

# Diplomatura d'Estadística

---

**Títol:** L'ENCERTEN ELS SONDEJOS ELECTORALS ?

**Autor:** Adrià Caballé Mestres

**Director:** Pere Grima Cintas

**Departament:** Departament d'Estadística i Investigació Operativa

**Convocatòria:** Juny de 2010



Facultat de Matemàtiques  
i Estadística

UNIVERSITAT POLITÈCNICA DE CATALUNYA

## ÍNDIX

1	INTRODUCCIÓ .....	6
2	RECOLLIDA DE LES DADES .....	8
2.1	Tria de les convocatòries electorals que volem estudiar .....	8
2.2	Procedència de les dades .....	8
2.3	Dificultats de la recerca.....	9
2.4	Les dades .....	10
2.4.1	Dades de les eleccions al Parlament de Catalunya .....	10
2.4.2	Dades de les eleccions al Congrés dels Diputats.....	14
3	ASSIGNACIÓ D'ESCONS EN FUNCIÓ DEL PERCENTATGE DE VOTS .....	18
3.1	Percentatge llindar .....	18
3.2	La llei de Hondt .....	18
4	SONDEJOS TEÒRICS .....	22
4.1	Metodologia teòrica per crear les mostres: ús de la distribució multinomial...	23
4.2	Tria de la mida mostral .....	24
4.2.1	Mides mostrals al Parlament de Catalunya .....	24
4.2.2	Mides mostrals al Congrés dels Diputats.....	28
4.3	Sondejos teòrics amb una correcció en les proporcions de vot.....	30
5	OBTENCIÓ DE LES MOSTRES PELS SONDEJOS TEÒRICS .....	31
5.1	Constants emprades : mides mostrals i proporcions de vots.....	31
5.2	Obtenció de les proporcions de vot.....	32
5.3	Metodologia aplicada per l'obtenció d'escons .....	32
6	DEFINICIÓ DE LES DISTÀNCIES .....	34
6.1	Selecció de les distàncies .....	34
6.2	Una única distància per sondeig: justificació .....	37
6.3	Càlcul de les distàncies .....	37
6.4	Distàncies entre els sondejos i els resultats reals.....	41
7	REPRODUCCIÓ DE LES DISTÀNCIES OBTINGUDES.....	43
7.1	Una altra via: Anàlisi de components principals .....	44
7.1.1	Gràfiques d'ACP en el Parlament de Catalunya .....	45
7.1.2	Gràfiques d'ACP en el Congrés dels Diputats .....	49
8	ANÀLISI DE LES DISTÀNCIES I COMPARACIONS .....	53

8.1	Evolució dels sondejos al llarg dels anys .....	53
8.1.1	Evolució dels sondejos al Parlament de Catalunya .....	53
8.1.2	Evolució dels sondejos al Congrés dels Diputats .....	55
8.2	Exactitud dels sondejos publicats .....	57
8.2.1	Exactitud dels sondejos al Parlament de Catalunya .....	58
8.2.2	Exactitud dels sondejos al Congrés dels Diputats .....	59
8.3	Tendències polítiques dels diaris .....	62
8.3.1	Tendències polítiques al Parlament de Catalunya .....	63
8.3.2	Tendències polítiques al Congrés dels Diputats .....	63
8.4	Anàlisi dels errors en les principals candidatures .....	66
9	CONCLUSIONS .....	75
10	BIBLIOGRAFIA .....	77

## PRESENTACIÓ

---

En el total dels quatre mesos que hem treballat per a la realització d'aquest projecte, han estat moltes les sorpreses, habilitats que he adquirit i també les dificultats que m'han sorgit, i m'han ajudat a créixer i madurar tant en l'aspecte personal com professional.

La idea de realitzar un projecte sobre aquesta temàtica va sorgir a partir de la lectura d'un document escrit pels professors Pedro Delicado (UPC) i Federic Ubina (UPF) amb títol "*¿cómo y cuánto fallan los sondeos electorales?*". Ens va semblar un problema molt interessant d'abordar i hem intentat il·lustrar, des de un punt de vista diferent al que ells havien proposat, de quina forma erren els sondejos electorals que publiquen els diaris de més tirada. Ho hem realitzat en dues de les eleccions que hi ha en l'Estat Espanyol. Una, més de caràcter estatal, com són les eleccions al Congrés dels Diputats (o eleccions generals). La segona, fa referència a les eleccions al Parlament de Catalunya.

El primer que hem tingut d'utilitzar són les dades de totes les convocatòries que s'han produït en els últims anys. Tant els resultats que es varen donar al dia de la celebració, com els sondejos publicats pels diaris una setmana abans del dia clau. Trobar les dades referents als resultats és fàcil. Entretenint-nos una mica en la web ja varem trobar un parell de pàgines que ens proporcionaven el que buscàvem. En canvi, no podem dir el mateix a l'hora de localitzar els sondejos. Hem tingut alguns problemes per aconseguir les diferents prediccions que varen fer al seu dia els diaris. La raó és que hi ha convocatòries que no es varen publicar i també cal tenir en compte que hi ha diaris que no tenen suficients anys de hemeroteca per cercar si en van fer o no en el seu dia.

El que hem buscat és la manera de comparar i quantificar de forma apropiada les diferències entre les prediccions dels diaris i els resultats que realment es varen donar. Hem creat una distància que ens quantifiqués d'una forma justa aquests errors de predicció.

També i per tal de saber quan podríem considerar un sondeig com a ben fet hem establert un sondeig teòric. Les enquestes que fan els diaris fan inferència de mostres petites respecte una població molt gran. Les proporcions de vot obtingudes de l'enquesta, lògicament, serà pràcticament impossible que siguin les que realment es donaran el dia de les eleccions. Existeix una variabilitat. Amb aquests sondejos teòrics es podrà esbrinar entre quins escons es poden moure els partits únicament tenint present el factor atzar. Aquests seran els sondejos ideals. Si les distàncies que presenten els diaris són inferiors a una fita superior que hem establert a partir d'aquests, direm que el pronòstic del diari és excel·lent i immillorable.

Tenint sempre present que hi ha un factor que els sondejos (per molt que aleatoritzin) no poden controlar, com són els individus enquestats que canviïn d'opinió o menteixin,

així hem pensat i hem realitzat un retoc en el sondeig teòric inicial. Aquest ja no serà un sondeig perfecte, però si un de més realista i prudent a l'hora de castigar els sondejos dels diaris. Ens proporciona una altra frontera per considerar si la predicció del diari és acceptable, atenent a les dificultats induïdes per aquest factors de soroll.

Per a la realització d'aquests sondejos teòrics ens hem endinsat en la mecànica que aplica la legislació espanyola per a l'assignació d'escons a partir de les proporcions de vot: la llei de Hondt.

Conegudes totes les distàncies, tant dels sondejos teòrics com dels diaris, hem mirat com s'han aproximat les prediccions a la realitat. Quins anys han estat millors i quins pitjors, les diferències entre els diaris, el seu posicionament a afavorir una certa tendència política, i molts altres aspectes que podreu descobrir en el treball.

Cal mencionar que, possiblement, és en aquest camp dels sondejos electorals on el ciutadà, sense cap mena d'experiència amb l'estadística, identifica als estadístics. S'equivoquen perquè erren estrepitosament o perquè no ho poden fer millor?

Un cop feta aquesta petita presentació del que ha estat aquest projecte, i al ser aquest treball l'última passa per aconseguir la diplomatura, m'agradaria agrair al conjunt de professorat que he tingut al llarg dels tres anys. Sempre n'hi ha alguns en concret que els hi tens més predilecció, però el cert és que de tots n'he pogut aprendre quelcom nou. Especialment donar les gràcies al meu director del projecte, el senyor Pere Grima, primer per haver tingut l'oportunitat de realitzar el projecte amb ell, i sobretot per haver pogut treballar en la seva elaboració amb tant bon clima entre els dos i sempre amb plena comoditat.

# 1 INTRODUCCIÓ

---

Fa molts anys ja de l'existència de varies eleccions arreu de l'estat espanyol. En la major part d'elles, s'ha pronosticat quins serien els escons per a cada un dels partits que s'hi presenten, abans que arribés el dia de les votacions. Aquestes prediccions les publiquen diferents diaris i es coneixen com a sondejos electorals.

Se'n presenten múltiples durant el període de transició entre dues convocatòries, però els que seran de més interès són els que es realitzen una setmana abans de les eleccions. El motiu principal és que aquests poden contrastar els pronòstics, amb els escons resultants 7 dies després. En canvi els altres, els que publiquen fora d'aquestes dates tant pròximes, no es pot saber del cert si fallen o no.

Els sondejos no són molt transcendents, o importants, pel dia a dia de la societat, però el cert és que cada vegada que se'n presenta un, és portada en els diaris més llegits del país. Per tant, i encara que només siguin per satisfer la curiositat de la població, són d'interès popular.

