
MEMORIA TÉCNICA DE LA ACTIVIDAD



“ESTADÍSTICA DE ITINERARIOS VITALES DE LOS RESIDENTES EN ANDALUCÍA”

Producto de difusión:

- ESTADÍSTICAS LONGITUDINALES DE SUPERVIVENCIA Y LONGEVIDAD EN ANDALUCÍA
2002-2016
- ESTADÍSTICAS LONGITUDINALES DE BIOGRAFÍAS REPRODUCTIVAS EN ANDALUCÍA
2002-2013

ÍNDICE

0. IDENTIFICACIÓN DE LA ACTIVIDAD.....	3
1. INTRODUCCIÓN.....	4
2. OBJETIVOS.....	9
3. METODOLOGÍA.....	11
4. PLAN DE ANÁLISIS, TABULACIÓN Y COMPILACIÓN.....	13
5. PLAN DE DIFUSIÓN.....	22
6. CRONOGRAMA.....	24

0. IDENTIFICACIÓN DE LA ACTIVIDAD

Código y denominación de la actividad

01.01.03 Estadísticas de itinerarios vitales de los residentes en Andalucía

Tipo de actividad

Actividad estadística y cartográfica

Clasificación de la actividad

- En función de su etapa de desarrollo: proyecto
- En función de su objeto: actividad de producción

Área temática

1. Población, Familia y Hogares

Subárea temática

1.01 Demografía y población

Sectores económicos u otros cubiertos por la actividad

-

Organismo responsable

Instituto de Estadística y Cartografía de Andalucía

Unidad ejecutora

Servicio de Estadísticas Demográficas y Sociales

Organismos colaboradores y convenio

Consejo Superior de Investigaciones Científicas

1. INTRODUCCIÓN

Área de estudio

Esta actividad estadística tiene encomendada la tarea de preparar y explotar la información contenida en la Base de Datos Longitudinal de Población de Andalucía (BDLPA), centrándose específicamente en los aspectos longitudinales de esta fuente. Su objeto de investigación son las trayectorias o "itinerarios vitales" que los distintos miembros de una población generan a lo largo de su vida. Conceptos como "ciclo vital", "desarrollo" o "envejecimiento", son connaturales con esta óptica. Una investigación longitudinal pretende no tanto conocer la frecuencia de determinados "sucesos" en determinado colectivo, sino las secuencias de eventos, sus interrelaciones, inter-dependencias y las duraciones de los intervalos inter-eventos, así como el efecto que sobre estas secuencias tienen determinadas características estructurales que delimitan grupos poblacionales y sociales.

A modo de ejemplo aclaratorio, una investigación estadística realizada desde una óptica transversal suele centrarse en las características de una población en un momento dado¹, o bien en conocer la intensidad de un determinado fenómeno en una año dado². Por el contrario una investigación realizada desde una perspectiva longitudinal se interesa en obtener las probabilidades de que ocurra un determinado evento asociado a una transición que determina la aparición de un nuevo estado³, así como en el tiempo medio que se ha tardado en alcanzarlo⁴. También es habitual preguntarse desde la óptica longitudinal sobre cuál es el impacto que pertenecer a un determinado colectivo tiene sobre estas trayectorias (por ejemplo, como de distinta es la movilidad residencial, las trayectorias de actividad económica, o la longevidad en los distintos colectivos definidos por su nivel de educación).

Marco conceptual: conceptos y definiciones

La actividad estadística que describe esta memoria técnica se encarga de procesar y preparar la información de la BDLPA desde una óptica "longitudinal" o de los "itinerarios vitales".

- Base de Datos Longitudinal de Población de Andalucía (BDLPA). Es una actividad cuyo objetivo es mantener una base de datos que contiene la información de las inscripciones padronales de los residentes en Andalucía desde 1998. La información proviene

¹Como por ejemplo la proporción de determinada edad, con un nivel estudios determinado y con una pensión no contributiva.

²Como por ejemplo la esperanza de vida o el índice sintético de fecundidad en un año dado.

³Como por ejemplo un determinado nivel de cualificación educativa, el primer empleo, el nacimiento de un hijo ...

⁴Como por ejemplo la duración media de una graduación, el tiempo medio de búsqueda del primer empleo, el intervalo intergenésico o la duración media de la vida para determinados colectivos.

principalmente del Registro de Población de Andalucía, aunque adicionalmente contiene información proveniente de otros registros poblacionales.

- El término "itinerario vital" proviene de las teorías sociológicas de "ciclo vital" que intentan encasillar las historias individuales en base a una sucesión relativamente simple de etapas y procesos por los que la vida de los individuos suelen transitar. Partiendo de alguna de las ideas recogidas en los desarrollos teóricos del "ciclo vital", como las de episodios y transiciones, se han readaptado los conceptos para su uso estadístico y, partir de su adecuada combinación, surge el concepto de "itinerario vital".
- "Apuntes registrales" son las unidades elementales de trabajo de esta actividad estadística que se almacenan en la BDLPA. Representa cada una de las distintas entradas de información de alguna de las fuentes usadas por el sistema y que incluyen su correspondiente marco temporal especificado.

Los "apuntes registrales" se pueden clasificar en dos tipos: "variaciones" y "observaciones". Se denominan "variaciones" a aquellos apuntes que informan de modificaciones en algún estado o característica del individuo ocurrido en una determinada fecha. Se denominan "observaciones" a aquellos apuntes que informan de determinadas características o estados poseídos por un individuo concreto en determinada fecha, aunque no se conozca la fecha exacta en que dicho estado se adquirió. Por ejemplo, la información recogida en un censo, suele tener la característica de observación.

