

ANÁLISIS ENCUESTAS ELECTORALES

Febrero 2017

nº 26

**Promoción 2016-2017 ADE UAM
Asignatura Análisis y Obtención de Datos**

El estudio se ha centrado en analizar las encuestas que tuvieron por objetivo determinar la intención de voto en las elecciones al parlamento español de 26 de Junio 2016, elecciones en las que el error cometido por las encuestas fue notable y que llevó a que se originaran numerosas críticas hacia los sondeos por diferentes sectores de la sociedad.

Para cada encuesta se definieron las variables más significativas que pudieran aportar información a nuestro objetivo de analizar los errores y encontrar regularidades estadísticas de los mismos, siendo destacable la relación entre el error cometido con los partidos y el diario que publica la encuesta

Equipo colaborador:

Alba Bravo Lobo

Jesús Cruz Rodríguez

Alfonso García Caja

Jesús Serna Cañas

Ana Beatriz Bento Coelho Martins de Castro

Jose Luis Casana Pascual

Belén Caro Poza

Juan José Martínez Sánchez

Belinda Rica

Manuel Fernández Calleja

Cristian Jiménez Alonso

Marta González Álvarez

Estefanía Clara

Nerea Palombi Montiel

Fabián Rehbach

Nicolás Vélez Arcos

Gonzalo Chen Jiang

Nuria Calderón Corredor

Inês Sofía Sousa Marques

Santiago Torres Cadavid

Irene Fernández García

Sara Morales López

Jaime de Ramón Benedito

Edita:

Instituto L.R.Klein – Centro Gauss

Facultad de CC.EE. y EE.

Universidad Autónoma de Madrid

28049 Madrid

Teléfono y Fax: 913974191

Correo Electrónico: klein.gauss@uam.es

Página Web: www.uam.es/klein/gauss

ISSN 1696-5035

Depósito Legal: M-30165-2003

© Todos los derechos reservados. Queda prohibida la reproducción total o parcial de esta publicación sin la previa autorización escrita del editor.

ÍNDICE

I.- PRESENTACIÓN	4
II.- ERRORES RECIENTES EN LA ENCUESTAS ELECTORALES	6
III.- BASE DE DATOS PARA EL ANÁLISIS	8
I. Estadísticos Descriptivos de las variables numéricas	9
IV.- PRINCIPALES RESULTADOS	10
II. Errores cometidos en las encuestas para cada uno de los partidos	11
III. Tipo de entrevista utilizada.....	12
IV. Tamaño muestral medio de los periódicos	13
V. Tamaño muestral medio de las empresas demoscópicas	14
VI. Error cometido en las encuestas según el diario que las publica	15
VII. Error cometido en las encuestas según la empresa que la realiza	16
VIII. Relación entre error de la encuesta y el tamaño muestral	17
IX. Relación entre error de la encuesta y días previo a las elecciones.....	18
X. Relación entre el trimestre y el ECM según la empresa que realiza las encuestas	19
XI. Error cometido con PP según la empresa que realiza la encuesta	20
XII. Error cometido con PSOE según la empresa que realiza la encuesta.....	21
XIII. Error cometido con UNIDOS PODEMOS según la empresa que realiza la encuesta	22
XIV. Error cometido con CIUDADANOS según la empresa que realiza la encuesta.....	23
XV. Dependencia entre error cuadrático medio por partido político y el periódico que publica la encuesta	24
V. CONCLUSIONES.....	25
VI. BIBLIOGRAFÍA.....	26
VII. ANEXO	27

I.-PRESENTACIÓN

Durante el primer semestre del curso 2016/17 y como trabajo de investigación de la asignatura ‘Obtención y Análisis de Datos’ del último año del perfil ‘Métodos de Planificación y Previsión Empresarial’ de la Facultad de Económicas de la Universidad Autónoma de Madrid, se ha realizado un estudio sobre la capacidad predictiva de las encuestas de intención de voto. El estudio, en el que han intervenido todos los alumnos de la asignatura, se centró en analizar las encuestas que tuvieron por objetivo determinar la intención de voto en las elecciones al parlamento español de 26 de Junio 2016, elecciones en las que el error cometido por las encuestas fue notable y que llevó a que se originaran numerosas críticas hacia los sondeos por diferentes sectores de la sociedad.

Estas páginas recogen los principales resultados y conclusiones de la investigación y si bien no pretendemos intervenir en la polémica suscitada por los errores cometidos en las encuestas, si quisiéramos aportar algo de luz al problema, con un trabajo que pretende ser objetivo y producto de un análisis riguroso y de carácter científico en el ámbito de las ciencias sociales.

La polémica no es exclusiva de nuestro país y recientemente podemos encontrar casos similares en otros países o incluso más llamativos por el tamaño del error y la menor dificultad estadística del pronóstico. Así en el Reino Unido y en el referéndum sobre el Brexit (23/6/2016), las encuestas anticipaban la pertenencia a Europa y el resultado fue contrario. En Colombia (2/10/2016) las encuestas anticipaban que el referéndum sobre el acuerdo de paz con las FARC sería positivo y el resultado fue el contrario. Finalmente, en las elecciones a la presidencia de Estados Unidos las encuestas anticipaban que Hillary Clinton ocuparía la Casa Blanca pero fue el candidato republicano Donald Trump quién ganó la presidencia. Todos estos fracasos hacen más relevante y oportuno un análisis sobre las encuestas de intención de voto en España.

