



# Estimación de Parámetros de Volatilidad Electoral empleando métodos de Inferencia Ecológica

Tesis para optar al grado  
de Magister en Estadística Aplicada

---

Autor:  
Pablo Sandoval

Directora: Silvia Ojeda.

Fecha finalización: Agosto, 2019.

Fecha aprobación: Marzo, 2021.



Estimación de Parámetros de Volatilidad Electoral empleando métodos de Inferencia Ecológica by Pablo Sandoval is licensed under a [Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

## Resumen

Esta tesis propone una estrategia para medir la composición del electorado en siete categorías de votantes tras dos procesos electorarios consecutivos, empleando para ello métodos de estimación con datos agregados (Inferencia Ecológica). Dos de estas siete categorías son 1) Votante estable: que sufraga por la misma opción política (partido o coalición) en ambos comicios; y 2) Votante transferido: que cambia su preferencia electoral entre la oferta instituida en la contienda electoral en ambos procesos.

Se realizó una revisión de antecedentes sobre métodos estadísticos para la estimación de matrices de transición electoral con datos agregados. Posteriormente, estos métodos fueron vinculados con el problema de la composición del electorado a través de un desarrollo algebraico de carácter inédito.

Conocer la composición del electorado en estas siete categorías resulta ser un asunto de alta relevancia en el campo de la sociología política, pues se relaciona con el tipo de vinculación entre el conjunto societario y el sistema de representación política. Al abarcar un periodo más extenso que sólo dos votaciones consecutivas, este tipo de evidencia empírica contribuye a juzgar la existencia de una estructura de clivajes sociales a la base del funcionamiento de un sistema democrático. Este campo de investigación es de relevancia también para la ciencia política: parte importante de las fortalezas institucionales de un sistema democrático, así como de las posibilidades de estructuración de una representación programática efectiva, se deciden en función del grado de regularidad de los patrones de votación.

La propuesta fue aplicada al caso de la competencia electoral entre coaliciones políticas en elecciones de cámara baja Chile para el periodo comprendido entre los años 1993 y 2009. Los resultados obtenidos resultaron altamente coincidentes con evidencia previa procedente de encuestas electorales. Este resultado constituye evidencia a favor de la posibilidad de efectuar esta clase de diagnóstico por medio de datos agregados, los que a diferencia de encuestas electorales, son de libre acceso en la mayor parte de los países del mundo, posibilitando, por tanto, un estudio comparativo y/o longitudinal de institucionalización partidista.

El documento fue generado con la librería knitr del software libre RStudio. Esta funcionalidad integra código  $\LaTeX$  para la elaboración de reportes, y código R para la realización de análisis estadístico, generación de tablas y gráficos. En el Anexo 5.1 se incluye el código del análisis, lo que posibilita la completa reproducibilidad de los resultados obtenidos.

# Índice general

<b>1. Planteamiento del Problema</b>	<b>2</b>
1.1. Estabilidad electoral y Desafección política . . . . .	2
1.2. Composición del Electorado . . . . .	4
1.3. Estimación con datos agregados: Inferencia Ecológica . . . . .	7
1.4. Síntesis . . . . .	10
1.5. Objetivos . . . . .	11
1.5.1. Objetivo General . . . . .	11
1.5.2. Objetivos Específicos . . . . .	11
<b>2. Desarrollo de la Propuesta</b>	<b>12</b>
2.1. Espacio Paramétrico del Problema . . . . .	12
2.1.1. Formalización del Espacio Paramétrico . . . . .	12
2.1.2. Parámetros de la Volatilidad Electoral . . . . .	16
2.1.3. Síntesis . . . . .	21
2.2. Métodos Estadísticos para la estimación de entradas de Matrices de transición electoral . . . . .	22
2.2.1. Matriz de Transición electoral . . . . .	22
2.2.2. Inferencia Ecológica . . . . .	22
2.2.3. Relación con Cadenas de Markov . . . . .	31
2.2.4. Síntesis . . . . .	33
2.3. Propuesta de Estimación con datos agregados . . . . .	34
2.3.1. Estimación de la Composición del Electorado con Métodos EI . . . . .	34
2.3.2. Acotación del Espacio Paramétrico . . . . .	37
2.3.3. Síntesis . . . . .	42
<b>3. Aplicación de la Propuesta a Datos reales</b>	<b>43</b>
3.1. Resumen de la Propuesta . . . . .	43
3.2. Aplicación al caso de Chile 1993-2009 . . . . .	44
3.2.1. Descripción Sintética del caso . . . . .	44
3.2.2. Datos . . . . .	45
3.3. Resultados . . . . .	48
3.3.1. Volatilidad Electoral Inter-Bloques 1993-2009 . . . . .	48
3.3.2. Aplicación de Métodos EI . . . . .	50
3.3.3. Comentarios . . . . .	58

3.4. Propuestas de mejora para el estudio de la Institucionalización partidista . . . . .	59
<b>4. Conclusiones</b>	<b>63</b>
<b>5. Anexos</b>	<b>65</b>
5.1. Trabajo con los Datos . . . . .	65
5.2. Espacio de Probabilidad del Estimador . . . . .	86
5.3. Métodos Montecarlo con Cadenas de Markov . . . . .	90
5.3.1. Terminología de Cadenas de Markov para la simulación Montecarlo . . . . .	90
5.3.2. Relación con Métodos de Inferencia Bayesiana . . . . .	92
5.4. El enfoque de la Institucionalización partidista de Mainwaring & Scully . . . . .	102
<b>Referencias Bibliográficas</b>	<b>104</b>

# Capítulo 1

## Planteamiento del Problema

### 1.1. Estabilidad electoral y Desafección política

En el transcurso de las últimas dos décadas, el estudio del grado de institucionalización de los sistemas de partidos se ha instalado como uno de los focos prioritarios de la investigación politológica en América Latina. En líneas generales, el concepto de institucionalización alude a la existencia de un conjunto de interacciones normadas en la competencia entre partidos, la existencia de al menos dos fracciones competidoras y una relativa continuidad de las mismas a lo largo del tiempo.

Mainwaring & Scully (1995), con foco inicialmente en los sistemas de partidos latinoamericanos, propusieron un *constructo* para medir este fenómeno bajo un *enfoque comparado*<sup>1</sup>. El constructo es integrado por 4 dimensiones –o variables latentes–: 1) estabilidad de las reglas y de la dinámica de contienda electoral; 2) arraigo social de los agentes de representación política (partidos y/o conglomerados); 3) legitimización de la democracia como régimen de gobierno; y 4) autonomía organizacional de los partidos. La traducción operacional del concepto dio lugar a un índice aditivo que combina información de las cuatro dimensiones que lo integran. En su versión inicial, la estabilidad fue medida a través del Índice de volatilidad electoral, mientras que para las otras tres dimensiones se recogía información de estudios de opinión pública y otros antecedentes que permitieran caracterizar el funcionamiento interno de los partidos.<sup>2</sup> A partir de entonces, numerosos estudios se han guiado por este marco conceptual para efectos del estudio de los sistemas de la región (Jones, 2005; Luna, 2008; Mainwaring, 1999; Mainwaring & Torcal, 2006; Mainwaring & Zoco, 2007; Shair-Rossenfield, 2009), lo que le ha valido también creciente aceptación para el estudio de los regímenes democráticos en otras regiones del mundo en vías de desarrollo (Luna, 2014).

Bajo la impronta de este marco de análisis, sin embargo, un tipo específico de sistema de partidos figura como caso paradójico. Se trata de sistemas que exhiben conjuntamente estabilidad a nivel de los resultados electorales agregados y desafección política en el plano de la vinculación de la ciudadanía con los agentes de representación democrática, como han sido entre otros los casos de Brasil y Chile (Zuco, 2015; Luna

---

<sup>1</sup>Esto es, en diferentes países.

<sup>2</sup>Los resultados de la primera aplicación dieron lugar a una tipología que distinguió entre “sistemas institucionalizados”, incluyendo a Venezuela, Costa Rica, Chile, Uruguay, Colombia y Argentina, “sistemas de partidos incipientes” y “sistemas de partidos hegemónicos en transición”.

& Altman, 2011); esto es, sistemas democráticos que ven convivir significativos índices de distanciamiento hacia los partidos políticos, deslegitimización de los mecanismos de representación democrática, e incluso descrédito de la democracia como régimen de gobierno, con una alta estabilidad de los patrones electorales.

La connotación paradójica asociada a este tipo de casos se origina en la adopción de dos supuestos en la traducción operacional del concepto de institucionalización partidista de Mainwaring y Scully.

La primera suposición consiste en que la volatilidad electoral observada entre eventos de votación consecutivos permite concluir acerca de la composición relativa del electorado entre votantes estables y no estables tras la realización del segundo proceso electoral. Por ejemplo, para Mainwaring & Zoco (2007) la volatilidad: “*refers to the aggregate turnover from one party to others, from one election to the next*” (Mainwaring & Zoco, 2007: 158). En consecuencia, de ser estable un sistema, la suma total de votos transferidos sería baja, lo que daría cuenta de la vigencia de lealtades electorales, o lo que viene a ser lo mismo, de la primacía del votante estable (que vota por la misma opción política) en la composición del electorado.

La volatilidad electoral se mide con el Índice de Pedersen (1979), cuya formulación viene dada por la siguiente expresión:

$$V = \frac{1}{2} \sum_{j=1}^k |v_{(t+1),j} - v_{(t),j}| \quad (1.1.1)$$

Donde  $v_{(\cdot),j}$  corresponde al porcentaje de votación relativa obtenida por el  $j$ -ésimo partido político en competencia en el proceso electoral, con  $j = 1, \dots, k$ , en tanto que  $t \in \mathbb{N}$  representa una primera realización del proceso, y  $t + 1$  la consecutiva. El valor del índice oscila entre 0 y 100, donde el máximo representa un 100 % de cambio entre los comicios en estudio.

El segundo supuesto es que los cuatro componentes del constructo se encuentran estrechamente relacionados entre sí, lo que vendría a permitir la medición de una solo de aquellos en representación de los restantes. En palabras de Mainwaring (1999): “*these four dimensions of institutionalization need not go together, but they almost always do. Conceptually, a party system could be fairly institutionalized along one dimension but empirically this is the exception*” (Mainwaring, 1999: 27). De acuerdo con Luna y Altman (2011), lo anterior equivale a suponer de modo implícito un alto grado de correlación lineal positiva entre los cuatro componentes que integran el constructo.

La adopción de este último supuesto ha permitido sostener el enfoque de investigación comparativa de la propuesta original, dificultado por la falta de comparabilidad de las fuentes de información necesarias para atribuir una puntuación a los componentes de enraizamiento, legitimidad y autonomía organizacional de los partidos. En cambio, la volatilidad electoral sí ofrecía una medida comparable, y cuyo cómputo resulta relativamente accesible en los diferentes sistemas analizados. Progresivamente, entonces, el diagnóstico de la institucionalización hubo de trasladarse a la medición del grado de volatilidad electoral del sistema.

Sintetizando, para el enfoque de la institucionalización partidista (PSI, por sus siglas en inglés), grados de volatilidad cercanos a cero se asumen indicativos de que la mayor parte del electorado mantiene estabilidad en su comportamiento de voto (supuesto 1), lo cual explicaría la regularidad de los resultados

de votación, haciéndose luego esta constatación heurísticamente extensible al diagnóstico del sistema en su conjunto (supuesto 2), es decir, abarcando el arraigo social de los agentes de representación, la legitimidad del sistema y la autonomía organizacional de los partidos. De allí que la existencia de sistemas de partidos que exhiben conjuntamente estabilidad electoral y desafección política figure, bajo esta impronta de análisis, como casos anómalos.

A la base del carácter paradójico otorgado a la coexistencia entre estabilidad electoral y desafección política subyace, entonces, la interpretación de la volatilidad electoral como indicador de la composición del electorado entre votantes estables y no estables tras la celebración del segundo de dos procesos de votación consecutivos.

El argumento que se desarrolla a continuación muestra que, antes que ceñirse exclusivamente al enfoque PSI, la interpretación de la volatilidad como indicador del comportamiento de voto no estable permanece hasta la actualidad como problema no resuelto en el escenario más amplio de la investigación politológica comparativa. Una descripción más detallada del enfoque PSI es incluida en el Anexo 5.4 del documento.

## 1.2. Composición del Electorado

Partiendo de la taxonomía de Sani (1986), el electorado existente en un sistema democrático tras la realización del segundo de dos eventos de votación consecutivos del mismo tipo que solicitan al electorado emitir una única preferencia entre el conjunto de candidaturas inscritas en la contienda electoral, puede clasificarse en las siguientes siete categorías:

1. Votante estable: votan por la misma opción política en ambas elecciones.
2. Votante transferido: votan por opciones políticas diferentes en ambos comicios.
3. Desmovilizado: vota en la primera elección y se abstiene en la segunda.
4. Movilizado: se abstiene en la primera elección y vota en la segunda.
5. Abstencionista permanente: se abstiene en ambas votaciones.
6. Nuevo concurrente: no pertenecían al padrón en la primera elección, se integran a éste en la segunda y votan por alguna opción política.
7. Nuevo abstencionista: no pertenecían al padrón en la primera elección, se integran a éste en la segunda, y se abstienen.

Por la “composición del electorado” se hace referencia aquí al peso relativo de cada una de las anteriores categorías una vez efectuada la segunda de las dos votaciones objeto de análisis. El estudio de esta composición es de relevancia en el campo de la sociología política, pues se relaciona con el tipo de vinculación entre el conjunto societario y el sistema de representación<sup>3</sup>. Al abarcar un periodo más extenso que sólo dos

---

<sup>3</sup>En función de la eventual primacía de alguno de los anteriores tipos de electores se puede diagnosticar la preponderancia de un tipo de vinculación u otro (ejemplo: político-programática, mercantil, clientelar; etc.), o bien la diversificación de este tipo de vinculación, en caso de que ninguno de las categorías resulte predominante.

votaciones consecutivas, este tipo de evidencia empírica contribuye a juzgar la existencia de una estructura de “clivajes sociales” a la base del funcionamiento del sistema democrático. Este campo de investigación es de relevancia también para la ciencia política. Parte importante de las fortalezas institucionales de un sistema democrático, así como de las posibilidades de estructuración de una representación programática efectiva, se deciden en función del grado de regularidad de los patrones de votación (Torcal, 2015)<sup>4</sup>.

Un denominador común al interés de la sociología y la ciencia política en este tema consiste en poder realizar esta clase de diagnóstico a nivel comparado, esto es, entre diferentes países. Cuando se habla de sostener un “enfoque comparativo” se hace alusión a un programa de investigación politológica enmarcada no sólo en la realidad de un país, sino en una región geográfica (por ejemplo, América Latina, Europa Oriental, etc.).

Entre las alternativas metodológicas para estimar esta composición figuran el uso de encuestas con preguntas retrospectivas de comportamiento de voto, o bien, recurrir a una muestra de panel sobre la cual realizar un seguimiento longitudinal a este tipo de conducta. Desafortunadamente, alrededor del mundo estas fuentes de información resultan ser escasas o intermitentes (Powell & Tucker, 2013; Shair-Rossenfield, 2008; Van del Meer et al., 2015), lo que significa un obstáculo para el abordaje comparativo del problema. Además, en caso de existir, las encuestas presentan problemas de fiabilidad de la respuesta por “recordación” u ocultamiento del comportamiento anterior por parte de los sujetos encuestados (Klima et al., 2017; Van de Ploeg et al., 2006); y tanto la metodología de panel, como las encuestas, no incorporan a individuos presentes en el padrón en sólo una de las dos votaciones<sup>5</sup>.

Frente a estas limitaciones, parte de la investigación internacional ha recurrido a cifras de volatilidad electoral. A priori, de constatarse drásticas variaciones en los resultados de votación en eventos electorales consecutivos, resulta intuitivo concluir que los votantes estables (tipo 1) no resultarán predominantes en la composición del electorado. Extendiendo esta intuición a la situación contraria, bajos índices de volatilidad electoral han sido entendidos como un resultado de la primacía de este último tipo de electores en dicha composición. Existen, no obstante, numerosos trabajos que refutan este último tipo de interpretación (Bartolini & Mair, 1990; Crewe & Denver, 1985; Roblizo, 2001). Por ejemplo, Luna & Altman (2011: 4-5) proporcionan un caso hipotético que permite ver que, en la situación límite, la volatilidad de un sistema puede ser nula, aún cuando la totalidad del electorado modifique su comportamiento de voto<sup>6</sup>. De allí, como ha señalado Roblizo (2001), que bajos –e inclusive nulos– niveles de volatilidad, pueden esconder un alto –pero auto-compensado– nivel de cambio en las preferencias de los electores.

Bartolini & Mair (1990) realizaron una importante contribución a este debate. Sintetizando aportacio-

---

<sup>4</sup>En sistemas con altos índices de volatilidad electoral, los agentes del sistema democrático (partidos y/o conglomerados) cuentan con menos incentivos para estructurar proyectos de representación política que gocen de gobernabilidad en el mediano plazo.

<sup>5</sup>Las encuestas pre o post la segunda de dos votaciones objeto de análisis no incorporan a individuos que abandonan el padrón previo a su celebración, presentes, sin embargo, en el primer comicio. Por su parte, las muestras de panel requerirían actualizarse permanentemente a efectos de representar el comportamiento de nuevos votantes.

<sup>6</sup>Esto ocurriría, suponiendo por ejemplo la presencia de dos opciones de representación, A y B, en dos votaciones sucesivas, además de ausencia de cambio en la composición del padrón electoral, en el caso que ambas concentraran un 50 % de las preferencias en la primera celebración del proceso, y perdieran luego la totalidad de su electorado, a favor de la otra, en la segunda.

nes previas, distinguieron a la “Volatilidad total” o neta, correspondiente a la medida por el Índice de Pedersen, de la “Volatilidad individual”, que definieron como el porcentaje de votantes no estables en la composición del electorado tras la realización del segundo de los comicios (lo que correspondería aquí al peso conjunto de los votantes 2 a 7 enunciados al inicio de esta sección). Del análisis de la evidencia empírica (encuestas o muestras de panel) disponible a la fecha, constataron una correlación lineal cercana a 0.7 entre ambas medidas. En paralelo a ello, realizaron un trabajo de simulación con matrices de transición electoral, lo cual arrojó un grado de correlación similar, de magnitud entre 0.6 y 0.7. Con base en todos estos antecedentes concluyeron que la volatilidad puede utilizarse como un indicador del porcentaje de votantes no estables: “*Such a degree of concurrence confirms the basic validity - given a sufficient number of cases- of considering aggregate volatility as one appropriate indicator of individual volatility. And while the relationship should be interpreted with some caution, it can be justified not only as a postulated heuristic device, but also as one which is both empirically and statistically likely*” (1990: 34).

Lo cierto es que la investigación politológica posterior se ha visto considerablemente influenciada, tanto por los hallazgos y desarrollo conceptual aportados por el trabajo de Bartolini y Mair, así como también por las consideraciones metodológicas que le sirvieron de sustento. En adelante, buena parte de los estudios realizados sobre este tema han recurrido a la interpretación de la volatilidad como indicador indirecto del voto no estable. Esta adopción se observa en los estudios de la institucionalización partidista en América Latina; en la investigación sobre el funcionamiento de los sistemas democráticos de Europa del Este, post caída del régimen soviético (Powell & Tucker, 2013; Sikk, 2005); en Europa central y occidental (Chiaramente & Emanuele, 2017; Dassonneville & Hooghe, 2011); y en otras regiones del mundo.

Sin embargo, llegado a este punto, cabe hacer notar que las conclusiones de Bartolini y Mair tomaron lugar en el marco de un estudio correlacional. En otras palabras, si la volatilidad agregada se encuentra altamente correlacionada con la volatilidad individual, cabría esperar entonces, por transitividad, que fenómenos altamente correlacionados con la volatilidad agregada también lo hagan con la individual.

Pese a ello, perduran en la actualidad dos tipos de indefiniciones en cuanto a su uso como indicador indirecto del voto no estable. En primer lugar, se identifica cierta persistencia por parte de algunos trabajos en otorgarle un alcance de medición indirecta del porcentaje de votantes no estables, o al menos, no explicitar el alcance otorgado en este ámbito. Por ejemplo, según Sikk (2005): “*In principle, volatility is an index that reflects the extent of personal vote transfers between subsequent elections (...) Even though there can be counter-examples where the correlation between personal and overall is poor, the index of volatility nevertheless approximates the level of individual vote shifts quite well*” (2005: 392), o de acuerdo con Noorudin y Chhibber (2008): “*Aggregate levels of electoral volatility, measured as the net change in vote shares for parties competing in the elections, are used as a convenient proxy for the cumulation of individual vote shifts*” (2008: 1073).

En segundo término, existe también una inclinación a identificar a la volatilidad individual como el porcentaje de votantes que modifican sus preferencias de votación entre la oferta electoral existente, vale decir, como el segundo tipo de votantes enunciados al inicio de esta sección, obviando de esta manera la existencia de los cinco restantes. Por ejemplo, según Powell & Tucker (2013) la volatilidad electoral, medida con el Índice de Pedersen: “*occurs when voters switch their votes between existing parties*” (2013: 123); o de acuerdo con las conclusiones del estudio comparado de Mainwaring et al. (2017), el manejo de la eco-

nomía por parte de los gobiernos: “*has no impact on vote transfers among established parties*” (2017: 624).

Este tipo de indefiniciones repercuten luego en los marcos analíticos existentes para el estudio de las democracias modernas. De acuerdo con Casal Bértoa (2016), diversos programas investigativos de nivel internacional confluyen en la volatilidad electoral como medida empírica del nivel de institucionalización de los sistemas de partidos. Uno de estos casos corresponde al enfoque de la institucionalización partidista de Mainwaring & Scully (1995), referenciado en la sección precedente. Para este marco de análisis, la existencia de sistemas de partidos electoralmente estables, pero crecientemente desarraigados de la ciudadanía, constituyen casos paradójicos, a consecuencia de la interpretación otorgada a la volatilidad electoral.

Con todo, recurrir a la volatilidad a efectos de determinar la composición del electorado en eventos de votación consecutivos resulta ser una alternativa problemática, al igual que ocurre con encuestas electorales o muestras de panel.

### 1.3. Estimación con datos agregados: Inferencia Ecológica

Una tercera alternativa para estimar la composición del electorado, hasta aquí no extensamente explorada en la investigación politológica, consiste en recurrir al modelamiento estadístico basado en la teoría de la probabilidad. En la literatura estadística existe un desarrollo generado a partir de mediados del siglo XX para la estimación de entradas de matrices de transición electoral, utilizando para ello datos agregados. Una matriz de transición electoral puede entenderse como una “tabla” cuyas filas contienen la proporción que cada agente de representación electoral (partidos o conglomerados políticos) retienen o pierden entre eventos de votación sucesivos, cuya suma de los elementos de cada fila es igual a 1.

En la actualidad, parte importante de estos modelos se agrupan bajo el rótulo de “Inferencia ecológica” (EI, por sus siglas en inglés), término con el que se hace referencia a la estimación de cantidades desagregadas (por ejemplo, a nivel individual) a partir del conocimiento de cifras censales, tanto en un nivel agregado superior (por ejemplo, nacional), como en unidades de desagregación intermedias (por ejemplo, provincias, municipios, etc.).

El uso del apelativo “ecológica” procede del trabajo de Robinson (1950), quien puso de relieve que las correlaciones observadas tomando como fuente de información estadísticas agregadas a nivel de sub-unidades territoriales, frecuentemente no coincidirán con las obtenidas al tomar como unidad de medición a la totalidad de los elementos que componen la población. Esto se debe a que tales cifras suelen ser promedios, y por consiguiente, tienen una dispersión asociada<sup>7</sup>.

A partir del trabajo de Robinson se denomina como “información ecológica” a aquella disponible en unidades agregadas (nacional, ciudades, provincias, etc.), dado el contexto de que se pretende interrelacionarla para extraer conclusiones acerca de los elementos que las componen. Los términos “variables ecológicas”

---

<sup>7</sup>Por ejemplo, si se efectúa la correlación entre cifras a nivel de provincias, tales como promedios de ingreso per cápita y de escolarización de la población, ello no implicará necesariamente que la correlación observada se replique al tomar como unidad de medición al universo de individuos que habitan dichas provincias. En torno a esto Goodman (1953: 663) provee un ejemplo que muestra que una alta correlación a nivel agregado puede tener en paralelo una nula correlación a nivel individual, aquello como caso extremo, para graficar el espectro de posibilidades involucradas en este tipo de problema.

o “correlación ecológica” son utilizados bajo la misma acepción. Posiblemente, la elección de este adjetivo tuvo por motivación la raíz etimológica de dicho término<sup>8</sup>, aunque esto no resulta totalmente claro desde el examen de la literatura disponible. El mismo Robinson introdujo el término “falacia ecológica” para denotar un tipo de error interpretativo originado de concluir atributos individuales a partir de relaciones agregadas.

En este trabajo los términos Inferencia ecológica, “Estimación con datos agregados” y “métodos EI” serán utilizados bajo el mismo significado.

En líneas generales, el problema de la estimación con datos agregados se compone de los siguientes elementos.

- Una población finita  $U$  de cardinalidad  $N$ .
- Dos variables categóricas  $W$  y  $Z$ , con  $R$  y  $C$  categorías, respectivamente, definidas sobre  $U$ .
- La distribución del número de elementos de la población en las categorías de cada variable es conocida por fuentes censales.
- Sin embargo, al proceder de medios censales diferentes –o como en el caso electoral, a consecuencia de la confidencialidad del voto– el número de elementos en la intersección de las categorías de ambas variables es desconocido.

Los valores desconocidos pueden ser representados al interior de una matriz con  $R$  filas (de rows) y  $C$  columnas. Para esquematizar se incluye la siguiente representación.

$$\begin{array}{cccc} \left[ \begin{array}{cccc} ? & \dots & \dots & ? \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ ? & \dots & \dots & ? \end{array} \right] & \begin{array}{l} w_1 \\ \dots \\ \dots \\ w_R \end{array} \\ z_1 & \dots & \dots & z_C & N \end{array}$$

Al dividir por  $N$  todas las cantidades al interior de la matriz, se obtiene el caso particular en que las magnitudes desconocidas –foco de la estimación– corresponden a los valores de la distribución conjunta de ambas variables categóricas, mientras que las sumas conocidas por filas y columnas equivalen a la distribución marginal de cada una de estas variables. En torno a esto último, la literatura especializada ha puesto de relieve que el conocimiento de la suma de los elementos por filas y columnas impone restricciones a las magnitudes desconocidas, las que se traducen en cotas superiores e inferiores para estas cantidades (Duncan & Davis, 1953). Proceder al cómputo de estas cotas permite arribar a intervalos con probabilidad 1 para las magnitudes desconocidas, los cuales, no obstante, resultan ser, si bien no triviales, demasiado amplios a efectos prácticos (King, 1997).

---

<sup>8</sup>Del griego: oikos: hogar, y logos: ciencia. Las primeras aplicaciones se refirieron a problemas de tipo demográfico, donde las unidades  $\{u_1, \dots, u_M\}$  corresponden a sub-unidades geográficas. Así, es posible que se haya querido relevar la pertenencia de los individuos en sus respectivas localizaciones geográficas.

Una segunda fuente de información presente en un problema EI deviene del conocimiento de la distribución de las mismas variables  $W$  y  $Z$  en un conjunto de unidades  $\{u_1, \dots, u_M\}$  que forman una partición de la población, cada uno de estos  $M$  subconjuntos con cardinalidad  $N_m$ ,  $m = 1, \dots, M$ . En aplicaciones donde los elementos de  $U$  resultan ser de tipo demográfico (individuos), la partición puede estar representada por provincias, regiones, municipios, o en general cualquier tipo de subdivisión territorial tal que forme una partición de  $U$ .

Haciendo uso de este último tipo de información, se han propuesto dos tipos de métodos para la estimación de entradas de matrices de transición electoral utilizando datos agregados.

1. Modelos Stayers-Movers. Se trata de determinar la matriz de transición electoral que minimiza algún criterio de discrepancia con respecto a los datos observados en la partición  $\{u_1, u_2, \dots, u_M\}$  recurriendo para ello a métodos de optimización con restricciones. Han sido desarrollados con foco exclusivo en la transición electoral, sin pretender alcanzar un método EI general.
2. Modelos con cotas determinísticas y supuestos distribucionales. Han buscado desarrollar métodos EI con aplicabilidad a diversos campos de investigación, no sólo en el ámbito electoral. Las contribuciones realizadas a la actualidad se rastrean en la discusión iniciada por Robinson (1950), y continuada luego por Duncan & Davis (1953), Goodman (1953; 1959), King (1997), Freedman (2001), King et al. (1999), Rosen et al. (2001), Greiner & Quinn (2009) y Klima et al. (2017), entre las principales.

Dentro del segundo tipo de modelos figuran una serie de aproximaciones que recurren a estructuras bayesianas jerárquicas (Greiner & Quinn, 2009; King et al., 1999; Rosen et al., 2001) y a métodos Montecarlo con Cadenas de Markov (Markov Chain Montecarlo, “MCMC”). De acuerdo con la evaluación de Klima et al. (2016), el modelo jerárquico Multinomial-Dirichlet de Rosen et al., implementado en la librería “eiPack” del software libre R ofrece estimaciones adecuadas en el contexto de problemas de transición electoral, aspecto comprobable a través de ejercicios de simulación: *“the hierarchical multinomial-Dirichlet model is superior to the other models in simulations with more realistic voter transition structures. This holds if the assumptions of the model are fulfilled, but also in the more realistic scenarios which do not comply with the distributional assumption of the model”* (2016: 153).

Por su parte, una aplicación reciente de modelos Stayers-Movers, junto a una descripción sintética de los mismos, se encuentra en el trabajo de Corominas et al. (2015). Parte de estos últimos métodos se encuentran implementados en la librería “vottrans” del mismo software R.

En ambos casos, los datos utilizados para la estimación corresponden a resultados de votación y cifras de participación electoral a nivel de las sub-unidades  $\{u_1, \dots, u_M\}$  para los dos comicios considerados. La distribución en las categorías de  $W$  corresponde aquí a resultados electorales y cifras de participación para el primer proceso, y en el caso de  $Z$ , a las cifras correspondientes al segundo.

Cabe señalar que este tipo de métodos se han concentrado en la estimación de las entradas de matrices de transición electoral, sin abordar el problema de la composición del electorado y su relación con los índices de volatilidad electoral. En este sentido, perdura hasta la actualidad lo constatado por Bartolini & Mair en relación a esta clase de modelos: *“the bulk of this literature focusses attention on the problem of how to use aggregate data to estimate individual volatility values, the goal being to measure the table*

*of transition probabilities derived from aggregate voting data, rather than to study the possible relation between the two” (1990: 30).*

Por su lado, los estudios politológicos basados en cifras de volatilidad electoral, entre los que se incluyen los marcos analíticos para el estudio del grado de institucionalización de los sistemas de partidos, no han recurrido a métodos de inferencia con datos agregados para estimar la composición del electorado.

## 1.4. Síntesis

En la investigación politológica internacional perduran dos tipos de indefiniciones en la interpretación de la volatilidad electoral: 1) su carácter como indicador del voto no estable, a saber, restringido a un plano correlacional, o bien como una medición indirecta de la segunda; y 2) los diferentes tipos de desplazamientos de votantes agrupados bajo el concepto de “volatilidad individual”: entre la oferta electoral existente, o en un espectro más amplio de posibilidades.

Este tipo de indefiniciones repercuten luego en los marcos analíticos existentes para el estudio de las democracias modernas, entre los cuales se encuentra la propuesta de Mainwaring & Scully para el estudio del grado de institucionalización de los sistemas de partidos. Para este enfoque la existencia de sistemas de partidos electoralmente estables, pero crecientemente desarraigados de la ciudadanía constituyen casos paradójicos.

En el trasfondo de la relación entre volatilidad agregada e individual subyace el problema de la composición del electorado en siete perfiles de votantes tras la realización de dos procesos electorales consecutivos. Debido al carácter secreto del voto esta composición es desconocida, pero puede ser estimada por medio de encuestas o muestras de panel. Esta clase de alternativas, sin embargo, presentan problemas de fiabilidad de respuesta por recordación u ocultamiento del comportamiento de voto, a la vez que resultan ser escasas y/o intermitentes a nivel internacional, lo que impone limitaciones para un abordaje comparativo y/o longitudinal del problema.

Frente a ello, una alternativa consiste en recurrir a métodos de estimación con datos agregados, lo que en la literatura estadística internacional recibe el nombre de “Inferencia ecológica”. Como principales ventajas asociadas al uso de este tipo de métodos figuran la libre disponibilidad de datos en la mayor parte de los países del mundo, y el encontrarse exentos de problemas de fiabilidad de encuestas o muestras de panel. En contrapartida, los desarrollos previos no han enfocado el problema de la composición del electorado, centrándose más bien en la estimación de entradas de matrices de transición electoral.

El propósito general del proyecto consiste en relacionar los métodos estadísticos de estimación de matrices de transición electoral utilizando datos agregados, con el problema de la composición del electorado tras eventos de votación consecutivos. Relacionar ambos tópicos vendría a posibilitar el diagnóstico comparativo de la composición del electorado en diversos sistemas de partidos y periodos electorales.

## 1.5. Objetivos

### 1.5.1. Objetivo General

Relacionar los métodos estadísticos de estimación de matrices de transición electoral por la vía del modelamiento con métodos de Inferencia ecológica, con el grado de volatilidad electoral constatado entre eventos de votación consecutivos, posibilitando así estimar la composición del electorado tras la celebración del segundo de aquellos procesos.

### 1.5.2. Objetivos Específicos

1. Formalizar el espacio muestral subyacente a la composición del electorado tras dos eventos electorarios consecutivos.
2. Desarrollar una formulación matricial del Índice de Pedersen como función de la composición relativa del electorado entre dos eventos de votación consecutivos.
3. Formalizar el espacio paramétrico del problema.
4. Describir los métodos de Inferencia Ecológica utilizados en la estimación de matrices de transición electoral y su relación con Cadenas de Markov.
5. Proponer un procedimiento general para la acotación del espacio paramétrico, que incorpore el método de las cotas determinísticas y otras relaciones entre las variables aleatorias involucradas en un problema de transición electoral.
6. Identificar alternativas para obtener valores esperados y varianzas de las variables aleatorias involucradas en el problema.
7. Incorporar el anterior desarrollo en las distribuciones previas de los modelos de Inferencia Ecológica jerárquicos para el caso de matrices  $RxC$  y evaluar los resultados obtenidos por medio de un ejercicio de simulación con el software R.
8. Aplicar la propuesta al caso de la contienda electoral entre coaliciones políticas en Chile entre los años 1990 y 2009.
9. A partir de lo anterior, proponer mejoras al abordaje empírico del estudio de la institucionalización partidista como marco analítico para el diagnóstico comparado de la calidad de la democracia.

## Capítulo 2

# Desarrollo de la Propuesta

### 2.1. Espacio Paramétrico del Problema

El desarrollo que se incluye en esta sección busca dar cumplimiento a los objetivos específicos número 3: Formalizar el espacio paramétrico del problema, y número 2: Desarrollar una formulación matricial del Índice de Pedersen como función de la composición del electorado. Las proposiciones 2.1.1 y 2.2.1 incluidas en este apartado son de carácter inédito.

#### 2.1.1. Formalización del Espacio Paramétrico

##### 2.1.1.1. Procesos Electorales

Las principales características que se hacen presentes en un proceso electoral pueden sintetizarse como sigue:

- La definición del tipo de autoridad de representación democrática a ser electa (gobernadores, diputados, etc.)
- Una serie de  $k$  partidos políticos, agrupados o no en conglomerados, a los cuales adscriben individuos habilitados por el proceso para recibir sufragios (candidatos); o bien,  $k - 1$  partidos y una opción adicional (“Otros”) bajo la cual pueden agruparse, a efectos del análisis, las candidaturas independientes.
- Una celebración del proceso eleccionario en un momento temporal determinado  $t : t \in \mathbb{N}$ .
- Un conjunto de votantes  $E_t$  o padrón electoral<sup>9</sup>.
- Un procedimiento de conteo de votos emitidos, junto con una regla –o aplicación– que permite traducir sumatorias de votos obtenidos en cargos de representación.

---

<sup>9</sup>Generalmente, la condición para formar parte del padrón dice relación con cumplir ciertos requisitos etarios, como la mayoría de edad, y no encontrarse inhabilitado para emitir un sufragio por alguna causa de tipo legal. En sistemas de inscripción no automática, se requiere, además de lo anterior, el registro voluntario del votante en dicho padrón.

Como componentes adicionales puede mencionarse: el establecimiento de la modalidad en que se los votantes materializan sus preferencias, a saber, presencial o de otro tipo; así como una periodicidad en la que el proceso electoral vuelve a celebrarse.

### 2.1.1.2. Tipo de Proceso electoral a examinar

Nos centraremos aquí en dos procesos electorales consecutivos con las siguientes características:

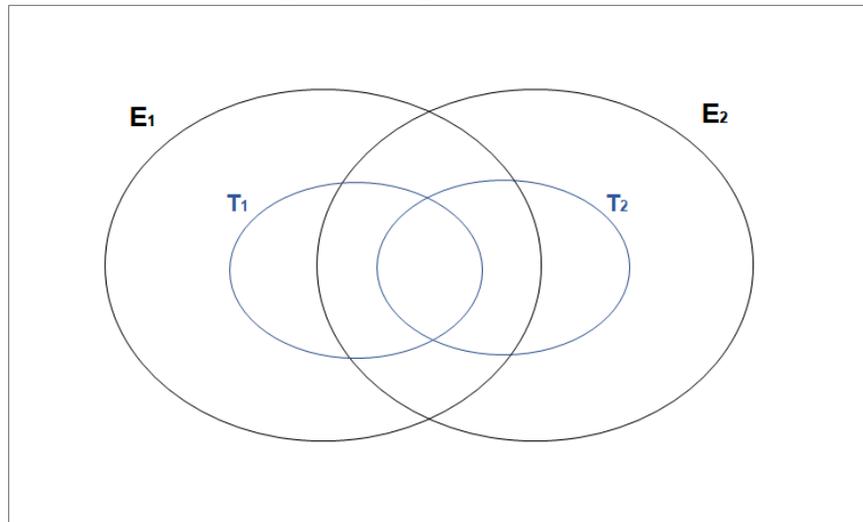
- Las autoridades democráticas a ser electas son las mismas en ambos procesos (mismo tipo de elecciones).
- Cada proceso eleccionario solicita a los votantes emitir una única preferencia.
- Los mismos  $k$  agentes de representación (partidos o coaliciones políticas) concentran la totalidad de la oferta electoral en ambos procesos.

