

Proyecciones en la noche electoral

por

JOSÉ MANUEL PAVÍA MIRALLES

Departamento de Economía Aplicada
Universidad de Valencia

PRUDENCIO MUÑIZ RODRÍGUEZ

Departamento de Economía Aplicada
Universidad de Valencia

JOSÉ ANTONIO ÁLVAREZ JAREÑO

Departamento de Economía Aplicada
Universidad de Valencia

RESUMEN

La década de los noventa en España no ha sido muy propicia para los sondeos electorales, lo que ha incrementado el nivel de incertidumbre que se vive en la noche electoral. En este artículo se ofrece un modelo basado en las relaciones que entre elecciones presentan las mesas electorales. El procedimiento permite realizar, durante el proceso de recuento, predicciones secuenciales acerca de la distribución final de votos y escaños sumamente fiables a partir de los votos escrutados en cada instante. El modelo se ejemplifica con una aplicación a las Elecciones a Cortes Valencianas de 1999.

Palabras clave: Elecciones, Predicciones secuenciales, Muestras no aleatorias.

Clasificación AMS: 62M20, 62P25, 92K10.

1. INTRODUCCIÓN

En la noche electoral la demanda de información sobre los resultados de la consulta es permanente. Desde el mismo instante en que se cierran los colegios electorales la actividad es frenética. En primer término, se dan a conocer las encuestas realizadas a salida de urna sobre voto emitido elaboradas para los distintos medios de comunicación y las formaciones políticas; le siguen las proyecciones obtenidas con los 100 primeros votos escrutados de un conjunto de mesas previamente seleccionadas como representativas; y, a la vez, comienzan a llegar los primeros resultados de mesas electorales escrutadas al centro logístico de recogida de información. El deseo por comprobar si los resultados de voto escrutado confirman o son compatibles con los datos de encuesta es evidente, por lo que los agentes encargados de comunicar los resultados demandan tal servicio.

Ahora bien, dadas las características especiales que presentan los datos provenientes de las primeras mesas escrutadas, las técnicas habituales de proyección, basadas en muestras aleatorias, no pueden ser empleadas. En efecto, para poder hacer estimaciones consistentes y robustas la muestra debería ser representativa de la población objeto de estudio o, al menos, en ausencia de poder garantizar la representatividad de la muestra, se exige que los datos constituyan un conjunto aleatorio de la población.

En el caso que nos ocupa, y especialmente al inicio del escrutinio, los datos disponibles no pueden ser considerados ni como representativos de la distribución de votos ni como aleatorios; dado que, como es conocido, habitualmente los primeros resultados que llegan a la Central de Datos pertenecen a pequeñas mesas de la zona rural, y difícilmente puede considerarse la distribución de votos de estas áreas similares a las distribuciones de votos de grandes mesas o del entorno urbano.

Ahora bien, a pesar de que los votos escrutados no pueden considerarse como representativos ni aleatorios de la distribución de votos, sí cabe esperar cierta consistencia entre elecciones. Así es:

- Por una parte, es previsible que mesas que habitualmente votan mayoritariamente a una opción política continúen haciéndolo en nuevas consultas;
- Mientras, por otra, cabe esperar que los posibles *desplazamientos* o trasvases de votos entre formaciones que se produzcan entre elecciones no se concentren en unas pocas mesas, sino que éstos se repartan de alguna manera a lo largo del territorio, aunque no necesariamente de manera uniforme.

De modo que el objetivo consistiría en proporcionar, a medida que los resultados van llegando al Centro de Datos, una secuencia de predicciones sobre el resultado final de la consulta, salvando el problema de la no aleatoriedad de los

datos con las armas que nos suministran las dos consideraciones anteriores. En el presente trabajo el problema se ataca desde una perspectiva de estadística clásica, sin embargo, para una aproximación bayesiana, en la que se integra en el proceso de proyecciones información proveniente de encuestas, se puede consultar Bernardo y Girón (1992) y Bernardo (1995).

El resto del documento está estructurado como sigue: en el apartado segundo se presenta el problema y algunas de las consideraciones realizadas; el apartado tercero está dedicado al planteamiento del modelo; en el apartado cuarto se detalla el proceso a seguir para emplear el modelo en situaciones empíricas y en el quinto los resultados obtenidos en su aplicación a las elecciones autonómicas a Cortes Valencianas de junio de 1999; por último, en el apartado sexto se ofrece un resumen y se realizan algunas consideraciones.