La primera fase del trabajo consistió en construir la base de datos con las encuestas publicadas en los medios de difusión escrita y que serían objeto de nuestro análisis. Por tanto nuestro estudio se ciñe a los sondeos publicados o referenciados en los diarios españoles, obviando los solamente publicados en otros medios como radio o televisión. Para ello se realizó una búsqueda exhaustiva que concluyó con la localización de un total de 40 estudios pertenecientes en su mayoría a los diarios de mayor difusión; ABC (7), El País (7), La Razón (9), El Periódico (6), El Mundo (5), más otros (6) no asignados a estos medios. Así mismo se identificó, entre otras variables la empresa o institución que realizó las encuestas y que fueron NC Report, GAD3, Metroscopia, GESOP, SIGMA DOS y CIS

Para los 40 estudios se definieron las variables más significativas que pudieran aportar información a nuestro objetivo de analizar los errores y encontrar regularidades estadísticas de los mismos. Concretamente se tomaron como variables el medio, fecha de realización y de publicación, tamaño de la muestra, error de muestreo, trabajo de campo, estratificación y la intención de voto estimada para cada uno de los principales partidos (PP, PSOE, PODEMOS/IU, CIUDADANOS). La intención de voto comparada con el resultado final nos permitió la estimación del error por partido así como una medida del error global de la encuesta mediante el cálculo del error cuadrático medio.

Con los datos disponibles se procedió a realizar el análisis que se agrupó según el procedimiento estadístico y las variables utilizadas, es decir el análisis descriptivo de la información, análisis de correlación, análisis Chi cuadrado, análisis de la varianza y análisis de regresión. Todos los análisis estadísticos se realizaron con el programa R, que por sus características de potencia y disponibilidad en código abierto ha sido el programa utilizado en la asignatura. Los principales resultados se muestran en estas páginas atendiendo a un criterio de simplicidad y sencillez expositiva para que sea accesible a cualquier lector sin estudios especializados y en el ANEXO I se recoge la base de datos completa. Esta base de datos se ofrece por si se quisiera comprobar o recalcular los análisis o incluso aportar, criticar o mejorar nuestras primeras conclusiones desde la ausencia de ocultamiento que debe prevalecer en cualquier estudio mínimamente científico.

Las conclusiones son múltiples y todas ellas se pueden encontrar en el punto dedicado a las mismas, pero si quisiera resaltar una que a mi juicio es más llamativa y que hace referencia al sesgo que se comete por en la sobreestimación o subestimación de la intención de voto según el periódico que ha publicado la encuesta. En ningún momento afirmamos que exista sesgo intencionado, lo que significaría manipulación, nos limitamos a señalar una conclusión que se extrae de las encuestas realizadas para las elecciones de junio de 2016 y que sería deseable contrastar con otros análisis y otras consultas electores.

Finalmente quiero insistir en que el trabajo es el resultado de un taller de investigación en la asignatura que imparto, que sus resultados son válidos y reales, y que en esta ocasión hemos decidido hacerlos públicos por si con ello pudiéramos arrojar algo de luz a la problemática que rodea las encuestas electorales. Mis felicitaciones a los alumnos participantes por el trabajo realizado.

José Vicéns Otero

Catedrático de Economía Aplicada

II.- ERRORES RECIENTES EN LAS ENCUESTAS ELECTORALES

Tras el conocimiento de los resultados de las segundas elecciones celebradas el 26 de junio de 2016, uno de los temas que más repercusión alcanzó fue el fracaso de las predicciones realizadas por las encuestas. La intención de voto plasmada en las encuestas, poco tuvo que ver con el voto real.

Ante esta situación, tras la lectura y recogida de opiniones y análisis de diversos artículos relativos a estos errores, podemos señalar como principales causas sobre estos errores las siguientes:

1. **Cambio de preferencias de voto.** La incertidumbre existente en el periodo pre-electoral ha llevado a una gran cantidad de los ciudadanos españoles a cambiar su intención original de voto, optando por los partidos tradicionales (Berta Barbet, doctora en ciencias políticas, junio 2016¹)
2. **Sobrestimación de electorado movilizado.** Las encuestas pronosticaban un alto grado de movilización, sobre todo en los partidos emergentes, que al final no se produjo demostrando que el electorado español es moderado y no de extremos (Narciso Michavila, presidente de GAD3, junio 2016²)
3. **Voto oculto.** Se pudo llegar al límite de que parte de la ciudadanía no mostrara la verdadera intención de voto en las encuestas, ya sea por miedo o por desconfianza hacia su interlocutor falseando el resultado (Jesús Andrés Lombardero, junio 2016³)
4. **Configuración de las muestras y tipo de muestreo.** La mayoría de encuestas electorales se realizan a través de teléfono fijo, dejando fuera de la muestra a un segmento importante de la población, que utiliza teléfono móvil, de edad más joven y que representa en mayor grado a los abstencionistas (Jesús Andrés Lombardero, junio 2016).
5. **Encuestas como arma propagandística.** Los partidos y los medios confían en que si un sondeo anticipa una opción ganadora respecto a las expectativas, muchos indecisos se apuntarán al “caballo ganador” (Amando de Miguel, sociólogo, junio 2016⁴)

Aumentando el radio de la preocupación acerca de los errores cometidos en las encuestas de intención de voto, conviene citar algunos otros casos ocurridos en otros países aunque bien es cierto que el caso español es el que aquí nos interesa

¹http://www.elconfidencial.com/elecciones-generales/2016-06-28/los-expertos-explican-el-fracaso-de-las-encuestas-y-el-pinchazo-de-unidos-podemos_1224451/

²http://www.elconfidencial.com/elecciones-generales/2016-06-28/los-expertos-explican-el-fracaso-de-las-encuestas-y-el-pinchazo-de-unidos-podemos_1224451/

³<http://www.lacuartacolumna.com/los-motivos-por-los-que-las-encuestas-han-supuesto-el-gran-fracaso-de-la-jornada-electoral/>

⁴<http://theobjective.com/elsubjetivo/amando-de-miguel/el-fracaso-de-las-encuestas-electorales/>

Atendiendo únicamente al año 2016 vemos diversos ejemplos. En Reino Unido, las encuestas apuntaban a que el referéndum realizado el 23 de junio de ese mismo año tendría un resultado favorable a la opción de mantenerse en la unión Europea, llegando a dar un porcentaje del 55% a la pertenencia en la UE. El resultado, aunque no homogéneo por territorios, fue justo el contrario. Con un 52%, Reino Unido iniciaba el proceso de salida de la Unión Europea.

