

Seminario da Rede Tecnoloxías e Análise de Datos Lingüísticos

Filoloxía e Lingüística Galega
Modelos de Optimización, Decisión, Estatística e
Aplicacións



DEPARTAMENTO DE ESTATÍSTICA,
ANÁLISE MATEMÁTICA E OPTIMIZACIÓN

TecAnDaLi

Tecnoloxías e Análise dos Datos Lingüísticos



INSTITUTO DA LINGUA GALEGA (ILG)

Xeografía lingüística - Apelidos

Os apelidos proporcionan indicadores relacionados con aspectos

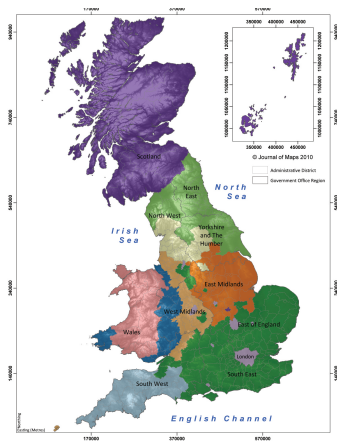
- demográficos, étnicos, culturais e xenéticos das poboacións.

A súa transmisión atende a características

- socio-económicas, culturais e xeográficas.

Problemas que podemos abordar

- Identificación de rexións de apelidos.
- Identificación de patróns espaciais subxacentes.



Extraída de Cheshire e Longley (2011). Rexións de apelidos en Gran Bretaña.



Actualidade — Nados e 1965 ou antes — Divisións dialectais de Asturias

- Dos resultados obtidos pode inferirse a existencia de vinculacións entre a distribución de apelidos e as variedades lingüísticas identificadas para o asturiano.
- Na rexionalización por apelidos de Asturias, atópanse bastantes semellanzas coas **divisións dialectais** propostas tradicionalmente (de esquerda a dereita, Galego Asturiano, Bable Occidental, Bable Central e Bable Oriental).



Proxecto de colaboración

- **Proposta:** construción das rexións de apelidos en España.
 - Destacar que ata agora non se realizou ningunha análise exhaustiva sobre os datos do censo.
 - Os estudos existentes son moi parciais e non realizan as correccións do noso método.

Distancias lingüísticas e onomásticas

Proxecto de colaboración

- **Proposta:** estudo comparado das distancias lingüísticas e das distancias onomásticas.



Calderón R, Hernández CL, Cuesta P, Dugoujon JM (2015) Surnames and Y-Chromosomal Markers Reveal Low Relationships in Southern Spain. *PLoS ONE* **10(4)**: e0123098. doi:10.1371/journal.pone.0123098



Rodríguez-Díaz R, Manni F, Blanco-Villegas MJ (2015) Footprints of Middle Ages Kingdoms Are Still Visible in the Contemporary Surname Structure of Spain. *PLoS ONE* **10(4)**: e0121472. doi:10.1371/journal.pone.0121472

Proxecto de colaboración

- **En curso:** construción das rexións de apelidos (segundo as propostas da literatura).
- **Proposta:** adaptación de novas medidas que teñan en conta a *abundancia* dos apelidos (adaptando medidas de biodiversidade).

As técnicas poderían adaptarse á construción de rexións dialectais (?).

- Traballaremos cos índices de Shannon e Simpson que foron creados para o ámbito da ecoloxía e irémolos adaptando no futuro á análise espacial da onomástica.

Patróns espaciais

Os apelidos poden clasificarse segundo a súa orixe como

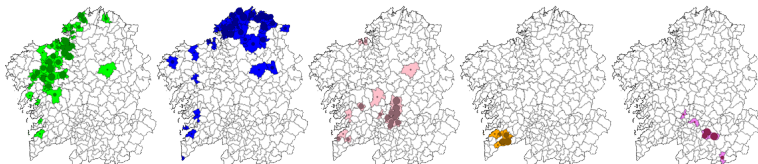
- Nomes comúns: Blanco.
- Patronímicos: derivados de nomes (Fernando \Rightarrow Fernández).
- Toponímicos: derivados de nomes de lugares (Cures).

A distribución espacial dos apelidos (especialmente a dos toponímicos) presenta un patrón espacial non aleatorio, que pode deberse a múltiples causas (moitas non observables)...

- Campo latente.

Para un apelido concreto, o obxectivo é atopar unha estrutura subxacente que explique a variación espacial (e posiblemente temporal) do apelido tendo en conta unha compoñente de veciñanza.

Distribución espacial dalgúns apelidos

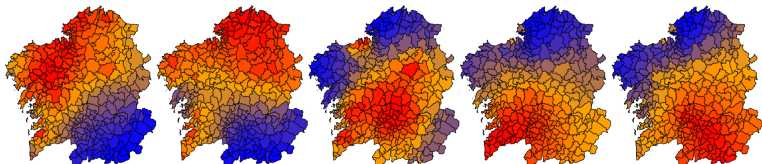


Localizacións en Galicia dos apelidos Miramontes, Galdo, Cudeiro, Farrapeira e Padreda. O tamaño do punto é proporcional ao número de persoas que levan ese apelido ponderado pola poboación do concello.

Proxecto de colaboración

- **En curso:** identificación dos patróns espaciais con técnicas da epidemioloxía.
- **Proposta:** extensión aos **patróns espaciais**; introdución de distancias non euclídeas.
- **Proposta:** extensión aos **patróns espazo–temporais**; introdución de distancias non euclídeas.

Modelo espacial (axuste con INLA)



Media a posteriori da parte estruturada do campo latente. Apelidos Miramontes, Galdo, Cudeiro, Farrapeira e Padreda (Censo 2011).



Besag, J., York, J. and Mollié, A. (1991) Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, **43**, 1–59.



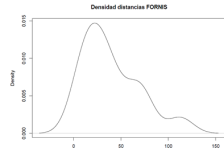
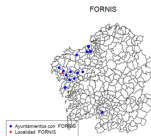
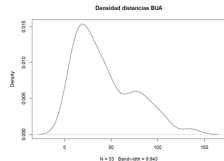
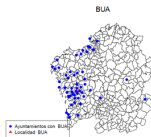
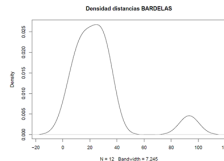
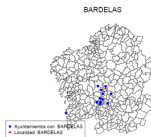
Rue, H., Martino, S. and Chopin, N (2009) Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations. *Journal of the Royal Statistical Society, Series B*, **71**, 319–392.



Rue, H. and Held, L. (2005) *Gaussian Markov Random Fields: Theory and Applications*. Chapman and Hall, CRC Press, London.

Proxecto de colaboración

- En curso: identificación dos patróns espaciais con técnicas da estatística non paramétrica.



Seminario da Rede Tecnoloxías e Análise de Datos Lingüísticos

Filoloxía e Lingüística Galega
Modelos de Optimización, Decisión, Estatística e
Aplicacións



DEPARTAMENTO DE ESTATÍSTICA,
ANÁLISE MATEMÁTICA E OPTIMIZACIÓN

TecAnDaLi

Tecnoloxías e Análise dos Datos Lingüísticos



INSTITUTO DA LINGUA GALEGA (ILG)