Para el desarrollo que se presenta a continuación denotaremos los periodos de realización de los procesos electorales como  $t = 1$  y  $t + 1 = 2$ .

### 2.1.1.3. Conjuntos de votantes

Sean  $E_1$  y  $E_2$  los conjuntos de votantes que integran el padrón electoral en las votaciones 1 y 2 respectivamente;  $e$  un elemento genérico en  $E_1 \cup E_2$ ; y  $T_{(t)}$  el subconjunto de votantes en  $E_t$  que votan por alguna de las  $k$  opciones de representación<sup>10</sup>. Al agregar a este último conjunto el subíndice “ $j$ ”,  $T_{(t)j} \subseteq T_{(t)}$  representa el conjunto en el proceso  $t$  que vota por la opción  $j$ , con  $t = 1, 2; j = 1, \dots, k$ . Para esquematizar se incluye el siguiente diagrama de Venn.

Figura 1: Diagrama de Venn I



<sup>10</sup>De “voting Turnout at time  $t$ ”.

A efectos del recuento electoral en cada uno de estos procesos electorales, se distinguen  $k + 2$  estados que pueden asociarse a todo elemento  $e \in E_1 \cup E_2$ :

- Un total de  $k$  estados que corresponden a votar por alguno de los  $k$  agentes de representación en competencia:  $e \in T_{(t)j}, T_{(t)j} \subseteq E_{(t)}, t = 1, 2, j = 1, \dots, k$ .
- Abstencionismo:  $e \in E_t \cap T_{(t)}^c$ .
- Ausencia del padron,  $e \notin E_t$ .

La comparación entre estos  $k + 2$  estados en ambas votaciones permite proponer la siguiente partición de  $E_1 \cup E_2$ .

**Proposición 2.1.1** *Los siguientes ocho subconjuntos forman una partición de  $E_1 \cup E_2$ .*

- $A_1 = \left\{ e \in (E_1 \cup E_2) / e \in \bigcup_{j=1}^k (T_{1j} \cap T_{2j}) \right\}$ : “Votantes estables”.
- $A_2 = \left\{ e \in (E_1 \cup E_2) / e \in \bigcup_{1 \leq j, l \leq k, l \neq j} (T_{1j} \cap T_{2l}) \right\}$ : “Votantes transferidos”.
- $A_3 = \left\{ e \in (E_1 \cup E_2) / e \in (E_2 \cap T_2^c \cap T_1) \right\}$ : “Votantes desmovilizados”.
- $A_4 = \left\{ e \in (E_1 \cup E_2) / e \in (E_2 \cap T_1^c \cap T_2) \right\}$ : “Votantes movilizados”.
- $A_5 = \left\{ e \in (E_1 \cup E_2) / e \in (E_1 \cap E_2 \cap T_1^c \cap T_2^c) \right\}$ : “Votantes abstencionistas permanentes”.
- $A_6 = \left\{ e \in (E_1 \cup E_2) / e \in [(E_2 - E_1)] \cap T_2 \right\}$ : “Votante nuevo movilizado”.
- $A_7 = \left\{ e \in (E_1 \cup E_2) / e \in [(E_2 - E_1)] \cap T_2^c \right\}$ : “Votante nuevo no concurrente”.
- $A_8 = \left\{ e \in (E_1 \cup E_2) / e \notin E_2 \right\} = \left\{ e \in (E_1 \cup E_2) / e \in (E_2 - E_1) \right\}$ : “Ex votantes”.

DEMOSTRACIÓN: Como el tipo de proceso electoral examinado requiere a los electores emitir una única preferencia, los conjuntos  $\{T_{11}, \dots, T_{1k}\}$  son disjuntos, al igual que  $\{T_{21}, \dots, T_{2k}\}$ . Luego, para  $j \neq l$ .

$$A_1 \cap A_2 = \left( \bigcup_{j=1}^k T_{1j} \cap T_{2j} \right) \cap A_2 = \bigcup_{j=1}^k ((T_{1j} \cap T_{2j}) \cap A_2) = \bigcup_{j=1}^k \left( (T_{1j} \cap T_{2j}) \cap \left( \bigcup_{1 \leq l, j \leq k} T_{1j} \cap T_{2l} \right) \right) = \emptyset$$

Ahora, la unión de estos dos sub-cunjuntos es:

$$\begin{aligned} A_1 \cup A_2 &= \left( \bigcup_{j=1}^k T_{1j} \cap T_{2j} \right) \cup \left( \bigcup_{1 \leq l, j \leq k} T_{1j} \cap T_{2l} \right) = \bigcup_{j=1}^k \left( \bigcup_{l=1}^k (T_{1j} \cap T_{2l}) \right) = \bigcup_{j=1}^k T_{1j} \cap \left( \bigcup_{l=1}^k T_{2l} \right) = \\ &= \left( \bigcup_{j=1}^k T_{1j} \right) \cap T_2 = T_1 \cap T_2 \end{aligned}$$

A continuación, procediendo de manera secuencial se llega a que  $A_j \cap \left(\bigcup_{i=1}^{j-1} A_i\right) = \emptyset$  para  $j = 3, 4, 5$ , lo que junto a  $A_1 \cap A_2 = \emptyset$  muestra que  $\{A_1, A_2, \dots, A_5\}$  son disjuntos. La unión de los primeros 5 subconjuntos es.

$$\bigcup_{i=1}^5 A_i = (T_1 \cap T_2) \cup [(E_1 \cap E_2) \cap (T_1 \cap T_2)^c] = E_1 \cap E_2$$

Por su parte para  $A_6$  y  $A_7$ :

$$\begin{aligned} A_6 \cap A_7 &= [(E_2 - E_1) \cap T_2] \cap [(E_2 - E_1) \cap T_2^c] = (E_2 - E_1) \cap (T_2 \cap T_2^c) = \emptyset \\ A_6 \cup A_7 &= [(E_2 - E_1) \cap T_2] \cup [(E_2 - E_1) \cap T_2^c] = (E_2 - E_1) \cap (T_2 \cup T_2^c) = E_2 - E_1 \end{aligned}$$

También:

$$(E_2 - E_1) \cap (E_1 \cap E_2) = (E_2 \cap E_1^c) \cap (E_1 \cap E_2) = E_2 \cap (E_1 \cap E_1^c) = \emptyset$$

Lo que muestra que  $A_6$  y  $A_7$  son disjuntos uno a uno con  $A_1, A_2, A_3, A_4$  y  $A_5$ .

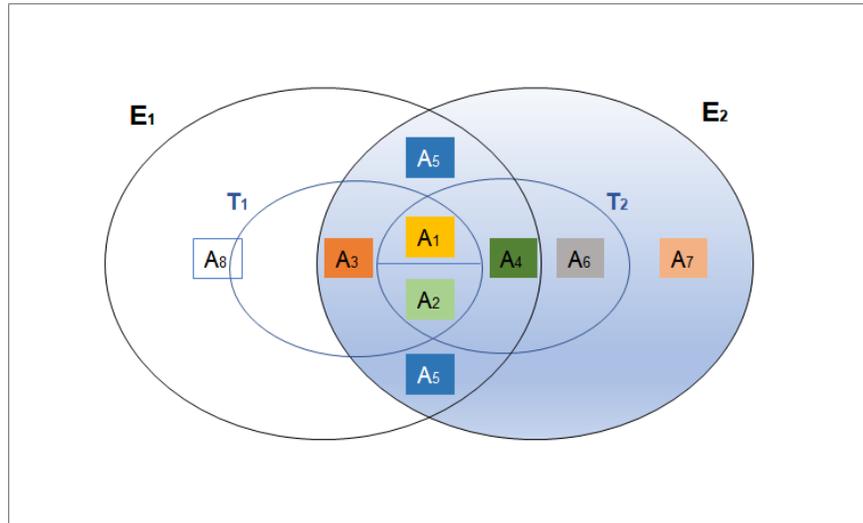
Finalmente,  $A_8$  está formado por los electores que no pertenecen a  $E_2$ . Claramente  $(E_1 - E_2) \cap (E_2 - E_1) = \emptyset = (E_1 - E_2) \cap (E_1 \cap E_2)$ , lo que muestra que  $A_8$  es disjunto de los conjuntos  $A_1$  a  $A_7$ .

La unión de los 8 subconjuntos equivale a:

$$(E_1 \cap E_2) \cup (E_2 \cap E_1^c) \cup (E_1 \cap E_2^c) = (E_1 \cap E_2) \cup [(E_1 \cup E_2) \cap (E_1 \cap E_2)^c] = E_1 \cup E_2$$

Por tanto, se tiene que estos ocho subconjuntos forman una participación de  $E_1 \cup E_2$ . El siguiente diagrama representa esta participación.

**Figura 2: Diagrama de Venn II**



#### 2.1.1.4. Espacio Paramétrico

Sean ahora  $N_t$  y  $n_t$ ,  $t = 1, 2$  la cardinalidad de  $E_t$  y  $T_{(t)}$  respectivamente. En particular, sea  $N_2 = \#\{E_2\}$ . Luego, como  $p_i, i = 1, 2, \dots, 7$  se representa a la proporción de votantes en  $E_1 \cup E_2 - \{A_8\} = E_2$  que componen los subconjuntos  $A_i, i = 1, 2, \dots, 7$ , esto es:

$$p_i = \frac{\#\{A_i\}}{N_2}, \text{ donde } \#\left\{\bigcup_{i=1}^7 A_i\right\} = N_2.$$

El problema consiste en estimar el vector de parámetros  $\theta = (p_1, p_2, \dots, p_7) \in \Theta$ , donde:

$$\Theta = \left\{ (p_1, p_2, \dots, p_7) \mid 0 \leq p_i \leq 1, \sum_{i=1}^7 p_i = 1 \right\} \quad (2.1.1)$$

Adicionalmente, se define como  $p_8 = \frac{\#\{A_8\}}{N_2}$ , al número de ex-votantes dividido por la cardinalidad del electorado en el segundo proceso electoral.

#### 2.1.2. Parámetros de la Volatilidad Electoral

En el marco de este proyecto las cantidades  $\{p_1, \dots, p_7, p_8\}$  se denominan “parámetros de la volatilidad electoral” debido a que la volatilidad constatada entre eventos de votación sucesivos es resultado de esta composición desconocida. A continuación se justifica el uso de esta denominación.

##### 2.1.2.1. Vector de Resultados electorales

Con la realización de un proceso electoral se obtiene un vector de conteos,  $\mathbf{d}_t$  de longitud  $k + 1$ , con entradas  $d_{tj} = \#\{T_{(t)j}\}$  para  $j = 1, \dots, k$ , y  $d_{t,k+1} = \#\{E_t \cap T_{(t)}^c\} = N_t - n_t$ .

Para la conversión del número total de sufragios en cargos de representación democrática, el abstencionismo<sup>11</sup> es excluido del conteo. La traducción de votos en cargos de representación utiliza “votaciones relativas”, valores que, a su vez, componen el “vector de resultados electorales” para el proceso eleccionario.

La votación relativa  $v_{(t)j}$ ,  $t = 1, 2, j = 1, \dots, k$ , obtenida por el  $j$ -ésimo agente de representación política en un proceso electoral, corresponde al cociente  $\frac{d_{(t)j}}{n_t}$ , donde  $n_t = \#\{T_{(t)}\} = \sum_{j=1}^k d_{tj}$ .

Por su parte, el vector de resultados electorales  $\mathbf{v}_t$  para el proceso es un vector con entradas  $\frac{d_{(t)j}}{n_t}$ ,  $j = 1, \dots, k$ .

##### 2.1.2.2. Forma matricial del Índice de Pedersen

Recordando, la ecuación del Índice de Pedersen (1979) dada previamente en (1.1.1) es la siguiente.

---

<sup>11</sup>Que corresponde a la suma de votos en “blanco” (es decir, sufragios que no marcan ninguna preferencia), de votos “nulos” (aquellos que marcan más opciones que las solicitadas, por ejemplo) y electores que pertenecen al padrón, pero no concurren a sufragar.

$$V = \frac{1}{2} \sum_{j=1}^k |v_{(t+1),j} - v_{(t),j}|$$

Esta formulación puede reescribirse en notación vectorial. Denotando como  $\mathbf{v}_t$  al vector de resultados electorales en el tiempo  $t$  y  $\mathbf{v}_{t+1}$  en  $t + 1$ , además  $\mathbf{1}$  el vector constante de unos de longitud  $k$ , entonces el Índice de Pedersen puede reescribirse como:

$$V = \frac{1}{2} \sum_{j=1}^k |v_{(t+1),j} - v_{(t),j}| = \frac{1}{2} (|\mathbf{v}_{t+1} - \mathbf{v}_t|)' \mathbf{1} = \frac{1}{2} \|\mathbf{v}_{t+1} - \mathbf{v}_t\|_1 \quad (2.1.2)$$

Es decir, la volatilidad corresponde a la mitad de la norma 1 del vector diferencia de resultados electorales.

Ahora bien, el vector de resultados electorales en  $t + 1$  puede obtenerse como transformación lineal del vector de resultados electorales precedente. Vale decir, existe una matriz  $\mathbf{M}$  cuadrada, de orden  $k$  tal que  $\mathbf{v}_{t+1} = \mathbf{M}\mathbf{v}_t$ .

Estructuraremos a continuación esta matriz  $\mathbf{M}$  como resultado de la composición del electorado. Esta estructuración se desarrolla en dos pasos.

**2.1.2.2.1. Paso 1** Definimos en primer lugar una matriz  $\mathbf{X} = (x)_{rc}$ ,  $r = 1, \dots, k + 2$ ,  $c = 1, \dots, k + 2$  como el producto de otras dos matrices cuadradas  $\Phi$  y  $\mathbf{B}$  de orden  $k + 2$  y un coeficiente escalar  $\frac{n_1}{N_2}$ ,  $\mathbf{X} = \frac{n_1}{N_2} \Phi \mathbf{B}$ .

Aquí la matriz representada como  $\Phi = (\phi)_{rc}$  se compone por  $k + 2$  vectores columna tales que  $\phi_{rc} \geq 0$  y  $\sum_{r=1}^{k+2} \phi_{rc} = 1$ ,  $\forall c = 1, \dots, k + 2$ . Las entradas en cada vector columna contienen la proporción que, correlativamente, los  $k$  agentes de representación política (columnas 1 a  $k$ ), el abstencionismo en el primer proceso (columna  $k + 1$ ), y los nuevos votantes (columna  $k + 2$ ), distribuyen en en los  $k + 2$  estados posibles para el segundo, que corresponden a:

- Votar por alguno de los  $k$  agentes de representación:  $e \in T_{2j}$ ,  $j = 1, \dots, k$ .
- Abstencionismo:  $e \in E_2 \cap T_2^c$ .
- Ausencia del padrón,  $e \notin E_2$ .

Estos  $k + 2$  estados figuran ordenados correlativamente en las filas de la matriz. Así, por ejemplo, la entrada  $\phi_{rc}$  para  $c = 1, \dots, k$ :

$$\{(\phi)_{rc}\}_{c=1}^k = \begin{cases} \frac{\#\{T_{1c} \cap T_{2r}\}}{\#\{T_{1c}\}} & \text{si } r = 1, \dots, k \\ \frac{\#\{T_{1c} \cap [E_2 \cap T_2^c]\}}{\#\{T_{1c}\}} & \text{si } r = k + 1 \\ \frac{\#\{T_{1c} \cap E_2^c\}}{\#\{T_{1c}\}} & \text{si } r = k + 2 \end{cases} \quad (2.1.3)$$

Las entradas en las columnas  $k + 1$  y  $k + 2$  de esta última matriz se obtienen al reemplazar en (2.1.3)  $T_{1c}$  por  $E_1 \cap T_1^c$  y  $E_1^c$ , respectivamente.

Por su parte, la matriz  $\mathbf{B}$  es diagonal, con entradas:

$$\mathbf{B} = \text{diag} \left( \frac{1}{n_1} \left[ \{d_{1j}\}_{j=1}^{k+1}, nv_2 \right] \right)$$

Donde  $nv_2 = \#(A_6 \cup A_7)$ .

A modo de ejemplo, para  $k = 2$ , esta última matriz sería:

$$\mathbf{B} = \frac{1}{n_1} \begin{bmatrix} d_{11} & 0 & 0 & 0 \\ 0 & d_{12} & 0 & 0 \\ 0 & 0 & N_1 - n_1 & 0 \\ 0 & 0 & 0 & nv_2 \end{bmatrix}$$

Luego, en el producto matricial  $\frac{n_1}{N_2} \Phi \mathbf{B}$ , el vector columna  $j$  dentro de las primeras  $k$  columnas es un vector con entradas:

$$\begin{aligned} & \frac{1}{N_2} [\phi_{1j}d_{1j} \quad \phi_{2j}d_{1j} \quad \dots \quad \phi_{k+1,j}d_{1j} \quad \phi_{k+2,j}d_{1j}]^t = \\ & = \frac{d_{1j}}{N_2} [\phi_{1j} \quad \phi_{2j} \quad \dots \quad \phi_{k+1,j} \quad \phi_{k+2,j}]^t, \quad j = 1, \dots, k \end{aligned} \quad (2.1.4)$$

Por su lado, los elementos en los vectores columna  $k + 1$  y  $k + 2$  de esta matriz son, respectivamente:

$$\frac{(N_1 - n_1)}{N_2} [\phi_{1,k+1} \quad \phi_{2,k+1} \quad \dots \quad \phi_{k+1,k+1} \quad \phi_{k+2,k+1}]^t \quad (2.1.5)$$

$$\frac{nv_2}{N_2} [\phi_{1,k+2} \quad \phi_{2,k+2} \quad \dots \quad \phi_{k+1,k+2} \quad \phi_{k+2,k+2}]^t \quad (2.1.6)$$

Si observamos, la traza de la submatriz obtenida al remover las últimas 2 filas y columnas de esta matriz corresponde a:

$$\sum_{r=1}^k x_{rr} = \frac{1}{N_2} \sum_{r=1}^k \phi_{rr} d_{1r} = p_1$$

Y en general:

$$\begin{aligned}
p_1 &= \sum_{r=1}^k x_{rr} \\
p_2 &= \left( \sum_{r=1}^k \sum_{c=1}^k x_{rc} \right) - p_1 \\
p_3 &= \sum_{c=1}^k x_{k+1,c} \\
p_4 &= \sum_{r=1}^k x_{r,k+1} \\
p_5 &= x_{k+1,k+1} \\
p_6 &= \sum_{r=1}^k x_{r,k+2} \\
p_7 &= x_{k+1,k+2} \\
p_8 &= \sum_{c=1}^{k+2} x_{k+2,c}
\end{aligned} \tag{2.1.7}$$

Entonces,  $\mathbf{X}$  resulta ser una matriz cuadrada de orden  $k + 2$ , de cuya combinación lineal de sus entradas se obtienen las magnitudes  $\{p_1, \dots, p_7, p_8\}$ .

**2.1.2.2.2. Paso 2** Tomando  $v_{(t)k}$ , la entrada  $k$  de  $\mathbf{v}_t$ , como categoría de referencia, se define una matriz  $\mathbf{C}$  con  $k + 2$  filas y  $k$  columnas, tal que  $\mathbf{C}^t = [\mathbf{C}_1 \ \mathbf{C}_2] = [\mathbf{I}_k \ \mathbf{C}_2]$ , donde  $\mathbf{I}_k$  es la matriz identidad de orden  $k$  y  $\mathbf{C}_2$  contiene valores nulos en todas sus entradas, a excepción de  $\{c_{2,k+1,k}\} = \left(\frac{N_t - n_t}{v_{(t)k}}\right) / n_t$  y también  $\{c_{2,k+2,k}\} = \left(\frac{nv_{t+1}}{v_{(t)k}}\right) / n_t$ , donde por  $nv_{t+1}$  se representa al número de nuevos votantes en el segundo proceso analizado. Ejemplificando nuevamente con  $k = 2$ :

$$\mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & \frac{N_t - n_t}{n_t v_{(t)2}} \\ 0 & \frac{nv_{t+1}}{n_t v_{(t)2}} \end{bmatrix}$$

Por su parte, sea  $\mathbf{A}$  otra matriz de  $k \times k + 2$  con  $\mathbf{A} = \frac{N_{t+1}}{n_{t+1}} [\mathbf{I}_k \ \mathbf{0}]$ , siendo ahora  $\mathbf{0}$  la submatriz cero de dimensiones  $k \times 2$ . Para  $k = 2$ :

$$\mathbf{A} = \frac{N_{t+1}}{n_{t+1}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Finalmente, con esta notación, sea  $\mathbf{M} = \mathbf{A}\mathbf{X}\mathbf{B}^{-1}\mathbf{C}$ , siendo  $\mathbf{B}$  no singular. Entonces, se tiene la siguiente proposición.

**Proposición 2.1.2** Sean dos procesos de votación consecutivos con vectores de resultados electorales  $\mathbf{v}_1$  y  $\mathbf{v}_2$ , tales que solicitan al electorado emitir una única preferencia y los mismos  $k$  agentes de representación política concentran la totalidad de la oferta electoral;  $V$  la volatilidad entre ambos procesos; y  $\mathbf{M} = \mathbf{A}(\mathbf{X})\mathbf{B}^{-1}\mathbf{C}$  una matriz cuadrada de orden  $k$ , definida según el desarrollo anterior. Entonces  $V = \frac{1}{2}\|(\mathbf{M}-\mathbf{I})\mathbf{v}_1\|_1$ .

DEMOSTRACIÓN:

$$T_2 = \bigcup_{j=1}^k T_{2j} = (A_1 \cup A_2 \cup A_4 \cup A_6) \Rightarrow T_{2j} \subseteq (A_1 \cup A_2 \cup A_4 \cup A_6), \quad j = 1, \dots, k \quad (2.1.8)$$

Entonces, de (2.1.8) y (2.1.3):

$$d_{2j} = \# \{T_{2j}\} = \phi_{jj}d_{1j} + \sum_{1 \leq c \leq k, c \neq j} \phi_{jc}d_{1c} + \phi_{j,k+1}(N_1 - n_1) + \phi_{j,k+2}nv_2 \quad (2.1.9)$$

Donde  $d_{1j}$ ,  $j = 1, \dots, k+1$  es la entrada  $j$  del vector  $\mathbf{d}_1$ , y  $\phi_{jc}$  el  $(j, c)$ -ésimo coeficiente dado por (2.1.3) y su extensión para  $k+1$  y  $k+2$ . En consecuencia, la  $j$ -ésima entrada del vector  $\mathbf{v}_2$  es:

$$v_{2j} = \frac{d_{2j}}{n_2} = \frac{1}{n_2} \left[ \phi_{jj}d_{1j} + \sum_{1 \leq c \leq k, c \neq j} \phi_{jc}d_{1c} + \phi_{j,k+1}(N_1 - n_1) + \phi_{j,k+2}nv_2 \right]$$

Ahora en el producto matricial  $\mathbf{M}\mathbf{v}_1$  se observa que:

$$\mathbf{M}\mathbf{v}_1 = \mathbf{A}\mathbf{X}\mathbf{B}^{-1}\mathbf{C}\mathbf{v}_1 = \mathbf{A}\mathbf{X}\mathbf{1} = \frac{n_1}{N_2} [\mathbf{A}\Phi\mathbf{B}] \mathbf{1}$$

Dado que  $\mathbf{A} = \frac{N_2}{n_2} [\mathbf{I}_k \quad \mathbf{0}]$ , simbolizando por  $\mathbf{a}'_j$  al vector fila  $j$  de  $\mathbf{A}$ , se tiene entonces que:

$$\begin{aligned} \frac{n_1}{N_2} \mathbf{a}'_j [(\Phi\mathbf{B}) \mathbf{1}] &= \frac{1}{n_2} \mathbf{a}'_j [\Phi(n_1\mathbf{B}) \mathbf{1}] = \\ &= \frac{1}{n_2} \left[ \phi_{jj}d_{1j} + \sum_{1 \leq c \leq k, c \neq j} \phi_{jc}d_{1c} + \phi_{j,k+1}(N_1 - n_1) + \phi_{j,k+2}nv_2 \right] = v_{2j} \end{aligned} \quad (2.1.10)$$

Como (2.1.10) se cumple  $\forall j = 1, \dots, k$ , se concluye que  $\mathbf{v}_2 = \mathbf{M}\mathbf{v}_1$ , y por consiguiente de (2.1.2.2):

$$V = \frac{1}{2}\|\mathbf{v}_2 - \mathbf{v}_1\|_1 = \frac{1}{2}\|\mathbf{M}\mathbf{v}_1 - \mathbf{v}_1\|_1 = \frac{1}{2}\|(\mathbf{M}-\mathbf{I})\mathbf{v}_1\|_1$$

Que en notación genérica resulta en:

$$V = \frac{1}{2}\|(\mathbf{M}-\mathbf{I})\mathbf{v}_t\|_1 \quad (2.1.11)$$

Entonces, la volatilidad electoral observada tras dos procesos de votación consecutivos, así como el vector de parámetros con la composición del electorado en siete categorías de votantes, son ambas funciones –la primera escalar, y la segunda vectorial– de los mismos valores al interior de la matriz aquí representada como  $\mathbf{X}$ . En consecuencia, la volatilidad electoral observada entre procesos electorales consecutivos es un resultado de la composición del electorado.

### 2.1.2.3. Inconveniencia de la Estimación por la vía muestral

Una alternativa inmediata para la estimación de las cantidades  $\{p_1, \dots, p_7\}$  viene dada por la vía muestral. En este caso se requeriría una muestra *i.i.d* de  $E_2$ . Sin embargo por los problemas de recordación, se tendrá que la muestra no procede del total de dicha población, sino de un subconjunto de aquella (encuestados que recuerdan su comportamiento de voto). A ello se suman problemas de fiabilidad de respuesta por “ocultamiento” del comportamiento de voto, así como la escasez y/o intermitencia de encuestas electorales o muestras de panel sobre este tema alrededor del mundo.

### 2.1.3. Síntesis

El espacio paramétrico del problema corresponde a:

$$\Theta = \left\{ (p_1, p_2, \dots, p_7) / 0 \leq p_i \leq 1, \sum_{i=1}^7 p_i = 1 \right\}$$

Donde  $p_i, i = 1, 2, \dots, 7$  es la proporción de votantes que componen el padrón tras dos elecciones sucesivas que pertenecen a cada una de las siete categorías enunciadas al inicio de la sección 1.2, de modo tal que  $p_1$  es la proporción de votantes estables;  $p_2$  proporción de votantes transferidos, ..., y  $p_7$  la fracción de electores nuevos abstencionistas.

En el marco de este proyecto dichas magnitudes son denominadas “parámetros de la volatilidad electoral” y se obtienen como combinación lineal de una o más entradas de una matriz aquí representada como  $\mathbf{X}$ , la cual es cuadrada de orden  $k + 2$ , donde  $k$  es el número de opciones políticas en competencia.

Esta matriz figura en la transformación lineal del vector de resultados electorales en el tiempo  $t$ , en su equivalente para  $t + 1$ . Definiendo tres matrices adicionales  $\mathbf{A}$ ,  $\mathbf{B}$  y  $\mathbf{C}$ , se demostró que, al hacer  $\mathbf{M} = \mathbf{A}(\mathbf{X})\mathbf{B}^{-1}\mathbf{C}$ , entonces  $\mathbf{v}_2 = \mathbf{M}\mathbf{v}_1$ , lo que conduce luego a una formulación matricial del índice de volatilidad electoral  $V = \frac{1}{2}\|(\mathbf{M}-\mathbf{I})\mathbf{v}_t\|_1$ . Ello muestra que la volatilidad electoral observada ocurre como un resultado de la composición del electorado en las siete categorías de votantes antes señaladas.

Las magnitudes al interior de la matriz  $\mathbf{X}$  pueden ser estimadas mediante el uso de encuestas o muestras de panel. Sin embargo, estas alternativas presentan problemas de fiabilidad, y a nivel internacional resultan ser escasas, lo que dificulta sostener un abordaje comparativo-longitudinal del problema.

## 2.2. Métodos Estadísticos para la estimación de entradas de Matrices de transición electoral

En esta sección se busca dar cumplimiento al cuarto objetivo específico del proyecto: Describir los métodos de Inferencia Ecológica para estimar matrices de transición electoral y su relación con Cadenas de Markov. Tanto las definiciones como los resultados que se presentan en el mismo fueron extraídos de la literatura previa. El único componente inédito es la demostración de la proposición 2.2.1, la cual no figuraba en el material examinado. También cabe señalar que los algoritmos bayesianos jerárquicos incluidos en la descripción son descritos en este apartado de manera resumida. Una descripción más detallada de los mismos se incluye en el Anexo 5.3.2 del presente documento.

### 2.2.1. Matriz de Transición electoral

Siendo  $k$  el número de agentes de representación política que concentran la totalidad de la oferta electoral en un proceso electoral en un momento  $t \in \mathbb{N}$ , y  $l$  el número de opciones para otro proceso en  $t + \Delta t$ ,  $\Delta t \geq 0$ , con vectores de resultados electorales  $\mathbf{v}_t$  y  $\mathbf{v}_{t+\Delta t}$ , respectivamente, se ha representado a una matriz de transición electoral  $\mathbf{R} = (r)_{ij}$  como una matriz con  $k$  filas y  $l$  columnas, que satisface las siguientes tres propiedades:

1.  $r_{ij} \geq 0, \forall ij, i = 1, \dots, k, j = 1, \dots, l$ .
2.  $\sum_{j=1}^l r_{ij} = 1, \forall i = 1, \dots, k$ .
3.  $\mathbf{v}_{t+\Delta t} = \mathbf{R}' \mathbf{v}_t$ .

Vale decir  $\mathbf{R}$  es una matriz con entradas no negativas, donde suma de los elementos de cada una de sus filas es igual a 1, y que al realizar el producto con el vector de resultados electorales en  $t$ , se obtiene el vector en  $t + \Delta t$ .

Las filas de la matriz de transición electoral representan las proporciones que cada agente de representación política retiene o pierde entre ambos eventos de votación.

### 2.2.2. Inferencia Ecológica

El problema de la Inferencia Ecológica involucra a una población finita  $U$  compuesta por  $N$  elementos, y  $\{u_1, \dots, u_M\}$  una participación de  $U$ , cada uno de estos  $M$  subconjuntos con cardinalidad  $N_m$ ,  $m = 1, \dots, M$ .

También, en el problema figuran dos variables categóricas  $W$  y  $Z$ , con  $R$  y  $C$  categorías respectivamente, definidas sobre  $U$  con distribución conocida  $g_1(W)$  y  $g_2(Z)$  y distribuciones  $g_{1m}(W)$  y  $g_{2m}(Z)$ ,  $m = 1, \dots, M$ , también conocidas.

El conocimiento de la distribución de estas variables proviene de medios censales. Por la misma vía se conoce las distribuciones de las mismas variables en las unidades que forman la partición. Sin embargo,

al proceder de fuentes censales diferentes, la distribución conjunta es desconocida.

Luego, sea  $g(W, Z)$  la distribución conjunta definida sobre  $U$ , desconocida, y  $g_m(W, Z)$ ,  $m = 1, \dots, M$  las conjuntas sobre  $\{u_1, \dots, u_M\}$ , también desconocidas. Los valores de la distribución conjunta  $g(W, Z)$  se pueden representar al interior de una matriz  $\mathbf{H}$  de dimensiones  $R \times C$ ,  $R \geq 2, C \geq 2$ . Una posible representación de esta matriz es la siguiente:

$$\mathbf{H} = \begin{bmatrix} ? & \dots & \dots & ? \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ ? & \dots & \dots & ? \\ z_1 & \dots & \dots & z_C & 1 \end{bmatrix} \begin{matrix} w_1 \\ \dots \\ \dots \\ w_R \\ 1 \end{matrix}$$

La mayor parte de las aplicaciones realizadas hasta la actualidad han ubicado en las entradas de esta matriz  $\mathbf{H}$  los valores de la distribución condicional de una de las dos variables dados los valores de la otra. Genéricamente, se ha buscado estimar la distribución de  $W$  dado  $Z = z$ . Esto puede entenderse como un caso particular del problema, dado que el desconocimiento de ambas distribuciones condicionadas procede del desconocimiento de la distribución conjunta.

A continuación, sea  $\mathbf{h}$  un vector de longitud  $n = RC$  cuyas entradas corresponden a los valores de la distribución conjunta  $g(W, Z)$ , tal que  $\mathbf{h} = [h_{11} \ h_{12} \ \dots \ h_{21} \ \dots \ h_{r1} \ \dots \ h_{RC}]^t$  y  $\mathbf{f}(\mathbf{h}) : \mathbb{R}^n \rightarrow \mathbb{R}^k$ ,  $k \geq 1$  una función de la distribución conjunta. Entonces, el objetivo de la Inferencia ecológica se puede plantear en los siguientes términos:

*Estimar  $\mathbf{f}(\mathbf{h})$  utilizando la información contenida en  $g_{1m}(W)$  y  $g_{2m}(Z)$ ,  $m = 1, \dots, M$ .*

Esta definición incluye el caso en que el objetivo de la estimación son las mismas entradas desconocidas de la matriz. En tal caso  $\mathbf{f}(\mathbf{h}) = \mathbf{h}$ .

Uno de los supuestos básicos de estos métodos es que  $\{u_1, \dots, u_M\}$  constituye una partición de la población  $U$ . Esto ocurre siempre que  $\bigcup_{m=1}^M u_m = U$  y que la definición de la población no involucre un componente temporal. De incluirlo, se presenta el problema que los elementos de la población pueden estar en dos subunidades diferentes en un momento u otro. Esto puede generar problemas de identificación de las distribuciones  $g_{1m}(W)$  y  $g_{2m}(Z)$ . Como señalan Klima et al. (2016) en aplicaciones a problemas de transición electoral es importante chequear la magnitud de tendencias tales como cambios de domicilio o de registro electoral. En caso de tratarse de cifras minoritarias, se suele proceder bajo el supuesto que las unidades de desagregación intermedia sí conforman una partición de  $U$ .

### 2.2.2.1. Modelos Stayers-Movers

Los modelos “Stayers-movers” buscan identificar la matriz  $\mathbf{R} = (r)_{ij}$ ,  $i = 1, \dots, k$ ;  $j = 1, \dots, l$  de transición electoral que minimiza alguno de los siguientes cuatro criterios de decisión:

$$\sum_{m=1}^M \sum_{j=1}^l \left( n_m v_{(t+\Delta t)jm} - \sum_{i=1}^k (n_m v_{(t)jm} r_{ij}) \right)^2 = \sum_{m=1}^M n_m^2 \sum_{j=1}^l \left( v_{(t+\Delta t)jm} - \sum_{i=1}^k (v_{(t)jm} r_{ij}) \right)^2 \quad (2.2.1)$$

$$\sum_{m=1}^M \sum_{j=1}^l \left( v_{(t+\Delta t)jm} - \sum_{i=1}^k (v_{(t)jm} r_{ij}) \right)^2 \quad (2.2.2)$$

$$\max_{\substack{m=1, \dots, M; \\ j=1, \dots, l}} n_m \left| v_{(t+\Delta t)jm} - \sum_{i=1}^k (v_{(t)jm} r_{ij}) \right| \quad (2.2.3)$$

$$\max_{\substack{m=1, \dots, M; \\ j=1, \dots, l}} \left| v_{(t+\Delta t)jm} - \sum_{i=1}^k (v_{(t)jm} r_{ij}) \right| \quad (2.2.4)$$

Sujeto a las restricciones dadas por la definición de matriz de transición electoral:

- $r_{ij} \geq 0, \forall i, j, i = 1, \dots, k, j = 1, \dots, l.$
- $\sum_{j=1}^l r_{ij} = 1, \forall i = 1, \dots, k.$
- $\mathbf{v}_{t+\Delta t} = \mathbf{R}' \mathbf{v}_t.$

Es decir, se aborda el tema como un asunto de optimización con restricciones, donde  $\mathbf{v}_{(\cdot)}$  es el vector de resultados electorales en el tiempo  $t, t + \Delta t$ . Los criterios (2.2.1) y (2.2.2) efectúan la minimización de discrepancias en norma euclídea, el primero tomando como unidad de medida el número de preferencias electorales, y el segundo a las proporciones que conforman los vectores de resultados electorales. En tanto, los criterios (2.2.3) y (2.2.4) recurren a la norma infinito.

Estas alternativas pertenecen a la clase denominada “estimadores de norma  $L_p$  en regresión lineal”. Una descripción de la existencia y unicidad de los mismos se encuentra en el trabajo de Agüero (1998: 94-95).

Como una de las principales limitaciones asociadas al uso de este tipo métodos en el contexto de la estimación con datos agregados figura la falta de una estimación por intervalo, al no tener un fundamento explícito en un componente distribucional.

### 2.2.2.2. Modelos con supuestos distribucionales

En la década de 1950 surgieron dos aportaciones para abordar el problema de la “falacia ecológica” previamente descrito por Robinson: el método de las cotas de Duncan & Davis (1953) y la Regresión ecológica de Goodman (1953). Las propuestas de carácter más reciente recogen elementos de ambas contribuciones, agregando un componente distribucional.

**2.2.2.2.1. Acotación de Duncan & Davis.** Duncan & Davis (1953) pusieron de relieve que las sumas conocidas por filas y columnas de la matriz aquí representada como  $\mathbf{H}$  imponen restricciones a sus entradas, las que se traducen en cotas superiores e inferiores para estas cantidades. Una primera demostración formal de este resultado para el caso de la distribución condicional de  $W|Z = z$  aparece en el trabajo de Horowitz & Manski (1995).

Por su parte, una traducción formal de este razonamiento al caso de la distribución conjunta de ambas variables categóricas puede plantearse en los siguientes términos.

**Proposición 2.2.1** Sea  $\mathbf{H} = (h)_{rc}, r = 1, \dots, R; c = 1, \dots, C$  una matriz cuyas entradas contienen la distribución conjunta de dos variables categóricas, con  $R$  y  $C$  categorías, respectivamente. Entonces  $\forall i, l$  tales que  $1 \leq i, l \leq R, i \neq l$  y también  $\forall j, k$  tales que  $1 \leq j, k \leq C, j \neq k$ :

$$\max \left( 0, \sum_{c=1}^C h_{ic} - \sum_k \sum_{r=1}^R h_{ik, k \neq j} \right) \leq h_{ij} \leq \min \left( \sum_{r=1}^R h_{rj}, \sum_{c=1}^C h_{ic} \right)$$

DEMOSTRACIÓN: Se parte del caso 2x2, es decir, de una matriz  $\mathbf{H}$  con  $R = 2$  filas y  $C = 2$  columnas. Para simplificar en este paso la notación, sean  $a$  y  $c$  las sumas de los elementos ubicados en las filas 1 y 2, respectivamente, tales que  $a + c = 1$ ; y sean  $b$  y  $d$  las sumas de las columnas 1 y 2, respectivamente, tales que  $b + d = 1$ .