2. PLANTEAMIENTO DEL PROBLEMA

Considérese una situación en la que n electores, repartidos en c circunscripciones, a razón de n_i (para $i = 1, 2, \dots, c$) en cada circunscripción, deben elegir e_i escaños, con e_i escaños en cada circunscripción, entre los candidatos de f formaciones políticas que concurren a unas elecciones. Sea m el número total de mesas en que los votantes se encuentran repartidos, con m_i el número de mesas en cada una de las c circunscripciones y n_{ij} el censo electoral o número de electores de la j -ésima mesa de la i -ésima circunscripción. Se admite que se dispone de información histórica para cada una de las m mesas, de modo que $\mathbf{R}_{ij}(0)$ representa la matriz $f \times A$ de resultados registrados en las A convocatorias(1) anteriores en la j -ésima mesa de la i -ésima circunscripción, siendo $r_{ijk}(0)_a$ el porcentaje de votos que obtuvo el k -ésimo partido político en la a -ésima convocatoria en tal mesa y $\mathbf{r}_{ijk}(0)$ el vector fila $1 \times A$ que recoge los datos referidos a la k -ésima formación.

Se supone que en el instante t del recuento se dispone de información acerca de $m_i(t)$ mesas, con $0 \leq m_i(t) \leq m_i$, de la circunscripción i -ésima, siendo \mathbf{r}_{ij} el vector $f \times 1$ de resultados, de componentes positivos y suma uno, cuya componente k -ésima, r_{ijk} , representa el porcentaje de votos válidos que ha obtenido la k -ésima formación en la mesa j -ésima de la i -ésima circunscripción.

Por tanto, el problema consiste en obtener aproximaciones, $\hat{\mathbf{r}}_{ij}$, a los vectores \mathbf{r}_{ij} para las $m_i - m_i(t)$ mesas donde no se dispone de datos escrutados, utilizando para ello la información pasada que suministran las matrices $\mathbf{R}_{ij}(0)$ y los resultados de la convocatoria presente, \mathbf{r}_{ij} , disponibles en cada instante.

(1) Una situación simplificada se daría cuando sólo se considerase como relevante la información de las últimas elecciones, donde $A = 1$.

De modo que, una vez estimados los porcentajes en las mesas no disponibles se procede a elevar el conjunto de resultados, los reales y los estimados, para derivar aproximaciones a los porcentajes en el conjunto de las circunscripciones y así poder obtener, aplicando el algoritmo adecuado(2) con las restricciones que en cada caso imponga la Ley Electoral, estimaciones de la distribución de escaños entre las diferentes formaciones políticas. En concreto, si se asume que las mesas disponibles son las $m_i(t)$ primeras, y se utiliza como coeficiente de elevación la importancia relativa de la mesa, medida en términos del número de votantes de la mesa sobre el número de votantes total de la circunscripción, $\omega_j = n_{ij}/n_i$, se tiene que la estimación del porcentaje de votos que obtienen las f formaciones en la i -ésima circunscripción viene dada por el vector $f \times 1 \hat{r}_i$ de la ecuación (1).

$$\hat{r}_i = \sum_{j=1}^{m_i(t)} \omega_j r_{ij} + \sum_{j=m_i(t)+1}^{m_i} \omega_j \hat{r}_{ij} \quad (1)$$

Ahora bien, al problema ya comentado en la introducción acerca de la imposibilidad de poder considerar como aleatoria a la muestra de mesas que en cada instante se dispone en el Centro de Datos, se agrega el inconveniente adicional que supone el cambio entre elecciones en la composición e, incluso, en el número de las mesas electorales. En efecto, a los cambios que se registran en las mesas que aparentemente se mantienen, debido a nuevas incorporaciones (por jóvenes que han alcanzado ya su mayoría de edad o nuevos residentes) y a bajas (por defunción o por cambio de residencia), hay que sumar las creaciones que se realizan de nuevas mesas y las fusiones y desdoblamientos que se producen entre las ya existentes.

Por tanto, para un correcto uso de la información es preciso establecer una correspondencia entre las mesas de las elecciones actuales y las pasadas. A fin de establecer esta correspondencia y para que en todas las mesas se disponga de información pasada es necesario establecer una serie de hipótesis o criterios. En las mesas dónde únicamente se registra un cambio de composición la asignación se realiza en forma directa, en la esperanza de que el comportamiento de salidas y entradas en el censo sean aleatorios; en las mesas resultado de fusión de dos o más mesas se plantea tomar como valores históricos la media ponderada de los resultados obtenidos en tales mesas; mientras que para las mesas provenientes de desdoblamientos la propuesta es asignar a todas ellas los valores de la mesa origen; finalmente, para las mesas de nueva creación, que generalmente se locali-

(2) La Ley de Jefferson-d'Hondt en el caso español.

zan en áreas de expansión de las ciudades, se sugiere tomar la media de la sección, el distrito o el municipio en el que se localiza.

3. FORMULACIÓN DEL MODELO

En el apartado anterior se ha expuesto que la solución al problema pasaba por predecir los resultados de las mesas no disponibles utilizando la información que suministraban los resultados de las mesas observadas y los resultados registrados en tales mesas en anteriores convocatorias electorales. La dificultad estriba en determinar como combinar tal información para alcanzar los objetivos propuestos.

