



Máster en Técnicas Estadísticas

# Dinámica del consumo de tabaco en Galicia: 2006-2020. Evaluación en diferentes escenarios

Paula Raña Míguez

Julio 2012



Este trabajo se presenta para su evaluación como Trabajo Fin de Máster, en la modalidad de prácticas en empresas, de la alumna Paula Raña Míguez.

Ha sido realizado bajo la dirección de Rosa María Crujeiras Casais, María Isolina Santiago Pérez y Mónica Pérez Ríos, que autorizan la entrega del mismo.

Santiago, a 29 de junio de 2012.

Rosa María Crujeiras Casais

María Isolina Santiago Pérez

Mónica Pérez Ríos





Dna. M<sup>a</sup> Isolina Santiago Pérez e Dna. Mónica Pérez Ríos, nomeadas pola Dirección Xeral de Innovación e Xestión da Saúde Pública da Consellería de Sanidade para exercer as funcións de titoras de Paula Raña Míguez no Convenio de Cooperación Educativa entre a Consellería de Sanidade e as Universidades de Santiago de Compostela, A Coruña e Vigo para a realización do seu traballo Fin de Máster, certifican que Paula Raña Míguez realizou satisfactoriamente o traballo titulado "Dinámica do consumo de tabaco en Galicia no período 2006-2020. Avaliación en diferentes escenarios", e que cumpriu os obxectivos establecidos no Plan de Traballo do Convenio. O traballo desenvolveuse no servizo de Epidemioloxía da Dirección Xeral de Innovación e Xestión da Saúde Pública entre febreiro e abril de 2012.

M<sup>a</sup> Isolina Santiago Pérez



Mónica Pérez Ríos

Santiago de Compostela, 28 de maio de 2012



# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>I Técnicas Estadísticas</b>	<b>5</b>
<b>2. Introducción a las Cadenas de Markov</b>	<b>7</b>
2.1. Caracterización de una CMTD . . . . .	8
2.1.1. Espacio de estados . . . . .	8
2.1.2. Matriz de transición . . . . .	8
2.1.3. Distribución inicial . . . . .	9
2.1.4. Distribución de la CMTD en la etapa $n$ -ésima . . . . .	9
2.1.5. Clasificación de estados . . . . .	11
2.1.6. Distribución estacionaria de una CMTD . . . . .	13
<b>3. Estimación en CMTD</b>	<b>15</b>
3.1. Estimación por máxima verosimilitud . . . . .	16
3.1.1. Cálculo de los estimadores mediante tablas . . . . .	17
3.1.2. Distribución asintótica de los estimadores . . . . .	17
3.1.3. Contrastes de hipótesis sobre la distribución en la etapa $n$ -ésima . . . . .	18
<b>4. Técnicas Bootstrap para CMTD</b>	<b>23</b>
4.1. Réplicas Bootstrap de una CMTD . . . . .	23
4.1.1. Intervalos de confianza Bootstrap para las probabilidades de transición . . . . .	25
4.1.2. Intervalos de confianza Bootstrap para la distribución en la etapa $n$ -ésima . . . . .	25
4.2. Suavización de la matriz de transición . . . . .	26
<b>II Material y Métodos</b>	<b>29</b>
<b>5. Fuente de datos</b>	<b>31</b>
5.1. Sistema de Información sobre Conductas de Riesgo (SICRI) . . . . .	31
5.1.1. Tamaño de la muestra . . . . .	32

## ÍNDICE GENERAL

5.1.2. Cuestionarios . . . . .	33
5.2. Fuentes de datos complementarias . . . . .	33
5.2.1. Programa-concurso Déixao e Gaña . . . . .	33
5.2.2. Modelo SimSmoke . . . . .	35
<b>6. Modelo general de simulación</b> . . . . .	<b>37</b>
6.1. Espacio de estados . . . . .	37
6.2. Determinación del estado de los individuos . . . . .	38
6.2.1. Determinación del estado actual . . . . .	38
6.2.2. Determinación del estado anterior . . . . .	38
6.3. Distribución inicial . . . . .	39
6.4. Matriz de transición . . . . .	40
6.4.1. Estimación de las probabilidades de transición . . . . .	40
6.4.2. Intervalos de confianza . . . . .	40
6.5. Distribución en la etapa $n$ -ésima . . . . .	41
6.6. Comparación con los datos observados . . . . .	41
<b>7. Modelos considerados</b> . . . . .	<b>43</b>
7.1. Modelo 1 . . . . .	43
7.1.1. Espacio de estados . . . . .	43
7.1.2. Transiciones . . . . .	44
7.1.3. Determinación del estado actual . . . . .	45
7.1.4. Determinación del estado actual . . . . .	45
7.2. Modelo 2 . . . . .	47
7.2.1. Espacio de estados . . . . .	47
7.2.2. Transiciones . . . . .	48
7.3. Modelo 3 . . . . .	49
7.3.1. Espacio de estados . . . . .	49
7.3.2. Transiciones . . . . .	49
<b>8. Modelo de predicción</b> . . . . .	<b>51</b>
8.1. Nunca fumadores . . . . .	52
8.2. Fumadores . . . . .	53
8.3. Exfumadores . . . . .	54
<b>9. Efecto de las intervenciones</b> . . . . .	<b>57</b>
<b>III Resultados</b> . . . . .	<b>61</b>
<b>10. Aplicación 2006-2011</b> . . . . .	<b>63</b>
10.1. Modelo 1. Resultados globales . . . . .	63
10.2. Modelo 1. Resultados por sexos . . . . .	68
10.3. Modelo 2. Resultados . . . . .	71
10.4. Modelo 3. Resultados . . . . .	71



## ÍNDICE GENERAL

<b>11. Proyecciones hasta 2020</b>	<b>73</b>
11.1. Predicción de prevalencias . . . . .	73
11.2. Efecto de las intervenciones . . . . .	82
<b>12. Líneas abiertas</b>	<b>93</b>
<b>A. Software</b>	<b>95</b>
<b>Bibliografía</b>	<b>97</b>

*ÍNDICE GENERAL*

# Capítulo 1

## Introducción

Actualmente el tabaquismo está considerado como el factor de riesgo susceptible de ser prevenido que más muertes causa en el mundo [1]. Se estimó que en el año 2000 el tabaco produjo casi 5 millones de muertes prematuras en el mundo, la mitad de las cuales sucedieron en los países desarrollados y la otra mitad en aquellos en vías de desarrollo, con 3,8 millones de varones fallecidos [2]. La mortalidad atribuida en los países en vías de desarrollo va en aumento y se espera que en el año 2030 más del 80 % de las muertes atribuidas al tabaco sucedan en países de ingresos medios y bajos [3].

Se estima que en el mundo fuman 1.250 millones de personas, de las que una quinta parte son mujeres. La evolución de la epidemia tabáquica es diferente en los países desarrollados y en aquellos en vías de desarrollo. El consumo de tabaco ha disminuido a lo largo de los últimos años en los países industrializados, tanto en hombres como en mujeres. Así, en países como el Reino Unido la prevalencia de consumo ha disminuido desde la década de 1960 pasando del 61 % en hombres y el 42 % en mujeres hasta el 23 % y el 21 %, respectivamente, en el año 2006. Pero a pesar de que la prevalencia de fumadores es menor, cada persona fumadora sigue consumiendo aproximadamente la misma cantidad de tabaco [3].

España es uno de los países europeos con una prevalencia más alta de consumo de tabaco. Según los datos del Eurobarómetro publicado en Mayo de 2010 [4], la prevalencia de consumo en población mayor de 14 años en 2009 es del 35 %, 6 puntos porcentuales por encima de la media europea. La última Encuesta Nacional de Salud estima la prevalencia del año 2006 en el 30 %, para la población de 16 años y más [5]. La prevalencia de consumo de tabaco es más alta entre los hombres (35 % versus 24 %), debido a la baja prevalencia entre las mujeres mayores de 54 años. La prevalencia de consumo, excepto en el grupo de 16 a 24 años, es siempre más elevada entre los hombres alcanzando diferencias de 18 puntos porcentuales en el grupo con edades comprendidas entre los 55 y los 64 años. La prevalencia de consumo alcanza valores máximos en España en el grupo de 25 a 35 años.

La edad media de inicio al consumo de tabaco en España es de 17,4 años [6] entre los fumadores diarios. Esta edad se retrasa en casi 2 años en los ocasionales. En las sociedades de consumo, en las que la adolescencia y la juventud es un estado, y en las que los jóvenes se caracterizan por presentar un cierto poder adquisitivo, el consumo de tabaco comienza en la temprana adolescencia y está ya consolidado a los 20 años. En estos países desarrollados es muy raro que la experimentación se produzca después de los 18 años y que el consumo diario se produzca después de los 20 [7].

En España, la prevalencia de consumo de tabaco está descendiendo desde la segunda parte de la década de 1970 [8] [9] (Figura 1.1). En los últimos 20 años se observa un descenso de casi 20 puntos porcentuales en la prevalencia de fumadores varones y una estabilización en la prevalencia de fumadoras (Figura 1.2).

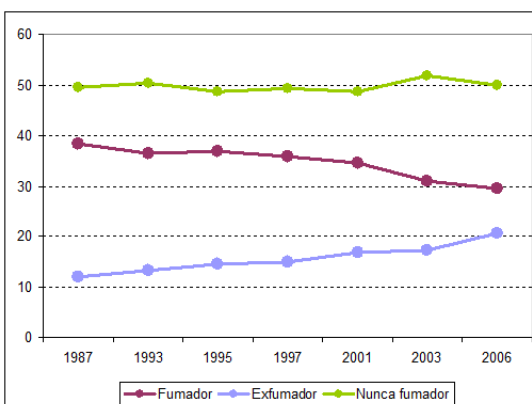


Figura 1.1: Prevalencia de consumo de tabaco en España (1987-2006).

Destaca la importante diferencia en la prevalencia de nunca fumadores en función del sexo y la tendencia decreciente que se observa en las mujeres y la creciente en los hombres; lo que refleja la continua (aunque moderada) incorporación de las mujeres al consumo de tabaco en las últimas 2 décadas y el descenso en la incorporación de los hombres. Por último la prevalencia de exfumadores siempre ha sido más elevada entre los hombres. En los últimos 20 años la prevalencia de exfumadores ha aumentado tanto en hombres como en mujeres (Figura 1.2).

En España se estima que el tabaco produce cada año aproximadamente 50.000 muertes [10].

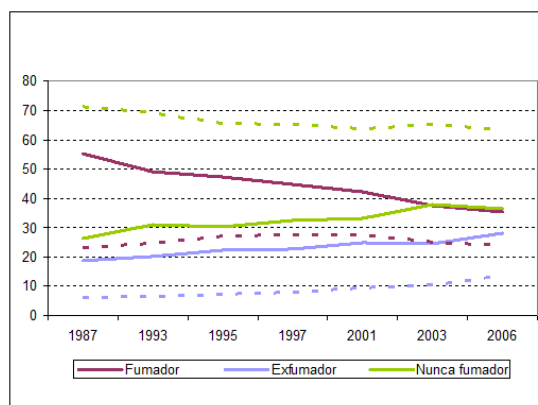


Figura 1.2: Prevalencia de consumo de tabaco en España en función del sexo (1987-2006). Las líneas sólidas reflejan la prevalencia en hombres y las punteadas en mujeres.

En Galicia, entre los años 2005 y 2010 ha descendido la prevalencia de fumadores, mayores de 16 años, de un 30 % a un 26 %. La prevalencia de nunca fumadores se mantiene estable, en torno a un 50 % a lo largo de los años considerados, por lo tanto el descenso en la prevalencia de fumadores se corresponde con un aumento en la de exfumadores (de 20 % a 24 %). Son valores similares a los obtenidos para toda España en el mismo período.

Por sexos, la diferencia en la prevalencia de fumadores entre hombres y mujeres no alcanza el 10 %. En el caso de los hombres, en el mismo período de tiempo se ha reducido la prevalencia de fumadores en 3 puntos porcentuales, pasando de 35 % a 32 %. Las mujeres mantienen una prevalencia de fumadoras más estable, con una reducción menor del 2 %, pasando de 23,2 % a 21,5 %. Al igual que en el resto de España, se ha venido observando, tanto en hombres como en mujeres, un aumento en la prevalencia de exfumadores.

A la vista del impacto que tiene el consumo de tabaco en España se han aprobado, especialmente en los últimos años, leyes que regulan el consumo. Estas tienen como objetivo potenciar el cese entre los fumadores y también la prevención, retrasando la edad de inicio. Para conseguirlo son eficaces las medidas de control del consumo que restringen la promoción y la publicidad, y aquellas que dificultan la accesibilidad y disponibilidad del tabaco, además de la prohibición de consumo en lugares públicos. En el año 2005 se aprobó la ley 28/2005 de medidas sanitarias frente al tabaquismo [11], que posteriormente fue modificada por la Ley 42/2010 [12] que estableció medidas más estrictas de control.

Conocer el impacto que estas leyes tienen en la evolución de la epidemia tabáquica de Galicia es complicado, ya que la tendencia previa de la prevalencia dificulta observar el efecto directo que pudiesen producir las leyes en el consumo. Estudios previos han mostrado como se puede valorar el impacto de las políticas de control de tabaquismo a partir de modelos basados en Cadenas de Markov ([14]-[18]).

En este trabajo se modeliza la dinámica de consumo de tabaco en Galicia mediante un modelo basado en las técnicas de Cadenas de Markov. Estas técnicas permiten hacer predicciones de la prevalencia de consumo de tabaco en la comunidad gallega. El objetivo es estudiar la evolución del consumo de tabaco en la población gallega (población de residentes en Galicia) mayor de 16 años y valorar el impacto de las leyes reguladoras del consumo en la población entre los años 2010-2020.

Parte I

**Técnicas Estadísticas**





## Capítulo 2

# Introducción a las Cadenas de Markov

En este capítulo se presenta una breve introducción a las Cadenas de Markov, repasando sus características y propiedades. Se comienza introduciendo el concepto de proceso estocástico, dentro del que se enmarcan las Cadenas de Markov como caso particular. Para una explicación más exhaustiva sobre formulación, caracterización y análisis de Cadenas de Markov véase [19] y [20].

De manera poco precisa, se puede entender como proceso estocástico a una familia de variables aleatorias definidas sobre un espacio de probabilidad. Dentro de los procesos estocásticos consideraremos los de espacio de estados discreto y en ellos las Cadenas de Markov en Tiempo Discreto (CMTD).

**Definición 2.1.** Una CMTD es una sucesión de variables aleatorias  $\{X_n, n \in \mathbb{N}\}$  que cumple:

1. Cada variable  $X_n$  toma valores en un conjunto finito o numerable  $E$ , que se denomina espacio de estados.
2. La sucesión de variables verifica la condición de Markov:

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n), \quad (2.1)$$

donde  $i_0, i_1, \dots, i_n$  denotan los estados en los que se encuentra la cadena en cada etapa.

La condición de Markov exige que la probabilidad de que la cadena de Markov se encuentre en un estado  $j$  en el instante  $(n + 1)$  dependa únicamente del estado en que se encontraba en el instante  $n$  y que esto se cumpla para cualquier etapa en que se encuentre la cadena. La condición de homogeneidad en el tiempo, según la cual la probabilidad de pasar de  $i$  a  $j$  es independiente de la

etapa en que se encuentre la cadena, hace que la probabilidad de transición de  $i$  a  $j$  sea:

$$p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(X_{m+1} = j | X_m = i), \quad \forall n, m \in \mathbb{N}. \quad (2.2)$$

Una CMTD que cumpla esta condición de homogeneidad en el tiempo se dice que tiene probabilidades de transición estacionarias. En este trabajo consideraremos este tipo de CMTD.

## 2.1. Caracterización de una CMTD

Para caracterizar una CMTD debemos definir los siguientes elementos:

- Espacio de estados: Es el conjunto de posibles estados en los que se puede encontrar la CMTD en cada etapa. Puede ser discreto o infinito numerable.
- Matriz de transición: Es la matriz que aglutina las probabilidades de transición de unos estados a otros.
- Distribución inicial de la cadena: Vector de probabilidades en la etapa inicial de la CMTD.

### 2.1.1. Espacio de estados

El espacio de estados,  $E$ , es el conjunto de posibles valores que puede tomar el proceso en cada una de sus etapas. Consideraremos un espacio de estados discreto con  $k$  estados,  $\#E = k$ .

### 2.1.2. Matriz de transición

La matriz de transición  $P$ , que recoge las probabilidades de transición entre estados, es de dimensión  $k \times k$ . Los elementos de esta matriz son  $\mathbb{P}(X_{n+1} = j | X_n = i) = p_{ij}$ ,  $i, j \in E$ ,  $\forall n \in \mathbb{N}$ , que indican la probabilidad de pasar al estado  $j$  desde el estado  $i$ , en cualquier etapa. Dada la homogeneidad de la cadena, la matriz es de la forma:

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1k} \\ \vdots & \ddots & \vdots \\ p_{k1} & \cdots & p_{kk} \end{pmatrix}.$$

Los elementos de la matriz  $P$  verifican que  $p_{ij} \geq 0$ ,  $\forall i, j \in E$ . Además  $\sum_{j \in E} p_{ij} = 1$ ,  $\forall i \in E$ , es decir, las sumas por filas de la matriz son uno. Esta propiedad es característica de las matrices estocásticas, a las que pertenecen las

matrices de transición. Además, las potencias de una matriz estocástica también lo son, lo que será de utilidad a la hora de calcular la distribución de la CMTD en etapas posteriores.

### 2.1.3. Distribución inicial

La distribución inicial de la cadena se expresa en forma de vector, en el que cada componente indica la probabilidad de que la cadena se encuentre en el estado  $i$  en el instante inicial. De esta forma se conoce el punto de partida del proceso. Se expresa como:

$$P^{(0)} = (p_1^{(0)}, \dots, p_i^{(0)}, \dots, p_k^{(0)}),$$

donde  $p_i^{(0)} = \mathbb{P}(X_0 = i)$ ,  $i \in E$ .

Además, la distribución inicial cumple:  $p_i^{(0)} \geq 0$  y  $\sum_{i \in E} p_i^{(0)} = 1$ .

### 2.1.4. Distribución de la CMTD en la etapa $n$ -ésima

Una vez definida la CMTD se puede obtener la distribución marginal de  $X_n$ , es decir, la distribución de la cadena en la etapa  $n$ -ésima.

Si se denota la probabilidad de que en la etapa  $n$  la Cadena de Markov se encuentre en el estado  $i$  por  $p_i^{(n)} = \mathbb{P}(X_n = i)$ , para cada etapa se tiene un vector  $P^{(n)} = (p_1^{(n)}, p_2^{(n)}, \dots, p_i^{(n)}, \dots, p_k^{(n)})$  que representa la probabilidad de que la cadena se encuentre en cada uno de los posibles estados en la etapa  $n$ . Para poder obtener esta distribución se deberá calcular previamente la matriz de probabilidades de transición en  $n$  etapas.

Partiendo de que se conoce la probabilidad de transición en una etapa dada por (2.2), el siguiente paso es obtener la probabilidad de transición en  $n$  etapas,  $p_{ij}^{(n)} = \mathbb{P}(X_{m+n} = j | X_m = i)$ . Dichas probabilidades de transición en  $n$  etapas se obtienen calculando la potencia  $n$ -ésima de la matriz de transición  $P$ .

Este resultado se obtiene a partir de la ecuación de Chapman-Kolmogorov (véase [20], Sección 1.2):

$$p_{ij}^{(m+n)} = \sum_{g \in E} p_{ig}^{(m)} p_{gj}^{(n)}, \quad i, j \in E. \quad (2.3)$$

Intuitivamente, para pasar del estado  $i$  al  $j$  en  $(m+n)$  etapas, se debe pasar por un estado  $g$  en  $m$  etapas y después ir desde  $g$  hasta  $j$  en las  $n$  etapas restantes.

La condición de Markov (2.1) implica que las dos partes de la transición de  $i$  a  $j$  son independientes, por lo que se puede escribir  $p_{ij}^{(m+n)}$  como el producto de las probabilidades de transición y por lo tanto  $P^{(m+n)} = P^{(m)}P^{(n)}$ .

Tomando  $n = 1$  en la ecuación (2.3) se obtiene:

$$p_{ij}^{(m+1)} = \sum_{g \in E} p_{ig}^{(m)} p_{gj}, \quad i, j \in E.$$

Se tiene entonces la siguiente condición recursiva:

$$\begin{aligned} P^{(1)} &= P, \\ P^{(2)} &= P^{(1)}P^{(1)} = P^2, \\ &\vdots \\ P^{(n)} &= P^{(n-1)}P = P^{(n-2)}P^2 = \dots = P^{(0)}P^n, \end{aligned}$$

La distribución de la cadena en la etapa  $n$  se calcula como:  $P^{(n)} = P^{(0)}P^n$ . Por lo tanto, la probabilidad de transición en  $n$  etapas,  $p_{ij}^{(n)} = \mathbb{P}(X_{m+n} = j | X_m = i)$  se obtiene de la  $n$ -ésima potencia de la matriz de transición  $P$ .

Para calcular la potencia  $n$ -ésima de la matriz de transiciones se utiliza la descomposición espectral de dicha matriz. De esta forma, basándose en la teoría matricial, se obtienen de forma sencilla las potencias.

La matriz de transiciones es cuadrada, ya que es de dimensión  $k \times k$ , siendo  $k$  el cardinal del espacio de estados. Se hace una diagonalización de la misma, que se puede escribir como:

$$P = V\Lambda V^{-1},$$

donde  $\Lambda$  es la matriz diagonal formada por los valores propios de  $P$ :

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_k \end{pmatrix},$$

y  $V$  es la matriz que contiene, por columnas, los vectores propios de  $P$ .

De esta forma, la potencia de  $P$  se corresponde con:

$$P^n = V\Lambda^n V^{-1},$$

donde los elementos de  $P^n$  son las probabilidades de transición en  $n$  etapas.

Basta entonces saber que la potencia de una matriz diagonal se calcula a partir de la potencia de cada elemento de la matriz:

$$\Lambda^n = \begin{pmatrix} \lambda_1^n & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_k^n \end{pmatrix}.$$

La distribución en la etapa  $n$  se reduce al producto de la distribución inicial de la cadena por la potencia correspondiente de la matriz de transición.

$$P^{(n)} = P^{(0)} P^n.$$

### 2.1.5. Clasificación de estados

Dentro del espacio de estados de una CMTD se puede hacer una clasificación de los estados según las transiciones permitidas entre ellos y las probabilidades de pasar de un estado a otro.

Para realizar dicha clasificación de los distintos tipos de estado es necesario definir los conceptos de probabilidad y tiempos de primera pasada.

**Definición 2.2.** (Probabilidad de primera pasada) Es la probabilidad de que, empezando en  $i$ , la cadena pase por primera vez por el estado  $j$  en la etapa  $n$ . Se denota por:

$$f_{ij}^{(n)} = \mathbb{P}(X_n = j, X_r \neq j \forall r < n \mid X_0 = i).$$

**Definición 2.3.** (Probabilidad de pasada) Se trata de la probabilidad de que la cadena llegue alguna vez al estado  $j$  partiendo del estado  $i$ .

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)} = \mathbb{P}(\exists n; X_n = j \mid X_0 = i).$$

**Definición 2.4.** (Tiempo de primera pasada) Es el número de etapa en que la cadena llega por primera vez al estado  $j$  cuando parte del estado  $i$ . Se determina por la siguiente variable aleatoria:

$N_{ij} = \{n^o \text{ de la primera etapa en la cual la cadena está en } j \text{ partiendo de } i\}.$

Se escribe entonces:

$$f_{ij}^{(n)} = \mathbb{P}(N_{ij} = n), f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)} = \mathbb{P}(N_{ij} < \infty).$$

Se establece la siguiente relación entre las probabilidades de transición en  $n$  etapas,  $p_{ij}^{(n)}$  y las probabilidades de primera pasada  $f_{ij}^{(n)}$ :

$$p_{ij}^{(n)} = f_{ij}^{(1)} p_{jj}^{(n-1)} + f_{ij}^{(2)} p_{jj}^{(n-2)} + \dots + f_{ij}^{(n-1)} p_{jj} + f_{ij}^{(n)}.$$

Se puede hacer, entonces, la siguiente clasificación en el espacio de estados  $E$  de una CMTD:

1. Estado recurrente: Un estado  $j \in E$  se dice que es recurrente si  $f_{jj} = 1$ , es decir, es seguro que la cadena va a volver al estado  $j$  una vez que ya ha llegado a él en alguna etapa. De otra forma, es recurrente, si y sólo si el número de veces que se espera que la cadena pase por él cuando ya parte de él es infinito.
2. Estado transitorio: Se dice que un estado  $j \in E$  es transitorio si no es recurrente, es decir,  $f_{jj} < 1$ . En este caso se espera un número finito de visitas a dicho estado.
3. Estado que comunica con otro estado: Se dice que un estado  $i \in E$  comunica con otro estado  $j \in E$ , si  $f_{ij} > 0$ , es decir, existe la posibilidad de que la cadena llegue al estado  $j$  partiendo del estado  $i$ . Se cumple también si alguna potencia de la matriz de transición otorga probabilidad no nula a la transición entre los estados.
4. Estados que intercomunican. Dos estados  $i, j \in E$  intercomunican si  $i$  comunica con  $j$  y  $j$  comunica con  $i$ .
5. Estado efímero: Se dice que un estado  $j \in E$  es efímero si  $p_{ij} = 0, \forall i \in E$ . No se puede llegar a él desde ningún otro, solo se puede salir desde él hacia cualquier otro.
6. Estado absorbente: Un estado  $j \in E$  se dice absorbente si es imposible abandonarlo, es decir,  $p_{jj} = 1$ . Una vez que se alcanza, la CMTD sólo puede mantenerse en él. Además, una CMTD se dice que es absorbente si tiene al menos un estado absorbente y desde cada uno de los estados es posible alcanzar alguno de los absorbentes en un número finito de etapas.