A continuación tomamos alguna entrada de la matriz, por ejemplo  $h_{11}$ . Luego notamos que:

$$\left. \begin{array}{l} h_{11} + h_{12} = a \Rightarrow h_{11} \leq a \\ h_{11} + h_{21} = b \Rightarrow h_{11} \leq b \end{array} \right\} \Rightarrow h_{11} \leq \min(a, b) \quad (2.2.5)$$

Empleando la misma lógica se llega a que  $h_{12} \leq \min(a, d)$ ,  $h_{21} \leq \min(b, c)$  y también  $h_{22} \leq \min(c, d)$ . Vinculado con esto último, dado que  $h_{11} + h_{12} = a$  y  $h_{11} + h_{21} = b$  se tiene que:

$$\left. \begin{array}{l} h_{11} = a - h_{12} \Rightarrow h_{11} \geq a - \min(a, d) \Rightarrow h_{11} \geq \max(0, a - d) \\ h_{11} = b - h_{21} \Rightarrow h_{11} \geq b - \min(b, c) \Rightarrow h_{11} \geq \max(0, b - c) \end{array} \right\} \Rightarrow h_{11} \geq \max(0, a - d, b - c) \quad (2.2.6)$$

Así, combinando (2.2.5) y (2.2.6) se concluye respecto a  $h_{11}$ :

$$\max(0, a - d, b - c) \leq h_{11} \leq \min(a, b)$$

Pero,  $a - d = b - c$  porque  $a + c = 1 = b + d$ , de modo que, una expresión más sintética de la acotación es:

$$\max(0, a - d) \leq h_{11} \leq \min(a, b) \quad (2.2.7)$$

El mismo tipo de razonamiento lleva a concluir:

- $\max(0, a - b) \leq h_{12} \leq \min(a, d)$
- $\max(0, c - d) \leq h_{21} \leq \min(b, c)$
- $\max(0, c - b) \leq h_{22} \leq \min(c, d)$

Ahora, para el caso general  $R \times C$ , notemos que toda matriz de dimensiones  $R \geq 2, C > 2$  o bien  $R > 2, C \geq 2$  puede representarse como una matriz particionada de 2x2, sumando los valores de las entradas que son objeto de agrupación.

Tomemos a modo genérico la entrada  $h_{ij}$  de la matriz particionada, donde  $1 \leq i, l \leq R, i \neq l$  y  $1 \leq j, k \leq C, j \neq k$ . La suma de los elementos ubicados en la fila  $i$  es  $\sum_{c=1}^C h_{ic}$  y la de la columna respectiva  $\sum_{r=1}^R h_{rj}$ . En tanto, la suma de los elementos ubicados en filas diferentes  $\sum_l \sum_{c=1}^C h_{lj, l \neq i}$  y en las restantes

columnas  $\sum_k \sum_{r=1}^R h_{ik, k \neq j}$ . Luego, empleando el mismo razonamiento que (2.2.5), (2.2.6) y (2.2.7), esta vez para un valor genérico  $h_{ij}$ .

$$h_{ij} \leq \min \left( \sum_{c=1}^C h_{ic}, \sum_{r=1}^R h_{rj} \right) \quad (2.2.8)$$

$$h_{ij} \geq \max \left( 0, \sum_{c=1}^C h_{ic} - \sum_k \sum_{r=1}^R h_{rk, k \neq j} \right) \quad (2.2.9)$$

$$h_{ij} \geq \max \left( 0, \sum_{r=1}^R h_{rj} - \sum_l \sum_{c=1}^C h_{lc, l \neq i} \right) \quad (2.2.10)$$

Finalmente, combinando (2.2.8), (2.2.9) y (2.2.10) se concluye que  $\forall i, j : i = 1, \dots, R; j = 1, \dots, C$ :

$$\max \left( 0, \sum_{c=1}^C h_{ic} - \sum_k \sum_{r=1}^R h_{rk, k \neq j}, \sum_{r=1}^R h_{rj} - \sum_l \sum_{c=1}^C h_{lc, l \neq i} \right) \leq h_{ij} \leq \min \left( \sum_{r=1}^R h_{rj}, \sum_{c=1}^C h_{ic} \right)$$

Que, como en (2.2.7), una versión simplificada resulta en:

$$\max \left( 0, \sum_{c=1}^C h_{ic} - \sum_k \sum_{r=1}^R h_{rk, k \neq j} \right) \leq h_{ij} \leq \min \left( \sum_{r=1}^R h_{rj}, \sum_{c=1}^C h_{ic} \right)$$

**2.2.2.2.2. Regresión Ecológica de Goodman.** En la actualidad, este método es empleando a modo referencial para evaluar los resultados ofrecidos por nuevas aproximaciones.

Goodman (1953) propuso una estrategia para estimar la distribución condicional de una de las dos variables categóricas, dados los valores de la otra. Esta propuesta se basa en el supuesto de las distribuciones condicionales de  $W$  dado  $Z = z$  son constantes en todas las unidades  $\{u_1, \dots, u_M\}$  (“Constancy assumption”). En el caso electoral esto vendría a significar que la misma matriz de transición opera en la transformación de los resultados de votación en todas las subunidades territoriales.

La propuesta de Goodman parte enfocando la distribución condicional de una variable  $W$  con dos categorías, dados los valores de otra variable  $Z$ , también de dos categorías; por ejemplo: Alfabetización (Sí/No) según Género; Acceso a un bien de consumo (Sí/No) según Residencia en la capital provincial (o regional); etc.

La distribución en una población finita de una variable con dos categorías puede representarse como una distribución de Bernoulli, reexpresando las categorías como presencia/ausencia de un atributo específico<sup>12</sup>. Al reformular el problema en dichos términos, la distribución de ambas variables en la población tendrán asociada una media.

- $\mu_W = \frac{1}{N} \sum_{i=1}^N W_i$

---

<sup>12</sup>Por ejemplo, alfabetizado=1, no alfabetizado=0; reside en la capital=1, no reside en la capital=0, etc.

- $\mu_Z = \frac{1}{N} \sum_{i=1}^N Z_i$

Donde  $N$  es la cardinalidad de la población  $U$ . Nótese que  $W$  y  $Z$  no son variables aleatorias, sino las variables en la población<sup>13</sup>. Se recurre al mismo tipo de notación para las medias condicionales:  $\mu_{W|Z=1}$  y  $\mu_{W|Z=0}$ . Al hacer  $y = \mu_W$ ,  $x = \mu_Z$ ,  $p = \mu_{W|Z=1}$  y  $q = \mu_{W|Z=0}$  se tiene, por el teorema de la probabilidad total:

$$y = px + q(1 - x) = q + (p - q)x$$

Del mismo modo, en las subunidades de la población se cumple:

$$\begin{aligned} y_1 &= q_1 + (p_1 - q_1)x_1 \\ y_2 &= q_2 + (p_2 - q_2)x_2 \\ &\vdots \\ y_M &= q_M + (p_M - q_M)x_M \end{aligned}$$

Denotamos ahora como  $\mathbf{y}$  al vector de medias de  $W$  de longitud  $M$  en las unidades de la partición  $\{u_1, \dots, u_M\}$ , y como  $\mathbf{x}$  su equivalente para el caso de  $Z$ . Luego, si  $q_m = q \forall m = 1, \dots, M$  y  $p_m = p \forall m = 1, \dots, M$  (“Constancy assumption”), entonces:

$$\begin{aligned} y_1 &= q + (p - q)x_1 \\ y_2 &= q + (p - q)x_2 \\ &\vdots \\ y_M &= q + (p - q)x_M \end{aligned}$$

Al hacer  $\mathbf{C} = [\mathbf{1} \quad \mathbf{x}]$ , donde  $\mathbf{1}$  es el vector constante de unos de longitud  $M$ :

$$\mathbf{y} = \mathbf{C} \begin{bmatrix} q \\ p - q \end{bmatrix}$$

Entonces, la solución de mínimos cuadrados  $\beta' = [a \ b] = (\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\mathbf{y}$  es tal que:

$$\beta = \begin{bmatrix} a \\ b \end{bmatrix} = (\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\mathbf{y} = (\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\mathbf{C} \begin{bmatrix} q \\ p - q \end{bmatrix} = \begin{bmatrix} q \\ p - q \end{bmatrix} \quad (2.2.11)$$

Así,  $\hat{q} = a = q$  y  $\hat{p} = b + \hat{q} = p$ .

Como se señaló anteriormente, en la actualidad este método es empleado principalmente como referencia comparativa, ampliándolo a mayores dimensiones de la matriz representada aquí como  $\mathbf{C}$  e imponiendo restricciones a los coeficientes (King, 1997; Klima et al., 2016). Otro criterio utilizado a modo referencial

---

<sup>13</sup>En cambio, la selección de un elemento  $Z_i$  sí resulta ser una variable aleatoria, en este caso, una variable de Bernoulli de parámetro  $\mu_Z$ .

se conoce como el “Neighborhood model” el cual se basa esta vez en el supuesto de que  $W$  y  $Z$  son independientes en todas las unidades intermedias  $\{u_1, \dots, u_M\}$ . Luego, los estimadores de las magnitudes desconocidas corresponden las sumas ponderadas de los valores obtenidos bajo el supuesto de independencia en cada unidad intermedia.

Una aplicación reciente del método de Goodman aplicado a problemas de transición electoral se encuentra en el trabajo de Russo (2014).

**2.2.2.2.3. Modelos generales para el caso  $R \times C$ .** La información disponible en un problema de estimación con datos agregados permite precisar intervalos con probabilidad 1 para las cantidades desconocidas, como resultado de la acotación de Duncan & Davis. Sin embargo, dichos intervalos resultan ser por lo general, si bien no triviales, demasiado amplios a efectos prácticos (King, 1997).

Llegado este punto, los abordajes del problema han recurrido a supuestos. En el modelo de Goodman el supuesto era coeficientes constantes en unidades de desagregación intermedia, y en el “Neighborhood model” la independencia entre ambas variables categóricas. Desde fines del siglo XX las anteriores suposiciones comienzan a ser desplazadas esta vez por supuestos de índole distribucional.

Entre las aproximaciones de carácter más reciente figuran una serie de modelos que recurren a estructuras bayesianas jerárquicas y supuestos de independencia condicional (King et al, 1999; Rosen et al., 2001; Greiner & Quinn, 2009).

La lógica tras este tipo de modelos en el contexto de estimación de matrices de transición electoral puede sintetizarse en los siguientes pasos.

1. Existen  $M$  unidades que forman una partición de una población finita  $U$ .
2. En cada unidad  $m = 1, \dots, M$  existe una matriz de transición de electoral que opera entre ambos procesos de votación. Las cantidades desconocidas al interior de estas matrices se representan como  $\beta_{rc}^m$ ,  $r = 1, \dots, R$ ,  $c = 1, \dots, C$ ,  $m = 1, \dots, M$ . Al tratarse de matrices de transición electoral, la suma de los elementos de cada fila es igual a 1.
3. Las cifras correspondientes al segundo proceso electoral en cada subunidad son modeladas bajo el supuesto que corresponden a realizaciones de  $M$  distribuciones multinomiales independientes.

El tercer paso conlleva la adopción del supuesto de independencia condicional: dado un conjunto de observaciones  $\{y_1, \dots, y_M\}$ , la distribución de  $y_m | \theta_m \stackrel{\perp}{\sim} p(\theta_m)$ ,  $m = 1, \dots, M$ . En este caso, esto se traduce en:

$$\{N_{1m}, \dots, N_{Cm}\} \stackrel{\perp}{\sim} \text{Multinomial} \left( N_m, \sum_{r=1}^R \beta_{r1}^m W_{rm}, \dots, \sum_{r=1}^R \beta_{rC}^m W_{rm} \right), \quad m = 1, \dots, M.$$

Continuando, el siguiente paso en la construcción de estos modelos es.

4. Las distribuciones multinomiales dependen de una estructura única en la población. Cada fila de las  $M$  matrices de transición electoral se asumen como realizaciones independientes por filas de una

distribución común a todas las subunidades de la población. En el modelo de Rosen et al. (2001) se utiliza en este paso la distribución de Dirichlet, y para el modelo de Greiner & Quinn (2009) una distribución normal.

Las distribuciones añadidas para las filas de las matrices de transición electoral corresponden en este caso a las distribuciones previas de los parámetros. En el contexto de la inferencia Bayesiana, una distribución previa permite incorporar conocimiento pre-existente acerca de las cantidades que se busca estimar. A estos efectos, se añade al modelo una medida de probabilidad  $\Pi$  para el mismo parámetro  $\pi(\theta)$ . Por su parte, la función de verosimilitud se ve como la distribución condicional de los datos dado el valor del parámetro.

$$L = f_{\theta}(x_1, \dots, x_n) = f(x_1, \dots, x_n | \theta)$$

De esta última expresión se tiene una distribución marginal para los datos  $m(\mathbf{x})$  la que en el caso continuo se obtiene integrando la siguiente expresión:

$$m(\mathbf{x}) = \int_{\Theta} \pi(\theta) f_{\theta}(\mathbf{x}) d\theta$$

Aquí el objetivo es obtener la distribución del parámetro condicional a los datos observados, o “distribución posterior”  $\pi(\theta | \mathbf{x})$ :

$$\pi(\theta | x_1, \dots, x_n) = \frac{\pi(\theta) f_{\theta}(\mathbf{x})}{m(\mathbf{x})}$$

Las inferencias en el modelo Bayesiano se basan en esta distribución posterior, la cual resume el conocimiento previo sobre el sistema y la evidencia muestral. Luego, la media de la distribución posterior opera como estimador puntual y en torno a aquella se obtiene una estimación por intervalo, bajo la forma de intervalos creíbles de Bayes<sup>14</sup>.

Retomando la descripción pretendida en esta sección, el quinto paso en los modelos bayesianos jerárquicos aplicados a problemas de estimación con datos agregados consiste en determinar la distribución posterior de los parámetros. A consecuencia de la adopción del supuesto de independencia condicional, estos modelos añaden un nivel adicional a la estructura distribucional, incorporando distribuciones “hiper-previas”.

5. Distribuciones hiper-previas: siendo la distribución previa y la posterior los dos primeros niveles del modelo, se añaden una o más distribuciones para modelar la pérdida de conocimiento previo sobre el valor del parámetro, originada por el modelamiento de observaciones con estructura de covarianza complejas.

En el modelo de Rosen et al. se utiliza en este nivel una distribución Gamma y en el de Greiner & Quinn una inversa de Wishart.

Finalmente, el sexto paso consisten en lo siguiente.

---

<sup>14</sup>Una descripción sintética del concepto de intervalos creíbles de Bayes es incluida en el Anexo 5.3.2.1 del documento.

6. Obtener la estimación de cada parámetro como la media de la distribución posterior de cada parámetro, condicional a los datos y los valores de los restantes parámetros. Ahora, dada la alta dimensionalidad paramétrica originada en la adopción del supuesto de independencia condicional, se recurre a estos efectos a métodos Montecarlo con Cadenas de Markov.

**2.2.2.2.3.1. El modelo de Rosen et al.** Como fue señalado con anterioridad, la evaluación de Klima et al. (2016) ubica al modelo jerárquico Multinomial-Dirichlet de Rosen et al. como aquel de mejor desempeño entre un conjunto de modelos de estimación con datos agregados para la inferencia de entradas de matrices de transición electoral. Una aplicación reciente de este modelo a problemas de transición electoral se encuentra en el trabajo de Gregor (2014).

El primer paso del modelo implica el supuesto de  $M$  distribuciones multinomial independientes.

$$\{N_{1m}, \dots, N_{Cm}\} \stackrel{\perp}{\sim} \text{Multinomial} \left( N_m, \sum_{r=1}^R \beta_{r1}^m W_{rm}, \dots, \sum_{r=1}^R \beta_{rC}^m W_{rm} \right), \quad m = 1, \dots, M. \quad (2.2.12)$$

En el segundo nivel se asume que los vectores  $\beta_r^m = (\beta_{r1}^m, \dots, \beta_{r,C-1}^m)^t$  siguen independientemente Distribución de Dirichlet. Nótese que para cada fila es una única distribución de Dirichlet.

$$\beta_r^m \stackrel{\perp}{\sim} \text{Dirichlet}(\alpha_{r1}, \dots, \alpha_{rC}), \quad r = 1, \dots, R.$$

En la elección de la distribución previa para modelar la incertidumbre respecto al valor del parámetro un concepto relevante es el de conjugación de distribuciones. Una distribución previa que de lugar a una distribución posterior de la misma forma funcional que la primera, aunque con diferentes parámetros, se denominan *previas conjugadas*. Esto ocurre en el caso de la distribución de Dirichlet como previa para el parámetro de la distribución Multinomial. Una descripción más extensa de este tema se incluye en el Anexo 5.3.2 del documento.

La incorporación del supuesto de independencia condicional permite arribar a una representación de la verosimilitud de los datos, con el costo de pérdida de conocimiento previo propia del enfoque general de estimación bayesiana. El modelo jerárquico se completa, entonces, con una distribución hiperprevia para las cantidades  $\alpha_{rc}$ , que corresponde en este caso a una misma distribución Gamma.

$$\alpha_{rc} \stackrel{i.i.d}{\sim} \text{Gamma}(\lambda_1, \lambda_2), \quad r = 1, \dots, R, c = 1, \dots, C. \quad (2.2.13)$$

Se arriba así, a una distribución conjunta. El objetivo es obtener la estimación  $\widehat{\beta}_{rc}^m$  en base a la media de la distribución posterior de cada parámetro, condicional a los datos y los valores de los restantes parámetros. Siguiendo la impronta de la estimación bayesiana, dichas cantidades se estiman con la media de las distribuciones:

$$\text{P} \left( \beta_{rc}^m | \{\beta_{jk}^m\}_{j \neq r}^{k \neq c}, N_{cm}, \alpha_{rc}, \lambda_1, \lambda_2 \right), \quad c = 1, \dots, C, m = 1, \dots, M \quad (2.2.14)$$

Ahora, dado el alto volumen de parámetros involucrados en el problema, obtener dichas distribuciones por la vía de la integración –para a continuación calcular la media posterior– no resulta factible. A estos efectos se recurre a métodos MCMC, en particular al “algoritmo Metropolis within Gibbs”. Por esta vía

se obtienen las estimaciones de  $\widehat{\beta}_{rc}^m$ . Luego, los estimadores para las entradas de la matriz corresponden a las sumas ponderadas para subíndices equivalentes de la estimaciones obtenidas para  $\widehat{\beta}_{rc}^m$ .

Entonces, sin necesidad de obtener cada distribución marginal vía integración, se obtiene una muestra de cada una de aquellas: “*the Metropolis algorithm is iterated, and the final value in this chain is treated as a deviate from the full conditional distribution*” (Rosen et al., 2001:6).

Existe, por tanto, una vinculación entre problemas que involucran la estimación de entradas de matrices de transición electoral y la teoría de Cadenas de Markov (“MC”). En primer lugar, la transición electoral en si misma puede describirse como una Cadena de Markov con espacio de estados no definido y probabilidades de transición no estacionarias. En segundo término, las MC tienen una vinculación funcional al ser utilizadas por los modelos bayesianos Jerárquicos para la simulación MCMC.

### 2.2.3. Relación con Cadenas de Markov

Una Cadena de Markov con espacio de estados discreto  $S$  y probabilidades de transición estacionarias  $p_{ij}$ ,  $i, j \in S$  es un proceso estocástico  $\{X_n : n = 1, \dots\}$  que satisface la propiedad de Markov:

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n)$$

A la distribución  $X_0$  se le denomina distribución inicial de la Cadena de Markov.

Por su parte, a las probabilidades  $p_{ij}(n, n+1) = p_{ij}(X_{n+1} = j | X_n = i)(n, n+1)$  se les llama probabilidades de transición del estado  $i$  en el tiempo  $n$ , al estado  $j$  en el tiempo  $n+1$ . Se dice que las probabilidades de transición  $p_{ij}(n, n+1)$  son estacionarias en el tiempo si no dependen de  $n$ . Por ejemplo:

$$p_{ij}(n, n+1) = p_{ij}(0, 1) = p_{ij}(1) = p_{ij}$$

Los valores  $p_{ij}$  se pueden representar al interior de una matriz denominada “estocástica”.

**Definición 2.2.1** Una matriz se denomina estocástica si satisface las siguientes dos propiedades:

1.  $p_{ij} \geq 0 \forall i, j \in S$ .
2.  $\sum_j p_{ij} = 1 \forall i$ .

Por consiguiente, los elementos que componen una Cadena de Markov con espacio de estados discretos y probabilidades de transición estacionarias son los siguientes:

1. Un espacio de estados  $S$ .
2. Una distribución de probabilidad inicial  $X_0$ .
3. Una matriz estocástica.

Una descripción más detalla de la terminología de Cadenas de Markov es incluida en Anexo 6.3.1 del documento.

### 2.2.3.1. Transición Electoral como Cadena de Markov

El espacio de estados no es fijo, sino dinámico, pues los agentes de representación política cambian a lo largo del tiempo. Interpretando los valores de la matriz de transición electoral  $\mathbf{R}$  como probabilidades, éstas no serán estacionarias en el tiempo. Por tanto no se cumple la propiedad de Markov. Como señalan Van del Ploeg et al. (2005) en los fenómenos sociales que evolucionan en el tiempo, incluidos dentro de éstos los procesos electorales, las cadenas de Markov no tienen aplicabilidad directa.

### 2.2.3.2. Uso de Métodos MCMC en modelos EI bayesianos Jerárquicos

En cambio, las Cadenas de Markov con probabilidades de transición estacionarias juegan un rol fundamental en la simulación en los métodos Montecarlo con Cadenas de Markov, los que son utilizados en modelos bayesianos Jerárquicos con aplicabilidad a problemas EI, y dentro de aquellos, a problemas de transición electoral.

En los modelos bayesianos jerárquicos las estimaciones se basan en la media de la distribución posterior de cada magnitud de interés, condicional a los datos y los restantes parámetros. En el modelo Multinomial-Dirichlet esto correspondía a:

$$\mathbb{P} \left( \beta_{rc}^m | \{\beta_{jk}\}_{j \neq r}^{k \neq c}, N_{cm}, \alpha_{rc}, \lambda_1, \lambda_2 \right), c = 1, \dots, C, m = 1, \dots, M$$

En el caso continuo, para obtener la distribución posterior de cada parámetro se requeriría integrar desde la distribución conjunta para determinar por esta vía la distribución marginal de interés. Pero en este caso se trata de un modelo sobre-parametrizado, es decir, contiene más parámetros que observaciones. Por ejemplo, para el caso de una versión del modelo que incluye una covariable sería necesario integrar una expresión como la siguiente.

$$\prod_{m=1}^M \prod_{c=1}^C \left( \sum_{r=1}^R \beta_{rc}^m X_{rm} \right)^{N_{cm}} \prod_{m=1}^M \prod_{r=1}^R \left\{ \frac{\Gamma \left( d_r \sum_{c=1}^C \exp(\gamma_{rc} + \delta_{rc} X_m) \right)}{\prod_{c=1}^C \Gamma(d_r \exp(\gamma_{rc} + \delta_{rc} X_m))} \prod_{c=1}^C \beta_{rc}^m d_r \exp(\gamma_{rc} + \delta_{rc} X_m) \right\} \exp \left( -\lambda \sum_{r=1}^R d_r \right)$$

Como se señala en Rosen et al.: “*Obtaining the marginals of this posterior distribution using high numerical integration is infeasible*” (2001: 5). Para afrontar este tipo de problema se utiliza el algoritmo Metrópolis within Gibbs, el cual, a su vez, combina dos algoritmos específicos.

- Muestreador de Gibbs: es un algoritmo que permite generar muestras de una distribución conjunta de difícil evaluación con herramientas disponibles. Para ello se requiere expresar la distribución de cada variable aleatoria –en el caso bayesiano, de cada parámetro, pues los parámetros se tratan como variables aleatorias– dada las demás como una distribución conocida. Luego, iterativamente se van simulando los valores. Llegado a la etapa  $N$  del procedimiento, se demuestra que, bajo condiciones generales de imparcialidad, las observaciones obtenidas convergen en distribución a la distribución conjunta objetivo.
- Algoritmo Metropolis-Hastings: las distribuciones condicionadas en el muestreador de Gibbs no siempre pueden identificarse como distribuciones conocidas, o con rutinas preestablecidas en los

softwares para simulación. El algoritmo Metropolis-Hastings construye una Cadena de Markov que tiene a la distribución condicionada de interés como distribución estacionaria.

Al alcanzar la convergencia de las Cadenas de Markov generadas con el algoritmo Metropolis se itera entonces el muestreador de Gibbs, obteniéndose en definitiva una muestra de la distribución conjunta y para cada distribución de acuerdo con (2.2.14).

Luego, si se ejecutan  $N$  Cadenas independientes, se obtendrá entonces una muestra *i.i.d.* de cada distribución marginal posterior, obteniendo entonces el cálculo de las medias posteriores e intervalos creíbles con fundamento en métodos Montecarlo basados en las leyes de los grandes números. Otra opción es ejecutar una única Cadena, lo que significa obtener como muestra para cada distribución observaciones correlacionadas. De ser este el caso, se requiere examinar el proceso obtenido como Serie de Tiempo, en este caso, tanto la convergencia a la distribución estacionaria, como asimismo la función de autocorrelación, siendo a partir de aquello, factible obtener las estimaciones como medias de las distribuciones marginales posteriores.

El algoritmo involucra detalles adicionales, que son descritos con mayor profundidad en el Anexo 5.3.2.4 del documento.

#### 2.2.4. Síntesis

Los métodos de inferencia ecológica buscan estimar una función de los valores de la distribución conjunta de dos variables categóricas  $W$  y  $Z$  definidas sobre una misma población finita  $U$ . En el caso electoral estas dos variables corresponden a estadísticas de resultados y participación electoral asociadas a dos eventos eleccionarios consecutivos. Para ello se utiliza la información de las distribuciones conocidas  $g_{1m}(W)$  y  $g_{2m}(Z)$ ,  $m = 1, \dots, M$ . Haciendo uso de este tipo de información, el método de las cotas de Duncan & Davis (1953) ofrece intervalos con probabilidad 1 para las magnitudes desconocidas, los cuales sin embargo, suelen resultar demasiado amplios a efectos prácticos. Como alternativa a esto último se han desarrollado dos tipos de métodos de estimación: 1) Modelos Stayers-movers, los cuales recurren a algoritmos de optimización con restricciones; y 2) Modelos con supuestos distribucionales, basados principalmente en estructuras bayesianas jerárquicos y métodos Montecarlo con Cadenas de Markov.

## 2.3. Propuesta de Estimación con datos agregados

En esta sección se procede a vincular los métodos de estimación de matrices de transición electoral utilizando datos agregados con el problema de la composición del electorado tras dos eventos electorales consecutivos. Junto a ello se propone una estrategia para la acotación del espacio paramétrico del problema, lo que posibilita el cumplimiento del cuarto objetivo específico del proyecto. Las proposiciones incluidas en el capítulo son todas de carácter inédito.

### 2.3.1. Estimación de la Composición del Electorado con Métodos EI

De acuerdo con el desarrollo incluido en el capítulo precedente, para estimar la composición del electorado utilizando datos agregados se requiere ubicar las magnitudes desconocidas al interior de una matriz cuya suma de elementos por filas y columnas es conocida. Esto se cumple para la matriz representada en este trabajo como  $\mathbf{X}$ .

#### 2.3.1.1. Propiedades de la matriz $\mathbf{X}$

Dentro de las propiedades que pueden deducirse preliminarmente de esta matriz figuran las siguientes.

1. Por construcción,  $x_{k+2,k+2} = 0$ . Debido a aquello, contiene  $(k+2)^2 - 1$  cantidades desconocidas.
2. Se cumple que  $\sum_{r=1}^{k+1} \sum_{c=1}^{k+2} x_{rc} = 1$ .
3. Son conocidas las sumatorias de sus entradas por filas y columnas.
4.  $tr(\mathbf{X}) \leq 1 - V^* + p_8$ , donde  $V^*$  es una magnitud identificable a partir del grado de variación de los resultados de votación.

La primera propiedad se obtiene de  $x_{k+2,k+2} = \frac{1}{N_2} \# \{(A_6 \cup A_7) \cap A_8 = \emptyset\} = 0$ . Dado que se requieren al menos dos agentes de representación en competencia, la matriz en cuestión tendrá como mínimo  $4^2 - 1 = 15$  magnitudes desconocidas. La segunda propiedad proviene del hecho que  $\sum_{r=1}^{k+1} \sum_{c=1}^{k+2} x_{rc} = \sum_{i=1}^7 p_i = 1$ .

La demostración de la cuarta propiedad se posterga a la sección siguiente. En cuanto a la tercera propiedad, se puede plantear lo siguiente.

**Proposición 2.3.1** Sea  $\mathbf{X} = \frac{n_1}{N_2} \Phi \mathbf{B}$ . Entonces se cumple que:

- Cada una de las primeras  $k$  columnas debe sumar  $\frac{d_{1j}}{N_2}$ .
- La suma de la columna  $k+1$  es  $\frac{N_1 - n_1}{N_2}$ .
- La suma de la columna  $k+2$  es  $\frac{nv_2}{N_2}$ .

DEMOSTRACIÓN: el vector en la columna  $j$  dentro de las primeras  $k$  columnas de  $\Phi \mathbf{B}$  es un vector con entradas:

$$\frac{1}{n_1} [\phi_{1j}d_{1j} \quad \phi_{2j}d_{1j} \quad \dots \quad \phi_{k+1,j}d_{1j} \quad \phi_{k+2,j}d_{1j}]^t \quad (2.3.1)$$

Cuya suma de elementos es:

$$\frac{1}{n_1} \sum_{r=1}^{k+2} \phi_{rj}d_{1j} = \frac{d_{1j}}{n_1} \sum_{r=1}^{k+2} \phi_{rj} = \frac{d_{1j}}{n_1} 1 = \frac{d_{1j}}{n_1}$$

Luego, como  $\mathbf{X} = \frac{n_1}{N_2} \Phi \mathbf{B}$ , la suma de los elementos en los primeros  $k$  vectores columna es:

$$\frac{n_1}{N_2} \left( \frac{d_{1j}}{n_1} \right) = \frac{d_{1j}}{N_2} \quad (2.3.2)$$

La demostración para el vector columna  $k+1$  y  $k+2$  sigue la misma lógica, observando que los elementos en los vectores columna  $k+1$  y  $k+2$  de  $\Phi \mathbf{B}$  son, respectivamente:

$$\frac{(N_1 - n_1)}{n_1} [\phi_{1,k+1} \quad \phi_{2,k+1} \quad \dots \quad \phi_{k+1,k+1} \quad \phi_{k+2,k+1}]^t \quad (2.3.3)$$

$$\frac{nv_2}{n_1} [\phi_{1,k+2} \quad \phi_{2,k+2} \quad \dots \quad \phi_{k+1,k+2} \quad \phi_{k+2,k+2}]^t \quad (2.3.4)$$

Que al proceder luego como en (2.3.2) llega a que los elementos en las columnas  $k+1$  y  $k+2$  de  $\mathbf{X}$  suman:  $\frac{N_1 - n_1}{N_2}$  y  $\frac{nv_2}{N_2}$ , respectivamente.

**Proposición 2.3.2** Sea  $\mathbf{X} = \frac{n_1}{N_2} \Phi \mathbf{B}$ . Entonces se cumple que:

- Cada una de las primeras  $k$  filas debe sumar  $\frac{d_{2j}}{N_2}$ .
- La suma de la fila  $k+1$  es  $\frac{N_2 - n_2}{N_2}$ .
- La suma de la fila  $k+2$  es  $\frac{N_1 + nv_2 - N_2}{N_2}$ .

DEMOSTRACIÓN: de acuerdo con la representación de los vectores columna de  $\Phi \mathbf{B}$  dada por (2.3.1), (2.3.3) y (2.3.4), una representación genérica de los elementos en cada uno de los primeros  $k$  vectores filas de  $\mathbf{X}$  es:

$$\frac{1}{N_2} [\phi_{j1}d_{11} \quad \phi_{j2}d_{12} \quad \dots \quad \phi_{j,k+1}(N_1 - n_1) \quad \phi_{j,k+2}nv_2] \quad (2.3.5)$$

Que junto con (2.1.9) lleva a que la suma de los elementos para los primeros  $k$  vectores fila sea igual a:

$$\frac{d_{2j}}{N_2}, j = 1, \dots, k \quad (2.3.6)$$

En el caso de la fila  $k+1$  notemos que:

$$\sum_{c=1}^{k+2} x_{k+1,c} = p_3 + p_5 + p_7 = \frac{1}{N_2} [\#\{A_3 \cup A_5 \cup A_7\}] = \frac{1}{N_2} [\#\{E_2 \cap T_2^c\}] = \frac{1}{N_2} [\#\{E_2 - T_2\}] =$$

$$= \frac{1}{N_2} [\#\{E_2\} - \#\{T_2\}] = \frac{N_2 - n_2}{N_2} \quad (2.3.7)$$

Finalmente, en el caso de la fila  $k + 2$  se parte de la siguiente identidad:

$$N_2 = N_1 + nv_2 - N_2 p_8 \Rightarrow p_8 = \frac{N_1 + nv_2 - N_2}{N_2} \quad (2.3.8)$$

Pero, por construcción,  $p_8 = \sum_{c=1}^{k+2} x_{k+2,c}$ , de modo que:

$$\sum_{c=1}^{k+2} x_{k+2,c} = \frac{N_1 + nv_2 - N_2}{N_2} \quad (2.3.9)$$

**Corolario 2.3.1** Sea  $\mathbf{X} = \frac{n_1}{N_2} \Phi \mathbf{B}$ . Entonces:

- La suma conjunta de las primeras  $k$  filas es  $\frac{n_2}{N_2}$
- La suma conjunta de las primeras  $k$  columnas es  $\frac{n_1}{N_2}$
- La suma conjunta de las primeras  $k + 1$  columnas es  $\frac{N_1}{N_2}$

DEMOSTRACIÓN: por (2.3.5), para la suma de los elementos ubicados en las primeras  $k$  filas se tiene:

$$\sum_{r=1}^k \sum_{c=1}^{k+2} x_{rc} = \sum_{r=1}^k \frac{d_{2r}}{N_2} = \frac{n_2}{N_2} \quad (2.3.10)$$

El caso de las primeras  $k$  columnas es análogo:

$$\sum_{c=1}^k \sum_{r=1}^{k+2} x_{rc} = \sum_{c=1}^k \frac{d_{1c}}{N_2} = \frac{n_1}{N_2} \quad (2.3.11)$$

De acuerdo con esto último, junto a (2.3.3) y (2.3.4) llevan a que la suma de las primeras  $k + 1$  columnas sea.

$$\sum_{c=1}^{k+1} \sum_{r=1}^{k+2} x_{rc} = \frac{n_1}{N_2} + \frac{N_1 - n_1}{N_2} = \frac{N_1}{N_2} \quad (2.3.12)$$

Para esquematizar, y suponiendo que compiten más de 2 opciones políticas en ambos procesos electorales, una posible representación de esta matriz en la siguiente.

$$\begin{array}{cccccc}
x_{11} & x_{12} & \dots & x_{1k} & x_{1k+1} & x_{1k+2} & \frac{d_{21}}{N_2} \\
x_{21} & x_{22} & \dots & x_{2k} & x_{2k+1} & x_{2k+2} & \frac{d_{22}}{N_2} \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots \\
x_{k1} & x_{k2} & \dots & \dots & \dots & x_{k,k+2} & \frac{d_{2k}}{N_2} \\
x_{k+1,1} & \dots & \dots & \dots & \dots & \dots & \frac{N_2 - n_2}{N_2} \\
\dots & \dots & \dots & \dots & \dots & \dots & \frac{N_1 + nv_2 - N_2}{N_2} \\
\hline
\frac{d_{11}}{N_2} & \frac{d_{12}}{N_2} & \dots & \frac{d_{1k}}{N_2} & \frac{N_1 - n_1}{N_2} & \frac{nv_2}{N_2} & 
\end{array}$$

Entonces, la estimación de las entradas de la matriz, y luego, del vector de parámetros, puede ser realizada empleando métodos EI.

### 2.3.1.2. Estrategia de Estimación

Las cantidades  $\{p_1, \dots, p_7, p_8\}$  son parámetros poblacionales. Por tanto, para cada par de procesos electorales consecutivos existirá una única matriz  $\mathbf{X}$  que contiene estos valores. Sin embargo, debido al carácter secreto del voto dichas magnitudes son desconocidas. En cambio, existe una cantidad inconmensurablemente alta de matrices que cumplen las restricciones por filas y columnas.

Un estimador puntual del vector de parámetros se obtiene entonces al seleccionar una matriz dentro del conjunto que satisfacen las restricciones por filas y columnas.

Los métodos EI previamente descritos posibilitan obtener una estimación de la transposición de la matriz aquí representada como  $\Phi$ . Sea  $\widehat{\Phi}_o$  (de “optimización”) la estimación obtenida con el método Stayers-Movers; y sea  $\widehat{\Phi}_h$  (de “hierarchical”) aquella obtenida con el modelo bayesiano Jerárquico Multinomial-Dirichlet. Luego, asociada a cada matriz  $\widehat{\Phi}_\bullet$  se propone como estimadores de las entradas de la matriz  $\mathbf{X}$ :

- $\widehat{\mathbf{X}}_o = \frac{n_1}{N_2} \widehat{\Phi}_o \mathbf{B}$
- $\widehat{\mathbf{X}}_h = \frac{n_1}{N_2} \widehat{\Phi}_h \mathbf{B}$

Donde los valores  $n_1$  y  $N_2$  son conocidos, al igual que las entradas de  $\mathbf{B}$ . Luego, los estimadores del vector de parámetros  $\hat{\theta} = (\hat{p}_1, \hat{p}_2, \dots, \hat{p}_7)$  se pueden obtener como combinaciones lineales de los valores de cada matriz obtenida  $\widehat{\mathbf{X}}_\bullet$ .