- Todos los apuntes registrales de un mismo individuo son enlazados e identificados por una clave numérica única (IDP).
- A todos los apuntes con un mismo IDP se les denominan "expediente registral".
- Denominaremos "secuencia registral" al conjunto de apuntes ordenados por criterios temporales y espaciales de un "expediente registral". Cada uno de los componentes de esta secuencia se encuentran relacionados con registros adyacentes por relación de precedencia y consecuencia. En el caso concreto de los apuntes padronales (o variaciones padronales) se distinguen claramente dos tipos de relaciones: diacrónicas y sincrónicas.
- Son "relaciones diacrónicas" los pares de variaciones (alta, baja) producidas en un mismo lugar y donde la fecha de variación de la baja es igual o superior a la del alta. Este tipo de relaciones, respetando unas determinadas reglas de coherencia, están directamente relacionados con los "episodios". Por el contrario, las "relaciones sincrónicas" son pares de variaciones (baja, alta) producidas en el mismo tiempo pero con diferentes estados. Están directamente relacionadas con las "transiciones" entre estados y relacionadas con el suceso o evento que ha producido la transición (por ejemplo una variación residencial).

- Asociado a las transiciones está el concepto de "episodio", que se define como un periodo con duración determinada, delimitado entre un par de sucesos "alta/baja" que ocurren en distinto tiempo (fecha de variación del alta igual o menor que la de la baja), pero en un mismo estado (por ejemplo en la misma residencia). Un episodio, conceptualmente, es un periodo durante el cual un individuo dado no ha modificado su permanencia a un estado concreto. Los episodios se definen en base a un espacio de estados concreto y este puede ser conceptualizado de diferente manera en base a un objetivo concreto, por lo tanto es posible definir distinto tipo de episodios con un sistema de información concreto.
- Si entre los diversos apuntes que componen una secuencia registral no hay inconsistencias, más concretamente si son capaces de superar un conjunto bien definido de "reglas de consistencia", hablaremos de "secuencia registra coherente" o "itinerario vital", y éste es el material básico de posteriores explotaciones estadísticas.

Antecedentes: en el tiempo / en otros ámbitos

La óptica longitudinal introduce la historia de los individuos que componen una generación. Sin embargo, por cuestiones prácticas, este tipo de análisis se realiza sólo sobre un componente específico (la mortalidad, la fecundidad, las migraciones...), lo que implícitamente supone la asunción de independencia entre comportamientos y homogeneidad de los efectivos generacionales. Este hecho limita en parte su potencia. Si bien en el análisis longitudinal la fecundidad actual depende de la del pasado, la asunción de independencia implica que otros componentes del pasado, tales como la experiencia de mortalidad o las migraciones de la cohorte, no influirán sobre la fecundidad. Por ejemplo, el efecto que el déficit de varones provocado por una guerra tendría sobre la fecundidad, no podría ser analizado fácilmente con esta metodología.

En respuesta a ésta y otras limitaciones del análisis clásico, en los años 80 surgió una metodología a la que los anglosajones denominan como "event history analysis" y los francófonos como "analyse biographique"⁵. La aportación de esta nueva perspectiva ha sido la introducción explícita del individuo, lo que en ciencias sociales se ha denominado "individualismo metodológico". Esta perspectiva asume que los comportamientos sociales son agregados de decisiones individuales y, por lo tanto, están básicamente determinados por la experiencia biográfica de cada individuo. En este paradigma metodológico se pretende analizar las sucesiones e interrelaciones de los sucesos (eventos) demográficos. La información para este tipo de análisis precisa de la existencia de registros longitudinales que contengan los tipos y los tiempos en que un individuo concreto padece estos sucesos.

⁵Courgeau D., Lelièvre E., Analyse démographique des biographies, Paris, Éditions de l'INED, 1989

En estos estudios el interés se desplaza hacia un conjunto de trayectorias individuales dentro de un gran número de estados. El objeto de análisis no son tanto los "sucesos" o "movimientos" sino las biografías individuales. Se busca averiguar cómo determinadas experiencias influyen sobre la vida posterior de los individuos. Y ya que éstos siguen trayectorias vitales complejas, se precisa disponer de información mucho más detallada sobre los sucesos relevantes que han afectado a un individuo concreto.

Esta actividad ha heredado parte de los contenidos de la actividad estadística "Registro de Población de Andalucía" (RPA) que fue incluida por primera vez en Plan Estadístico 2003-2006, donde ya se hacía referencia a la investigación estadística de los "itinerarios vitales". Sin embargo, durante la primera etapa de desarrollo de esta actividad se priorizaron los aspectos transversales y no es hasta el siguiente Plan Estadístico 2007-2012, concretamente en el Programa Estadístico 2008 cuando se incluye la actividad y comienzan a desarrollarse los aspectos longitudinales de "investigación de itinerarios" de la actividad RPA. Desde entonces ha estado incluida ininterrumpidamente en todos los programas anuales ofreciendo sus primeros resultados en julio del 2014 con el producto "Estadísticas Longitudinales de Supervivencia y Longevidad en Andalucía, 2002-2010".

Esta actividad estadística tuvo en principio su modelo en el "Estudio Demográfico Longitudinal de INE" que fue incluido en la planificación estadística de 1999, pero fue dado de baja en el año 2012, habiendo dejado su huella en actividades tan relevantes como el Censo de 2011, donde se aplicó por primera vez una estrategia basada en el aprovechamiento de fuentes administrativas.

Hay pocas experiencias en España de reconstrucción de trayectorias vitales integrando fuentes estadísticas y administrativas. El "Registro estadístico de población de Euskadi" ha sido una de las actividades estadísticas que al integrar información de registros estadísticos y administrativos a nivel de individuo entronca con esta actividad estadística. No obstante, esta actividad estadística hasta ahora ha tenido mayoritariamente una orientación transversal.