En la introducción ya se comentó que era previsible cierta consistencia entre elecciones, de forma que mesas que presentaban porcentajes altos (bajos) de votos para un partido político continuasen presentando altos (bajos) porcentajes en nuevas convocatorias, lo que podría sugerir la existencia de una relación lineal entre los porcentajes obtenidos en cada mesa por cada formación. Es decir, un modelo del tipo:

$$r_{ijk} = \alpha_k + r_{ijk}(0) \beta_k + e_{ijk}, \quad \text{para } k = 1, 2, \dots, f \text{ y } j = 1, 2, \dots, m_i \quad (2)$$

con α_k y β_k vectores paramétricos de órdenes 1 y $A \times 1$, respectivamente, y específicos para cada partido y los e_{ijk} valores aleatorios normales de media cero y con correlaciones, a fin de simplificar el modelo, constantes por partidos para cada circunscripción y nulas entre mesas, es decir, $E(e_{ijk} e_{ijk'}) = \sigma_{kk'}(i)$ y $E(e_{ijk} e_{ij'k}) = 0$. Esto implica que consideradas aisladamente las perturbaciones asociadas a cada una de las f relaciones lineales anteriores son independientes y homocedásticas, y que el elemento (k, k') de la matriz, V_i , de correlaciones entre partidos, en la i -ésima circunscripción, es igual a $\sigma_{kk'}(i)$.

Por tanto, si se expresa en notación matricial las ecuaciones de (2) se tiene que éstas se pueden expresar de modo compacto mediante:

$$r_{ij} = \alpha + R_{ij}(0) \bullet \beta + e_{ij}, \quad \text{para } j = 1, 2, \dots, m_i \quad (3)$$

donde $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_f]'$ es un vector $f \times 1$ de coeficientes, $\beta = [\beta_1, \beta_2, \dots, \beta_f]$ es una matriz $A \times f$ de coeficientes, \bullet representa el símbolo del producto fila por columna de matrices(3), y $E(e_{ij} e'_{ij}) = V_i$ y $E(e_{ij} e'_{ij'}) = O_{f \times f}$, donde $O_{f \times f}$ es una matriz $f \times f$ de ceros. Obsérvese, sin embargo, que el sistema anterior tiene perturbaciones linealmente

(3) Este operador funciona del siguiente modo: dada una matriz A $n \times m$ y una matriz B $m \times n$, el resultado es un vector $n \times 1$ cuya i -ésima componente es igual al producto matricial de la i -ésima fila de A por la i -ésima columna de B .

dependientes dado que la suma de los porcentajes obtenidos por las f formaciones para todas las mesas cualquiera que sea la convocatoria es uno, por lo que se está ante un sistema de ecuaciones generalizado con matriz de varianzas-covarianzas singular, cuyo modo de tratamiento se puede encontrar, por ejemplo, en Theil (1971, pp. 282-9), que sobre los parámetros se concreta en la restricción:

$$1 = \sum_{k=1}^f \alpha_k + \sum_{k=1}^f \sum_{a=1}^A \beta_{ka}, \quad (4)$$

donde β_{ka} es la a -ésima componente de β_k .

Si bien la expresión (3) permite observar claramente el hecho de que el sistema presenta restricciones, la anterior reorganización de la información no resulta muy operativa, por lo que se procederá a reordenar las $f \cdot m_i$ relaciones de (2) de otro modo. Para ello, se denota por \mathbf{p}_{ik} al vector $m_i \times 1$ cuya componente j -ésima es igual a r_{ijk} (es decir, es el vector de porcentajes obtenidos en todas las mesas de la i -ésima circunscripción por la k -ésima formación), mediante $\mathbf{P}_{ik}(0)$ a la matriz $m_i \times A$ cuya fila j -ésima es igual a $r_{ijk}(0)$ (es decir, la matriz de porcentajes obtenidos en las A convocatorias anteriores por la k -ésima formación en todas las mesas de la i -ésima circunscripción) y por \mathbf{e}_{ik} al vector $m_i \times 1$ cuya componente j -ésima es igual a e_{ijk} . Adicionalmente, se nota mediante $\mathbf{p}_i = [\mathbf{p}'_{i1}, \mathbf{p}'_{i2}, \dots, \mathbf{p}'_{if}]'$ al vector $(m_i \cdot f) \times 1$ de porcentajes de todos los partidos en todas las mesas, mediante $\mathbf{e}_i = [\mathbf{e}'_{i1}, \mathbf{e}'_{i2}, \dots, \mathbf{e}'_{if}]'$ al vector $m_i \times 1$ de perturbaciones, por $\boldsymbol{\chi} = \boldsymbol{\alpha} \otimes \mathbf{1}_{m_i \times 1}$ al vector $(m_i \cdot f) \times 1$ de que recoge los coeficientes α_k -donde \otimes representa el símbolo del producto de Kronecker y $\mathbf{1}_{m_i \times 1}$ es un vector $m_i \times 1$ de unos- y mediante $\boldsymbol{\gamma} = \text{block-diag}[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_f]$ a la matriz $(A \cdot f) \times f$ de coeficientes de escala y por $\mathbf{P}_i(0) = \text{block-diag}[\mathbf{P}_{i1}(0), \mathbf{P}_{i2}(0), \dots, \mathbf{P}_{if}(0)]$ a la matriz $(m_i \cdot f) \times (A \cdot f)$ de porcentajes históricos de votos registrados -siendo block-diag el operador que dadas unas matrices construye una matriz bloque diagonal con las matrices de entrada como bloques-. De modo que con esta notación se tiene que las ecuaciones de (2) ó (3) pueden expresarse como el siguiente sistema de ecuaciones generalizadas:

$$\mathbf{p}_i = \boldsymbol{\chi} + \mathbf{P}_i(0) \boldsymbol{\gamma} + \mathbf{e}_i, \quad (5)$$

con matriz de varianzas-covarianzas $E(\mathbf{e}_i \mathbf{e}_i') = \mathbf{V}_i \otimes \mathbf{I}_{m_i}$ -donde \mathbf{I}_{m_i} es una matriz identidad de orden $m_i \times m_i$ -, es decir, $E(\mathbf{e}_{ik} \mathbf{e}_{ik}') = \sigma_{kk}(i) \mathbf{I}_{m_i}$ y que verifica la restricción:

$$(\mathbf{1}_{1 \times m_i} \otimes \mathbf{I}_f) \mathbf{p}_i = \mathbf{1}_{m_i \times 1}, \quad (6)$$

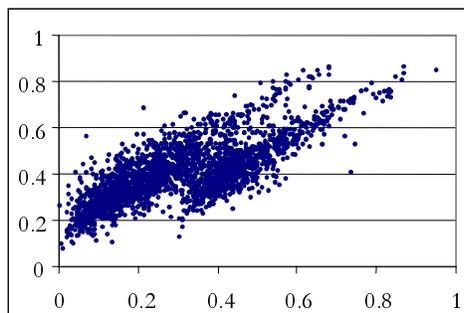
es decir, la suma de los porcentajes en cada una de las mesas es uno, que a su vez implica la restricción (4).

Ahora bien, cualquiera que sea la expresión que se considere, un modelo de este tipo supone admitir que los coeficientes de cambio (α_k y β_k) y las correlaciones entre partidos son iguales en todas las mesas. Sin embargo, como se comentó es la introducción, es razonable suponer que los trasvases que se originan entre elecciones entre formaciones no se producen con la misma intensidad en todas las mesas, existiendo mesas con votantes más *fieles* y mesas con votantes más propensos a cambiar su opción política, con una amplia gama de grados de fidelidad. En efecto, observando la Figura 1, que muestra mediante una nube de puntos los porcentajes de votos obtenidos en cada una de las 2410 mesas de la circunscripción de Valencia por (a) Partido Popular (PP) y (b) Partido Socialista Obrero Español (PSOE) en las Elecciones a Cortes Valencianas de 1991 y 1995, se intuye claramente que los cambios no son uniformes ni por partidos ni a lo largo de todo el territorio. Por lo que, para una correcta formulación del modelo es preciso identificar y agrupar a las distintas mesas según su propensión al cambio.

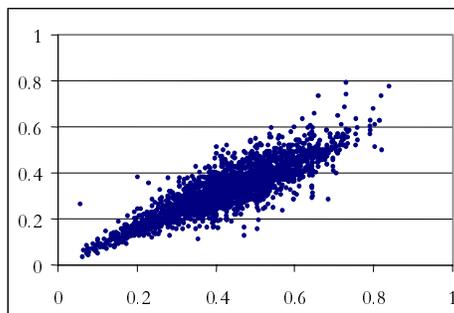
Figura 1

PORCENTAJE DE VOTOS OBTENIDOS EN LAS MESAS DE LA CIRCUNSCRIPCIÓN DE VALENCIA EN LAS ELECCIONES AUTONÓMICAS DE 1995 Y 1991

(a) Partido Popular



(b) Partido Socialista Obrero Español



Fuente: Elaboración propia a partir de Abacus (1991, 1995).

Para ello, y partiendo del nivel de la circunscripción -dado que, en general, tal es el ámbito relevante para la determinación del número de escaños(4) que obtiene cada formación- se clusterizan o agrupan las distintas mesas que componen cada

(4) Variable principal en la mayoría de convocatorias electorales.

circunscripción creando grupos de mesas más homogéneos en cuanto a propensión al cambio(5).