### 2.3.2. Acotación del Espacio Paramétrico

De la definición del espacio paramétrico del problema dada en (2.1.3):

$$\Theta = \left\{ (p_1, p_2, \dots, p_7) / 0 \leq p_i \leq 1, \sum_{i=1}^7 p_i = 1 \right\}$$

La tercera propiedad de la matriz aquí simbolizada como  $\mathbf{X}$ , junto con lo establecido por la acotación de Duncan & Davis (1953), que fue descrita en la sección 2.2.2.2.1, permiten proponer un procedimiento para la acotación de  $\Theta$ .

### 2.3.2.1. Acotación a partir de sumas por filas y columnas conocidas

Notemos que la acotación de Duncan & Davis vale para toda matriz con entradas no negativas y sumas conocidas por filas y columnas.

Dado aquello, el primer paso en la estrategia propuesta es llevar las cantidades desconocidas a una representación de matriz 2x2, sumando las entradas que son objeto de agrupación. A modo genérico:

$$\begin{bmatrix} p_i & \dots \\ \dots & \dots \end{bmatrix}$$

De la tercer propiedad de la matriz  $\mathbf{X}$ , cualquier representación 2x2 de aquella será también una matriz con sumas por filas y columnas conocidas. Luego, simbolizando como en la acotación de Duncan & Davis como  $a$  y  $c$  las sumas de los elementos ubicados en las filas 1 y 2, respectivamente; y  $b$  y  $d$  las sumas de las columnas 1 y 2, respectivamente, se llega en este caso a que.

$$\max(0, a - d) \leq p_i \leq \min(a, b)$$

**Proposición 2.3.3**  $\frac{1}{N_2} \max(0, [(n_2 + n_1) - (N_1 + nv_2)]) \leq (p_1 + p_2) \leq \frac{1}{N_2} \min(n_1, n_2)$

DEMOSTRACIÓN: las cantidades  $p_1$  y  $p_2$  se obtienen como sumatorias de los elementos ubicados en las primeras  $k$  filas y  $k$  columnas de la matriz:

$$\sum_{r=1}^k \sum_{c=1}^k x_{rc} = p_1 + p_2$$

Al ubicar esta suma en una matriz 2x2 del siguiente modo.

$$\begin{bmatrix} p_1 + p_2 & \dots \\ \dots & \dots \end{bmatrix}$$

Se tiene por el corolario de la tercer propiedad de la matriz que:

- $a = \sum_{r=1}^k \sum_{c=1}^{k+2} x_{rc} = \frac{n_2}{N_2}$
- $b = \sum_{c=1}^k \sum_{r=1}^{k+2} x_{rc} = \frac{n_1}{N_2}$
- $c = \sum_{c=1}^{k+2} \sum_{r=k+1}^{k+2} x_{rc} = \frac{N_2 - n_2}{N_2} + \frac{N_1 + nv_2 - N_2}{N_2} = \frac{N_1 + nv_2 - n_2}{N_2}$
- $d = \sum_{r=1}^{k+2} \sum_{c=k+1}^{k+2} x_{rc} = \frac{N_1 - n_1 + nv_2}{N_2}$

Así, en este caso:

$$\begin{aligned} \max(0, a - d) \leq (p_1 + p_2) \leq \min(a, b) &\Rightarrow \\ \Rightarrow \frac{1}{N_2} \max(0, [(n_2 + n_1) - (N_1 + nv_2)]) \leq (p_1 + p_2) \leq \frac{1}{N_2} \min(n_1, n_2) &\quad (2.3.13) \end{aligned}$$

**Proposición 2.3.4**  $\frac{1}{N_2} \max(0, [(N_2 + n_1) - (N_1 + n_2 + nv_2)]) \leq p_3 \leq \frac{1}{N_2} \min(n_1, N_2 - n_2)$

DEMOSTRACIÓN: la cantidad  $p_3$  se forma como suma de los elementos ubicados en la fila  $k + 1$  y primeras  $k$  columnas de la matriz.

$$\sum_{c=1}^k x_{k+1,c} = p_3$$

Llevando la representación de la matriz a una forma 2x2.

$$\begin{bmatrix} p_3 & \dots \\ \dots & \dots \end{bmatrix}$$

Entonces, en este caso:

- $a = \sum_{c=1}^{k+2} x_{k+1,c} = \frac{N_2 - n_2}{N_2}$
- $b = \sum_{r=1}^{k+2} \sum_{c=1}^k x_{rc} = \frac{n_1}{N_2}$
- $c = \sum_{r \neq k+1} \sum_{c=1}^{k+2} x_{rc} = \frac{N_1 + nv_2 + n_2 - N_2}{N_2}$
- $d = \sum_{c=k+1}^{k+2} \sum_{r=1}^{k+2} x_{rc} = \frac{N_1 - n_1 + nv_2}{N_2}$

Por consiguiente, se tiene aquí:

$$\begin{aligned} \max(0, a - d) \leq p_3 \leq \min(a, b) &\Rightarrow \\ \Rightarrow \frac{1}{N_2} \max(0, [(N_2 + n_1) - (N_1 + n_2 + nv_2)]) \leq p_3 \leq \frac{1}{N_2} \min(n_1, N_2 - n_2) &\quad (2.3.14) \end{aligned}$$

Para los restantes parámetros  $p_i, i = 4, \dots, 7$  se procede del mismo modo, permitiendo establecer en general que:

$$\begin{aligned} \frac{1}{N_2} \max(0, [(n_2 + n_1) - (N_1 + nv_2)]) \leq (p_1 + p_2) &\leq \frac{1}{N_2} \min(n_1, n_2) \\ \frac{1}{N_2} \max(0, [(N_2 + n_1) - (N_1 + n_2 + nv_2)]) \leq p_3 &\leq \frac{1}{N_2} \min(n_1, N_2 - n_2) \\ \frac{1}{N_2} \max(0, [n_2 - n_1 - nv_2]) \leq p_4 &\leq \frac{1}{N_2} \min(n_2, N_1 - n_1) \\ \frac{1}{N_2} \max(0, [N_2 - n_2 - n_1 - nv_2]) \leq p_5 &\leq \frac{1}{N_2} \min(N_2 - n_2, N_1 - n_1) \\ \frac{1}{N_2} \max(0, [n_2 - N_1]) \leq p_6 &\leq \frac{1}{N_2} \min(n_2, nv_2) \\ \frac{1}{N_2} \max(0, [N_2 - n_2 - N_1]) \leq p_7 &\leq \frac{1}{N_2} \min(N_2 - n_2, nv_2) \end{aligned} \quad (2.3.15)$$

### 2.3.2.2. Acotación superior a partir del nivel de Volatilidad electoral observado

La volatilidad electoral observada entre procesos de votación consecutivos ofrece, indirectamente, información adicional para acotar el espacio paramétrico. Intuitivamente, a medida que la volatilidad electoral se incrementa, el porcentaje de votantes que se mantienen en el mismo estado en ambos comicios habrá de disminuir. En el límite, si la volatilidad asume su valor máximo, la cardinalidad del conjunto  $A_1$  habrá de ser 0.

Para traducir esta idea intuitiva en una proposición operativa, se requiere vincular la volatilidad electoral observada con la matriz  $\mathbf{X}$ . Se comienza proponiendo una condición *necesaria y suficiente* para  $V = 0$ .

**Proposición 2.3.5**  $V = 0 \Leftrightarrow d_{2j} = cd_{1j} \forall j = 1, \dots, k$ , donde  $d_{1j}$  y  $d_{2j}$  corresponden a la entrada  $j$  de  $\mathbf{d}_1$  y  $\mathbf{d}_2$  respectivamente y  $c \geq 0$  una constante.

DEMOSTRACIÓN: Podemos reescribir la ecuación escalar del Índice de Pedersen en los siguientes términos.

$$V = \frac{1}{2} \sum_{j=1}^k \Delta v_j$$

Donde  $\Delta v_j = |v_{(t+1),j} - v_{(t)j}|$ .

Como  $\Delta v_j \geq 0$ , entonces  $V = 0 \Leftrightarrow \Delta v_j = 0 \forall j = 1, \dots, k$ . Pero  $\Delta v_j = |v_{(t+1),j} - v_{(t)j}| = \left| \frac{d_{2j}}{n_2} - \frac{d_{1j}}{n_1} \right| \Rightarrow V = 0 \Leftrightarrow \frac{d_{2j}}{n_2} = \frac{d_{1j}}{n_1} \forall j = 1, \dots, k$ .

Notemos que  $n_1 = \sum_{j=1}^k d_{1j}$  y también  $n_2 = \sum_{j=1}^k d_{2j}$ . Luego, si  $d_{2j} = cd_{1j} \forall j = 1, \dots, k$  entonces  $n_2 = c \sum_{j=1}^k d_{1j}$ . Por tanto:

$$\frac{d_{2j}}{n_2} = \frac{cd_{1j}}{c \sum_{j=1}^k d_{1j}} = \frac{d_{1j}}{n_1} \forall j = 1, \dots, k \Rightarrow V = 0$$

A la inversa, si  $V = 0$  entonces  $\frac{d_{2j}}{n_2} = \frac{d_{1j}}{n_1} \forall j = 1, \dots, k \Rightarrow d_{2j} = \frac{n_2}{n_1} d_{1j} = cd_{1j}$ .

De lo anterior se desprende una condición *suficiente* para  $V = 0$  a partir de la matriz  $\mathbf{X}$ .

**Proposición 2.3.6** Si  $\mathbf{X}$  es simétrica, entonces  $V = 0$ .

DEMOSTRACIÓN: Sea  $\Delta v_j^* = \left| \sum_{r=1}^{k+2} x_{rj} - \sum_{c=1}^{k+2} x_{jc} \right|$ . Es claro que si  $\Delta v_j^* = 0 \forall j = 1, \dots, k$ , entonces  $d_{1j} = d_{2j} \forall j = 1, \dots, k \Rightarrow n_2 = n_1 \Rightarrow \frac{d_{2j}}{n_2} = \frac{d_{1j}}{n_1} \forall j = 1, \dots, k \Rightarrow V = 0$ .

Ahora  $\Delta v_j^* = \left| \sum_{r=1}^{k+2} x_{rj} - \sum_{c=1}^{k+2} x_{jc} \right| = \left| \sum_{c=1}^{k+2} (x_{cj} - x_{jc}) \right|, c \neq j$ .

De esta manera se ve que si  $\mathbf{X}$  es simétrica, entonces:

$$x_{jc} = x_{cj} \Rightarrow \left| \sum_{c=1}^{k+2} (x_{cj} - x_{jc}) \right| = 0 \Rightarrow d_{2j} = 1d_{1j} = d_{1j} \forall j = 1, \dots, k \Rightarrow V = 0 \quad (2.3.16)$$

**Corolario 2.3.2** si  $\mathbf{X}$  es diagonal, entonces  $V = 0$ .

DEMOSTRACIÓN: se obtiene directamente a partir del hecho de que toda matriz diagonal es simétrica.

En base a esto último se puede ver más claramente la confusión existente al equiparar volatilidad agregada, o neta, y volatilidad individual –esta última definida en los términos de Bartolini & Mair (1990)–, es decir, al no restringir el uso de la primera al estatus de indicador indirecto de la segunda, sino que extenderlo al de “estimador indirecto”. La interpretación estándar en los estudios de la institucionalización partidista tiende a asumir implícitamente que cuando  $V$  es próxima a cero, entonces  $\mathbf{X}$  resulta –en la terminología empleada en este proyecto– *próxima* a ser diagonal. Aquello equivale a incurrir en una falacia ecológica.

La definición  $\Delta v_j^*$  sirve a la acotación superior de  $p_1$  y  $p_5$ . Para ello se define la siguiente medida:

$$V^* = \frac{1}{2} \sum_{j=1}^{k+2} \Delta v_j^* \quad (2.3.17)$$

Nótese que en este caso la sumatoria se entiende hasta  $k + 2$ .

**Proposición 2.3.7**  $p_1 + p_5 \leq 1 - V^* + p_8$

DEMOSTRACIÓN: Como  $x_{rc} \geq 0$ :

$$\begin{aligned} \Delta v_1^* &= |(x_{12} - x_{21}) + \dots + (x_{1k+2} - x_{k+2,1})| \leq (x_{12} + x_{21}) + \dots + (x_{1k+2} + x_{k+2,1}) \\ \Delta v_2^* &= |(x_{21} - x_{12}) + \dots + (x_{2k+2} - x_{k+2,2})| \leq (x_{21} + x_{12}) + \dots + (x_{2k+2} + x_{k+2,2}) \\ &\vdots \\ \Delta v_{k+2}^* &\leq (x_{k+2,1} + x_{1,k+2}) + \dots + (x_{k+2,k+1} + x_{k+1,k+2}) \end{aligned}$$

Luego, al realizar la sumatoria:

$$\begin{aligned} &\sum_{j=1}^{k+2} \Delta v_j^* \leq 2p_2 + 2p_3 + 2p_4 + 2p_6 + 2p_7 + 2p_8 \Rightarrow \\ \Rightarrow &\frac{1}{2} \sum_{j=1}^{k+2} \Delta v_j^* \leq \left( \sum_{i=2}^8 p_i \right) - p_5 \Rightarrow V^* \leq \left( \sum_{i=2}^8 p_i \right) - p_5 \Rightarrow V^* + p_5 \leq \sum_{i=2}^8 p_i \end{aligned} \quad (2.3.18)$$

Pero  $p_1 = 1 - \sum_{i=2}^7 p_i$ , de manera que:

$$V^* + p_5 \leq 1 - p_1 + p_8 \Rightarrow p_1 + p_5 \leq 1 - V^* + p_8 \quad (2.3.19)$$

De momento no resulta estrictamente necesario reescribir la última acotación en términos de  $V$ , porque  $V^*$  se puede calcular directamente con información disponible, dada la tercer propiedad de la matriz  $\mathbf{X}$  enunciada al inicio de la sección 2.3.1. Es decir, se obtendría simplemente la misma acotación, escrita en

otros términos.

Como corolario, de esto último se obtiene la demostración de la cuarta propiedad de la matriz  $\mathbf{X}$

**Corolario 2.3.3** *Sean dos procesos de votación consecutivos con vectores de resultados electorales  $\mathbf{v}_1$  y  $\mathbf{v}_2$ , tales que solicitan al electorado emitir una única preferencia y los mismos  $k$  agentes de representación política concentran la totalidad de la oferta electoral; y sea  $\mathbf{X} = \frac{n_1}{N_2} \Phi \mathbf{B}$ . Entonces  $tr(\mathbf{X}) \leq 1 - V^* + p_8$ .*

DEMOSTRACIÓN: Se obtiene de (2.3.19) junto con el hecho de que  $x_{k+2,k+2} = 0$ .

También, lo anterior permite extender la acotación del espacio paramétrico expresada previamente en (2.3.15).

$$\begin{aligned}
0 \leq p_1 &\leq \min \left( \frac{n_1}{N_2}, \frac{n_2}{N_2}, 1 - V^* + p_8 \right) \\
0 \leq p_2 &\leq \frac{1}{N_2} \min(n_1, n_2) \\
\frac{1}{N_2} \max(0, [(n_2 + n_1) - (N_1 + nv_2)]) &\leq (p_1 + p_2) \leq \frac{1}{N_2} \min(n_1, n_2) \\
\frac{1}{N_2} \max(0, [(N_2 + n_1) - (N_1 + n_2 + nv_2)]) &\leq p_3 \leq \frac{1}{N_2} \min(n_1, N_2 - n_2) \\
\frac{1}{N_2} \max(0, [n_2 - n_1 - nv_2]) &\leq p_4 \leq \frac{1}{N_2} \min(n_2, N_1 - n_1) \\
\frac{1}{N_2} \max(0, [N_2 - n_2 - n_1 - nv_2]) &\leq p_5 \leq \min \left( \frac{N_2 - n_2}{N_2}, \frac{N_1 - n_1}{N_2}, 1 - V^* + p_8 \right) \\
\frac{1}{N_2} \max(0, [n_2 - N_1]) &\leq p_6 \leq \frac{1}{N_2} \min(n_2, nv_2) \\
\frac{1}{N_2} \max(0, [N_2 - n_2 - N_1]) &\leq p_7 \leq \frac{1}{N_2} \min(N_2 - n_2, nv_2)
\end{aligned} \tag{2.3.20}$$

### 2.3.3. Síntesis

Las sumas por filas y columnas de la matriz representada en este trabajo como  $\mathbf{X}$  resultan ser cifras conocidas. Ello posibilita la estimación de las entradas de la matriz, y luego del vector de parámetros, con métodos EI. Bajo estas condiciones, tanto el modelo bayesiano como el método de optimización ofrecen una estimación de la matriz  $\Phi$ . Luego, asociada a cada  $\widehat{\Phi}_\bullet$ , se propone como estimador  $\widehat{\mathbf{X}}_\bullet = \frac{n_1}{N_2} \widehat{\Phi}_\bullet \mathbf{B}$ , con  $n_1$  y  $N_2$  conocidos, al igual que las entradas de  $\mathbf{B}$ .

Las mismas sumas conocidas permiten establecer una acotación del espacio paramétrico, como resultado de la aplicación del método de Duncan & Davis (1953). Asimismo, la volatilidad electoral observada conduce a una acotación superior adicional para  $p_1$  y  $p_5$ . Combinando ambos procedimientos se obtiene una acotación completa del espacio de parámetros, dado por (2.3.20). Luego, esta acotación sirve para evaluar los resultados obtenidos con el uso de los métodos de estimación con datos agregados empleados en este proyecto.

## Capítulo 3

# Aplicación de la Propuesta a Datos reales

En este capítulo se aplica la propuesta a datos reales, correspondientes a la competencia entre conglomerados políticos en Chile en elecciones de Diputados celebradas entre los años 1993 y 2009. La elección de este caso responde a que para el mismo periodo señalado se cuenta con un antecedente comparativo sobre el mismo tema generado a partir del análisis de encuestas electorales con preguntas retrospectivas (Avendaño & Sandoval, 2013), lo que permite evaluar el grado de concordancia de los resultados alcanzados. Con ello se da cumplimiento al octavo objetivo específico del proyecto. Asimismo, se enuncia una alternativa para asociar una región de probabilidad a los resultados obtenidos con la aplicación de los métodos empleados en este estudio (objetivos específicos 1 y 6). Finalmente, sobre la base de los hallazgos alcanzados se propone mejoras al abordaje empírico del fenómeno de la institucionalización partidista (objetivo específico 9).

### 3.1. Resumen de la Propuesta

La utilización de métodos EI que fueron descritos en la sección 2.2 permite obtener una estimación de la matriz representada en este proyecto como  $\Phi$ . Luego, para cada par de procesos electorales consecutivos una estimación de la matriz  $\mathbf{X}$  se puede obtener como:

- $\widehat{\mathbf{X}}_o = \frac{n_1}{N_2} \widehat{\Phi}_o \mathbf{B}$
- $\widehat{\mathbf{X}}_h = \frac{n_1}{N_2} \widehat{\Phi}_h \mathbf{B}$

A continuación, el vector de parámetros  $\hat{\theta} = (\hat{p}_1, \hat{p}_2, \dots, \hat{p}_7)$  es estimado como combinaciones lineales de los valores de cada matriz obtenida  $\widehat{\mathbf{X}}_\bullet$ .

Para evaluar el ajuste de los resultados obtenidos con ambos métodos se recurre a la acotación del espacio paramétrico dada en la sección 2.3.2, así como también a antecedentes previos procedentes de encuestas electorales.

## 3.2. Aplicación al caso de Chile 1993-2009

Los datos utilizados para el análisis corresponden a resultados electorales y cifras de composición del padrón de votantes a nivel de municipios en Chile entre los años 1993 y 2009.

### 3.2.1. Descripción Sintética del caso

En Chile, los Municipios constituyen la menor unidad de la división político-administrativa del Estado. En este plano, Chile se subdivide desde fines de la década de 1980 en regiones, provincias y municipios. Las autoridades regionales (intendentes) y provinciales (gobernadores) son designadas por el poder ejecutivo central. En cambio, las autoridades municipales (Gobierno local) son elegidas por medio de procesos electorarios.

En conjunto, los municipios constituyen una partición del país a nivel geográfico y, por tanto, también demográfico. Para diferenciar a la institucionalidad misma del Gobierno local (el Municipio), de la unidad territorial y demográfica respectiva, se suele utilizar el término “comuna” para referir a la segunda. El número de comunas existentes al año 1993 ascendió a 335, cifra que se incrementó posteriormente como resultado de la subdivisión de algunas de las originales. En 1997 y 2001 este número ascendió a 342, mientras que a partir de 2005 son 346.

Tres tipos de elecciones, en términos de las autoridades a ser electas, fueron efectuadas en el periodo: presidenciales, parlamentarias (Senadores y Diputados) y municipales (Alcaldes y “Concejales”).

En este trabajo se analizan resultados de elecciones de Diputados, las cuales fueron celebradas con regularidad de 4 años desde el periodo de restitución del régimen democrático, el año 1989. Este tipo de proceso electoral solicitaba al electorado emitir una única preferencia en cada comicio. Para las elecciones de Diputados entre 1989 y 2009 rigió el llamado sistema electoral “Binominal”, el cual incentivó la competición entre conglomerados políticos, antes que entre partidos.

La unidad electoral para la elección de Diputados corresponde a la Circunscripción, la cual agrupa una o más comunas. No obstante esto último, en el país, los resultados de cada elección se ofrecen desagregados para su consulta pública a nivel regional, provincial y también comunal<sup>15</sup>.

Finalmente, por el lado de la participación electoral, a lo largo del periodo el sistema electoral en Chile establecía como requisitos para integrar el electorado la mayoría de edad (mayor a 18 años) y la inscripción voluntaria en el Padrón electoral. Sin embargo, al figurar en el padrón, participar en los diferentes procesos electorales se convertía en obligatorio (Inscripción voluntaria y voto obligatorio). Con posterioridad al periodo aquí estudiado el sistema se modificó a un sistema de inscripción automática (a partir del requisito etario) y voto voluntario.

---

<sup>15</sup>Ver por ejemplo: <https://historico.servel.cl/SitioHistorico>

## 3.2.2. Datos

### 3.2.2.1. Tipo de Datos y Estrategia de Obtención

**3.2.2.1.1. Resultados Electorales.** Corresponden a los resultados de las elecciones de Diputados celebradas los años 1993, 1997, 2001, 2005 y 2009 desagregados a nivel comunal. Estos datos son de dominio público y pueden obtenerse directamente del sitio web del Servicio Electoral de Chile (SERVEL) desde la dirección <https://www.servel.cl/elecciones-de-diputados-1989-al-2013-por-circunscripcion-electoral/>.

Estos resultados se presentan por Listas, diferenciadas con letras (A,B,C.,etc.), las cuales agrupaban partidos políticos pertenecientes a diferentes coaliciones. El análisis se realiza a nivel de la competencia en Bloques Políticos, a efectos de la comparación con resultados previos generados con encuestas electorales (Avenidaño & Sandoval, 2013). Los partidos existentes en el periodo pueden agruparse en cuatro opciones.

- Centro izquierda, conformada por cuatro partidos principales: Partido Demócrata Cristiano (PDC), Partido Socialista (PS), Partido por la Democracia (PPD) y Partido Radical (PR). Agrupada orgánicamente bajo la denominación de “Concertación de Partidos por la Democracia”.
- Derecha: integrada por dos partidos principales: Unión Demócrata Independiente (UDI) y Renovación Nacional (RN). Agrupada, a efectos electorales, en la denominada “Alianza por Chile”.
- Izquierda extraparlamentaria<sup>16</sup>: representada principalmente por las candidaturas adheridas al Partido Comunista (PC) y Partido Humanista (PH), además de otros movimientos de izquierda.
- Otros: candidaturas no adheridas a ninguna de las anteriores agrupaciones de partidos.

La lista, como código, cumplía meramente una función administrativa, vale decir, un grupo de partidos podía obtener, indistintamente, una letra u otra en un proceso y el siguiente. La agrupación listas-conglomerados, para cada elección se presenta en el Anexo 5.1 del presente documento.

**3.2.2.1.2. Padrón electoral comunal.** El mismo SERVEL dispone de información de Composición del Padrón de votantes desagregado a nivel comunal para cada año a partir de 1988. A diferencia de los resultados de votación, estos datos no se encuentran disponibles directamente en medios on-line. Por esta razón fueron requeridos al organismo recurriendo a la Ley 20.285 (conocida como “Ley de transparencia”)<sup>17</sup>, obteniéndose en definitiva por dicha vía. Estos datos informaban:

- Número de electores que integran el padrón comunal al 31 de diciembre de cada año.
- Total de cancelaciones del registro electoral comunal, en relación a la re-apertura del padrón electoral en el país, el año 1987.
- Número de Cancelaciones debidas a cambios de comuna de residencia, en comparación también al padrón comunal del año 1987.

---

<sup>16</sup>La vigencia del sistema binominal tuvo el efecto de concentrar la representación entre las dos principales coaliciones existentes durante este periodo, la “Concertación de Partidos por la Democracia” y la “Alianza por Chile” (anteriormente “Unión por Chile”), significando un obstáculo prácticamente infranqueable para las candidaturas no adheridas a aquellas.

<sup>17</sup>La mayor parte de estos fueron obtenidos mediante el traspaso físico de los libros de cada año, existentes en las dependencias del Organismo público.

Así, denotando como  $N_{(t)m}$  al número de electores en el padrón comunal el año de la elección  $t = 1993, \dots, 2009$  para la comuna  $m : 1, 2, \dots, M_{(t)}$ ;  $n_{(t)m}$  el número de votos obtenidos por las candidaturas adheridas al conjunto de las diferentes listas en la unidad  $m$ ;  $c_{8(t)m}$  al número de cancelaciones del padrón para cada comuna respecto al proceso electoral anterior, y  $nv_{(t+1)m}$ , se dispone de información como la siguiente:

$x_{11m}$	$x_{12m}$	...	$x_{1k,m}$	$x_{1k+1,m}$	$x_{1k+2,m}$	$\frac{d_{21m}}{N_{2m}}$
$x_{21m}$	$x_{22m}$	...	...	...	...	$\frac{d_{22m}}{N_{2m}}$
...	...	...	...	...	...	...
...	...	...	...	...	...	$\frac{d_{2k,m}}{N_{2m}}$
...	...	...	...	...	...	$\frac{N_{2m} - n_{2m}}{N_{2m}}$
...	...	...	...	...	...	$\frac{c_{8m}}{N_{2m}}$
$\frac{d_{11m}}{N_{2m}}$	$\frac{d_{12m}}{N_{2m}}$	...	$\frac{d_{1k,m}}{N_{2m}}$	$\frac{N_{1m} - n_{1m}}{N_{2m}}$	$\frac{nv_{2m}}{N_{2m}}$	

Donde las entradas  $x_{ijm}, m : 1, \dots, M; i : 1, \dots, k + 2, j : 1, \dots, k + 2$  son también desconocidas.

No obstante, cabe señalar que en las estadísticas comunales las sumas tanto de la fila  $k + 2$  como de la columna  $k + 2$  no corresponden a cifras exactas, a diferencia de las correspondientes a las demás filas y columnas, producto de limitaciones de los datos, respecto de lo cual se hace referencia a continuación.

### 3.2.2.2. Ajuste de Datos

Una vez delimitado el periodo de análisis, los resultados de votación fueron cargados en un marco de datos. Por su parte, para cada proceso electoral se obtuvo información del padrón comunal. Fusionando ambas fuentes de datos, y una vez chequeada la consistencia de la información, se obtuvo 4 marcos de datos, correspondientes a cada par de procesos electorales.

A partir de lo anterior se procedió a la generación de variables, correspondientes a: 1) votación obtenida por opción política en cada año; 2) abstencionismo; 3) número de ex-votantes; y 4) número de nuevos votantes en cada comuna.

La medición de las variables tipo 1 y 2 de las recién señaladas en cada unidad comunal es exacta. En cambio, las cifras de ex y nuevos votantes contiene error de medición, originado en:

- El número de cancelaciones totales es informado cada año respecto de las inscripciones en la reapertura del padrón electoral, el año 1987. Debido a falta de algunos registros comprendidos entre los años 1988 y 1992, las cancelaciones totales en 1993 contienen error, lo que repercute en impresiones a la hora de determinar las cifras para los años posteriores.
- El número de nuevos votantes en cada unidad municipal fue determinado utilizando la información de cancelaciones, partiendo de la identidad (2.3.8) dada en la sección 2.3.1.

Aún cuando las falencias anteriores no se hicieran presentes, la validez de estas mediciones se vería de todos modos afectada por el fenómeno de cambio de unidad comunal de residencia. A pesar de aquello, se trata de cifras minoritarias, por lo cual es posible asumir que, a efectos del problema estudiando, las

comunas sí forman una partición del electorado para cada par de elecciones analizadas.

Finalmente, para el análisis fue necesario excluir algunas comunas que experimentaron divisiones entre algunos pares de elecciones, debido a que la información de composición del padrón comunal presentaba imprecisiones en este aspecto. Para el par de procesos electorales 1993-1997 fueron incluidos un total de 329 municipios de 342 existentes en 1997, lo que congrega un 91.5 % del electorado de dicho año a nivel nacional. Fueron excluidas tanto las comunas que se dividían, como las nuevas. Se operó de la misma forma para el par de elecciones 2001-2005, siendo incorporadas al análisis 338 de 346 comunas existentes en 2005, lo que significó el 96.6 % del electorado nacional de dicho año. Para los pares 1997-2001 y 2005-2009 no fue necesario realizar este tipo de ajustes, al no presentarse divisiones de comunas, por tanto, se trabajó con información de la totalidad del electorado presente en tales años.

Este tipo de decisiones operacionales incidieron, asimismo, en el cálculo de las cotas determinísticas. Sin embargo, al excluir un volumen menor de información tratarse de cifras minoritarias en comparación al volumen total de información, puede asumirse que los resultados obtenidos no se ven afectados sustancialmente por la adopción de las mismas.

Mayor detalle del procedimiento empleado es incluido en el Anexo 5.1 del documento.

### 3.2.2.3. Uso de librerías de R: vottrans y eiPack

Luego, se procedió al uso de las librerías eiPack<sup>18</sup> para el modelo Multinomial-Dirichlet sin convariable, dada por la opción ei.MD.bayes, y vottrans<sup>19</sup> para el modelo de optimización. Inicialmente, ambas librerías fueron utilizadas recurriendo a sus valores predefinidos.

Para el modelo de optimización, la alternativa predefinida equivale a la minimización ponderada en norma euclídea sujeta a las restricciones de la definición de matriz de transición electoral, dada previamente en (2.2.1), es decir:

$$\sum_{m=1}^M \sum_{j=1}^l \left( n_m v_{(t+\Delta t)jm} - \sum_{i=1}^k (n_m v_{(t)jm} r_{ij}) \right)^2 = \sum_{m=1}^M n_m^2 \sum_{j=1}^l \left( v_{(t+\Delta t)jm} - \sum_{i=1}^k (v_{(t)jm} r_{ij}) \right)^2$$

Se decidió utilizar esta última opción debido a que los Municipios varían considerablemente en el país a nivel de su tamaño demográfico.

Por su lado, posterior al procesamiento inicial de la información, la opción ei.MD.bayes fue ajustada a valores de burnin=100.000 y thin=1.000 –contemplados éstos entre los evaluados por Klima et al. (2016: 157)–, a efectos de la convergencia de los algoritmos MCMC.

Ambas alternativas arrojaron como solución una estimación de la matriz aquí representada como  $\Phi$ , a partir de lo cual se obtuvo una estimación de las matrices  $\mathbf{X}$ , y a continuación, de los vectores de parámetros respectivos.

<sup>18</sup>Ver: <https://cran.r-project.org/web/packages/eiPack/eiPack.pdf>

<sup>19</sup><https://cran.r-project.org/web/packages/vottrans/vottrans.pdf>

El código completo en R para la lectura y ajuste de datos, aplicación de modelos y cálculo de cotas determinísticas se presenta en el Anexo 5.1 del documento, lo cual permite la completa reproducibilidad de los resultados obtenidos.

### 3.3. Resultados

#### 3.3.1. Volatilidad Electoral Inter-Bloques 1993-2009

El periodo estudiando se extiende entre los años 1993 y 2009. Cada cuatro años fueron realizadas elecciones de Cámara Baja a nivel nacional. Las elecciones tuvieron lugar, entonces, los años 1993, 1997, 2001, 2005 y 2009 bajo el marco del sistema electoral antes descrito.

El número de votantes que acudieron a sufragar en dichos procesos, junto al número de electores que formaban parte del Padrón electoral en cada uno de aquellos<sup>20</sup>, se presentan en la siguiente tabla:

**Tabla 1: Participación electoral Elecciones de diputados Chile 1993-2009**

Año	Concurrencia	Padrón	Participación
1993	6.738.859	8.085.493	83.3 %
1997	5.795.773	8.077.743	71.7 %
2001	6.144.003	8.075.446	76.1 %
2005	6.601.811	8.220.897	80.3 %
2009	6.482.856	8.282.186	78.3 %

Fuente: Elaboración propia con cifras de SERVEL

**Tabla 2: Resultados Elecciones de diputados Chile 1993-2009**

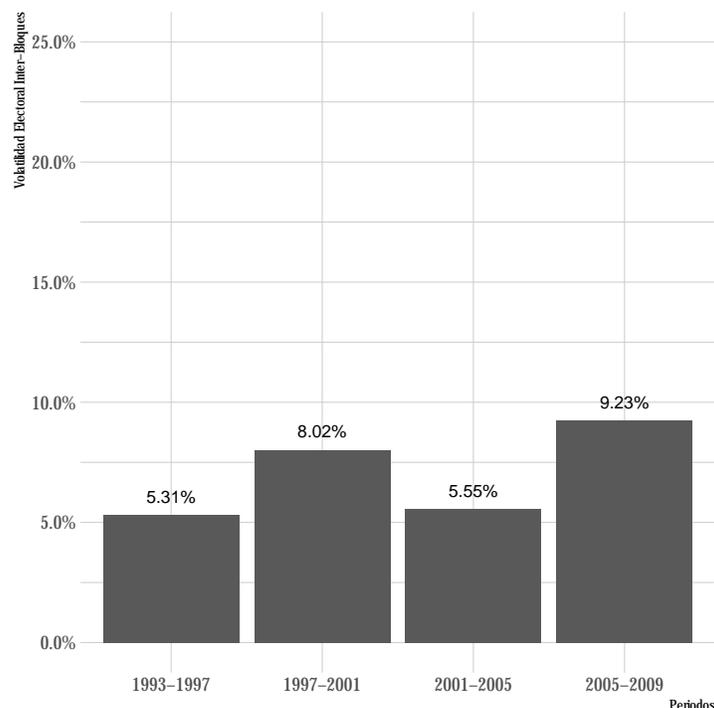
Agente	1993	1997	2001	2005	2009
Centro-izquierda	55.4 %	50.5 %	47.9 %	51.8 %	45.3 %
Derecha	36.7 %	36.3 %	44.3 %	38.7 %	44.3 %
Izq. extra-P	7.8 %	10.4 %	6.4 %	7.4 %	4.7 %
Otros	0.1 %	2.8 %	1.5 %	2.1 %	5.7 %
Total	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %

Fuente: Elaboración propia con cifras de SERVEL

Luego, en base a estos resultados, la volatilidad inter-bloques con la fórmula de Pedersen:

<sup>20</sup>Recuérdese que se trataba de un Sistema de Inscripción voluntaria y voto obligatorio.

**Figura 3: Volatilidad inter-bloques por Periodos**



El rango del Índice oscila entre 0 y 100 %, por tanto, se trata de cifras de volatilidad bastante moderadas. Para una parte de la investigación internacional, tal grado de regularidad de los resultados de votación respondía a la preponderancia del votante estable en la composición del electorado a lo largo del periodo. En gran medida sobre la base de esto último, el sistema chileno era catalogado en el periodo como uno de los más institucionalizados de la región.

### **3.3.1.1. La Paradoja de la Estabilidad**

Los bajos índices de volatilidad contrastaban con un progresivo distanciamiento de la ciudadanía hacia los agentes de representación política, lo que se manifestaba en un drástico descenso en la identificación partidaria, inscripción y participación electoral, confianza en las instituciones democráticas, etc. (Luna & Altman, 2011; Luna & Rossenblatt, 2012). Se trataba de una confluencia paradójica. Los cánones interpretativos difundidos por el enfoque PSI llevaban a concluir, a partir de las cifras de volatilidad electoral, la primacía del votante estable en la composición del electorado. La evidencia de los estudios de opinión pública y cifras de participación electoral, en cambio, planteaban serias interrogantes acerca de lo anterior<sup>21</sup>.

<sup>21</sup>En torno a este asunto, Luna (2008) se refería: “La presencia de un sistema de partidos débil en cuanto a su relación con la sociedad civil y sumamente fuerte en cuanto a reproducir lealtades electorales (tal como señala el índice de volatilidad), es aparentemente paradójica. En otras palabras, no parece usual que un sistema de partidos estable e institucionalizado a nivel electoral conviva con una ciudadanía crecientemente alejada de los partidos y sin simpatías partidarias cristalizadas” (2008: 88).

El estudio de Avendaño & Sandoval (2013; 2016) propuso una explicación a la confluencia entre estabilidad electoral y desafección política. Recurriendo a encuestas con preguntas retrospectivas de comportamiento de voto (realizadas a lo largo del periodo por el Centro de Estudios de la Realidad Contemporánea, CERC), la investigación concluyó la siguiente composición electorado respecto a las elecciones de diputados.

**Tabla 3: Composición del Electorado 1997-2009**

Perfiles	1997	2001	2005	2009
1. Estable	58.2 %	60.2 %	53.1 %	54.3 %
2. Transferido	6.5 %	8.1 %	9.9 %	8.8 %
3. Desmovilizado	12.5 %	10.4 %	13.8 %	8.3 %
4. Movilizado	7.2 %	8.2 %	17.1 %	19.8 %
5. Abstencionismo permanente	14.3 %	10.4 %	4.7 %	5.3 %
6. Nuevo concurrente	3.9 %	3.1 %	5.2 %	12.5 %
7. Nuevo abstencionista	1.3 %	0.3 %	1.4 %	6.9 %
Total	100.0 %	100.0 %	100.0 %	100.0 %

Fuente: Elaboración propia en base a Avendaño y Sandoval (2013: 154-157)

De acuerdo con esta evidencia, el votante estable a lo largo del periodo tuvo menor preponderancia que la sugerida en función de los cánones interpretativos del enfoque PSI y la estabilidad electoral a nivel agregado se producía como resultado de la auto-compensación del voto no estable, estimulada por la vigencia del sistema electoral binominal. No obstante, para el tratamiento de los datos fue preciso excluir respuestas con problemas de recordación (u otros), trabajando en definitiva con cerca del 75 % de cada muestra (2013: 152).