En la actualidad, este proyecto es pionero en España en la puesta en producción de información estadística longitudinal, habiendo sido requeridos en algunas ocasiones para suministrar asesoría técnica a otras comunidades autónomas que están interesadas en la inclusión de actividades estadísticas similares en sus planes estadísticos.

Justificación y utilidad

Esta actividad es complementaria a la actividad 01.01.01 "Base de Datos Longitudinal de Población de Andalucía (BDLPA)", desarrollando las tareas de producción y difusión estadística que la actividad anterior no incluye. Por ello, comparte con la BDLPA las justificaciones y utilidades. En concreto la producción de estadística de itinerarios vitales permite realizar análisis e investigaciones estadísticas similares a las obtenidas por otras operaciones estadísticas mucho más complejas y costosas, como las muestras tipo panel o grandes encuestas retrospectivas, con un coste mucho menor y un nivel de detalle y precisión mucho mayor.

La investigación intersectorial de las relaciones y dependencias entre tipos de sucesos y determinadas parcelas y características personales son factibles con esta fuente. Por ejemplo, el impacto y la interrelación entre itinerarios educativos y laborales, las relaciones entre ciclo familiar y migraciones o el impacto que sobre la salud y la longevidad pueden tener las historias y las características del hogar de convivencia de los individuos.

En general, para la investigación causal sobre los determinantes de procesos y comportamientos sociales, las fuentes longitudinales como la BDLPA, son más adecuadas que las fuentes transversales clásicas.

Con esta actividad se cubren los siguientes objetivos generales del Plan Estadístico y Cartográfico de Andalucía 2013-2020:

- Producir la información estadística y cartográfica de calidad requerida para la ejecución y seguimiento de las políticas europeas, nacionales y autonómicas en el ámbito de competencia de la Junta de Andalucía, contribuyendo a la planificación en múltiples materias.
- Dotar a la ciudadanía de la información suficiente y objetiva que permita la evaluación de las políticas ejecutadas por la Junta de Andalucía y sus entes instrumentales.

2. OBJETIVOS

Objetivo general

Conocer las interrelaciones y dependencias que se producen entre las dimensiones del desarrollo humano: cualificación, biografía laboral, trayectoria familiar, condiciones de salud, etc

Objetivos específicos

1. Construir y depurar las tablas de episodios residenciales, a nivel de vivienda, para conseguir las trayectorias espacio-temporales del cambio residencial en Andalucía.
2. Construir a partir de las trayectorias residenciales individuales las tablas de relaciones de convivencia de la población andaluza, lo que permitirá describir en detalle las dinámicas de formación, crecimiento y disolución de los hogares y familias en Andalucía desde el comienzo del siglo. El cuadro necesario para investigar las relaciones de cohabitación y parentesco que se den en el interior de los hogares andaluces se deberá completar con la reconstrucción de las redes de parentesco: ascendencia, fraternidad, conyugalidad a partir de la información contenida en el MNP y en los cuadernillos censales.
3. Asignar las unidades de convivencia (hogares), estudiadas en el apartado anterior, a las unidades físicas geo-referenciadas (viviendas), de tal manera que se permita el estudio longitudinal, no solo de los personas y familias, sino también de las denominaciones y demarcaciones administrativas que cubre un territorio.
4. Definir tablas de episodios en otros espacios de estados distintos del residencial, conforme se vaya incorporando información suficiente proveniente de ámbitos específicos, como las relativas a trayectorias laborales, formativas o sanitarias.
5. Diseñar los productos de difusión adecuados para hacer accesible la información estadística generada en esta actividad a la administración, investigadores y público en general. Entre otros posibles resultados, se tiene previsto la difusión de varios productos centrados en el análisis longitudinal de los temas clásicos de la literatura demográfica:
 - a. La longevidad de la población andaluza.
 - b. Trayectorias familiares: formación y ampliación de familias en Andalucía.
 - c. Migraciones y cambio residencial

Ejes transversales

- Cohesión: La actividad es pertinente al eje de cohesión porque estudia variables que contribuyen a corregir desigualdades económicas, sociales o territoriales.

- Sostenibilidad: Para esta actividad no es aplicable el eje transversal de sostenibilidad.
- Innovación: Para esta actividad no es aplicable el eje transversal de innovación.
- Género: Esta actividad es pertinente al eje transversal de género ya que estudia variables relativas a personas.
- Evaluación de la eficacia de las políticas públicas: La actividad es pertinente a este eje en cuanto produce información de calidad requerida para la ejecución y el seguimiento de políticas europeas, nacionales y autonómicas en el ámbito de competencia de la Junta de Andalucía contribuyendo a la planificación en múltiples materias.

3. METODOLOGÍA

Ámbito de estudio

Población objeto de estudio: Población que durante algún periodo posterior al 1-1-1998 ha fijado su residencia habitual en Andalucía.

Unidad investigada: La unidad de investigación son las personas residentes en Andalucía y los episodios y eventos obtenidos a partir de la depuración de sus apuntes registrales almacenados en la BDLPA.

Los "itinerarios vitales" son, básicamente, una secuencia coherente de "episodios" y "transiciones" que quedan definidos por el "espacio de estados" y los sucesos relevantes definidos para un objetivo de investigación concreto.

Periodicidad en la recogida: Continua

Desagregación territorial máxima alcanzada: La información relativa a la residencia del individuo se recoge por medio de descriptores postales con el máximo nivel de detalle a nivel de vivienda. En cuanto a la información que se difunde se realiza, siempre que se pueda, a nivel inframunicipal.