Hi ha una sensació generalitzada arreu del país que s'equivoquen molt. De fet, són meres hipòtesis sense una base forta on agafar-se. Es pretén quantificar aquestes sensacions que té la població i no quedar-se únicament amb opinions no contrastades. S'ha de tenir en compte que l'elaboració d'una bona enquesta per a la recollida de vots és verdaderament complicada. Aquest estudi no entra en detall en com ho organitzen, si bé es poden destacar a priori algun dels motius que la poden fer imprecisa:

- En tot l'estat espanyol hi ha molts territoris amb dret a vot. No només hi ha diferents tendències polítiques en cada autonomia. Cada província és diferent. Fins i tot cada ciutat i poble pensen de forma heterogènia. Aconseguir una mostra representativa que presti la importància adequada a cada un dels pobles és molt complexa i, a més, té un cost molt elevat.
- Hi ha partits que són mal vistos, o si més no, no queda gaire bé manifestar el desig a votar-los. Tot i que les enquestes són anònimes i ningú sabrà per quin partit t'has inclinat, hi ha una tendència a mentir.
- Per últim, és molt típic que a una setmana vista, algun votant que s'ha decantat en l'enquesta per un dels partits canviï de preferència en darrer moment. Aquest fet pot ser degut a les campanyes electorals, o als fets extra polítics que es produeixen en aquesta setmana d'impàs. Un exemple molt clar seria el de les eleccions generals de març de 2004, quan amb els atemptats produïts a l'estació d'Atocha (Madrid), el Partit Popular va sortir desfavorit de forma molt evident. Aquest cas el podrem repassar en l'estudi.

La comparació entre sondejos i resultats es realitza per mitjà d'una distància. Hi ha varies vies per establir-ne una, i no n'existeix una sola de vàlida. S'ha triat una de les opcions intentant que fos comprensible pel gran públic. És a partir d'aquesta que es poden analitzar numèricament com de allunyats estan cada un dels sondejos publicats respecte dels valors reals.

És d'interès saber si hi ha diaris que s'aproximen més que d'altres a la realitat, o si pel contrari, no hi ha diferències abismals entre ells. A més, comprovar si hi ha la tendència d'algun d'ells a pronosticar a l'alça per un partit en concret. No cal obviar que se'ls atribueix sovint una determinada ideologia política. Dos exemples coneguts podrien ser el del diari El Mundo, pel Partit Popular, i El País pel Partit Socialista.

També es vol comprovar si, entre totes les eleccions que estudiem, hi ha anys millors que d'altres en el pronòstic, o fins i tot, si hi ha una millora en quant a precisió a mesura que han anat passant els anys.

Un altre punt que prestem atenció consisteix en la comparació dels dos tipus d'eleccions que s'analitzen amb la mateixa mecànica (recordem que s'estudien les eleccions al Parlament de Catalunya i al Congrés dels Diputats). Pretenem saber si els sondejos són igual de precisos a Catalunya que a l'Estat (hem de tenir present que en tot l'Estat hi ha molts més votants i per tant regions molt més heterogènies).

Podria ser una mica injust i insuficient només comparar els sondejos amb els resultats reals. El que fan els sondejos és fer inferència d'una part molt petita de la població. Les mostres són de l'ordre de 1.000-2.000 individus sobre més de 5 milions de votants a Catalunya i de 2.000-12.000 individus sobre més de 30 milions en tot l'estat Espanyol. Seria més oportú considerar que al tractar amb mostres tant petites la comparació fos considerant mides de mostra similars. Perquè així sigui, es podrien desenvolupar una sèrie de sondejos teòrics amb els paràmetres reals. Es a dir, es pot pensar en simular la realització de moltes enquestes a partir de les proporcions de vots de cada partit que es van donar el dia de les eleccions (les suficients per poder distingir l'aleatorietat del procés). Hauran de ser instaurades tenint en compte totes les circumscripcions. En el cas de les del Congrés dels Diputats en són 52 i en el Parlament de Catalunya 4.

Un cop aconseguides les mostres, es vol desenvolupar una metodologia per comparar-les amb els sondejos. D'aquesta manera s'esbrinarà si les diferències que es veien en la comparació amb els resultats reals són per pur atzar (on es podria considerar com a bo el sondeig), o si pel contrari, les diferències són massa grans per ser degudes al factor aleatori (ara el sondeig seria incorrecte).

Les metodologies de comparacions entre els sondejos i els resultats reals, tenint en compte els sondejos teòrics, estan acompanyades d'una bona representació gràfica. D'aquesta forma es reconeixerà fàcilment per a qualsevol lector quant i com falla un sondeig.

## **2 RECOLLIDA DE LES DADES**

---

### **2.1 Tria de les convocatòries electorals que volem estudiar**

Des de la restauració de la democràcia, que data de l'any 1976, s'han succeït tota una sèrie d'eleccions en el conjunt de l'Estat Espanyol. Sobretot se'n destaquen 3 tipus.

Una és la del Congrés dels Diputats. Es realitzen votacions a tot Espanya des de 1977 i ja s'han dut a terme 10 eleccions.

Una de les altres eleccions són les de les autonomies. Cada una de les 17 que hi ha en el total de l'estat té la seva pròpia metodologia i data de realització. Per proximitat es pot destacar la del Parlament de Catalunya. Es dóna des de 1980 i, fins ara, consta de 8 eleccions.

Les últimes que caldria esmentar són les eleccions municipals. A Catalunya hi ha 4 diputacions, on hi consten 946 municipis. La data de l'escrutini és la mateixa per a tots. Fa des de 1979 que es van instaurar i ja se n'han succeït 8 celebracions.

En totes tres, l'interval entre dues convocatòries és de 4 anys. Malgrat que hi ha casos on aquest temps s'ha escurçat.

En aquest treball únicament s'estudia a fons les dues primeres, les del Congrés dels Diputats i les del Parlament de Catalunya. S'ha descartat l'estudi de les municipals ja que són molts municipis els que s'haurien d'haver analitzat i, a molts d'ells, no es disposava ni tant sols dels sondejos. L'únic que es podria haver inclòs en el treball correspon a les eleccions de Barcelona, però hem decidit que, de moment, no era procedent.

Una altra possibilitat hagués estat fixar-se amb processos electorals de fora Espanya, i comparar-los amb els d'aquí. Aquest ja seria un treball molt més llarg i costós de realitzar. Podria ser un bon projecte pel futur.

### **2.2 Procedència de les dades**

La realització dels sondejos no l'executa directament els diaris. El que fan és subcontractar una empresa especialitzada en aquest tipus d'enquesta. Se'n destaquen els següents instituts :

OPINA, Demoscopia, Gesop, NOXA, Vox Pública, DYM, Metroscopia, SOFEMASA, IOPE-ETMAR i METRA/SEIS.



Primerament es tenia la intenció d' analitzar únicament els diaris El País, La Vanguardia i El Periódico, atès que desperten un interès més gran a tot Catalunya. Tot i això i donat que hi ha altres diaris de caràcter més estatal que també proposen un sondeig, s'ha decidit incloure'ls en l'estudi. Aquests són l'ABC i El Mundo.

En el cas del Congrés dels Diputats s'han trobat els resultats de les eleccions en la pàgina web del "Ministerio del Interior" (link: [www.MIR.es](http://www.MIR.es)). Els sondejos s'han cercat en l'hemeroteca de cada un dels diaris esmentats.

Pel que fa a les eleccions del Parlament de Catalunya s'han localitzat tant els resultats com la major part de sondejos en la mateixa pàgina web. És la que disposa la Generalitat de Catalunya (link: [www.gencat.cat](http://www.gencat.cat)) en l'apartat de governació. Els sondejos que no estaven en la pàgina s'han cercat en la hemeroteca corresponent a cada diari.

### **2.3 Dificultats de la recerca**

A primera vista, sembla relativament fàcil localitzar aquestes dades a Internet però, el cert és que no és tant trivial. En la pàgina web de la Generalitat de Catalunya estan molt ben especificats els resultats i prediccions, però, no hi són tots. Aquestes dades que manquen, precisament no les destaca la web perquè són complicades d'aconseguir.

Ho són per diferents motius:

- Moltes vegades l'edició escrita dels diaris el dia que presenten les prediccions no informen de tot el que han trobat en l'enquesta. Hi ha ocasions que només donen els possibles escons dels dos partits principals, o fins i tot la distància amb punts entre un i altre. Aquest va ser el cas de l'edició d'El País el dia 2 de març de 2008 en les eleccions generals. En altres casos només informen de les proporcions de vots totals a cada partit polític. Aleshores és impossible passar aquests percentatges a escons (per fer-ho es necessiten les proporcions estimades a cada circumscripció), que al final és el que es té que menester.
- Un altre problema són les hemeroteques d'alguns diaris. No tenen edicions antigues. S'accedeix en El Periódico i El Mundo. Aquests només ofereixen els últims 6 i 12 anys respectivament. En conseqüència, molts dels possibles sondejos que varen publicar al seu dia abans d'aquestes dates són molt difícils d'aconseguir.
- A més, hi ha anys que no es varen publicar les prediccions. Aquests solen ser els primers anys de democràcia, on la tradició de fer sondejos encara no estava tant arrelada.

