

demogal: paquete para proyectar flujos y stocks de población

RESUMEN

El objetivo de este trabajo es el desarrollo de un paquete en el lenguaje de programación R, **demogal**, que permita la elaboración de proyecciones de la población para uno o más espacios geográficos, basándose en el método clásico de las componentes y utilizando el modelo de proyección multirregional. El paquete demogal incorpora en la elaboración de la proyección la posibilidad de consistencia con un espacio superior o no.

Palabras y frases clave: proyección de población, paquete R, demografía.

1. Introducción

La elaboración de indicadores demográficos y proyecciones de población es una tarea habitual de las oficinas de estadística pública porque, en todo tipo de planificación, es fundamental conocer la estructura demográfica actual y las perspectivas de evolución de la población. Hoy en día existen herramientas informáticas que permiten el cálculo de indicadores y de las tablas de mortalidad de manera sencilla, como pueden ser **Epidat**[2], hojas de cálculo, etc. Estas herramientas, en algunos casos, no tienen la flexibilidad necesaria para adaptarse a cualquier metodología de cálculo, en otros casos, no permiten procesar grandes conjuntos de datos y, lo que es más importante, cuando se maneja gran cantidad de información, la mayor parte de ellas no permiten automatizar el proceso de cálculo, etc. Teniendo en cuenta todas estas dificultades que plantean las herramientas existentes, desde el Instituto Galego de Estatística (IGE) se pensó en aprovechar la flexibilidad de R para desarrollar un paquete que permitiese automatizar el cálculo de todas las funciones demográficas que periódicamente se utilizan en el IGE.

2. Metodología

La elaboración de estas proyecciones de población está basada en el método clásico de componentes[3], el cual, a partir de una población residente en un cierto espacio geográfico y de los datos observados para cada una de las componentes demográficas básicas (mortalidad, fecundidad y migración), permite obtener la población en periodos futuros bajo ciertas hipótesis sobre la evolución de las componentes, que son las que determinan el crecimiento de la población y su estructura por edades. Se utiliza el modelo de proyección multirregional propuesto por Willekens[1], que proporciona las cifras de población residente y las cifras proyectadas de nacimientos, defunciones y movimientos migratorios que tendrán lugar en cada espacio considerado y para cada uno de los periodos del horizonte proyectivo, guardando todo ello la necesaria coherencia entre flujos y stocks demográficos y la debida consistencia interterritorial. En el paquete demogal se contemplan dos enfoques:

- Proyección de la población para uno o más espacios geográficos aplicando el modelo multirregional.
- Proyección de la población para varios espacios geográficos mediante el modelo multirregional, pero imponiendo que los datos agregados del conjunto de espacios coincidan con los de una proyección preexistente del área formada por este conjunto de espacios (consistencia con el espacio superior).

El método empleado proyecta cada uno de los fenómenos demográficos por separado, para después unirlos en la obtención de la proyección de la población. Las hipótesis que se hayan planteado sobre cada uno de estos fenómenos demográficos se aplican a la población en el periodo de partida t_0 , obteniendo la población en el periodo t_{0+1} . Actuando así, de forma iterada, se llega al horizonte de la proyección, t_{0+nn} . Para la obtención de la proyección de la población será necesario disponer de información de nacimientos, defunciones y movimientos migratorios correspondiente a los últimos n periodos: $t_{0-n}, t_{0-(n-1)}, \dots, t_{0-1}$.

3. El paquete demogal

En el paquete demogal hay un total de 13 nuevas funciones que pueden ser de interés para las personas que se dedican a hacer indicadores y proyecciones de distintos tipos de poblaciones. Las funciones se presentan de manera resumida en la Tabla 1.