En Colombia, el 2 de octubre de 2016 se celebró un referéndum sobre un acuerdo de paz con el grupo guerrillero de las FARC. Igualmente, las encuestas apuntaban a que había una mayoría, llegando incluso al 68%, que votaría afirmativamente ese acuerdo de paz. El resultado fue de nuevo contrario a lo que las encuestas apuntaban y se rechazó el acuerdo de paz con un 50,22% de votos negativos.

El último fracaso fue el de las elecciones de EEUU celebradas el día 8 de noviembre, donde las encuestas apuntaban a que la candidata del partido demócrata, Hillary Clinton, se haría con la presidencia de EEUU. Aunque aquí las encuestas no se equivocaron en cuanto a victoria en porcentaje de votos, ya que Hillary superó a Donald Trump por más de 2 millones de votos, si se equivocaron en la persona que ocuparía la Casa Blanca ya que, finalmente, fue el candidato del partido republicano quien se hizo con la presidencia.

A partir de ahora, habiendo plateado la situación en la que nos situamos, comienza nuestro verdadero estudio. Este recoge la elaboración de un informe que redacte los conocimientos necesarios para llevarlo a cabo, con los que luego poder obtener y tratar los datos empleados en la elaboración de nuestra base de datos sobre la intención de voto previa a las elecciones generales del 20 de junio 2016.

El objetivo fundamental de este análisis es dar una respuesta de por qué las encuestas han cometido tanto error, para lo cual utilizaremos sobre todo información secundaria, estudios realizados por organismos competentes y artículos redactados por expertos en la materia. Todos los datos serán tabulados y recogidos en una base de datos, para su posterior utilización. Como medios para la realización de la investigación, nos serviremos de las herramientas informáticas R y Excel.

III.- BASE DE DATOS PARA EL ANÁLISIS

Para la construcción de la base de datos que nos permitiría analizar los resultados de las encuestas electorales y su fiabilidad hemos seguido los siguientes pasos:

- a) Establecer las principales fuentes de información secundaria, que en nuestro caso son los principales medios de comunicación que publican las encuestas electorales. Hemos seleccionado los medios de prensa escrita más influyentes como pueden ser El País, El Mundo, El Periódico, ABC y La Razón, así como otros que hemos considerado importantes también como El Diario, EuropaPress o El Confidencial, que tendrán menor impacto al recogerse menores encuestas de estos medios.

El número de encuestas recogidas de cada medio de comunicación han sido las siguientes: ABC 7 encuestas, El País 7 encuestas, La Razón 9 encuestas, El Mundo 5 encuestas, El Periódico 6 encuestas, El Diario 2 encuestas, Europa Press 1 encuesta, El Confidencial 1 encuesta, Cadena Ser 1 encuesta e Índice de Opinión Pública 1 encuesta. El total de encuestas recogidas ha sido de 40.

- b) Búsqueda de las encuestas en los diferentes medios de comunicación que las publican, y búsqueda, a su vez, de los datos técnicos de cada una de las encuestas. Este proceso se realizó a través de las publicaciones de cada encuesta que están asociadas a su ficha técnica, que representó la forma en la que se realizó su desarrollo.
- c) Establecimiento de los aspectos más relevante de las encuestas. En nuestro caso se han seleccionado como características más relevantes: El **medio que publica** la encuesta y la **empresa encargada de realizarla, tipo de encuesta** (Personal, telefónica, online etc.), la **muestra** utilizada y su forma de selección (**tipo de muestreo**: Aleatorio simple, estratificado), el **tipo de estratos** utilizados en casos de muestreo aleatorio estratificado, **la fecha de publicación** y la **fecha de realización** de la encuesta, y por último, los **resultados de los partidos políticos** que cada una de las encuestas han establecido, así como su **error cuadrático** cometido con respecto a los resultados finales. (Para ver más detallado las variables utilizados utilizar el anexo)
- d) Recoger los aspectos más relevantes establecidos en un formato que permite trabajar con los datos obtenidos. Se utiliza una hoja de cálculo, en este caso Excel, que es una herramienta global que nos permita integrar los datos en plataformas de tratamiento de datos como puede ser Eviews o el programa R.

La base de datos puede consultarse en el ANEXO I y los principales estadísticos descriptivos así como los análisis de relación y dependencia se expondrán en las páginas siguientes.

I.- Estadísticos Descriptivos de las variables numéricas

Variable	Media	Mediana	Desviación típica	Coefficiente de variación	Recorrido Intercuartílico
Diferencia Dias	47,17	27,5	42,65	0,94	44,25
Error de muestreo	2,83	2,88	0,75	0,26	0,88
Tamaño de la muestra	1869,62	1213	2593,51	1,39	925
Error cuadrático medio	7,04	6,34	3,01	0,43	4,50
Error sobre PP	3,73	3,78	1,11	0,30	1,65
Error sobre PSOE	1,43	1,56	0,85	0,59	0,82
Error con UNIDOS PODEMOS	-3,05	-3,45	1,20	-0,39	1,72
Error con CIUDADANOS	-1,93	-1,85	1,15	-0,59	1,00
Error con OTROS	-0,18	-0,44	1,25	-6,94	1,82

De forma resumida y a efectos de presentación de los datos utilizados podemos decir que las encuestas analizadas se han realizado con una anticipación de 47 días y un error de muestreo del 2,8%, con un tamaño muestral medio de 1869,6. Los errores medios cometidos fueron más elevados para el PP (3,7%) y UNIDOS PODEMOS (-3%), mientras que se cometieron menos errores con el PSOE (1,4%) y con otros partidos (-0,18%).