També cal mencionar que tot i que s'estan analitzant els sondejos una setmana abans de la celebració, no tots els diaris presenten els seus pronòstics el mateix dia. Cal afegir a la recerca una feina extra d'anàlisis dels dies propers a aquests últims 7 dies de campanya.

## 2.4 Les dades

En aquest apartat es presenten en taules resum les dades que s'utilitzen en el treball. S'hi especifiquen les proporcions de vot i els escons totals a cada partit, tant de les eleccions generals com de les de la Generalitat. No només s'hi destaca els resultats que es varen donar a les urnes. També hi ha les prediccions que presenten els diaris i el interval al 95% que s'obté al realitzar els sondejos perfectes que introduïrem en següents temes (tema 3, 4 i 5).

En les taules, les dades absents marcades amb el símbol “ – “ signifiquen casos on el diari en qüestió no va presentar un sondeig, o bé que un partit polític no va tenir representació. Pel que fa a les marcades amb el símbol “ \* “ representen casos on el diari sí que va presentar un sondeig però no va publicar les proporcions de vot

### 2.4.1 Dades de les eleccions al Parlament de Catalunya

En el Parlament de Catalunya s'han produït 8 disputes d'eleccions per la repartició dels diputats. Es reparteixen 135 escons. Cada una de les 4 províncies tenen fixats un nombre concret de diputats que s'assignen segons les votacions en els seus municipis. La primera de totes data de l'any 1980 i l'última de l'any 2006.

Diferents partits polítics han canviat el seu nom durant aquest període. És el cas del Partit Popular (abans Aliança Popular), o també, Iniciativa per Catalunya - els Verds (on durant molt temps les dues candidatures van estar separades).

També destaquen partits com el PSUC (Partit Socialista Unificat de Catalunya). Aquest va tenir un pes important a la Generalitat durant els primers anys de la democràcia (de 1980 - 1984). En les següents eleccions ja no es va presentar. Atès que tenia una tendència política similar als d'Iniciativa per Catalunya- Els Verds, s'ha usat una sola columna per representar les proporcions i diputats que varen aconseguir. No s'ha cregut oportú fer el mateix amb dos partits, molt minoritaris, i que només van formar part al Parlament en una ocasió cada un. El motiu és que són dos tendències polítiques molt contràries i totalment independents. Aquests són el CDS i CPC.

Atès que els partits polítics solen tenir un nom massa llarg, serà molt habitual usar el seu acrònim per referir-se a ells. Coneixem Convergència i Unió com a CIU, el Partit Socialista de Catalunya com a PSC, el Partit Popular com a PP, Esquerra Republicana de Catalunya com a ERC, Iniciativa per Catalunya- Els Verds com a IC-IV (o també IV i prou), el Partit Socialista Unificat de Catalunya com a PSUC, el Centre Democràtic Social com a CDS i Ciutadans Per la Ciutadania com a CPC.

Malgrat es tenia la intenció d' estudiar totes les 8 convocatòries, no es té constància de cap sondeig dels diaris de les primeres eleccions catalanes. Per tant es prescindirà de l'anàlisi de l'any 1980 per falta de dades. Les taules de proporcions de vot i d'escons de totes les convocatòries estan en la taula 2.4.1.1 i la taula 2.4.1.2 respectivament.

*taula 2.4.1.1 Proporcions de vot dels diferents partits en totes les convocatòries electorals al Parlament de Catalunya.*

ANY	Sondejos/ Resultats	CIU	PSC	PP	ERC	IC-IV (PSUC*)	CDS	CPC
1984	Reals	46,80	30,11	7,70	4,41	5,58*	-	-
	Perfectes	44,47	27,95	6,48	3,47	4,49*	-	-
		49,15	32,31	8,98	5,40	6,71*	-	-
	La Vang.	37,90	30,10	4,50	4,00	2,90*	-	-
	El Periód.	*	*	*	*	*	-	-
	El País	36,90	32,00	9,90	7,30	8,60*	-	-
	ABC	*	*	*	*	*	-	-
El Mundo	*	*	*	*	*	-	-	
1988	Reals	45,72	29,78	5,31	4,14	7,76	3,83	-
	Perfectes	43,31	27,54	4,23	3,21	6,43	2,93	-
		48,12	32,03	6,40	5,14	9,14	4,81	-
	La Vang.	48,00	29,00	6,00	4,00	6,00	3,00	-
	El Periód.	50,20	27,90	5,30	3,40	6,30	4,20	-
	El País	49,30	27,70	6,10	3,10	5,40	3,60	-
	ABC	*	*	*	*	*	-	-
El Mundo	*	*	*	*	*	-	-	
1992	Reals	46,19	27,55	5,97	7,96	6,50	-	-
	Perfectes	43,67	25,27	4,78	6,63	5,25	-	-
		48,69	29,86	7,21	9,32	7,82	-	-
	La Vang.	47,90	28,20	5,61	6,50	7,00	-	-
	El Periód.	*	*	*	*	*	*	*
	El País	46,50	27,60	6,30	6,20	7,10	-	-
	ABC	*	*	*	*	*	*	*
El Mundo	*	*	*	*	*	*	*	
1995	Reals	40,95	24,88	13,09	9,49	9,71	-	-
	Perfectes	38,65	22,90	11,50	8,14	8,26	-	-
		43,26	26,93	14,72	10,86	11,16	-	-
	La Vang.	46,00	21,00	11,00	10,00	10,00	-	-
	El Periód.	46,50	23,20	8,40	8,30	9,40	-	-
	El País	45,40	23,20	11,70	7,20	11,60	-	-
	ABC	44,70	24,30	10,40	9,20	7,80	-	-
El Mundo	45,20	23,80	11,40	8,80	7,40	-	-	
1999	Reals	37,70	37,85	9,51	8,76	2,51	-	-
	Perfectes	35,35	35,44	8,08	7,36	1,88	-	-
		40,01	40,21	10,95	10,02	3,54	-	-
	La Vang.	39,00	36,00	10,50	7,50	4,00	-	-
	El Periód.	40,50	30,30	11,30	9,70	5,60	-	-
	El País	40,10	36,60	10,70	7,70	3,20	-	-
	ABC	37,20	32,90	12,10	8,90	6,70	-	-
El Mundo	40,40	35,70	10,20	7,70	2,60	-	-	

ANY	Sondejos/ Resultats	CIU	PSC	PP	ERC	IC-IV (PSUC*)	CDS	CPC
2003	Reals	30,94	31,16	11,89	16,44	7,28	-	-
	Perfectes	28,79	29,01	10,38	14,70	6,03	-	-
		33,16	33,36	13,45	18,22	8,55	-	-
	La Vang.	33,10	36,20	9,50	13,00	6,90	-	-
	El Periód.	30,90	36,50	10,00	13,70	6,90	-	-
	El País	32,50	34,50	10,50	12,50	6,50	-	-
	ABC	33,20	36,10	10,50	12,00	7,50	-	-
El Mundo	32,20	35,00	12,40	12,80	5,90	-	-	
2006	Reals	31,52	26,82	10,65	14,03	9,52	-	3,03
	Perfectes	29,23	24,56	9,12	12,31	8,01	-	2,16
		33,83	29,01	12,26	15,74	11,02	-	3,94
	La Vang.	28,40	34,00	10,20	12,50	10,90	-	-
	El Periód.	32,00	29,00	11,30	13,20	10,50	-	-
	El País	33,50	29,90	12,10	14,80	8,40	-	-
	ABC	33,70	27,80	10,20	14,80	8,90	-	-
El Mundo	*	*	*	*	*	*	*	

taula 2.4.1.2 Escons dels diferents partits en totes les convocatòries referents a les eleccions al Parlament de Catalunya.