Variables

Las principales variables de estudio para esta actividad, depende en cada momento del tema de estudio que se aborde, analizando éste según:

- Características demográficas básicas: sexo, fecha, lugar de nacimiento, lugar de residencia, nivel de estudios, etc..
- Fechas de ocurrencia u observación de los hechos relevantes, así como información auxiliar relativa a las características propias del evento.
- Características del hogar de convivencia
- Variables socio-económicas, como por ejemplo la relación con la actividad, recogidas en distintas fuentes estadísticas y administrativas, según el caso.

Recogida de información

La información de base para la realización de la actividad procede de la explotación de la Base de Datos Longitudinal de Andalucía (BDLPA), actividad estadística responsabilidad del Instituto de Estadística y Cartografía de Andalucía. Dicha actividad se nutre de las inscripciones en los padrones municipales de Andalucía, los boletines estadísticos del Movimiento Natural de Población y los Censos de Población y Viviendas. La integración de esta información permite la

incorporación de variables recopiladas en una fuente a las de otra, que o bien no disponía de dichas variables o bien las tenía recogidas con un menor nivel de detalle. Este proceso de integración y reutilización de información es la que permite el desarrollo de esta actividad, cuyos productos de difusión amplían y complementan la producción estadística clásica, con un coste menor que el que supondría diseñar nuevos circuitos de recogida de información.

Restricciones y alternativas

El desarrollo de esta actividad se basa en la BDLPA, y ésta a su vez, depende en gran parte de la información suministrada por el Instituto Nacional de Estadística (INE) en base al convenio de colaboración estadística INE-IECA, por lo que es necesario el mantenimiento de los circuitos de intercambio de la información padronal entre ambos organismos.

El desarrollo y la ampliación de la actividad sobre nuevos ámbitos dependerá de que se establezcan nuevos acuerdos y convenios de colaboración para la explotación estadística de los registros administrativos de base poblacional.

Pueden existir valores fuera de rango o diferentes codificaciones de una misma variable en función de la fuente de procedencia que generen conflictos en la integración de la información. Ésto se corrige, aunque solo parcialmente, usando fuentes ajenas a la información padronal. Para el resto, se opta por hacer imputaciones en función de los valores válidos de personas de características y entornos similares.

4. PLAN DE ANÁLISIS, TABULACIÓN Y COMPILACIÓN

Preparación y tratamiento base de la información

El ciclo de producción puede esquematizarse a grandes rasgos en una serie de procesos y tratamientos que se describen a continuación.

El procesado de la información comienza con la carga de los ficheros con la información base, constituida por un conjunto de registros que contiene la información a nivel individual e incluyen algún tipo de clave que permita su identificación. A una unidad individual de información nominalizada se denomina "apunte registral". Los ficheros fuente son cargados en tablas adecuadas a su contenido sobre la BDLPA. Durante el proceso de carga se consigue de manera directa la identificación (asignación del IDP) en todos aquellos registros donde se produce una fusión directa, basándonos en el conjunto de claves definidas como identificadoras personales. Si hay una concordancia perfecta de los campos identificadores se enlazarán con alguna versión de persona pre-existente en el sistema. Dado que la "identificación por fusión directa" no es exhaustiva y con algunos ficheros se obtienen resultados insuficientes es preciso implantar, en un segundo paso, una identificación secundaria donde se aplican técnicas de enlace más sofisticadas: medidas de proximidad y enlaces probabilísticos. En esta fase se consiguen unos niveles de enlace aceptables, en todo caso superiores al 95%.

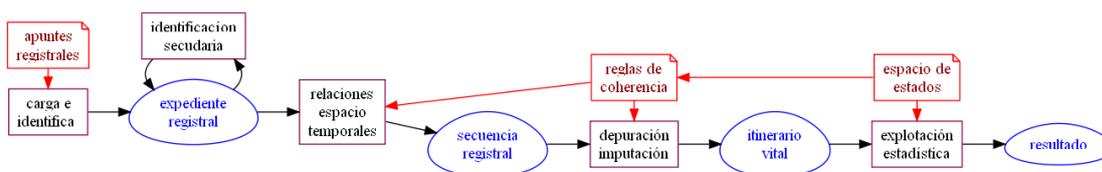


Ilustración 1. Esquema del proceso de producción de los itinerarios desde las fuentes primarias.

La identificación (asignación de un identificador de persona IDP) nos permite agrupar toda la información referida a un mismo individuo en un agregado lógico que denominamos "expediente registral". Trabajando con los expedientes registrales es posible establecer una serie de relaciones de continuidad entre los apuntes registrales. Para el caso de las variaciones padronales podemos clasificar estas relaciones como diacrónicas y sincrónicas. Es preciso que exista un conjunto sencillo de reglas de coherencia que dependen del objeto de la investigación y no exclusivamente de la calidad de los datos. Aplicando estas reglas a una secuencia de variaciones determinada permite obtener las relaciones de continuidad y precedencia adecuada para reconstruir los itinerarios individuales.

Las relaciones diacrónicas y sincrónicas entre los pares de variaciones determinan una ordenación espacio-temporal de los apuntes de un expediente, a la que denominamos "secuencia registral". Esta secuencia completa puede ser coherente o no en base a las reglas de coherencia

definidas para el evento estudiado. Las secuencias no coherentes deberán ser depuradas y, en su caso, imputadas hasta obtener su coherencia. Una secuencia registral coherente se convierte en un "itinerario vital" que puede ser explotado estadísticamente. Si no es posible obtener una secuencia consistente se retira total o parcialmente para el análisis dicho expediente. Estos expedientes retirados representan para el sistema un problema similar al de las no localizaciones o negativas a responder en las operaciones muestrales, por la posibilidad de introducir sesgos en el análisis.