### 3.3.2. Aplicación de Métodos EI

En primer lugar se obtienen las cotas determinísticas identificadas en la sección 2.3 y resumidas en (2.3.20). Posteriormente se aplican los modelos a los datos obtenidos.

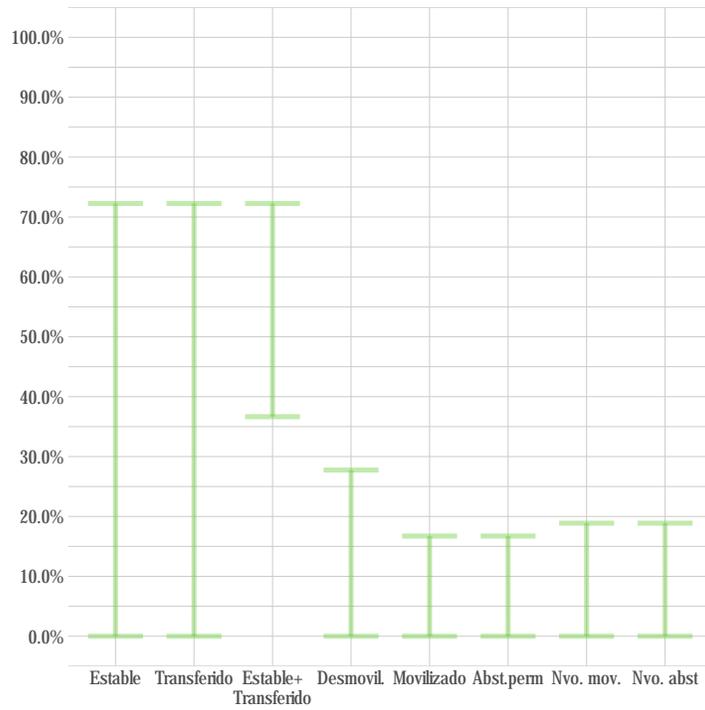
#### 3.3.2.1. Cálculo de Cotras determinísticas

Las siguientes tablas muestran las cotas determinísticas para la composición del electorado los años 1997, 2001, 2005 y 2009.

**Tabla 4: Cotas determinísticas por Parámetro 1997**

Perfiles	Cota inferior	Cota superior
1. Estable	0.0 %	72.3 %
2. Transferido	0.0 %	72.3 %
Estable + Transferido	36.7 %	72.3 %
3. Desmovilizado	0.0 %	27.7 %
4. Movilizado	0.0 %	16.7 %
5. Abstencionismo permanente	0.0 %	16.7 %
6. Nuevo concurrente	0.0 %	18.9 %
7. Nuevo abstencionista	0.0 %	18.9 %

**Figura 4: Rangos por parámetros 1997**



**Tabla 5: Cotas determinísticas por Parámetros 2001**

Tipo de votante	Cota inferior	Cota superior
1. Estable	0.0 %	71.8 %
2. Transferido	0.0 %	71.8 %
Estable + Transferido	38.0 %	71.8 %
3. Desmovilizado	0.0 %	23.9 %
4. Movilizado	0.0 %	27.9 %
5. Abstencionismo permanente	0.0 %	23.9 %
6. Nuevo concurrente	0.0 %	10.1 %
7. Nuevo abstencionista	0.0 %	10.1 %

**Tabla 6: Cotas determinísticas por Parámetros 2005**

Tipo de votante	Cota inferior	Cota superior
1. Estable	0.0 %	75.4 %
2. Transferido	0.0 %	75.4 %
Estable + Transferido	46.0 %	75.4 %
3. Desmovilizado	0.0 %	19.0 %
4. Movilizado	0.0 %	23.6 %
5. Abstencionismo permanente	0.0 %	19.0 %
6. Nuevo concurrente	0.0 %	11.4 %
7. Nuevo abstencionista	0.0 %	11.4 %

**Tabla 7: Cotas determinísticas por Parámetros 2009**

Tipo de votante	Cota inferior	Cota superior
1. Estable	0.0 %	79.7 %
2. Transferido	0.0 %	79.7 %
Estable + Transferido	50.0 %	79.7 %
3. Desmovilizado	0.0 %	20.2 %
4. Movilizado	0.0 %	18.7 %
5. Abstencionismo permanente	0.0 %	18.7 %
6. Nuevo concurrente	0.0 %	11.1 %
7. Nuevo abstencionista	0.0 %	11.1 %

### 3.3.2.2. Aplicación de Modelos

Procedemos a continuación a aplicar los modelos EI descritos en la sección 2.2 con los datos obtenidos.

**3.3.2.2.1. Periodo 1993-1997.** Los resultados de ambos métodos de estimación resultan coincidentes con los informados por los antecedentes previos. Las principales diferencias entre el modelo de optimización y la encuesta ocurren para el votante transferido (11 % versus 6.5 %), y para el movilizado (0 versus 7 %); en

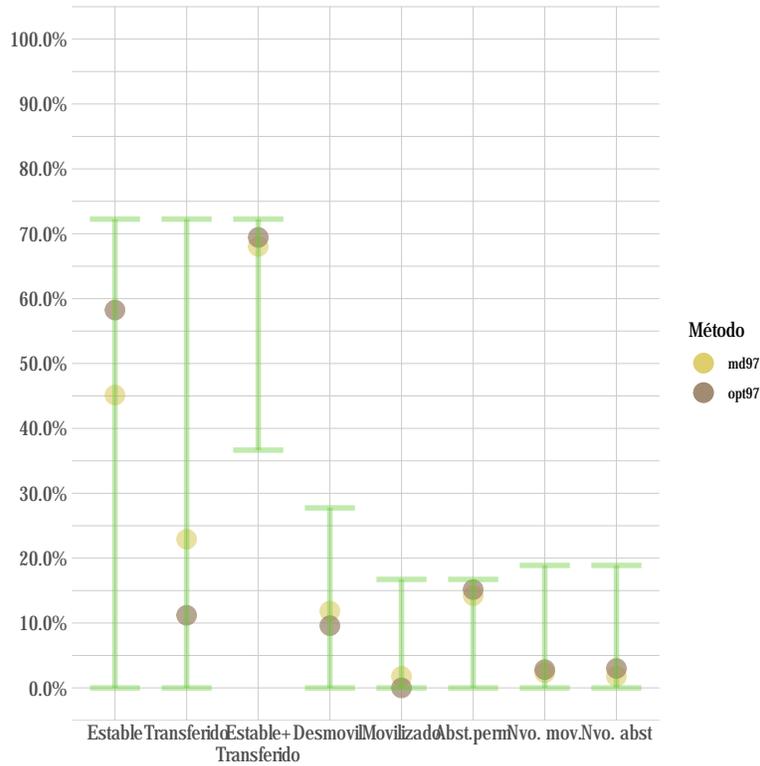
las restantes categorías, las estimaciones ofrecidas por ambas alternativas resultan prácticamente idénticas. Por su parte, el modelo bayesiano difiere de las cifras de las encuestas electorales para la estimación de votantes estables y transferidos.

**Tabla 8: Composición del Electorado estimada con Métodos EI 1997**

Tipo de votante	MD-Bayes	Optimización	Encuesta
1. Estable	45.1 %	58.2 %	58.2 %
2. Transferido	22.9 %	11.2 %	6.5 %
3. Desmovilizado	11.8 %	9.6 %	12.5 %
4. Movilizado	1.8 %	0.0 %	7.2 %
5. Abstencionismo permanente	14.2 %	15.1 %	14.3 %
6. Nuevo concurrente	2.3 %	2.8 %	3.9 %
7. Nuevo abstencionista	1.8 %	3.0 %	1.3 %
Total	100.0 %	100 %	100.0 %

También, las estimaciones ofrecidas por los dos métodos de inferencia ecológica se encuentran dentro de los rangos de la acotación del espacio paramétrico correspondiente al par de elecciones parlamentarias 1993-1997. En el gráfico que se presenta a continuación los puntos de color púrpura corresponden a las cifras estimadas por el modelo de optimización, y las de color naranja a las ofrecidas por el modelo MD-Bayes.

Figura 5: Comparación estimaciones v/s Rangos 1997



**3.3.2.2.2. Periodos 1997-2001, 2001-2005 y 2005-2009** Si bien en menor intensidad que en el periodo anterior, los resultados siguen mostrando coincidencia entre los resultados del modelo de optimización y el obtenido previamente con encuestas electorales. Por su lado, las estimaciones del modelo bayesiano siguen diferenciándose de las otras dos, principalmente en lo referido al peso del votante estable y el transferido, esto a excepción del año 2005.

Para estos tres pares de procesos electorales, las estimaciones ofrecidas por ambos modelos de inferencia ecológica vuelven a ubicarse en los rangos de la acotación del espacio paramétrico. Entre el modelo de optimización y la información de encuestas la principal congruencia se encuentra en la estimación del porcentaje de votantes estables. Ambos métodos coinciden en estimar a este perfil en el rango 50 % a 65 % a lo largo de todo el periodo analizado. Para el año 2009, la estimación del votante estable vuelve a ser prácticamente idéntica.

Por su parte, sucesivas aplicaciones del modelo bayesiano muestran, para los diferentes pares de procesos de votación, resultados consistentes, lo que sugiere la convergencia de los algoritmos MCMC a efectos del muestreo de la distribución posterior. Con los parámetros establecidos para estos modelos (burnin=100.000 y thin=1.000) se observa que las cadenas de Markov se comportan como un proceso de ruido blanco para

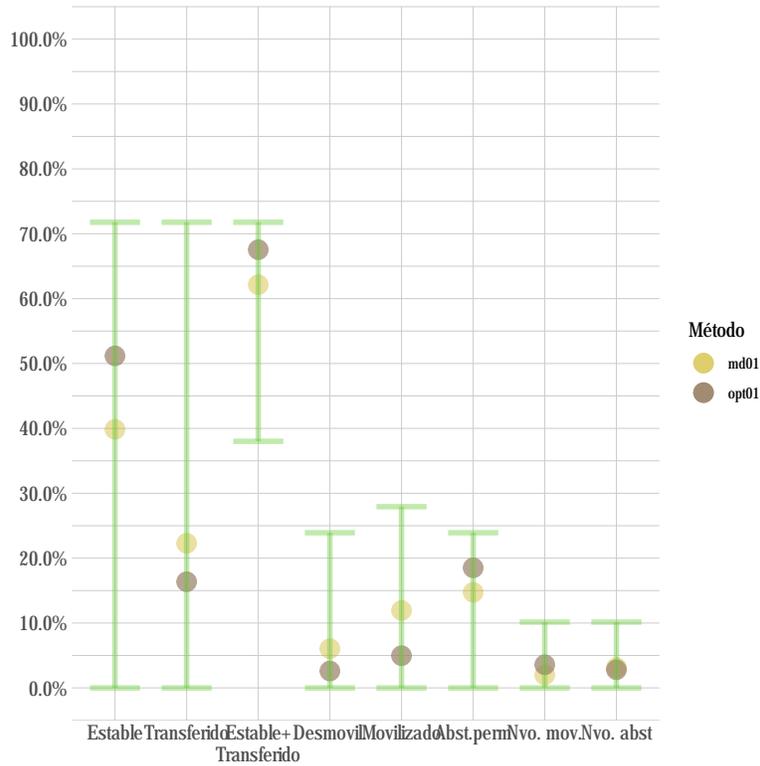
comunas donde el padrón comunal se compone por 10 mil o menos votantes, no así para aquellas de mayor tamaño. Estos resultados se presentan en el Anexo 5.1 del documento. Dada la estructura jerárquica del modelo, la evaluación de la convergencia debiera realizarse para cada parámetro y cada par de procesos electorales<sup>22</sup>.

**Tabla 9: Composición del Electorado estimada con Métodos EI, 2001**

Tipo de votante	MD-Bayes	Optimización	Encuesta
1. Estable	39.8 %	51.2 %	60.2 %
2. Transferido	22.3 %	16.4 %	8.1 %
3. Desmovilizado	6.0 %	2.6 %	10.4 %
4. Movilizado	11.9 %	5.0 %	8.2 %
5. Abstencionismo permanente	14.7 %	18.5 %	10.4 %
6. Nuevo concurrente	2.0 %	3.6 %	3.1 %
7. Nuevo abstencionista	3.1 %	2.8 %	0.3 %
Total	100.0 %	100.0 %	100.0 %

<sup>22</sup>Alrededor de 12.140 parámetros para cada par de procesos electorales.

**Figura 6: Comparación estimaciones v/s Rangos 2001**



**Tabla 10: Composición del Electorado estimada con Métodos EI, 2005**

Tipo de votante	MD-Bayes	Optimización	Encuesta
1. Estable	54.9 %	62.7 %	53.1 %
2. Transferido	16.4 %	10.6 %	9.9 %
3. Desmovilizado	2.1 %	0.5 %	13.8 %
4. Movilizado	4.8 %	1.0 %	17.1 %
5. Abstencionismo permanente	15.3 %	16.8 %	4.7 %
6. Nuevo concurrente	4.9 %	6.7 %	5.2 %
7. Nuevo abstencionista	1.6 %	1.6 %	1.4 %
Total	100.0 %	100.0 %	100.0 %

Figura 7: Comparación estimaciones v/s Rangos 2005

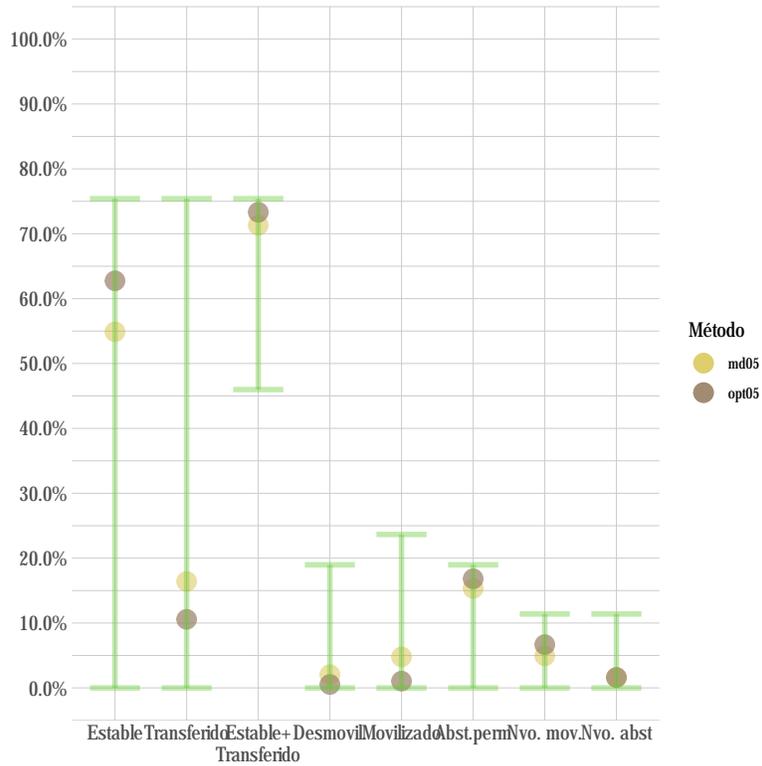
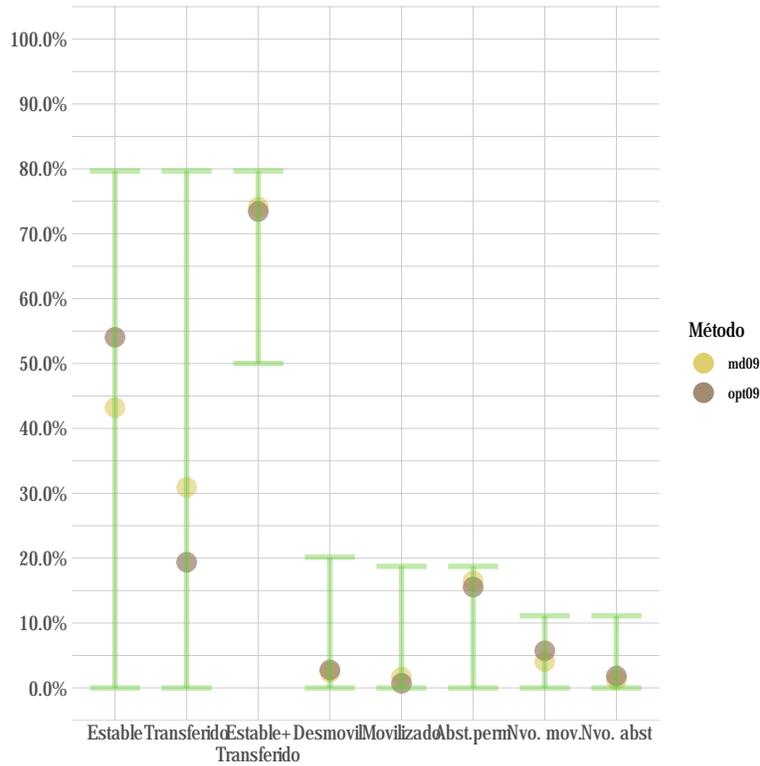


Tabla 11: Composición del Electorado estimada con Métodos EI, 2009

Tipo de votante	MD-Bayes	Optimización	Encuesta
1. Estable	43.2 %	54.0 %	54.3 %
2. Transferido	30.9 %	19.4 %	8.8 %
3. Desmovilizado	2.4 %	2.8 %	8.3 %
4. Movilizado	1.7 %	0.7 %	19.8 %
5. Abstencionismo permanente	16.5 %	15.5 %	5.3 %
6. Nuevo concurrente	4.1 %	5.7 %	12.5 %
7. Nuevo abstencionista	1.2 %	1.8 %	6.9 %
Total	100.0 %	100.0 %	100.0 %

Figura 8: Comparación estimaciones v/s Rangos 2009



La principal diferencia entre los resultados del modelo de optimización y el antecedente de la encuesta viene dado por el peso atribuido al votante movilizado. El modelo de optimización lo ubica consistentemente cerca de 0, mientras, según la encuesta, se habría incrementado sostenidamente a lo largo del periodo analizado.

Pese a lo anterior, puede considerarse que el método de optimización empleado resulta, en lo sustancial, coincidente con los antecedentes comparativos previos. Esta similitud puede considerarse como un hecho poco probable, a no ser que ambos métodos se encuentren midiendo el mismo fenómeno.

### 3.3.3. Comentarios

Obtener una estimación de la proporción de votantes en las siete categorías enumeradas en la sección 2.1 mediante el uso de métodos EI equivale a seleccionar una matriz  $\hat{\mathbf{X}}_i$  dentro del conjunto que satisfacen las propiedades de la matriz representada en este trabajo como  $\mathbf{X}$ . Los métodos empleados en este estudio ofrecen una alternativa para aproximar una matriz con estas propiedades. Sin embargo, tal selección no garantiza el cumplimiento de todas estas condiciones. Esto se debe a que:

- El fenómeno de cambio de unidad de residencia conduce a que las mediciones del número de ele-

mentos en los subconjuntos  $A_6$ ,  $A_7$  y  $A_8$  no sea exacta.

- Eventual exclusión de unidades de intermedias que se dividen o emergen entre pares de procesos consecutivos.
- Falta de especificación de algunas condiciones en la construcción de modelos.

Como se muestra en el anexo 5.1 del documento, la solución obtenida por ambos métodos no satisface la condición  $x_{k+2,k+2} = 0$ , al no figurar en las especificaciones en la construcción de los mismos. A pesar de esto último, tanto el modelo de optimización, como los resultados de la aplicación del modelo jerárquico, ofrecieron estimaciones consistentes con la acotación del espacio paramétrico del problema.

En el segundo tipo de modelo, sin embargo, el algoritmo MCMC no logró exhibir convergencia para comunas de mayor tamaño demográfico, pese a incrementar drásticamente los parámetros predefinidos por el mismo (burnin=1.000 y thin=1).

Por su parte, los resultados de la aplicación del modelo de optimización mostraron, en lo estructural, el mismo tipo de estimación que los antecedentes previos obtenidos con encuestas electorales. En particular, la mayor coincidencia se produce para el par de elecciones de Diputados realizadas entre 1993 y 1997<sup>23</sup>. Esta coincidencia, en particular en las cifras de votantes estables, constituye evidencia a favor de la posibilidad de utilizar métodos de Inferencia ecológica para obtener estimaciones de este tipo de fenómenos.

No obstante, y como se hizo mención en la sección 2.2.3.1, una de las principales limitaciones asociadas al uso de métodos Stayers-movers consiste en la falta de una estimación por intervalo, al no tener un fundamento explícito en un componente distribucional, como sí lo tienen, en cambio, las aproximaciones bayesianas, bajo la forma de intervalos plausibles de Bayes. Vincular este tipo de métodos a una base distribucional constituye un tema en sí mismo, cuyo abordaje excede largamente a los objetivos pretendidos con este proyecto. En su lugar, se propone como punto de partida para futuros desarrollos en esta materia el uso de una distribución Hipergeométrica multivariada truncada. Esta justificación es incluida en el Anexo 5.2 del documento.

### 3.4. Propuestas de mejora para el estudio de la Institucionalización partidista

Como fue señalado en la sección introductoria, para el enfoque de la institucionalización partidista de Mainwaring & Scully (1995) –ampliamente difundido para el estudio de los sistemas políticos en América Latina– la existencia de sistemas electoralmente estables, pero crecientemente desarraigados de la ciudadanía, figuran como casos paradójales. Esta connotación se origina en la adopción del supuesto de que los cuatro componentes que integran el constructo se encuentran estrechamente relacionados entre sí. De

---

<sup>23</sup>Para los restantes periodos, las cifras estimadas son algo menos coincidentes. En esto último puede incidir el comportamiento de los mismos electores frente a las encuestas realizadas con posterioridad al año 1997. El incremento de la desafección política acontecida en el país analizado puede haber conducido a una menor predisposición a participar de estas mediciones, sesgando entonces las estimaciones; o bien, a un mayor ocultamiento del comportamiento de voto real. En cualquier caso, esto último no constituye un asunto que pueda ser zanjado con el tipo de evidencia aquí provista.

ser éste el caso, la medición de uno sólo de aquellos vendría a posibilitar la caracterización del fenómeno en su conjunto. La volatilidad electoral constituye la medida más recurrentemente empleada en el marco de este tipo de estudios. Su utilización en este contexto ha permitido, además, sostener el enfoque de investigación comparativa pretendido por la propuesta inicial.

Sin embargo, la interpretación de la volatilidad electoral en el contexto más amplio de la investigación politológica internacional adolece de indefiniciones, las que repercuten luego en los marcos interpretativos que recurren a ésta en calidad de indicador del grado de inestabilidad del comportamiento del electorado. Con el desarrollo incluido en este proyecto se ha abordado algunas de estas indefiniciones, a partir de lo cual es posible plantear mejoras al abordaje empírico del fenómeno de la institucionalización partidista. Estas propuestas buscan conciliar el componente sustantivo del fenómeno de la institucionalización, con la posibilidad de mantener un enfoque de investigación comparativa en torno a esta materia.

En primer término, cabe distinguir claramente a la volatilidad electoral agregada de la proporción de votantes no estables que componen el electorado tras dos votaciones consecutivas. Si bien en la práctica ambas medidas pueden encontrarse correlacionadas (Bartolini & Mair, 1990), la primera no debe utilizarse para estimar la magnitud de la segunda, pues su empleo con este propósito conlleva el riesgo de incurrir en una falacia ecológica, en particular cuando un sistema exhibe estabilidad en el plano de los resultados de votación. En lugar de aquello, se propone restringir el uso de la volatilidad agregada exclusivamente a la medición del grado de variación de los patrones de votación agregados, y en lo referido al comportamiento del electorado, recurrir a métodos de estimación con datos agregados.

Como ha señalado Torcal (2015), el grado de variabilidad de los patrones de votación en el tiempo es un determinante del nivel de institucionalización de un sistema democrático, pues, entre otros aspectos, incide en el cálculo estratégico de los agentes de representación política al momento de estructurar la oferta de representación: en contextos altamente volátiles las elites políticas disponen de menor margen de acción para fundamentar la representación en un componente programático y para mantener un cuerpo organizacional perduradero. Por ende, la estabilidad electoral como atributo de la institucionalización tiene un valor intrínseco, con independencia de su relación con el comportamiento del electorado.

Pese a lo anterior, la estabilidad electoral por sí sola no resulta suficiente para catalogar a un sistema como institucionalizado. De hacerse presente, es importante también que tal estabilidad provenga de una vinculación efectiva entre el electorado y los agentes de representación política. Como indican los antecedentes previos referidos en la sección introductoria de este informe, un escenario de volatilidad nula puede ocurrir aun cuando la totalidad del electorado modifique sus preferencias de votación; o, como se desprende del desarrollo incluido en este proyecto, desde otro tipo de escenarios, como por ejemplo la preeminencia de la desafección política bajo la forma de abstencionismo. El indicador que puede servir para diagnosticar este componente es el porcentaje de votantes estables, el cual puede ser estimado recurriendo a métodos de Inferencia ecológica como los que fueron utilizados en este proyecto. Así, tanto la volatilidad agregada como el grado de estabilidad del voto se refieren a aspectos sustantivos del fenómeno de la institucionalización, y que pueden ser medidos con datos agregados en diferentes sistemas de partidos.

Aún con esto último, la estabilidad, tanto de los patrones de votación agregados, como del comportamiento del electorado, pueden tomar lugar en contextos de marcada disimilitud al momento de juzgar

el grado institucionalización de un sistema concreto. Un caso hipotético mencionado por Luna (2014) es la primacía de un vínculo clientelar hacia liderazgos de signo populista. Así, es importante también analizar la naturaleza del tipo de vinculación entre el electorado y el sistema de representación política. A efectos del abordaje empírico del problema este tipo de información puede ser obtenida desde estudios de opinión pública, o a partir de información como la ofrecida por el proyecto V-Dem<sup>24</sup> (Bizarro et al., 2017).

Con todo, diferenciar claramente a la volatilidad electoral del porcentaje de votantes no estables, tanto desde su rol en la configuración de un escenario de institucionalización partidista, como asimismo en lo referido a la vía empírica para sostener su medición comparativa, permite sortear la connotación paradójica atribuida a sistemas electoralmente estables, pero desarraigados de la ciudadanía. Bajo esta estrategia, dicha configuración resulta un escenario particular, de entre otros posibles. Si ambas medidas son complementadas con información relativa al tipo de vinculación entre electorado y sistema de representación, se configura la triada estabilidad agregada/estabilidad del comportamiento de voto/vinculación programática, susceptible de ser diagnosticada comparativamente en diferentes sistemas electorales.

Finalmente, para la aplicabilidad de la propuesta aquí desarrollada en diferentes sistemas democráticos, un problema a enfrentar es la posible discontinuidad de los agentes de representación política. En el caso que fue aquí analizado, las mismas coaliciones de partidos mantuvieron presencia a lo largo de todo el periodo.

Se pueden presentar dos tipos de problemas: 1) Surgimiento de agentes de representación nuevos y/o Disolución de otros preexistentes; 2) Fusiones y/o división de partidos en  $t + 1$ . El trabajo de Casal Bértoa et al. (2017) enumera una serie de alternativas que han sido utilizadas por la investigación comparativa internacional para abordar este tipo de situaciones.

Para el primer problema, una opción utilizada hoy en día consiste, simplemente, en establecer en un valor igual a 0 la votación precedente (Agentes nuevos), o la sucesiva (Disolución). Esta estrategia es empleada actualmente para diferenciar a la “volatilidad tipo A” que se identifica como aquella causada por la entrada o salida de agentes de representación, de la volatilidad tipo B, que se produce entre los agentes presentes en ambos procesos electorales (Powell & Tucker, 2013).

Supongamos que un total de  $k$  agentes de representación operaron en  $t$  o en  $t + 1$ . De éstos,  $o < k$  no perduran en el segundo evento, mientras que otros  $n < k$  emergen en dicha instancia. Al hacer 0 la votación posterior, o precedente, la ecuación del Índice de Volatilidad puede expresarse como:

$$V = \frac{1}{2} \sum_{j=1}^k |v_{(t+1),j} - v_{(t),j}| = \frac{1}{2} \sum_{j=1}^o v_{(t)j} + \frac{1}{2} \sum_{j=o+1}^{o+n} v_{(t+1)j} + \frac{1}{2} \sum_{j=o+n+1}^k |v_{(t+1)j} - v_{(t)j}|$$

Definiendo la “volatilidad tipo A”:  $V_A = \frac{1}{2} \sum_{j=1}^o v_{(t)j} + \frac{1}{2} \sum_{j=o+1}^{o+n} v_{(t+1)j}$ , y como volatilidad “tipo B”  $V_B = \frac{1}{2} \sum_{j=o+n+1}^k |v_{(t+1)j} - v_{(t)j}|$ , resulta  $V = V_A + V_B$ , que es la propuesta de Powell & Tucker (2013).

Entonces, definir como 0 la votación en el proceso de no presencia por parte de un agente de representación, constituye una alternativa razonable, y de hecho actualmente utilizada, para afrontar este tipo de

---

<sup>24</sup><https://www.v-dem.net/es/>

discontinuidad.

En el caso de divisiones o fusiones, supongamos en primer lugar que un agente de representación B existente en un proceso  $t$  se divide en otros dos C y D en  $t+1$ . En esta situación, una alternativa razonable es considerar en conjunto a C y D como sucesor de B en  $t+1$ , y luego, de perdurar como tales, considerarlos por separado en  $t+2$ . Se puede emplear una lógica similar en el caso de que dos opciones C y D diferentes en  $t$  se fusionen en otra opción B en  $t+1$ ; o de presentarse cambios de denominación de alguno(s) de los agentes de representación.

Sea cual sea el caso, el abordaje del segundo tipo de discontinuidad requiere explicitar las decisiones operacionales adoptadas en torno a la materia. Recurriendo a este tipo de decisiones operacionales, la estrategia desarrollada a lo largo de este proyecto resulta compatible también con el problema de la discontinuidad de los agentes de representación política.

Es decir, para cada operacionalización de los agentes de representación presentes entre dos procesos electorales consecutivos del mismo tipo que solicitan a los votantes emitir una única preferencia, tal que se puedan identificar los mismos  $k$  agentes de representación concentrando la totalidad de la oferta electoral en ambos procesos, existirá una única matriz  $\mathbf{X}$  que determina el grado de volatilidad electoral observado, y el vector de parámetros que contiene el peso relativo de los siete perfiles de votantes enumerados en este trabajo.

## Capítulo 4

# Conclusiones

Con lo realizado se logró relacionar los métodos estadísticos para la estimación de matrices de transición electoral con el estudio de la composición del electorado tras dos eventos de votación consecutivos, proveyendo evidencia a favor de la posibilidad de efectuar el diagnóstico comparativo de esta composición por medio de datos agregados.

Sobre esta base se propuso como mejoras al abordaje empírico de la Institucionalización partidista: 1) Restringir el uso de la volatilidad electoral a la medición del grado de regularidad de los patrones de votación agregados; y 2) Medir la estabilidad del comportamiento de los votantes por medio de métodos de estimación con datos agregados; todo lo cual permitió dar cumplimiento al noveno objetivo específico del proyecto. Con esto se logra sortear el carácter paradójico hasta ahora atribuido a sistemas electoralmente estables, pero desarraigados de la ciudadanía, configurándose bajo este marco como un escenario particular, entre otros posibles.

Este relacionamiento fue efectuado mediante un desarrollo algebraico de carácter inédito. En la sección 2.1.1 se formalizó el espacio paramétrico del problema, dando cumplimiento al tercer objetivo específico del proyecto. En el apartado 2.1.2 se justificó el uso de la denominación “parámetros de la volatilidad electoral” para denotar a la proporción de votantes en cada una de las categorías enumeradas en el proyecto, lo que condujo a una formulación matricial del Índice de Pedersen (objetivo específico N°2). En tanto, en la sección 2.3.2 se desarrolló una acotación del espacio paramétrico (objetivo específico específico N°5).

La revisión de antecedentes previos permitió identificar dos tipos de métodos de Inferencia ecológica con aplicabilidad a problemas de transición electoral: 1) Modelos stayers-movers, los cuales recurren a algoritmos de optimización de discrepancias sujetas a las restricciones dadas por la definición de matriz de transición electoral; y 2) Modelos con supuestos distribucionales, los cuales recurren a estructuras bayesianas jerárquicas y algoritmos MCMC. En la sección 2.2 se procedió a describir estos métodos y su relación con Cadenas de Markov. Así, se dio cumplimiento al cuarto objetivo específico del proyecto.

Sobre la base de esto último, se propuso como estimadores de la matriz aquí representada como  $\mathbf{X}$ :  $\widehat{\mathbf{X}}_o$  (la estimación obtenida con el método de optimización), y  $\widehat{\mathbf{X}}_h$  (modelo bayesiano jerárquico).

La propuesta fue aplicada al caso de la competencia entre coaliciones políticas en Chile para las elecciones de cámara baja realizadas entre los años 1993 y 2009. Esto permitió dar cumplimiento al octavo objetivo específico del proyecto. Los resultados obtenidos con la aplicación de ambos modelos resultaron consistentes con evidencia previa procedente de encuestas electorales. En particular, el método de optimización exhibió la mayor concordancia con los antecedentes previos. En el caso del modelo bayesiano, junto con ofrecer estimaciones más distantes a las otras dos, los algoritmos MCMC no lograron exhibir convergencia para todas las unidades de medición (comunas), pese a incrementar drásticamente los parámetros predefinidos. Esto llevó a atribuir, para el caso aquí analizado, un mejor desempeño al uso de métodos de optimización.

Como principal ventaja asociada al uso de este tipo de métodos en lo relacionado con el problema de estudio aquí abordado figura la libre disponibilidad de datos en la mayor parte de los países del mundo, lo que posibilita sostener un enfoque de investigación comparativa en torno a esta materia. En contrapartida, como desventajas se cuentan:

- Un supuesto básico es que las unidades forman una partición de la población, lo cual es transgredido por el fenómeno del cambio de residencia.
- Los modelos bayesianos jerárquicos, además de descansar en supuestos de difícil verificación, pueden llegar a requerir una considerable capacidad de hardware para ser implementados, y aun con ello, no lograr convergencia en los algoritmos MCMC.
- Los métodos de optimización no ofrecen una estimación por intervalo, al no tener un fundamento explícito en un componente distribucional.

Con respecto a último punto, una alternativa que puede ser abordada en futuros desarrollos consiste en asumir a la solución del método de optimización como una realización de una distribución hipergeométrica multivariada truncada. Este planteamiento, aunque de carácter general, se vincula con el cumplimiento de los objetivos específicos N°1: Formalizar el espacio muestral subyacente a la composición del electorado tras dos eventos electorarios consecutivos; y N°6: Identificar alternativas para obtener valores esperados y varianzas de las variables aleatorias involucradas en el problema. El único objetivo específico que no fue cumplido fue el N°5: Incorporar el anterior desarrollo en las distribuciones previas de los modelos de Inferencia Ecológica jerárquicos para el caso de matrices  $R \times C$  y evaluar los resultados obtenidos por medio de un ejercicio de simulación con el software R, al escapar en definitiva al foco de los resultados pretendidos con este estudio. No obstante, puede considerarse, al igual que los Objetivos 1 y 6, como una alternativa a explorar en futuros desarrollos teóricos.

El cumplimiento de la mayor parte de los objetivos específicos del proyecto permitió dar cumplimiento al Objetivo geneneral del proyecto. Además, la propuesta aquí desarrollada resulta compatible con el problema de la discontinuidad de los agentes de representación política por emergencia/disolución, fusiones/divisiones o cambios de rótulo. Se provee, así, de un marco de evidencia empírica que puede servir de base en el ámbito politológico a un estudio comparativo de mayor escala, y en lo referido al ámbito estadístico, a nuevas formalizaciones sobre el tema de la estimación con datos agregados.

# Capítulo 5

## Anexos

### 5.1. Trabajo con los Datos

Requiere el empleo de las librerías del software R: `openxlsx`, `plyr`, `dplyr`, `car`, `tidyr`, `reshape`, `quadprog`, `scales`, `rlist`, `ggplot2`, `bayesplot`, `coda`, `vottrans` y `eiPack`.

#### Lectura de Datos

Los datos de resultados electorales se encuentran disponibles en formato Excel en el sitio web del organismo y pueden ser descargados desde <https://www.servel.cl/elecciones-de-diputados-1989-al-2013-por-circunscripcion-electoral/>. Una vez guardado, se puede configurar en R el directorio de trabajo con la función `setwd()`, para luego leer los datos con la librería `openxlsx` y guardados en un data frame denominado “reselect0”.

```
library(openxlsx)
reselect0<-read.xlsx("resultados_elecciones_diputados.xlsx",sheet=1)
```

Por su parte, los datos obtenidos por medio de procedimientos de acceso a la información pública fueron traspasados a un archivo Excel “Datos Padron comunal DEF.xlsx”, leídos con la misma librería, y guardados en un data frame con nombre `regcomunal`.

```
regcomunal<-read.xlsx("Datos Padron comunal DEF.xlsx",sheet=1)
```

Un ejemplo de cómo estos datos se encontraban disponibles en los registros del Servicio Electoral se presenta en la siguiente imagen<sup>25</sup>.