Con respecto a la clasificación de los estados se tiene el siguiente teorema (véase [20], Sección 1.3):

**Teorema 2.5.** *(Teorema de descomposición) El espacio de estados puede dividirse en dos subconjuntos disjuntos, estados transitorios y estados recurrentes.*

1. *Los estados transitorios son inaccesibles desde los estados recurrentes.*
2. *Los estados recurrentes pueden dividirse de manera única en clases de equivalencias establecidas por la relación de equivalencia  $i \sim j$ .*

- a) Si dos estados  $i, j$  están en la misma clase  $f_{ij} = f_{ji} = 1$ .  
 b) Si dos estados están en clases distintas  $f_{ij} = f_{ji} = 0$ .

Se definen además los siguientes conceptos, en función de la clasificación de los estados:

**Definición 2.6.** (Conjuntos de estados)

1. Un subconjunto de estados  $C \subset E$ , se dice que es cerrado si cada estado de  $C$  comunica exclusivamente con estados de  $C$ .
2. Un subconjunto de estados  $C \subset E$ , se dice que es irreducible si todos los estados de  $C$  intercomunican entre sí.

**Definición 2.7.** (Clasificación de estados recurrentes)

1. Un estado recurrente  $j \in E$  se dice que es recurrente positivo si  $\mathbb{E}(N_{jj}) < \infty$ , es decir, el número esperado de la primera etapa en la que la cadena pasa por  $j$  partiendo de  $j$  es finito. Se espera volver al estado  $j$  en un número finito de etapas.
2. Un estado recurrente  $j \in E$  se dice que es recurrente nulo si  $\mathbb{E}(N_{jj}) = \infty$ . Por ser recurrente es seguro que la cadena va a volver a pasar por el estado  $j$  si parte de él, pero lo hará en un número infinito de etapas.

A partir de esta definición de estados recurrentes se sabe que, si dos estados recurrentes  $i, j$  intercomunican, o son ambos positivos o son ambos nulos. Además, en una CMTD finita e irreducible todos los estados son recurrentes positivos.

### 2.1.6. Distribución estacionaria de una CMTD

La distribución estacionaria asociada a una CMTD tiene la propiedad de que, si  $X_0$  tienen una determinada distribución, todas las  $X_n$  tiene la misma distribución. Al considerarse probabilidades de transición homogéneas, la transición de un estado a otro no depende de la etapa en que se produzca. Esto hace que las distribuciones marginales de la cadena se estabilicen, por lo que no hay variaciones entre ellas en diferentes etapas una vez que se alcanza la distribución “en equilibrio”.

**Definición 2.8.** (Distribución estacionaria) Una distribución  $\Pi = \{\pi_i\}_{i \in E}$  sobre  $E$  se dice estacionaria respecto de una Cadena de Markov con matriz de transición  $P$  si verifica que  $\Pi P = \Pi$ .

De la misma forma,  $\Pi$  es una distribución estacionaria si  $\sum_{i \in E} \pi_i = 1$  y se verifica:

$$\pi_i = \sum_{j \in E} \pi_j p_{ij}, \quad \forall i \in E.$$

Los valores de la distribución estacionaria,  $\{\pi_i\}_{i \in E}$  se pueden interpretar como la proporción final de tiempo que la cadena ha pasado en cada estado a lo largo de su evolución, con probabilidad 1. Se puede decir que es también la proporción, a largo plazo, de etapas en que la cadena se encuentra en el estado  $i$  a lo largo de su evolución si ha partido de  $i$ , o de otro estado recurrente que intercomunica con  $i$ .

La distribución estacionaria asigna probabilidad 0 a los estados transitorios o recurrentes nulos. Además, la condición necesaria y suficiente para que una CMTD tenga distribución estacionaria es que tenga alguna subcadena recurrente positiva, como se ve en el siguiente teorema:

**Teorema 2.9.** (*Teorema de Levinson*) *Si  $P$  es la matriz de transición de una CMTD irreducible, existe distribución estacionaria respecto de  $P$  si y sólo si la cadena es recurrente positiva. En este caso, la distribución estacionaria es única y se tiene que:*

$$\pi_j = \frac{1}{\mathbb{E}[N_{jj}]}.$$

Si existen varias subcadenas recurrentes positivas, la cadena tendrá tantas distribuciones estacionarias como combinaciones lineales convexas de dichas distribuciones.

En la descripción de los métodos desarrollados (Parte II), se matizarán las caracterizaciones de CMTD vistas en este capítulo, si bien el análisis de los resultados no abarcará todos los aspectos revisados.



## Capítulo 3

# Estimación en CMTD

En la mayoría de las situaciones prácticas, las probabilidades de transición de una CMTD son desconocidas y deben estimarse en base a observaciones de la misma. En numerosos artículos se han estudiado estimadores para dichas probabilidades de transición, desarrollando el estimador de máxima verosimilitud utilizado en este trabajo. Véase [21], [22], [23] y [24].

A partir de la observación de la cadena a lo largo de una serie de etapas, se pretende estimar las probabilidades de transición, es decir, los elementos de la matriz de transición que determina la CMTD. Estas probabilidades tienen que cumplir las condiciones siguientes:

- Ser mayores o iguales a cero.

$$p_{ij} \geq 0.$$

- La suma de las filas de la matriz debe ser uno. La suma de las probabilidades de pasar del estado  $i$  a todos los otros estados posibles es uno.

$$\sum_{j \in E} p_{ij} = 1.$$

Sean  $i = 1, 2, \dots, k$  los estados de la CMTD. Sean  $t = 0, 1, \dots, n$  los tiempos de observación. Denotemos por  $p_{i_{t-1}i_t}$  la probabilidad de estar en el estado  $i_t$  en la etapa  $t$ , sabiendo que estaba en el estado  $i_{t-1}$  en la etapa  $(t-1)$ .

La observación de una CMTD supone una secuencia de estados en los que la cadena está en cada etapa. Para las etapas  $t = 0, 1, \dots, n$  se situará en los estados:  $i_0, i_1, \dots, i_n$ . Dado el estado inicial  $i_0$ , hay  $k^n$  posibles secuencias, correspondientes a las  $k^n$  posibles ordenaciones de las transiciones.

Las probabilidades de cada secuencia del tipo  $S = \{i_0, i_1, \dots, i_n\}$ , donde  $X_0 = i_0, X_1 = i_1, \dots, X_n = i_n$ , dado que las probabilidades de transición son estacionarias, son:

$$\mathbb{P}(S) = p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

Se denota por  $n_{i_0}$  al número de individuos en el estado  $i_0$  en la etapa 0 (etapa inicial) y por  $n_{i_{t-1} i_t}$  al número de individuos en el estado  $i_{t-1}$  en la etapa  $(t-1)$  y en el estado  $i_t$  en la etapa  $t$ .

Sea  $n_S = n_{i_0, i_1, \dots, i_n}$  el número de individuos cuya secuencia de estados es  $S = \{i_0, i_1, \dots, i_n\}$ . En una cadena de Markov con probabilidades de transición estacionarias, el conjunto  $n_S, \forall S \in E$  forma un conjunto de estadísticos suficientes (véase [21]). La probabilidad, en el espacio que describe las secuencias de la cadena, de un conjunto de secuencias ordenadas dada para los  $n_S$  individuos es:

$$\prod_S \mathbb{P}(S)^{n_S} = \prod_S [p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}]^{n_{i_0, i_1, \dots, i_n}} = \prod_{i, j} p_{ij}^{n_{ij}}. \quad (3.1)$$

### 3.1. Estimación por máxima verosimilitud

Las probabilidades de transición estacionarias  $p_{ij}$  (2.2) pueden estimarse maximizando la probabilidad (3.1) con respecto a  $p_{ij}$ , sujeto a las restricciones de que  $p_{ij} \geq 0$  y que  $\sum_{j \in E} p_{ij} = 1$ , cuando  $n_{ij}$  son las observaciones actuales.

El estimador de máxima verosimilitud de  $p_{ij}$  es entonces:

$$\hat{p}_{ij} = \frac{n_{ij}}{n_i^*}, \quad \text{donde } n_i^* = \sum_{j \in E} n_{ij}, \quad (3.2)$$

siendo  $\hat{P}$  la matriz de transiciones de dimensión  $k \times k$  estimada por máxima verosimilitud, cuyos elementos son los  $\hat{p}_{ij}$ .

Para realizar las estimaciones es necesario conocer el estado actual y el estado anterior de los individuos observados, de forma que se pueda contabilizar el cambio en, al menos, una etapa. Se denotarán por  $n_{ij}$  las observaciones de transiciones, el recuento del número de individuos que pasan del estado  $i$  al estado  $j$  en una etapa.

Se trata, en la práctica, de hacer recuentos del número de individuos que pasan del estado  $i$  al estado  $j$  en las observaciones de una etapa, para todos los estados posibles. Este número se divide por la suma de todos los individuos que salen del estado  $i$  considerado en cada caso.

Es conocido (véase [21] y [22]) que los estimadores de máxima verosimilitud siguen un proceso similar a los estimadores de las probabilidades para una distribución multinomial para cada  $i \in E$ . Son probabilidades obtenidas de forma similar a  $n$  muestras independientes, donde cada una consiste en  $n_i^*$  pruebas multinomiales con probabilidades  $p_{ij}$ .

### 3.1.1. Cálculo de los estimadores mediante tablas

Para calcular las probabilidades a partir de tablas de contingencia se pueden definir dos variables, estado anterior y estado actual. Si para cada realización de la cadena se obtienen observaciones de ambas variables, una para la variable estado anterior y otra para la variable estado actual, se puede hacer el cálculo de los estimadores a partir de la tabla de contingencia de estas dos variables.

La tabla de contingencia de estado anterior y estado actual contabilizará el número de veces que la cadena pasa del estado  $i$  al estado  $j$ . Por lo que cada elemento  $(i, j)$  de la tabla será  $n_{ij}$ . Calculando las proporciones por filas de esa tabla se obtiene el número de individuos que pasan del estado  $i$  al  $j$  dividido por el número de individuos que salen del estado  $i$ , coincidiendo con la definición del estimador (3.2). Además, es obvio que la suma de cada fila será uno.

Por lo tanto, la matriz de transición estimada por máxima verosimilitud se puede obtener a partir de la tabla de contingencia de las variables estado anterior y estado actual, calculando las distribuciones marginales por filas.

Cuando las observaciones de las transiciones, es decir, de las variables estado anterior y estado actual en base a las que se calculan los estimadores provienen de datos obtenidos mediante muestreo, es necesario incluir en el cálculo el efecto del diseño muestral como se verá en las Secciones 6.3 y 6.4.

### 3.1.2. Distribución asintótica de los estimadores

En [21] se estudia la distribución asintótica de los estimaciones de máxima verosimilitud de las probabilidades de transición (3.2). Dicha distribución asintótica viene dada por:

$$\sqrt{n_i^*}(\hat{p}_{ij} - p_{ij}) \longrightarrow N(0, p_{ij}(1 - p_{ij})), \quad \forall i, j \in E,$$

cuando  $n_i^* \longrightarrow \infty$ .

De manera compacta, el resultado anterior se puede expresar para todas las probabilidades de transición a partir de la vectorización de la matriz  $P$ . En concreto, si denotamos por  $P_v$  el vector de longitud  $k \times k$  formado por las filas de la matriz  $P$  concatenadas, y por  $\hat{P}_v$  al correspondiente vector de las estimaciones, se tiene que:

Cuando  $n \longrightarrow \infty$ ,

$$\sqrt{n}(\hat{P}_v - P_v) \longrightarrow N(0, \Sigma_P).$$

Debe cumplirse además la condición de que, cuando  $n \longrightarrow \infty$ , es decir, cuando crece el número total de transiciones, lo hace también el número total

de transiciones partiendo de cada estado. De ahora en adelante se supone esta condición en los resultados asintóticos.

Además,  $\Sigma_P$ , que es la matriz de covarianzas de los estimadores de las probabilidades de transición, es una matriz cuadrada  $k^2 \times k^2$ , con  $k$  el número de estados de la cadena, que viene dada por:  $(\Sigma_P)_{(ij,kl)} = \delta_{ij}P_{ij}(\delta_{jl} - P_{il})$ . donde  $\delta_{ij}$  es la función indicadora (vale 1 si los estados  $i$  y  $j$  coinciden y cero en otro caso).

Como consecuencia de la distribución asintótica de los estimadores se pueden construir intervalos de confianza. Por ejemplo, con un nivel de confianza  $(1 - \alpha)$ , el intervalo de confianza vendría dado por:

$$\hat{p}_{ij} \pm z_{1-\alpha/2} \sqrt{\frac{\widehat{\Sigma}_P(ij, ij)}{n_i^*}} = \hat{p}_{ij} \pm z_{1-\alpha/2} \sqrt{\frac{\hat{p}_{ij}(1 - \hat{p}_{ij})}{n_i^*}}. \quad (3.3)$$

siendo  $z_\alpha$  el cuantil  $\alpha$  en la  $N(0,1)$ .

Entonces, para cada probabilidad de transición estimada por máxima verosimilitud, de forma individual, se establecen intervalos de confianza basados en la teoría asintótica de dichos estimadores. Se podrían obtener también regiones de confianza para subvectores de probabilidades de transición, utilizando teoría sobre formas cuadráticas en vectores normales.

### 3.1.3. Contrastes de hipótesis sobre la distribución en la etapa $n$ -ésima

Conociendo la distribución en una determinada etapa de una CMTD, puede interesar contrastar si esa distribución coincide con otra dada. Esto es útil, por ejemplo, para validar los modelos realizados con CMTD en los casos en que se conocen las distribuciones reales en ciertas etapas. De esta forma se puede ver si el ajuste es bueno o no, comparando la distribución obtenida por la CMTD con sus valores esperados. En [25] y [26] se estudiaron distintos contrastes para la bondad de ajuste de una distribución multinomial, que en el contexto que nos ocupa, pueden ser utilizados para contrastar una determinada distribución sobre  $P^{(n)}$ , la distribución de la cadena en la etapa  $n$ -ésima.

Dado que la distribución de la cadena en la etapa  $n$ -ésima puede considerarse como una distribución multinomial de dimensión  $k$ , siendo  $k$  el número de estados, y tamaño  $N$ , donde  $N$  viene determinado por el número de observaciones de la cadena en base a las que se obtienen las estimaciones de sus probabilidades de transición, y la probabilidad de cada uno de los estados es  $p_i^{(n)} = \mathbb{P}(X_n = i)$ ,  $i \in E$ .

Para contrastar la hipótesis nula de que la distribución de la cadena en la etapa  $n$ -ésima toma los valores  $P^{(n)} = (p_{10}^{(n)}, \dots, p_{k0}^{(n)})$ , con  $\sum_{i=1}^k p_{i0}^{(n)} = 1$ ,

consideraremos un test de razón de verosimilitudes sobre una distribución multinomial. Se plantea el siguiente contraste:

$$H_0 : P^{(n)} = \left( p_{10}^{(n)}, \dots, p_{k0}^{(n)} \right)$$

frente a la alternativa general

$$H_a : P^{(n)} = \wp, \quad \text{con } \wp = (\wp_1, \dots, \wp_k), \quad \sum_{i=1}^k \wp_i = 1.$$

Bajo la hipótesis alternativa  $H_a : P^{(n)} = \wp$ , la masa de probabilidad de una muestra  $\vec{x} = (x_1, \dots, x_k)$ , donde  $\sum_{i=1}^k x_i = N$ , viene dada por:

$$\mathbb{P}_{H_a}(\vec{x}) = N! \prod_{i=1}^k \frac{\wp_i^{x_i}}{x_i!}. \quad (3.4)$$

Por tanto, la función de verosimilitud, definida sobre los vectores de masa de probabilidad  $\wp / \wp_i \geq 0, \sum \wp_i = 1 \forall i = 1, \dots, k$ , es la siguiente:

$$l(\wp) = N! \prod_{i=1}^k \frac{\wp_i^{x_i}}{x_i!}.$$

Para obtener la expresión del test de razón de verosimilitudes es necesario obtener la expresión de la verosimilitud utilizando los estimadores de máxima verosimilitud de las probabilidades de transición. El estadístico de contraste viene dado por:

$$T = -2 \sum_{i=1}^k x_i \log \frac{p_{i0}}{\hat{p}_i},$$

que bajo ciertas condiciones de regularidad sigue una distribución asintótica  $\chi_{k-1}^2$ , con  $(k-1)$  grados de libertad.

La aproximación del estadístico por  $\chi^2$  empeora cuantas más categorías,  $k$ , se consideren (véase [26]). La mala aproximación del estadístico se traduce en un mal ajuste de los momentos. La mejora más sencilla consiste en multiplicar el estimador por un factor de corrección que fuerce el ajuste en los primeros momentos de la distribución. Una primera propuesta consiste en utilizar  $T_1$  en lugar del estadístico  $T$ :

$$T_1 = \frac{T}{q_1}, \quad \text{donde } q_1 = 1 + \frac{1}{6N(k-1)} \left( \sum_{i=1}^k p_{i0}^{-1} - 1 \right).$$

Esta corrección fue mejorada en [26], utilizando la siguiente:

$$T_2 = \frac{T}{q_2}, \text{ donde } q_2 = 1 + \frac{1}{6N(k-1)} \left( \sum_{i=1}^k p_{i0}^{-1} - 1 + \frac{1}{N} \sum_{i=1}^k (p_{i0}^{-1} - p_{i0}^{-2}) \right).$$

Mediante estas correcciones se consigue reducir el error de aproximación del estadístico de contraste a la distribución  $\chi^2$ , al reducir el error de aproximación de sus momentos. Además, la distribución del estadístico no se altera con las correcciones vistas.

En ocasiones las correcciones propuestas no son suficientes, el tamaño muestral y el número de categorías dificultan que los resultados de los test se correspondan realmente con las distribuciones. En [25] se estudian varios test de bondad de ajuste para la multinomial. Dichos test tienen la misma estructura, con una divergencia potencial, de forma que es posible estudiar de forma simultánea varias expresiones del mismo test. En concreto se estudian los siguientes estadísticos:

- Estadístico  $\chi^2$ :  $\chi^2 = \sum_{i=1}^k (X_i - n\varphi_i)^2 / n\varphi_i$ .
- Estadístico de razón de verosimilitudes:  $G^2 = 2 \sum_{i=1}^k X_i \ln(X_i / n\varphi_i)$ .
- Estadístico de Freeman-Tukey:  $T^2 = 4 \sum_{i=1}^k (\sqrt{X_i} - \sqrt{n\varphi_i})^2$ .
- Estadístico  $\chi^2$  modificado de Neyman:  $NM^2 = \sum_{i=1}^k (X_i - n\varphi_i)^2 / X_i$ .
- Estadístico modificado de razón de verosimilitudes:  $GM^2 = 2 \sum_{i=1}^k n\varphi_i \ln(n\varphi_i / X_i)$ .

Para estos test se mantienen las hipótesis nula y alternativa planteadas anteriormente.

Estos estadísticos forman parte de la misma familia de divergencias, además son asintóticamente equivalentes y siguen una distribución  $\chi^2$  con  $k-1$  grados de libertad.

Cuando los dos primeros momentos de la distribución de los estadísticos, bajo  $H_0$ , no son cercanos a  $k-1$  y  $k^2-1$ , respectivamente, se proponen correcciones para los momentos de orden uno y orden dos. De esta forma el estadístico corregido tendrá media  $k-1$  y varianza  $2(k-1)$ , por lo que se mejora la aproximación, al menos, en los primeros momentos.

Se ha visto que las probabilidades de transición de una CMTD se pueden estimar por máxima verosimilitud cuando son desconocidas. Para realizar dichas

estimaciones es necesario contar con alguna observación de la cadena, conociendo el estado de un determinado número de individuos en dos etapas consecutivas para poder determinar la transición que llevaron a cabo. En base a estas observaciones se realizan los recuentos necesarios para obtener las estimaciones. Estas observaciones pueden proceder de muestras complejas, como las utilizadas para obtener los datos analizados en este trabajo, mediante muestreo estratificado.

Los test de bondad de ajuste, asumiendo una distribución multinomial y aproximando su distribución por  $\chi^2$  no tienen en cuenta el efecto del diseño muestral. En [27] se estudia el efecto de las muestras complejas sobre las distribuciones asintóticas del test de bondad de ajuste  $\chi^2$ , viendo que en muestras estratificadas el estadístico  $\chi^2$  es asintóticamente conservador.

Se proponen correcciones basadas en el efecto del diseño. Para realizar estas correcciones es necesario conocer la matriz de covarianzas de las distribuciones utilizadas en el contraste, lo que no siempre es posible.





## Capítulo 4

# Técnicas Bootstrap para CMTD

Basándose en técnicas bootstrap se pueden obtener intervalos de confianza para las probabilidades de transición de la cadena y para la distribución de la misma en cualquier etapa  $n$ . En [28], [24] y [29] se utilizan técnicas bootstrap aplicables a CMTD.

Sean  $X_1, \dots, X_n$  observaciones de una CMTD con matriz de transición  $P$  y sea  $\hat{P}$  el estimador de máxima verosimilitud de  $P$ . Se genera una cadena bootstrap  $X_1^*, \dots, X_n^*$ , generando una CMTD con matriz de transición estimada por máxima verosimilitud  $\hat{P}^*$ .

En los trabajos citados se ha probado que las estimaciones bootstrap en este ámbito son consistentes y se ha estudiado su distribución y propiedades.

### 4.1. Réplicas Bootstrap de una CMTD

Se muestra a continuación el proceso por el cual se generan las réplicas bootstrap de la cadena, basadas en la matriz de probabilidades de transición estimadas. Se pretenden generar  $B$  muestras bootstrap  $X_1^{*(b)}, \dots, X_n^{*(b)}$ , con  $b = 1, \dots, B$  a partir de la distribución inicial de la cadena,  $P^{(0)}$  y de la matriz de transiciones, estimada por máxima verosimilitud,  $\hat{P}$ .

Se considera, al igual que en resultados anteriores, que tanto la distribución inicial como las filas de la matriz de transiciones pueden asimilarse a una distribución multinomial.

Paso 1:

Para iniciar la CMTD se generan aleatoriamente  $B$  multinomiales de tamaño  $n$ , igual al tamaño muestral original, con probabilidades iguales a la distribución inicial, utilizando la misma parametrización que en (3.4). Cada uno de estos  $B$  vectores será una muestra observada en una etapa de la cadena.

$$P^{(0)} \longrightarrow \begin{matrix} X_0^{*(1)} \\ \vdots \\ X_0^{*(B)}. \end{matrix}$$

Paso 2:

Realizamos una transición de la cadena aplicando, a cada una de las  $B$  muestras generadas como distribución inicial, la matriz de transición estimada por máxima verosimilitud,  $\hat{P}$ .

$$\begin{matrix} X_0^{*(1)} & \xrightarrow{\hat{P}} & X_1^{*(1)} & \dots & X_n^{*(1)} \\ X_0^{*(2)} & \xrightarrow{\hat{P}} & X_1^{*(2)} & \dots & X_n^{*(2)} \\ \vdots & & \vdots & & \vdots \\ X_0^{*(B)} & \xrightarrow{\hat{P}} & X_1^{*(B)} & \dots & X_n^{*(B)} \end{matrix}$$

Paso 3:

En este momento podemos aplicar la definición del estimador de máxima verosimilitud (3.2) para la matriz de transición y obtener  $B$  matrices de transición bootstrap, una para cada muestra.

$$\begin{matrix} X_1^{*(1)} & \dots & X_n^{*(1)} & \longrightarrow & \hat{P}^{*(1)} \\ X_1^{*(2)} & \dots & X_n^{*(2)} & \longrightarrow & \hat{P}^{*(2)} \\ \vdots & & \vdots & & \vdots \\ X_1^{*(B)} & \dots & X_n^{*(B)} & \longrightarrow & \hat{P}^{*(B)} \end{matrix}$$

#### 4.1.1. Intervalos de confianza Bootstrap para las probabilidades de transición

Se han construido  $B$  matrices de transición a partir de la CMTD. Vectorizando cada una de las  $B$  matrices, de la misma forma que en la Sección 3.1.2., obtenemos  $B$  valores para cada probabilidad de transición.

$$\begin{array}{cccc} \hat{P}_v^{*(1)} & \hat{P}_v^{*(2)} & \dots & \hat{P}_v^{*(B)} \\ \hline p_{11}^{*(1)} & p_{11}^{*(2)} & \dots & p_{11}^{*(B)} \\ p_{12}^{*(1)} & p_{12}^{*(2)} & \dots & p_{12}^{*(B)} \\ \vdots & \vdots & & \vdots \end{array}$$

Basta con ordenar los valores correspondientes a cada probabilidad y calcular los cuantiles de probabilidades  $\alpha/2$  y  $(1 - \alpha/2)$  para obtener intervalos de confianza de nivel  $(1 - \alpha)$  para cada una de las probabilidades de transición de la CMTD.

#### 4.1.2. Intervalos de confianza Bootstrap para la distribución en la etapa $n$ -ésima

Se pueden obtener intervalos de confianza bootstrap para la distribución de la cadena en cualquier etapa  $n$ . Anteriormente se han generado  $B$  matrices de transición a partir de la CMTD, con las que se obtienen intervalos de confianza para las probabilidades de transición. Estas  $B$  matrices sirven también para obtener nuevos intervalos de confianza.

En el Capítulo 2 se ha visto que la distribución de la CMTD en una etapa  $n$  se calcula a partir de la matriz de transición y de la distribución inicial como  $P^{(n)} = P^{(0)} P^n$ .

Para cada una de las  $B$  matrices de transición obtenidas, junto con la distribución inicial, se puede calcular la distribución de la cadena generada en la etapa  $n$ -ésima de la forma:

$$\begin{array}{rcl} (\hat{P}^{*(1)})^{(n)} & = & P^{(0)} (\hat{P}^{*(1)})^n \\ (\hat{P}^{*(2)})^{(n)} & = & P^{(0)} (\hat{P}^{*(2)})^n \\ \vdots & & \vdots \\ (\hat{P}^{*(B)})^{(n)} & = & P^{(0)} (\hat{P}^{*(B)})^n \end{array}$$

Se tienen entonces  $B$  valores para la distribución  $n$ -ésima de la CMTD. Se ordenan los  $B$  valores correspondiente a cada estado de la cadena y se calculan los cuantiles de probabilidades  $\alpha/2$  y  $(1 - \alpha/2)$  respectivamente. Se obtienen intervalos de confianza para la distribución de la cadena en la etapa  $n$ -ésima, de forma individual para cada estado.

