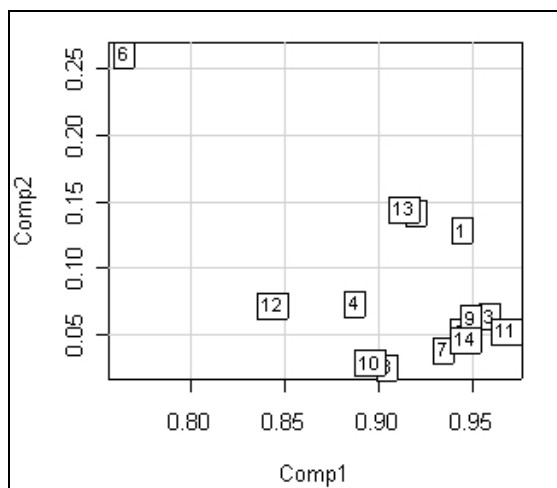
No seguinte gráfico (Gráfico 2) preséntanse as posicións dos factores de cada área sobre os dous primeiros eixes da *AFM*. Obsérvase a elevada correlación entre os primeiros factores das diferentes áreas co primeiro factor da análise global, mentres que para o segundo e terceiro esta correlación perde intensidade e asóciase máis ó segundo factor.

Gráfico 2.- Proxección das columnas na *AFM*



A continuación (Gráfico 3), vense as posicións das áreas sobre os dous primeiros factores. Como se pode comprobar, o primeiro factor separa claramente a Coruña Oriental e Lugo Sur da área de Pontevedra Occidental, mentres que o factor 2 separa dúas áreas que, aínda que pertencen á mesma provincia, caracterízanse por ter unhas características marcadamente diferentes: Lugo Norte e Lugo Sur

Gráfico 3.- Posicións das áreas sobre os dous primeiros factores da *AFM*



Parecería lóxico pensar que a disposición das áreas no gráfico anterior debería estar relacionada cos ingresos medios das áreas nas que se subdividiu este estudo. Neste senso, se comparamos o Gráfico 3 co Cadro 2 onde se presentan a media de ingresos netos mensuais por área obsérvase como hai bastante correlación, é dicir áreas con ingresos altos aparecen máis próximas no gráfico.

Cadro 2.- Media de netos mensuais por áreas

Código	Áreas	Media de ingresos netos mensuais (€)
7	Comarca de Lugo	1.308,02
4	Comarca de A Coruña	1.305,17
3	Comarca de Santiago	1.277,28
14	Comarca de Vigo	1.158,66
13	Comarca de Pontevedra	1.124,03
10	Comarca de Ourense	1.103,49
8	Lugo Norte	1.101,49
11	Pontevedra Occidental	1.082,24
2	Coruña Occidental	1.034,46
5	Comarca de Ferrol	995,06
1	Coruña Oriental	994,51
12	Pontevedra Oriental	948,81
9	Ourense Occidental e Oriental	883,55
6	Lugo Sur	883,40

4. CONCLUSIÓNS

O estudio dos fogares galegos agrupados en torno ás 14 áreas xeográficas en función das características que permite observa-la *Enquisa de Condicións de Vida das Familias* supón un paso máis no coñecemento da estrutura socioeconómica galega. Así, o emprego de técnicas factoriais de análise de táboas múltiples configúrase como un aspecto salientable neste tipo de estudos.

En canto ós resultados obtidos e, establecidas as cautelas previas de toda análise estatística, especialmente polo que se refire á limitación a un número determinado de variables e o seu carácter estático, o que é indubidable é a súa utilidade na análise económica aplicada. Ademais, cando se obteñen resultados que coinciden, ou polo menos non difiren en grande medida, das intuicións que se podían ter a priori, a sensación de comodidade está presente de modo importante.