---

<sup>25</sup>Comuna: Hualaihué, Año: 2009. Fuente: SERVEL, 2009: pp. 373

Figura 9: Ejemplo registro oficial de información

REPÚBLICA DE CHILE  
SERVICIO ELECTORAL

COMUNA HUALAIHUÉ

INSCRIPCIÓN ELECTORAL VIGENTE AL 31 DE DICIEMBRE 2009

REGION : 10 DE LOS LAGOS  
PROVINCIA : PALENA  
CIRC. SENATORIAL : DECIMA SEPTIMA  
DISTRITO : 58  
JUNTA ELECTORAL : PALENA  
JUNTA INSCRIPTORA : HUALAIHUÉ  
SUBLEJA : SULECHA  
CANTON : HUALAIHUÉ PUESTO

	REGIONAL	NACIONAL
VARONES	3.015	1,54
MUJERES	2.548	1,24
TOTAL INSCRITOS	5.563	1,39

A) DE INSCRITOS VIGENTES POR GRUPOS ETAREOS

	18-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-++	TOTAL
VARONES	1.398	4.444	7.664	9.654	13.734	16.654	11.644	8.964	7.304	5.344	4.284	3.814	2.494	2.754	100.004
MUJERES	1.458	6.324	10.014	10.284	13.274	13.894	11.074	7.654	5.974	4.914	3.654	2.754	3.024	2.894	100.004
TOTAL	2.856	10.768	17.678	19.938	27.028	30.548	22.718	16.558	13.278	10.258	7.938	6.568	5.518	5.648	200.008

B) DE INSCRITOS EXTRANJEROS

	VARONES	MUJERES
TOTAL	57	3

C) DE INSCRITOS ANALFABETOS Y NO VIDENTES

	ANALFABETOS	NO VIDENTES
VARONES	134	1
MUJERES	211	4

D) DE CANCELACIONES DE INSCRIPCIONES SEGUN CAUSAL LEGAL (25 FEBRERO 1997 - 31 DICIEMBRE 2009)

	CAMBIO DOMICILIO	DEFUNCION	INHABIL. CONSTIT.	OTRAS CAUSAS	TOTAL CANCELACIONES
VARONES	1.455	412	88	242	2.217
MUJERES	1.044	231	4	217	1.496
TOTAL	2.499	643	92	459	3.713

E) DE INSCRITOS PROVENIENTES DE OTRAS COMUNAS

REGION	VARON	MUJER	TOTAL
XV	4	2	6
I	2	3	5
II	1	2	3
XII	1	2	3
IV	6	3	9
V	27	33	60
R.H.	104	117	221
VI	5	2	7
VII	8	8	16
VIII	61	38	99
IX	189	44	233
XIV	59	41	100
X	1.961	990	2.951
XI	42	18	60
XIII	44	22	66
TOTAL	1.914	1.324	3.238

F) DE EMIGRACION ELECTORAL A OTRAS COMUNAS

REGION	VARON	MUJER	TOTAL
XV	1	2	3
I	2	1	3
II	3	5	8
III	1	0	1
IV	5	2	7
V	26	19	45
R.M.	63	47	110
VI	7	2	9
VII	7	7	14
VIII	38	19	57
IX	109	19	128
XIV	20	11	31
X	1.135	881	1.986
XI	44	8	52
XII	19	24	43
TOTAL	1.460	1.047	2.507

## Ajuste de Datos

### Resultados Electorales

**Delimitación del Periodo.** Se procedió a filtrar los datos para el periodo 1993-2009.

```
library(plyr)
library(dplyr)
reselect0<-reselect0 %>% filter(Año.de.Elección>=1993, Año.de.Elección<=2009)
```

**Unificación categoría Nulo/Blanco.** La variable Lista deja como valor perdido para votos nulos y blancos. Estos valores fueron llevados a una categoría NB

```
library(car)
reselect0$Lista<-as.factor(Recode(reselect0$Lista, "' '='NB'; ' ' '='NB'; NA='NB'"))
```

**Rectificación de Datos.** Los datos presentaban algunos errores desde su origen, los que fueron rectificados en el data frame respectivo.

```
id1_93<-as.numeric(rownames(subset.data.frame(reselect0,
xor(Nombre.Circunscripción.Electoral=="CHIGUAYANTE",
```

```

Nombre.Circunscripción.Electoral=="SAN PEDRO DE LA PAZ") & Año.de.Elección==1993)))
id2_93<-as.numeric(rownames(subset.data.frame(reselect0,
Nombre.Circunscripción.Electoral=="CONCON" & Año.de.Elección==1993)))
id3_93<-as.numeric(rownames(subset.data.frame(reselect0,
Nombre.Circunscripción.Electoral=="SAN RAFAEL" & Año.de.Elección==1993)))
id4_93<-as.numeric(rownames(subset.data.frame(reselect0,
Nombre.Circunscripción.Electoral=="PADRE LAS CASAS" & Año.de.Elección==1993)))

reselect0[c(id1_93), "Comuna"]<-c(rep("CONCEPCION", length(id1_93)))
reselect0[c(id2_93), "Comuna"]<-c(rep("VIÑA DEL MAR", length(id2_93)))
reselect0[c(id3_93), "Comuna"]<-c(rep("PELARCO", length(id3_93)))
reselect0[c(id4_93), "Comuna"]<-c(rep("TEMUCO", length(id4_93)))

```

Corregido lo anterior, se creó un nuevo data frame:

```

library(tidyr)
library(reshape)
reselct<-cast(reselect0 %>% complete(Comuna, Año.de.Elección,
Lista, fill=list(Votos.Totales=0)), Comuna~Lista+Año.de.Elección, fun="sum")

```

La comuna Navarino cambió de nombre a Cabo de Hornos el año 2001. Para el ajuste previo al procesamiento definitivo se dejó como Cabo de Hornos a lo largo de todo el proceso analizado.

```

reselct[reselct$Comuna=="CABO DE HORNOS", c("A_1993", "B_1993", "C_1993", "D_1993",
"CI_1993", "NB_1993")]<-reselct[reselct$Comuna=="NAVARINO", c("A_1993", "B_1993", "C_1993",
"D_1993", "CI_1993", "NB_1993")]

head(subset.data.frame(reselct, Comuna=="NAVARINO")) ## registro 182 ##
reselct<-reselct[-182,]

```

Finalmente, fueron completados algunos datos faltantes en el Excel original, datos que fueron verificados desde el sitio web del Tribunal Calificador de Elecciones (TRICEL)<sup>26</sup>.

```

reselct[reselct$Comuna=="QUILLECO",
c("A_1997", "B_1997", "C_1997", "CI_1997", "D_1997", "E_1997",
"NB_1997")]<-c(225, 1936, 2119, 0, 196, 100, 1055)

reselct[reselct$Comuna=="SAN FERNANDO",
c("A_1997", "B_1997", "C_1997", "CI_1997", "D_1997", "E_1997",
"NB_1997")]<-c(625, 10859, 8800, 0, 4032, 2485, 5723)

reselct[reselct$Comuna=="CHIMBARONGO",

```

<sup>26</sup><http://www.tribunalcalificador.cl/resultados-electorales/>

```
c("A_1997", "B_1997", "C_1997", "CI_1997", "D_1997", "E_1997",
"NB_1997")]<-c(482, 2344, 5650, 0, 886, 4284, 2885)

reselct[reselct$Comuna=="QUILACO",
c("A_1997", "B_1997", "C_1997", "CI_1997", "D_1997", "E_1997",
"NB_1997")]<-c(65, 481, 1329, 0, 81, 37, 330)
```

## Composición del Padrón Electoral Comunal

Al igual que en el caso anterior, los datos de la comuna Navarino fueron llevados al registro de la comuna Cabo de Hornos.

```
regcomunal[regcomunal$Comuna=="CABO DE HORNOS",
c("insc93", "ct93", "cd93", "insc97", "ct97", "cd97")]<-regcomunal[regcomunal$Comuna=="
NAVARINO", c("insc93", "ct93", "cd93", "insc97", "ct97", "cd97")]

head(subset.data.frame(regcomunal, Comuna=="NAVARINO")) ##registro en fila 285 ##
regcomunal<-regcomunal[-285,]
```

## Unificación de Datos y Creación de Variables

### Unificación de Datos

```
regcomunal$Comuna<-toupper(regcomunal$Comuna)
datos<-merge(reselct, regcomunal, by="Comuna")
```

### Creación de Variables

**Número de electores en Padrón comunal por año.** Corresponde a los valores representados previamente como  $N_{(t)m}$ ,  $t = 1993, \dots, 2009$ ,  $m = 1, \dots, M_{(t)}$ .

```
datos$N93<-datos$insc93
datos$N97<-datos$insc97
datos$N01<-datos$insc01
datos$N05<-datos$insc05
datos$N09<-datos$insc09
```

**Número de votos válidamente emitidos.** Primero se asocian las cifras por listas a los agentes de representación presentes en el periodo. La equivalencia Lista-Agentes de representación se presenta en las siguientes tablas.

**Tabla 12: Equivalencia Lista-Conglomerado Elección de Diputados año 1993**

Agente de Representación	Listas
Centro izquierda	D
Derecha	B
Izquierda extra-parlamentaria	A y C
Otros	Candidaturas Independientes

**Tabla 13: Equivalencia Lista-Conglomerado Elección de Diputados año 1997**

Agente de Representación	Listas
Centro izquierda	C
Derecha	B
Izquierda extra-parlamentaria	A y D
Otros	Candidaturas Independientes y Lista E

**Tabla 14: Elección año 2001**

Agente de Representación	Listas
Centro izquierda	E
Derecha	C
Izquierda extra-parlamentaria	A y B
Otros	D y Candidaturas Independientes (CI)

**Tabla 15: Elección año 2005**

Agente de Representación	Listas
Centro izquierda	B
Derecha	D
Izquierda extra-parlamentaria	C
Otros	Lista A y Candidaturas Independientes

**Tabla 16: Elección año 2009**

Agente de Representación	Listas
Centro izquierda	A
Derecha	B
Izquierda extra-parlamentaria	C
Otros	D y Candidaturas Independientes

En R, lo anterior se realizó de la siguiente manera.

```
reselct$cizq93<-reselct$D_1993
reselct$der93<-reselct$B_1993
reselct$izqex93<-apply(reselct[,c("A_1993", "C_1993")], 1, sum)
```

```

reselct$otros93<-reselct$CI_1993

reselct$cizq97<-reselct$C_1997
reselct$der97<-reselct$B_1997
reselct$izqex97<-apply(reselct[,c("A_1997", "D_1997")], 1, sum)
reselct$otros97<-apply(reselct[,c("CI_1997", "E_1997")], 1, sum)

reselct$cizq01<-reselct$E_2001
reselct$der01<-reselct$C_2001
reselct$izqex01<-apply(reselct[,c("A_2001", "B_2001")], 1, sum)
reselct$otros01<-apply(reselct[,c("CI_2001", "D_2001")], 1, sum)

reselct$cizq05<-reselct$B_2005
reselct$der05<-reselct$D_2005
reselct$izqex05<-reselct$C_2005
reselct$otros05<-apply(reselct[,c("A_2005", "G_2005")], 1, sum)

reselct$cizq09<-reselct$A_2009
reselct$der09<-reselct$B_2009
reselct$izqex09<-reselct$C_2009
reselct$otros09<-apply(reselct[,c("D_2009", "G_2009")], 1, sum)

```

Luego se calculan las cantidades representadas previamente como  $n_{(t)m}$ ,  $t = 1993, \dots, 2009$ ,  $m = 1, \dots, M_{(t)}$ .

```

reselct$n93<-apply(reselct[,c("cizq93", "der93", "izqex93", "otros93")], 1, sum)
reselct$n97<-apply(reselct[,c("cizq97", "der97", "izqex97", "otros97")], 1, sum)
reselct$n01<-apply(reselct[,c("cizq01", "der01", "izqex01", "otros01")], 1, sum)
reselct$n05<-apply(reselct[,c("cizq05", "der05", "izqex05", "otros05")], 1, sum)
reselct$n09<-apply(reselct[,c("cizq09", "der09", "izqex09", "otros09")], 1, sum)

```

**Abstencionismo.** Fueron representados con el prefijo “abs”.

```

datos$abs93<-datos$N93-datos$n93
datos$abs97<-datos$N97-datos$n97
datos$abs01<-datos$N01-datos$n01
datos$abs05<-datos$N05-datos$n05
datos$abs09<-datos$N09-datos$n09

```

**Ex y Nuevos votantes.** Se utiliza el prefijo “tipo8” para ex votantes, y “nv” para los nuevos. Las cifras fueron obtenidas con la siguiente sintaxis.

```

datos$tipo8_97<-datos$ct97-datos$ct93
datos$tipo8_01<-datos$ct01-datos$ct97
datos$tipo8_05<-datos$ct05-datos$ct01
datos$tipo8_09<-datos$ct09-datos$ct05
datos[datos$Comuna=="VITACURA","tipo8_97"]<-0 ##Dato inconsistente desde su origen##

datos$nv_97<-datos$N97+datos$tipo8_97-datos$N93
datos$nv_01<-datos$N01+datos$tipo8_01-datos$N97
datos$nv_05<-datos$N05+datos$tipo8_05-datos$N01
datos$nv_09<-datos$N09+datos$tipo8_09-datos$N05

```

## Ajuste Final para Procesamiento entre procesos electorales

Para cada par de procesos electorales objeto de análisis fueron excluidas aquellas comunas que experimentaron escisiones, o bien, emergieron en el periodo. Esto ocurrió para los pares 1997-1993 y 2001-2005.

```

datos9793<- datos %>% filter(!Comuna %in% c("CONCEPCION", "CHIGUAYANTE",
"SAN PEDRO DE LA PAZ", "VIÑA DEL MAR", "CONCON", "PELARCO", "SAN RAFAEL",
"CHILLAN", "CHILLAN VIEJO", "TEMUCO", "PADRE LAS CASAS", "PEÑAFLO",
"PADRE HURTADO")) %>% filter(!Comuna %in% c("ALTO HOSPICIO", "HUALPEN",
"CHOLCHOL", "ALTO BIOBIO"))%>% select(Comuna, contains("97"),contains("93"))

datos0197<- datos %>% filter(!Comuna %in% c("ALTO HOSPICIO", "HUALPEN",
"CHOLCHOL", "ALTO BIOBIO"))%>% select(Comuna, contains("01"),contains("97"))

datos0501<- datos %>% filter(!Comuna %in% c("ALTO HOSPICIO", "HUALPEN",
"CHOLCHOL", "ALTO BIOBIO")) %>% filter(!Comuna %in% c("IQUIQUE", "TALCAHUANO",
"NUEVA IMPERIAL", "SANTA BARBARA")) %>% select(Comuna, contains("05"),contains("01"))

datos0905<- datos %>% select(Comuna, contains("09"),contains("05"))

```

## Aplicación de Modelos

El modelo de optimización genera siempre el mismo resultado. El modelo bayesiano, en cambio, al usar métodos MCMC para el muestreo de la distribución posterior, arroja estimaciones diferentes, pero consistentes entre sí. En este último caso los parámetros burnin=100.000 y thin=1.000 adoptados para la convergencia del método requieren una considerable capacidad de hardware<sup>27</sup>.

En lugar de ejecutarse nuevamente, los resultados pueden ser leídos desde los archivos adjuntos, habiendo especificado el directorio de trabajo, por medio del siguiente código,

<sup>27</sup>El algoritmo fue procesado con un equipo portátil con procesador Intel Core I7 y 20GB de memoria RAM, demorando la ejecución del algoritmo ei.md.Bayes entre 30 y 60 minutos

```
require(rlist)
md9397<-list.load('md9397.rds')
md9701<-list.load('md9701.rds')
md0105<-list.load('md0105.rds')
md0509<-list.load('md0509.rds')
```

## Periodo 1997-1993

```
library(eiPack)
md9397<-summary(ei.MD.bayes(cbind(cizq97,der97,izqex97,otros97,abs97,tipos8_97) ~
cbind(cizq93,der93,izqex93,otros93,abs93,nv_97),
burnin=100000, thin=1000, data=datos9793))

require(rlist)
list.save(md9397$draws, 'md9397.rds')

library(vottrans)
R0_9397<-as.matrix(cbind(rowSums(datos9793[,c("cizq93","der93","izqex93",
"otros93","abs93","nv_97")])),datos9793[,c("cizq93","der93","izqex93",
"otros93","abs93","nv_97")]))

R1_9397<-as.matrix(cbind(rowSums(datos9793[,c("cizq97","der97","izqex97",
"otros97","abs97","tipos8_97")])),datos9793[,c("cizq97","der97","izqex97","otros97",
"abs97","tipos8_97")]))

Phi_9397<-vottrans(R0_9397,R1_9397)

md9397<-list.load('md9397.rds')
X_h97<-as.matrix(rbind(md9397$Cell.counts[1:6,1],
md9397$Cell.counts[7:12,1],
md9397$Cell.counts[13:18,1],md9397$Cell.counts[19:24,1],
md9397$Cell.counts[25:30,1],md9397$Cell.counts[31:36,1]))
X_h97<-(1/sum(X_h97[1:5,]))*X_h97
colnames(X_h97)<-c("V1","V2","V3","V4","V5","V6")

p1_h97<-sum(diag(X_h97[1:4,1:4]))
p2_h97<-sum(X_h97[1:4,1:4])-sum(diag(X_h97[1:4,1:4]))
p3_h97<-sum(X_h97[5,1:4])
p4_h97<-sum(X_h97[1:4,5])
p5_h97<-X_h97[5,5]
p6_h97<-sum(X_h97[1:4,6])
p7_h97<-X_h97[5,6]
```

```

N2<-sum(datos9793$insc97)
n1B<-diag(apply(datos9793[,c("cizq93","der93","izqex93","otros93",
"abs93","nv_97")],2,sum))

X_o97<-t(Phi_9397)%*%n1B)*(1/N2)

p1_o97<-sum(diag(X_o97[1:4,1:4]))
p2_o97<-sum(X_o97[1:4,1:4])-sum(diag(X_o97[1:4,1:4]))
p3_o97<-sum(X_o97[5,1:4])
p4_o97<-sum(X_o97[1:4,5])
p5_o97<-X_o97[5,5]
p6_o97<-sum(X_o97[1:4,6])
p7_o97<-X_o97[5,6]

p_h97<-c(p1_h97,p2_h97,p3_h97,p4_h97,p5_h97,p6_h97,p7_h97)
p_o97<-as.numeric(c(p1_o97,p2_o97,p3_o97,p4_o97,p5_o97,p6_o97,p7_o97))

```

## Periodos 2001-1997, 2005-2001, 2009-2005

```

#### Periodo 2001-1997 ###

md9701<-summary(ei.MD.bayes(cbind(cizq01,der01,izqex01,otros01,abs01,
tipo8_01) ~ cbind(cizq97,der97,izqex97,otros97,abs97,nv_01),
burnin=100000, thin=1000, data=datos0197))
list.save(md9701$draws, 'md9701.rds')

R0_9701<-as.matrix(cbind(rowSums(datos0197[,c("cizq97","der97","izqex97",
"otros97","abs97","nv_01")]),datos0197[,c("cizq97","der97","izqex97",
"otros97","abs97","nv_01")]))

R1_9701<-as.matrix(cbind(rowSums(datos0197[,c("cizq01","der01","izqex01",
"otros01","abs01","tipo8_01")]),datos0197[,c("cizq01","der01","izqex01",
"otros01","abs01","tipo8_01")]))

Phi_9701<-vottrans(R0_9701,R1_9701)

md9701<-list.load('md9701.rds')
X_h01<-as.matrix(rbind(md9701$Cell.counts[1:6,1],
md9701$Cell.counts[7:12,1],
md9701$Cell.counts[13:18,1],md9701$Cell.counts[19:24,1],
md9701$Cell.counts[25:30,1],md9701$Cell.counts[31:36,1]))

```

```

X_h01<- (1/sum(X_h01[1:5,]))*X_h01
colnames(X_h01)<-c("V1", "V2", "V3", "V4", "V5", "V6")

p1_h01<-sum(diag(X_h01[1:4,1:4]))
p2_h01<-sum(X_h01[1:4,1:4])-sum(diag(X_h01[1:4,1:4]))
p3_h01<-sum(X_h01[5,1:4])
p4_h01<-sum(X_h01[1:4,5])
p5_h01<-X_h01[5,5]
p6_h01<-sum(X_h01[1:4,6])
p7_h01<-X_h01[5,6]

N2<-sum(datos0197$insc01)
n1B<-diag(apply(datos0197[,c("cizq97", "der97", "izqex97", "otros97",
"abs97", "nv_01")], 2, sum))

X_o01<- (t(Phi_9701)%*%n1B)*(1/N2)

p1_o01<-sum(diag(X_o01[1:4,1:4]))
p2_o01<-sum(X_o01[1:4,1:4])-sum(diag(X_o01[1:4,1:4]))
p3_o01<-sum(X_o01[5,1:4])
p4_o01<-sum(X_o01[1:4,5])
p5_o01<-X_o01[5,5]
p6_o01<-sum(X_o01[1:4,6])
p7_o01<-X_o01[5,6]

p_h01<-c(p1_h01, p2_h01, p3_h01, p4_h01, p5_h01, p6_h01, p7_h01)
p_o01<-c(p1_o01, p2_o01, p3_o01, p4_o01, p5_o01, p6_o01, p7_o01)

#### Periodo 2005-2001 ###

md0105<-summary(ei.MD.bayes(cbind(cizq05, der05, izqex05, otros05,
abs05, tipo8_05) ~ cbind(cizq01, der01, izqex01, otros01, abs01, nv_05),
burnin=100000, thin=1000, data=datos0501))
list.save(md0105$draws, 'md0105.rds')

R0_0105<-as.matrix(cbind(rowSums(datos0501[,c("cizq01", "der01", "izqex01",
"otros01", "abs01", "nv_05")]), datos0501[,c("cizq01", "der01", "izqex01",
"otros01", "abs01", "nv_05")]))

R1_0105<-as.matrix(cbind(rowSums(datos0501[,c("cizq05", "der05", "izqex05",
"otros05", "abs05", "tipo8_05")]), datos0501[,c("cizq05", "der05", "izqex05",
"otros05", "abs05", "tipo8_05")]))

```

```

Phi_0105<-vottrans(R0_0105,R1_0105)

md0105<-list.load('md0105.rds')
X_h05<-as.matrix(rbind(md0105$Cell.counts[1:6,1],
md0105$Cell.counts[7:12,1],
md0105$Cell.counts[13:18,1],md0105$Cell.counts[19:24,1],
md0105$Cell.counts[25:30,1],md0105$Cell.counts[31:36,1]))
X_h05<-(1/sum(X_h05[1:5,]))*X_h05
colnames(X_h05)<-c("V1", "V2", "V3", "V4", "V5", "V6")

p1_h05<-sum(diag(X_h05[1:4,1:4]))
p2_h05<-sum(X_h05[1:4,1:4])-sum(diag(X_h05[1:4,1:4]))
p3_h05<-sum(X_h05[5,1:4])
p4_h05<-sum(X_h05[1:4,5])
p5_h05<-X_h05[5,5]
p6_h05<-sum(X_h05[1:4,6])
p7_h05<-X_h05[5,6]

N2<-sum(datos0501$insc05)
n1B<-diag(apply(datos0501[,c("cizq01", "der01", "izqex01", "otros01",
"abs01", "nv_05")], 2, sum))

X_o05<-t(Phi_0105)%*%n1B)*(1/N2)

p1_o05<-sum(diag(X_o05[1:4,1:4]))
p2_o05<-sum(X_o05[1:4,1:4])-sum(diag(X_o05[1:4,1:4]))
p3_o05<-sum(X_o05[5,1:4])
p4_o05<-sum(X_o05[1:4,5])
p5_o05<-X_o05[5,5]
p6_o05<-sum(X_o05[1:4,6])
p7_o05<-X_o05[5,6]

p_h05<-c(p1_h05,p2_h05,p3_h05,p4_h05,p5_h05,p6_h05,p7_h05)
p_o05<-c(p1_o05,p2_o05,p3_o05,p4_o05,p5_o05,p6_o05,p7_o05)

#### Periodo 2009-2005 ####

md0509<-summary(ei.MD.bayes(cbind(cizq09,der09,izqex09,otros09,
abs09,tipo8_09) ~ cbind(cizq05,der05,izqex05,otros05,abs05,nv_09),
burnin=100000, thin=1000, data=datos0905))
list.save(md0509$draws, 'md0509.rds')

```

```

R0_0509<-as.matrix(cbind(rowSums(datos0905[,c("cizq05","der05",
"izqex05","otros05","abs05","nv_09")]),datos0905[,c("cizq05","der05",
"izqex05","otros05","abs05","nv_09")]))

R1_0509<-as.matrix(cbind(rowSums(datos0905[,c("cizq09","der09",
"izqex09","otros09","abs09","tipo8_09")]),datos0905[,c("cizq09",
"der09","izqex09","otros09","abs09","tipo8_09")]))

Phi_0509<-vottrans(R0_0509,R1_0509)

md0509<-list.load('md0509.rds')
X_h09<-as.matrix(rbind(md0509$Cell.counts[1:6,1],
md0509$Cell.counts[7:12,1],
md0509$Cell.counts[13:18,1],md0509$Cell.counts[19:24,1],
md0509$Cell.counts[25:30,1],md0509$Cell.counts[31:36,1]))
X_h09<-(1/sum(X_h09[1:5,]))*X_h09
colnames(X_h09)<-c("V1","V2","V3","V4","V5","V6")

p1_h09<-sum(diag(X_h09[1:4,1:4]))
p2_h09<-sum(X_h09[1:4,1:4])-sum(diag(X_h09[1:4,1:4]))
p3_h09<-sum(X_h09[5,1:4])
p4_h09<-sum(X_h09[1:4,5])
p5_h09<-X_h09[5,5]
p6_h09<-sum(X_h09[1:4,6])
p7_h09<-X_h09[5,6]

N2<-sum(datos0905$insc09)
n1B<-diag(apply(datos0905[,c("cizq05","der05","izqex05",
"otros05","abs05","nv_09")],2,sum))

X_o09<-t(Phi_0509)%*%n1B)*(1/N2)

p1_o09<-sum(diag(X_o09[1:4,1:4]))
p2_o09<-sum(X_o09[1:4,1:4])-sum(diag(X_o09[1:4,1:4]))
p3_o09<-sum(X_o09[5,1:4])
p4_o09<-sum(X_o09[1:4,5])
p5_o09<-X_o09[5,5]
p6_o09<-sum(X_o09[1:4,6])
p7_o09<-X_o09[5,6]

p_h09<-c(p1_h09,p2_h09,p3_h09,p4_h09,p5_h09,p6_h09,p7_h09)
p_o09<-c(p1_o09,p2_o09,p3_o09,p4_o09,p5_o09,p6_o09,p7_o09)

```

## Estimaciones

```
round(X_o97,4); round(X_h97,4)
```

```
##      V1      V2      V3      V4      V5      V6
## 1 0.3032 0.0082 0.0285 0e+00 0.0000 0.0262
## 2 0.0028 0.2561 0.0000 0e+00 0.0000 0.0000
## 3 0.0516 0.0006 0.0229 4e-04 0.0000 0.0000
## 4 0.0199 0.0000 0.0000 3e-04 0.0000 0.0021
## 5 0.0601 0.0310 0.0049 0e+00 0.1516 0.0300
## 6 0.0261 0.0100 0.0103 0e+00 0.0157 0.1306
##      V1      V2      V3      V4      V5      V6
## [1,] 0.2750 0.0650 0.0111 1e-04 0.0058 0.0094
## [2,] 0.0801 0.1522 0.0102 1e-04 0.0073 0.0072
## [3,] 0.0292 0.0131 0.0242 2e-04 0.0041 0.0052
## [4,] 0.0162 0.0032 0.0004 0e+00 0.0006 0.0017
## [5,] 0.0457 0.0630 0.0096 1e-04 0.1421 0.0178
## [6,] 0.0174 0.0092 0.0111 1e-04 0.0073 0.1476
```

```
round(X_o01,4); round(X_h01,4)
```

```
##      V1      V2      V3      V4      V5      V6
## 1 0.2686 0.0324 0.0346 0.0062 0.0127 0.0099
## 2 0.0754 0.2015 0.0000 0.0098 0.0274 0.0228
## 3 0.0000 0.0036 0.0391 0.0016 0.0028 0.0012
## 4 0.0000 0.0000 0.0000 0.0025 0.0069 0.0018
## 5 0.0104 0.0156 0.0000 0.0000 0.1851 0.0280
## 6 0.0081 0.0070 0.0009 0.0002 0.0445 0.0377
##      V1      V2      V3      V4      V5      V6
## [1,] 0.2234 0.0407 0.0429 0.0071 0.0472 0.0032
## [2,] 0.0833 0.1736 0.0076 0.0085 0.0512 0.0127
## [3,] 0.0201 0.0056 0.0011 0.0018 0.0163 0.0034
## [4,] 0.0044 0.0008 0.0001 0.0004 0.0048 0.0007
## [5,] 0.0219 0.0264 0.0110 0.0012 0.1473 0.0314
## [6,] 0.0095 0.0130 0.0120 0.0013 0.0126 0.0500
```

```
round(X_o05,4); round(X_h05,4)
```

```
##      V1      V2      V3      V4      V5      V6
## 1 0.3082 0.0486 0.0208 0.0019 0.0033 0.0373
## 2 0.0000 0.2869 0.0000 0.0018 0.0000 0.0274
## 3 0.0324 0.0000 0.0276 0.0002 0.0000 0.0006
## 4 0.0000 0.0000 0.0000 0.0048 0.0072 0.0015
## 5 0.0036 0.0000 0.0000 0.0017 0.1682 0.0162
```

```
## 6 0.0155 0.0000 0.0000 0.0001 0.0579 0.0310
##          V1      V2      V3      V4      V5      V6
## [1,] 0.2821 0.0716 0.0151 0.0019 0.0293 0.0199
## [2,] 0.0349 0.2394 0.0045 0.0020 0.0119 0.0233
## [3,] 0.0169 0.0079 0.0250 0.0008 0.0050 0.0052
## [4,] 0.0054 0.0029 0.0001 0.0024 0.0016 0.0010
## [5,] 0.0105 0.0067 0.0014 0.0020 0.1533 0.0159
## [6,] 0.0099 0.0069 0.0022 0.0014 0.0354 0.0486
```

```
round(X_o09,4); round(X_h09,4)
```

```
##          V1      V2      V3      V4      V5      V6
## 1 0.2594 0.0381 0.0348 0.0050 0.0070 0.0098
## 2 0.0322 0.2704 0.0067 0.0034 0.0000 0.0342
## 3 0.0185 0.0000 0.0046 0.0001 0.0000 0.0132
## 4 0.0549 0.0000 0.0000 0.0060 0.0000 0.0000
## 5 0.0233 0.0000 0.0033 0.0012 0.1554 0.0184
## 6 0.0241 0.0000 0.0096 0.0011 0.0249 0.0354
##          V1      V2      V3      V4      V5      V6
## [1,] 0.2258 0.0769 0.0311 0.0070 0.0054 0.0080
## [2,] 0.0949 0.2046 0.0114 0.0034 0.0065 0.0261
## [3,] 0.0237 0.0057 0.0011 0.0005 0.0024 0.0032
## [4,] 0.0457 0.0075 0.0012 0.0006 0.0026 0.0032
## [5,] 0.0104 0.0070 0.0047 0.0020 0.1650 0.0125
## [6,] 0.0119 0.0068 0.0094 0.0034 0.0056 0.0580
```

```
Perfiles=c("Estable","Transferido","Desmovilizado","Movilizado",
"Abstencionismo permanente","Nuevo concurrente","Nuevo asbtencionista")
```

```
Res97<-data.frame(Perfiles,MD97=p_h97,Opt97=p_o97); Res97
```

```
##          Perfiles      MD97      Opt97
## 1          Estable 0.45132308 5.824122e-01
## 2          Transferido 0.22915462 1.119305e-01
## 3          Desmovilizado 0.11839757 9.592418e-02
## 4          Movilizado 0.01777131 -1.775978e-18
## 5 Abstencionismo permanente 0.14212983 1.515505e-01
## 6          Nuevo concurrente 0.02345317 2.821176e-02
## 7          Nuevo asbtencionista 0.01777042 2.997079e-02
```

```
Res01<-data.frame(Perfiles,MD01=p_h01,Opt01=p_o01); Res01
```

```
##          Perfiles      MD01      Opt01
## 1          Estable 0.39849201 0.51175904
```

```

## 2      Transferido 0.22289856 0.16363003
## 3      Desmovilizado 0.06047982 0.02607902
## 4      Movilizado 0.11940133 0.04968700
## 5 Abstencionismo permanente 0.14733952 0.18512271
## 6      Nuevo concurrente 0.02001412 0.03573200
## 7      Nuevo asbtencionista 0.03137465 0.02799020

Res05<-data.frame(Perfiles,MD05=p_h05,Opt05=p_o05); Res05

##          Perfiles      MD05      Opt05
## 1      Estable 0.54889012 0.627442915
## 2      Transferido 0.16412823 0.105681724
## 3      Desmovilizado 0.02058483 0.005311667
## 4      Movilizado 0.04780705 0.010439459
## 5 Abstencionismo permanente 0.15325189 0.168201933
## 6      Nuevo concurrente 0.04942648 0.066702772
## 7      Nuevo asbtencionista 0.01591140 0.016219531

Res09<-data.frame(Perfiles,MD09=p_h01,Opt09=p_o09); Res09

##          Perfiles      MD09      Opt09
## 1      Estable 0.39849201 0.540476107
## 2      Transferido 0.22289856 0.193757226
## 3      Desmovilizado 0.06047982 0.027758817
## 4      Movilizado 0.11940133 0.007029476
## 5 Abstencionismo permanente 0.14733952 0.155379211
## 6      Nuevo concurrente 0.02001412 0.057217728
## 7      Nuevo asbtencionista 0.03137465 0.018381435

```

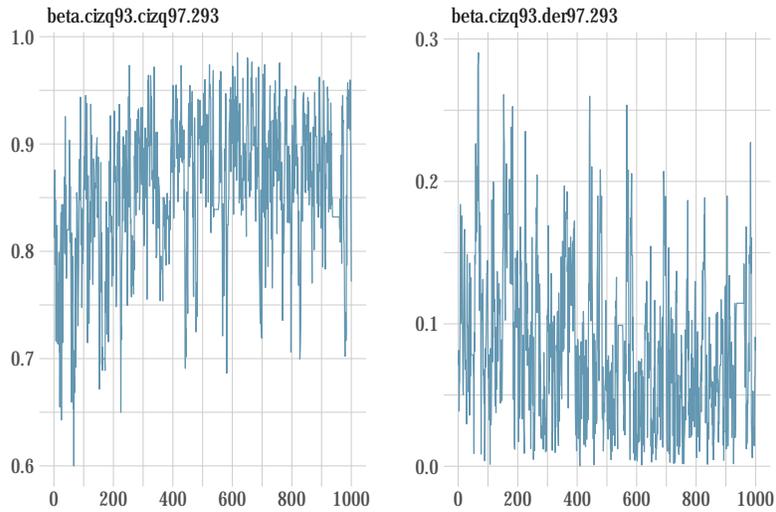
## Convergencia de métodos MCMC

```

require(bayesplot)
mcmc_trace(md9397$Beta, pars = c("beta.cizq93.cizq97.293", "beta.cizq93.der97.293")) +
theme_ipsum() +
theme(legend.position="none")

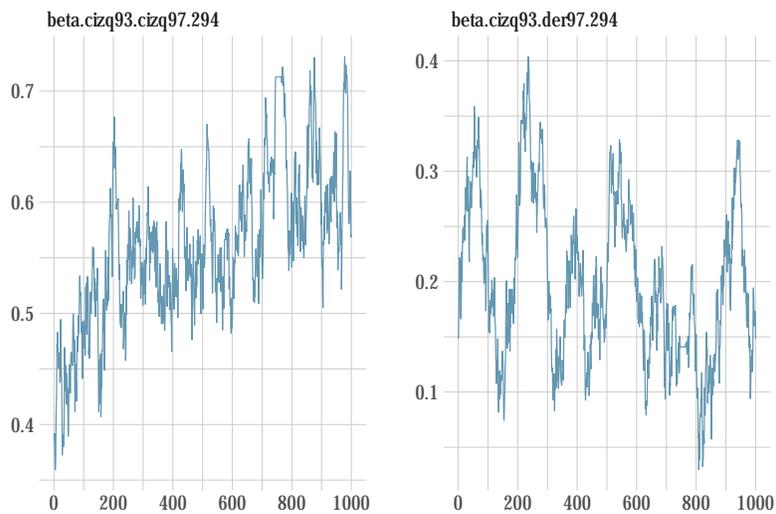
```

Figura 10: Comuna: Santa María (N° votantes en Padrón comunal 1997=7.301)



```
mcmc_trace(md9397$Beta, pars = c("beta.cizq93.cizq97.294", "beta.cizq93.der97.294")) +  
theme_ipsum() + theme(legend.position="none")
```

Figura 11: Comuna: Santiago (N° votantes en Padrón comunal 1997=151.335)



## Cálculo de Cotas determinísticas

```
cizq_93<-sum(datos9793$cizq93)
der_93<-sum(datos9793$der93)
izqex_93<-sum(datos9793$izqex93)
otros_93<-sum(datos9793$otros93)
abs_93<-sum(datos9793$abs93)

cizq_97<-sum(datos9793$cizq97)
der_97<-sum(datos9793$der97)
izqex_97<-sum(datos9793$izqex97)
otros_97<-sum(datos9793$otros97)
abs_97<-sum(datos9793$abs97)
nv_97<-sum(datos9793$nv_97)
tipo8_97<-sum(datos9793$tipo8_97)

N_93<-sum(datos9793$N93)
N_97<-sum(datos9793$N97)
n_93<-sum(datos9793$n93)
n_97<-sum(datos9793$n97)

V_ast97<-(1/(2*N_97))*(abs(cizq_97-cizq_93)+abs(der_97-der_93)+abs(izqex_97-izqex_93)+
abs(otros_97-otros_93)+abs(abs_97-abs_93)+abs(nv_97-tipo8_97))

cotas97<-as.matrix(rbind(c(0,min(n_97/N_97,n_93/N_97,((N_93+nv_97)/N_97)-V_ast97)),
c(0,min(n_97/N_97,n_93/N_97)),
(1/N_97)*c(max(0,(n_97+n_93)-(N_93+nv_97)),min(n_97,n_93)),
(1/N_97)*c(max(0,(N_97+n_93)-(N_93+n_97+nv_97)),min(n_93,N_97-n_97)),
(1/N_97)*c(max(0,(n_97-n_93-nv_97)),min(n_93,N_93-n_93)),
(1/N_97)*c(max(0,(N_97-n_97-n_93-nv_97)),min(N_97-n_97,N_93-n_93)),
(1/N_97)*c(max(0,(n_97-N_93)),min(n_97,nv_97)),
(1/N_97)*c(max(0,(N_97-n_97-N_93)),min(N_97-n_97,nv_97))))))

cizq_97<-sum(datos0197$cizq97)
der_97<-sum(datos0197$der97)
izqex_97<-sum(datos0197$izqex97)
otros_97<-sum(datos0197$otros97)
abs_97<-sum(datos0197$abs97)

cizq_01<-sum(datos0197$cizq01)
der_01<-sum(datos0197$der01)
izqex_01<-sum(datos0197$izqex01)
otros_01<-sum(datos0197$otros01)
```

```

abs_01<-sum(datos0197$abs01)
nv_01<-sum(datos0197$nv_01)
tipo8_01<-sum(datos0197$tipo8_01)

N_97<-sum(datos0197$N97)
N_01<-sum(datos0197$N01)
n_97<-sum(datos0197$n97)
n_01<-sum(datos0197$n01)

V_ast01<-(1/(2*N_01))*(abs(cizq_01-cizq_97)+abs(der_01-der_97)+abs(izqex_01-izqex_97)+
abs(otros_01-otros_97)+abs(abs_01-abs_97)+abs(nv_01-tipo8_01))

cotas01<-as.matrix(rbind(c(0,min(n_01/N_01,n_97/N_01,((N_97+nv_01)/N_01)-V_ast01)),
c(0,min(n_01/N_01,n_97/N_01)),
(1/N_01)*c(max(0,(n_01+n_97)-(N_97+nv_01)),min(n_01,n_97)),
(1/N_01)*c(max(0,(N_01+n_97)-(N_97+n_01+nv_01)),min(n_97,N_01-n_01)),
(1/N_01)*c(max(0,(n_01-n_97-nv_01)),min(n_97,N_97-n_97)),
(1/N_01)*c(max(0,(N_01-n_01-n_97-nv_01)),min(N_01-n_01,N_97-n_97)),
(1/N_01)*c(max(0,(n_01-N_97)),min(n_01,nv_01)),
(1/N_01)*c(max(0,(N_01-n_01-N_97)),min(N_01-n_01,nv_01))))

cizq_01<-sum(datos0501$cizq01)
der_01<-sum(datos0501$der01)
izqex_01<-sum(datos0501$izqex01)
otros_01<-sum(datos0501$otros01)
abs_01<-sum(datos0501$abs01)

cizq_05<-sum(datos0501$cizq05)
der_05<-sum(datos0501$der05)
izqex_05<-sum(datos0501$izqex05)
otros_05<-sum(datos0501$otros05)
abs_05<-sum(datos0501$abs05)
nv_05<-sum(datos0501$nv_05)
tipo8_05<-sum(datos0501$tipo8_05)

N_01<-sum(datos0501$N01)
N_05<-sum(datos0501$N05)
n_01<-sum(datos0501$n01)
n_05<-sum(datos0501$n05)

V_ast05<-(1/(2*N_05))*(abs(cizq_05-cizq_01)+abs(der_05-der_01)+abs(izqex_05-izqex_01)+
abs(otros_05-otros_01)+abs(abs_05-abs_01)+abs(nv_05-tipo8_05))