Dada la definición de los errores, los signos positivos indican subestimación de la encuesta mientras que signos negativos sobreestimación en la intención de voto. Los datos indican que la intención de voto del PP fue subestimada por las encuestas mientras que para UNIDOS PODEMOS las encuestas anticipaban una intención de voto superior al voto finalmente obtenido.

La mediana de la diferencia de días nos indica que la mitad de las encuestas analizadas se publicaron 27,5 días antes de las elecciones y la otra mitad después de ese tiempo.

El recorrido intercuartílico, al igual que otros estadísticos, indica que hay mayor dispersión en los errores cometidos con el PP (1,65) y con UNIDOS PODEMOS (1,72), mientras que PSOE ha sido el partido con menor dispersión en los errores con un recorrido intercuartílico de (0,82).

IV.- PRINCIPALES RESULTADOS

A continuación se muestran los resultados obtenidos. En primer lugar analizando las diferentes variables de forma individual tenemos:

1. Errores para cada uno de los partidos analizados
2. Tipo de entrevista utilizada
3. Tamaño muestral medio
4. Error cometido en las encuestas

Y en segundo lugar analizando las relaciones entre variables, con el objetivo de encontrar posibles relaciones significativas entre estas, encontramos:

1. Relaciones entre variables cuantitativas, analizadas a través del coeficiente de correlación
 - a) Relación entre error de la encuesta y el tamaño muestral
 - b) Relación entre error de la encuesta y días previos a las elecciones
 - c) Relación entre los errores de cada partido y días previos a las elecciones
 - d) Relación entre el trimestre y el ECM según la empresa que realiza las encuestas
2. Relaciones entre variables cualitativas, analizadas mediante Chi cuadrado
 - a) Error cometido con el PP según el medio que publica la encuesta
 - b) Error cometido con el PP según la empresa que realiza la encuesta
 - c) Error cometido con el PSOE según el medio que publica la encuesta
 - d) Error cometido con el PSOE según la empresa que realiza la encuesta
 - e) Error cometido con UNIDOS PODEMOS según el medio que publica la encuesta
 - f) Error cometido con UNIDOS PODEMOS según la empresa que realiza la encuesta
 - g) Error cometido con CIUDADANOS según el medio que publica la encuesta
 - h) Error cometido con CIUDADANOS según la empresa que publica la encuesta
3. Relaciones entre variables numéricas y cualitativas, analizadas con ANOVA
 - a) Error por partido según el medio que publica la encuesta
- 4.- Los análisis de regresión para explicar el error y tomando como variables explicativas el tamaño muestral y el tiempo o número de días previos no han obtenido resultados significativos y no muestran relación entre las variables.

II.- Errores cometidos en las encuestas para cada uno de los partidos analizados

Partido Político	Errores Medios	 Errores Medios 	DT	Coef. de Variación
PP	3,73	3,73	1,11	0,30
PSOE	1,43	1,52	0,85	0,59
UNIDOS PODEMOS	-3,05	3,05	1,20	-0,39
CIUDADANOS	-1,93	1,99	1,15	-0,59
OTROS	-0,18	1,04	1,25	-6,94

En términos promedios para PP y PSOE los medios han infra estimado el número de votos y por el contrario para UNIDOS PODEMOS, CIUDADANOS y OTROS, los medios han sobre estimado el número de votos. El mayor error se cometió para el PP con casi un 4% de error medio y para UNIDOS PODEMOS(3,6%) superando así el error de muestro estadístico esperado.

El error de las encuestas analizadas se focaliza principalmente en estos dos partidos y mientras que al PP le adjudicaron menos votos de los que finalmente obtuvo a UNIDOS PODEMOS le asignaron en exceso.

Analizando dicho error en términos absolutos vemos como la media de errores cometidos por los medios aumenta para algunos partidos sobre los que se ha cometido menos error.

III. Tipo de entrevista utilizada

Trabajo de Campo	Media Muestral	% Muestra
Entrevista Telefónica	1471,86	92%
Entrevista Personal	6932,67	8%

La práctica totalidad de las encuestas, con la excepción de las realizadas por el CIS, se han realizado telefónicamente con un tamaño muestra medio de 1472 sujetos, mientras que para las entrevistas personales se realizaron mayor número de entrevistas medio (6933 sujetos).

IV.- Tamaño muestral medio de los periódicos

Medios	Nº Sondeos	Tamaño Muestral Medio
La razón	8	2062,50
ABC	7	1074,14
El País	7	1471,43
El Periódico	6	1358,00
El Mundo	5	1100,00

Excluyendo las encuestas del CIS, de carácter público y no encargadas por ningún medio, el mayor tamaño muestral corresponde a los sondeos llevados a cabo por La Razón (2063 entrevistas) y los de menor tamaño muestral a ABC (1074 entrevistas) y El Mundo(1100 entrevistas).

Sin embargo, y como se verá más adelante, el tamaño muestral no guarda relación con el error cuadrático medio cometido

V.- Tamaño muestral medio de las empresas demoscópicas

Empresa	Nº Sondeos	Tamaño Muestral Medio
NC Report	8	2062,50
GAD3	7	1074,14
Metroscopia	7	1471,43
GESOP	6	1358,00
SIGMA DOS	5	1100,00
CIS	2	9990,50

Si analizamos los tamaños muestrales según las empresas que realizaron las diferentes encuestas destacan las llevadas a cabo por el CIS, institución de la que se recogieron 2 sondeos con 17488 y 2493 entrevistas respectivamente.