Para la clusterización de las mesas se utiliza la información contenida en las matrices $R_{ij}(0)$ disponibles. Es conveniente seleccionar pocos clusters o emplear un algoritmo de agrupación jerárquico, pues en otro caso, sobre todo al inicio del proceso de predicción cuando el número de observaciones es muy bajo, se corre el riesgo de tener clusters con un número excesivamente bajo de mesas disponibles, lo que dificultaría la realización de estimaciones para las mesas de tales grupos. Con un procedimiento de agrupación jerárquico, por el contrario, si el número de mesas disponibles en un cluster no supera cierto umbral mínimo se puede trabajar con el nivel de agrupación inmediatamente superior que resuelva tal eventualidad y tener, de ese modo, en todos los clusters un número suficientemente grande de observaciones. Por tanto, si se denota por g_i al número total de grupos en que se divide la i -ésima circunscripción y mediante $g_i(t) \leq g_i$ al número de clusters para los que se dispone de mesas suficientes en el instante(6) t , entonces lo que se propone es plantear las relaciones lineales de (2), (3) ó (5) en cada cluster. Es decir, para $g = 1, 2, \dots, g_i(t)$ se tendría un conjunto de ecuaciones dada por (7).

$$r_{ijk} = \alpha_{kg} + r_{ijk}(0) \beta_{kg} + e_{ijk}, \quad \text{para } k = 1, 2, \dots, f \text{ y } j = 1, 2, \dots, m_i \quad (7)$$

donde el subíndice g denota el cluster y el resto de elementos tienen la misma interpretación que la ecuación (2) y (4), salvo que ahora la matriz de varianzas-covarianzas de perturbaciones entre formaciones es diferente por circunscripción y cluster, es decir, $E(e_{ijk} e_{ijk'}') = V_{ig}$.

Se tendría, por tanto, en cada cluster, un sistema de ecuaciones distinto cuyos parámetros $-\alpha(g)$'s, $\beta(g)$'s y $\sigma(g)$'s- sería preciso estimar a partir de las mesas del cluster disponibles en cada instante y de ese modo poder estimar los porcentajes para las mesas no observadas y así poder realizar proyecciones sobre el resultado final de las elecciones.

(5) Bernardo y Girón (1992) consideran la distribución geográfica-administrativa como la variable relevante para la agrupación. De forma que razonan que el nivel de uniformidad entre los resultados irá decreciendo a medida que se asciende en la secuencia: sección, distrito, municipio, circunscripción. Esta hipótesis, sin embargo, no se ve corroborada, en general, por la realidad, dado que, aunque resulte sorprendente, es posible encontrar mesas de una misma sección con actitudes hacia el cambio bien diferenciadas.

(6) Si el número de clusters no es muy elevado $g_i(t)$ coincide con g_i en un estadio inicial del proceso de recuento.

4. PROCEDIMIENTO DE PROYECCIÓN

Una vez comenzado el recuento, los datos escrutados llegan de forma fluida y con gran celeridad, por lo que es necesario tener todo el proceso informatizado para poder hacer frente al torrente de información que recibe el Centro de Datos. Es fundamental tener todos los elementos que intervienen en el proceso perfectamente estructurados y tener previstas las posibles eventualidades que puedan ir apareciendo, como, por ejemplo, el retraso sistemático en la llegada de las mesas de una circunscripción concreta.

Cada mesa debe estar perfectamente identificada a fin de establecer adecuadamente la correspondencia entre la información histórica de la misma y sus datos actuales de forma que cada vez que se proceda a realizar un corte en el recuento para realizar proyecciones se puedan construir apropiadamente en cada cluster las ecuaciones de (7). Con los datos disponibles se estiman los parámetros del modelo y a partir de éstos las mesas faltantes, para a continuación hacer uso de la expresión (1) y obtener estimaciones de los porcentajes que en cada circunscripción obtendría cada formación política. Estimaciones puntuales que es necesario flexibilizar a fin de reducir las probabilidades de error, recurriéndose para ello a la utilización de intervalos de confianza. Intervalos, que debido a la hipótesis de normalidad asumida para las perturbaciones de las relaciones lineales, se obtendrían a partir de la distribución normal de probabilidades.

No obstante, en general, el interés en la mayoría de las consultas electorales que se celebran en el estado español se centra, más que en conocer los porcentajes de votos que obtendrá cada partido, en saber cual será la composición o reparto de escaños que se producirá entre los partidos que concurren a las elecciones, siendo, por tanto, ésta la variable relevante a aproximar. Así, habrá que aplicar sobre los porcentajes estimados las particularidades de la Ley Electoral correspondiente para poder disponer de las ansiadas proyecciones.

Ahora bien, para cada uno de los porcentajes no se dispone de un único valor, sino que está disponible todo un intervalo, por lo que la pregunta que emerge de forma obvia es: ¿Qué valor utilizar de cada formación para realizar proyecciones?

Una posible respuesta a la pregunta anterior, que es la que se ha implementado en la aplicación que se presenta en el apartado siguiente, está basada en el propio significado de los intervalos de confianza en este contexto. En efecto, a un nivel de confianza $1-\alpha$, el límite inferior representa el porcentaje mínimo de votos que obtendrá tal formación mientras el límite superior el porcentaje máximo. Por lo que, se puede calcular el número de escaños mínimo y el máximo que obtendría cada partido político situándose, para el mínimo, en la peor de las circunstancias y, para el máximo, en la mejor.