```

```

cotas05<-as.matrix(rbind(c(0,min(n_05/N_05,n_01/N_05,((N_01+nv_05)/N_05)-V_ast05)),
c(0,min(n_05/N_05,n_01/N_05)),
(1/N_05)*c(max(0,(n_05+n_01)-(N_01+nv_05)),min(n_05,n_01)),
(1/N_05)*c(max(0,(N_05+n_01)-(N_01+n_05+nv_05)),min(n_01,N_05-n_05)),
(1/N_05)*c(max(0,(n_05-n_01-nv_05)),min(n_01,N_01-n_01)),
(1/N_05)*c(max(0,(N_05-n_05-n_01-nv_05)),min(N_05-n_05,N_01-n_01)),
(1/N_05)*c(max(0,(n_05-N_01)),min(n_05,nv_05)),
(1/N_05)*c(max(0,(N_05-n_05-N_01)),min(N_05-n_05,nv_05))))

cizq_05<-sum(datos0905$cizq05)
der_05<-sum(datos0905$der05)
izqex_05<-sum(datos0905$izqex05)
otros_05<-sum(datos0905$otros05)
abs_05<-sum(datos0905$abs05)

cizq_09<-sum(datos0905$cizq09)
der_09<-sum(datos0905$der09)
izqex_09<-sum(datos0905$izqex09)
otros_09<-sum(datos0905$otros09)
abs_09<-sum(datos0905$abs09)
nv_09<-sum(datos0905$nv_09)
tipo8_09<-sum(datos0905$tipo8_09)

N_05<-sum(datos0905$N05)
N_09<-sum(datos0905$N09)
n_05<-sum(datos0905$n05)
n_09<-sum(datos0905$n09)

V_ast09<-(1/(2*N_09))*(abs(cizq_09-cizq_05)+abs(der_09-der_05)+abs(izqex_09-izqex_05)+
abs(otros_09-otros_05)+abs(abs_09-abs_05)+abs(nv_09-tipo8_09))

cotas09<-as.matrix(rbind(c(0,min(n_09/N_09,n_05/N_09,((N_05+nv_09)/N_09)-V_ast09)),
c(0,min(n_09/N_09,n_05/N_09)),
(1/N_09)*c(max(0,(n_09+n_05)-(N_05+nv_09)),min(n_09,n_05)),
(1/N_09)*c(max(0,(N_09+n_05)-(N_05+n_09+nv_09)),min(n_05,N_09-n_09)),
(1/N_09)*c(max(0,(n_09-n_05-nv_09)),min(n_05,N_05-n_05)),
(1/N_09)*c(max(0,(N_09-n_09-n_05-nv_09)),min(N_09-n_09,N_05-n_05)),
(1/N_09)*c(max(0,(n_09-N_05)),min(n_09,nv_09)),
(1/N_09)*c(max(0,(N_09-n_09-N_05)),min(N_09-n_09,nv_09))))

cotas_97<-data.frame(Perfiles=factor(seq(1,8,1), levels=c(1:8), labels=c("Estable",
"Transferido","Estable+Transferido","Desmovil.", "Movilizado", "Abst.perm", "Nvo. mov.",
"Nvo. abst")),cotas97); colnames(cotas_97)<-c("Perfiles","Minm","Maxm"); cotas_97

```

```
##           Perfiles      Minm      Maxm
## 1           Estable 0.0000000 0.7225545
## 2           Transferido 0.0000000 0.7225545
## 3 Estable+Transferido 0.3665209 0.7225545
## 4           Desmovil. 0.0000000 0.2774455
## 5           Movilizado 0.0000000 0.1672408
## 6           Abst.perm 0.0000000 0.1672408
## 7           Nvo. mov. 0.0000000 0.1887927
## 8           Nvo. abst 0.0000000 0.1887927
```

```
cotas_01<-data.frame(Perfiles=factor(seq(1,8,1), levels=c(1:8), labels=c("Estable",
"Transferido","Estable+Transferido","Desmovil.","Movilizado","Abst.perm","Nvo. mov.",
"Nvo. abst")),cotas01); colnames(cotas_01)<-c("Perfiles","Minm","Maxm"); cotas_01
```

```
##           Perfiles      Minm      Maxm
## 1           Estable 0.0000000 0.7176882
## 2           Transferido 0.0000000 0.7176882
## 3 Estable+Transferido 0.3800726 0.7176882
## 4           Desmovil. 0.0000000 0.2391919
## 5           Movilizado 0.0000000 0.2792980
## 6           Abst.perm 0.0000000 0.2391919
## 7           Nvo. mov. 0.0000000 0.1014374
## 8           Nvo. abst 0.0000000 0.1014374
```

```
cotas_05<-data.frame(Perfiles=factor(seq(1,8,1), levels=c(1:8), labels=c("Estable",
"Transferido","Estable+Transferido","Desmovil.","Movilizado","Abst.perm","Nvo. mov.",
"Nvo. abst")),cotas05); colnames(cotas_05)<-c("Perfiles","Minm","Maxm"); cotas_05
```

```
##           Perfiles      Minm      Maxm
## 1           Estable 0.0000000 0.7540963
## 2           Transferido 0.0000000 0.7540963
## 3 Estable+Transferido 0.4598340 0.7540963
## 4           Desmovil. 0.0000000 0.1897331
## 5           Movilizado 0.0000000 0.2364964
## 6           Abst.perm 0.0000000 0.1897331
## 7           Nvo. mov. 0.0000000 0.1139364
## 8           Nvo. abst 0.0000000 0.1139364
```

```
cotas_09<-data.frame(Perfiles=factor(seq(1,8,1), levels=c(1:8), labels=c("Estable",
"Transferido","Estable+Transferido","Desmovil.","Movilizado","Abst.perm","Nvo. mov.",
"Nvo. abst")),cotas09); colnames(cotas_09)<-c("Perfiles","Minm","Maxm"); cotas_09
```

```
##           Perfiles      Minm      Maxm
## 1           Estable 0.0000000 0.7967854
```

## 2	Transferido	0.0000000	0.7967854
## 3	Estable+Transferido	0.5001418	0.7967854
## 4	Desmovil.	0.0000000	0.2015195
## 5	Movilizado	0.0000000	0.1873482
## 6	Abst.perm	0.0000000	0.1873482
## 7	Nvo. mov.	0.0000000	0.1109906
## 8	Nvo. abst	0.0000000	0.1109906

## 5.2. Espacio de Probabilidad del Estimador

La matriz  $\mathbf{X}$  con los verdaderos valores subyacentes a la transición electoral entre dos procesos de votación consecutivos es una matriz dentro de las que satisfacen las siguientes condiciones.

1.  $\mathbf{v}_2 = \mathbf{A}\mathbf{X}\mathbf{B}^{-1}\mathbf{C}\mathbf{v}_1$ .

2.  $x_{rc} \in \left\{0, \frac{1}{N_2}, \dots, \frac{N_2-1}{N_2}, 1\right\}$ , si  $r = 1, \dots, k+1$

3.  $x_{rc} \in \left\{0, \frac{1}{N_2}, \dots, \frac{N_1+nv_2-N_2}{N_2}\right\}$ , si  $r = k+2$

4.  $\sum_{c=1}^{k+2} x_{rc} =$

$$\begin{cases} \frac{d_{2j}}{N_2} & \text{si } j = r = 1, \dots, k \\ \frac{N_2-n_2}{N_2} & \text{si } r = k+1 \\ \frac{N_1+nv_2-N_2}{N_2} & \text{si } r = k+2 \end{cases}$$

5.  $\sum_{r=1}^{k+2} x_{rc} =$

$$\begin{cases} \frac{d_{1j}}{N_2} & \text{si } j = c = 1, \dots, k \\ \frac{N_1-n_1}{N_2} & \text{si } c = k+1 \\ \frac{nv_2}{N_2} & \text{si } c = k+2 \end{cases}$$

6.  $x_{k+2,k+2} = 0$ .

7.  $tr(\mathbf{X}) \leq 1 - V^* + p_8$ .

Para valores  $N_t, n_t, nv_2$  y  $d_{tj}, t = 1, 2, j = 1, \dots, k$  determinados.

Existe un conjunto de matrices que, sin contener los verdaderos valores, cumplen con todas estas propiedades. Como consecuencia de las condiciones 2 a 5 recién enunciadas, éste resulta ser un conjunto finito.

Sea  $\Omega$  el conjunto de matrices que satisfacen estas siete propiedades. Las entradas de cada matriz en  $\Omega$  se representan a modo genérico como  $(x^*)_{rc}$ . También sea  $\mathcal{A}$  una  $\sigma$ -álgebra sobre  $\Omega$ .

Luego, sobre  $\Omega$  se define la variable aleatoria  $X_i \in \left\{[l_i, u_i] \cap \left\{0, \frac{1}{N_2}, \dots, \frac{N_2-1}{N_2}, 1\right\}\right\}$ , donde  $l_i$  y  $u_i$  corresponden, respectivamente, a las cotas inferiores y superiores dadas por (2.3.20) en la sección 2.3.2 para  $i = 1, \dots, 7$ .

Cada matriz  $\mathbf{X}^*$  tiene asociado un número finito de combinaciones en las que puede asumir los valores en sus entradas.

La probabilidad del evento  $X_i = x_i$  se puede modelar entonces como el el número de combinaciones que conducen a la solución  $\mathbf{X} = \mathbf{X}^*/X_i = x_i$ . Al definir la siguiente función de probabilidad  $\mathbb{P}$  como:

$$\mathbb{P}(X_i = x_i) = \frac{\#\{\mathbf{X} = \mathbf{X}^* / X_i = x_i\}}{\sum \#\{\mathbf{X} = \mathbf{X}^*\}} / \mathbf{X}^* \in \Omega$$

Entonces  $(\Omega, \mathcal{A}, \mathbb{P})$  resulta ser un espacio de probabilidad para cada  $i = 1, 2, \dots, 7$ . Luego, el evento  $N_2 X_i = N_2 x_i$  puede representarse como una suma de variables aleatorias con distribución Hipergeométrica multivariada truncada.

La función de probabilidad para una distribución hipergeométrica multivariada, para una variable  $N_j$  restringida al intervalo  $[b + 1, c]$  viene dada (Johnson et al., 1997: 177-179) por:

$$\begin{aligned} P(n_1, n_2, \dots, n_k) &= \frac{1}{\sum_{l=b+1}^c \binom{m_l}{l} \binom{m-m_j}{n-l}} \prod_{i=1}^k \binom{m_i}{n_i} \\ &= \frac{1}{\sum_{l=b+1}^c \frac{m_j^{(l)} (m-m_j)^{(n-l)}}{l!(n-l)!}} \prod_{i=1}^k \frac{m_i^{n_i}}{n_i!} \end{aligned} \quad (5.2.1)$$

En el caso de la distribución hipergeométrica multivariada doblemente truncada para un conjunto de variables  $N_1, N_2, \dots, N_s$ , con  $N_j$  restringida al intervalo  $[b_j + 1, c_j]$ ,  $j = 1, \dots, s$  la función de probabilidad es la siguiente.

$$\begin{aligned} P(n_1, n_2, \dots, n_k) &= \frac{1}{\sum_{l_1=b_1+1}^{c_1} \dots \sum_{l_s=b_s+1}^{c_s} \left\{ \prod_{i=1}^s \binom{m_i^{(l_i)}}{l_i} \right\} \left( \binom{m - \sum_{i=1}^s m_o}{n - \sum_{i=1}^s l_i} \right)} \prod_{i=1}^k \binom{m_i}{n_i} \\ &= \frac{1}{\sum_{l_1=b_1+1}^{c_1} \dots \sum_{l_s=b_s+1}^{c_s} \left\{ \prod_{i=1}^s \frac{m_i^{(l_i)}}{l_i!} \right\} \frac{(m - \sum_{i=1}^s m_i)^{(n - \sum_{i=1}^s l_i)}}{(n - \sum_{i=1}^s l_i)!}} \prod_{i=1}^k \frac{m_i^{(n_i)}}{n_i!} \end{aligned} \quad (5.2.2)$$

**Ejemplo 5.2.1** La situación más simple se presenta si  $k = 2$  y  $\bigcup_{i=3}^8 A_i = \emptyset$ . En tal caso  $E_2 = E_1 = T_1 = T_2$ . Una matriz genérica en este conjunto se vería de esta forma.

$$\begin{bmatrix} x_{11}^* & x_{12}^* & 0 & 0 \\ x_{21}^* & x_{22}^* & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Supongamos  $N_2 = 100$  y que se observan los siguientes marginales (filas para elección 2, columnas para elección 1).

$$\begin{bmatrix} x_{11}^* & x_{12}^* & 0 & 0 \\ x_{21}^* & x_{22}^* & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0,7 & 0,3 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} 0,4 \\ 0,6 \\ 0 \\ 0 \end{matrix} = \frac{1}{100} \begin{bmatrix} c_{11} & c_{12} & 0 & 0 \\ c_{21} & c_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 70 & 30 & 0 & 0 & 100 \end{bmatrix} \begin{matrix} 40 \\ 60 \\ 0 \\ 0 \end{matrix}$$

Podemos despejar las incógnitas en términos de la entrada 1,1.

- $c_{11} + c_{12} = 40 \Rightarrow c_{12} = 40 - c_{11}$
- $c_{21} = 70 - c_{11}$
- $c_{22} = c_{11} - 10$

También, la lógica de las cotas nos dice en este caso que:

$$\max(0, 40 - 30) \leq c_{11} \leq \min(40, 70) \Rightarrow c_{11} \in \mathbb{Z} \cap [10, 40]$$

Dada la simplificación inducida con el ejemplo, podemos en este caso identificar todas las matrices del conjunto solución. Nos quedamos aquí con la parte a priori no nula de cada matriz, que corresponden en este ejemplo a las primeras 2 filas y columnas.

$$\Omega = \left\{ \begin{bmatrix} 10 & 30 \\ 60 & 0 \end{bmatrix}, \begin{bmatrix} 11 & 29 \\ 59 & 1 \end{bmatrix}, \dots, \begin{bmatrix} 40 & 0 \\ 30 & 30 \end{bmatrix} \right\}$$

Resultan ser entonces  $40 - 10 + 1 = 31$  matrices que componen el conjunto solución.

Existe un número de combinaciones en las cuales cada matriz en  $\Omega$  puede asumir los valores de sus entradas. Por ejemplo, para la primera matriz de las enumeradas por extensión.

$$\# \left\{ \begin{bmatrix} 10 & 30 \\ 60 & 0 \end{bmatrix} \right\} = \binom{70}{10} \binom{30}{0} = \frac{70!}{10!(70-10)!} \frac{30!}{0!(30-0)!} = \frac{70!}{10!60!}$$

También, en el segundo y último caso de los enumerados.

$$\# \left\{ \begin{bmatrix} 11 & 29 \\ 59 & 1 \end{bmatrix} \right\} = \frac{70!}{11!59!} \frac{30!}{29!}$$

$$\# \left\{ \begin{bmatrix} 40 & 0 \\ 30 & 30 \end{bmatrix} \right\} = \frac{70!}{40!30!}$$

Luego, se puede atribuir una medida de probabilidad a cada matriz.

$$\mathbb{P} \left( \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} i & 40-i \\ 70-i & i-10 \end{bmatrix} \right) = \frac{\binom{70}{i} \binom{30}{i-10}}{\sum_{i=10}^{40} \binom{70}{i} \binom{30}{i-10}} \quad (5.2.3)$$

Esta expresión coincide exactamente con la forma general dada en (5.2.1).

Operando hacia atrás se llega a que  $X_1 \in [0,1; 0,7]$  y que:

$$\mathbb{P}(X_1 = 0,1) = \frac{\binom{70}{10} \binom{30}{0}}{\sum_{i=10}^{40} \binom{70}{i} \binom{30}{i-10}}$$

Para  $N_t \rightarrow \infty$ , como ha de ser por lo general el caso en este tipo de problemas, los factoriales involucrados pueden ser aproximados con la fórmula de Stirling.

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n = \sqrt{2\pi n} n^n \exp(-n)$$

Esto sugiere la existencia de una estructura Gaussiana de nivel asintótico subyacente en la solución general.

## 5.3. Métodos Montecarlo con Cadenas de Markov

### 5.3.1. Terminología de Cadenas de Markov para la simulación Montecarlo

Los resultados de teoría de Cadenas de Markov presentados en esta sección son extraídos de Rincón (2012) y Evans & Rosenthal (2014).

**Teorema 5.3.1** *Ecuación de Chapman-Kolmogorov.* Sea  $\{X_n, n \in \mathbb{N}\}$  una cadena de Markov con espacio de estados discreto y matriz de transición  $\mathbf{P}$ . Entonces para cualquier par de tiempos  $r$  y  $n$ , tal que  $0 \leq r \leq n$ ,  $p_{ij}(n) = \sum_k p_{ij}(r)p_{kj}(n-r)$ .

DEMOSTRACIÓN:

$$p_{ij}(n) = \mathbb{P}(X_n = j | X_0 = i) = \frac{\mathbb{P}(X_n = j | X_0 = i)}{\mathbb{P}(X_0 = i)}.$$

Por el teorema de la probabilidad total:

$$\begin{aligned} \frac{\mathbb{P}(X_n = j | X_0 = i)}{\mathbb{P}(X_0 = i)} &= \sum_k \frac{\mathbb{P}(X_n = j, X_r = k, X_0 = i)}{\mathbb{P}(X_0 = i)} = \\ &= \sum_k \frac{\mathbb{P}(X_n = j | X_r = k, X_0 = i) \mathbb{P}(X_r = k, X_0 = i)}{\mathbb{P}(X_0 = i)} \end{aligned}$$

Dado que  $\{X_n\}$  cumple con la propiedad de Markov, lo anterior es igual a:

$$\begin{aligned} &= \sum_k \mathbb{P}(X_n = j | X_r = k) \mathbb{P}(X_r = k | X_0 = i) = \\ &= \sum_k \mathbb{P}(X_{n-r} = j | X_0 = k) \mathbb{P}(X_r = k | X_0 = i) = \sum_k p_{ij}(r)p_{kj}(n-r) \end{aligned}$$

**Teorema 5.3.2** Si  $\mathbf{P}$  es la matriz de transición en un paso, entonces  $p_{ij}(n) = (\mathbf{P}^n)_{ij}$ .

DEMOSTRACIÓN: utiliza inducción sobre  $\mathbb{N}$ . Para  $n = 1$ ,  $p_{ij}(1) = (\mathbf{P})_{ij}$ . Ahora se supone la proposición cierta para  $n = m$  cualquiera, esto es,  $p_{ij}(m) = (\mathbf{P}^m)_{ij}$ . Entonces se debe demostrar que la afirmación se cumple para  $m + 1$ . Por la ecuación de Chapman-Kolmogorov:

$$p_{ij}(m+1) = \sum_k p_{ik}(1)p_{kj}(m) = \sum_k p_{ik}(1)(\mathbf{P}^m)_{kj} = (\mathbf{P}\mathbf{P}^m)_{ij} = (\mathbf{P}^{m+1})_{ij}$$

**Definición 5.3.1** Sean  $i, j \in S$  dos estados de una Cadena de Markov  $\{X_n, n \in \mathbb{N}\}$ . Entonces se dice que:

- $j$  es accesible desde  $i$  si  $\exists n \geq 0$  tal que  $p_{ij}(n) > 0$ , lo que se representa como  $i \rightarrow j$ .
- $i$  y  $j$  son estados comunicantes si se cumple  $i \rightarrow j$  y  $j \rightarrow i$ , relación que a su vez se simboliza como  $i \leftrightarrow j$ .

**Definición 5.3.2** Se dice que una Cadena de Markov  $\{X_n, n \in \mathbb{N}\}$  con espacio de estados discreto  $S$  es “irreducible” si  $i \leftrightarrow j \forall i, j \in S$ .

**Definición 5.3.3** Sea  $i \in S$ , entonces el periodo de un estado  $i$  se simboliza como  $d(i)$  se define como:

- $d(i) = m.c.d \{n \geq 1, p_{ii}(n) > 0\}$ .
- Se define  $d(i) = 0$  cuando  $p_{ii} = 0 \forall n \geq 1$
- Se dice que un estado  $i \in S$  es “aperiódico” si  $d(i) = 1$ .

**Definición 5.3.4** Se dice que una Cadena de Markov  $\{X_n, n \in \mathbb{N}\}$  con espacio de estados discreto  $S$  es “aperiódica” si  $d(i) = 1 \forall i \in S$ .

El carácter irreducible y la aperiocidad de una Cadena de Markov juega un rol central en el estudio del comportamiento del proceso en el largo plazo. Más en específico, se busca saber si la cadena converge a alguna distribución cuando  $n$  tiende a infinito. De existir, dicha distribución se denomina “estacionaria” para la Cadena.

**Definición 5.3.5** Una distribución de probabilidad  $\pi = (\pi_1, \pi_2, \dots)$  se dice estacionaria para una Cadena  $X_n$  con matriz de transición  $\mathbf{P} = p_{ij}$  si:

$$\pi_j = \sum_i \pi_i p_{ij} \quad \forall j \in S.$$

Matricialmente, la distribución es estacionaria si:  $\pi^t = \pi^t \mathbf{P}$ , vale decir si  $\pi$  es un vector propio de la matriz de transición transpuesta, asociado a un autovalor de 1.

Ahora, del álgebra lineal se sabe que toda matriz cuadrada puede tener más de 1 vector propio, asociados a diferentes valores propios, no necesariamente de valor 1. En el caso de una matriz estocástica puede ocurrir que:

- No existencia: no existe ninguna distribución de probabilidad que cumpla  $\pi^t = \pi^t \mathbf{P}$
- Existencia múltiple
- Existencia única.

Bajo determinadas condiciones, se garantiza que la matriz de transición tiene un sólo vector propio, que es una distribución de probabilidad válida y que el proceso converge a la distribución estacionaria independientemente de la distribución inicial.

**Teorema 5.3.3** Toda Cadena de Markov que sea:

- Irreducible
- Aperiódica
- Recurrente positiva

Tiene una única distribución estacionaria con  $\pi_j = \lim_{n \rightarrow \infty} p_{ij}(n), \forall j \in S$ .

Centrando el foco en la distribución límite, una forma de verificar la estacionaridad consiste en lo siguiente.

**Definición 5.3.6** Una Cadena de Markov es reversible respecto a una distribución  $\{\pi_i\}$  si, para todo  $i, j \in S$  se verifica  $\pi_i p_{ij} = \pi_j p_{ji}$

**Teorema 5.3.4** Si una Cadena de Markov es reversible respecto a  $\{\pi_i\}$ , entonces  $\{\pi_i\}$  resulta ser una distribución estacionaria para la Cadena.

DEMOSTRACIÓN: Con base en la hipótesis de reversibilidad se calcula  $\forall j \in S$ .

$$\sum_{i \in S} \pi_i p_{ij} = \sum_{i \in S} \pi_j p_{ji} = \pi_j \sum_{i \in S} p_{ji} = \pi_j (1) = \pi_j$$

Hasta aquí nos hemos concentrado en Cadenas de Markov definidas en un espacio de estados discreto. Dichas ideas se extienden al caso continuo. Denotando a las probabilidad de transición como  $q(x, y)$ , se requiere:

$$\int q(x, y) dy = 1$$

En tal caso, se obtiene una representación continua, tanto para la ecuación de Chapman-Kolmogorov, la distribución estacionaria del proceso, como así también para la condición de reversibilidad. En el marco del desarrollo contenido en esta sección, dichas ideas son representadas de modo general para el caso discreto.

Con todo, utilizando la reversibilidad como herramienta fundamental se han construido algoritmos que permiten realizar simulación Montecarlo en problemas de alta dimensionalidad paramétrica. El principal se conoce como “algoritmo Metropolis-Hastings”. Estos métodos son utilizados por los modelos bayesianos jerárquicos con aplicabilidad a problemas EI para el caso con  $R$  filas y  $C$  columnas.

### 5.3.2. Relación con Métodos de Inferencia Bayesiana

Los resultados que se presentan en esta sección fueron extraídos de Evans & Rosenthal (2014), Wackerly et al. (2010), Walsh (2004) y desde ejemplos disponibles en la web.

#### 5.3.2.1. Intervalos creíbles de Bayes

Partiendo de la distribución posterior, como posibles estimadores para una función  $\psi(\theta)$  figuran:

- La moda de la distribución posterior:  $\widehat{\psi(\theta)}$
- La media posterior:  $\mathbb{E}[\psi(\theta|\mathbf{x})]$

**Definición 5.3.7** Siendo  $\gamma$  una probabilidad, un intervalo  $C(x)$  plausible de Bayes, con valor inferior  $l(x)$  y valor superior  $u(x)$ , es un intervalo que satisface la condición:

$$\Pi(\psi(\theta) \in C(x)|x) = \Pi(\theta : l(x) \leq \psi(\theta) \leq u(x)|x) \geq \gamma$$

Para el caso en que la distribución posterior sea totalmente continua, el intervalo creíble se obtiene<sup>28</sup> como:

$$\int_{l(x)}^{u(x)} \pi(\theta|x) d\theta = \gamma$$

### 5.3.2.2. Conjugación de Distribuciones

**Definición 5.3.8** Una variable aleatoria  $Y$  tiene distribución Beta con parámetros  $\alpha_1 > 0$  y  $\alpha_2 > 0$  si su función de densidad es la siguiente:

$$f(Y) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} Y^{\alpha_1-1} (1-Y)^{\alpha_2-1}, \quad 0 \leq Y \leq 1$$

Para una variable  $Y$  con distribución Beta de parámetros  $\alpha_1 > 0$  y  $\alpha_2 > 0$  se cumple:

- $E(Y) = \frac{\alpha_1}{\alpha_1 + \alpha_2}$
- $Var(Y) = \frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)^2 (\alpha_1 + \alpha_2 + 1)}$

La distribución de Dirichlet corresponde a la generalización multivariada de la distribución Beta:

**Definición 5.3.9** Se dice que una serie de  $k$  variables aleatorias  $Y_1, \dots, Y_k$  siguen distribución de Dirichlet con parámetros  $\alpha_1 > 0, \dots, \alpha_{k+1} > 0$  si su función de densidad conjunta viene dada por la siguiente expresión:

$$\begin{aligned} f_{Y_1, \dots, Y_k}(Y_1, \dots, Y_k) &= \frac{\Gamma(\alpha_1 + \dots + \alpha_{k+1})}{\Gamma(\alpha_1) \dots \Gamma(\alpha_{k+1})} Y_1^{\alpha_1-1} \dots Y_k^{\alpha_k-1} \left(1 - \sum_{j=1}^k Y_j\right)^{\alpha_{k+1}-1} \\ &= \frac{\Gamma\left(\sum_{j=1}^{k+1} \alpha_j\right)}{\prod_{j=1}^{k+1} \Gamma(\alpha_j)} \left(1 - \sum_{j=1}^k Y_j\right)^{\alpha_{k+1}-1} \prod_{j=1}^k Y_j^{\alpha_j-1}, \quad Y_j \geq 0, \quad j = 1, \dots, k, \quad \sum_{j=1}^k Y_j = 1 \end{aligned}$$

Se demuestra que la distribución marginal de  $Y_j, j = 1, \dots, k$  es Beta con parámetros  $\alpha_j, \alpha_{-j}$  donde  $\alpha_{-j} = 1 - \alpha_j$ . Así:

$$f_{Y_j}(Y_j) \sim Beta(\alpha_j, \alpha_{-j}) \quad (5.3.1)$$

Por tanto se cumplen las expresiones para medias y varianzas para una distribución Beta, de manera tal que:

<sup>28</sup>Al igual que en la construcción de intervalos de confianza en la inferencia muestral, existen infinitas alternativas de formar el intervalo creíble. Una opción general es distribuir el valor  $1 - \gamma$  en igual proporción en ambas colas de la distribución posterior. Otra alternativa se conoce como “intervalos HPD” o intervalos de máxima densidad posterior.

- $\mathbb{E}(Y_j) = \frac{\alpha_j}{\alpha_j + \alpha_{-j}}$
- $Var(Y_j) = \frac{\alpha_j \alpha_{-j}}{(\alpha_j + \alpha_{-j})^2 (\alpha_j + \alpha_{-j} + 1)}$

La siguiente definición es extraída textualmente de Evans & Rosenthal (2014:447).

**Definición 5.3.10** Una familia de distribuciones previas  $\{\pi_\tau : \tau \in \Upsilon\}$  para el parámetro  $\theta$  del modelo  $\{f_\theta : \theta \in \Theta\}$  es conjugada, si para todos los datos  $s \in S$  y todos los  $\tau \in \Upsilon$ , la distribución posterior  $\pi_\tau(\cdot | s)$  pertenece a  $\{\pi_\tau : \tau \in \Upsilon\}$ .

**Teorema 5.3.5** Sea  $X_1, \dots, X_n$  variables aleatorias independientes e idénticamente distribuidas de una población Bernoulli de parámetro  $\theta : \theta \in [0, 1]$ , desconocido. Entonces tomando como distribución previa para  $\theta \sim Beta(\alpha_1, \alpha_2)$  se tiene que la distribución Beta es conjugada para la distribución de Bernoulli.

DEMOSTRACIÓN: Dado que  $\{X_1, \dots, X_n\}$  es una muestra aleatoria de una población Bernoulli( $\theta$ ), la versolimitud viene dada por:

$$L(X_1, \dots, X_n; \theta) = \prod_{i=1}^n f_\theta(X_i) = \prod_{i=1}^n f(X_i | \theta) = \prod_{i=1}^n \theta^{X_i} (1 - \theta)^{1 - X_i} = \theta^{\sum_{i=1}^n X_i} (1 - \theta)^{n - \sum_{i=1}^n X_i} \quad (5.3.2)$$

Por su lado, como previa para el parámetro:

$$\pi(\theta) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \theta^{\alpha_1 - 1} (1 - \theta)^{\alpha_2 - 1} \quad (5.3.3)$$

Luego, la distribución conjunta:

$$\begin{aligned} h(\mathbf{x}, \theta) &= \pi(\theta) f(x_1, x_2, \dots, x_n | \theta) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \theta^{\alpha_1 - 1} (1 - \theta)^{\alpha_2 - 1} \theta^{\sum_{i=1}^n X_i} (1 - \theta)^{n - \sum_{i=1}^n X_i} \\ &= \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \theta^{n\bar{x} + \alpha_1 - 1} (1 - \theta)^{n(1 - \bar{x}) + \alpha_2 - 1} \end{aligned} \quad (5.3.4)$$

A continuación se observa que la distribución marginal de la información muestral es:

$$\begin{aligned} m(\mathbf{x}) &= \int_{\Theta} \pi(\theta) f_\theta(\mathbf{x}) d\theta = \int_0^1 \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \theta^{n\bar{x} + \alpha_1 - 1} (1 - \theta)^{n(1 - \bar{x}) + \alpha_2 - 1} d\theta = \\ &= \frac{\Gamma(\alpha_1 + \alpha_2) \Gamma(n\bar{x} + \alpha_1) \Gamma(n[1 - \bar{x}] + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2) \Gamma(n + \alpha_1 + \alpha_2)} \end{aligned} \quad (5.3.5)$$

En consecuencia, la distribución posterior de  $\theta$  corresponde a:

$$\begin{aligned} \pi(\theta | x_1, \dots, x_n) &= \frac{\pi(\theta) f_\theta(\mathbf{x})}{m(\mathbf{x})} = \frac{\frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \theta^{n\bar{x} + \alpha_1 - 1} (1 - \theta)^{n(1 - \bar{x}) + \alpha_2 - 1}}{\frac{\Gamma(\alpha_1 + \alpha_2) \Gamma(n\bar{x} + \alpha_1) \Gamma(n[1 - \bar{x}] + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2) \Gamma(n + \alpha_1 + \alpha_2)}} \\ &= \frac{\Gamma(n + \alpha_1 + \alpha_2)}{\Gamma(n\bar{x} + \alpha_1) \Gamma(n[1 - \bar{x}] + \alpha_2)} \theta^{n\bar{x} + \alpha_1 - 1} (1 - \theta)^{n(1 - \bar{x}) + \alpha_2 - 1} \end{aligned} \quad (5.3.6)$$

Que resulta ser una distribución Beta con parámetros  $\alpha_1^* = n\bar{x} + \alpha_1$  y  $\alpha_2^* = n[1 - \bar{x}] + \alpha_2$ . Por tanto, se concluye la conjugación.

Un hecho que se destaca con frecuencia en tópicos de inferencia Bayesiana es que la distribución posterior resulta proporcional al producto de la previa y la función de verosimilitud, esto debido a que:

$$\pi(\theta|x_1, \dots, x_n) = \frac{\pi(\theta)f_\theta(x)}{m(x)} = \frac{\pi(\theta)f_\theta(x)}{\int_{-\infty}^{\infty} \pi(\theta)f_\theta(x)d\theta}$$

Como  $\int_{-\infty}^{\infty} \pi(\theta)f_\theta(x)d\theta$  es libre de  $\theta$ , lo anterior se puede representar también como:

$$\pi(\theta|x_1, \dots, x_n) = c(x_1, \dots, x_n) \pi(\theta)f_\theta(x) \Rightarrow \pi(\theta|x_1, \dots, x_n) \propto \pi(\theta)f_\theta(x) \quad (5.3.7)$$

Aquí  $c(x_1, \dots, x_n) = \int_{-\infty}^{\infty} \pi(\theta)f_\theta(x)d\theta$  y opera como constante inversa de normalización. Con base en este resultado, se tiene el siguiente teorema.

**Teorema 5.3.6** Sean  $X_1, \dots, X_n$  variables aleatorias con distribución conjunta Multinomial de parámetro  $\{\theta_1, \dots, \theta_k\}$ , tal que  $\sum_{i=1}^k \theta_i = 1$ , y que como distribución previa para  $\{\theta_1, \dots, \theta_{k-1}\}$  se adopta  $\sim$  Dirichlet  $(\alpha_1, \dots, \alpha_k)$ ,  $\alpha_i \geq 0 \forall i = 1, \dots, k$ . Entonces se tiene que la distribución de Dirichlet es conjugada para la distribución Multinomial.

DEMOSTRACIÓN: Una versión equivalente de la verosimilitud Multinomial viene dada por  $\prod_{j=1}^k \theta_j^{x_j}$ . Como  $\{\theta_1, \dots, \theta_{k-1}\} \sim$  Dirichlet  $(\alpha_1, \dots, \alpha_k)$ , la distribución conjunta  $h(\mathbf{x}, \theta)$  resulta en:

$$\begin{aligned} h(\mathbf{x}, \theta) &= \frac{\Gamma(\alpha_1 + \dots + \alpha_k)}{\Gamma(\alpha_1)\dots\Gamma(\alpha_k)} \theta_1^{\alpha_1-1} \dots \theta_{k-1}^{\alpha_{k-1}-1} \left(1 - \sum_{j=1}^{k-1} \theta_j\right)^{\alpha_k-1} \prod_{j=1}^k \theta_j^{x_j} \\ &= \frac{\Gamma(\alpha_1 + \dots + \alpha_k)}{\Gamma(\alpha_1)\dots\Gamma(\alpha_k)} \theta_1^{x_1+\alpha_1-1} \dots \theta_{k-1}^{x_{k-1}+\alpha_{k-1}-1} \left(1 - \sum_{j=1}^{k-1} \theta_j\right)^{x_k+\alpha_k-1} \end{aligned} \quad (5.3.8)$$

Que, como función de  $\theta$  es proporcional entonces a:

$$\pi(\theta|x_1, \dots, x_n) \propto \theta_1^{x_1+\alpha_1-1} \dots \theta_{k-1}^{x_{k-1}+\alpha_{k-1}-1} \left(1 - \sum_{j=1}^{k-1} \theta_j\right)^{x_k+\alpha_k-1}$$

Se concluye entonces que la distribución posterior  $\pi(\theta|x_1 \dots x_n)$  es Dirichlet de parámetro  $\{\alpha_1 + x_1, \dots, \alpha_k + x_k\}$  lo cual demuestra la conjugación.

### 5.3.2.3. Modelos Bayesianos Jerárquicos

Surgen para el tratamiento de datos con estructura de covarianza compleja, por ejemplo, cuando las observaciones muestrales no pueden asumirse independientes. A estos efectos, se añade dos componentes adicionales al modelo bayesiano:

1. Supuesto de independencia condicional: dado un conjunto de observaciones  $\{y_1, \dots, y_K\}$ , la distribución de  $y_k|\theta_k \stackrel{\perp}{\sim} p(\theta_k)$ ,  $k = 1, \dots, K$ .
2. Distribuciones hiper-previas: siendo la distribución previa y la posterior los dos primeros niveles del modelo, se añaden una o más distribuciones para modelar la pérdida de conocimiento previo sobre el valor del parámetro, originada por el modelamiento de observaciones con estructura de covarianza complejas.