VI.- Error cometido en las encuestas según el diario que las publica

Medios	Media ECM
La razón	5,81
ABC	4,94
El País	10,50
El Periódico	8,67
El Mundo	5,07

El País, La Razón y El Periodico son los diarios que mayor error cuadrático medio comenten mientras que El Mundo y ABC son los medios que menor error presentan. El análisis de la varianza concluye en que las diferencias son significativas por lo que el tamaño del error cometido si depende del diario que publica la encuesta (F= 6,391 p-value: 0,000874)

Como puede apreciarse la ordenación de mayor a menor error sería El País, El Periodico, La Razón , El Mundo y ABC, mientras que la ordenación del cuadro anterior de menor a mayor tamaño muestral es ABC,El Mundo, El Periodico, El País y La razón con lo que ya podemos afirmar la ausencia de relación entre estas dos variables.

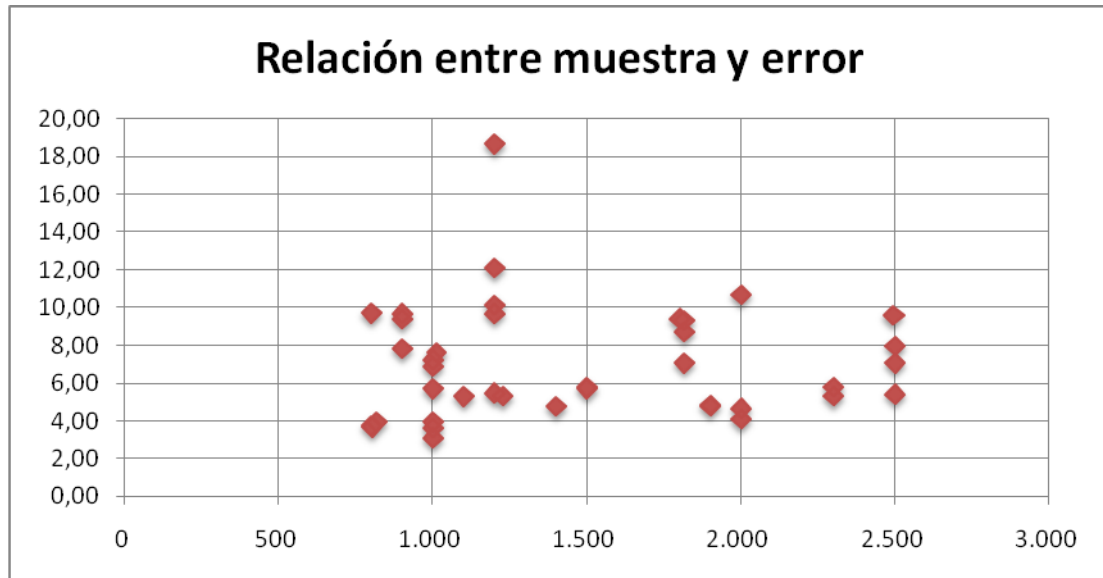
VII.- Error cometido en las encuestas según la empresa que las realiza

Empresa	Media ECM
NC Report	5,81
GAD3	4,94
Metroscopia	10,50
GESOP	8,67
SIGMA DOS	5,07
CIS	8,79

Metrocopia, CIS y Gesop son las empresas demoscópicas que mayor error cuadrático medio comenten mientras que GAD 3 y Sigma Dos son las empresas que menor error presentan. El análisis de la varianza concluye en que las diferencias son significativas por lo que el error cometido si depende de la empresa que realiza la encuesta.(F=5,45, P-Value=0,00118

Al igual que con los medios, no se encuentra relación en el caso de la empresa y/o institución que realiza la encuesta entre error y tamaño muestral

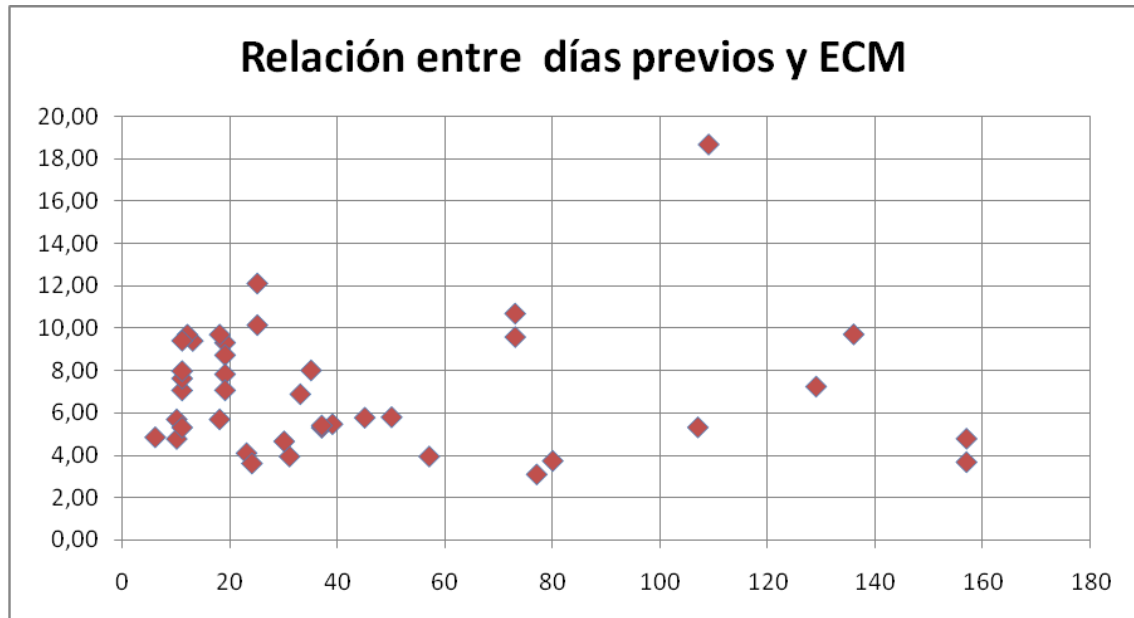
VIII.- Relación entre error de la encuesta y el tamaño muestral