ANY	Sondejos/ Resultats	CIU	PSC	PP	ERC	IC-IV (PSUC*)	CDS	CPC
1984	Reals	72	41	11	5	6*	-	-
	Perfectes	69-76	39-45	8-12	2-7	4-8*	-	-
		La Vang.	66-68	44-46	9-10	6	7-8*	-
	El Periód.	*	*	*	*	*	-	-
	El País	56-58	43-45	13-14	9-10	9-11*	-	-
	ABC	*	*	*	*	*	*	*
	El Mundo	*	*	*	*	*	*	*
1988	Reals	69	42	6	6	9	3	-
	Perfectes	66-74	38-45	4-8	1-7	7-11	0-5	-
		La Vang.	73-76	39-41	7-8	4-5	7-8	0-2
	El Periód.	78	38	6	3	7	3	-
	El País	73-76	38-40	6-7	7-9	0-3	2-3	-
	ABC	*	*	*	*	*	*	*
	El Mundo	*	*	*	*	*	*	*
1992	Reals	70	40	7	11	7	-	-
	Perfectes	67-74	36-43	5-9	9-13	5-9	-	-
		La Vang.	69-72	40-42	5-7	8-9	8	-
	El Periód.	*	*	*	*	*	*	*
	El País	71-72	39-40	8	8	7-8	-	-
	ABC	*	*	*	*	*	*	*
	El Mundo	*	*	*	*	*	*	*

RECOLLIDA DE LES DADES

ANY	Sondejos/ Resultats	CIU	PSC	PP	ERC	IC-IV (PSUC*)	CDS	CPC
1995	Reals	60	34	17	13	11	-	-
	Perfectes	57-64	31-37	15-19	11-15	9-13	-	-
	La Vang.	69-71	28-30	11-14	11-14	9-13	-	-
	El Periód.	69-72	31-35	11-12	10-12	11-12	-	-
	El País	66-70	29-30	14-16	10-11	10-12	-	-
	ABC	62-65	33-34	13-15	10-12	11	-	-
	El Mundo	65-68	31-35	15-16	11-12	8-9	-	-
1999	Reals	56	52	12	12	3	-	-
	Perfectes	54-60	50-56	10-14	10-14	0-3	-	-
	La Vang.	56-58	51-52	13	9-10	4	-	-
	El Periód.	57-63	40-46	14-15	12-15	5-6	-	-
	El País	58-60	48-50	13-14	10	3	-	-
	ABC	53-55	45-47	15-16	12-13	7-8	-	-
	El Mundo	58-62	48-50	12-14	10	3	-	-
2003	Reals	46	42	15	23	9	-	-
	Perfectes	43-49	39-46	13-17	21-26	6-10	-	-
	La Vang.	48-49	49-50	12	18-19	6-7	-	-
	El Periód.	45-47	50-51	12-13	19-20	7-8	-	-
	El País	47-49	48-50	13	17-19	7	-	-
	ABC	49	48-49	13	18-19	6	-	-
	El Mundo	46-49	46-49	15-16	18-20	5-6	-	-
2006	Reals	48	37	14	21	12	-	3
	Perfectes	44-52	35-42	12-16	18-24	10-14	-	0-4
	La Vang.	50-52	38-40	13	17-19	13-15	-	0
	El Periód.	47-49	39-41	14-15	19-20	12-13	-	0
	El País	45	42-43	14	20	13-14	-	0
	ABC	50-52	39	13-14	21	10-11	-	0
	El Mundo	50-52	39	13-14	21	10-11	-	0

### 2.4.2 Dades de les eleccions al Congrés dels Diputats

En el Congrés dels Diputats s'han realitzat 10 disputes d'eleccions. Es reparteixen 350 diputats distribuïts en les 52 circumscripcions. Cada una d'elles se li adjudiquen uns certs escons, que posen en joc el dia de l'escrutini. La primera celebració fou el 1977 i l'última el 2008.

Hi ha forces partits polítics que tenen representació a la cambra, alguns més que en les eleccions de Catalunya. Tot i això, en són pocs els realment importants (partits que acostumen a treure més de 5 diputats) i que s'han repetit en cada convocatòria. N'hi ha 5 i són el Partido Socialista Obrero Español, el Partido Popular, Izquierda Unida, Convergència i Unió i el Partido Nacionalista Vasco. D'altres que també han tingut etapes amb presència al Congrés com el Centro Democrático i Social (en els primers anys) i Esquerra Republicana de Catalunya (en els darrers anys) s'han de tenir en compte i constaran els pronòstics i resultats on han destacat. La resta de partits que també han obtingut algun escó, però que són més minoritaris, estan ajuntats en un sol grup anomenat en la taula 2.4.2.1 i la taula 2.4.2.2 com a *Altres* i prendran importància en totes les convocatòries..

De la mateixa manera que en el Parlament, es sol substituir els noms complets dels partits per les seves sigles. El Partido Socialista Obrero Español és el PSOE, el Partido Popular és el PP, Izquierda Unida el coneixem com IU, Convergència i Unió com a CIU, el Partido Nacionalista Vasco com a PNV, ERC és Esquerra Republicana de Catalunya i CDS el Centro Democrático y Social.

No es té constància de l'existència d'alguns dels sondejos publicats pels diaris. No s'ha trobat les dades corresponents a les tres primeres convocatòries. Per tant, es prescindirà del seu estudi. Únicament s'analitzen a fons el succeït en les darreres 7 celebracions.

Les taules de proporcions de vot i d'escons de totes les convocatòries corresponents a les eleccions generals es troben en la taula 2.4.2.1 i la taula 2.4.2.2 respectivament.

taula 2.4.2.1 Proporcions de vot dels diferents partits en totes les convocatòries electorals al Congrés dels Diputats.

ANY	Sondej/ Result.	PSOE	PP	CIU	IU	PNV	ALTRES	CDS	ERC
1986	Reals	44,06	25,97	5,02	4,63	1,53	4,09	9,22	-
	Perfectes	43,12 44,97	25,18 26,80	4,55 5,42	4,31 5,16	1,33 1,67	3,63 4,27	8,69 9,79	-
	La Vang.	-	-	-	-	-	-	-	-
	El Periód.	42,4	24,1	4,7	4,9	1,7	2,8	11,7	-
	El País	-	-	-	-	-	-	-	-
	ABC	*	*	*	*	*	*	*	-
	El Mund.	-	-	-	-	-	-	-	-
1989	Reals	39,60	25,79	5,04	9,07	1,24	4,86	7,89	-
	Perfectes	38,68 40,54	24,99 26,62	4,58 5,47	8,50 9,65	1,07 1,40	4,56 5,23	7,38 8,42	-
	La Vang.	41,0	24,0	4,8	8,8	2,0	*	9,0	-
	El Periód.	-	-	-	-	-	-	-	-
	El País	41,6	22,8	5,0	8,0	1,5	12,2	9,1	-
	ABC	-	-	-	-	-	-	-	-
	El Mund.	*	*	*	*	*	*	*	-
1993	Reals	38,78	34,76	4,94	9,55	1,24	5,40	-	-
	Perfectes	37,88 39,68	33,89 35,60	4,52 5,35	8,96 10,11	1,07 1,38	5,11 5,83	-	-
	La Vang.	35,0	35,6	5,0	11,4	1,3	10,8	-	-
	El Periód.	34,5	32,5	5,5	10,1	1,5	15,2	-	-
	El País	34,5	34,8	5,0	10,7	1,4	12,9	-	-
	ABC	*	*	*	*	*	*	-	-
	El Mund.	34,6	35,9	4,9	11,4	1,5	3,9	-	-
1996	Reals	37,63	38,79	4,60	10,54	1,27	4,42	-	-
	Perfectes	36,76 38,50	37,93 39,64	4,19 5,00	9,85 11,01	1,11 1,41	4,17 4,83	-	-
	La Vang.	35,0	41,0	4,5	10,5	1,5	7,5	-	-
	El Periód.	33,5	41,2	4,5	11,7	*	7,6	-	-
	El País	33,4	42,4	4,2	12,3	*	7,7	-	-
	ABC	32,5	44,1	4,2	12,8	*	6,4	-	-
	El Mund.	*	*	*	*	*	*	-	-
2000	Reals	34,16	44,52	4,19	5,45	1,53	5,39	-	-
	Perfectes	33,22 35,08	43,57 45,49	3,79 4,59	5,16 6,10	1,34 1,67	5,03 5,81	-	-
	La Vang.	37,2	41,8	4,2	8,0	1,3	7,5	-	-
	El Periód.	37,5	41,7	4,2	7,5	1,6	*	-	-
	El País	*	*	*	*	*	*	*	-
	ABC	36,7	40,5	4,3	7,3	1,4	10,3	-	-
	El Mund.	37,3	41,9	4,5	6,5	1,7	5,9	-	-

ANY	Sondej/ Result.	PSOE	PP	CIU	IU	PNV	ALTRES	CDS	ERC
2004	Reals	42,59	37,71	3,23	4,96	1,63	3,45	-	2,52
	Perfectes	41,67	36,84	2,87	4,66	1,47	3,16	-	2,16
		43,51	38,56	3,59	5,54	1,78	3,71	-	2,78
	La Vang.	39,2	41,4	3,0	6,3	2,5	5,9	-	2,1
	El Periód.	37,3	42,5	3,2	7,1	1,5	6,5	-	1,9
	El País	38,0	42,0	3,0	6,3	1,7	7,1	-	1,9
	ABC	37,2	42,2	3,1	7,0	1,5	6,9	-	2,1
El Mund.	37,6	42,1	3,4	5,3	1,8	7,8	-	2,0	
2008	Reals	43,87	39,94	3,03	3,77	1,19	2,94	-	1,16
	Perfectes	42,94	39,03	2,69	3,24	1,05	2,6	-	0,92
		44,82	40,84	3,38	3,99	1,14	3,18	-	1,36
	La Vang.	43,5	39,5	2,5	4,4	1,3	7,0	-	1,8
	El Periód.	*	*	*	*	*	*	-	*
	El País	42,9	38,8	2,6	4,8	1,5	6,0	-	1,9
	ABC	42,0	40,0	**	4,1	**	14,9	-	**
El Mund.	43,4	39,3	3,2	4,4	1,5	13,3	-	1,9	

taula 2.4.2.2 Escons dels diferents partits en totes les convocatòries referents a les eleccions al Congrés dels diputats.