El conjunto de reglas de coherencia a aplicar a una secuencia registral es dependiente del espacio de estados que precise una investigación estadística concreta. El espacio de estados define las diferentes situaciones en las que se puede encontrar un individuo concreto en un instante cualquiera de su itinerario vital. El espacio de estados se organiza en base a ámbitos o dimensiones independientes, por ejemplo las trayectorias educativas y las residenciales emplean su propio subconjunto de estados, y el producto cartesiano de los estados educativos y residenciales serían el espacio de estados de una investigación sobre trayectorias educativas-residenciales. Las reglas de coherencia podrían en último punto ser representadas por una matriz de transiciones entre los diferentes estados que componen un sistema, donde solo estarían activadas las casillas con los saltos entre estados permitidos por las reglas. Esta dependencia implica que la producción de los itinerarios (los objetos de los que se parte para la explotación estadística) es dependiente del objetivo de la investigación. Es decir, una explotación que precise gran precisión necesitará un espacio de estados muy detallado, lo que implica un conjunto mucho más extenso y estricto de reglas de coherencia. Por ejemplo, una secuencia registral puede ser coherente para un objetivo dado y sin embargo incoherente para otro, como sería el caso de un fichero de itinerarios construido para estudiar la movilidad, residencia a nivel municipal si se pretendieran estudiar los cambios de domicilio a nivel de vivienda. Para esta profundización en el nivel de detalle, habría que generar un nuevo fichero de itinerarios basado en nuevas reglas de coherencia aplicadas a nivel de vivienda y no al de municipio y, por tanto, mucho más estrictas.

Sistemas de codificación y nomenclaturas utilizados

Los sistemas de codificación y nomenclaturas utilizados en esta actividad estadística son heredados de la actividad estadística 01.01.01-Base de Datos Longitudinal de Población de Andalucía, tales como las explotaciones estadísticas de los padrones municipales, Movimiento Natural del Población o los Censos de Población. Por lo tanto la codificación empleada coincide, en general, con los estándares utilizados habitualmente por el sistema estadístico en España y en particular por las codificaciones utilizadas por el INE.

La lista de estándares y codificaciones le INE pueden verse en el siguiente enlace:

<http://www.ine.es/ss/Satellite?>

[L=es_ES&c=Page&cid=1254735839296&p=1254735839296&pagename=MetodologiaYEstandares%2FINELayout#](http://www.ine.es/ss/Satellite?L=es_ES&c=Page&cid=1254735839296&p=1254735839296&pagename=MetodologiaYEstandares%2FINELayout#)

Criterios y variables de clasificación

La clasificación de la posible tabulación se ha diseñado delimitando cada una de las tablas en base a tres criterios:

1. Criterio temático:

Determina el objetivo y secundariamente las variables (aunque no su nivel de detalle) a incluir en las tablas. Las características demográficas sexo, edad y estado civil, características transversales que se incluirán prácticamente en todas las tablas. Además de ellas se consideran algunas variables correspondientes a otras categorías temáticas: características socio-económicas personales, estructuras de convivencia, calidad del entorno, características relativas a la actividad económica, características relativas a la movilidad.

2. Criterio territorial:

Se han definido agrupaciones con diferente nivel de detalle territorial: Andalucía y sus provincias y municipios, unidades territoriales del Plan de Ordenación Territorial de Andalucía (POTA), secciones censales y celdas regulares.

3. Colectivo sub-poblacional:

Determina el subconjunto de la población a la que hace referencia una tabulación concreta. Dado que determinadas variables sólo están disponibles en un determinado colectivo, la información de cada una de ellas está referida a la población filtrada para las categorías para las que se dispone de información.

Los tres criterios de filtrado usados para generar una tabla concreta son los siguientes:

- Total de población censada en viviendas familiares y colectivas: Dado que sobre los colectivos se recogió un cuestionario muy simplificado, los conjuntos temáticos que se podrán aplicar a este colectivo serán necesariamente limitados. En este colectivo solo se dispone de variables demográficas básicas y nacionalidad.
- Población en viviendas familiares: Sobre esta población se dispone del grueso de las variables contenidas en el cuestionario censal.
- Población que había trabajado la semana anterior a la entrevista censal: Una serie de preguntas relativas a la actividad económica de la población solo se aplicó sobre el colectivo que había trabajado la semana anterior, por lo que solo es posible investigar estas características en este grupo.

Análisis realizado y metodología

La explotación estadística de este tipo de fuentes se puede plantear bajo distintas alternativas. En un acercamiento directamente relacionado con el análisis de supervivencia por ejemplo, las unidades de investigación son los individuos, categorizados por un conjunto de covariables invariables en el tiempo y una sola variable dependiente: el tiempo hasta la ocurrencia del evento de interés o censura. La censura se produce por cierre de estudio o por pérdida o salida del espacio de observación, por ejemplo una emigración fuera de Andalucía.

Este esquema tiene la ventaja de su relativa sencillez y la potencia y robustez de sus resultados. Por otra parte tiene también importantes limitaciones, siendo la más evidente la imposibilidad de trabajar con más de una escala de tiempo y de no disponer de un método sencillo para incluir variables tiempo dependientes, que cambian a lo largo de la historia individual. Si se precisa trabajar con variables tiempo-dependientes o usar varias escalas de tiempo hay que optar por subdividir los seguimientos individuales en subepisodios definidos por los distintos valores que toman las variables tiempo dependientes dentro de cada episodio genérico. En estos casos, las unidades de investigación no son ya los individuos participantes, sino los distintos episodios de seguimiento que cada uno de los participantes generan a lo largo de su historia individual.