Es decir, para estimar el número mínimo (máximo) de escaños de un partido político se propone tomar el límite inferior (superior) del intervalo de confianza obtenido para su porcentaje estimado de votos y los límites superiores (inferiores) para el resto de formaciones. De modo que, aplicando el algoritmo de reparto adecuado sobre tales porcentajes se estimarían el número de escaños mínimo y máximo que obtendría cada formación a un nivel de confianza de al menos $1-\alpha$.

5. PROYECCIONES A CORTES VALENCIANAS DE 1999

En los apartados anteriores se ha expuesto un procedimiento que permite obtener, a medida que se va recibiendo información de votos escrutados, proyecciones sobre la distribución de votos y/o escaños en unas elecciones. En este apartado se presentan los resultados obtenidos, tras aplicar la metodología anteriormente descrita, en las elecciones a Cortes Valencianas celebradas en Junio de 1999.

En efecto, el pasado 13 de Junio de 1999 se celebraron, al igual que en otras doce Comunidades Autónomas y junto a elecciones municipales y al Parlamento europeo, en la Comunidad Valenciana elecciones a Cortes Autonómicas. En ellas más de tres millones de valencianos, repartidos en tres circunscripciones (1.096.759 en Alicante, 384.195 en Castellón y 1.857.306 en Valencia) y 5125 mesas electorales (1864 en Alicante, 652 en Castellón y 2609 en Valencia), estaban convocados para elegir 89 escaños (30 en Alicante, 22 en Castellón y 37 en Valencia) entre 16 formaciones políticas.

Ahora bien, a pesar de ser 16 las formaciones que se presentaban a las elecciones, no todas ellas tenían posibilidades realistas de obtener representación parlamentaria. En concreto, y a priori, sólo cinco presentaban posibilidades razonables de acceder al parlamento: Partido Popular (PP), Partit Socialista del País Valencià - Partido Socialista Obrero Español - progresistas (PSOE), Esquerra Unida del País Valencià - Izquierda Unida (EU-IU), Unió Valenciana (UV) y Bloc Nacionalista Valencià - Els Verds (BLOC), por lo que se ha limitado la información que se presenta en la Tabla 1 a estas formaciones.

La Ley Electoral a Cortes Valencianas establece que el reparto de escaños ha de realizarse en cada circunscripción de acuerdo con el algoritmo de Jefferson-d'Hondt entre aquellas formaciones que hayan obtenido en el conjunto de la Comunidad Valenciana al menos un 5% de los votos válidos. Esta condición del techo del 5%, específica de la Ley Electoral Valenciana (DOGV, 1987), dificulta notablemente las proyecciones. En efecto, por una parte, al coincidir en una misma cita elecciones municipales, autonómicas y europeas (siendo este el orden de recuento), el retraso en la disponibilidad de la información de las mesas de tamaño elevado es acumulativo, lo que puede provocar que se disponga de un número elevado de

mesas para algunos clusters cuando de otros casi no se dispone de información. Por otra parte, y más importante, uno de los interrogantes fundamentales que existía en estas elecciones era saber si tanto UV como BLOC-Els Verds superarían el límite del 5% (de hecho obtuvieron el 4.71% y el 4.56% respectivamente, superando en algunas circunscripciones a formaciones que finalmente obtuvieron representación parlamentaria), lo que en gran medida podría condicionar el reparto de escaños.

Con los condicionantes anteriores y con el objetivo de predecir la composición del Parlamento Autonómico Valenciano a medida que se iba recibiendo información en el Centro de Datos se realizaron 126 cortes o proyecciones, el primero con sólo un 0.18% del censo escrutado y el último con ya un 96.08% escrutado. En la Tabla 1 se presentan a modo de ejemplo algunas de las proyecciones realizadas.