ANY	Sondej/ Result.	PSOE	PP	CIU	IU	PNV	ALTRES	CDS	ERC
1986	Reals	184	105	18	7	6	11	19	-
	Perfectes	180-191	101-112	16-20	4-9	5-7	6-11	16-23	-
		La Vang.	-	-	-	-	-	-	-
	El Periód.	180	98	17	5	7	13	30	-
	El País	-	-	-	-	-	-	-	-
	ABC	174	93	18	9	10	37	9	-
	El Mund.	-	-	-	-	-	-	-	-
1989	Reals	175	107	18	17	5	14	14	-
	Perfectes	165-177	102-113	18-21	16-22	4-6	10-15	12-18	-
		La Vang.	172-178	104-107	17-18	13	7-8	12-19	14-15
	El Periód.	-	-	-	-	-	-	-	-
	El País	176-182	95-104	18	15-16	5-7	12-16	16-19	-
	ABC	-	-	-	-	-	-	-	-
	El Mund.	168	103	17	16	5	18	23	-
1993	Reals	159	141	17	18	5	10	-	-
	Perfectes	149-161	136-147	16-19	15-21	4-7	10-16	-	-
		La Vang.	142-146	144-148	18-19	24-25	6-7	9-11	-
	El Periód.	133-152	134-151	19-21	19-24	6-7	11-17	-	-
	El País	135-151	141-155	18-20	20-22	6	13-17	-	-
	ABC	134-144	144-153	16-17	26-28	6-7	11-14	-	-
	El Mund.	132-141	147-155	18-19	25-27	5-6	12-14	-	-



RECOLLIDA DE LES DADES

ANY	Sondej/ Result.	PSOE	PP	CIU	IU	PNV	ALTRES	CDS	ERC
1996	Reals	141	156	16	21	5	11	-	-
	Perfectes	136-148	151-162	14-17	17-23	4-6	9-12	-	-
	La Vang.	135-145	160-170	14-15	19-21	5-6	7-8	-	-
	El Periód.	119-129	165-175	14-16	18-27	-	12-15	-	-
	El País	118-128	170-178	13-15	24-27	-	13-15	-	-
	ABC	117-125	176-184	13-14	22-28	-	10-13	-	-
	El Mund.	113-123	170-179	14-15	25-29	-	14-16	-	-
2000	Reals	125	183	15	8	7	12	-	-
	Perfectes	123-133	175-185	14-18	5-10	6-7	9-15	-	-
	La Vang.	136-140	166-170	14-15	12-14	6	9-12	-	-
	El Periód.	135-140	165-170	13-14	9-12	7-8	13-14	-	-
	El País	131-139	165-171	14-15	9-13	6-7	12-14	-	-
	ABC	138-144	158-164	15-16	9-10	6-7	8-12	-	-
	El Mund.	137-143	164-170	16	9-11	6-7	9-12	-	-
2004	Reals	164	148	10	5	7	8	-	8
	Perfectes	157-168	143-154	9-12	4-8	6-9	6-9	-	7-9
	La Vang.	143-147	162-167	10	9-10	7	8-9	-	5-6
	El Periód.	135-140	169-173	10-11	10-12	7	10-14	-	6-7
	El País	134-141	168-172	10-11	11	7-8	8-9	-	7
	ABC	133-137	174-177	10-11	8-10	7	7	-	6-7
	El Mund.	138-144	168-173	11	6-8	7	7-9	-	5-6
2008	Reals	169	154	10	2	6	6	-	3
	Perfectes	164-175	148-158	9-13	2-4	5-7	4-7	-	1-4
	La Vang.	162-167	152-156	8-9	4-5	6-7	5-7	-	6
	El Periód.	156-160	156-160	8-9	4-6	6-7	9-12	-	5-6
	El País	165-169	148-154	9	4	7	6-7	-	5-6
	ABC	-	-	-	-	-	-	-	*
	El Mund.	157-171	148-161	9-11	4	7	3-8	-	5-6

### 3 ASSIGNACIÓ D'ESCONS EN FUNCIÓ DEL PERCENTATGE DE VOTS

---

Cada una de les circumscripcions té assignats uns diputats per repartir. En cada una d'elles hi ha diferents candidatures que es presenten a les eleccions i que el ciutadà les pot votar el dia de la celebració. Un cop sabuts els vots que ha aconseguit cada partit, s'han de distribuir els diputats. La forma que es fa és senzilla d'entendre i s'analitza en aquest apartat.

#### 3.1 Percentatge llindar

Un partit només opta a l'assignació d'escons si supera un  $s$  % del total dels vots (contant tots els vots a candidatures i els vots en blancs). Aquest és el percentatge llindar, i en les eleccions que ens ocupen, està fixat al 3%.

En conseqüència, a un territori que distribueix uns certs escons, els partits que no sobrepassin aquest límit obtindran zero escons ja d'inici.

Aquesta norma doncs, afavoreix als partits amb més votants. Si es fes la repartició dels diputats sense tenir en compte la frontera, partits més minoritaris tindrien opcions a aconseguir algun escó en detriment dels partits més forts.

#### 3.2 La llei de Hondt

La llei de Hondt és la fórmula que ha adoptat la legislació electoral espanyola pel repartiment d'escons. Cada un dels  $K$  partits que han fet candidatura obtenen un nombre de vots  $(v_1, v_2, \dots, v_k)$ . Hi ha  $N$  escons a repartir i es formen els anomenats "quocients de Hondt":

$$v_i, v_i/2, \dots, v_i/N,$$

on "i" distingeix cada una de les candidatures que ha obtingut vots.

Per tant, s'obté una taula de vots  $K \times N$  on hi ha tots els quocients de totes les candidatures (mirar taula 3.2.1). Cada un dels  $N$  quocients més grans se'ls hi atribueix un escó.

La peculiaritat d'aquesta llei és que té biaix. És a dir, si es fessin múltiples mostres aleatòries tenint en compte les proporcions reals, el promig de totes les prediccions (la

mitjana dels escons previstos de cada partit) no coincidiria amb el resultat final del dia de la celebració.

**taula 3.2.1** Taula de quocients de Hondt dels partits que han aconseguit vots.

Quocients	partit 1	partit 2	partit 3	...	partit k
$q_1$	$v_1$	$v_2$	$v_3$	...	$v_k$
$q_2$	$v_1/2$	$v_2/2$	$v_3/2$	...	$v_k/2$
$q_3$	$v_1/3$	$v_2/3$	$v_3/3$	...	$v_k/3$
...	...	...	...	...	...
$q_N$	$v_1/N$	$v_2/N$	$v_3/N$	...	$v_k/N$

Anem a veure un parell d'exemples que il·lustrin el funcionament de la llei i els problemes que presenta:

El primer data de l'any 2004. Es tracta de la repartició d'escons en les eleccions al Congrés dels Diputats a la província de La Rioja. Es repartien 4 diputats. El Partido Popular va aconseguir 2 dels diputats, amb un 49,94% dels vots, i el Partido Socialista els 2 restants, amb un 43,97% dels vots. La tercera força va ser Izquierda Unida, que tot i no aconseguir cap escó va acumular el 2,76% dels vots a la província. Al ser aquest percentatge inferior al 3% dels vots totals, Izquierda Unida no va poder optar a cap diputat.

Per tant, la taula de quocients dels tres partits amb més suport de la població és la següent:

**taula 3.2.2** Taula de quocients de Hondt de les eleccions al Congrés dels Diputats l'any 2004 en la província de La Rioja

quocients	PP	PSOE	IU
$q_1$	$v_1 = 92.441$	$v_2 = 81.390$	$v_3 = 5.115$
$q_2$	$v_1/2 = 46.221$	$v_2/2 = 40.695$	$v_3/2 = 2.557,5$
$q_3$	$v_1/3 = 30.814$	$v_2/3 = 27.130$	$v_3/3 = 1.705$
$q_4$	$v_1/N = 23.110$	$v_2/N = 20.348$	$v_3/N = 1.278,7$

Els quocients marcats són els més grans. Per tant, s'assignaran 2 escons a cada partit. El límit inferior per aconseguir un escó el marca el quocient més baix que n'ha aconseguit un. En aquest cas són 40.695 vots.

En aquesta província, en l'any 2004, si es realitzen moltes rèpliques aleatòries de les proporcions reals no surt el biaix que comentàvem abans. Per això anem a veure un altre exemple on s'hi produeix aquest fet.