## 4.2. Suavización de la matriz de transición

Las probabilidades de transición de una CMTD se estiman, mediante el método de máxima verosimilitud, en base a observaciones de la misma. Dichas observaciones se obtienen de una muestra de la que se conoce su estado en, al menos, dos etapas para poder determinar la transición ocurrida. Algunos estados tienen probabilidad muy baja de aparecer en la muestra, bien porque sean casos raros o porque el tamaño muestral no es lo suficientemente grande. Esto supone que, en ocasiones, algunas probabilidades de transición estén infraestimadas o incluso se les asigne valor cero. Una probabilidad nula en una determinada transición significa que dicha transición es imposible, por lo que asignar probabilidad cero a transiciones que ocurren en la realidad induce errores en el modelo al obviar ciertas situaciones posibles.

La utilización de estimadores suavizados para la matriz de transiciones, como los estudiados en [29], resuelven este problema al asignar valores positivos a todas las probabilidades para variables discretas. Este proceso se utiliza cuando no es posible aumentar el tamaño muestral hasta que la muestra sea representativa de todos los estados considerados.

En el caso en que algunas probabilidades de transición sean cero debido a que es imposible en la realidad que se produzcan, este método le asignaría también probabilidad positiva. Por lo tanto, solo tiene sentido utilizar la suavización de la matriz cuando todas las transiciones entre estados son posibles y, en el caso de que alguna sea cero, sea debido a que no se observa en la muestra.

En un primer momento se consideran estimadores de la forma:

$$\tilde{P} = (1 - \alpha)\hat{P} + \alpha\theta,$$

donde  $\hat{P}$  es la matriz de transiciones estimada por máxima verosimilitud, con  $\alpha \in [0, 1]$  y  $\theta$  es un vector de probabilidades.

Otra forma de suavizar la matriz de transiciones es utilizar el método Kernel de la siguiente forma:

$$\tilde{P}_{ij} = \sum_{g \in E} \hat{P}_{ig} K \left( \frac{i-g}{h} \right),$$

$$K \left( \frac{i-g}{h} \right) = \begin{cases} h & \text{si } g = j \\ \frac{1-h}{k-1} & \text{si } g \neq j, \end{cases}$$

donde  $\frac{1}{k} \leq h \leq 1$ .

Esta formulación es equivalente a la primera ya que:

$$\tilde{P}_{ij} = h\hat{P}_{ij} + (1 - \hat{P}_{ij})\frac{1-h}{k-1} \iff \tilde{P}_{ij} = \frac{kh-1}{k-1}\hat{P}_{ij} + \frac{1-h}{k-1},$$

tomando  $\alpha = k \left( \frac{1-h}{k-1} \right)$  y  $\theta = (1/k, \dots, 1/k)$ .  $k$  representa el número de estados de la cadena y  $h$  es el parámetro de suavizado.

La dificultad en estos métodos radica en la elección del parámetro de suavizado,  $h$ .



Parte II

Material y Métodos





## Capítulo 5

# Fuente de datos

Los modelos basados en Cadenas de Markov en Tiempo Discreto (CMTD) necesitan datos que permitan calcular las transiciones entre los distintos estados considerados. En este caso concreto, en el que se define un modelo que predice prevalencias de consumo de tabaco, se necesita conocer el comportamiento de un grupo de individuos con respecto al consumo del mismo en un determinado momento temporal. Este comportamiento individual determinará la evolución de la CMTD.

La fuente principal de datos para el modelo fue el Sistema de Información sobre Conductas de Riesgo en Galicia, de ahora en adelante SICRI. De forma complementaria se utilizaron datos procedentes del programa-concurso “Déixao e Gaña” y del modelo SimSmoke.

### 5.1. Sistema de Información sobre Conductas de Riesgo (SICRI)

El SICRI es un sistema de información que se puso en marcha en Galicia en el año 2005 en el servicio de Epidemiología de la Dirección Xeral de Innovación e Xestión da Saúde Pública (DXIXSP). Se basa en la realización de encuestas telefónicas empleando un sistema CATI (computed-assisted telephone interview), con periodicidad anual, sobre prácticas preventivas y conductas de riesgo para la salud entre la población residente en Galicia.

#### Características principales de las encuestas

- Ámbito poblacional: mayores de 15 años residentes en Galicia.
- Ámbito geográfico: todo el territorio gallego.
- Ámbito temporal: para cada una de las encuestas el período de referencia es el año correspondiente.

En este trabajo se emplearon las encuestas realizadas en los años 2006, 2007, 2010 y 2011. A continuación se detallan sus características principales, que se resumen en la Figura 5.1.

### **SICRI 2006**

- Marco de muestreo: directorio telefónico que contiene a los titulares de una línea de teléfono fija.
- Tipo de muestreo: bietápico con estratificación de las unidades de primera etapa (UPE). Las UPE son los números telefónicos, que representan a las viviendas familiares. Las unidades de segunda etapa, y unidades de análisis, son las personas de 16 años de edad y más residentes en las viviendas seleccionadas. Dentro de cada vivienda se selecciona a una única persona. Las UPE se estratifican en función del municipio al que pertenecen, utilizando un doble criterio: el tamaño de población, según el Padrón de 2005, y el nivel de ingresos, con el objetivo de asegurar una adecuada representación de los ámbitos rural y urbano. Se definen cuatro estratos. Además, las unidades de segunda etapa se estratifican en función del sexo y grupo de edad (16-24,25-44,45-64,65 y más). Se definen ocho estratos.

### **SICRI 2007, 2010 y 2011**

- Marco de muestreo: Tarjeta Sanitaria. La encuesta es telefónica, por lo que se reduce el marco de muestreo al listado de individuos de 16 años y más que tienen teléfono en la base de la Tarjeta Sanitaria. La cobertura estimada es del 93,5 % de la población.
- Tipo de muestreo: aleatorio estratificado. Las unidades de muestreo y análisis son los individuos de 16 años y más que tienen teléfono en la base de la Tarjeta Sanitaria. Esta población se estratifica en función del sexo y grupo de edad (16-24,25-44,45-64,65 y más). Se definen ocho estratos.

#### **5.1.1. Tamaño de la muestra**

En cualquiera de las encuestas SICRI empleadas en este trabajo, el tamaño de muestra necesario se determinó con el objetivo de garantizar una adecuada representatividad en los ocho estratos definidos por sexo y grupo de edad. En cada uno de los estratos, el tamaño de muestra se calculó para una prevalencia del 50 %, un error absoluto del 3,5 % con un nivel de confianza del 95 % y un efecto del diseño de 1,25. El tamaño de población se obtuvo del Padrón del año anterior, que por ejemplo en 2005 contaba con 2.421.978 habitantes. Para cada grupo, resultó un tamaño teórico de  $n=980$  que supuso un total de 7840 entrevistas. En la Figura 5.1 se indica el tamaño de muestra de cada encuesta.

	SICRI 2006	SICRI 2007	SICRI 2010	SICRI 2011
Marco de muestreo	Directorio telefónico	Tarjeta Sanitaria	Tarjeta Sanitaria	Tarjeta Sanitaria
Tamaño de muestra	7841	7819	7845	7844
Número de preguntas	106	43	70	91
Relacionadas con el consumo de tabaco	66	15	26	26

Figura 5.1: Características principales de las distintas encuestas SICRI.

### 5.1.2. Cuestionarios

Todas las encuestas enmarcadas en el SICRI están estructuradas en bloques de preguntas agrupadas según la temática y hay varios bloques de preguntas comunes a todos los estudios. Los bloques comunes en todas las encuestas son:

1. Información sociodemográfica: sexo, edad, nivel de estudios, situación laboral y estado civil.
2. Consumo de tabaco: diferentes preguntas que permiten estimar la prevalencia de consumo de tabaco, la edad de experimentación y consolidación del hábito, el tipo de tabaco consumido y la cantidad, intentos de dejar de fumar entre los fumadores y el tiempo que llevan sin fumar los exfumadores.
3. Exposición al humo ambiental del tabaco: diferentes preguntas que permiten estimar la prevalencia de exposición en diferentes lugares.
4. Actividad física: descripción detallada de las distintas actividades realizadas por la población en relación con el ejercicio físico.
5. Medidas antropométricas autodeclaradas: peso y talla.

## 5.2. Fuentes de datos complementarias

La información aportada por el SICRI no permite la construcción completa del modelo. Esta carencia condiciona el uso de otras fuentes complementarias, que se detallan a continuación.

### 5.2.1. Programa-concurso Déixao e Gaña

Debido a que las encuestas del SICRI no permiten estimar las tasas de recaídas entre los exfumadores fue necesario explorar la existencia de otras fuentes de datos. Así se valoró obtener la información de otros estudios o estimarla a

partir de datos del programa-concurso “Déixao e Gaña”. Se decidió su estimación a partir de la información aportada por “Deixao e Gaña” ya que los otros estaban realizados en poblaciones con historias tabáquicas diferentes a las de Galicia y, por tanto, no se pueden aplicar a nuestra comunidad.

El “Deixao e Gaña” es un programa-concurso destinado a fumadores que quieren dejar de fumar. Sus antecedentes están en el Minnesota Heart Program, un estudio que empezó en Minnesota en la década de 1980 y que entre sus actividades incluía una competición para dejar de fumar. Esta competición consiguió unas tasas de participación que llegaron en ocasiones al 7 % y que estaba acompañada de una abstinencia a los 6 meses que superó el 20 %. Estos resultados animaron la realización de otras actividades de promoción de la salud basadas en competiciones o concursos con un diseño similar, en las que se fueron modificando diferentes parámetros como, por ejemplo, la edad de la población que podía participar o la duración del concurso. Estos programas fueron los precursores del “Quit & Win” y del “Deixao e Gaña” que se celebra en Galicia, organizado por la DXIXSP, desde el año 1998 con una periodicidad bianual y que en 2010 celebró su 7ª edición. El programa-concurso está dirigido a fumadores mayores de edad que se comprometen, a partir de una fecha fija, a dejar de fumar durante 30 días pudiendo conseguir a cambio diferentes premios.

El procedimiento del concurso es el siguiente. Los fumadores interesados en dejar de fumar se inscriben al concurso, que previamente se publicita en toda Galicia. Al inicio se realiza un control para comprobar que los participantes son fumadores y se les explica que, para poder optar a los premios, deben mantenerse sin fumar durante el mes que dura el concurso. A los que lo solicitan se les proporciona ayuda para dejar de fumar. Una vez transcurrido el mes se selecciona a los ganadores, que deben haber estado todo el mes sin fumar, y se les realizan las pruebas pertinentes orientadas a comprobar que no fuman. Para conocer la efectividad del programa, un año después de que acaba se selecciona una muestra aleatoria de los participantes fumadores y se les pregunta por su relación en ese momento con el consumo de tabaco. En concreto, se quiere saber cuál es la “abstinencia mantenida” (es decir, el porcentaje de personas que no fumaron desde el inicio del concurso), al mes y a los 12 meses, y cuál la “abstinencia puntual” a los 12 meses (es decir, el porcentaje de personas que no fuman en el momento de la entrevista, independientemente de que hayan fumado o no, en algún momento del período de tiempo evaluado).

En este trabajo para la estimación de la tasa de recaída se empleó la información del “Deixao e Gaña” del año 2006. A partir del número de participantes en el concurso que habían vuelto a fumar al cabo de un año se estimó el porcentaje de recaídas empleado en los modelos.

### 5.2.2. Modelo SimSmoke

El modelo SimSmoke [13] es uno de los métodos de simulación más utilizados para estudiar la evolución de la prevalencia de consumo de tabaco. Este modelo se divide en varios módulos: un módulo de población y un módulo de tabaco basados en Cadenas de Markov, un módulo de estudio de muertes atribuidas al tabaco y un último módulo de políticas de control de tabaquismo. Se realizó una adaptación de este modelo, con excepción del módulo que estudia las muertes atribuidas al tabaquismo, para predecir la prevalencia de consumo de tabaco.

#### Módulo de población

El modelo poblacional del SimSmoke estudia la evolución de la población de partida en base a nacimientos y muertes, aplicando Cadenas de Markov de primer orden en tiempo discreto. El objetivo de este módulo es conocer qué población abandona el modelo debido a la mortalidad y, a su vez, qué población se incorpora debido a la natalidad.

Las fuentes de datos disponibles permiten conocer la evolución de la población sin necesidad de aplicar un modelo poblacional. En concreto, se utilizaron las siguientes fuentes:

- Instituto Galego de Estatística (IGE): Este organismo realiza proyecciones de población a corto y largo plazo a partir del censo. Dichas proyecciones se pueden obtener segregadas por sexos, por edad y por situación geográfica. En este trabajo se utiliza la proyección de la población gallega de 15 años, por sexos, entre 2011 y 2020.
- Sistema de información sobre conductas de risco na xuventude (SIX), año 2007: Este sistema de información supone una adaptación del SICRI a los jóvenes, siendo su población objetivo los alumnos gallegos matriculados en 4º de ESO, en este caso concreto en el curso escolar 2007-2008. Según la Consellería de Educación se estimó que el tamaño de dicha población, en ese curso, era de unos 22.400 alumnos. La muestra, de 2.124 alumnos, se seleccionó mediante muestreo unietápico de conglomerados con estratificación en las unidades de primera etapa, que eran las aulas de 4º de ESO. Las unidades de análisis fueron los alumnos matriculados en las aulas seleccionadas. De este estudio se extrae el porcentaje de fumadores de 15 años en Galicia.
- Cancer Prevention Study II (CPS-II): Este estudio, llevado a cabo por la American Cancer Society, comenzó en el año 1982. Se trata de un estudio prospectivo de mortalidad en el que se realiza el seguimiento de 1,2 millones de voluntarios que residen en los 50 estados de Estados Unidos, junto con el distrito de Columbia y Puerto Rico. Cada participante realiza un cuestionario confidencial con sus datos personales, historial médico, antecedentes de enfermedades, hábitos alimentarios, consumo de alcohol y

tabaco y actividad física. Durante 24 años de seguimiento de esta cohorte (1982-2006) se han producido 491.188 muertes, de las que se conoce su causa en un 99,3%. A partir de este estudio se extraen las tasas de mortalidad utilizadas en este trabajo, diferenciados por sexo, edad y relación con el tabaco.

### **Módulo de políticas de control de tabaquismo.**

El modelo SimSmoke permite estimar el impacto de las políticas de control de tabaquismo en la evolución de la epidemia tabáquica en un contexto en el que además se tiene en cuenta la dinámica poblacional. Este modelo se desarrolla en 4 módulos, uno de ellos es el "Policy Module". En este módulo se describe el impacto que diferentes políticas, adoptadas con el objetivo de disminuir el consumo de tabaco y la exposición al humo producido durante su combustión, tienen en la tasa de inicio y/o en la tasa de cese. Se consideran distintos tipos de intervenciones, como leyes centradas en el establecimiento de espacios sin humo, incrementos en los impuestos sobre el tabaco, advertencias del impacto en la salud del consumo de cigarrillos en las cajetillas o campañas en los medios de comunicación. Los datos publicados por los autores proceden de revisiones de la literatura especializada o de paneles u opiniones de expertos. Se identificaron cuáles de las políticas incluidas en el SimSmoke habían sido adoptadas en España y se valoró el impacto que tuvieron en Galicia. Se incluyó en el modelo el impacto de los incrementos en los impuestos del tabaco, campañas dirigidas a los jóvenes, advertencias globales sobre salud y las políticas de aire limpio que afectan a los lugares de trabajo, restaurantes y otros.

## Capítulo 6

# Modelo general de simulación

En una primera etapa se ajusta un modelo general basado en CMTD que permite estudiar la evolución de la prevalencia de consumo de tabaco. Se modeliza así su evolución dentro de una población fija, sin mortalidad y sin reemplazo.

El modelo de simulación generado se basa en CMTD como las presentadas en el Capítulo 2. Se considera como distribución inicial de la cadena la correspondiente al año 2006, considerando que cada año constituye una etapa. La CMTD, como se ha visto en el Capítulo 2, queda caracterizada por su distribución inicial, su matriz de transición y un espacio de estados.

### 6.1. Espacio de estados

La definición del espacio de estados puede variar en función de las categorías consideradas. Dentro de cada espacio se puede adaptar la definición de cada estado para conocer la predicción de la prevalencia de tabaco en un determinado grupo. En este caso se han considerado tres espacios de estados distintos, que se describen con detalle en capítulos posteriores.

En el modelo se toman como referencia las prevalencias de consumo de tabaco. De modo que los estados se definen a partir de la clasificación de un individuo según su relación con el consumo del mismo en fumador, nunca fumador y ex-fumador. Estas tres categorías definen el espacio de estados del primer modelo considerado. Son entonces los estados de referencia en el desarrollo de ese primer modelo y de otros modelos posteriores.

Teniendo en cuenta la epidemiología del tabaco en Galicia, en el diseño de los espacios de estados para los distintos modelos se han tenido en cuenta las siguientes hipótesis:

- Los nunca fumadores empiezan a fumar antes de los 25 años. A partir de

esta edad, los nunca fumadores se mantienen como tal el resto de su vida.

- Los fumadores abandonan el consumo a partir de los 25 años. Entre los más jóvenes no se produce abandono.
- Los exfumadores pueden recaer en el consumo. Sin embargo, si el tiempo que llevan sin fumar supera el año, ya no se producen esas recaídas y se convierten en exfumadores de larga duración.
- Los exfumadores de más de 65 años, no van a recaer en el consumo. A esa edad, si un fumador deja de fumar lo hace de forma definitiva.

## 6.2. Determinación del estado de los individuos

Una vez que se plantea el espacio de estados, el siguiente paso es estimar las probabilidades de transición. El método utilizado para realizar estas estimaciones es el de máxima verosimilitud (3.2). Para ello necesitamos conocer, para cada individuo considerado de la encuesta del SICRI 2006, el estado al que pertenece en ese momento (estado actual) y el estado del que proviene (estado anterior) que en este caso hace referencia al estado en el año 2005.

### 6.2.1. Determinación del estado actual

De la encuesta del SICRI del año 2006 se puede obtener la relación de cada individuo de la muestra con respecto al consumo de tabaco. La primera clasificación se basa en la variable “habito3”, que divide a los individuos de la muestra en fumadores, nunca fumadores y exfumadores, lo que coincide con los estados del modelo más sencillo. Para los otros dos modelos se combina esta clasificación con otras variables de la encuesta, como la edad, para asignar a cada individuo uno de los estados de la CMTD. Las preguntas de la encuesta proporcionan también información con respecto a la historia de consumo de cada individuo. En la Figura 6.1 se muestra un extracto de las encuestas.

### 6.2.2. Determinación del estado anterior

Para poder estimar las probabilidades de transición, que conforman la matriz de transición del modelo, es necesario definir para cada individuo de la muestra del SICRI 2006 (del que ya se conoce su estado en ese momento) su estado en la etapa anterior, en el año 2005.

Los datos obtenidos de las preguntas del SICRI, así como los datos complementarios obtenidos del concurso “Déixao e Gaña”, permiten hacer esa clasificación del estado anterior de los individuos de la muestra.



P.8.- ¿A que idade empezou a fumar de forma regular?  
 Aos \_\_\_ anos Nr/Nc 99  
 (Se pregunta a los que alguna vez fumaron)

P.43.- ¿Canto tempo hai que deixou de fumar? LEMBRAR: UNHA SOA RESPOSTA

-Días.....	1	→ P.44.1.	Cantos días? ___	días	→ Pasar a P.45
-Semanas.....	2	→ P.44.2.	Cantas semanas? ___	semanas	→ Pasar a P.45
-Meses.....	3	→ P.44.3.	Cantos meses? ___	meses	→ Pasar a P.45
-Años.....	4	→ P.44.4.	Cantos anos? ___	anos	→ Pasar a P.45
-Ns/Nc.....	9	→ Pasar a P.45			

(Se pregunta a los que actualmente son exfumadores)

Figura 6.1: Extracto del cuestionario

A partir de las preguntas del cuestionario de la Figura 6.1 y de otros datos como la edad, se puede identificar la procedencia de un individuo, por ejemplo, si un fumador proviene de un estado previo como nunca fumador, es decir, si empezó a fumar durante el último año o si proviene de un fumador, es decir, ya fumaba hace un año.

Se puede determinar también si un exfumador proviene de un estado previo como fumador, es decir, dejó de fumar durante el último año, o bien proviene de un exfumador, si ya era exfumador hace un año.

El SICRI no permite saber si un fumador proviene de un exfumador que ha vuelto a fumar durante el último año, lo que sería una recaída. El concurso “Déixao e Gaña” proporciona información del porcentaje de individuos que, siendo exfumadores en la etapa anterior, ya que dejaron de fumar al menos durante el mes del concurso, han vuelto a fumar al cabo de un año.

### 6.3. Distribución inicial

Al conocer la clasificación de cada individuo según su estado actual, que es su estado en el año 2006, se puede conocer ya la distribución inicial de la CMTD. Esto se debe a que se considera el año 2006 como año de partida de la cadena, es decir, como etapa inicial.

A partir de la clasificación de los individuos en los distintos estados del modelo, se estiman las proporciones que conforman la distribución inicial. Se trata entonces de tener para cada estado la probabilidad de que un individuo de la muestra se encuentre en el estado indicado en la etapa cero, año 2006. La suma de las probabilidades de pertenecer a cada estado será uno.

Para calcular los estimadores se debe tener en cuenta el diseño de la muestra que se ha utilizado. El diseño de la muestra, en el caso del 2006, es un diseño bietápico con estratificación en las unidades de primera etapa. Para estimar las proporciones se utiliza un estimador de razón y las unidades muestrales se ponderan por el inverso de la probabilidad de selección. Entonces, el estimador de razón viene dado por:

$$\hat{\Theta} = \frac{\sum_{h_i} W_{h_i} X_{h_i}}{\sum_{h_i} W_{h_i}}$$

Donde  $h$  indica el estrato e  $i$  el individuo,  $X_{h_i}$  es el valor de la característica que se estima (0 – 1) en el individuo  $i$  del estrato  $h$  y  $W_{h_i}$  es el factor de elevación o ponderación del individuo  $i$  del estrato  $h$ . En este caso el valor de la característica se corresponde con la pertenencia de cada individuo a los estados de la CMTD.

## 6.4. Matriz de transición

En este momento se tiene un conjunto de individuos con una clasificación según su estado anterior y su estado actual, relativos a los años 2005 y 2006. Esta clasificación permite estimar las probabilidades de transición que conforman la matriz de transición de la cadena.

### 6.4.1. Estimación de las probabilidades de transición

El método se basa en hacer un recuento, de forma que se tiene para cada uno de los estados el número de individuos que, en el 2005, se encontraban en él. De cada uno de estos grupos se conoce también el número de individuos que, saliendo de ese estado, llegan a cada uno de los estados en 2006. La probabilidad entonces de pasar de un estado  $i$  a un estado  $j$  en una etapa es el cociente entre el número de individuos que estando en el estado  $i$  en el año 2005, pasan al estado  $j$  en el año 2006, y el número de individuos que salen del estado  $i$  en el año 2005 (3.2). Dichas estimaciones, como se ha visto en la Sección 3.1.1., se puede obtener a partir de la tabla de contingencia de las variables “estado anterior” y “estado actual”.

Se debe tener en cuenta, de nuevo, que los datos provienen de una encuesta por muestreo. Por lo tanto, para estimar las probabilidades de transición, que se basan en recuentos y proporciones, se utiliza el procedimiento del estimador de máxima verosimilitud (3.2), pero adaptado al diseño muestral mediante el estimador de razón descrito en el apartado 6.3.

### 6.4.2. Intervalos de confianza

Las probabilidades de transición se estiman a partir de los datos y es posible calcular intervalos de confianza, una vez determinada la distribución de dichas proporciones (que es normal, pero no son independientes unas de otras).

La varianza del estimador de razón utilizado se calcula utilizando una aproximación por linealización de Taylor y, a partir de ella, se obtienen intervalos de confianza del 95 % mediante la expresión:

$$\hat{\Theta} \pm 1.96\sqrt{\hat{V}ar(\hat{\Theta})}$$

Anteriormente se ha visto la forma de obtener también intervalos de confianza para las mismas probabilidades a partir de la distribución asintótica de los estimadores de máxima verosimilitud de las probabilidades de transición (3.3) y mediante técnicas bootstrap (Sección 4.1.1.).

## 6.5. Distribución en la etapa $n$ -ésima

A partir del espacio de estados, la matriz de transición y la distribución inicial, es posible aplicar la teoría de CMTD para calcular la distribución de la cadena en cualquier etapa  $n$ , lo que permite obtener predicciones para cualquier año posterior a 2006. Para cada etapa se obtiene una distribución de los individuos en cada uno de los estados, de forma que es posible conocer la evolución que siguen a lo largo de los años.

En este caso no se utilizan las distribuciones calculadas para hacer predicciones, sino para comprobar el buen funcionamiento del modelo. Al calcular estas distribuciones en diversas etapas se ve la evolución de la población considerada en la etapa inicial, siendo una población fija, sin considerar mortalidad ni natalidad. Esto hace que la población vaya envejeciendo con el paso de los años, lo que permite observar su evolución con respecto al consumo del tabaco, pero impide hacer predicciones a largo plazo. Las predicciones se harán en base a las transiciones calculadas utilizando las fórmulas del modelo SimSmoke (Capítulo 8), en las que sí se observa el comportamiento dinámico de la población.

Para la distribución de la cadena en la etapa  $n$ -ésima se construyen intervalos de confianza basados en técnicas Bootstrap (Sección 4.1.2.).

## 6.6. Comparación con los datos observados

Las primeras distribuciones calculadas son las de las etapas 1, 4 y 5, correspondientes a los años 2007, 2010 y 2011. Se conoce, según el espacio de estados que estemos considerando, la distribución estimada a partir de la encuesta correspondiente del SICRI de los años 2007, 2010 y 2011. Por lo tanto, para estas etapas, es posible comparar los resultados que proporciona el modelo con los observados en el SICRI.