Así pues, los episodios son duraciones durante las cuales el individuo no modifica su estado. Por ejemplo, si se define como relevante para una determinada investigación la edad en años cumplidos, un itinerario que fuera seguido durante 9 años a partir de su aniversario generaría 9 episodios distintos de un año de duración, con una edad cumplida diferente y con distintas propiedades de ocurrencia en cada uno de los subepisodios.

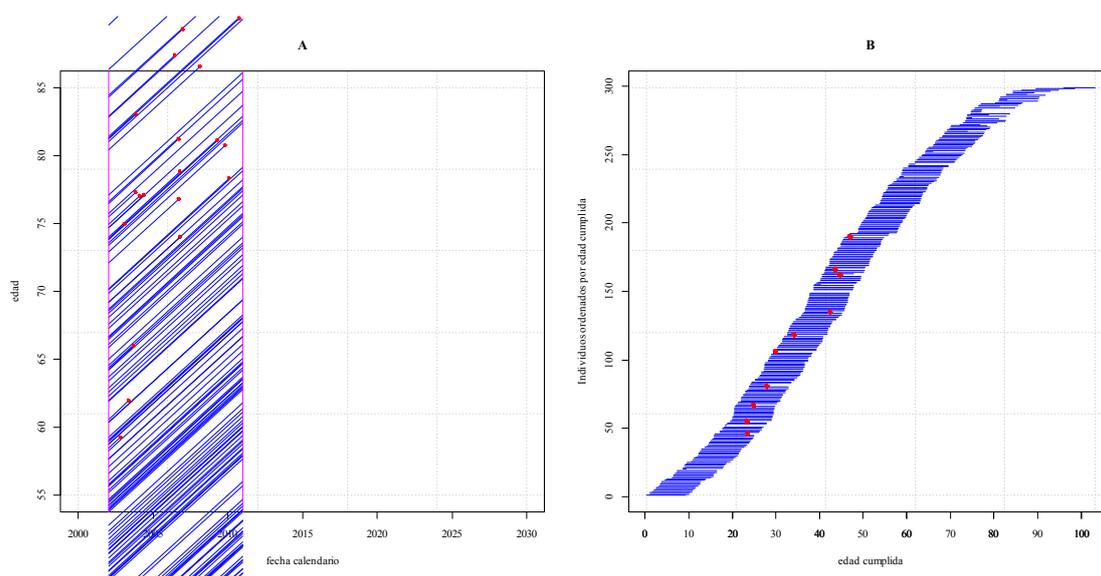
Los episodios se definen en función de un espacio de estados explícito y en base a los objetivos de la investigación en curso y de la calidad y nivel de detalle de la información de la fuente que se utilice.

Por ejemplo, en una de las explotaciones estadísticas diseñadas dentro de esta actividad, sobre la investigación longitudinal de la longevidad en la cohorte censal de 2001, se ha seleccionado a la población de nacionalidad española que residía en Andalucía durante la operación censal de 2001 y se le ha seguido durante varios años, estudiándose sus trayectorias en un sencillo espacio de estados que queda definido en torno a tres dimensiones: (1) el estado vital de los participantes (vivo o fallecido), (2) su lugar de residencia (Andalucía a distinto detalle territorial o fuera de Andalucía) y (3) dos escalas de tiempo (el periodo y la edad). Estas son las únicas variables durante el periodo de seguimiento definido para este estudio que pueden ser tratadas como variables tiempo dependientes⁶.

⁶Evidentemente es posible realizar diseños mucho más complejos, introduciendo más variables tiempo dependientes. Conforme los sistemas de información se vayan consolidando será posible incluir nuevas dimensiones, tiempo dependientes, como el estado civil, la actividad económica, el nivel educativo....

En el gráfico 1.A se ha representado una pequeña muestra de los episodios de este estudio sobre un diagrama de Lexis. Se aprecia una serie de episodios observados durante el periodo de seguimiento del estudio. Todas las líneas de vida comienzan el 1-1-2002, fecha de inicio del seguimiento y la mayoría acaban en la fecha de cierre del estudio (en este ejemplo, 31-12-2010). Unos pocos episodios finalizan previamente a la fecha de estudio, bien por fallecimiento, evento marcado como un punto rojo en el diagrama, bien por emigración hacia fuera de Andalucía, sin una marca específica.

Gráfico 1. Representación una muestra de 300 seguimientos de la cohorte censal sobre un diagrama de Lexis y una representación alternativa con una sola escala de tiempo



Si del gráfico 1.A suprimimos la escala de fecha calendario y nos quedamos solo con la escala de edad cumplida, y representamos los episodios ordenados por edad el 1-1-2002, tenemos el esquema de seguimientos del gráfico 1.B muy útil para el análisis de supervivencia. Los episodios vitales se encuentran censurados por la derecha, bien con censuras tipo I no aleatoria por fin de estudio (tras 9 años de seguimiento) o bien con censura tipo II o aleatoria por emigración fuera de Andalucía. Los episodios están truncados por la izquierda a la edad al comienzo del estudio el 1-1-2002.