Function	Description
demogal.proj	Esta es la función principal del paquete que proyecta la población, los nacimientos y las defunciones a partir de datos de población, nacimientos, defunciones y migraciones de periodos anteriores.
mortable	Permite calcular las funciones de la tabla de mortalidad completa o abreviada por sexo, para diferentes espacios geográficos y diferentes periodos de tiempo a partir de los datos de defunciones y población.
project	Proyecta diferentes magnitudes de la fecundidad en base a una relación log-lineal de las mismas y el tiempo. Se utilizan n periodos para hacer la proyección.
tr	Calcula el Índice sintético de fecundidad y el Índice sintético de emigración.
PMortality	Proyecta las tasas de mortalidad y los supervivientes de la tabla de mortalidad para el espacio de proyección.
PFertility	Proyecta el Índice sintético de fecundidad y las tasas de fecundidad para el espacio de proyección.
ExtEm	Proyecta las tasas de emigración al exterior del espacio de proyección.
PMortalityReg	Proyecta y regionaliza las tasas de mortalidad para el espacio de proyección, sexo, edad y todo el periodo de proyección.
PFertilityReg	Proyecta y regionaliza las tasas de fecundidad para el espacio de proyección, edad y todo el horizonte de proyección.
ExtEmReg	Proyecta y regionaliza las tasas de emigración al exterior del espacio de proyección.
ExtIn	Proyecta y regionaliza las inmigraciones procedentes del exterior.
IntEm	Proyecta las tasas de emigración interna en todo el periodo de proyección.
projection	Proyecta la población para el espacio de proyección, sexo, edad y todo el periodo de proyección.

Tabla 1: Nuevas funciones del paquete demogal.

El flujo de utilización de las funciones se muestra en la Figura 1. En lo que sigue se presentan un par de ejemplos ilustrativos del uso de las funciones descritas en la Tabla 1.

4. Ejemplo de uso de demogal

El paquete demogal permite realizar proyecciones de la población para uno o más espacios geográficos. En el caso de que se considere más de un espacio geográfico se puede optar por la consistencia con un espacio superior (unión de cada uno de los espacios de proyección) o no.

Se presenta en primer lugar un ejemplo de la proyección de Galicia del año 2016 al 2030. El paquete *demogal* requiere el input de un parámetro: $nn = 15$ (número de años que se van a proyectar). Para llevar a cabo esta proyección se necesita además información de los fenómenos demográficos en periodos anteriores, así como la población del periodo base. Se utiliza por tanto la siguiente información:

- Población del año 2015 por sexo y grupos quinquenales de edad (*poboacion*).

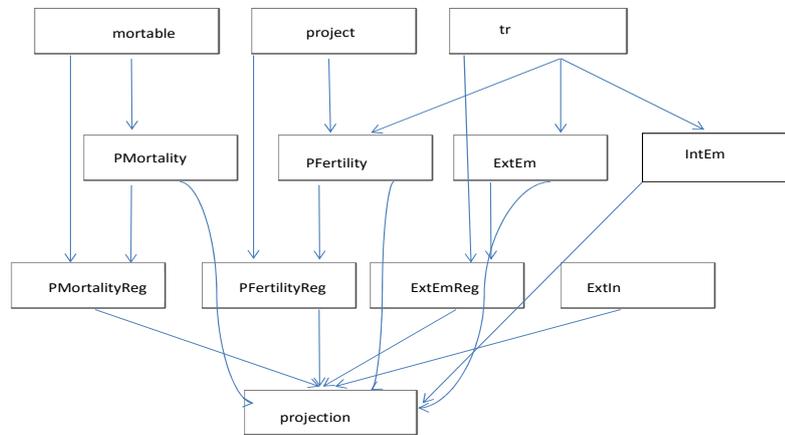


Figura 1: Flujo de ejecución del paquete.

- Nacimientos por sexo y grupos quinquenales de edad de la madre para los 10 años anteriores (*data.nac.ga*).
- Mortalidad, población y defunciones por sexo y grupos quinquenales de edad para los 10 años anteriores (*data.ga*).
- Emigrantes al exterior por sexo y grupos quinquenales de edad, para los últimos 3 años (*emig.out.ga*).
- Inmigrantes procedentes del exterior por sexo y grupos quinquenales de edad, para los últimos 3 años (*imig.out*).