En un primer momento, interesa calcular la distribución en 2007, primera etapa, para comparar los resultados predichos con los resultados del SICRI. De esta forma se puede ver si se producen desajustes muy grandes que indiquen algún fallo en la construcción del modelo.

El año 2011 tiene un interés particular. Se trata del primer año de aplicación de la ley 42/2010 [12], por lo que sirve para comprobar su impacto. Se pueden comparar los resultados en el año 2010 y 2011 para comprobar si se producen cambios destacables entre estos dos años. Dado que el modelo no tiene en cuenta, de momento, modificaciones debidas a las leyes, un buen ajuste del mismo a los datos del SICRI 2011 daría a entender que no se han producido cambios significativos en la prevalencia del tabaco durante ese año. No se habría modificado ni el hábito ni la tendencia de los jóvenes a empezar a fumar, o la de los fumadores a dejar de fumar, por lo que se podría decir que la ley no ha tenido efecto en la población a corto plazo.

Para realizar estas comparaciones se utilizarán los test vistos en la Sección 3.1.3.

# Capítulo 7

## Modelos considerados

Se han considerado tres modelos diferentes de simulación, diferenciándose entre sí en la definición del espacio de estados. El primero de ellos considera un espacio de estados según el cual se divide a la población en nunca fumadores, fumadores y exfumadores. A partir de este primer modelo se desarrollan otros dos en los que se amplía el número de estados y se divide a la población por grupos de edad.

### 7.1. Modelo 1

#### 7.1.1. Espacio de estados

$$\Omega = \{NF, F, ExF\}$$

##### Definición de los estados

- NF: nunca fumadores de cualquier edad.
- F: fumadores de cualquier edad, de los que se conoce la edad a la que empezaron a fumar, siempre que sea de menos de 25 años. Como excepción, se incluyen todos los fumadores de más de 65 años, independientemente de la edad a la que empezaran a fumar. Se incluye en esta definición a los menores de 25 años que se declaran exfumadores, ya que en esas edades solo se consideran nunca fumadores o individuos que han fumado alguna vez.
- ExF: exfumadores de más de 25 años, siempre que hayan dejado de fumar después de los 25 años.

El estado nunca fumadores es un estado transitorio y comunica con el de fumadores. A su vez, los estados fumadores y exfumadores intercomunican y son recurrentes.

### 7.1.2. Transiciones

A partir del espacio de estados se definen también las transiciones entre ellos. La estimación de las probabilidades de transición se basa en los cambios permitidos en los individuos entre dos etapas.

En la Figura 7.1 se representa el espacio de estados con las transiciones permitidas entre ellos.

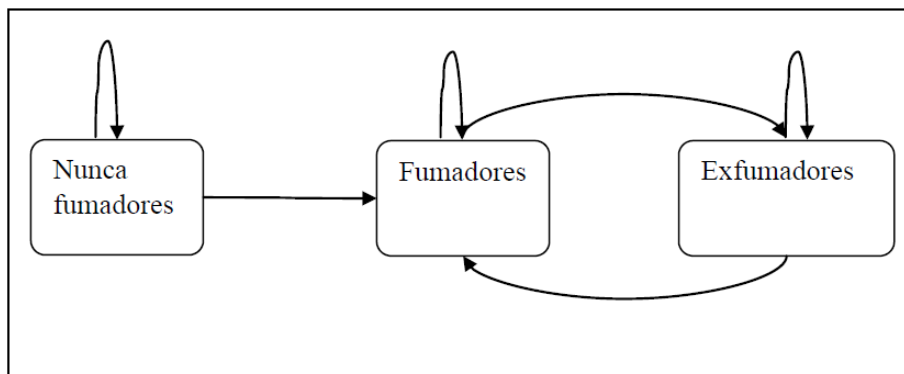


Figura 7.1: Representación del espacio de estados y transiciones en el Modelo 1.

#### Transiciones posibles

- $NF \rightarrow NF$ : nunca fumadores en un año que no se inician en el hábito de fumar al año siguiente (durante una etapa).
- $NF \rightarrow F$ : INICIO. Nunca fumadores que han empezado a fumar durante el último año, siempre que tengan menos de 25 años.
- $F \rightarrow F$ : fumadores que conservan el hábito de una etapa a la siguiente, aunque puede que hayan dejado de fumar temporalmente durante los meses intermedios entre dos etapas.
- $F \rightarrow ExF$ : CESE. Fumadores que han dejado de fumar durante el último año y se mantienen como no fumadores, siempre que tengan al menos 25 años.
- $ExF \rightarrow F$ : RECAÍDA. Exfumadores que, llevando menos de un año sin fumar, han vuelto a fumar. Deben tener entre 25 y 65 años para que se permita esta transición.
- $ExF \rightarrow ExF$ : exfumadores que siguen sin fumar.

### Transiciones no permitidas en una etapa

- $NF \rightarrow ExF$ : para pasar de nunca fumador a exfumador tienen que haber pasado por el estado fumador al menos una vez. Se necesitan 2 etapas.
- $F \rightarrow NF$ :  $NF$  son los nunca fumadores, por lo que si alguna vez fumaron solo pueden seguir siendo fumadores o, si lo dejan, pasar a ser exfumadores.
- $ExF \rightarrow NF$ : Se es exfumador si alguna vez se ha fumado, por lo que no existe la posibilidad de volver a ser nunca fumador.

Además de estas restricciones, tanto en la definición de los estados como en las transiciones permitidas y no permitidas, se considera también que, debido a los datos disponibles para las recaídas y a las características del proceso de abandono del hábito, dentro del estado “exfumadores” solo se permite que sufran recaídas aquellos que llevan menos de un año sin fumar.

Entonces la transición duraría dos etapas y sería:

$$\text{Fumador} \rightarrow \text{Exfumador} \rightarrow \text{Fumador}$$

Además, las recaídas se producen entre los mayores de 25 años y los menores de 65.

#### 7.1.3. Determinación del estado actual

En este modelo el estado actual se determina de forma muy sencilla a partir de la clasificación que hace la variable “habito3” en nunca fumadores, fumadores y exfumadores. Se parte del SICRI 2006, del que se excluye a los siguientes individuos:

- Fumadores de entre 25 y 65 años que no han respondido a la pregunta: “¿A que idade empezou a fumar de forma regular?” ( $n=46$ )
- Entre los fumadores que sí respondieron, se elimina a los que empezaron a fumar con más de 25 años, independientemente de su edad actual, salvo en los mayores de 65. ( $n=61$ )
- Exfumadores que, teniendo más de 25 años, no han respondido a la pregunta de: “¿Canto tempo hai que deixou de fumar?” ( $n=271$ )
- Conociendo los años que llevan sin fumar, se elimina también a los que hayan dejado de fumar definitivamente con menos de 25 años, sabiendo que su edad actual es de más de 25 años. ( $n=107$ )

Se parte de un tamaño de muestra de  $n=7841$ , por lo que quedan  $n=7356$  individuos para este análisis. De ellos, 3644 son hombres y 3712 son mujeres.

#### 7.1.4. Determinación del estado actual

Para estimar las probabilidades de transición es necesario determinar el estado del que proceden los individuos que se han clasificado en la etapa actual, 2006. Se razona la asignación del estado anterior, 2005, a partir del estado actual y de variables auxiliares obtenidas del SICRI.

En la Figura 7.2 se muestran las transiciones entre dos etapas.

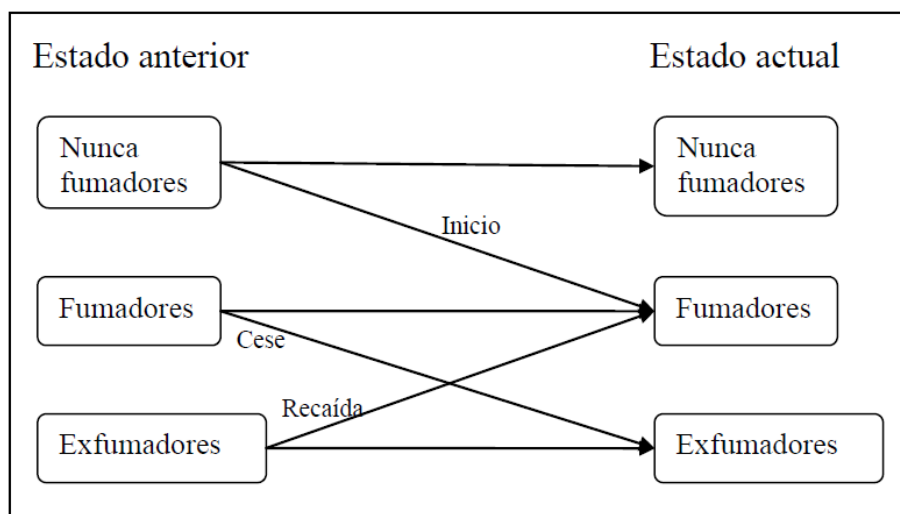


Figura 7.2: Representación de las transiciones entre dos etapas en el Modelo 1.

Los que son nunca fumadores en el estado actual, solo pueden proceder del estado nunca fumadores.

Los fumadores, como se ve en la Figura 7.2, pueden proceder de nunca fumadores, de fumadores o de exfumadores. En principio no se considera la posibilidad de que procedan de exfumadores, que sería el caso de las recaídas, ya que los datos que se tienen de esa transición no proceden de la encuesta. Las clasificaciones relativas a las recaídas se realizan en último lugar. Por lo tanto, en este momento, los fumadores pueden ser de dos tipos: los que empezaron a fumar durante el último año y los que ya fumaban en el año anterior.

En el primero de los casos el fumador procedería de un nunca fumador. Para determinar a los individuos que se encuentran en esta situación se compara su edad actual con la edad a la que empezaron a fumar de forma regular. Si la edad que tiene el fumador durante el último año coincide con la edad de inicio en el hábito, procede de un nunca fumador. En otro caso, procede de un fumador.



Nótese que cada persona habrá cumplido años durante el último año desde la fecha de realización de la encuesta, por eso es necesario comparar la edad de inicio con la edad actual y con la edad actual, menos un año.

En el caso de los exfumadores la información que se conoce es el tiempo que llevan sin fumar. Se distingue, por un lado, a los que llevan menos de un año sin fumar y, por lo tanto, en la etapa anterior eran fumadores. Por otro lado, los que lleven más de un año sin fumar ya proceden de exfumadores.

### Determinación de las recaídas

Falta por determinar a los individuos que sufrieron una recaída. Para ello la información que se obtiene del concurso “Déixao e Gaña” nos permite conocer el porcentaje de fumadores al cabo de un año entre los que dejaron de fumar durante el concurso. Es decir, el porcentaje de los que recaen en el hábito después de haber sido exfumadores, que es de un 50,1%. Estos exfumadores se corresponden con los que llevan sin fumar menos de un año, que son los que pueden sufrir recaídas. El proceso se representa en la Figura 7.3, y tiene los siguientes pasos:

- Se calcula el número de exfumadores que actualmente llevan sin fumar al menos un año pero menos de dos, es decir, responden que llevan sin fumar 1 año (se les llamará ExF1). Se contabilizan en  $\#(\text{ExF1})$ , que es un dato conocido. En la etapa anterior llevarían sin fumar menos de un año, por lo que su estado sería equivalente a los participantes del “Déixao e Gaña”.
- Siendo exfumador de menos de un año en la etapa anterior (se les llamará ExF0) puede pasar a ser fumador (sufrir una recaída) o seguir sin fumar. Si se sabe que recaen el  $r\%$  de los exfumadores de menos de un año (ExF0), son el  $(100 - r)\%$  los que siguen sin fumar (ExF1).  $\#(\text{ExF0})$  es un dato desconocido que no se puede determinar a partir de los datos del SICRI.
- Como se conoce  $r$  y  $\#(\text{ExF1})$ , que es el  $(100 - r)\%$  de  $\#(\text{ExF0})$ , se puede calcular el número de exfumadores de menos de un año en la etapa anterior,  $\#(\text{ExF0})$ , mediante

$$\#(\text{ExF0}) = \frac{\#(\text{ExF1}) * 100}{100 - r}$$

De esta forma se determina el número de exfumadores susceptibles de sufrir una recaída.

- Se calcula el número de recaídas,  $R$ , calculando el  $r\%$  de  $\#(\text{ExF0})$
- Entre los que en un principio se clasificaron como fumadores que proceden de fumadores, que tienen más de 25 años y menos de 65, se sortea aleatoriamente el número de recaídas,  $R$ .

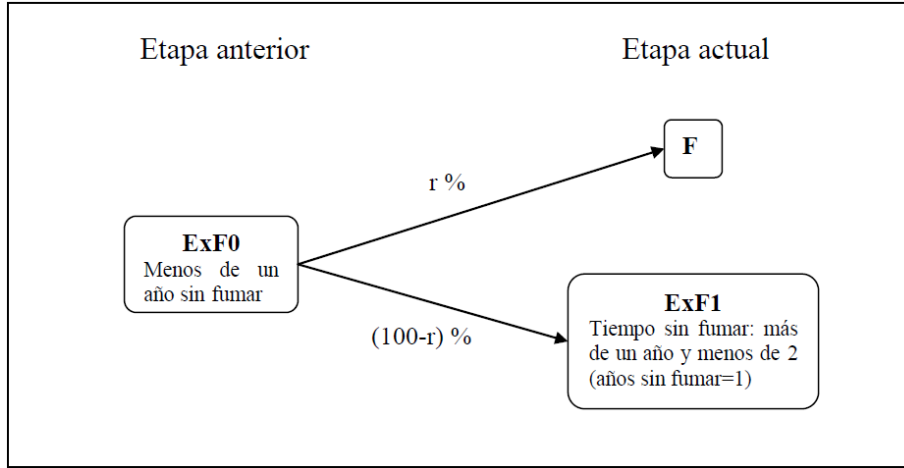


Figura 7.3: Representación de la transición relativa a las recaídas.

- A los que resultan del sorteo se les cambia el estado anterior de fumadores por el de exfumadores. De esta forma la transición que realizan es de exfumadores a fumadores y, por lo tanto, se corresponden con las recaídas.

A partir de este momento se siguen los pasos indicados en el modelo general, calculando las distribuciones en las etapas correspondientes y comparándolas con la distribución observada en las encuestas del SICRI. En el capítulo de resultados se muestran las predicciones obtenidas, así como los resultados del test de razón de verosimilitudes que compara las distribuciones del modelo con las de la encuesta, los intervalos de confianza y las intervenciones.

## 7.2. Modelo 2

### 7.2.1. Espacio de estados

En este segundo modelo se divide a los exfumadores en dos estados, exfumadores de corta y de larga duración. Los de corta duración (ExF) son los que llevan sin fumar menos de un año. Los que lleven sin fumar más de un año son exfumadores de larga duración (ExFL). Los fumadores y nunca fumadores son estados análogos a los del primer modelo.

$$\Omega = \{NF, F, ExF, ExFL\}$$

Este modelo se obtiene de forma casi inmediata del anterior. En el modelo anterior, dentro del mismo estado de exfumadores, se clasificaba a los individuos de distinta forma en función del tiempo que llevaran sin fumar, permitiendo recaídas o no. Parece lógico, entonces, dividir este estado en dos. De esta forma,

los exfumadores que llevan más de un año sin fumar tienen determinado su estado anterior y siguiente, sin variaciones debidas a las recaídas, ya que son los exfumadores de corta duración los que pueden sufrir recaídas. Se puede entender que los exfumadores de larga duración son exfumadores permanentes, ya no cambian de estado. En la Figura 7.4 se muestran los estados y las posibles transiciones entre ellos.

En esta ocasión se tiene un estado absorbente, el de exfumadores de larga duración. Nunca fumador comunica con fumador, fumador y exfumador de corta duración intercomunican y exfumador de corta duración comunica con exfumador de larga duración.

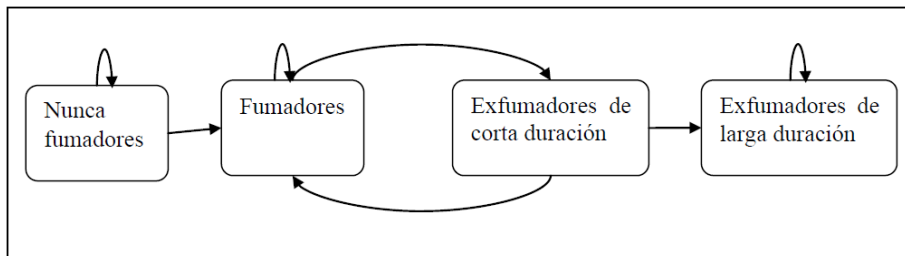


Figura 7.4: Representación del espacio de estados y transiciones en el Modelo 2.

### 7.2.2. Transiciones

Para este modelo se siguen las mismas consideraciones que en el Modelo 1. Los cambios se producen en los estados nuevos. Los exfumadores de corta duración son los que pueden recaer o, si siguen sin fumar, pasar a ser exfumadores de larga duración. De esta forma un individuo no puede estar en el mismo estado de exfumador de corta duración durante dos etapas consecutivas.

Los exfumadores de larga duración constituyen un estado absorbente ya que, una vez que se alcanza no se vuelve a salir de él. Además, en el proceso de dejar de fumar se requieren al menos dos etapas para llegar a ser exfumador de larga duración ya que es obligatorio el paso previo por el estado de exfumadores de corta duración. En la Figura 7.5 se muestran las transiciones permitidas en una etapa.

Las recaídas en este modelo se obtienen de la misma forma que en el Modelo 1, salvo que en esta ocasión ya habremos determinado el conjunto de individuos que pueden sufrir una recaída, los exfumadores de corta duración.

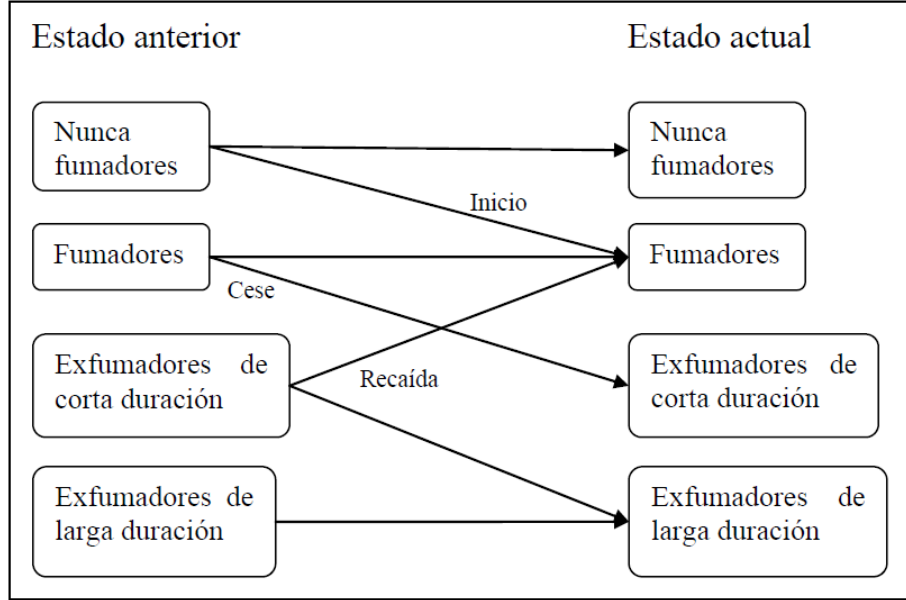


Figura 7.5: Representación de las transiciones entre dos etapas en el Modelo 2.

### 7.3. Modelo 3

#### 7.3.1. Espacio de estados

Basándose en la clasificación del Modelo 2 en nunca fumadores, fumadores, exfumadores de corta y de larga duración, se crean nuevos estados introduciendo los grupos de edad. En las encuestas del SICRI se consideran los siguientes grupos:

Grupos de edad: 1:  $[16, 25)$ , 2:  $[25, 45)$ , 3:  $[45, 65)$ , 4:  $[65, \infty)$

Por lo que el espacio de estados es:

$$\Omega = \{NF1, NF2, NF3, NF4, F1, F2, F3, F4, \\ ExF2, ExF3, ExFL2, ExFL3, ExF4\}$$

Para los grupos de edad intermedios se repiten los cuatro estados del Modelo 2. En el primer grupo de edad se tienen solo dos estados, los nunca fumadores y los que alguna vez han fumado. En el grupo cuarto se considera un único estado de exfumadores. En la Figura 7.6 se muestra el espacio de estados y transiciones de este modelo.

Los estados NF4 y ExF4 son estados absorbentes, mientras que los demás

son transitorios. En la Figura 7.6 se pueden ver los estados que comunican con otros e intercomunican entre sí.

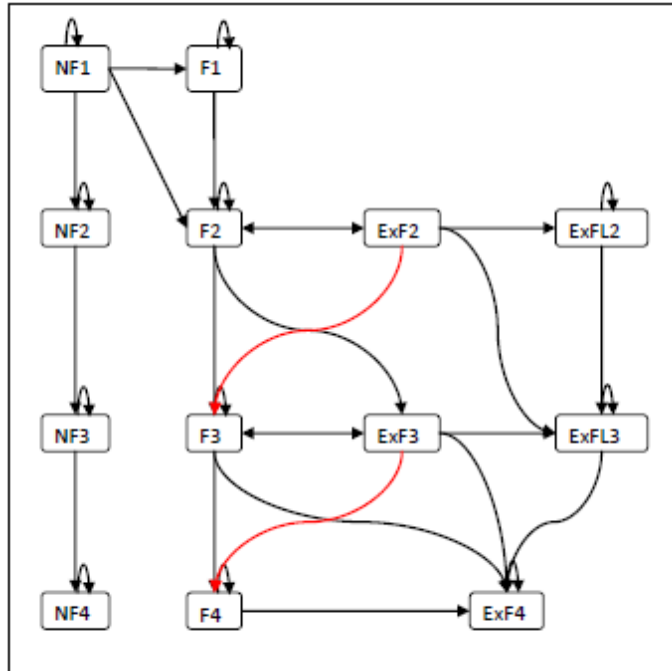


Figura 7.6: Representación del espacio de estados y transiciones en el Modelo 3.

### 7.3.2. Transiciones

Se consideran transiciones de los siguientes tipos:

- Debidas a la edad: en muchos casos la transición se produce de un estado al estado análogo en el grupo de edad siguiente. Estas transiciones se deben al incremento de la edad en los individuos al pasar de una etapa a la siguiente.
- Debidas a un cambio de estado: independientemente de los grupos de edad, se producen transiciones análogas a los modelos anteriores.
- Combinadas: se trata de aquellas transiciones en la que se produce a la vez un cambio del grupo de edad y un cambio en el estado.

En este modelo se producen recaídas solamente en los estados de exfumadores de corta duración de los grupos de edad intermedios, que se corresponden con

los mayores de 25 años y menores de 65. En la Figura 7.7 se muestran las transiciones permitidas en una etapa.

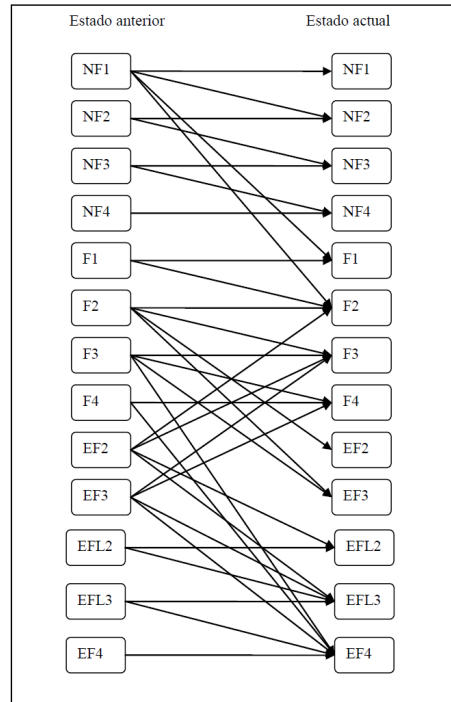


Figura 7.7: Representación de las transiciones entre dos etapas en el Modelo 3.

## Capítulo 8

# Modelo de predicción

En este capítulo se desarrolla un modelo de predicción. Basándose en los modelos de CMTD anteriores se busca hacer predicciones incorporando la dinámica poblacional para evitar que la población incluida en el estudio envejezca sin considerar la mortalidad. Se parte del Modelo 3, en el que se consideran 13 estados en función de la relación con el tabaco y de los grupos de edad y se busca obtener el número de individuos en cada uno de los 13 estados en los años comprendidos entre 2010 y 2020.

Se necesita conocer las probabilidades de transición relativas al Modelo 3. Para ello se actualiza dicho modelo con los datos del SICRI 2010. De esta forma se calcula una nueva matriz de transiciones y se considera como distribución inicial la relativa a dicho año 2010. Las transiciones de este modelo nos proporcionan la información relativa al envejecimiento de la población, por los cambios de grupo de edad, y a la evolución de la relación de los individuos con el tabaco.

Conociendo la distribución inicial en 2010, dada por la CMTD, y el total poblacional en Galicia en el mismo año se obtiene el número de individuos de la población clasificados en cada uno de los estados considerados. A partir de estos valores, aplicando las fórmulas que se verán a continuación, se calcula el número de individuos en cada estado para cada año hasta 2020, obteniendo las prevalencias asociadas a la población de cada año.

En la formulación mediante la que se obtienen las predicciones, además de las transiciones entre estados, se incorpora información relativa a la natalidad y a la mortalidad. La edad de partida para el estudio es de 16 años. Esto supone que a partir de la primera etapa del modelo la población envejece un año en cada etapa y por lo tanto se extinguen los individuos de edad menor. Para suplir este envejecimiento se debe incorporar cada año a la población de 15 años del año anterior. De esta forma se controla la dinámica poblacional relativa a la natalidad. Además, para el primer grupo de edad (16-24) los estados considerados son dos, nunca fumadores y los individuos que alguna vez fumaron (nos

referiremos a ellos como fumadores en general). Por lo tanto, la incorporación de la población al modelo se produce en esos dos estados. Se dispone de la proporción de nunca fumadores entre los jóvenes de 15 años procedente del estudio SIX. Los datos de población considerados provienen de la proyección del censo realizada por el IGE por edad y sexo hasta el año 2020. En base a estos datos se procede de la siguiente forma:

- Para cada año posterior a 2010 se obtiene el número de individuos de 15 años del año anterior.
- Se aplica la proporción de nunca fumadores sobre dicha población.
- Para el año considerado se añade el número de nunca fumadores y de fumadores de 16 años.