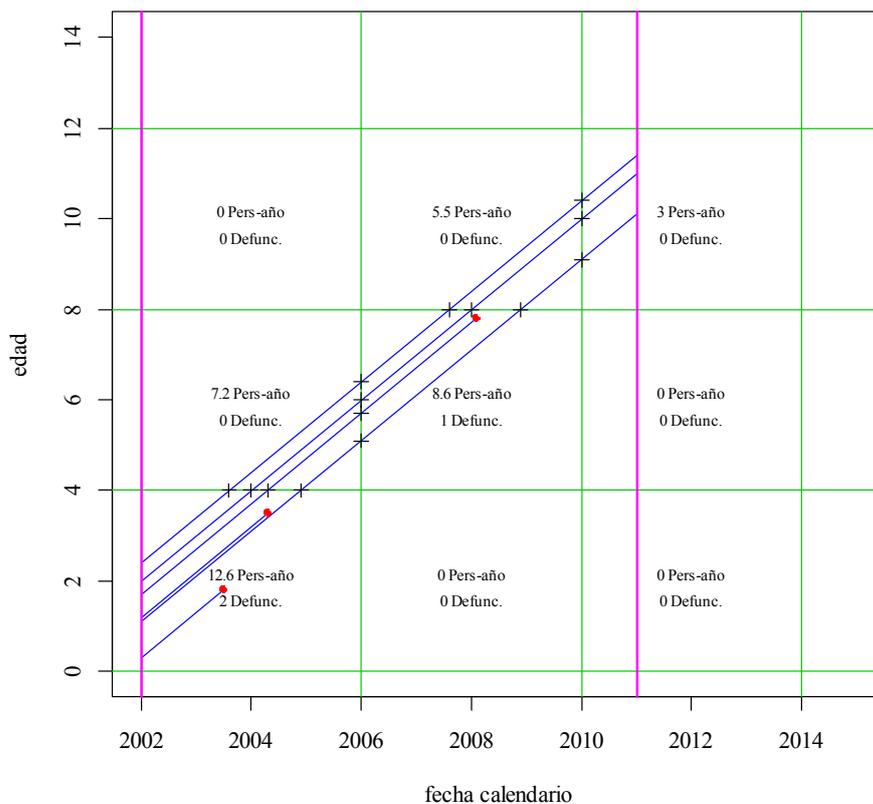
Para facilitar el análisis de esta información es habitual tener que redimensionar en unidades discretas las escalas de tiempo, con gran importancia predictora, como la edad y periodo. Si bien es posible trabajar con ellas como variables continuas si se suministra la información en forma de microdatos, si se precisa difundir tablas estadísticas agregadas de eventos o de tiempo de exposición a riesgo es necesario transformarlas en unidades discretas, tales como años cumplidos o periodos anuales.

Para discretizar las escalas de tiempo se definen una serie de fronteras temporales en cada una de las escalas a partir de las cuales se establece su partición. A nivel práctico, este paso consiste

en transformar las variables continuas (por ejemplo, la edad) en funciones escalonadas, con la implicación de asumir que el efecto de la variable temporal es constante en todo el intervalo que se ha discretizado. Por ejemplo, una persona con 40 años de edad cumplida se considera con igual propensión a fallecer el día que cumple esta edad, como 364 días después. Estas fronteras, definidas sobre las escalas edad y periodo, dibujan una malla sobre el diagrama de Lexis que permite dividir los episodios en subepisodios con edad y periodos ya invariables (al haber sido transformados a una escala discreta).

En el diagrama de Lexis adjunto (gráfico 2) se han representado seis episodios vitales de un estudio longitudinal: tres de ellos han sufrido un evento defunción representado con un punto rojo y los otros tres han sobrevivido hasta el fin de periodo de seguimiento. Sobre este conjunto de episodios se muestra la malla (edad, periodo) utilizada para discretizar las escalas temporales, con objeto de derivar una tabla de eventos y personas-año. Para simplificar la representación se han elegido grupos de edad y periodo cuatrienales.

Gráfico 2 Recorte de episodios al discretizar las escalas de tiempo edad y periodo



Las líneas verticales violeta marcan el comienzo y el fin del periodo de seguimiento y la rejilla verde representa los tiempos frontera que se han usado en cada escala de tiempo para

discretizarla. Las cruces negras sobre las líneas de vida azules representan las intersecciones de las líneas de vida con las fronteras de las categorías temporales discretas (grupos de edad y de periodo). En este sentido las cruces son transiciones entre los estados definidos por la escala de tiempo discreta (por ejemplo grupo de edad). Los episodios vitales originales quedan divididos por estas marcas en subepisodios durante los cuales se considera que los efectos de la edad y del periodo no van a cambiar significativamente (a nivel práctico se toman como constantes a lo largo de toda la duración del subepisodio).

En el centro de los cuadrados delimitados por esta rejilla se han colocado las sumas de las duraciones de los episodios (personas-año) y el número de eventos producidos en los intervalos temporales discretos definidos en cada una de las celdillas de la malla.

Es posible diseñar mallas con más de dos escalas de tiempo, introduciendo por ejemplo la dimensión generación de nacimiento, o mallas más sencillas por ejemplo definiendo solo grupos de edad (y no haciendo relevante el efecto periodo).

Control de calidad

Además de los controles de calidad que sistemáticamente se realizan en las actividades estadísticas que genera la BDLPA y la depuración que se realiza dentro de la propia actividad BDLPA, la construcción de los "itinerarios vitales" conlleva la necesidad de sus propios controles de calidad y consistencia. En particular, es necesario prestar atención a los aspectos relacionados con la secuencia temporal y a las reglas de precedencia que existen entre eventos: por ejemplo nadie puede morir antes de nacer, o tampoco salir de Andalucía si previamente no ha nacido aquí o ha llegado por inmigración.

El conjunto de reglas de consistencia definen la mayor o menor complejidad de los procesos de depuración necesarios para generar secuencias registrales coherentes. Este conjunto de reglas variará en complejidad y rigidez en función de los objetivos de investigación y la extensión del espacio de estados que se definan. Por ejemplo, una secuencia del subsistema de variaciones padronales puede ser coherente pero tornarse incoherente cuando se le agregan los eventos provenientes del MNP. Igualmente una secuencia puede ser coherente para un determinado nivel de detalle (por ejemplo desagregación municipal) pero ser incoherente a una mayor profundidad de detalle (digamos a nivel de vivienda).