A continuación se presenta un breve resumen de las funciones utilizadas:

```

R> ###PROYECCIÓN DE GALICIA
R> nn=15
R> #PROYECTAR MORTALIDAD
R> m.rate.pr=PMortality(data.ga, nn)[[1]]
R> #PROYECTAR FERTILIDAD
R> f.rate.pr=PFertility(data.nac.ga,nn)[[2]]
R> #PROYECTAR EMIGRACIONES
R> u=ExtEm(emig.out.ga,nn)
R> #PROYECTAR INMIGRACIONES
R> prov=12
R> inmig=27000
R> v=as.data.frame(cbind(prov,inmig))
R> uin=ExtIn(inmig.out=imig.out,nn=nn,v_inmig=v)
R> uin.ga=apply(uin,c(2,3,4),sum)
R> uin.ga=array(uin.ga,c(1,2,22,15))
R> dimnames(uin.ga)=dimnames(u[[1]])
  
```

```

R> ###SE DEFINE la matrix de transición (r)
R> r<-expand.grid(unique(poboacion[,1]),unique(poboacion[,2]),unique(poboacion[,3])[2:21])
R> r[,4]=0.2
R> names(r)=c("espazo", "sexo", "Idadeq", "dato")

R> #Lista con las emigraciones y las inmigraciones. En este caso no hay migraciones internas
R> migration.ga=list(u[[1]],uin.ga,NULL)
R> ### SE PROYECTA LA POBLACIÓN DE GALICIA
R> m=projection(nn,r,poboacion,m.rate.pr,f.rate.pr,migration.ga,relation=NULL)

```

En la Figura 2, 3 se presentan los resultados de la evolución de la proyección de la población y de los nacimientos y defunciones de manera gráfica.

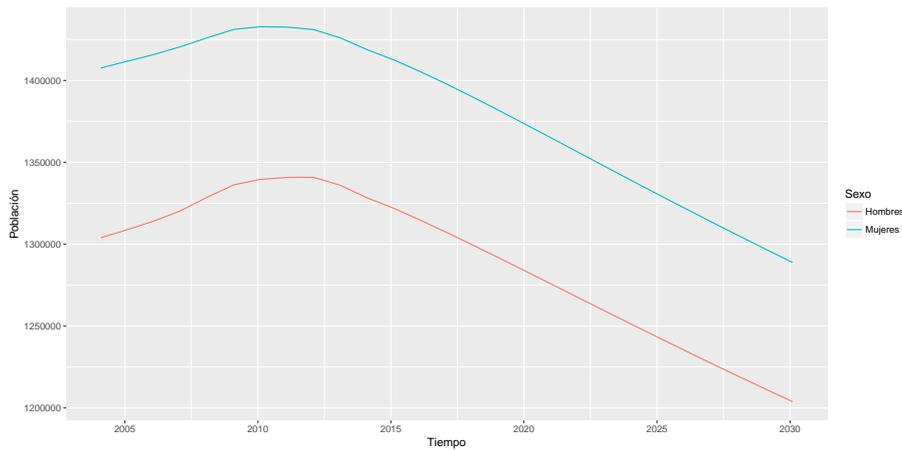


Figura 2: Detalle de la proyección de la población por sexo.