El modelo necesita incorporar, además, las tasas de mortalidad correspondientes. Se utilizan tasas de mortalidad por grupos de edad, por sexo y por la relación con el tabaco. De esta forma se distingue la tasa de mortalidad de los fumadores de la tasa de mortalidad de no fumadores (se incluyen en esta definición tanto los nunca fumadores como los exfumadores). Además, en el grupo de edad más joven (16-24) la tasa de mortalidad se asume nula.

En base a los datos anteriores se obtiene el número de individuos en cada uno de los 13 estados para cada año entre 2011 y 2020, partiendo de la distribución de la población en el año 2010. Para ello se utilizan las fórmulas mostradas a continuación. Dichas fórmulas, que proceden del modelo SimSmoke [13] se basan en que el número de individuos en un determinado estado en un año se corresponde con el número de individuos que, durante la transición desde año anterior, han llegado a dicho estado y que han “sobrevivido”. Se dice que han “sobrevivido” ya que, además de aplicar las probabilidades de transición correspondientes, se aplican las tasas de mortalidad asociadas a cada caso.

En la formulación se distinguen los tres estados principales en la relación con el tabaco, los nunca fumadores, fumadores y exfumadores. Dentro de cada uno se adaptan las características necesarias, dadas por los distintos grupos de edad y sus transiciones asociadas.

## 8.1. Nunca fumadores

Se consideran los años  $t = 2010, \dots, 2020$ . El número de nunca fumadores del primer grupo de edad (16-24 años) en el año  $(t + 1)$  se corresponde con el número de nunca fumadores del mismo grupo de edad del año  $t$  que, durante esa transición no cambian de estado, a los que se añade el número de nunca fumadores de 16 años correspondiente.



$$NF_1^{t+1} = NF_1^t \mathbb{P}(NF_1 \rightarrow NF_1) + NF_{15}^t.$$

Para los grupos de edad  $i = 2, \dots, 4$  se tiene la siguiente fórmula:

$$NF_{i+1}^{t+1} = NF_{i+1}^t (1 - M(NF_{i+1})) \mathbb{P}(NF_{i+1} \rightarrow NF_{i+1}) + NF_i^t (1 - M(NF_i)) \mathbb{P}(NF_i \rightarrow NF_{i+1}).$$

Se denota por  $M(j)$  a la tasa de mortalidad asociada al estado  $j$  y por  $\mathbb{P}(i \rightarrow j) = p_{ij}$  la probabilidad de transición entre los estados  $i$  y  $j$ .

De esta forma, para los grupos de edad 2, 3 y 4, los nunca fumadores de un año se obtienen a partir de los nunca fumadores del año anterior, tanto del mismo grupo de edad como del anterior, que no han empezado a fumar. En todas las transiciones de los grupos de edad 2, 3 y 4 se incorporan las tasas de mortalidad asociadas.

## 8.2. Fumadores

Los fumadores del primer grupo de edad en el año  $(t + 1)$  proceden de los fumadores del año  $t$  que mantienen su estado y de los nunca fumadores del primer grupo de edad que han empezado a fumar durante esa transición. Se incorporan los fumadores de 16 años.

$$F_1^{t+1} = F_1^t \mathbb{P}(F_1 \rightarrow F_1) + NF_1^t \mathbb{P}(NF_1 \rightarrow F_1) + F_{15}^t.$$

En los grupos siguientes los fumadores proceden de fumadores que siguen fumando, tanto del mismo grupo de edad como del anterior, y de exfumadores que han recaído en el hábito. En cualquier caso se consideran aquellos individuos que no han muerto, puesto que se le aplican las tasas de mortalidad de cada grupo de edad.

Fumadores del grupo 2:

$$F_2^{t+1} = F_2^t (1 - M(F_2)) \mathbb{P}(F_2 \rightarrow F_2) + F_1^t \mathbb{P}(F_1 \rightarrow F_2) + ExF_2^t (1 - M(NF_2)) \mathbb{P}(ExF_2 \rightarrow F_2).$$

Los fumadores del grupo 2 proceden de su mismo grupo de edad o del grupo de edad anterior.

Fumadores del grupo 3:

$$F_3^{t+1} = F_3^t (1 - M(F_3)) \mathbb{P}(F_3 \rightarrow F_3) + F_2^t (1 - M(F_2)) \mathbb{P}(F_2 \rightarrow F_3) + ExF_3^t (1 - M(NF_3)) \mathbb{P}(ExF_3 \rightarrow F_3) + ExF_2^t (1 - M(NF_2)) \mathbb{P}(ExF_2 \rightarrow F_3).$$

En los grupos 2 y 3 se tiene el estado de exfumadores de corta duración, en los que se permiten recaídas. Por lo tanto los fumadores en el año  $t+1$  son fumadores que siguen fumando o bien exfumadores que han recaído; en ambos casos se aplica la tasa de mortalidad asociada.

Fumadores del grupo 4:

$$F_4^{t+1} = F_4^t(1 - M(F_4))\mathbb{P}(F_4 \rightarrow F_4) + F_3^t(1 - M(F_3))\mathbb{P}(F_3 \rightarrow F_4) + \\ ExF_3^t(1 - M(NF_3))\mathbb{P}(ExF_3 \rightarrow F_4).$$

En el grupo 4 se tiene una única categoría de exfumadores, en la que no se permiten las recaídas. Por lo tanto los fumadores provienen de otros fumadores o de exfumadores de corta duración del grupo 3 que han recaído y a la vez han cambiado de grupo de edad durante la transición entre los años considerados.

### 8.3. Exfumadores

Los exfumadores de corta duración de los grupos 2 y 3 proceden de los fumadores que han dejado de fumar durante esa transición y que no han muerto.

$$ExF_2^{t+1} = F_2^t(1 - M(F_2))\mathbb{P}(F_2 \rightarrow ExF_2).$$

Dado que en el grupo 1 no se tiene una categoría de exfumadores, los exfumadores de corta duración del grupo 2 solo pueden proceder de fumadores de su mismo grupo de edad.

$$ExF_3^{t+1} = F_3^t(1 - M(F_3))\mathbb{P}(F_3 \rightarrow ExF_3) + F_2^t(1 - M(F_2))\mathbb{P}(F_2 \rightarrow ExF_3).$$

En el grupo 3 los exfumadores pueden proceder tanto de fumadores del grupo 3 como del grupo 2 que han cambiado de grupo de edad en esa transición.

Los exfumadores de larga duración de los grupos 2 y 3 proceden de exfumadores de corta duración que no han vuelto a fumar y de los exfumadores de larga duración de año anterior, que ya no cambian de estado en relación con el tabaco.

$$ExFL_2^{t+1} = ExF_2^t(1 - M(NF_2))\mathbb{P}(ExF_2 \rightarrow ExFL_2) + \\ + ExFL_2^t(1 - M(NF_2))\mathbb{P}(ExFL_2 \rightarrow ExFL_2).$$

$$ExFL_3^{t+1} = ExF_3^t(1 - M(NF_3))\mathbb{P}(ExF_3 \rightarrow ExFL_3) + \\ ExF_2^t(1 - M(NF_2))\mathbb{P}(ExF_2 \rightarrow ExFL_3) + \\ ExFL_3^t(1 - M(NF_3))\mathbb{P}(ExFL_3 \rightarrow ExFL_3) + \\ ExFL_2^t(1 - M(NF_2))\mathbb{P}(ExFL_2 \rightarrow ExFL_3).$$

Los exfumadores del grupo 4, dado que es la única categoría de exfumadores de ese grupo, proceden de fumadores que dejan de fumar y exfumadores de distintas categorías que siguen sin fumar.

$$\begin{aligned}
 Exf_4^{t+1} = & ExF_4^t(1 - M(NF_4))\mathbb{P}(ExF_4 \rightarrow ExF_4) + \\
 & ExFL_3^t(1 - M(NF_3))\mathbb{P}(ExFL_3 \rightarrow ExF_4) + \\
 & ExF_3^t(1 - M(NF_3))\mathbb{P}(ExF_3 \rightarrow ExF_4) + \\
 & F_4^t(1 - M(F_4))\mathbb{P}(F_4 \rightarrow ExF_4) + \\
 & F_3^t(1 - M(F_3))\mathbb{P}(F_3 \rightarrow ExF_4).
 \end{aligned}$$

Los exfumadores del grupo 4 pueden proceder de su mismo grupo de edad o del anterior.



## Capítulo 9

# Efecto de las intervenciones

Existen actualmente varias intervenciones que afectan al consumo de tabaco. Por una parte la venta de tabaco está sujeta a unos impuestos especiales, al margen del IVA, como los que afectan a las bebidas alcohólicas. Estos impuestos son susceptibles de incrementarse, tanto para aumentar la recaudación por parte del Estado como para reducir la venta de tabaco. Por otra parte, se han venido desarrollando campañas que buscan concienciar a la gente del peligro que supone fumar y que tienen el objetivo de disminuir el inicio y favorecer el cese. Además, en los últimos años se han aprobado distintas leyes de aire limpio, que restringen ciertos espacios al humo del tabaco.

Se considera que, siguiendo los modelos desarrollados en los capítulos anteriores, las intervenciones afectan a las transiciones entre los diferentes estados. En concreto, a las transiciones que marcan el inicio y el cese en el hábito de fumar. El cambio en las probabilidades de transición correspondientes se mide en términos de un porcentaje de cambio estimado ( $PC$ ), basado en estudios empíricos y en la opinión de un panel de expertos.

Las diferentes intervenciones tienen objetivos distintos, por lo que están dirigidas a un determinado sector de la población o, en algunos casos, a la población en general. El público objetivo de las distintas medidas hace que el cambio producido no afecte siempre a las mismas transiciones. Por ejemplo, medidas dirigidas a los jóvenes afectarán a la probabilidad de transición que marca el inicio en el hábito. En esas edades no se producen cambios en el cese del hábito ya que, debido a las características estudiadas en los modelos anteriores, no se contempla la posibilidad de tener exfumadores de menos de 25 años. De la misma forma, si el público objetivo es la población de mediana edad, el cambio se producirá en la probabilidad de dejar de fumar pero no en la de inicio ya que, a esas edades, se establece el supuesto de que no se empieza a fumar.

Las políticas que se tienen en cuenta, con su correspondiente porcentaje de cambio y población afectada son las siguientes [13]:

- Impuestos especiales sobre la venta de tabaco.  
Afectan a toda la población. El porcentaje de cambio entre los jóvenes (16-24) es de un 0.3 % sobre el inicio. A partir de 25 años afecta en un 0.15 % al cese. En el mejor de los casos estos porcentajes pueden alcanzar un 0.6 % y un 0.25 % respectivamente.
- Campañas dirigidas a los jóvenes.  
Afectan a la probabilidad de inicio en un 2.5 %. Realizar un mayor esfuerzo en estas campañas, aumentando su publicidad y su impacto, puede incrementar el porcentaje de cambio hasta un 25 %
- Advertencias globales sobre la salud.  
Están dirigidas a toda la población. Afectan en un 1 % tanto al inicio como al cese. En el mejor de los casos, utilizando la mayor difusión posible para estas campañas, puede llegar a un porcentaje de cambio de un 2 % en el cese y de un 6 % en el inicio.
- Políticas de aire limpio. Se tiene en cuenta la ley 42/2010 [12], que entró en vigor en enero de 2011, por lo que interviene en la primera transición del modelo de predicción. El porcentaje de cambio se mide según las distintas áreas afectadas:
  - Lugares de trabajo. Afecta en un 6 % al cese, excluyendo a la población de más de 65 años (grupo de edad 4).
  - Hostelería. Afecta, en global, en un 2 % tanto al inicio como al cese.
  - Otros. Un 1 % añadido tanto al inicio como al cese. Un aumento en las restricciones y en la publicidad de las leyes puede incrementar el porcentaje de cambio en un 0.05 %.

Estos porcentajes de cambio afectan, como se ha indicado, a las probabilidades de transición entre los diferentes estados considerados. En concreto, se modifican las probabilidades de transición que marcan el inicio y el cese en el hábito de fumar, lo que supone modificar la matriz de transiciones de la CMTD considerada. Se utiliza el espacio de estados del Modelo 3 (Sección 7.3.), dado que en él se dividen los estados en función del grupo de edad. De esta forma se pueden identificar las probabilidades que se alteran en función de las políticas en activo. Dichas modificaciones se realizan de la siguiente forma:

$$\mathbb{P}(\text{inicio en el grupo de edad 1}) * (1 - PC_i).$$

$$\mathbb{P}(\text{cese en los grupos 2,3 y 4}) * (1 + PC_j).$$

La probabilidad de inicio en el hábito, que se produce sólo entre los más jóvenes, se reduce según el porcentaje correspondiente a la política  $i$  ( $PC_i$ ). De la misma forma, la probabilidad de dejar de fumar para los grupos de edad a partir de 25 años, aumenta en función del porcentaje de cambio asociado a la intervención  $j$  ( $PC_j$ ).

En la Sección 2.1.2. se ha visto que la suma de las filas de la matriz de transición de una CMTD es uno. Al modificar algunas de las probabilidades de transición se incumple esta condición, por lo que es necesario reajustar el resto de probabilidades para mantener las características de la CMTD.

Por lo tanto, la reducción en la probabilidad de inicio para el grupo de edad 1, supone un aumento en la probabilidad de seguir siendo nunca fumador del mismo grupo de edad. Dicho aumento es exactamente la reducción asociada a las intervenciones.

De la misma forma, para cada grupo de edad, un aumento en la probabilidad de dejar de fumar supone una reducción de la misma magnitud en la probabilidad de seguir fumando.

Se han indicado diferentes intervenciones, cada una con un objetivo y un porcentaje de cambio estimado. Estas intervenciones suelen actuar en conjunto o aplicarse en el mismo período de tiempo. Esto implica que en un determinado año pueden estar en activo varias de las políticas vistas. Cuando esto ocurre se asume un efecto multiplicativo en el porcentaje de cambio. Por ejemplo, si están en activo las políticas  $i$  y  $j$ , con porcentajes de cambio asociados  $PC_i$  y  $PC_j$ , respectivamente, el efecto sobre las probabilidades de transición será:

$$(1 - PC_i) * (1 - PC_j) \text{ o bien, } (1 + PC_i) * (1 + PC_j).$$

Se realizará la modificación de las probabilidades de transición del Modelo 3 (Sección 7.3.) utilizado posteriormente en el modelo de predicción. Se obtendrán nuevas predicciones para los años 2010-2020 teniendo en cuenta las nuevas probabilidades de transición. En este caso se aplicarán todas las políticas descritas anteriormente, considerando que su efecto se mantiene constante a lo largo de los años en los que se realiza la predicción de prevalencias. Se consideran en principio los porcentajes de cambio normales, comparando su efecto con el escenario correspondiente a los mayores porcentajes posibles. De esta forma se puede ver si un mayor esfuerzo en las intervenciones consideradas supone un cambio significativo en las prevalencias.





Parte III

Resultados



## Capítulo 10

# Aplicación 2006-2011

En este capítulo se muestran los resultados obtenidos de la aplicación del Modelo 1 (Sección 7.1.) a los datos del SICRI. Se toma como etapa inicial el SICRI 2006 y se calcula la distribución en las etapas correspondientes a los años 2007, 2010 y 2011. Se muestran en primer lugar los resultados globales, para luego comparar los resultados divididos por sexos.

### 10.1. Modelo 1. Resultados globales

En este modelo se consideran tres estados: Nunca Fumador (NF), Fumador (F) y Exfumador (ExF). En primer lugar se calcula la matriz de transiciones en base a las observaciones, utilizando los estimadores de máxima verosimilitud. Se muestra en el Cuadro 10.1 dicha matriz junto con sus intervalos de confianza al 95 % calculados según el estimador muestral (Sección 6.4.2.), la distribución asintótica de las estimaciones (3.3.) y el método bootstrap (4.1.1.), respectivamente.

Se obtiene también la distribución inicial de la cadena, es decir, la distribución de la población estimada en el SICRI entre los tres estados considerados en el año 2006.

La distribución inicial correspondiente a este modelo es:

$$P^{(0)} = (0.587, 0.229, 0.184)$$

Se aprecia que la mayoría de la población está situada en el estado de nunca fumadores. Además, hay más fumadores que exfumadores, aunque los niveles son similares.

	NF		F		ExF	
NF	0.994	(0.992,0.995) (0.991,0.996) (0.992,0.996)	0.006	(0.005,0.008) (0.004,0.009) (0.004,0.008)	0,000	
F	0.000		0.944	(0.931,0.956) (0.933,0.955) (0.935,0.953)	0.056	(0.044,0.069) (0.045,0.067) (0.047,0.065)
Exf	0.000		0.055	(0.041,0.070) (0.043,0.068) (0.043,0.065)	0.945	(0.930,0.959) (0.932,0.957) (0.935,0.957)

Cuadro 10.1: Matriz de transiciones del Modelo 1 con intervalos de confianza para las probabilidades de transición clásicos, asintóticos y bootstrap, respectivamente.

El siguiente paso es obtener la distribución en distintas etapas. En el Cuadro 10.2 se muestra el resultado obtenido en la primera etapa, año 2007. Se calcula la distribución de la cadena en esa etapa, con sus correspondientes intervalos de confianza obtenidos mediante técnicas bootstrap. Se muestra además el valor obtenido a partir del SICRI del mismo año.

Año 2007	NF	F	ExF
Distribución modelo	0.583	0.230	0.187
Intervalo de confianza	(0.582,0.585)	(0.226,0.234)	(0.182,0.190)
Distribución SICRI	0.552	0.262	0.185

Cuadro 10.2: Distribución en el año 2007 dada el Modelo 1, con su intervalo de confianza al 95 % y distribución SICRI 2007.

En una etapa no es posible apreciar cambios notables entre las prevalencias, por lo que las distribuciones en los años 2006 y 2007 son muy similares. Destaca el buen ajuste que tiene la distribución obtenida con el modelo en 2007 comparada con la distribución obtenida del SICRI. La mayor diferencia es de un 3%. En el caso del estado de exfumadores la diferencia es inapreciable, como se puede ver en la Figura 10.1.

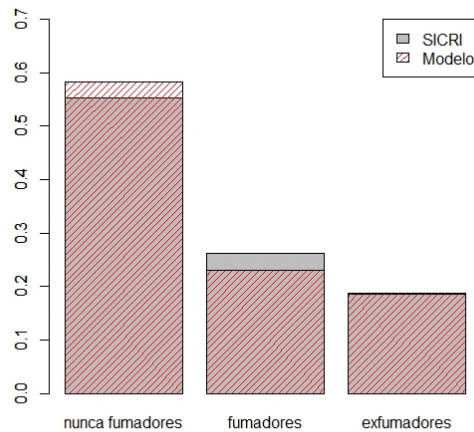


Figura 10.1: Comparación de distribuciones del Modelo 1 y del SICRI en 2007

La siguiente etapa considerada es la cuarta, correspondiente al año 2010, cuyos resultados se muestran en el Cuadro 10.3. En este año se aprecia una ligera subida en las prevalencias de fumadores y exfumadores, con una consecuente bajada en los nunca fumadores con respecto al 2007. Son cambios muy pequeños que no suponen una variación en la tendencia. Sin embargo, en los resultados del SICRI 2010 es la prevalencia de exfumadores la que aumenta con respecto a la obtenida en 2007.

Año 2010	NF	F	ExF
Distribución modelo	0.572	0.233	0.194
Intervalo de confianza	(0.567,0.577)	(0.225,0.246)	(0.183,0.202)
Distribución SICRI	0.553	0.246	0.201

Cuadro 10.3: Distribución en el año 2010 dada el Modelo 1, con su intervalo de confianza al 95 % y distribución SICRI 2010.

En esta etapa el ajuste de las prevalencias obtenidas del SICRI 2010 por parte del modelo mediante CMTD, mostrado en la Figura 10.2 es mejor que en la primera etapa, siendo sus diferencias menores del 2 % en todos los estados.

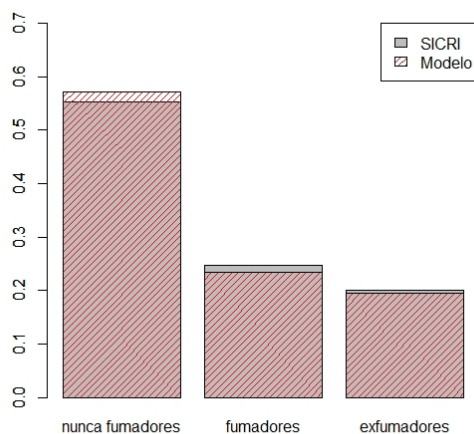


Figura 10.2: Comparación de distribuciones del Modelo 1 y del SICRI en 2010

El último año considerado en esta parte de validación es el 2011. El 2011 es el primer año de aplicación de la ley de tabaco 42/2010 [12], por lo que tiene gran interés ver si en las encuestas obtenidas durante ese año se aprecian cambios en las prevalencias, que se muestran en el Cuadro 10.4.

Año 2011	NF	F	ExF
Distribución modelo	0.569	0.235	0.197
Intervalo de confianza	(0.563,0.575)	(0.220,0.251)	(0.184,0.211)
Distribución SICRI	0.551	0.235	0.214

Cuadro 10.4: Distribución en el año 2011 dada el Modelo 1, con su intervalo de confianza al 95 % y distribución SICRI 2011.

De nuevo las variaciones producidas son muy leves, como se aprecia en la Figura 10.3. La distribución de la CMTD en esta quinta etapa ajusta muy fielmente la prevalencia de fumadores. Sin embargo sobreestima en menos de un 2 % la prevalencia de nunca fumadores y a la vez infraestima la prevalencia de exfumadores. A pesar de eso, el ajuste es el más fiel a los resultados SICRI 2011. El aumento en la prevalencia de exfumadores se podría atribuir a la entrada en vigor de la ley de tabaco.

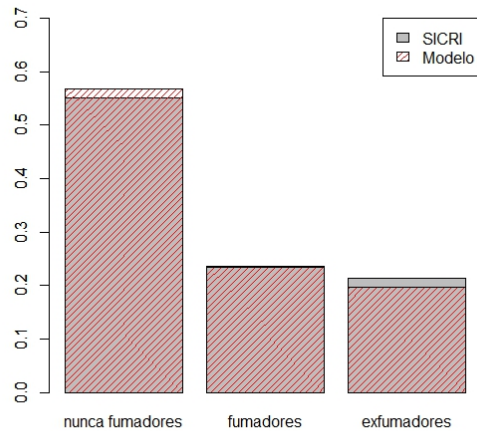


Figura 10.3: Comparación de distribuciones del Modelo 1 y del SICRI en 2011

Podemos ver, además, la evolución de las prevalencias en estas etapas, situando los valores del SICRI de los años respectivos, que se muestran como puntos en la Figura 10.4. Se aprecia la estabilidad que hay a lo largo de los años, lo que concuerda con las pequeñas diferencias vistas en los resultados. Se debe tener en cuenta que los resultados del SICRI se obtienen a partir de encuestas, por lo que tienen pequeñas variaciones entre distintas ediciones, lo que provoca que el modelo sea fiel a las prevalencias observadas en el año de partida, 2006.

Se aprecia una leve sobreestimación de los nunca fumadores, debido a que en el año 2006 sus niveles eran los más altos de los años 2006-2011, así como una infraestimación de las prevalencias de fumadores. Se observa una tendencia creciente en las prevalencias de fumadores y exfumadores, con un consecuente descenso de nunca fumadores.

En general, el ajuste obtenido mediante la CMTD a los resultados SICRI es bueno. Dichas diferencias se han contrastado utilizando los test vistos en la Sección 2.1.3. Sin embargo, las características de los datos, procedentes de una muestra compleja, dificultan el uso de dichos contrastes por lo que los resultados obtenidos por los mismos no serían fiables.

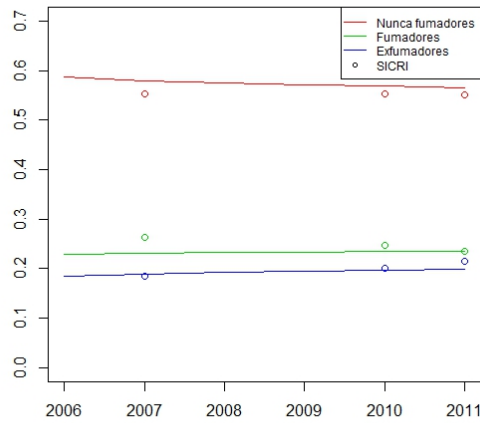


Figura 10.4: Evolución de las prevalencias entre 2006 y 2011. Se muestra con un punto la prevalencia correspondiente obtenida del SICRI.

## 10.2. Modelo 1. Resultados por sexos

En esta sección se analizan los resultados anteriores divididos por sexos, de forma que se puedan apreciar los distintos comportamientos en hombres y mujeres con respecto al tabaco.

En primer lugar, se obtiene la distribución inicial, correspondiente al año 2006. Se puede observar en la Figura 10.5 que entre los hombres hay más fumadores que entre las mujeres. En ambos casos el estado mayoritario es el de nunca fumadores, aunque en los hombres los valores son mucho más bajos que en las mujeres. Las diferencias entre la prevalencia de nunca fumadores y fumadores es mucho más acusada en el caso de las mujeres, en torno a un 50 %, mientras que en los hombres apenas supera 10 %. Así mismo, los niveles de fumadores y exfumadores son muy similares en el sector masculino.