S'agafen les mateixes eleccions de l'any 2004. A la província de Lleida. També es repartien 4 diputats. El Partido Socialista en va aconseguir 2 amb un 29,57% dels vots. La segona força va ser Convergència i Unió amb 1 diputat i el 29,46% dels vots. El tercer partit va ser ERC amb 1 diputat i el 21,48% dels vots. El Partido Popular i els Verds també van tenir forces votants (14,62% i 2,96% dels vots respectivament) però no van aconseguir cap escó.

Ara la taula de quocients és la següent:

**taula 3.2.3** *Taula de quocients de Hondt de les eleccions al Congrés dels Diputats l'any 2004 en la província de Lleida*

Quocients	PSOE	CIU	ERC	PP
$q_1$	$V_1 = 68.971$	$V_2 = 68.735$	$V_3 = 50.104$	$V_4 = 34.116$
$q_2$	$V_1/2 = 34.486$	$V_2/2 = 34.368$	$V_3/2 = 25.052$	$V_4/2 = 17.058$
$q_3$	$V_1/3 = 22.991$	$V_2/3 = 22.912$	$V_3/3 = 16.701$	$V_4/3 = 11.372$
$q_4$	$V_1/4 = 17.243$	$V_2/4 = 17.184$	$V_3/4 = 12.526$	$V_4/4 = 8.529$

Per molt poc, el PSOE s'endu l'últim diputat en joc. Si es fan múltiples mostres (s'ha fet amb 10.000) tenint en compte els paràmetres (o proporcions) que ha obtingut cada partit, es pot comprovar que les mitjanes d'escons que presenten cada partit no es corresponen als escons reals.

**taula 3.2.4** *Mitjanes dels escons de les 10.000 rèpliques del sondeig teòric al Congrés dels Diputats l'any 2004 en la província de Lleida.*

Partit	Mitjana
PSOE	1,35
CIU	1,336
ERC	1
PP	0,315
IV	0

Com es pot apreciar en la taula 3.2.4, al fer múltiples rèpliques, i ja que les dues proporcions de vots del PSC i el CIU són tant semblants, una bona part d'elles és CIU, i també el PP, qui aconsegueix l'últim escó que hi ha en joc. Aquest fet fa fer esbiaixades les prediccions. En definitiva, és molt fàcil que en comptes de predir el darrer escó pel PSC sigui CIU qui el guanyi. O fins i tot, el Partit Popular. És més, és més probable que s'assigni malament el segon diputat en un d'aquests dos partits que no pas correctament. Aquest problema passa a moltes de les diputacions que estudiem i farà que el total d'escons que es reparteixen en tot el territori no coincideixi amb les mitjanes dels sondejos perfectes.

No tenim constància que els sondejos tinguin en compte aquesta circumstància. Per predir de forma adequada els punts per cada partit i corregir aquest biaix, una opció seria aplicant tècniques de simulació. Fixar les proporcions de vot dels sondejos com els veritables paràmetres i fer múltiples replicues del sondeig. Aconseguir les mitjanes de diputats de cada partit i restar la diferència d'aquestes, respecte els diputats pronosticats inicialment pels sondejos amb aquests mateixos pronòstics.

$$Sf_j = Si_j - (Si_j - \overline{M}_j)$$

*On  $Si_j$  són els escons del sondeig inicial,  $\overline{M}_j$  és la mitjana dels sondejos simulats i  $Sf_j$  el sondeig després de la modificació. El subíndex  $j$  diferencia cada un dels partits.*

## 4 SONDEJOS TEÒRICS

---

Com ha de ser un sondeig per considerar-lo perfecte? Hi ha diferents aspectes que s'han de tenir en compte.

Un primer pas podria ser determinar quant individus entren en la mostra. És obvi que quants més enquestats hi hagin, millor es podran estimar les proporcions de vot dels partits. Tenint present això s'aconseguiria un sondeig ideal quan es preguntés la intenció de vot a tots els individus que componen el cens electoral. El que passa és que els recursos econòmics no són il·limitats. Aquest és el motiu de prendre una mostra de mida similar a les que utilitzen els sondejors. S'ha decidit que aquesta fos de 3.000 a les eleccions al Parlament de Catalunya i 20.000 a les Eleccions Generals. És una fita sempre superior a les que presenten els diaris (que es mouen en mides de entre 1.000 i 2.000 individus a Catalunya i entre 3.000 i 10.000 a Espanya) per establir aquesta idea de perfecció. No és més gran perquè tant el sondeig ideal com els dels diaris siguin comparables.

Un altre aspecte a considerar és que en un sondeig perfecte els individus preguntats diuen sempre la veritat. És a dir, la intenció de vot que proposen en el sondeig és realment el que votaran. D'aquesta manera s'evita errar en tots aquells casos que per múltiples raons no voten el mateix el dia de les eleccions del que tenien previst, o almenys el que havien manifestat, una setmana abans. Aquests canvis d'opinió poden fer fallar als sondejors dels diaris. En un de perfecte no. Un cas que podria il·lustrar aquesta diferència, per molt exagerat que sigui, podria ser un votant català que pensés anar a votar a ERC el dia de les eleccions. Un dia abans del dia clau es trenca les dues cames fent un experiment. Donada la seva situació aquesta persona decideix no participar en les votacions. Mentre que el sondeig del diari inclourà el vot per ERC, el sondeig perfecte no el tindrà en compte. També és coneguda la tendència a mentir sobre la predilecció d'algun partit polític. Sobretot es dona amb el Partit Popular a Catalunya. Els votants del partit de dretes solen tenir una certa prudència en manifestar el seu vot als populars. Per pal·liar aquesta situació és sabut que molts dels sondejors dels diaris corregeixen a l'alça els escons pronosticats al PP.

Una última consideració. El sondeig teòric és una mostra totalment aleatòria. És a dir, tots els individus que poden votar tenen les mateixes probabilitats d'entrar a la mostra. En els sondejors dels diaris això és complicat que es produeixi. Ho és perquè trobar cada una de les persones, que poden venir de qualsevol municipi, per petit que sigui, que el componen per atzar és una feina extremadament costosa. No ho especifiquen molt bé com distribueixen la mostra aquests, però segur que l'assignació de les enquestes no és totalment aleatòria.

En definitiva, es considera el sondeig perfecte com una mostra aleatòria de 3.000 o 20.000 individus (depenen del tipus d'elecció) on tot és ideal. De fet el que s'està fent és aconseguir exactament el que varen votar el dia de les eleccions unes certes persones elegides a l'atzar. Es fa com una radiografia de la realitat.

Si en comptes de fer una sola mostra amb aquesta metodologia se'n fessin milers d'elles, es tindria constància del rang de diputats que podrien sortir de fer un sondeig el màxim de ben fet possible. S'ha decidit fer-ne 10.000. És una xifra àmplia i suficient per guanyar a l'atzar.

D'aquesta manera es podran comparar els sondejos amb aquestes mostres i es veurà si la distància amb el resultat real és pel fet d'estar fent inferència de mostres de mida petita, i que per tant és l'atzar qui provoca l'error de predicció. O si pel contrari el sondeig s'allunya significativament de les distàncies que presenten les mostres dels sondejos perfectes i es podria intuir que l'atzar no es suficient com per considerar el sondeig vàlid.

Un cop feta la reflexió sobre aquests temes, es pot començar a proposar com es pot aconseguir aquest sondeig ideal.

#### **4.1 Metodologia teòrica per crear les mostres: ús de la distribució multinomial**

La llei multinomial és una distribució estadística que es basa en la distribució binomial. De fet, és una generalització d'aquesta.

La binomial compta el número d'èxits que hi ha hagut en  $N$  proves. Per tant, hi ha 2 possibles respostes en una prova: èxit o fracàs. La suma d'èxits i fracassos dona el número total d'experiments. La probabilitat de les dues opcions és coneguda. La diferència amb la multinomial recau en que aquesta no només té dos possibles resultats sinó que té  $k$  possibles resultats, on  $k$  és un nombre natural establert. D'aquesta forma cada una de les proves pot tenir múltiples respostes i cada una té fixada una probabilitat. La suma d'esdeveniments dels  $k$  possibles resultats donen el número total d'experiments. A més, la suma de probabilitats de sortir cada un dels " $k$ " resultats és 1.

Es pot suposar que cada esdeveniment és un vot d'un ciutadà. Que els  $k$  possibles resultats són els  $k$  partits que es presenten en les eleccions (també s'inclouria l'abstenció). I per últim, les  $k$  probabilitats són les proporcions de vots reals que ha obtingut cada un d'ells el dia de les eleccions (mirar taula 4.1.1) .

Si es simulen moltes proves (de similar mida que els sondejos) segons la distribució multinomial amb l'elecció dels paràmetres que s'ha comentat, s'obté una petita mostra

aleatòria. Aquesta s'hagués pogut obtenir en fer una enquesta a tantes persones com proves es fan elegides a l'atzar de tota la població del territori que s'està pronosticant.