No se ha encontrado relación entre el tamaño de la muestra y el error cometido, medido este mediante el error cuadrático medio (ECM). La correlación entre el ECM y el tamaño muestral es de 0,054 (con un p-valor=0,74), lo que demuestra la ausencia de relación entre ambas variables.

Cabe señalar que en el gráfico no se incluye las encuestas del CIS por problemas de escala.

IX. Relación entre error de la encuesta y días previos a las elecciones



Analizando los días que faltan para las elecciones desde el momento de realización de la encuesta con el ECM, no se encuentra una relación significativamente. El coeficiente de correlación entre ambas variables es de 0,01 con un p-valor de 0,9491.

X.- Relación entre el trimestre y el ECM según la empresa que realiza las encuestas

Empresa	Primer trimestre(Enero-Marzo)				Segundo trimestre(Abril-Junio)				Total
	Enero	Febrer o	Marz o	Trimestre	Abril	Mayo	Junio	Trimestre	
NC Report	4,78	-	5,32	5,05	10,69	5,41	4,10	5,90	5,81
GAD 3	3,68	9,71	-	6,69	3,84	3,94	4,81	4,25	4,94
Metroscopia	-	-	18,67	18,67	7,97	5,74	10,34	9,13	10,50
GESOP	-	-	-	-	-	-	7,43	7,43	8,67
SIGMA DOS CIS	-	7,24	-	7,24	3,09	-	5,00	4,52	5,07
TOTAL	4,23	8,48	11,99	8,23	6,50	5,69	7,48	6,83	7,04

La falta de relación entre la antelación de la encuesta y el error cuadrático medio se observa también en este cuadro con resultados por meses y trimestres ya que en general según las encuestas se aproximan a las elecciones no disminuyen los errores cometidos.

El coeficiente de correlación entre los trimestres y el error cuadrático medio es de 0,16 con un p-value de 0,2993.

XI.- Error cometido con PP según la empresa que realiza la encuesta

		Error cometido con PP (*)				
		Sobrevalora Mucho	Sobrevalora	Ajusta	Infravalora	Infravalora Mucho
Empresa	CIS	0%	0%	0%	0%	100%
	GAD3 (ABC)	0%	0%	0%	43%	57%
	GESOP (El Periódico)	0%	0%	0%	0%	100%
	Metroscopia (El País)	0%	0%	0%	0%	100%
	NC Report (La Razón)	0%	0%	0%	88%	12%
	SIGMA DOS (El Mundo)	0%	0%	0%	60%	40%

(*) El error se ha recodificado en los cinco niveles siguientes: “Sobrevalora Mucho” < -3%, -3%< “Sobrevalora” < -2%, -2%< “Ajusta” <2%, 2%< “Infravalora” <3% y “Infravalora Mucho”> 3%.

La práctica totalidad de las empresas y periódicos tendieron a infravalorar los resultados del PP y le asignaron unos resultados inferiores, en más del 3%, a los finalmente registrados. Esto fue más acusado en el caso del CIS, Gesop y Metroscopia, mientras que NC Report y Sigma 2 infravaloraron en un porcentaje menor, entre el 2% y el 3%.

XII.- Error cometido con PSOE según la empresa que realiza la encuesta

		Error cometido con PSOE (*)				
Empresa		Sobrevalora Mucho	Sobrevalora	Ajusta	Infravalora	Infravalora Mucho
	CIS	0%	0%	100%	0%	0%
	GAD3 (ABC)	0%	0%	100%	0%	0%
	GESOP (El Periódico)	0%	0%	100%	0%	0%
	Metroscopia (El País)	0%	0%	29%	71%	0%
	NC Report (La Razón)	0%	0%	100%	0%	0%
	SIGMA DOS (El Mundo)	0%	0%	40%	60%	0%

(*) El error se ha recodificado en los cinco niveles siguientes: "Sobrevalora Mucho" < -3%, -3% < "Sobrevalora" < -2%, -2% < "Ajusta" < 2%, 2% < "Infravalora" < 3% y "Infravalora Mucho" > 3%.

Los resultados para el PSOE fueron los más ajustados salvo en el caso de Metroscopia y Sigma dos que tendieron ligeramente a infravalorarlos.

XIII.- Error cometido con UNIDOS PODEMOS según la empresa que realiza la encuesta

		Error cometido con UNIDOS PODEMOS (*)				
Empresa		Sobrevalora Mucho	Sobrevalora	Ajusta	Infravalora	Infravalora Mucho
	CIS	50%	50%	0%	0%	0%
	GAD3 (ABC)	43%	0%	57%	0%	0%
	GESOP (El Periódico)	83%	17%	0%	0%	0%
	Metroscopia (El País)	57%	14%	29%	0%	0%
	NC Report (La Razón)	75%	12,5%	12,5%	0%	0%
	SIGMA DOS (El Mundo)	40%	20%	40%	0%	0%

(*) El error se ha recodificado en los cinco niveles siguientes: "Sobrevalora Mucho" < -3%, -3% < "Sobrevalora" < -2%, -2% < "Ajusta" < 2%, % < "Infravalora" < 3% y "Infravalora Mucho" > 3%.