Tabla 1
EJEMPLOS DE PREDICCIONES DE PORCENTAJES DE VOTOS Y ASIGNACIÓN DE ESCAÑOS DURANTE EL ESCRUTINIO

| | | PP | PSOE | EU-IU | UV | BLOC |
|--|------------|--------------|--------------|-------------|-------------|-------------|
| Escrutado al 1.89% del censo (21h55) | Porcentaje | 49.63 ± 0.92 | 32.98 ± 0.87 | 6.29 ± 0.45 | 4.12 ± 0.37 | 4.90 ± 0.40 |
| | Horquilla | 48 - 52 | 31 - 35 | 4 - 6 | 0 | 0 - 4 |
| | Escaños | 50 | 34 | 5 | 0 | 0 |
| Escrutado al 3.43% del censo (22h04) | Porcentaje | 49.54 ± 0.51 | 34.70 ± 0.40 | 6.15 ± 0.18 | 4.09 ± 0.27 | 4.61 ± 0.28 |
| | Horquilla | 49 - 50 | 33 - 35 | 4 - 6 | 0 | 0 |
| | Escaños | 50 | 35 | 4 | 0 | 0 |
| Escrutado al 10.07% del censo (22h20) | Porcentaje | 48.81 ± 0.42 | 34.29 ± 0.38 | 6.06 ± 0.17 | 4.24 ± 0.16 | 4.70 ± 0.16 |
| | Horquilla | 49 - 50 | 35 | 4 - 5 | 0 | 0 |
| | Escaños | 50 | 35 | 4 | 0 | 0 |
| Escrutado al 58.08% del censo (23h47) | Porcentaje | 48.41 ± 0.33 | 34.57 ± 0.30 | 6.17 ± 0.14 | 4.47 ± 0.12 | 4.50 ± 0.13 |
| | Horquilla | 49 - 50 | 34 - 35 | 5 | 0 | 0 |
| | Escaños | 49 | 35 | 5 | | |
| Escrutinio al 80.34% del censo (01h12) | Porcentaje | 48.39 ± 0.32 | 34.30 ± 0.29 | 6.14 ± 0.13 | 4.58 ± 0.12 | 4.47 ± 0.12 |
| | Horquilla | 49 | 35 | 5 | 0 | 0 |
| | Escaños | 49 | 35 | 5 | 0 | 0 |
| Escrutado al 100% del censo (05h00) | Porcentaje | 48.14 | 34.10 | 6.08 | 4.71 | 4.56 |
| | Escaños | 49 | 35 | 5 | 0 | 0 |

Nota: Para los porcentajes se presentan intervalos de estimación al 95.5% de confianza. Escaños representa el valor más probable.

Fuente: Elaboración propia.

Como se puede observar en la Tabla 1, donde se ofrecen algunas de las proyecciones realizadas, ya en una fase muy temprana del proceso de recuento -al

1.89% escrutado en la Comunidad Valenciana (1.75% en Alicante, 4.13% en Castellón y 1.52% en Valencia)- las estimaciones que se ofrecen son muy fiables. Las primeras proyecciones, sin embargo, presentaban unas horquillas algo amplias debido al hecho de que las predicciones que se obtenían en esos instantes para BLOC-Els Verds situaban el total de votos para este partido en el conjunto de las tres circunscripciones en un entorno del 5%, sin duda favorecido por la mayor celeridad con que se iban conociendo los resultados de la circunscripción de Castellón, donde finalmente fue la tercera fuerza política con el 5.32% de los votos válidos.

No obstante, esta situación no se mantuvo durante mucho tiempo, dado que sólo 9 minutos después y con un 3.43% escrutado (2.79% en Alicante, 7.88% en Castellón y 2.90% en Valencia), las proyecciones que se obtenían descartaban tal eventualidad. A pesar de ello, se ha remarcar las dificultades que habría acarreado sobre el proceso de proyecciones si UV y/o BLOC-Els Verds se hubiesen mantenido en el entorno del 5% durante buena parte del recuento.

Finalmente, es preciso destacar dos aspectos que se derivan del análisis de la Tabla 1. Por una parte, a pesar de que la mejor estimación de composición del Parlament converge rápidamente, se observa que las horquillas de predicción permanecen abiertas, aunque sólo sea por un único escaño, hasta bien avanzado el proceso de recuento. Este hecho es debido a que en la circunscripción de Castellón el último escaño se decidió al final del recuento por sólo 1.825 votos, apenas un 0.68% de los votos válidos de la circunscripción. Por otro lado, se advierte que las primeras proyecciones tienden a sobrestimar el porcentaje de votos de PP y subestimar las de UV, aunque paulatinamente las proyecciones convergen a los resultados finales. Esta circunstancia se explicaría por el hecho de que el recuento en la circunscripción de Valencia, donde las mesas tienen un tamaño medio superior (712 censados, frente a aproximadamente 589 de Alicante y Castellón), presentaba un mayor retraso y precisamente en esta circunscripción PP obtuvo sólo el 47.06% de los votos y UV, fue la tercera fuerza política, con el 6.47% de los votos válidos.

6. CONSIDERACIONES FINALES

En este documento se ofrece un procedimiento que, durante el proceso de recuento de la noche electoral, permite realizar proyecciones acerca de la distribución de votos y escaños entre las formaciones que concurren a las elecciones. El método se sustenta en las relaciones o consistencia entre elecciones que presentan las mesas electorales, utiliza como información básica los resultados, por mesas electorales, de convocatorias anteriores y los votos escrutados en cada instante, y

permite que en pocos segundos se disponga de predicciones sumamente fiables del resultado final de la consulta.

Además, este método, comparado con otros que también pueden ser implementados con el mismo objetivo como las encuestas a salida de urna y los muestreos en mesas representativas, tiene como principal virtud su bajo coste, a lo que se suma su enorme robustez y alta fiabilidad.