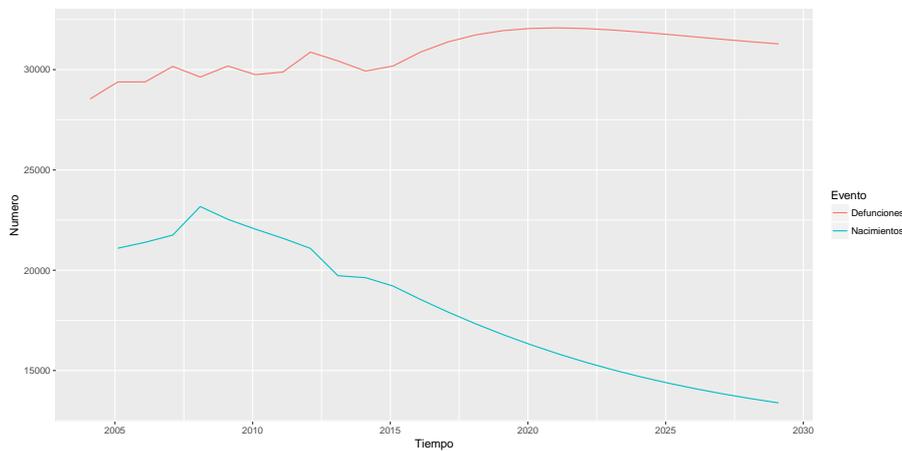


Figura 3: Detalle de la proyección de los nacimientos y las defunciones.

Se considera ahora un segundo ejemplo en el que se pretende hacer una proyección de la población de las provincias gallegas que sea consistente con la proyección hecha en el primer ejemplo para Galicia. La consistencia con el espacio superior se indica en la función de proyección con el dataframe *relation* y, al igual que en el primer ejemplo, se requiere también el input del parámetro $nn = 15$. Por lo que respecta a la información necesaria para realizar la proyección, en este caso se necesitará además de la misma información de partida que en el primer ejemplo pero relativa a cada provincia, la siguiente información:

- Población, nacimientos y defunciones por sexo y grupos de edad proyectados para Galicia.
- Fichero que relacione los espacios geográficos de proyección (provincias) con el espacio superior al que pertenecen (Galicia) (*relation*).

A continuación se presenta un breve resumen de las funciones utilizadas:

```
R> year = 2015
R> pp = m[[1]]
R> p = array(0, dim = c(dim(pp[[1]]), length(pp)))
R> dimnames(p)[1:3] = dimnames((pp)[[1]])
R> dimnames(p)[[4]] = ((year + 1):(year + nn))
R> for (i in 1:length(pp)) {
+   p[, , , i] = pp[[i]]
+ }
R> pop.sup.pr = adply(p, c(1, 2, 3, 4))
R> names(pop.sup.pr) = c("espazo", "sexo", "Idadeq", "ano", "pop")
R> b.sup.pr = adply(m[[2]], c(1, 2, 3))
R> names(b.sup.pr) = c("espazo", "sexo", "ano", "nac")
R> m.sup.pr = adply(m[[3]], c(1, 2, 3, 4))
R> names(m.sup.pr) = c("espazo", "sexo", "Idadeq", "ano", "def")
R> prox_m_ga <- PMortality(data.ga, nn)
R> prox_m_proReg <- PMortalityReg(data.prov, prox_m_ga[[2]], relation,
+   reg = "W")
R> tfr.sup = PFertility(data.nac.ga, nn)
R> prox_n_pro <- PFertilityReg(data.nac.prov, tfr.sup[[1]], nn,
+   relation)
R> internas = IntEm(emig.int_prov, emig.intc_prov, nn)
R> u = ExtEm(emig.out.ga, nn)
R> iset = u[[2]]
R> ise = u[[3]]
R> uu = ExtEmReg(emig.out, iset, ise, relation, nn)
R> prov = 12
R> inmig = 27000
R> v = as.data.frame(cbind(prov, inmig))
R> uin2 = ExtIn(imig.out, relation, nn, v_inmig = v)
R> m_prov = projection(nn, r, pop_prov, prox_m_proReg, prox_n_pro,
+   migration = list(uu, uin, internas), relation, pop.sup.pr,
+   b.sup.pr, m.sup.pr)
```

Referencias

- [1] Willekens, F.J. and Drew, P. (1984), A Multiregional Model for Regional Demographic Projection. *Demographic Research and Spatial Policy*, Ch.15, pp. 309-331 - Academic Press, London.
- [2] Epidat: programa para análise epidemiolóxica de datos. Versión 4.0, xuño 2012. Consellería de Sanidade, Xunta de Galicia, España; Organización Panamericana de la Salud (OPS-OMS). Disponible en: <http://dxsp.sergas.es>.
- [3] Vinuesa, Julio y Puga, Dolores (2007), Técnicas y ejercicios de Demografía