Unidos Podemos fue sobrevalorado por todas las encuestas con resultados resultaros inferiores a los estimados. Dentro de esta sobrevaloración general, Gad 3 y Sigma 2 serían las empresas con menor error.

XIV.- Error cometido con CIUDADANOS según la empresa que realiza la encuesta

		Error cometido con CIUDADANOS (*)				
Empresa		Sobrevalora Mucho	Sobrevalora	Ajusta	Infravalora	Infravalora Mucho
	CIS	0%	50%	50%	0%	0%
	GAD3 (ABC)	0%	0%	100%	0%	0%
	GESOP (El Periódico)	0%	67%	33%	0%	0%
	Metroscopia (El País)	57%	29%	14%	0%	0%
	NC Report (La Razón)	0%	12%	88%	0%	0%
	SIGMA DOS (El Mundo)	0%	40%	60%	0%	0%

(*) El error se ha recodificado en los cinco niveles siguientes: "Sobrevalora Mucho" < -3%, -3% < "Sobrevalora" < -2%, -2% < "Ajusta" < 2%, % < "Infravalora" < 3% y "Infravalora Mucho" > 3%.

Ciudadanos obtuvo unas estimaciones bastante ajustadas a los resultados finales, salvo en el caso de Metroscopia que en general sobreestimó la intención de voto. Por su parte Gad3 y NC Report cometieron un bajo error en la estimación de este partido

XV.- Dependencia entre error cuadrático medio por partido políticos y el periódico que publica la encuesta

	ABC	La Razón	El País	El Periódico	El Mundo	F	P-Value
PP	3,40	2,72	4,42	4,88	3,11	6,92	0,00
PSOE	1,06	1,41	1,88	1,53	1,40	0,77	0,55
UNIDOS PODEMOS	-2,37	-3,50	-3,10	-3,35	-2,42	1,18	0,34
CIUDADANOS	-1,35	-1,19	-3,49	-2,2	-1,49	6,99	0,00

Los datos indican que existen diferencias significativas de los errores que se comenten para un partido concreto según el periódico que publica o encarga la encuesta, excepto en el caso del PSOE y UNIDOS PODEMOS

Según las medias de error que se presentan en la tabla, podemos ver que algunos periódicos sobreestiman (signo -) y otros subestiman (signo +) los votos que recibiría cada partido. Los diferentes periódicos sobreestimaron a CIUDADANOS, y subestimaron a PP, siendo significativas estas diferencias según el medio que publica la encuesta. Así El País y El Periódico sobrestimaron más al PP que a otros partidos mientras que CIUDADANOS lo sobrestimaron estos mismos medios tratados.

V.-CONCLUSIONES

1. En general se observa un error elevado en las encuestas realizadas y que supera al error teórico asignado según las fórmulas del muestreo aleatorio simple publicados en las diferentes fichas de las encuestas
2. La bibliografía existente señala como principales causas de los errores los cambios de preferencia de voto, la sobreestimación del electorado movilizado, la presencia de voto oculto, la configuración de las muestras y métodos de muestreos y la utilización de las encuestas para atraer electores.
3. Los errores medios cometidos fueron más elevados para el PP (3,7%) y UNIDOS PODEMOS (-3%), mientras que se cometieron menos errores con el PSOE (1,43) y con otros partidos (-0,18). Signos negativos indican sobre-estimación de resultados de la encuesta mientras que signos positivos sub-estimación.
4. Se presenta mayor dispersión o menor coincidencia de las encuestas con las estimaciones para PP y UNIDOS PODEMOS, mientras que PSOE ha sido el partido con menor dispersión y mayor coincidencia.
5. En general para el PP y PSOE las encuestas infraestimaron el número de votos y por el contrario para UNIDOS PODEMOS, CIUDADANOS sobreestimaron el resultado. El mayor error se cometió para el PP con casi un 4% de error medio.
6. La práctica totalidad de las encuestas, con la excepción de las realizadas por el CIS, se han realizado telefónicamente, lo que podría ser una fuente del sesgo o error cometido.
7. Las encuestas publicadas por El País y El Periodico fueron las que mayor error cuadrático medio cometieron y por el contrario las publicadas por El Mundo y ABC las que menor error presentan.
8. No se ha encontrado relación significativa entre el tamaño de la muestra y el error cometido, medido este mediante el error cuadrático medio (ECM), si bien ningún tamaño muestral es inferior a 800 entrevistas.
9. No existe relación entre los días que transcurren desde que se realiza la encuesta y las elecciones con el error cometido
10. Existe dependencia fuerte entre el error cometido con el PP y el medio que publica la encuesta, menor error en el caso de La Razón y El Mundo y mayor en El País y el Periódico.
11. El error cometido con CIUDADANOS depende del medio que publica la encuesta. Las encuestas de El País son las que más error cometen (sobrestiman) mientras que ABC y La Razón se ajustan más a los resultados reales.

VI.-BIBLIOGRAFÍA

ABC. (2016). Obtenido de <http://abc.es/>

Advice Strategic Consultans. (2016). Obtenido de <http://advicestrategicconsultans.com/>

Cadena SER. (2016). Obtenido de <http://cadenaser.com/>

El Confidencial. (2016). Obtenido de <http://www.elconfidencial.com/>

El Diario. (2016). Obtenido de <http://www.eldiario.es/>

El Mundo. (2016). Obtenido de <http://www.elmundo.es/>

El País. (2016). Obtenido de <http://www.elpais.com/>

El Periódico. (2016). Obtenido de <http://www.elperiodico.com/es/>

Europa Press. (2016). Obtenido de <http://www.europapress.es/>

Índice de Opinión Pública. (2016). Obtenido de <http://www.simplelogica.com/iop/indices-opinion-publica.asp>

La Razón. (2016). Obtenido de <http://larazon.es/>

Vicéns, J. (2010). *Obtención y análisis de datos*. Madrid.