**taula 4.1.1** *Matriu on si destaquen el número de vots per atzar de cada una de les candidatures presentades tenint en compte unes probabilitats a priori.*

Candidatures	Número de vots en n individus	Probabilitat (proporcions)
Abstenció	$v_a$	$p_a$
partit 1	$v_1$	$p_1$
partit 2	$v_2$	$p_2$
partit 3	$v_3$	$p_3$
...	...	...
partit k	$v_k$	$p_k$
<b>Total</b>	$v_a + \sum_{i=1}^k v_i = n$	$p_a + \sum_{i=1}^k p_i = 1$

## 4.2 Tria de la mida mostral

Els sondejos electorals que publiquen els diaris, en gran majoria, sols indiquen quina és la mida de la mostra total per establir els pronòstics. No esmenten com han organitzat totes aquestes enquestes en els diferents territoris. Per això, i en l'afany de trobar la metodologia que faci el màxim de perfecte possible el sondeig es valora com s'han d'organitzar les mides mostrals segons les diferents circumscripcions.

Cal recordar que s'enquesten en cada una de les 10.000 mostres a 3.000 i 20.000 individus (al Parlament i al Congrés respectivament).

### 4.2.1 Mides mostrals al Parlament de Catalunya

#### a) Una primera proposta:

Partint del cas més extrem que es podria donar al cas en la distribució binomial. Es tracta de tenir tant la probabilitat d'èxit (p) com de fracàs (q) iguals. Per tant, valdran 0,5. És obvi que és l'esdeveniment més complicat de predir a través d'una mostra ja que és el que té més variabilitat. La variància d'una binomial és :

$$np(1-p) = npq$$



El màxim d'aquesta s'aconsegueix quan aquestes dues probabilitats valen 0,5. Mirar la figura 4.2.1 .

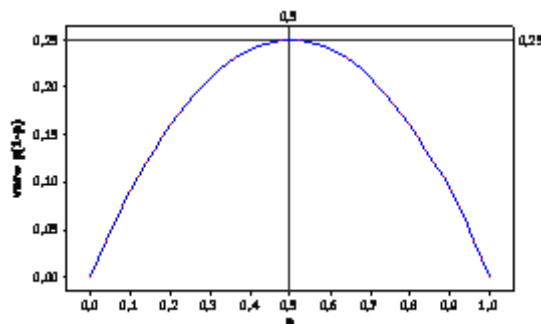


figura 4.2.1 Variància en funció de la probabilitat  $p$  i  $q$

Al tractar-se, en el cas que ens ocupa, de la distribució multinomial i no la binomial, es fixen a 0 totes les proporcions excepte la dels dos primers partits (que valdran 0,5) :

$$( p_1 = 0,5; p_2 = 0,5; p_k = 0, k = 3, \dots, m)$$

Mitjançant l'aproximació a la distribució normal es poden trobar quants individus s'haurien d'agafar en la mostra. Té en compte el total de la població censada (o amb dret a vot) i el marge d'error que es vol.

Atès que es vol utilitzar un total de 3.000 proves, aquesta fita es fixa aproximadament de manera que proporcioni una referència. La confiança escollida és del 95 %.

La fórmula per trobar la mida mostral és la següent :

$$n = \frac{Z^2_{\alpha/2} P(1-P)N}{d^2 N + Z^2_{\alpha/2} P(1-P)} + 0,5$$

On "d" és el marge d'error i "N" la població censada

1. Es fixa el marge d'error en  $\pm 0,0358$ .
2.  $Z_{\alpha/2} = 1,959963985$ .

La població censada a Barcelona l'any 2006 era de 3.991.904 persones.

**Mida estimada = 750**

La població censada a Girona l'any 2006 era de 483.543 persones.

**Mida estimada = 749**

La població censada a Lleida l'any 2006 era de 311.037 persones.

**Mida estimada = 748**

La població censada a Tarragona l'any 2006 era de 534.790 persones.

**Mida estimada = 749**

Ja s'observa que canvia molt poc la mida en les diferents províncies. D'aquí la importància de desenvolupar una bona mostra representativa. Per no haver de fer aquests càlculs a totes les eleccions i a la vista de les poques diferències de mida de les diferents províncies que s'aprecien, es podria fixar a les 4 les mateixes simulacions (o enquestes): 750.

b) La proposta final:

Tot aquest raonament seria el més correcte si totes les circumscripcions repartissin el mateix nombre d'escons. Això no és així. Amb els vots de tots els individus empadronats a algun dels municipis de Barcelona s'assignen 85 diputats. En el cas dels vots dels Gironins en són 17 diputats. A Lleida se'n adjudiquen 15. Per últim, a Tarragona en són 18.

Amb la primera metodologia proposada, tot i tenir els mateixos rangs de proporcions de vot en les diferents províncies, el rang d'escons que poden tenir els partits a Barcelona és molt més gran que a Girona, Lleida i Tarragona.

Per demostrar-ho s'ha creat un petit programa amb R que comparés aquests rangs en diferents combinacions de mides. Una és fixar 750 individus a cada territori. Aquesta és

la que es pretén refutar. Les següents ja són augmentant la mida mostral a la província de Barcelona en detriment de les mides de les altres 3 circumscripcions.

Per quantificar quina de les múltiples situacions s'ajusta millor, ens inventem una distància. El que s'ha de fer primer és ordenar, de menys a més escons, els 10.000 sondejos perfectes (s'explica en detall la seva obtenció en el tema 5). Això es fa per cada partit. S'aconsegueix un interval d'escons per cada partit al 95% si es pren com a nivell baix la mostra 250 (de la ordenada). El nivell alt el conté la mostra 9.750. La distància que s'agafa és la suma de totes les diferències entre nivell alt i baix. Com menors siguin més petits seran els intervals de confiança i millor s'estimaran els escons per cada partit. Per tant es tindran tants sumands com partits polítics han obtingut representació.

**taula 4.2.1** *Sondejos ideals ordenats de menys a més escons en cada un dels partits*

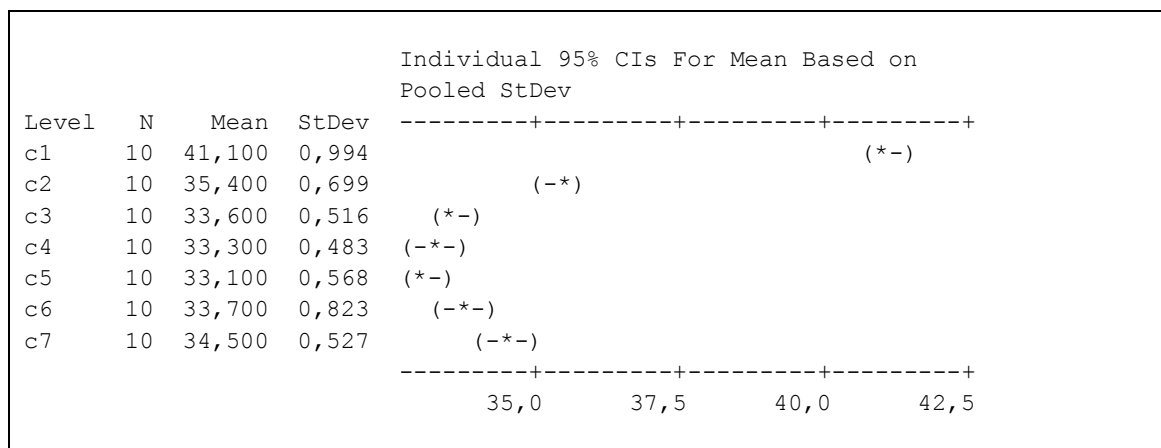
Nº sondeig	Partit 1	Partit 2	...	Partit n
1	$E_{1,1}$	$E_{1,2}$	...	$E_{1,n}$
...	...	...	...	...
250	$E_{250,1}$	$E_{250,2}$	...	$E_{250,n}$
...	...	...	...	...
9.750	$E_{9.750,1}$	$E_{9.750,2}$	...	$E_{9.750,n}$
...	...	...	...	...
10.000	$E_{10.000,1}$	$E_{10.000,2}$	...	$E_{10.000,n}$

$$\text{Dist.} = \sum_{i=1}^n (E_{9750i} - E_{250i})$$

En el cas de les eleccions del Parlament de Catalunya de l'any 2006 hi ha 6 partits que han aconseguit representació a la Generalitat. Per tant es tindran 6 sumands. Es busca la combinació que menys rang de diputats dona. Els resultats obtinguts es troben en la següent taula :

**taula 4.2.2** *Distàncies obtingudes en les múltiples combinacions proposades.*

Combinacions	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Mean
{750,750,750,750}	41	41	41	42	42	42	39	42	41	40	41,1
{1200,600,600,600}	35	35	36	36	34	36	35	35	36	36	35,4
{1500,500,500,500}	33	33	34	34	34	34	34	34	33	33	33,6
{1650,450,450,450}	33	33	33	34	33	33	34	33	33	34	33,3
{1800,400,400,400}	33	33	33	34	33	33	34	32	33	33	33,1
{1950,350,350,350}	34	34	33	33	35	35	33	33	33	34	33,7
{2100,300,300,300}	35	34	34	35	35	34	35	34	35	34	34,5



**figura 4.2.2** Sortida parcial de Minitab de l'ANOVA amb factors les diferents combinacions i resposta la distància.

Clarament la combinació de 750 individus (c1) a cada província és la pitjor. Entre les altres hi ha poques diferències. La combinació que té la mitjana més baixa és la 5 (33,1), amb distribució: {1800,400,400,400}. Es determina aquesta com a l'assignació d'enquestes perfectes segons els territoris. No hi ha diferències significatives en prendre aquesta o la combinació 4 {1650,450,450,450 }, però si és cert que queden valor més rodons i còmodes de tractar. El programa que hem realitzat es pot veure en l'annex 2 (sobretot fixar-se amb el text en cursiva).