En efecto, cualquiera de los métodos de encuesta habitualmente utilizados conllevan un gran despliegue logístico y de personal que encarecen enormemente su aplicación, mientras que para el empleo de este método basta con tener línea abierta con el ordenador que centraliza el recuento y tener previamente organizada toda la información previa, de modo que un pequeño equipo podría hacerse cargo de todo el trabajo.

Asimismo, y teniendo en cuenta que los resultados que se ofrecen tras el cierre de las urnas por los medios de comunicación provenientes de encuestas a pie de urna están fuertemente influenciados por *errores* de respuesta y que en la historia reciente (Autonómicas-99, Generales-2000) recibieron fuertes críticas, la notable robustez que presenta el método propuesto, que se puede observar en la Tabla 1, aboga a favor de su uso.

Por otro lado, confrontando los resultados que se ofrecen en la Tabla 1 con aquellos que pudieran provenir de estimaciones basadas en mesas representativas es preciso comentar tres cuestiones sumamente significativas: primero, no está claro que con el procedimiento de mesas representativas se disponga con mayor prontitud de proyecciones; segundo, el error asociado a las predicciones no es necesariamente menor, y, tercero, caso de registrarse algún retraso o problema con alguna de las mesas la calidad de las predicciones basadas en muestras representativas se ve notablemente afectada, no así el método propuesto.

En otro ámbito, al comparar este procedimiento con el propuesto por Bernardo y Girón (B&G), basado en inferencia bayesiana, aparece una certeza y una duda. Por una parte, el método expuesto no precisa, al contrario del de B&G, de estimaciones iniciales para la distribución de votos. Mientras, por otro lado, observando los resultados publicados en Bernardo (1995, p. 21, Tabla 1.3), donde se ofrecen predicciones únicamente al 77% y 91% de voto escrutado para las elecciones a Cortes Valencianas de 1995, surge cierta incertidumbre relacionada con la robustez del método B&G. Así es, en las elecciones autonómicas de 1995 se produjo un gran vuelco electoral en el panorama político de la Comunidad Valenciana que fue acompañado la noche electoral de una escasa convergencia del voto escrutado hasta bien avanzado el proceso de recuento lo que, sin duda, dificulta las proyecciones con bajos porcentajes de censo escrutados. Sin embargo, al aplicar la

metodología propuesta en este manuscrito sobre las elecciones autonómicas de 1995, empleando la información de las de 1991, con la secuencia en que se recibieron los datos en tales elecciones, como mecanismo para validar el método antes de su utilización, se comprobó que las predicciones que se hubiesen obtenido hubieran sido altamente satisfactorias, incluso con muy bajos porcentajes de censo escrutado.

Finalmente, interesa destacar un aspecto importante relativo al hecho de que las estimaciones electorales, contrariamente a lo que ocurre en muchos otros problemas de inferencia dónde no es posible medir la calidad o adecuación de las estimaciones, sí pueden ser contrastadas con los resultados ciertos, por lo que al ser posible cuantificar la bondad de las estimaciones obtenidas permiten calibrar la eficacia de la metodología empleada. De modo que, a la vista de los resultados obtenidos, el método propuesto podría ser exportado con garantía a otro tipo de problemas de naturaleza diferente que presentasen características similares al aquí tratado.

REFERENCIAS

ABACUS (1991). «*Base de Datos de Resultados de Elecciones a Corts Valencianes de 1991*». Valencia: Abacus, S.L..

ABACUS (1995). «*Base de Datos de Resultados de Elecciones a Corts Valencianes de 1995*». Valencia: Abacus, S.L..

BERNARDO, J.M. Y F.J. GIRÓN (1992). «*Robust Sequential Prediction from Non-random Samples: the Election Night Forecasting Case*». *Bayesian Statistics*, 4, 61-77.

BERNARDO, J.M. (1995). «*Probing Public Opinion: the State of Valencia Experience*». *Proceedings Third Workshop on Bayesian Statistics in Science and Technology*, Ed. R.Kass, 3-27.

DIARIO OFICIAL DE LA GENERALITAT VALENCIANA (1987). «*Ley 1/87, Ley Electoral Valenciana*». DOGV, nº 561, 31 de marzo de 1987.

THEIL, H. (1971). «*Principles of Econometrics*». London: John Wiley & Sons.

PROJECTIONS AT THE ELECTION NIGHT

SUMMARY

The nineties were not in Spain a propitious decade for the opinion polls. This fact has made the level of uncertainty at the electoral night increase. In this paper a model to make projections during the process of inventory about the distribution of votes and seats is offered. The method sustains itself in the relationships that show the polling stations between elections. By using the results of previous summons and the votes polled in every moment, predictions highly reliable of the final results can be obtained in a few seconds. The model is exemplified by an application to the elections for Cortes Valencianas of 1999.