VII.-ANEXO

Anexo I. Base de Datos

MEDIO	FECHA PUBLICACION	EMPRESA	FECHA REALIZACION	DIFERENCIA DIAS	MUESTRA	ERROR %	MUESTREO	TRABAJO DE CAMPO	ESTRATOS	ECM sin abstención	PP	PSOE	UNIDOS PODEMOS	CIUDADANOS	OTROS
El País	13-mar-16	Metroscopia	09-mar-16	109	1.200	2,90	Aleatorio Simple	Telefónica	CCAA, Edad	18,67	7,03	-0,44	-1,10	-6,45	0,96
El País	22-may-16	Metroscopia	18-may-16	39	1.200	2,90	Aleatorio Simple	Telefónica	CCAA, Sexo, Edad	5,47	3,13	2,46	-2,10	-2,45	-1,04
El País	01-may-16	Metroscopia	01-jun-16	25	1.200	2,90	Aleatorio Simple	Telefónica	CCAA, Sexo, Edad	10,14	4,03	2,36	-3,60	-3,85	1,06
El País	05-jun-16	Metroscopia	01-jun-16	25	1.200	2,90	Aleatorio Simple	Telefónica	CCAA, Municipio, Sexo, Edad	12,11	4,53	2,46	-4,50	-3,55	1,06
El País	12-jun-16	Metroscopia	08-jun-16	18	1.200	5,00	Aleatorio Simple	Telefónica	CCAA, Sexo, Edad	9,69	4,13	1,86	-4,30	-2,85	1,16
Cadena Ser	02-jun-16	MyWord	24-may-16	33	1000	3,10	Aleatorio simple	Entrevistas online	Sexo, edad, socioeconómico, Zona Nielsen y hábitat	6,88	3,83	1,96	-3,70	-1,15	-0,94
La Razón	17-mar-16	NC Report	11-mar-16	107	2.300	2,10	Aleatorio proporcional	Telefónica	CCAA, Edad	5,32	3,73	0,56	-2,80	-2,05	0,56
La Razón	09-may-16	NC Report	07-may-16	50	2.300	2,10	Aleatorio proporcional	Telefónica	Municipio, Sexo, Edad	5,79	2,83	1,86	-3,60	-1,95	0,86
La Razón	06-jun-16	NC Report	03-jun-16	23	2.000	2,20	Aleatorio Simple	Telefónica		4,10	2,53	1,26	-3,40	-0,85	0,46
La Razón	14-abr-16	NC Report	14-abr-16	73	2.000	2,24	Aleatorio Simple	Telefónica	Municipio	10,69	2,33	1,46	-5,70	-1,45	3,36

MEDIO	FECHA PUBLICACION	EMPRESA	FECHA REALIZACION	DIFERENCIA DIAS	MUESTRA	ERROR %	MUESTREO	TRABAJO DE CAMPO	ESTRATOS	ECM sin abstención	PP	PSOE	UNIDOS PODEMOS	CIUDADANOS	OTROS	Índice de Opinión Pública
La Razón	23-may-16	NC Report	20-may-16	37	2.500	2,24	Aleatorio Simple	Telefónica	CCAA, Municipio, Edad	5,40	2,63	1,56	-3,80	-1,45	1,06	20-jun-16
La Razón	30-may-16	NC Report	27-may-16	30	2.000	2,24	Aleatorio Simple	Telefónica	CCAA, Municipio, Edad	4,65	2,53	1,36	-3,60	-1,15	0,86	20-jun-16
La Razón	25-ene-16	NC Report	21-ene-16	157	1.900	2,30	Aleatorio Simple	Telefónica	Sexo, Edad	4,78	2,33	1,56	-1,30	1,05	-3,64	20-jun-16
La Razón	16-may-16	NC Report	12-may-16	45	1.500	2,50	Aleatorio Estratificado	Telefónica		5,77	2,83	1,66	-3,80	-1,65	0,96	20-jun-16
EL MUNDO	18-jun-16	SIGMA DOS	08-jun-16	18	1.500	2,58	Aleatorio Simple	Telefónica	Edad	5,69	2,53	2,66	-3,70	-1,05	-0,44	20-jun-16
EL MUNDO	22-feb-16	SIGMA DOS	18-feb-16	129	1.000	3,16	Aleatorio Simple	Telefónica	Edad	7,24	5,23	-0,44	-1,70	-2,25	-0,84	20-jun-16
EL MUNDO	23-abr-16	SIGMA DOS	10-abr-16	77	1.000	3,16	Aleatorio Simple	Telefónica	Edad	3,09	3,23	-0,24	-0,40	-2,15	-0,44	20-jun-16
EL MUNDO	05-jun-16	SIGMA DOS	02-jun-16	24	1.000	3,16	Aleatorio Simple	Telefónica		3,61	2,03	2,36	-2,60	-0,95	-0,84	20-jun-16
EL MUNDO	19-jun-16	SIGMA DOS	16-jun-16	10	1.000	3,16	Aleatorio Simple	Telefónica	Edad	5,69	2,53	2,66	-3,70	-1,05	-0,44	20-jun-16
Simple Lógica	20-jun-16		15-jun-16	11	1.226	2,86	Aleatorio Simple	Telefónica	Edad	5,31	3,83	0,16	-2,40	-2,35	0,76	20-jun-16