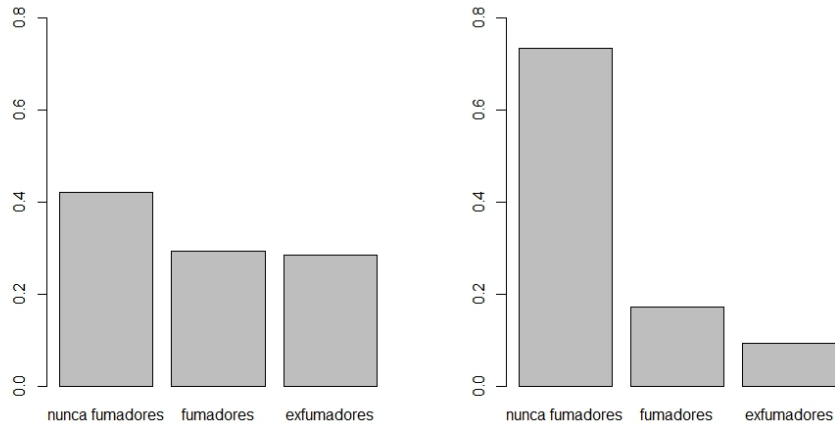


Figura 10.5: Distribución inicial del Modelo 1 para hombres y mujeres, respectivamente.

Se calcula la distribución en la primera etapa de la CMTD, con sus intervalos de confianza y se compara de nuevo con los resultados del SICRI 2007. En esta ocasión las diferencias de ajuste son mayores que en el caso global, como se ve en el Cuadro 10.5. Esto se debe a que el tamaño muestral se ha reducido, ya que se consideran dos modelos independientes en función del sexo. Esta reducción conlleva un peor ajuste de las estimaciones. A pesar de eso, las diferencias son menores del 5 %, por lo que se considera una buena aproximación.

Año 2007. Hombres	NF	F	ExF
Distribución modelo	0.417	0.296	0.288
Intervalo de confianza	(0.415,0.419)	(0.290,0.302)	(0.281,0.292)
Distribución SICRI	0.393	0.314	0.293
Año 2007. Mujeres	NF	F	ExF
Distribución modelo	0.733	0.171	0.097
Intervalo de confianza	(0.731,0.734)	(0.166,0.174)	(0.093,0.101)
Distribución SICRI	0.696	0.216	0.089

Cuadro 10.5: Distribución en el año 2007 dada el Modelo 1, con su intervalo de confianza al 95 % y distribución SICRI 2007,divididos por sexos.

Se ve en la Figura 10.6 que en ambos casos se sobreestima la prevalencia de nunca fumadores, igual que ocurría en el caso global, y se infraestima la de

fumadores.

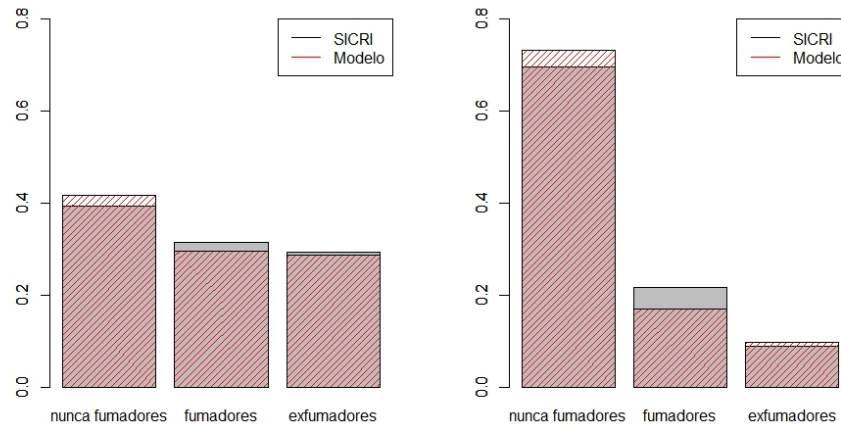


Figura 10.6: Comparación de distribuciones del Modelo 1 y del SICRI en 2007 para hombres y mujeres respectivamente.

La distribución de la CMTD en la cuarta etapa reduce considerablemente las diferencias con los resultados SICRI 2010, especialmente en el caso de los hombres, como se ve en el Cuadro 10.2. Tanto en hombres como en mujeres se aprecia un pequeño descenso en la prevalencia de nunca fumadores, que se corresponde con un ascenso en la de exfumadores.

Año 2010. Hombres	NF	F	ExF
Distribución modelo	0.403	0.302	0.295
Intervalo de confianza	(0.396,0.413)	(0.281,0.319)	(0.276,0.314)
Distribución SICRI	0.393	0.306	0.301
Año 2010. Mujeres	NF	F	ExF
Distribución modelo	0.724	0.170	0.106
Intervalo de confianza	(0.717,0.729)	(0.155,0.182)	(0.094,0.118)
Distribución SICRI	0.697	0.192	0.110

Cuadro 10.6: Distribución en el año 2010 dada el Modelo 1, con su intervalo de confianza al 95 % y distribución SICRI 2010,divididos por sexos.

En el caso de los hombres las diferencias son siempre inferiores al 1%, por lo que en la Figura 10.7 las diferencias son inapreciables.

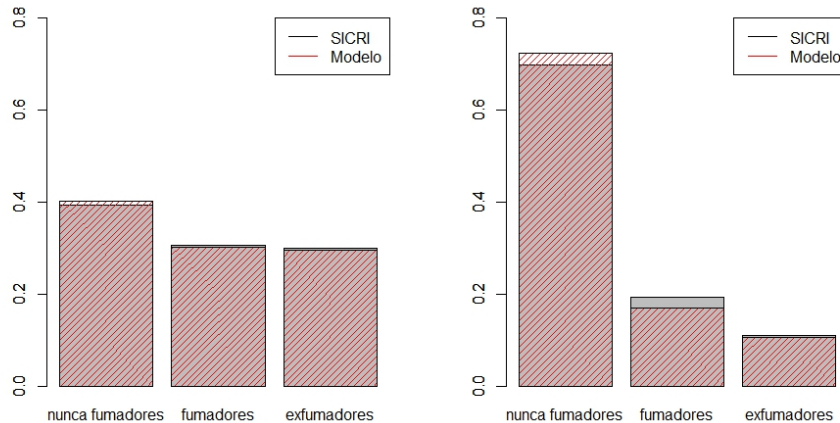


Figura 10.7: Comparación de distribuciones del Modelo 1 y del SICRI en 2010 para hombres y mujeres respectivamente.

La última etapa considerada es la quinta, correspondiente al 2011. En este año se aprecia en el Cuadro 10.2 un cambio en la tendencia de los hombres. Por primera vez, la prevalencia de exfumadores supera a la de fumadores en los resultados obtenidos del SICRI 2011. Este cambio puede atribuirse a la ley de tabaco que entró en vigor ese mismo año.

Año 2011. Hombres	NF	F	ExF
Distribución modelo	0.398	0.304	0.297
Intervalo de confianza	(0.387,0.410)	(0.281,0.328)	(0.276,0.321)
Distribución SICRI	0.400	0.287	0.314
Año 2011. Mujeres	NF	F	ExF
Distribución modelo	0.721	0.170	0.109
Intervalo de confianza	(0.711,0.729)	(0.154,0.191)	(0.093,0.123)
Distribución SICRI	0.697	0.192	0.110

Cuadro 10.7: Distribución en el año 2011 dada el Modelo 1, con su intervalo de confianza al 95 % y distribución SICRI 2011, divididos por sexos.

En ambos sexos, la prevalencia de exfumadores aumenta con respecto al año anterior. Las diferencias entre la distribución de la CMTD y el SICRI 2011 son mayores que en el año anterior en el caso de los hombres, esta mayor diferencia puede deberse a que la CMTD no refleja los cambios que ha podido provocar la

ley de tabaco en las prevalencias. Sin embargo en las mujeres, como se ve en la Figura 10.8 no se aprecian los mismos cambios.

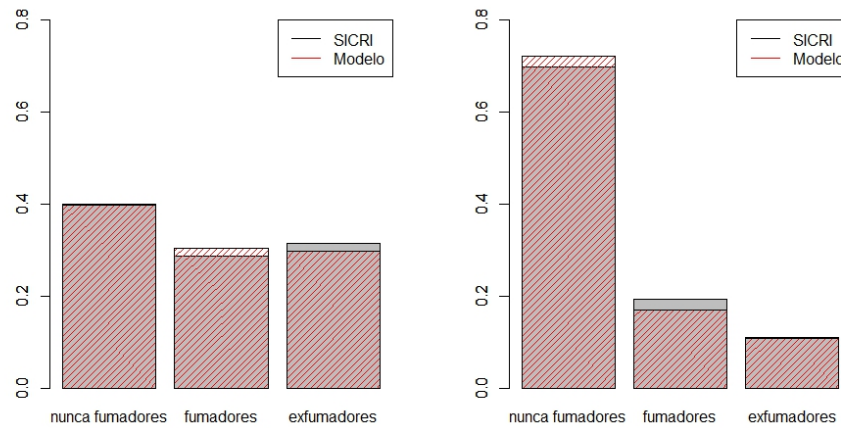


Figura 10.8: Comparación de distribuciones del Modelo 1 y del SICRI en 2011 para hombres y mujeres respectivamente.

En la evolución de las distribuciones de la CMTD a lo largo de las etapas, Figura 10.9, no se aprecian grandes cambios. Se observa una tendencia ligeramente creciente en las prevalencias de exfumadores en ambos sexos y descendiente en nunca fumadores. La prevalencia de mujeres fumadoras se mantiene constante.

En este modelo se obtienen las probabilidades de transición en base a los cambios de estado observados en el período 2005-2006 y dichas probabilidades se mantienen constantes, dado que se trata de una CMTD estacionaria. Por lo tanto no se ven cambios destacables en las prevalencias entre los años 2010 y 2011. Para que el modelo pueda reflejar alteraciones en las tendencias debidas a la ley de tabaco y otras intervenciones, deben modificarse las probabilidades de transición como se ha visto en el Capítulo 9.

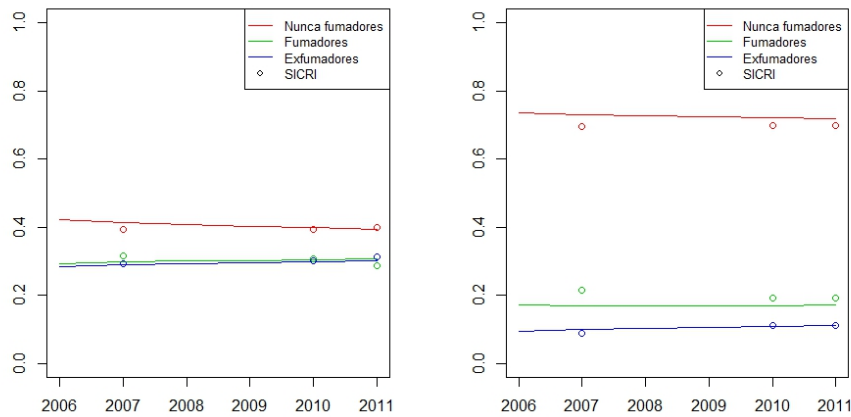


Figura 10.9: Evolución de las prevalencias, para hombres y mujeres respectivamente, en los años 2006-2011. Se muestra con puntos la prevalencia obtenida del SICRI.

### 10.3. Modelo 2. Resultados

En este apartado se muestran, de forma gráfica, los resultados obtenidos con el Modelo 2. Este modelo se distingue del anterior solamente en la definición de dos estados de exfumadores, de corta y de larga duración. La prevalencia de exfumadores de corta duración es muy baja, por lo que, en general, los resultados son análogos a los obtenidos con el Modelo 1.

Las diferencias entre las distribuciones de la CMTD y los resultados del SICRI para cada etapa siguen siendo muy pequeñas, por lo que se considera que este modelo también tiene un buen ajuste. La ventaja de este modelo con respecto al anterior es que permite diferenciar los estados de exfumadores en función del tiempo que llevan sin fumar y determinar de forma más sencilla las recaídas.

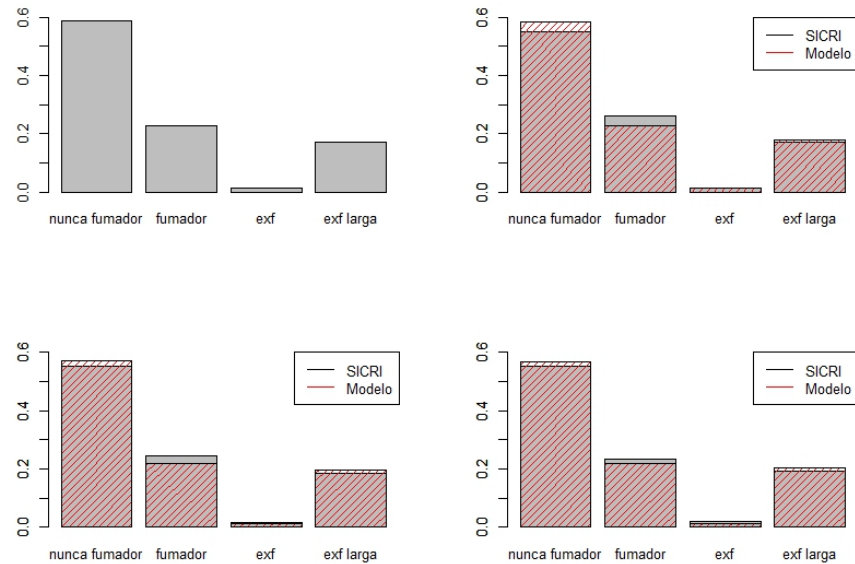


Figura 10.10: Resultados del Modelo 2. En la primera fila se muestra la distribución inicial y la distribución en el año 2007. En la segunda fila se muestra la distribución en el año 2010 y 2011 respectivamente.

## 10.4. Modelo 3. Resultados

El Modelo 3 introduce la edad en la definición del espacio de estados. Se tienen los mismos estados que en el Modelo 2 pero se dividen en función del grupo de edad. En el modelo no se tiene en cuenta mortalidad ni natalidad, por lo que se estudia una población fija, sin reemplazo.

Se puede ver en los siguientes gráficos la comparación entre la distribución de la CMTD y el SICRI. Se aprecia que, con el paso de las etapas, se van sobreestimando los estados correspondientes a los grupos de mayor edad, mientras que los demás se van infraestimando. En el año 2007, por ser la primera etapa calculada, no se aprecian tantas diferencias y se ve un buen ajuste entre las distribuciones.

Partiendo de la población estudiada en 2006, se estudia su evolución a lo largo de las etapas. Al cabo de un cierto número de etapas los individuos habrán envejecido tantos años como etapas. Dado que la edad está implícita en la definición de los estados, esto supone que la población se irá trasladando de

grupo de edad con el paso de los años, dejando vacíos los grupos más jóvenes. Al no tener en cuenta la mortalidad, la población irá acumulándose en los grupos de mayor edad. Este traslado de la población hace que las distribuciones de la CMTD y el SICRI no sean comparables, ya que la clasificación de los individuos por grupos de edad no se mantiene.

A partir de las probabilidades de transición obtenidas en este modelo se hacen las predicciones vistas en el Capítulo 8, en las que sí se tiene en cuenta la dinámica poblacional para evitar este desfase de grupos de edad entre etapas.

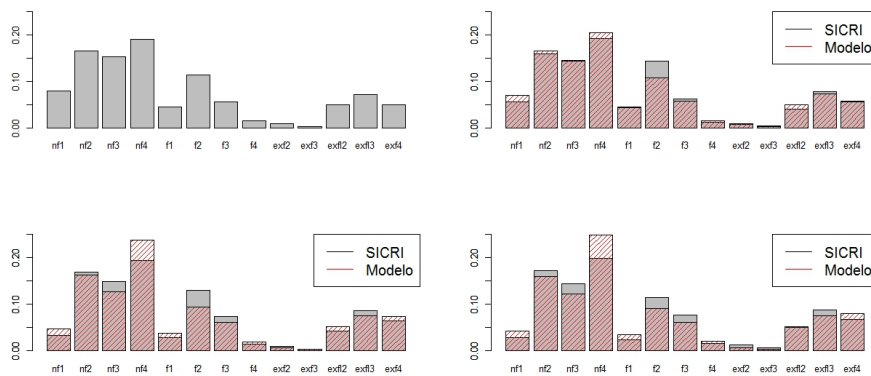


Figura 10.11: Resultados del Modelo 3. En la primera fila se muestra la distribución inicial y la distribución en el año 2007. En la segunda fila se muestra la distribución en el año 2010 y 2011 respectivamente





## Capítulo 11

# Proyecciones hasta 2020

En este capítulo se muestran los resultados procedentes de la predicción de prevalencias de consumo de tabaco desde el año 2010. En la primera parte se muestran las predicciones partiendo de los datos obtenidos en el SICRI 2010. En la segunda parte se estudia el efecto de las intervenciones sobre las predicciones calculadas, distinguiendo distintos escenarios según el impacto esperado de las mismas. Todos los resultados se muestra segregados por sexos.

### 11.1. Predicción de prevalencias

Para predecir las prevalencias de consumo de tabaco desde 2010 a 2020 se utilizan las fórmulas desarrolladas en el Capítulo 8. Se toma como distribución inicial la resultante de aplicar el Modelo 3 de CMTD, en el que se define el espacio de estados en función del grupo de edad.

Para obtener las prevalencias adaptadas a la población de cada año se introducen tasas de mortalidad y se incorpora en cada año una cantidad de población relativa a los nacimientos. Dado que la población considerada en el estudio es de más de 16 años, en cada etapa posterior a 2010 se incorporará la población de 16 años, obtenida de la proyección del censo del IGE, supliendo así la “extinción” de los grupos de edad. Dicha población entra en los estados de nunca fumadores y fumadores del grupo de edad 1 (16-24), por lo que se utiliza el porcentaje de fumadores de 15 años obtenido del estudio SIX, aplicado a la población de 15 años del año anterior al considerado, que será la población de 16 años que entra en el estudio. La proyección de la población de 15 años para los años 2010 a 2019, que se utiliza para las predicciones entre 2011 y 2020, dividida por sexos, se muestra en el Cuadro 11.1.

Las predicciones de las prevalencias se hacen por grupos de edad y por sexos. En primer lugar se muestran los resultados obtenidos para el grupo de edad 1 (16-24). En este grupo se consideran dos estados, nunca fumadores y fumadores,

incluyendo en estos a los exfumadores.

Año	2010	2011	2012	2013	2014	2015	2016	2017
Hombres	11154	10913	10538	10386	10562	10580	10909	11158
Mujeres	10615	10323	9996	9849	9923	9983	10229	10483
	2018	2019						
Hombres	10848	11358						
Mujeres	10211	10708						

Cuadro 11.1: Proyección de la población de 15 años entre 2010 y 2019.

Tanto en hombres como en mujeres la tendencia durante estos años es creciente en la prevalencia de fumadores, como se ve en los Cuadros 11.2 y 11.3. En el año 2020, con respecto al 2010, se habrá producido un incremento de, aproximadamente, un 15 % en los hombres fumadores y de un 10 % en mujeres fumadoras. Dicho incremento supone un descenso en la prevalencia de nunca fumadores de la misma magnitud.

Año	NF1		F1	
2010	0.606		0.393	
2011	0.568	(0.548,0.592)	0.431	(0.408,0.452)
2012	0.537	(0.514,0.565)	0.462	(0.435,0.486)
2013	0.511	(0.483,0.545)	0.488	(0.455,0.517)
2014	0.491	(0.462,0.528)	0.508	(0.472,0.538)
2015	0.476	(0.442,0.511)	0.523	(0.489,0.558)
2016	0.465	(0.432,0.505)	0.534	(0.495,0.568)
2017	0.458	(0.424,0.500)	0.541	(0.500,0.576)
2018	0.451	(0.413,0.497)	0.548	(0.503,0.587)
2019	0.448	(0.412,0.490)	0.551	(0.510,0.588)
2020	0.446	(0.412,0.481)	0.553	(0.519,0.588)

Cuadro 11.2: Predicciones de prevalencias para hombres del grupo de edad 16-24.

En la Figura 11.1 se aprecian las tendencias descritas. Además, cabe destacar que, en torno al año 2014, la proporción de fumadores superará a la de nunca fumadores en el caso de los hombres. Para las mujeres las tendencias se mantienen, de forma menos pronunciada y siendo siempre mayor la proporción de nunca fumadoras. En ambos casos parece que hacia 2020 las prevalencias se estabilizan.

Año	NF1		F1	
2010	0.647		0.352	
2011	0.624	(0.596,0.656)	0.375	(0.344,0.404)
2012	0.604	(0.568,0.644)	0.395	(0.356,0.432)
2013	0.588	(0.542,0.635)	0.411	(0.365,0.458)
2014	0.576	(0.538,0.626)	0.423	(0.374,0.462)
2015	0.566	(0.530,0.624)	0.433	(0.376,0.470)
2016	0.559	(0.522,0.612)	0.440	(0.388,0.478)
2017	0.555	(0.516,0.605)	0.444	(0.395,0.484)
2018	0.551	(0.515,0.596)	0.448	(0.404,0.485)
2019	0.549	(0.510,0.593)	0.450	(0.407,0.490)
2020	0.548	(0.509,0.592)	0.451	(0.408,0.491)

Cuadro 11.3: Predicciones de prevalencias para mujeres del grupo de edad 16-24.

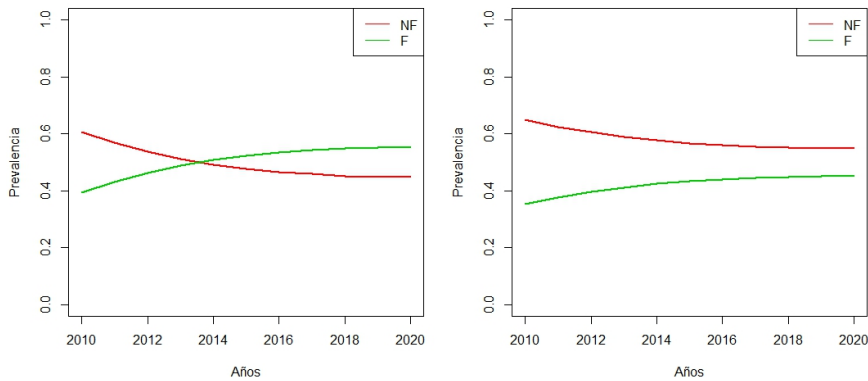


Figura 11.1: Predicciones de prevalencias para hombres y mujeres, respectivamente, para el grupo de edad 16-24.

En los grupos de edad 2 (25-44) y 3 (45-64) se consideran cuatro estados, nunca fumadores, fumadores, exfumadores de corta duración y exfumadores de larga duración. En ambos grupos, la proporción de exfumadores de corta duración es muy baja, por lo que son inapreciables cambios en su prevalencia.

En el grupo de 25-44 años, al contrario que en el grupo más joven, desciende la prevalencia de fumadores. En los hombres desciende en torno a un 10 % (Cuadro 11.4) frente a un 4 % en las mujeres (Cuadro 11.5). El descenso en los fumadores se corresponde, como es de esperar, con un ascenso en las prevalencias de exfumadores.

Año	NF2		F2	
2010	0.428		0.420	
2011	0.439	(0.433,0.447)	0.404	(0.389,0.417)
2012	0.449	(0.440,0.459)	0.389	(0.378,0.404)
2013	0.457	(0.443,0.468)	0.377	(0.364,0.397)
2014	0.465	(0.450,0.478)	0.367	(0.349,0.387)
2015	0.471	(0.455,0.485)	0.357	(0.338,0.378)
2016	0.477	(0.462,0.493)	0.349	(0.333,0.374)
2017	0.483	(0.465,0.497)	0.341	(0.324,0.362)
2018	0.488	(0.470,0.505)	0.335	(0.318,0.356)
2019	0.493	(0.475,0.511)	0.329	(0.312,0.350)
2020	0.497	(0.480,0.519)	0.323	(0.303,0.348)
Año	ExF2		ExLF2	
2010	0.032		0.118	
2011	0.029	(0.020,0.039)	0.126	(0.120,0.134)
2012	0.028	(0.021,0.036)	0.131	(0.123,0.142)
2013	0.027	(0.018,0.036)	0.136	(0.127,0.148)
2014	0.026	(0.020,0.035)	0.140	(0.128,0.156)
2015	0.026	(0.019,0.035)	0.144	(0.130,0.163)
2016	0.025	(0.018,0.033)	0.147	(0.132,0.165)
2017	0.024	(0.018,0.032)	0.150	(0.133,0.166)
2018	0.024	(0.018,0.032)	0.152	(0.133,0.168)
2019	0.023	(0.018,0.031)	0.153	(0.134,0.171)
2020	0.023	(0.018,0.030)	0.155	(0.134,0.174)

Cuadro 11.4: Predicciones de prevalencias para hombres del grupo de edad 25-44.

En la Figura 11.2 se aprecia que, en el caso de los hombres, además de la prevalencia de exfumadores, también aumenta la de nunca fumadores. Para las mujeres, la prevalencia de nunca fumadores se mantiene más estable, llegando incluso a descender ligeramente.

Año	NF2		F2	
2010	0.529		0.320	
2011	0.528	(0.523,0.537)	0.313	(0.300,0.322)
2012	0.526	(0.517,0.536)	0.306	(0.288,0.320)
2013	0.524	(0.514,0.538)	0.301	(0.283,0.317)
2014	0.521	(0.510,0.536)	0.297	(0.279,0.310)
2015	0.518	(0.506,0.535)	0.293	(0.274,0.307)
2016	0.515	(0.501,0.531)	0.290	(0.272,0.309)
2017	0.512	(0.496,0.529)	0.287	(0.266,0.306)
2018	0.509	(0.493,0.528)	0.284	(0.261,0.306)
2019	0.505	(0.487,0.528)	0.282	(0.259,0.303)
2020	0.502	(0.487,0.524)	0.280	(0.259,0.300)

Año	ExF2		ExLF2	
2010	0.026		0.122	
2011	0.025	(0.019,0.032)	0.132	(0.128,0.139)
2012	0.025	(0.016,0.034)	0.141	(0.132,0.149)
2013	0.024	(0.018,0.034)	0.149	(0.138,0.162)
2014	0.024	(0.016,0.033)	0.156	(0.145,0.172)
2015	0.023	(0.016,0.032)	0.163	(0.153,0.179)
2016	0.023	(0.014,0.032)	0.170	(0.157,0.185)
2017	0.023	(0.016,0.031)	0.176	(0.161,0.194)
2018	0.023	(0.015,0.030)	0.182	(0.166,0.201)
2019	0.022	(0.016,0.031)	0.188	(0.171,0.203)
2020	0.022	(0.016,0.032)	0.194	(0.176,0.211)

Cuadro 11.5: Predicciones de prevalencias para mujeres del grupo de edad 25-44.