Así cando a *ACM*, diferencia por un lado as áreas de Pontevedra Oriental, Coruña Oriental, Coruña Occidental xunto con Lugo Sur e, por outro, as comarcas de Lugo, Coruña e Vigo, estes resultados semellan coherentes ou non sorprendivos. Igualmente a clasificación dos fogares en tres grupos diferenciados, pode ser cualificada como moi adecuada a efectos prácticos. Así, o primeiro deles vén vinculado a ingresos moi elevados e dotacións vencelladas ás novas tecnoloxías, o segundo aparece relacionado con niveis altos de ingresos e certas comodidades do fogar, namentres que o último situaría ós fogares de baixos niveis de ingresos e con carencia de boa parte das variables estudias.

Pola súa banda, as conclusións extraíbles da *AFM* son, se cabe, aínda de maior interese. Así, traballáronse con tres factores, dos que dous, eran claramente interpretables: posesión ou non dalgunha das variables e nivel de ingresos e ausencia dalgunha característica moi pouco frecuente na mostra. A partir de aquí, pode extraerse a diferenciación da área de Lugo sur, que se configura como a menos correlacionada co resto de áreas, mentres que no tocante ás semellanzas, son as comarcas de Vigo e Coruña as que amosan unha maior correlación, seguidas das comarcas de Santiago coa área de Pontevedra occidental.

Estas características coinciden xa non coa suposición que se puidera ter previo o estudio, senón que na contrastación co nivel de ingresos obsérvase que se trata de variables relacionadas en boa medida. Así pois, este primeiro paso ou aproximación na obtención dunha tipoloxía dos fogares galegos permite unha separación dos mesmos nunha dobre vertente, por un lado segundo teñan ou non determinadas variables (que están á súa vez vencelladas ó seu nivel de ingresos) e, por outro lado, segundo a área xeográfica onde se localicen. En concreto, en canto ás variables, poderíase falar de tres situacións: dotación de novas tecnoloxías, dotación de comodidades do fogar e ausencia de comodidades básicas

Polo que se refire ó emprazamento xeográfico como é frecuente neste tipo de estudos, a polarización entre o mundo rural e o urbano aparece como característica dos fogares galegos.

Finalmente, resta plantexar as posibilidades e perspectivas que se abren tras esta análise. Certamente son moitos e moi variados os camiños que poden ser tomados, se ben gustaríanos destacar dous deles. En primeiro lugar, a aplicación e contrastación con outro tipo de técnicas estatísticas como pode se-lo método *Statis* proposto por Escoufier, Y., e, por outra banda, a inclusión do compoñente dinámico nas variables empregadas de tal xeito que a perspectiva temporal fose dalgún xeito contemplada.

5. REFERENCIAS BIBLIOGRÁFICAS

- Abascal, E., Fernández, K., Landaluce, M.I., Modroño, J.I. (2001). “Diferentes aplicaciones de las técnicas factoriales de análisis de tablas múltiples en las investigaciones mediante encuestas” en *Metodología de Encuestas*. Vol. 3, Núm 2, pp. 251-279
- Abascal, E., Landaluce, M.I. (2000). *Análisis factorial múltiple como técnica de estudio de la estabilidad interna de los resultados de un Análisis de Componentes Principales. Una aplicación empírica*. Questiiio
- Escoufier B., Pagés, J. (1992). *Análisis factoriales simples y múltiples: objetivos métodos e interpretación*. Servicio Editorial Universidad del País Vasco. Bilbao
- Escoufier B., Pagés, J. (1994). “Multiple Factor Analysis” en *Computational Statistics and Data Analysis*, 18, 121-140
- IGE (2001). *Enquisa de condicións de vida das familias. Módulo de características básicas dos fogares*. Xunta de Galicia, Santiago de Compostela
- Landaluce, M.I., Fernández, K., Modroño, J.I. (1999). “Reflexiones sobre el Uso Comparativo del Análisis Factorial Múltiple (AFM) y de la metodología Statis para el Análisis de Tablas Múltiples” en *Methodologica*, Vol 7, pp. 37-65
- Lima Santos, J.M. (2002) “Tipologia de territorios rurais: uma análise multivariada” en *Actas das IX Jornadas de Classificacao e Analise de Dados – JOCLAD 2002*. Lisboa.