Se visualizan las variaciones padronales como transiciones, dentro de un espacio de estados discreto. Por ejemplo, una defunción es una transición del estado vivo a muerto, un cambio de domicilio es un cambio de estado de residir en la vivienda 1 a residir en la vivienda 2. En los sistemas de registro de variaciones, como el padronal, una transición es la combinación de dos sucesos: uno de baja en el estado precedente y otro de alta subsecuente que ocurren

sincrónicamente en la misma fecha de variación. Ocasionalmente es posible que, en función de las fronteras del sistema, no sea posible observar los dos pares de sucesos. Por ejemplo, en Andalucía una variación residencial de Madrid a Sevilla se observará exclusivamente como un evento de alta; al contrario, una variación residencial de Sevilla a Madrid aparecerá en el sistema con solo la notificación de la baja.

La consistencia temporal de una sucesión de sucesos de altas y bajas es visualizable si superponemos éstos en un doble eje temporal, uno para las altas y otro para las bajas. Una sucesión de altas y bajas es temporalmente consistente si al dibujar los segmentos horizontales con comienzo en la diagonal a la fecha de alta y terminación en la vertical de la fecha de baja, el resultado es un gráfico perfectamente escalonado. La visualización de la consistencia espacio-temporal requeriría incluir una tercera dimensión donde se representara el espacio físico. La tercera dimensión representaría una determinada localización, un municipio, un portal o una vivienda. En cada caso el nivel de rigor exigido a la consistencia sería mayor. Una secuencia registral consistente se visualiza con un itinerario formado por una sucesión continua de segmentos dirigidos (episodios y transiciones).

Es evidente que no todas las secuencias registrales son directamente coherentes. La transformación de una secuencia registral incoherente en otra coherente se realiza mediante técnicas de depuración e imputación de datos.

Garantía de secreto estadístico

Con el fin de salvaguardar el Secreto Estadístico se han considerado distintos niveles de detalle en las variables analizadas. Por tanto, en tabulaciones con varias variables, no todas ofrecen el máximo nivel de detalle de forma que se dificulta la identificación directa o indirecta de individuos. En cuanto a los microdatos, se pone a disposición una muestra del 10% de la población analizada, suficiente para la obtención de resultados significativos pero no para la identificación de individuos.

Normas y estándares de aplicación

Entre las listas estándares aprobadas por el INE para clasificaciones se utilizan las de municipios, provincias, comunidades, países y otros territorios.

Estos estándares pueden consultarse en el siguiente enlace:

http://www.ine.es/ss/Satellite?L=es_ES&c=Page&cid=1254735839296&p=1254735839296&pagename=MetodologiaYEstandares%2FINELayout#

Posibles resultados (Plan de tabulación, mapas, fotos,..)

El plan de tabulación depende del fenómeno que se estudie en cada momento. De forma general, a partir de las tablas extraídas de la BDLPA y tras su depuración estadística para suprimir inconsistencias, los resultados que se ofrecen son:

1. Un conjunto de tablas estadísticas sobre el conjunto de la población con seguimiento efectivo.
2. Una serie de ficheros muestrales de microdatos, preparados para suministrar a usuarios especializados toda la información necesaria que pudieran necesitar en sus respectivos proyectos que no haya sido satisfecha por las tablas estadísticas del plan de tabulación previo.
3. La estimación de algún indicador para área pequeña y su representación sobre mapas del conjunto de Andalucía.

Niveles de desagregación territorial alcanzados

Inframunicipal

Unidades de medida utilizadas

- Personas-año: tiempos de exposición a riesgo de sufrir el evento considerado experimentados por la población objeto de análisis.
- Eventos: número de observaciones ocurridas del fenómeno estudiado.

Indicadores que se van a construir

Tasas brutas: cociente entre el número de casos observados y las personas-tiempo

Otros indicadores específicos, dependientes del fenómeno a estudiar: por ejemplo, razones de mortalidad estandarizadas suavizadas.

5. PLAN DE DIFUSIÓN

- **Denominación de la publicación**

Estadísticas longitudinales de supervivencia y longevidad en Andalucía, 2002-2016

Organismo difusor

Instituto de Estadística y Cartografía de Andalucía

Tipo de producto de difusión

Tablas, microdatos, gráficos, mapas

Medio de difusión

Internet

<http://www.juntadeandalucia.es/institutodeestadisticaycartografia/longevidad/index.htm>

Periodicidad de la publicación

No periódica

Periodicidad de difusión de resultados

No periódica

- **Denominación de la publicación**

Estadísticas longitudinales de biografías reproductivas en Andalucía, 2002-2013

Organismo difusor

Instituto de Estadística y Cartografía de Andalucía

Tipo de producto de difusión

Tablas, microdatos, gráficos, mapas

Medio de difusión

Internet

<http://www.juntadeandalucia.es/institutodeestadisticaycartografia/fecundidad/index.htm>

Periodicidad de la publicación

No periódica

Periodicidad de difusión de resultados

No periódica

Perfil de los receptores

Investigadores sociales y económicos, así como gestores y evaluadores de políticas públicas, en general con formación superior en alguna disciplina relacionada con los estudios de población.

Además se mantiene el compromiso, como en otras actividades estadísticas, de confeccionar parte de los productos de difusión pensando en el interés de cualquier tipo de usuario, usando tecnología de difusión accesible para un personal no especializado en análisis de datos.

Disponibilidad de bases de datos

Esta actividad permite elaborar resultados a medida en el Banco de Datos estadístico de Andalucía (BADEA)

Posibilidad de peticiones específicas a demanda o disponibilidad de microdatos

Existe disponibilidad de microdatos en la página donde se difunde los resultados.

6. CRONOGRAMA

Cada año se abordará algún estudio, pero no se puede definir un cronograma genérico ya que dependerá del evento o fenómeno estudiado.