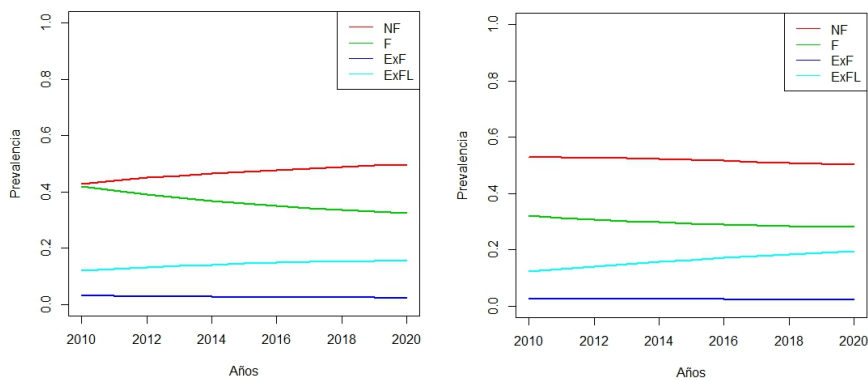


Figura 11.2: Predicciones de prevalencias para hombres y mujeres, respectivamente, para el grupo de edad 25-44.

En el grupo 3 (45-64) la situación de hombres y mujeres es muy diferente. Por un lado, en los hombres, la prevalencia más alta corresponde a los exfumadores de larga duración, superando en torno a un 10 % la prevalencia de los otros estados, excepto el de exfumadores de corta duración (Cuadro 11.6). Esta prevalencia, además, se mantiene prácticamente constante a lo largo de los años. Las prevalencias de nunca fumadores y fumadores son muy similares en los primeros años, siendo más alta la de fumadores. Sin embargo, a partir de 2014 la prevalencia de fumadores es inferior a la de nunca fumadores.

En la mujeres las tendencias son más pronunciadas que en los hombres. Se produce un aumento constante, en torno a un 8 %, tanto en fumadoras como en exfumadoras de larga duración, lo que supone un descenso en nunca fumadoras de un 18 % (Cuadro 11.7).

Año	NF3		F3	
2010	0.276		0.298	
2011	0.278	(0.270,0.285)	0.295	(0.282,0.307)
2012	0.280	(0.267,0.292)	0.291	(0.278,0.309)
2013	0.283	(0.270,0.301)	0.288	(0.270,0.311)
2014	0.286	(0.272,0.307)	0.284	(0.267,0.304)
2016	0.292	(0.274,0.310)	0.275	(0.253,0.301)
2017	0.296	(0.275,0.317)	0.271	(0.246,0.293)
2018	0.299	(0.279,0.319)	0.266	(0.241,0.291)
2019	0.302	(0.283,0.321)	0.262	(0.239,0.286)
2020	0.306	(0.283,0.329)	0.257	(0.236,0.282)
Año	ExF3		ExFL3	
2010	0.020		0.405	
2011	0.019	(0.012,0.027)	0.406	(0.397,0.416)
2012	0.019	(0.012,0.029)	0.407	(0.396,0.421)
2013	0.019	(0.013,0.028)	0.408	(0.391,0.426)
2014	0.019	(0.012,0.026)	0.410	(0.392,0.431)
2015	0.018	(0.011,0.027)	0.411	(0.392,0.431)
2016	0.018	(0.012,0.025)	0.413	(0.395,0.433)
2017	0.018	(0.011,0.026)	0.414	(0.395,0.438)
2018	0.018	(0.012,0.025)	0.415	(0.397,0.439)
2019	0.017	(0.012,0.024)	0.417	(0.396,0.443)
2020	0.017	(0.010,0.022)	0.418	(0.395,0.445)

Cuadro 11.6: Predicciones de prevalencias para hombres del grupo de edad 45-64.

Año	NF3		F3	
2010	0.661		0.177	
2011	0.643	(0.636,0.651)	0.188	(0.181,0.199)
2012	0.625	(0.616,0.639)	0.198	(0.188,0.211)
2013	0.607	(0.598,0.625)	0.208	(0.196,0.219)
2014	0.590	(0.579,0.610)	0.218	(0.201,0.231)
2015	0.573	(0.562,0.595)	0.226	(0.208,0.240)
2016	0.557	(0.544,0.583)	0.235	(0.214,0.246)
2017	0.541	(0.525,0.570)	0.243	(0.222,0.261)
2018	0.526	(0.510,0.556)	0.250	(0.226,0.267)
2019	0.511	(0.495,0.541)	0.257	(0.233,0.279)
2020	0.497	(0.480,0.527)	0.263	(0.238,0.287)

Año	ExF3		ExFL3	
2010	0.007		0.154	
2011	0.007	(0.003,0.012)	0.160	(0.154,0.167)
2012	0.007	(0.003,0.012)	0.167	(0.157,0.175)
2013	0.008	(0.003,0.013)	0.175	(0.161,0.185)
2014	0.008	(0.003,0.014)	0.182	(0.166,0.195)
2015	0.008	(0.003,0.014)	0.190	(0.174,0.206)
2016	0.009	(0.004,0.018)	0.197	(0.178,0.215)
2017	0.009	(0.004,0.015)	0.205	(0.187,0.225)
2018	0.009	(0.004,0.016)	0.213	(0.193,0.237)
2019	0.010	(0.004,0.017)	0.220	(0.198,0.244)
2020	0.010	(0.004,0.017)	0.228	(0.204,0.250)

Cuadro 11.7: Predicciones de prevalencias para mujeres del grupo de edad 45-64.

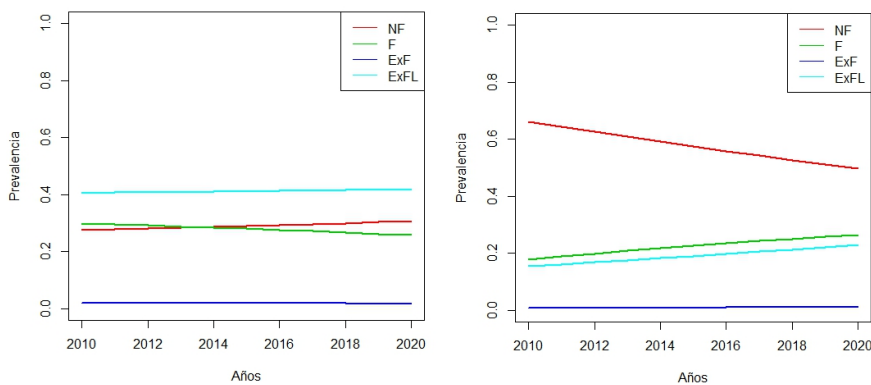


Figura 11.3: Predicciones de prevalencias para hombres y mujeres, respectivamente, para el grupo de edad 45-64.

En la Figura 11.3 se aprecia claramente el contraste entre la estabilidad de las prevalencias en hombres y las pronunciadas tendencias en mujeres. En el caso de los hombres se ve claramente que, entre 2013 y 2014 se cruzan las prevalencias de fumadores y nunca fumadores. En las mujeres, a pesar del fuerte descenso en nunca fumadoras, se siguen manteniendo en niveles mucho más altos que los otros estados.

En el grupo de edad 4 (65 y más), se consideran tres estados, nunca fumadores, fumadores y exfumadores. De nuevo la situación entre hombres y mujeres es totalmente distinta.

Cabe destacar que, en el caso de las mujeres, el tamaño muestral obtenido en el SICRI es muy bajo comparado con los demás grupos, lo que lleva a realizar peores estimaciones de las probabilidades de transición y, consecuentemente, a realizar predicciones erróneas en ese grupo. Ante esta situación se ha optado por aplicar a las mujeres del grupo de edad 4 las probabilidades de transición obtenidas en el grupo 3. En el caso de los hombres este problema no es tan acusado, siendo también las probabilidades de transición entre esos dos grupos de edad similares.

Las prevalencias en hombres del grupo de edad 4 son bastante estables. Se produce un aumento en los fumadores, que son mayoría en este grupo, de un 8%, como se ve en el Cuadro 11.1. Dicho incremento se traduce en un descenso tanto en nunca fumadores como en exfumadores.

Año	NF4	F4	ExF4
2010	0.397	0.098	0.504
2011	0.390 (0.383,0.398)	0.091 (0.082,0.101)	0.518 (0.508,0.528)
2012	0.384 (0.375,0.394)	0.086 (0.075,0.098)	0.529 (0.519,0.542)
2013	0.378 (0.368,0.392)	0.080 (0.068,0.094)	0.540 (0.528,0.553)
2014	0.374 (0.362,0.387)	0.076 (0.065,0.089)	0.549 (0.537,0.563)
2015	0.370 (0.356,0.383)	0.071 (0.061,0.085)	0.557 (0.544,0.573)
2016	0.366 (0.352,0.381)	0.068 (0.055,0.080)	0.565 (0.551,0.580)
2017	0.363 (0.351,0.378)	0.064 (0.054,0.076)	0.571 (0.557,0.589)
2018	0.361 (0.348,0.375)	0.061 (0.051,0.072)	0.577 (0.562,0.592)
2019	0.358 (0.346,0.372)	0.058 (0.048,0.070)	0.582 (0.568,0.597)
2020	0.356 (0.342,0.372)	0.055 (0.045,0.067)	0.587 (0.571,0.602)

Cuadro 11.8: Predicciones de prevalencias para hombres del grupo de edad 65 y más.



Año	NF4	F4	ExF4
2010	0.935	0.016	0.048
2011	0.927 (0.922,0.932)	0.018 (0.014,0.022)	0.053 (0.050,0.058)
2012	0.920 (0.914,0.928)	0.020 (0.015,0.026)	0.058 (0.053,0.065)
2013	0.914 (0.906,0.923)	0.022 (0.015,0.027)	0.063 (0.056,0.071)
2014	0.908 (0.899,0.917)	0.023 (0.016,0.030)	0.067 (0.061,0.076)
2015	0.902 (0.893,0.913)	0.025 (0.017,0.030)	0.072 (0.064,0.081)
2016	0.897 (0.887,0.907)	0.026 (0.018,0.033)	0.076 (0.069,0.085)
2017	0.891 (0.881,0.903)	0.027 (0.019,0.035)	0.080 (0.073,0.090)
2018	0.886 (0.875,0.897)	0.028 (0.019,0.035)	0.085 (0.076,0.095)
2019	0.880 (0.872,0.893)	0.029 (0.020,0.035)	0.089 (0.079,0.101)
2020	0.875 (0.866,0.887)	0.030 (0.020,0.037)	0.093 (0.084,0.104)

Cuadro 11.9: Predicciones de prevalencias para mujeres del grupo de edad 65 y más.

En la Figura 11.4 puede apreciarse la gran diferencia en los niveles de prevalencia de mujeres. Las nunca fumadoras superan en 2010 el 90 %, descendiendo un 5 % hacia 2020 (Cuadro 11.1). Además, es mayor la proporción de exfumadoras que de fumadoras, aunque en ambos casos son valores muy pequeños.

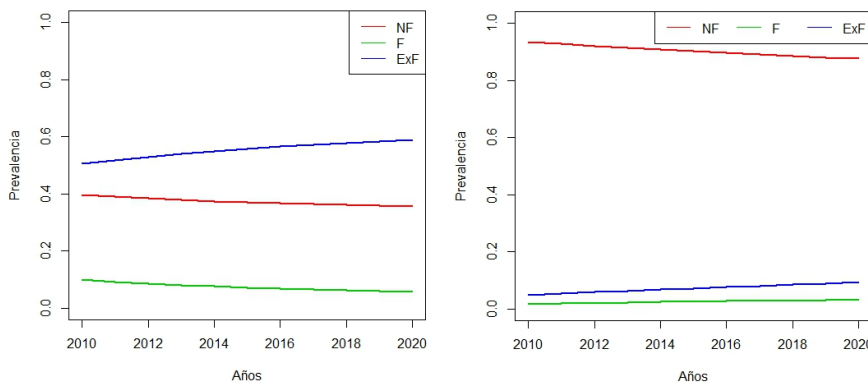


Figura 11.4: Predicciones de prevalencias para hombres y mujeres, respectivamente, para el grupo de edad 65 y más.

Por último, se analizan las predicciones de prevalencia por sexos, pero considerando todos los grupos de edad de forma conjunta. Se consideran los tres estados principales, de forma agregada: nunca fumadores, fumadores y exfumadores. Las prevalencias, en este caso, se mantienen más estables que al dividir las por grupos de edad.

La principal diferencia entre hombres y mujeres es que, en el caso de los hombres, las distancias entre las prevalencias de los distintos estados es muy pequeña en comparación con las mujeres, en las que destaca el dominio de la proporción de nunca fumadoras, como se ve en los Cuadros 11.1 y 11.1

Año	NF	F	ExF
2010	0.394	0.306	0.298
2011	0.393 (0.391,0.396)	0.300 (0.295,0.306)	0.306 (0.302,0.311)
2012	0.392 (0.390,0.395)	0.293 (0.288,0.300)	0.314 (0.309,0.319)
2013	0.392 (0.389,0.395)	0.286 (0.280,0.294)	0.321 (0.315,0.326)
2014	0.392 (0.389,0.396)	0.280 (0.272,0.287)	0.327 (0.321,0.334)
2015	0.392 (0.389,0.397)	0.273 (0.264,0.282)	0.333 (0.327,0.340)
2016	0.393 (0.390,0.398)	0.267 (0.256,0.276)	0.339 (0.331,0.347)
2017	0.394 (0.390,0.400)	0.261 (0.250,0.270)	0.344 (0.337,0.353)
2018	0.395 (0.391,0.401)	0.255 (0.245,0.264)	0.349 (0.343,0.358)
2019	0.396 (0.392,0.402)	0.249 (0.239,0.258)	0.354 (0.347,0.361)
2020	0.397 (0.393,0.404)	0.244 (0.233,0.253)	0.358 (0.351,0.366)

Cuadro 11.10: Predicciones de prevalencias totales para hombres.

Año	NF	F	ExF
2010	0.697	0.192	0.109
2011	0.693 (0.691,0.695)	0.192 (0.189,0.198)	0.114 (0.111,0.118)
2012	0.688 (0.685,0.691)	0.192 (0.187,0.198)	0.118 (0.114,0.123)
2013	0.684 (0.679,0.687)	0.192 (0.187,0.199)	0.123 (0.119,0.128)
2014	0.680 (0.675,0.682)	0.191 (0.187,0.199)	0.127 (0.124,0.133)
2015	0.676 (0.671,0.678)	0.191 (0.188,0.199)	0.132 (0.127,0.138)
2016	0.672 (0.667,0.674)	0.190 (0.186,0.199)	0.136 (0.133,0.144)
2017	0.669 (0.663,0.670)	0.189 (0.185,0.198)	0.140 (0.137,0.149)
2018	0.666 (0.659,0.666)	0.188 (0.185,0.197)	0.145 (0.141,0.153)
2019	0.662 (0.654,0.662)	0.188 (0.185,0.197)	0.149 (0.146,0.156)
2020	0.659 (0.650,0.658)	0.187 (0.185,0.198)	0.153 (0.149,0.161)

Cuadro 11.11: Predicciones de prevalencias totales para mujeres.

En ambos casos, la tendencia más pronunciada es la de exfumadores, que aumenta un 5%, aproximadamente, en los dos sexos. En hombres, además, se aprecia un descenso de fumadores más pronunciado que en las mujeres, donde se mantiene prácticamente constante, como se ve en la Figura 11.5.

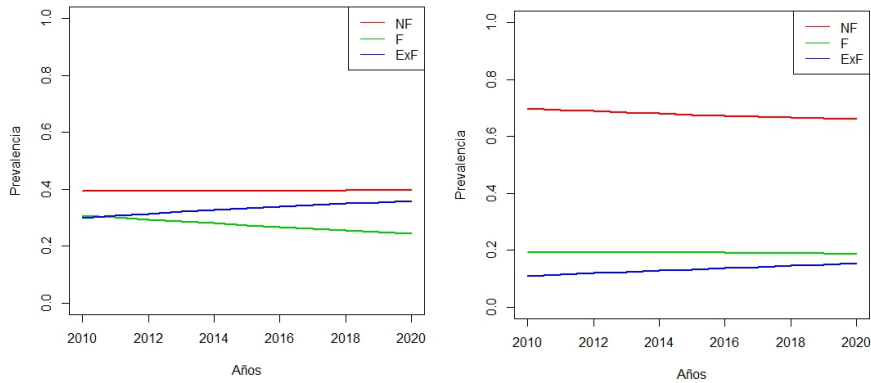


Figura 11.5: Predicciones de prevalencias totales para hombres y mujeres, respectivamente.

## 11.2. Efecto de las intervenciones

En este apartado se muestran los resultados de la aplicación de diferentes intervenciones sobre las predicciones obtenidas anteriormente, entre 2010 y 2020. Se utilizan las políticas detalladas en el Capítulo 9, de forma conjunta, con las que se modifican las probabilidades de transición entre los estados. En concreto, se modifican las probabilidades de inicio y cese en el hábito de fumar y se adapta el resto de probabilidades para cumplir las condiciones de una CMTD.

Se consideran los mismo estados que en el apartado anterior, divididos por sexos y grupos de edad. Se muestra la comparación entre las predicciones normales y las predicciones considerando el efecto de las intervenciones. Además, dentro de las intervenciones se consideran dos escenarios. Se modeliza el efecto normal de cada intervención y su efecto máximo, lo que permite observar en qué medida aumentar el impacto de ciertas políticas se traduce en un mayor efecto y concienciación sobre la población.

En primer lugar se muestran los resultados obtenidos sobre el grupo de edad de 16-24 años. Este es el grupo de edad en el que se produce el inicio en el hábito. Es de vital importancia reducir el número de jóvenes que empiezan a fumar de cara a reducir la prevalencia de fumadores en la población. Existen políticas especiales dirigidas a este grupo de edad con el objetivo de advertirles del riesgo que conlleva empezar a fumar, así como la dificultad de abandonar el hábito.

En todos los resultados, de aquí en adelante, se muestran los valores obte-

nidos para las predicciones de las prevalencias sin intervención, el valor bajo el efecto normal de las intervenciones (int.) y el valor bajo el efecto máximo posible de las mismas intervenciones (int. max.).

Año	NF	NF int.	NF int. max.	F	F int.	F int. max.
2010	0.607	0.607	0.607	0.393	0.393	0.393
2011	0.569	0.572	0.584	0.431	0.428	0.416
2012	0.537	0.542	0.565	0.463	0.458	0.435
2013	0.512	0.518	0.548	0.488	0.482	0.452
2014	0.492	0.500	0.536	0.508	0.500	0.464
2015	0.476	0.485	0.527	0.524	0.515	0.473
2016	0.466	0.475	0.520	0.534	0.525	0.480
2017	0.458	0.469	0.516	0.542	0.531	0.484
2018	0.452	0.462	0.513	0.548	0.538	0.487
2019	0.449	0.460	0.512	0.551	0.540	0.488
2020	0.447	0.458	0.512	0.553	0.542	0.488

Cuadro 11.12: Predicciones de prevalencias para hombres de edades 16-24 con intervenciones.

En el caso de los hombres de este grupo de edad (16-24), se produce el mayor cambio de todos los estados y grupos de edad considerados. Se aprecia una reducción de hasta el 6% en la prevalencia de fumadores (Cuadro 11.12). Con respecto al año 2010, supondría un aumento en la proporción de fumadores de apenas un 10%, frente al 15% que aumentaría si no se considera ninguna intervención.

Año	NF	NF int.	NF int. max.	F	F int.	F int. max.
2010	0.648	0.648	0.648	0.352	0.352	0.352
2011	0.624	0.625	0.627	0.376	0.375	0.373
2012	0.605	0.606	0.610	0.395	0.394	0.390
2013	0.589	0.590	0.596	0.411	0.410	0.404
2014	0.576	0.578	0.585	0.424	0.422	0.415
2015	0.567	0.569	0.576	0.433	0.431	0.424
2016	0.560	0.562	0.570	0.440	0.438	0.430
2017	0.555	0.558	0.566	0.445	0.442	0.434
2018	0.551	0.553	0.562	0.449	0.447	0.438
2019	0.550	0.552	0.560	0.450	0.448	0.440
2020	0.549	0.551	0.559	0.451	0.449	0.441

Cuadro 11.13: Predicciones de prevalencias para mujeres de edades 16-24 con intervenciones.

Las intervenciones tienen un efecto muy leve en las mujeres, en las que apenas desciende, en el mejor de los casos, un 1% la prevalencia de fumadoras en 2020, con respecto a las predicciones previas (Cuadro 11.13).

En la Figura 11.6 se aprecia que, en el mejor de los casos, se conseguiría que en el año 2020 la proporción de fumadores no fuera superior a la de nunca fumadores.

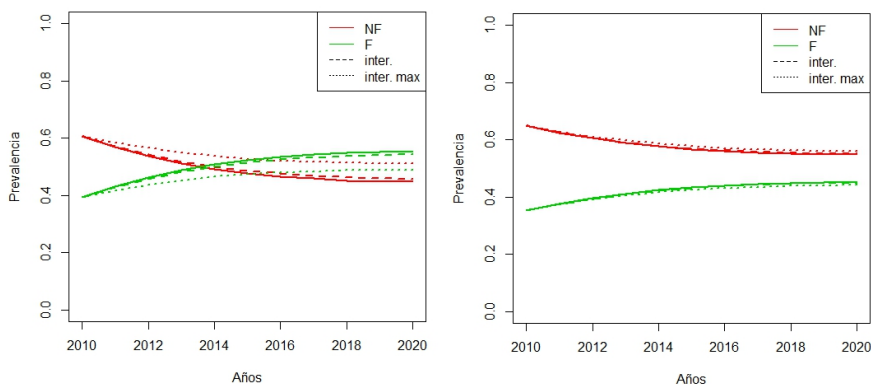


Figura 11.6: Predicciones de prevalencias para el grupo de edad 16-24 con intervenciones, para hombres y mujeres, respectivamente.

En los grupos de edad siguientes (25 y más), no hay inicio en el hábito de fumar. El efecto de las intervenciones es, por lo tanto, aumentar la probabilidad de dejar de fumar. Debería apreciarse un descenso en la prevalencia de fumadores, en favor de un aumento en la de exfumadores.

En el grupo de edad 2 (25-44), la reducción máxima sobre el valor de la predicción de la prevalencia de fumadores a 2020 es de un 2% en los hombres, aunque si se considera el efecto normal de las intervenciones dicha reducción no llegaría al 1% (Cuadro 11.14). En las mujeres los efectos son más leves, en general, en todos los grupos de edad. En este caso, la reducción en la prevalencia de fumadoras no supera el 1% aún aplicando el efecto máximo de las intervenciones (Cuadro 11.15).

En la Figura 11.7 se aprecia ese descenso en la prevalencia de fumadores hombres, con un consecuente aumento de exfumadores. En el caso de las mujeres las diferencias son muy pequeñas. Como es de esperar, en ambos casos la prevalencia de nunca fumadores se mantiene en los mismos niveles.

Año	NF	NF int.	NF int. máx.	F	F int.	F int. máx.
2010	0.429	0.429	0.429	0.420	0.420	0.420
2011	0.440	0.440	0.439	0.404	0.402	0.399
2012	0.450	0.449	0.448	0.390	0.387	0.384
2013	0.458	0.457	0.457	0.378	0.374	0.370
2014	0.465	0.464	0.464	0.367	0.363	0.357
2015	0.472	0.471	0.471	0.358	0.353	0.346
2016	0.478	0.477	0.477	0.349	0.344	0.336
2017	0.483	0.482	0.484	0.342	0.336	0.327
2018	0.488	0.487	0.490	0.335	0.329	0.319
2019	0.493	0.492	0.495	0.329	0.322	0.311
2020	0.498	0.497	0.501	0.324	0.316	0.304
Año	ExF	ExF int.	ExF int. max.	ExFL	ExFL int.	ExFL int. max.
2010	0.032	0.032	0.032	0.119	0.119	0.119
2011	0.030	0.033	0.036	0.126	0.126	0.125
2012	0.029	0.031	0.035	0.132	0.133	0.133
2013	0.028	0.030	0.033	0.137	0.139	0.140
2014	0.027	0.029	0.032	0.141	0.144	0.146
2015	0.026	0.028	0.031	0.145	0.148	0.152
2016	0.025	0.028	0.030	0.148	0.152	0.156
2017	0.025	0.027	0.029	0.150	0.155	0.160
2018	0.024	0.026	0.029	0.152	0.158	0.163
2019	0.024	0.026	0.028	0.154	0.160	0.166
2020	0.023	0.025	0.027	0.155	0.162	0.168

Cuadro 11.14: Predicciones de prevalencias para hombres de edades 25-44 con intervenciones.

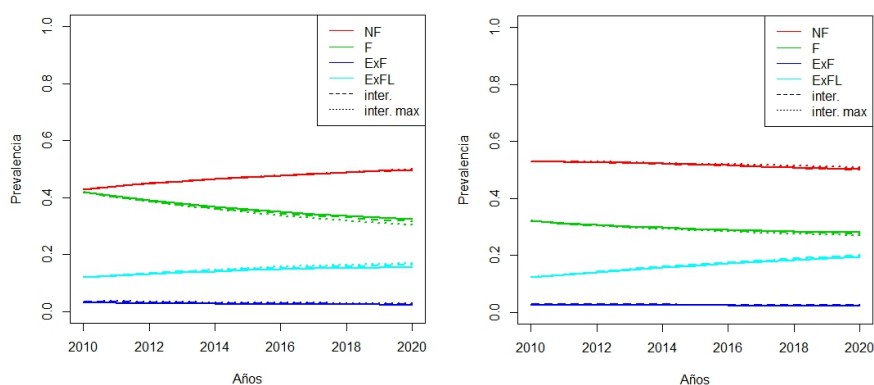


Figura 11.7: Predicciones de prevalencias para el grupo de edad 25-44 con intervenciones, para hombres y mujeres, respectivamente.