### Ejemplo 5.3.1

El siguiente ejemplo fue obtenido de un material disponible en la web<sup>29</sup> e involucra un problema donde las observaciones son el número de decesos  $y_k$  producidos por una enfermedad cancerígena observados en una unidad demográfica  $\{y_1, \dots, y_K\}$ ,  $k = 1, \dots, K$  y un intervalo temporal determinado. Siendo:

- $y_k$ : número de decesos observados en localidad  $k$  durante el periodo de observación.
- $n_k$ : número de personas residentes en la localidad  $k$  dividido por 100.000.
- $\theta_k$ : valor esperado de decesos en localidad  $k$  por cada 100 mil residentes.

Las observaciones  $\{y_1, \dots, y_K\}$  no pueden asumirse independientes. Para abordar el problema de la estimación del valor esperado  $\theta$  de decesos en la población que agrupa las diferentes unidades demográficas se adopta un modelo bayesiano jerárquico. El modelo adopta como supuestos:

$$y_k \stackrel{\perp}{\sim} \text{Poisson}(n_k \theta_k).$$

$$\theta_k \stackrel{i.i.d}{\sim} \text{Gamma}(\alpha, \beta)$$

$$\alpha \sim \text{Unif}(0, a_0)$$

$$\beta \sim \text{Unif}(0, b_0)$$

Las distribuciones de  $\alpha$  y  $\beta$  corresponden en este caso a las “hiper-previas”. Se asume que  $\alpha$  y  $\beta$  son independientes, para algunos valores no nulos de  $a_0$  y  $b_0$ .

La distribución posterior aquí es:

$$\begin{aligned} \mathbb{P}(\theta, \alpha, \beta | y) &\propto \prod_{k=1}^K [p(y_k | n_k \theta_k) p(\theta_k | \alpha, \beta)] p(\alpha) p(\beta) \\ &\propto \prod_{k=1}^K \left[ e^{-n_k \theta_k} \theta_k^{y_k} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta_k^{\alpha-1} e^{-\theta_k \beta} \right] I(0 < \alpha < a_0) I(0 < \beta < b_0) \end{aligned}$$

<sup>29</sup>Ver: <https://www.youtube.com/watch?v=RJjcBwRiR38&t=1919s>

#### 5.3.2.4. Algoritmo Metropolis within Gibbs

Como se puede apreciar en el ejemplo anterior, como consecuencia de la adopción del supuesto de independencia condicional, los modelos bayesianos jerárquicos involucran un alto número de parámetros, uno o más por observación. Dado que bajo este enfoque los parámetros en sí mismos se modelan como variables aleatorias, y que los estimadores de Bayes se basan principalmente en la media de la distribución posterior, una vía para obtener las estimaciones pretendidas puede consistir en recurrir a métodos Montecarlo desde la distribución conjunta posterior.

**5.3.2.4.1. Muestreador de Gibbs.** El muestreador de Gibbs es un algoritmo para obtener estimaciones desde una distribución conjunta de difícil evaluación con métodos de simulación estándar.

De acuerdo con la descripción incluida en Evans y Rosenthal (2014: 438), el problema de partida consiste en pretender generar muestras de una distribución conjunta de variables aleatorias  $(X_1, \dots, X_k) \in \mathbb{R}^k$ . El problema involucra asimismo la posibilidad de generar cada una de las distribuciones condicionales completas  $X_i | X_{-i} = x_{-i}$ , donde  $X_{-i} = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_k)$ . A continuación el algoritmo procede de la siguiente forma:

1. Especifica un valor inicial  $(x_{1(0)}, \dots, x_{k(0)})$  para  $(X_1, \dots, X_k)$ .
2. Para  $N > 0$  y para cada  $i = 1, \dots, k$  genera  $X_{i(N)}$  sobre la base de la distribución condicionada.

$$(x_{1(N)}, \dots, x_{i-1(N)}, x_{i+1(N-1)}, \dots, x_{k(N-1)}) .$$

Bajo condiciones generales de imparcialidad se demuestra (Evans & Rosenthal, 2014: 439) que  $(x_{1(N)}, \dots, x_{k(N)})$  converge en distribución a la distribución conjunta de  $(X_1, \dots, X_k)$  cuando  $N \rightarrow \infty$ .

#### *Ejemplo 5.3.2*

Supóngase que se quiere obtener una observación de la distribución conjunta dada por:

$$f(x, h) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \sqrt{\frac{\lambda}{2\pi}} h^{\alpha - \frac{1}{2}} \exp\left(\frac{-h}{b} - \frac{\lambda h((x - \mu)^2)}{2}\right)$$

La labor de simulación se ve obstaculizada, al tratarse de una distribución con rutinas no predefinida en softwares estadísticos. Como alternativa a ello se observa que:

$$h|x \sim \text{Gamma}(\alpha, \beta)$$

$$x|h \sim N\left(\mu, \frac{1}{\lambda h}\right)$$

Para valores  $\alpha$  y  $\beta$  definidos. Entonces, la simulación de  $h_{(1)}$  permite luego obtener por simulación un valor  $x_{(1)}$ , y éste a su vez  $h_{(2)}$ . Como resultado de iterar el algoritmo, bajo condiciones de imparcialidad  $(x_{(N)}, h_{(N)})$  converge en distribución a  $f(x, h)$  cuando  $N \rightarrow \infty$ .

### Ejemplo 5.3.3

Continuando con el ejemplo de modelo bayesiano jerárquico dado en la sección precedente, para la distribución conjunta  $\mathbb{P}(\theta, \alpha, \beta|y)$  puede descomponerse en distribuciones marginales (o “full conditionals”):

Para  $\theta_k$ :  $p(\theta_k|y, \theta_{-k}, \alpha, \beta) \propto p(\theta, \alpha, \beta|y)$

$$\propto e^{-n_k \theta_k} \theta_k^{y_k} \theta_k^{\alpha-1} e^{-\theta_k \beta} = \theta_k^{y_k + \alpha - 1} e^{-\theta_k (n_k + \beta)} \propto \text{Gamma}(y_k + \alpha, n_k + \beta)$$

Para  $\alpha$ :  $p(\alpha|y, \theta, \beta) \propto p(\theta, \alpha, \beta|y)$

$$\propto \prod_{k=1}^K \left[ \theta_k^{\alpha-1} \frac{\beta^\alpha}{\Gamma(\alpha)} \right] I(0 < \alpha < a_0) = \left( \prod_{k=1}^K \theta_k \right)^\alpha \beta^{n\alpha} \Gamma(\alpha)^{-K} I(0 < \alpha < a_0)$$

Y finalmente, para  $\beta$ :  $p(\beta|y, \theta, \alpha) \propto p(\theta, \alpha, \beta|y)$

$$\propto \prod_{k=1}^K [e^{-\theta_k \beta} \beta^\alpha] I(0 < \beta < b_0) = \beta^{n\alpha} e^{-\beta \sum_{k=1}^K \theta_k} I(0 < \beta < b_0)$$

En este caso, para implementar el algoritmo del muestreador de Gibbs se requeriría proceder de la siguiente manera:

1. Muestrear  $\theta^{(j+1)} \sim p(\theta|y, \alpha^{(j)}, \beta^{(j)})$ 
  - Para  $k = 1, \dots, K$  muestrear  $\theta_k^{(j+1)} \sim p(\theta_k|y_k, \alpha^{(j)}, \beta^{(j)})$
2. Muestrear  $\alpha^{(j+1)} \sim p(\alpha|\theta^{(j+1)}, \beta^{(j)})$
3. Muestrear  $\beta^{(j+1)} \sim p(\beta|\theta^{(j+1)}, \alpha^{(j+1)})$

No obstante, para implementar este algoritmo se requiere expresar la distribuciones marginales como distribuciones conocidas, que se encuentren implementadas en softwares estadísticos. En el ejemplo esto no resulta factible, dado que  $p(\alpha|y, \theta, \beta)$  no es una distribución conocida, mientras que  $p(\beta|y, \theta, \alpha)$  se identifica como una distribución Gamma truncada, de difícil implementación. Es en este tipo de problemas que entra a jugar el algoritmo Metropolis-Hastings para la simulación Montecarlo con Cadenas de Markov.

**5.3.2.4.2. Algoritmo Metropolis-Hastings.** Construye una cadena de Markov que tiene como distribución estacionaria una distribución objetivo, de difícil implementación en algoritmos de simulación preestablecidos, como por ejemplo en el caso anterior  $p(\alpha|y, \theta, \beta)$ .

Como fue señalado con anterioridad, se ejemplifica el método general con el caso discreto. El algoritmo inicia en un espacio de estados  $S$  y una distribución de probabilidad  $\{\pi_i\}$  definida sobre  $S$ , que corresponde a la distribución de la que se pretende muestrear. A continuación se seleccionan probabilidades de transición  $\{q_{ij} : i, j \in S\}$ , tales que  $q_{ij} \geq 0$  y  $\sum_{j \in S} q_{ij} = 1, \forall i$ , la que se denomina “distribución propuesta”.

Para un valor  $X_n = i$  el algoritmo obtiene  $X_{n+1} = j$  de la siguiente manera:

1. Selecciona  $Y_{n+1} = j$  según las probabilidades de transición  $\{q_{ij}\}$ .
2. Establece como probabilidad de aceptación  $\alpha_{ij} = \min \left\{ 1, \frac{\pi_j q_{ji}}{\pi_i q_{ij}} \right\}$ .
3. Con probabilidad  $\alpha_{ij}$ , acepta  $X_{n+1} = Y_{n+1} = j$ , caso contrario rechaza el valor propuesto  $Y_{n+1}$  con probabilidad  $1 - \alpha_{ij}$  y establece  $X_{n+1} = i$ , vale decir, permanece en el estado  $i$ .

**Teorema 5.3.7** *El algoritmo Metropolis-Hastings conduce en una cadena de Markov  $\{X_n\}$  a que tenga a  $\{\pi_i\}$  como distribución estacionaria.*

DEMOSTRACIÓN:

$$\begin{aligned} \mathbb{P}(X_{n+1} = j | X_n = i) &= q_{ij} \alpha_{ij} = q_{ij} \min \left\{ 1, \frac{\pi_j q_{ji}}{\pi_i q_{ij}} \right\} = \min \left\{ q_{ij}, \frac{\pi_j q_{ji}}{\pi_i} \right\} \Rightarrow \\ &\Rightarrow \pi_i \mathbb{P}(X_{n+1} = j | X_n = i) = \min \{ \pi_i q_{ij}, \pi_j q_{ji} \} \end{aligned} \quad (5.3.9)$$

De modo análogo:

$$\begin{aligned} \mathbb{P}(X_{n+1} = i | X_n = j) &= q_{ji} \alpha_{ji} = q_{ji} \min \left\{ 1, \frac{\pi_i q_{ij}}{\pi_j q_{ji}} \right\} = \min \left\{ q_{ji}, \frac{\pi_i q_{ij}}{\pi_j} \right\} \Rightarrow \\ &\Rightarrow \pi_j \mathbb{P}(X_{n+1} = i | X_n = j) = \min \{ \pi_j q_{ji}, \pi_i q_{ij} \} \end{aligned} \quad (5.3.10)$$

De esta manera se llega a que:

$$\pi_i \mathbb{P}(X_{n+1} = j | X_n = i) = \min \{ \pi_i q_{ij}, \pi_j q_{ji} \} = \pi_j \mathbb{P}(X_{n+1} = i | X_n = j) \quad (5.3.11)$$

Así, se tiene que  $\{\pi_i\}$  es reversible para la Cadena de Markov  $\{X_n\}$ , y por consiguiente, según el Teorema de la distribución reversible, es una distribución estacionaria para el proceso.

Cabe hacer notar que la estacionariedad proviene, no de la distribución propuesta  $\{q_{ij}\}$ , sino de la combinación de ésta con la regla de aceptación  $\alpha_{ij} = \min \left\{ 1, \frac{\pi_j q_{ji}}{\pi_i q_{ij}} \right\}$ . De hecho, como se señala en Evans y Rosenthal la distribución  $\{q_{ij}\}$  puede incluso no tener distribución estacionaria (2014: 678).

El algoritmo Metropolis-Hastings es la generalización del algoritmo originalmente denominado Metropolis, en la cual se requiere que la distribución propuesta sea simétrica, esto es:  $q_{ij} = q_{ji}, \forall i, j \in S$ .

Con todo, ejecutar la Cadena de Markov generada con el algoritmo en un periodo suficientemente extenso para alcanzar la convergencia, permite obtener una observación  $X_i$  de la distribución objetivo. Al repetir el proceso en las mismas condiciones  $N$  veces permitirá obtener una muestra *i.i.d.* desde dicha distribución. Luego, por las leyes de los grandes números:

$$\frac{1}{N} \sum_{i=1}^N X_i \xrightarrow{P} \mathbb{E}[X], \text{ cuando } N \rightarrow \infty.$$

Donde  $X$  es una variable con distribución  $\{\pi_i\}$ .

La teoría de Cadenas de Markov y el algoritmo Metropolis-Hastings se han desarrollado en esta descripción principalmente tomando como referencia el caso discreto. No obstante, el algoritmo aplica asimismo al caso de variables aleatorias continuas. El siguiente ejemplo es extraído nuevamente de Evans y Rosenthal (2014: 680-681).

**Ejemplo 5.3.4**

Se pretende generar una muestra *i.i.d.* de la distribución con función de densidad  $Cf(y)$ , donde:

$$f(y) = e^{-y^4}(1 + |y|)^3 \text{ y } C = 1 / \int_{-\infty}^{\infty} f(y) dy.$$

Para generar una Cadena de Markov que tenga esta distribución estacionaria se utiliza como distribución propuesta una distribución normal  $N(x, 1)$  con media  $x$  y varianza igual a 1. Dado  $X_n = x$ , se selecciona  $Y_{n+1}$  como  $Y_{n+1} \sim N(x, 1)$ . Así, la función de densidad propuesta:

$$q(x, y) = (2\pi)^{-1/2} e^{-(y-x)^2/2}$$

Por su parte, como probabilidad de aceptación:

$$\begin{aligned} \alpha(x, y) &= \min \left\{ 1, \frac{Cf(y)q(x, y)}{Cf(x)q(x, y)} \right\} \\ &= \min \left\{ 1, \frac{Ce^{-y^4}(2\pi)^{-1/2}e^{-(y-x)^2/2}}{Ce^{-x^4}(2\pi)^{-1/2}e^{-(x-y)^2/2}} \right\} \\ &= \min \left\{ 1, e^{-y^4+x^4} \right\} \end{aligned}$$

Dado  $X_n = x$  el algoritmo calcula  $X_{n+1}$  de la siguiente manera:

1. Genera  $Y_{n+1} \sim N(x, 1)$ .
2. Establece  $y = Y_{n+1}$  y calcula  $\alpha(x, y)$ .
3. Con probabilidad  $\alpha(x, y)$  establece  $X_{n+1} = Y_{n+1} = y$ . En cualquier otro caso, con probabilidad  $1 - \alpha(x, y)$  establece  $X_{n+1} = X_n = x$ .

Una alternativa a ejecutar  $N$  cadenas independientes, consiste en implementar sólo una, y evaluar el comportamiento de la misma como serie de tiempo, para así obtener una muestra desde aquella. En tal caso, se requiere evaluar, tanto la convergencia del proceso a la distribución estacionaria, como así también la función de autocorrelación del proceso.

**5.3.2.4.3. Evaluación de la Cadena de Markov obtenida como Serie de Tiempo.** Evidencia a favor de la convergencia del método se obtiene cuando la cadena se exhibe como un proceso de “Ruido blanco”, en tal caso se dice que el proceso se encuentra bien mezclado (“well mixed”), descartándose las observaciones previas. El periodo previo a la convergencia se denomina “burn-in”. Otros procedimientos para evaluar la convergencia son descritos en Walsh (2004: 15), Congdon (2010: 20-26) y Tanner (1996: 152-161).

Una desventaja asociada a ejecutar una sola cadena es que no se obtienen observaciones independientes, sino que éstas se encontrarán correlacionadas. No obstante, de acuerdo con un importante resultado de la teoría de Series de Tiempo, si las observaciones provienen de un proceso estacionario de carácter autorregresivo, realizaciones correlacionadas proveerán aún una representación insesgada de atributos de la distribución subyacente, si el tamaño de muestra es suficientemente grande. Más en concreto, para un proceso autorregresivo de orden 1 (AR1):

$$X_t = \mu + \phi_1 (X_{t-1} - \mu) + \epsilon, \text{ donde } \epsilon \text{ es un proceso de ruido blanco } \epsilon \sim N(0, \sigma^2)$$

En este caso la función de autocorrelación  $\rho_k = \rho_1^k$ , donde  $\rho_1 = \phi_1$ . Bajo estas condiciones:

$$\mathbb{E}[\bar{X}_n] = \mathbb{E}[X_t] = \mu, \text{ ya que } \mathbb{E}[X_t - \mu] = 0, \text{ donde } \bar{X}_n = \frac{1}{n} \sum_{t=1}^n X_t.$$

Con error estándar  $SE(\bar{X}_n)$ :

$$SE(\bar{X}_n) = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{1 + \rho_1}{1 - \rho_1}}$$

La librería eiPack del software libre R implementa este modelo. Establece las primeras 1.000 observaciones como burn-in, predefine a partir de entonces la cadena como un proceso Autoregresivo de orden 1 (AR1) (thin), y se obtiene una muestra de tamaño 1.000 (sample size). No obstante, los dos primeros parámetros, así como el tamaño de la muestra obtenida con el método, pueden modificarse en función de la examinación de la serie de tiempo obtenida.

## 5.4. El enfoque de la Institucionalización partidista de Mainwaring & Scully

Mainwaring & Scully (1995), prestando especial atención a las recurrentes dificultades que las democracias de la región han experimentado desde el siglo XIX a la hora de establecer una democracia efectiva, propusieron centrar el estudio del funcionamiento de los sistemas políticos latinoamericanos en torno al grado de asentamiento de las prácticas asociadas a la vigencia de un sistema democrático afianzado. En continuidad con los planteamientos de Huntington (1968), el énfasis está puesto en aspectos como la legitimidad del régimen democrático, la efectividad de las normas y el grado de arraigo de las prácticas democráticas en tanto factores determinantes de la calidad de la democracia. Sobre la base de lo anterior, Mainwaring & Scully identificaron cuatro ámbitos de relevancia para la dinámica de la institucionalización partidista en los sistemas de la región.

- Estabilidad de las reglas y de la dinámica de competencia entre partidos.
- Enraizamiento de los partidos en la sociedad, lo cual quiere decir que los partidos mantienen una base de apoyo electoral relativamente estable en el tiempo.
- El tercer componente consiste en el hecho de que los actores políticos otorgan legitimidad al proceso electoral y a los partidos, dando relevancia a las dinámicas eleccionarias como vía exclusiva de acceso del poder.
- Finalmente, el cuarto componente se refiere a la autonomía organizacional de los partidos, de modo que aquellos no se ven supeditados a intereses de líderes ambiciosos, sino que, por el contrario, cuentan con un estatus independiente y valor propio.

Tras esta propuesta es posible reconocer tres propósitos. En primer lugar se pone de relieve la necesidad de llevar a cabo un estudio de la fortaleza institucional de los regímenes democráticos latinoamericanos, como condición previa a su análisis comparativo con otros sistemas, tales como los europeos<sup>30</sup>. Por ejemplo, la comparación de aspectos tales como el *número efectivo de partidos* (NEP) o el grado de *polarización ideológica* puede llevar a conclusiones artificiales si no se reconoce el grado en que uno y otro sistema se encuentran institucionalizados (Mainwaring, 1999). En consecuencia, el grado de institucionalización del sistema democrático debe constituir el foco preferente con miras al estudio de la *calidad de la democracia* en los sistemas latinoamericanos. En segundo término, se pretende desarrollar un enfoque de investigación comparada en el análisis de la calidad de la democracia en América latina, tarea que según los autores se encontraba escasamente explorada hasta la fecha. En tercer lugar, se persigue introducir el estudio de la institucionalización de los sistemas de partidos como un continuo, a diferencia de clasificación dicotómica previamente elaborada por Sartori (1976) entre “sistema” y “no sistema”.

La importancia de la institucionalización ha sido refrendada en varios trabajos posteriores. En torno a ello se ha señalado que un sistema de partidos institucionalizado favorece una representación programática efectiva y otorga un marco de certidumbre al funcionamiento del sistema democrático (Torcal, 2015). En cambio, en sistemas no institucionalizados los partidos tienen escasas raíces en la sociedad; son fluidos, lo cual origina un alto marco de inestabilidad; algunos partidos aparecen, mientras otros se evaporan;

---

<sup>30</sup>Que junto con los canadienses y estadounidenses, son denominados “sistemas avanzados”.

existen escasas barreras de entrada para contener la irrupción de personalismos; tienden a primar lógicas populistas por sobre una representación programática; y en definitiva, son en mayor medida susceptibles a la ocurrencia de (nuevos) quiebres institucionales.

A partir de su divulgación, la propuesta analítica de Mainwaring & Scully se ha mostrado altamente influyente como marco interpretativo general del estudio de los sistemas democrático latinoamericanos, aun cuando no ha estado exenta de críticas. Artiga (1998) propuso diferenciar a la “fluidez” de la volatilidad electoral, entendiendo a la primera como el grado en que los partidos permanecen a lo largo del tiempo en la competencia electoral. A su juicio *“no diferenciar estas dos facetas del mismo problema (la institucionalización de los sistemas de partidos) puede conducir a errores metodológicos o volver redundantes algunos análisis”* (1998: 125). Lo cierto es que la introducción del concepto de fluidez puso de relieve cierta ambigüedad de la transición operacional de la dimensión “estabilidad”, que en su versión inicial se formulaba como “estabilidad de las reglas y en la naturaleza de la competición entre partidos”, vale decir, no sólo regularidad de los resultados electorales.

Un segundo ámbito que ha sido objeto de críticas se refiere a la rigidez heurística originada en la adopción del supuesto de interrelación entre las dimensiones que estructuran la conceptualización de sistema democrático institucionalizado, precepto que da lugar a que el caso de regímenes electoralmente estables, pero desenraizados del electorado, adquieran un carácter paradójico. Luna & Altman (2011), así como Zucco (2014), coinciden en utilizar el término “estabilidad sin raíces” o “estabilidad no institucionalizada” para denotar la coexistencia de bajos índices de volatilidad electoral, con crecientes niveles de desconexión del sistema de partidos con la ciudadanía. Como señalan Luna & Altman: *“the current structure of the concept implies an “all good things go together” assumption, and thus neglects potential trade-offs between different dimensions of the concept. Electorally stable but otherwise weak party systems are theoretically neglected in the structure of the concept. Intertemporal trade-offs between dimensions are also obscured”* (2011: 23). Referido al mismo asunto, Zucco señala: *“la relación entre una mayor institucionalización y el mejor funcionamiento de las democracias podría muy bien existir en promedio y en todos los sistemas de gobierno aunque no pueda apreciarse en un país concreto. Sin embargo, el estudio de esta relación en el caso brasileño sirve para recordarnos que la misma debe probarse empíricamente, y no presuponerse desde una visión puramente normativa”* (Zucco, 2014: 17).

Pese a la relevancia de este último tema, en lo posterior el enfoque de la institucionalización se ha enfocado en el problema enunciado inicialmente por el análisis de Artiga, y relevado con posterioridad a escala internacional por Powell & Tucker (2013), respecto a la discontinuidad de los agentes de representación existente en un sistema democrático. Powell & Tucker diferenciaron la volatilidad “tipo B”, que definieron como aquella que *“occurs when voters switch their votes between existing parties”*, de la volatilidad “tipo A”, que es originada a partir de la entrada y/o salida de partidos o bloques desde el sistema político (Powell & Tucker, 2013: 123). Una clasificación similar fue propuesta por Mainwaring et al. (2017), distinguiendo la *“extra system electoral volatility”* –equivalente a la volatilidad tipo A en la anterior nomenclatura–, de la volatilidad *“within system”* (tipo B). El problema de la estabilidad desenraizada, en cambio, no fue objeto de un análisis más extendido. Es esta desconsideración aquello que motivó la realización de este proyecto.

# Bibliografía

- [1] Agüero, Ysela. 1998. “Estimadores norma  $L_p$  en regresión lineal”. *Revista del Instituto de Investigación de la Facultad de Ciencias Matemáticas UNMSM*, 1(1): 91-102.
- [2] Artiga, Álvaro. 1998. “Fluidez y volatilidad en la institucionalización de los sistemas de partidos (notas metodológicas para su medición)”. *América latina hoy* (19): 125-129.
- [3] Andreadis, Ioannis y Theodore G. Chadjipadelis. 2009. “A method for the estimation of voter transition rates”. *Journal of Elections Public Opinion and Parties* 19(2): 203–218.
- [4] Avendaño, Octavio y Pablo Sandoval. 2016. “Desafección política y estabilidad de los resultados electorales en Chile, 1993-2009”. *Perfiles Latinoamericanos* 24(47): 175-198.
- [5] Avendaño, Octavio y Pablo Sandoval. 2013. “Volatilidad electoral y desafección política en las elecciones parlamentarias chilenas 1993-2009”. *Revista de Sociología de la Universidad de Chile* 28: 141-164.
- [6] Bartolini, Stefano y Peter Mair. 1990. *Identity, competition, and electoral availability. The stabilisation of European electorates 1885-1985*. Cambridge: Cambridge University Press.
- [7] Bizarro, Fernando, Allen Hicken y Darin Self. 2017. *The V-Dem Party Institutionalization Index: a new global indicator (1900-2015)*, Working Paper Series 2017:48. The Varieties of democracy Institute.
- [8] Casal Bértoa, Fernando. 2018. “The Three Waves of Party System Institutionalisation Studies: A Multi- or Uni-Dimensional Concept?”. *Political Studies Review* 16(1): 60-72.
- [9] Casal Bértoa, Fernando, Kevin Deegan-Krause y Tim Haughton. 2017. “The volatility of volatility: Measuring change in party vote shares”. *Electoral studies* 50: 142:156.
- [10] Chiaramonte, Alessandro y Vincenzo Emanuele. 2017. “Party system volatility, regeneration and de-institutionalization in Western Europe (1945–2015)”. *Party politics* 23(4): 376-388.
- [11] Congdon, Peter. 2010. *Applied Bayesian Hierarchical Methods*, New York: CRC Press, Taylor and Francis Group.
- [12] Corominas, Albert, Amaia Lusa y M Dolors Calvet. 2015. “Computing voter transitions: The elections for the Catalan Parliament, from 2010 to 2012”. *Journal of Industrial Engineering and Management*, 8(1): 122-136.
- [13] Crewe, Ivor y David Denver. 1985. *Electoral change in western democracies*, London: Croom Helm.

- [14] Dassonneville, Ruth y Marc Hooghe. 2011. *Mapping Electoral Volatility in Europe. An analysis of trends in electoral volatility in European democracies since 1945*, Leuven: KU Leuven Centre for Political Research.
- [15] Duncan, Otis Dudley y Beverly Davis. 1953. “An alternative to ecological correlation”. *American Sociological Review* 18(6): 665-666.
- [16] Evans, Michael J. y Jeffrey S. Rosenthal. 2014. *Probabilidad y Estadística. La ciencia de la incertidumbre*, Barcelona: Reverté.
- [17] Forcina, Antonio y Davide Pellegrino. 2019. “Estimation of voter transitions and the ecological fallacy”, *Quality and Quantity* 53: 1859–1874.
- [18] Freedman, David A. 2001. “Ecological Inference and the Ecological Fallacy” Pp. 4027-4030 en *International encyclopedia of the social and behavioral sciences*, editado por Neil J. Smelser y Paul B. Baltes, New York: Elsevier, Vol. 6.
- [19] Frogner, Charlie y Tomasso Poggio. 2019. “Fast and Flexible Inference of Joint Distributions from their Marginals”, *Proceedings of the 36th International Conference on Machine Learning*, PMLR 97: 2002-2011.
- [20] Goodman, Leo. 1959. “Some alternatives to ecological correlation”. *American Journal of Sociology* 64(6): 610-625.
- [21] Goodman, Leo. 1953. “Ecological regression and behaviour of individuals”. *American Sociological Review* 18(6): 663-664.
- [22] Gregor, Kamil. 2014. “Comparing the structure of voting transitions around the presidential elections in the Czech Republic and Slovakia”. *European Electoral Studies* 9(1): 3–20.
- [23] Greiner, D. James y Kevin M. Quinn. 2009. “RxC ecological inference: bounds, correlations, flexibility and transparency of assumptions”. *Journal of the Royal statistical society* 172, part.1: 67-81.
- [24] Horowitz, Joel L. y Charles F. Manski. 1995. “Identification and Robustness with contaminated and corrupted data”. *Econometrica* 63(2): 281-302.
- [25] Huntington, Samuel P. 1968. *Political order in changing societies*. New Haven: Yale University Press.
- [26] Johnson, Norman L., Samuel Kotz y Narayanaswamy Balakrishnan. 1997. *Discrete multivariate distributions*. John Wiley & Sons Inc.
- [27] Joignant, Alfredo y Miguel A. López. 2005. “Le comportement électoral au Chili: paradoxes et pré-somptions sur la continuité ou la rupture de l’orientation du vote”. *Problèmes d’Amérique Latine* 56: 63-80.
- [28] Jones, Mark. 2005. “The role of parties and party systems in the policymaking process”. Presentado en State reform, Washington D.C.: Banco Interamericano de Desarrollo. Febrero-Marzo de 2005.
- [29] King, Gary. 1997. *A solution to the ecological inference problem. Reconstructing individual behaviour from aggregate data* Princeton, NJ: Princeton University.

- [30] King, Gary, Ori Rosen y Martin A. Tanner. 1999. “Binomial-Beta Hierarchical Models for Ecological Inference”. *Sociological Methods and Research* 28(1): 61-90.
- [31] King, Gary, Ori Rosen y Martin A. Tanner. 2004. *Ecological inference. New methodological strategies* New York: Cambridge University Press.
- [32] Klima, André, Paul W. Thurner, Cristoph Molnar, Thomas Schlesinger y Herlmut Küchenhoff. 2016. “Estimation of voter transitions based on ecological inference: an empirical assessment of different approaches”. *Advances in statistical analysis* 100: 133-159.
- [33] Klima, André, Thomas Schlesinger, Paul W. Thurner y Herlmut Küchenhoff. 2019. “Combining Aggregate Data and Exit Polls for the Estimation of Voter Transitions”. *Sociological Methods and Research* 48(2): 296-325.
- [34] Lau, O., R. T. Moore y M. Kellermann. 2007. “eiPack: RxC ecological inference and higher-dimension data management”. *R News* 7(2): 43–47.
- [35] Luna, Juan Pablo. 2014. “Party system institutionalization. Do we need a new concept?”. *Studies in Comparative International Development* 49(4): 403-425.
- [36] Luna, Juan Pablo. 2008. “Partidos políticos y sociedad en Chile. Trayectoria histórica y mutaciones recientes”. Pp. 75-124 en *Reforma de los partidos políticos en Chile*, editado por Jorge Navarrete, Ignacio Walker, Arturo Fontaine, Cristián Larroulet. Santiago, Chile: CEP-PNUD.
- [37] Luna, Juan Pablo y David Altman. 2011. “Uprooted but stable. Chilean party system and the concept of party system institutionalization”. *Latin America Politics and Society* 56(2): 1-28.
- [38] Luna, Juan P. y Fernando Rossemblat. 2012. “¿Notas para una autopsia? Los partidos políticos en el Chile actual”. Pp. 115-265 en *Democracia con Partidos, Informe para la reforma de los partidos políticos en Chile*, editado por Díaz, Francisco y Lucas Sierra, Santiago de Chile: CEP/CIEPLAN.
- [39] Mainwaring, Scott. 1999. *Rethinking party systems in the third wave of democratization: The case of Brazil*, Stanford University Press.
- [40] Mainwaring, Scott, Carlos Gervasoni y Annabella España-Najera. 2017. “Extra- and within-system electoral volatility”. *Party politics* 23(6): 623-635.
- [41] Mainwaring, Scott y Timothy Scully. 1995. *Building Democratic Institutions: Party Systems in Latin America*. Stanford University Press.
- [42] Mainwaring, Scott y Mariano Torcal. 2006. “Party system institutionalization and party system theory after the third wave of democratization”. En *Handbook of party politics* editado por Richard S. Katz y William J. Crotty, London: Sage.
- [43] Mainwaring, Scott y Eduarne Zoco. 2007. “Political sequences and stabilization of interparty competition: electoral volatility in old and new democracies”. *Party politics* 13(152): 155-178.
- [44] Manski, Charles F. 2018. “Credible ecological inference for medical decisions with personalized risk assessment”. *Quantitative Economics* 9: 541-569.

- [45] Muzellec, Boris, Richard Nock, Giorgio Patrini y Frank Nielsen. 2017. “Tsallis Regularized Optimal Transport and Ecological Inference”. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*.
- [46] Noorudin, Irfan y Pradeep Chhibber. 2008. “Unstable Politics. Fiscal Space and Electoral Volatility in the Indian States”. *Comparative Political Studies* 41(8): 1069-1091.
- [47] Ocaña, Francisco. 2007. “An approximation problem in computing electoral volatility”. *Applied Mathematics and Computation* 192(2): 299-310.
- [48] Ocaña, Francisco y Pablo Oñate. 1999. “Índices e indicadores del sistema electoral y del sistema de partidos. Una propuesta informática para su cálculo”. *Revista Española de Investigaciones Sociológicas* 99(86): 23-45.
- [49] Pedersen, Morgens N. 1979. “The dynamics of West European party systems: Changing patterns of electoral volatility”. *European Journal of Political Research* 7: 1-26.
- [50] Plesia, Carolina y Lorenzo De Sio. 2018. “An evaluation of the performance and suitability of RxC methods for ecological inference with known true values”. *Quality and Quantity* 52: 669-683.
- [51] Powell, Eleanor Neff y Joshua Tucker. 2013. “Revisiting electoral volatility in post communist countries: New data, new results and new approaches”, *British Journal of Political Science* 44: 123-147.
- [52] Rincón, Luis. 2012. *Introducción a los Procesos Estocásticos*. México DF: UNAM, Facultad de Ciencias, Departamento de Matemáticas.
- [53] Robinson, William S. 1950. “Ecological correlations and the behavior of individuals”. *American Sociological Review* 15(3): 351–357.
- [54] Roblizo, M. 2001. *Transición a la democracia y evolución del comportamiento electoral en Bulgaria 1989-1994*. Cuenca: Ediciones Universidad de Castilla-La Mancha.
- [55] Rosen, Ori, Wenxin Jiang, Gary King y Martin A. Tanner, A. 2001. “Bayesian and Frequentist Inference for Ecological Inference: The RxC Case”, *Statistica Neerlandica* 55: 134-156.
- [56] Russo, Luana. 2014. “Estimating floating voters: a comparison between the ecological inference and the survey methods”. *Quality and Quantity* 28: 1667-1683.
- [57] Sani, Giacomo. 1986. “Los desplazamientos del electorado: anatomía del cambio”. Pp. 1-26 en *Crisis y cambio: electores y partidos en la España de los años ochenta*. editado por Juan J. Linz y Jose R. Montero, Madrid: Centro de estudios constitucionales.
- [58] Sartori, Giovanni. 1975. *Parties and party system. A framework for analysis*. Cambridge: Cambridge University Press.
- [59] Servicio Electoral de la República de Chile [SERVEL]. 1993. *Estadísticas de Inscripciones Vigentes por Grupos etareos 1993*, Santiago de Chile: SERVEL.
- [60] SERVEL. 1997(a). *Informe comunal de inscripciones electorales vigentes. Tomo I*, Santiago: SERVEL.

- [61] SERVEL. 1997(b). *Informe comunal de inscripciones electorales vigentes. Tomo II*, Santiago: SERVEL.
- [62] SERVEL. 2001. *Informe comunal de inscripciones electorales vigentes. Al 31 de diciembre de 2001*, Santiago: SERVEL.
- [63] SERVEL. 2005. *Informe comunal de inscripciones electorales vigentes. Al 31 de diciembre de 2005*, Santiago: SERVEL.
- [64] SERVEL. 2009. *Informe comunal de inscripciones electorales vigentes. Al 31 de diciembre de 2009*, Santiago: SERVEL.
- [65] Shair-Rossenfield, Sarah. 2008. *Assessing the causes and effects of electoral volatility: Party system fragmentation, time and executive turnover*, Thesis (MA)-The University of North California, Chapel Hill.
- [66] Sikk, Allan. 2005. “How unstable? Volatility and the genuinely new parties in Eastern Europe”. *European Journal of Political Research* 44: 391–412.
- [67] Tanner, Martin A. 1996. *Tools for statistical inference: methods for the exploration of posterior distributions and likelihood functions*, New York: Springer.
- [68] Torcal, Mariano. 2015. “Institucionalización de sistemas de partidos. Concepto, medición, procesos y consecuencias”. Pp. 7-18 en *Sistemas de partidos en América Latina. Causas y consecuencias de un equilibrio inestable* editado por Mariano Torcal, Barcelona: Antrophos.
- [69] Van der Meer, Tom WG, Erika van Elsas, Rozemarijn Lubbe y Wouter van der Brug. 2015. “Are volatile voters erratic, whimsical or seriously picky? A panel study of 58 waves into the nature of electoral volatility (The Netherlands 2006–2010)”. *Party Politics* 21(1): 100-114.
- [70] Van der Ploeg, Carin, Frank Van de Pool y Jarl Kampen. 2009. *A Comparison of Different Estimation Methods of voter transtions with an Application in the Dutch National Elections of 2003 and 2006*, Statistics Netherlands, The Hague/Heerlen. Discussion paper (09030).
- [71] Visser, Max. 1994. “Beyond the ecological fallacy. The Duncan Davis technique of ecological inference”. *Quality and Quantity* 28: 435-444.
- [72] Wackerly, Dennis D., William Mendehall y Richard L. Scheaffer. 2010. *Estadística Matemática con Aplicaciones*, México DF: CENAGE Learning.
- [73] Walsh, Bruce. 2004. *Markov Chain Montecarlo and Gibbs Sampling*. Lecture Notes for EEB 581, version 26.
- [74] Zucco, Cesar. 2015. “Estabilidad sin raíces: Institucionalización de sistemas de partidos en Brasil”. Pp. 78-107 en *Sistemas de partidos en América Latina. Causas y consecuencias de un equilibrio inestable*, editado por Mariano Torcal, Barcelona: Antrophos.