Año	NF	NF int.	NF int. máx.	F	F int.	F int. máx.
2010	0.530	0.530	0.530	0.320	0.320	0.320
2011	0.529	0.528	0.529	0.313	0.311	0.310
2012	0.527	0.526	0.528	0.307	0.305	0.303
2013	0.525	0.523	0.526	0.302	0.299	0.297
2014	0.522	0.520	0.524	0.297	0.294	0.292
2015	0.519	0.517	0.521	0.293	0.290	0.287
2016	0.516	0.513	0.518	0.290	0.286	0.283
2017	0.513	0.510	0.516	0.287	0.283	0.279
2018	0.509	0.506	0.513	0.285	0.280	0.276
2019	0.506	0.503	0.510	0.282	0.278	0.273
2020	0.503	0.499	0.507	0.280	0.275	0.270
Año	ExF	ExF int.	ExF int. max.	ExFL	ExFL int.	ExFL int. max.
2010	0.027	0.027	0.027	0.123	0.123	0.123
2011	0.026	0.028	0.029	0.132	0.132	0.132
2012	0.025	0.028	0.028	0.141	0.142	0.141
2013	0.025	0.027	0.027	0.149	0.151	0.150
2014	0.024	0.027	0.027	0.157	0.159	0.158
2015	0.024	0.026	0.026	0.164	0.167	0.166
2016	0.024	0.026	0.026	0.170	0.175	0.173
2017	0.023	0.025	0.025	0.177	0.182	0.180
2018	0.023	0.025	0.025	0.183	0.188	0.186
2019	0.023	0.025	0.025	0.189	0.195	0.192
2020	0.023	0.025	0.025	0.194	0.201	0.198

Cuadro 11.15: Predicciones de prevalencias para mujeres de edades 25-44 con intervenciones.

En el grupo de edad 3 (45-64) el efecto es análogo al del grupo de edad anterior, aunque ligeramente más leve. La reducción, tanto en hombres como en mujeres, de la prevalencia de fumadores en el 2020, con respecto a las predicciones sin intervenciones, alcanza en el mejor de los casos un 1,5 %, como se ve en los Cuadros 11.16 y 11.17.

En la Figura 11.8 se aprecia que, en este caso, apenas hay diferencias entre el efecto normal de las intervenciones y su efecto máximo. Parece que, a estas edades, un mayor impacto de las políticas no supone claramente una mayor reducción en la prevalencia de fumadores.

Año	NF	NF int.	NF int. máx.	F	F int.	F int. máx.
2010	0.276	0.276	0.276	0.298	0.298	0.298
2011	0.278	0.279	0.278	0.295	0.293	0.291
2012	0.281	0.281	0.280	0.292	0.288	0.286
2013	0.284	0.284	0.283	0.288	0.282	0.281
2014	0.287	0.287	0.285	0.284	0.277	0.275
2015	0.290	0.291	0.288	0.280	0.272	0.270
2016	0.293	0.294	0.291	0.275	0.266	0.264
2017	0.296	0.297	0.294	0.271	0.261	0.258
2018	0.299	0.301	0.297	0.267	0.255	0.253
2019	0.303	0.304	0.300	0.262	0.250	0.247
2020	0.306	0.307	0.303	0.258	0.245	0.242
Año	ExF	ExF int.	ExF int. max.	ExFL	ExFL int.	ExFL int. max.
2010	0.020	0.020	0.020	0.405	0.405	0.405
2011	0.020	0.022	0.025	0.406	0.407	0.406
2012	0.020	0.022	0.024	0.407	0.409	0.409
2013	0.019	0.021	0.024	0.409	0.412	0.413
2014	0.019	0.021	0.023	0.410	0.415	0.416
2015	0.019	0.020	0.023	0.412	0.417	0.419
2016	0.019	0.020	0.022	0.413	0.420	0.423
2017	0.018	0.020	0.022	0.415	0.423	0.426
2018	0.018	0.019	0.021	0.416	0.425	0.429
2019	0.018	0.019	0.021	0.417	0.427	0.432
2020	0.017	0.018	0.020	0.419	0.429	0.435

Cuadro 11.16: Predicciones de prevalencias para hombres de edades 45-64 con intervenciones.

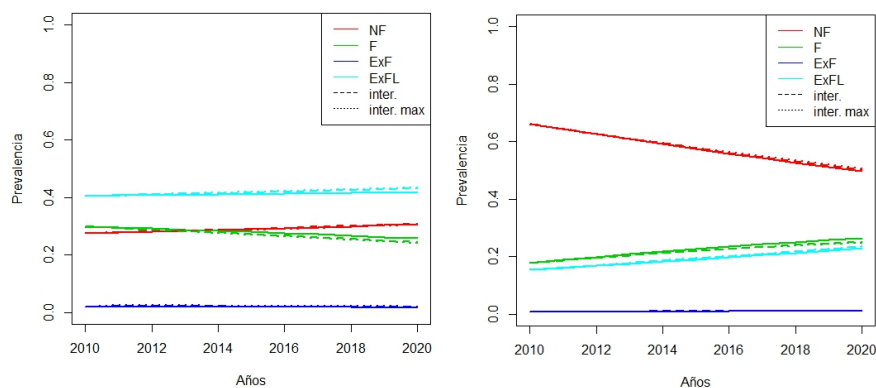


Figura 11.8: Predicciones de prevalencias para el grupo de edad 45-64 con intervenciones, para hombres y mujeres, respectivamente.



Año	NF	NF int.	NF int. máx.	F	F int.	F int. máx.
2010	0.661	0.661	0.661	0.177	0.177	0.177
2011	0.643	0.644	0.644	0.188	0.187	0.186
2012	0.625	0.627	0.627	0.199	0.196	0.195
2013	0.608	0.610	0.610	0.209	0.204	0.204
2014	0.591	0.594	0.594	0.218	0.212	0.212
2015	0.574	0.577	0.578	0.227	0.220	0.219
2016	0.558	0.561	0.562	0.235	0.227	0.226
2017	0.542	0.546	0.547	0.243	0.234	0.232
2018	0.527	0.531	0.533	0.250	0.240	0.238
2019	0.512	0.517	0.519	0.257	0.245	0.243
2020	0.498	0.503	0.505	0.263	0.250	0.248
Año	ExF	ExF int.	ExF int. max.	ExFL	ExFL int.	ExFL int. max.
2010	0.007	0.007	0.007	0.154	0.154	0.154
2011	0.007	0.008	0.008	0.161	0.161	0.161
2012	0.008	0.009	0.009	0.168	0.169	0.169
2013	0.008	0.009	0.009	0.175	0.177	0.177
2014	0.009	0.009	0.009	0.183	0.185	0.185
2015	0.009	0.010	0.010	0.190	0.193	0.193
2016	0.009	0.010	0.010	0.198	0.202	0.202
2017	0.010	0.010	0.010	0.206	0.210	0.210
2018	0.010	0.010	0.011	0.213	0.219	0.219
2019	0.010	0.011	0.011	0.221	0.227	0.227
2020	0.010	0.011	0.011	0.229	0.236	0.236

Cuadro 11.17: Predicciones de prevalencias para mujeres de edades 45-64 con intervenciones.

En el último grupo de edad (65 y más) la prevalencia de fumadores es la más baja de toda la población. La probabilidad de dejar de fumar afecta a un menor número de individuos y el efecto de las intervenciones, por lo tanto, será menos perceptible que en edades más jóvenes.

En los hombres se observa un descenso en la prevalencia de fumadores de apenas un 1% en 2020, como se ve en el Cuadro 11.18. En las mujeres las variaciones son inapreciables (Cuadro 11.2).

En la Figura 11.9 apenas se distingue la diferencia entre las predicciones normales y aquellas que consideran el efecto de las intervenciones. Se trata del grupo de edad menos influenciado por las intervenciones.

Año	NF	NF int.	NF int. máx.	F	F int.	F int. máx.
2010	0.397	0.397	0.397	0.098	0.098	0.098
2011	0.390	0.391	0.390	0.092	0.090	0.090
2012	0.384	0.385	0.385	0.086	0.083	0.082
2013	0.379	0.380	0.380	0.081	0.077	0.075
2014	0.374	0.376	0.375	0.076	0.071	0.070
2015	0.370	0.372	0.371	0.072	0.066	0.065
2016	0.367	0.369	0.368	0.068	0.062	0.061
2017	0.364	0.366	0.364	0.065	0.058	0.057
2018	0.361	0.363	0.362	0.061	0.055	0.053
2019	0.359	0.361	0.359	0.058	0.052	0.050
2020	0.357	0.359	0.357	0.056	0.049	0.047

Año	ExF	ExF int.	ExF int. max.
2010	0.504	0.504	0.504
2011	0.518	0.519	0.520
2012	0.530	0.532	0.533
2013	0.540	0.543	0.545
2014	0.550	0.553	0.555
2015	0.558	0.562	0.564
2016	0.565	0.569	0.572
2017	0.572	0.576	0.579
2018	0.578	0.582	0.585
2019	0.583	0.587	0.591
2020	0.588	0.592	0.596

Cuadro 11.18: Predicciones de prevalencias para hombres de edades 65 y más con intervenciones.

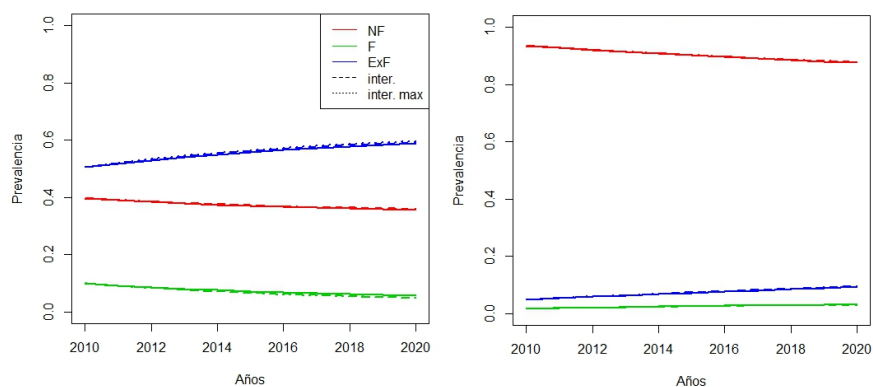


Figura 11.9: Predicciones de prevalencias para el grupo de edad 65 y más con intervenciones, para hombres y mujeres, respectivamente. En el gráfico de mujeres se sigue la misma leyenda que en el de hombres.

Año	NF	NF int.	NF int. máx.	F	F int.	F int. máx.
2010	0.935	0.935	0.935	0.016	0.016	0.016
2011	0.928	0.928	0.928	0.018	0.018	0.018
2012	0.921	0.921	0.921	0.020	0.020	0.020
2013	0.915	0.915	0.915	0.022	0.022	0.022
2014	0.909	0.909	0.909	0.024	0.023	0.023
2015	0.903	0.903	0.903	0.025	0.025	0.025
2016	0.897	0.897	0.897	0.026	0.026	0.026
2017	0.892	0.892	0.892	0.028	0.027	0.027
2018	0.886	0.886	0.887	0.029	0.028	0.028
2019	0.881	0.881	0.881	0.030	0.028	0.028
2020	0.876	0.876	0.876	0.031	0.029	0.029

Año	ExF	ExF int.	ExF int. max.
2010	0.049	0.049	0.049
2011	0.054	0.054	0.054
2012	0.059	0.059	0.059
2013	0.063	0.063	0.063
2014	0.068	0.068	0.068
2015	0.072	0.073	0.073
2016	0.076	0.077	0.077
2017	0.081	0.081	0.081
2018	0.085	0.086	0.086
2019	0.089	0.090	0.090
2020	0.094	0.095	0.095

Cuadro 11.19: Predicciones de prevalencias para mujeres de edades 65 y más con intervenciones.

En global, el efecto de las intervenciones apenas supondría un aumento en la prevalencia de nunca fumadores. El descenso máximo en la prevalencia de fumadores sería de un 2 % en los hombres y de un 0,5 % en las mujeres, como se ve en los Cuadros 11.2 y 11.2 respectivamente. Así mismo, el aumento en la prevalencia de exfumadores superaría el 1 % en los hombres, mientras que en las mujeres no llegaría al 0,5 %.

En la Figura 11.10 se aprecia el pequeño descenso que se produce en la prevalencia de fumadores hombres, con su correspondiente aumento en exfumadores y, muy levemente, en nunca fumadores. En el caso de las mujeres los cambios son inapreciables.

Año	NF	NF int.	NF int. máx.	F	F int.	F int. máx.
2010	0.394	0.394	0.394	0.307	0.307	0.307
2011	0.393	0.394	0.394	0.300	0.298	0.295
2012	0.392	0.393	0.395	0.293	0.290	0.286
2013	0.392	0.393	0.395	0.287	0.282	0.277
2014	0.392	0.393	0.396	0.280	0.275	0.268
2015	0.393	0.394	0.397	0.274	0.267	0.260
2016	0.393	0.395	0.398	0.267	0.260	0.252
2017	0.394	0.396	0.399	0.261	0.254	0.245
2018	0.395	0.397	0.401	0.255	0.247	0.237
2019	0.396	0.399	0.403	0.250	0.241	0.231
2020	0.398	0.400	0.404	0.244	0.235	0.224

Año	ExF	ExF int.	ExF int. max.
2010	0.299	0.299	0.299
2011	0.307	0.308	0.310
2012	0.314	0.317	0.319
2013	0.321	0.325	0.328
2014	0.328	0.332	0.336
2015	0.334	0.339	0.343
2016	0.340	0.345	0.350
2017	0.345	0.350	0.356
2018	0.350	0.356	0.362
2019	0.354	0.360	0.367
2020	0.358	0.365	0.371

Cuadro 11.20: Predicciones de prevalencias para hombres con intervenciones.

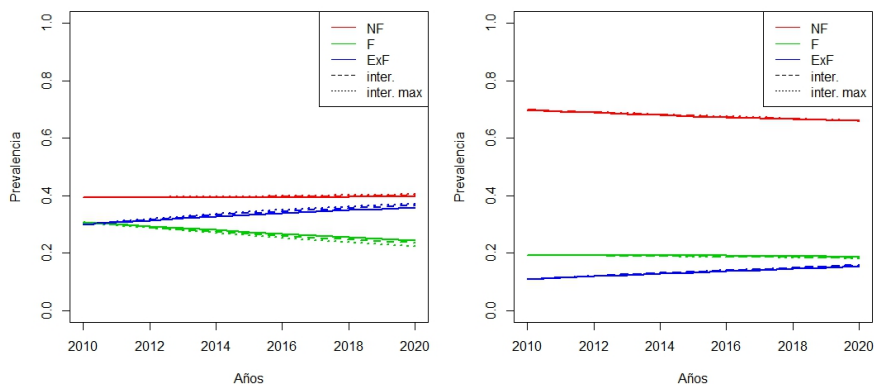


Figura 11.10: Predicciones de prevalencias para hombres y mujeres, respectivamente, con intervenciones.

Año	NF	NF int.	NF int. máx.	F	F int.	F int. máx.
2010	0.698	0.698	0.698	0.193	0.193	0.193
2011	0.693	0.693	0.694	0.193	0.192	0.191
2012	0.689	0.689	0.689	0.193	0.191	0.190
2013	0.685	0.685	0.686	0.192	0.190	0.190
2014	0.680	0.681	0.682	0.192	0.189	0.189
2015	0.677	0.677	0.678	0.191	0.188	0.187
2016	0.673	0.673	0.675	0.190	0.187	0.186
2017	0.669	0.669	0.671	0.190	0.186	0.185
2018	0.666	0.666	0.668	0.189	0.185	0.184
2019	0.663	0.663	0.665	0.188	0.184	0.183
2020	0.659	0.659	0.662	0.187	0.183	0.182

Año	ExF	ExF int.	ExF int. max.
2010	0.109	0.109	0.109
2011	0.114	0.115	0.115
2012	0.119	0.120	0.120
2013	0.123	0.125	0.125
2014	0.128	0.130	0.130
2015	0.132	0.135	0.135
2016	0.137	0.140	0.139
2017	0.141	0.144	0.144
2018	0.145	0.149	0.148
2019	0.149	0.153	0.153
2020	0.153	0.158	0.157

Cuadro 11.21: Predicciones de prevalencias para mujeres con intervenciones.

En conclusión, el efecto conjunto de las intervenciones detalladas en el Capítulo 9 supone una pequeña reducción de la prevalencia de fumadores. El efecto es más notable en los hombres, sobre todo entre los más jóvenes, aunque en global no supone un gran cambio. En general, las tendencias, en un plazo de 10 años, no se ven alteradas en función de las políticas implantadas en materia de tabaco.

Cabe destacar la reducción observada entre los jóvenes que empiezan a fumar. Un mayor impacto en las campañas dirigidas a este público sí se traduce en una mayor reducción en la prevalencia de fumadores. Deberían, por lo tanto, incentivarse las políticas que tienen como público objetivo a los menores de 25 años, favoreciendo una reducción en el número de jóvenes fumadores que, en un futuro, supondría una reducción en la prevalencia global de fumadores.

Conforme avanza la edad de los individuos, se vuelven menos receptivos ante las intervenciones, por lo que apenas se modifican las prevalencias. El efecto de políticas dirigidas al público adulto, como la ley de tabaco con la que se prohíbe fumar en los lugares de trabajo no repercute en un descenso importante de la prevalencia de fumadores.



## Capítulo 12

# Líneas abiertas

A lo largo de este trabajo se han encontrado dificultades en la aplicación de algunas de las técnicas expuestas en los Capítulos 3 y 4. Estas dificultades vienen provocadas por las características de los datos utilizados o bien por la falta de procedimientos específicos para CMTD. El estudio detallado de las mismas excede los límites y objetivos del trabajo aquí presentado, por lo que se plantean como líneas abiertas para futuras investigaciones.

Es el caso de la suavización de la matriz de transiciones vista en la Sección 4.2. En la aplicación de suavizadores tipo núcleo a CMTD no se han encontrado procedimientos que justifiquen la elección del parámetro de suavizado  $h$ . Se pueden plantear criterios bajo los que la elección de dicho parámetro se realice basándose en la distribución de la CMTD en la etapa  $n$ -ésima.

Los contrastes de hipótesis vistos en la Sección 3.1.3. se han probado sobre los resultados obtenidos de la aplicación de los métodos desarrollados. Los resultados de los mismos no pueden ser considerados determinantes ya que las características de los datos, procedentes de una muestra compleja, alteran las propiedades de los mismos. Se han indicado correcciones que adaptan los estadísticos utilizados en los contrastes teniendo en cuenta el diseño muestral. Para poder realizar dichas correcciones sería necesario un estudio más pormenorizado de los estimadores de máxima verosimilitud de las probabilidades de transición, para poder extraer las propiedades necesarias para dichas correcciones.





# Apéndice A

## Software

En este trabajo se han implementado los métodos vistos en los paquetes estadísticos R y Stata. Stata es el programa estadístico utilizado en la DXIXSP para el estudio del SICRI, así mismo, los datos de las diversas ediciones del SICRI se almacenan en bases de datos tipo `.dta`.

En R se han utilizado las siguientes librerías:

- `foreign`: paquete que permite la lectura en R de archivos de base de datos tipo `.dta`, propios de Stata.
- `survey`: paquete que permite realizar estimaciones incluyendo el diseño muestral. Se ha utilizado para obtener las estimaciones de las probabilidades de transición según el diseño muestral del SICRI.

Además de estas librerías se han implementado funciones para realizar los siguientes procedimientos:

- Cálculo de la distribución inicial de la CMTD. Obtención del estimador de máxima verosimilitud de la matriz de transiciones. Cálculo de la distribución de la CMTD en la etapa  $n$ -ésima.
- Cálculo de intervalos de confianza para las probabilidades de transición, basados en el diseño muestral y en la distribución asintóticas de los estimadores.
- Implementación del algoritmo Bootstrap visto en el Capítulo 4, junto con el cálculo de los intervalos de confianza para las probabilidades de transición y para la distribución de la CMTD en la etapa  $n$ -ésima.
- Implementación de los contrastes de hipótesis vistos en la Sección 3.1.3. con sus respectivas correcciones, a excepción de la propuesta por Rao [27].
- Implementación del estimador suavizado de la matriz de transición.

- Implementación de las fórmulas necesarias para realizar las predicciones de prevalencias vistas en el Capítulo 8.
- Implementación de las modificaciones vistas en el Capítulo 9 para estudiar el impacto de las políticas.

# Bibliografía

- [1] World Health Organization (2008). *WHO Report on the Global Tobacco Epidemic, 2008: the MPOWER package*. Geneva.
- [2] Ezzati M, Lopez AD. (2003). Estimates of global mortality attributable to smoking in 2000. *Lancet*. 2003 Sep 13;362(9387):847-52.
- [3] Shafey O, Eriksen M, Ross H, Mackay J. (2009) The tobacco atlas.3<sup>a</sup> ed. *American Cancer Society* ,Atlanta.
- [4] European Commission. (2010) *Special Eurobarometer 332. Tobacco*. Brussels.
- [5] Ministerio de Sanidad y Consumo. *Encuesta Nacional de Salud de España, 2006*. Disponible en: <http://www.ine.es/>
- [6] Ministerio de Sanidad y Consumo. *Encuesta Nacional de Salud 2003*. Disponible en: <http://www.ine.es/inebase>
- [7] U.S Department of Health and Human Services, (1994). Preventing tobacco use among young people. A report of the Surgeon General, 1994. *Public Health Service, Center for Disease Control and Prevention, Office on Smoking and Health*. Atlanta, Georgia.
- [8] Grupo de trabajo de tabaco de la Sociedad Española de Epidemiología. (2009). *Evaluación del impacto de la Ley de medidas sanitarias frente al tabaquismo*. Madrid.
- [9] Fernández E, Schiaffino A, García M, Saltó E, Villalbí J, Borrás J. (2003). *Prevalencia de consumo de tabaco en España entre 1945 y 1995. Reconstrucción a partir de Encuestas Nacionales de Salud*. *Med Clin (Barc)*. 2003;120(1):14-6.
- [10] Lopez MJ, Perez-Rios M, Schiaffino A, Nebot M, Montes A, Ariza C, et al. (2007). Mortality attributable to passive smoking in Spain, 2002. *Tob Control*. 2007 Dec;16(6):373-7.
- [11] Ley 28/ 2005 de 26 de diciembre, de medidas sanitarias frente al tabaquismo y reguladora de la venta, el suministro, el consumo y la publicidad de los productos del tabaco (BOE n<sup>o</sup> 309, del martes 27 de diciembre de 2005).

- [12] Ley 42/2010, de 30 de diciembre, por la que se modifica la Ley 28/2005, de 26 de diciembre, de medidas sanitarias frente al tabaquismo y reguladora de la venta, el suministro, el consumo y la publicidad de los productos del tabaco (BOE nº 318, del viernes 31 de diciembre de 2010).
- [13] Pacific Institute for Research and Evaluation. Cancer Intervention and Surveillance Modeling Network. Lung Cancer Model Profiles. SimSmoke. Disponible en: <http://cisnet.cancer.gov/lung/profiles.html>
- [14] Levy, D. T., Cummings, K. M., Hyland, A. (2000) Increasing Taxes as a Strategy to Reduce Cigarette Use and Deaths: Results of a Simulation Model. *American Health Foundation and Academic Press*.
- [15] Killeen, P. R. (2011) Markov model of smoking cessation *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108.
- [16] Méndez, D., Warner, K. E., Courant, P. N. (1998) Has Smoking Cessation Ceased? Expected Trends in the Prevalence of Smoking in the United States. *American Journal of Epidemiology*, Vol. 148, No. 3.
- [17] Nagelhout, G. E., Levy, D. T., Blackman, K., Currie, L., Clancy, L., Willemssen, M. C. (2011) The effect of tobacco control policies on smoking prevalence and smoking-attributable deaths. Findings from the Netherlands SimSmoke Tobacco Control Policy Simulation Model. *Addiction*, 107.
- [18] Méndez, D., Alshanteety, O., Warner, K. E. (2012) The potential impact of smoking control policies on future global smoking trends. *Tobacco Control*.
- [19] Kulkarni, V. G. (1995) *Modeling and Analysis of Stochastic Systems*. Text in Statistical Science, Chapman & Hall/CRC, Boca Raton, Florida.
- [20] Durrett, R. (1999) *Essentials of Stochastic Processes*. Springer, New York, USA.
- [21] Anderson, T. W. and Goodman, L.A. (1956) Statistical Inference about Markov Chains. *The Annals of Mathematical Statistics*. Vol. 28, No. 1.
- [22] Bath, B. R. (1960) Maximum Likelihood Estimation for Positively Regular Markov Chains. *Sankhya: The Indian Journal of Statistics*, Vol. 22, No. 3/4.
- [23] Bath, B. R. (1961) Some Properties of Regular Markov Chains. *The Annals of Mathematical Statistics*, Vol. 32, No. 1.
- [24] Teodorescu, J. (2009) Maximum Likelihood Estimation for Markov Chains. arXiv:0905.4131v1 [stat.CO].
- [25] Cressie, N. and Read, T. R. C. (1984) Multinomial Goodness-of-Fit Tests. *Journal of the Royal Statistical Society. Series B*, vol. 46, No. 3.

- [26] Smith P. J. et al. (1981) Approximating the Moments and Distribution of the Likelihood Ratio Statistic for Multinomial Goodness of Fit. *Journal of the American Statistical Association*, Vol. 76, No. 375.
- [27] Rao, J. N. K. and Scott A. J. (1981) The Analysis of Categorical Data From Complex Sample Surveys: Chi-Squared Tests for Goodness of Fit and Independence in Two-Way Tables. *Journal of the American Statistical Association*, Vol. 76, No. 374.
- [28] Kulperger, R. J. and Prakasa Rao, B. L. S. (1989) Bootstrapping a Finite State Markov Chain. *Sankhya: The Indian Journal of Statistics, Series A*, Vol. 51, No. 2.
- [29] Titterington, D. M. (1980) A Comparative Study of Kernel-Based Density Estimates for Categorical Data. *Technometrics*, Vol. 22, No. 2.