#### 4.2.2 Mides mostrals al Congrés dels Diputats

En les eleccions generals no es pot aplicar la mateixa metodologia que en el Parlament per detectar quina combinació de mides s'ha d'assignar a cada circumscripció. Es tenen masses diputacions i amb elles infinitat de combinacions que no es podrien representar amb un programa com el que hem utilitzat anteriorment.

Per aquest motiu s'ha decidit recercar com ho plantegen els sondejos més prestigiosos. Aquest és el que declara el CIS (Centro d'Investigación Sociològica) en cada convocatòria. És sense cap mena de dubte l'enquesta que disposa de més recursos. Molts més que els sondejos dels diaris que analitzem.

En la pàgina web del CIS es poden obtenir els procediments que han desenvolupat en la realització del mostreig. Ens fixem amb el de l'any 2004. En aquest estudi varen enquestar a 24.140 individus repartits en tot l'estat. Sense entrar en detall en les tècniques de mostreig que varen emprar, en la taula 4.2.2.1 s'hi destaquen les proporcions d'enquestats que varen prendre en cada circumscripció, així com les mides que s'utilitzaran. També es remarquen els escons que assigna cada província (serà útil en el següent tema)

**taula 4.2.2.1** *taula de proporcions, mides i escons que utilitzarem per aconseguir els sondejos teòrics al Congrés dels Diputats*

Circumscripció	proporcions	mides	Escons
A Corunya	2,4	480	9
Alava	2,2	440	4
Albacete	1,53	306	4
Alacant	2,17	434	11
Almería	1,63	326	5
Astúries	2,4	480	8
Avila	0,99	198	3
Badajoz	1,68	336	6
Illes Balears	1,96	392	8
Barcelona	4,79	958	31
Burgos	1,45	290	4
Càceres	1,6	320	4
Cádiz	2,54	508	9
Castelló	1,66	332	5
Ceuta	0,94	188	1
Ciudad Real	1,61	322	5
Córdoba	2,03	406	7
Cuenca	1,28	256	3
Girona	1,72	344	6
Granada	2,07	414	7
Guadalajara	1,33	266	3
Guipúscoa	2,88	576	6
Huelva	1,62	324	5
Huesca	1,28	256	3
Jaen	1,86	372	6
La Rioja	1,49	298	4
Las Palmas	2,18	436	8
León	1,71	342	5
Lleida	1,63	326	4
Lugo	1,61	322	4
Madrid	4,87	974	35
Málaga	2,58	516	10
Melilla	0,94	188	1
Murcia	2,18	436	9
Navarra	2,1	420	5
Ourense	1,57	314	4
Palència	1,08	216	3
Pontevedra	2,21	442	7
Salamanca	1,49	298	4
Santa Cruz de Tenerife	2,1	420	7

Circumscripció	proporcions	mides	Escons
Cantabria	1,93	386	5
Segovia	1,04	208	3
Sevilla	2,87	574	12
Soria	1,04	208	3
Tarragona	2,08	416	6
Teruel	1,57	314	3
Toledo	1,64	328	5
Valencia	3,14	628	16
Valladolid	1,62	324	5
Vizcaya	2,36	472	9
Zamora	1,32	264	3
Saragossa	2,03	406	7

### 4.3 Sondejos teòrics amb una correcció en les proporcions de vot

Un dels objectius del treball és determinar quan un sondeig està realment ben fet i quan no es pot considerar correcte. Per determinar en quin dels dos casos estem davant és molt exigent únicament prestar atenció als sondejos perfectes.

Considerar que en una setmana vista, la intenció de vot del ciutadà canvia d' un partit a un altre, amb un tant per cent petit és raonable, ja sigui perquè l'individu enquestat no diu la veritat o perquè ha canviat realment el seu pensament. Contra això, els sondejos dels diaris no hi poden fer res. S'ha aplicat als sondejos teòrics una correcció a la proporció de vot del partit principal. No és fàcil assignar una fita concreta a aquest canvi. Suposem que és del 3% dels vots. Aquest percentatge el guanya el partit i repercuteix negativament a la segona candidatura amb més representació (que perd un 3%).

D'aquesta manera les distàncies entre els teòrics i els reals seran més grans. Aleshores, els sondejos que es situïn a prop, encara que a més distància respecte als reals del interval que presenten els perfectes, es podrien considerar vàlids (tenint en compte les dificultats de realitzar tals prediccions). Aquests serien uns sondejos ideals amb un canvi en la intenció de vot entre els dos primers partits del 3%.

Per adjudicar en quin percentatge de canvi en la proporció de vot es troben els sondejos que superen la fita màxima del interval dels teòrics amb un retoc del 3%, s'estudia també, quines distàncies presentaran els sondejos perfectes si el canvi és del 6%.

## 5 OBTENCIÓ DE LES MOSTRES PELS SONDEJOS TEÒRICS

---

S'ha comentat com és fa per aconseguir els escons a partir de les proporcions de vot. També com ha de ser un sondeig perfecte i la base teòrica de la seva creació mitjançant l'ús de la distribució multinomial. En aquest tema s'explica com s'han aplicat aquests coneixements per obtenir les mostres.

S'han realitzat uns programes amb el software R que engloben tota la metodologia de creació dels sondejos perfectes. De fet consten tants programes com eleccions s'estudien.

### 5.1 Constants emprades : mides mostrals i proporcions de vots

El que es constaten primer són una sèrie de constants. Unes fan referència a mides mostrals. S'ha de fixar quants individus formen la mostra (en són 3.000 al Parlament de Catalunya i 20.000 al Congrés dels Diputats) i com es distribueixen segons la circumscripció. Es necessita una constant per cada mida mostral de la província (a Catalunya, 1.800 individus a Barcelona i 400 a totes les altres localitats). També cal determinar quantes rèpliques es fan del sondeig. En són 10.000 en les dues eleccions i és una altra constant.

En un altre bloc hi ha les que fan referència a les proporcions de vot. Aquestes són el motiu de fer un programa per elecció. De fet és l'únic que varia en cada Script. És obvi ja que en cada un dels anys de celebració hi ha hagut proporcions de vot diferents en els múltiples partits. Hi ha tantes constants per cada partit com circumscripcions se'n deriven (4 a Catalunya i 52 a Espanya). A més a més dels vots als partits, també s'ha de considerar l'abstenció a cada un d'aquests territoris.

En el Parlament, hi ha 5 partits que es repeteixen a totes les eleccions. Són CIU, PSC, PP, ERC i IC-IV (o PSUC) ja que varen aconseguir escons a totes les convocatòries. Hi ha un sisè que es marca com a *Altres* i que es destaca només quan aconsegueix algun escó un partit diferent dels esmentats anteriorment. Aquests són els partits CDS i CPC, que van aconseguir escons en 1988 i 2006 respectivament. L'últim que es diferencia és per precisar la resta de vots que no han aconseguit els 5 o 6 partits principals. Siguin vots per altres candidatures o vots en blanc.

En el Congrés, hi ha 6 partits fixes en totes les convocatòries. El PSOE, el PP, CIU, IU i el PNV. El sisè el formen candidatures més minoritàries. Aquest s'identifica com a *Altres* i estableixen la suma de diputats aconseguits entre tots aquests partits. Igual que passava a Catalunya hi ha dos partits més que han tingut una importància remarcable en

el Congrés durant un període curt. No constaran a totes les eleccions, però si que s'hauran de tenir presents en les que han destacat. Estem parlant del CDS i ERC

## 5.2 Obtenció de les proporcions de vot

Un cop establertes les constants, es creen una sèrie de matrius. Dues per cada província. En la primera hi hauran les proporcions de vot de cada partit (i l'abstenció) en aquell territori. La matriu és de  $N \times M$  on  $N$  són les rèpliques que fem (10.000) i  $M$  el total de partits que varen aconseguir escons a les eleccions juntament amb les dues columnes que determinen els vots a altres candidatures i ciutadans que s'han abstingut. La segona és quasi una còpia de la primera. Varia pel fet que només té en compte les proporcions que superen el llindar del 3% de vots. Els que no ho aconseguen valen 0. A més, ja no conté aquestes dues últimes columnes.

**taula 5. 1** *Proporcions de vot dels partits amb representació al Parlament en els sondejos perfectes.*

Most.	Barna.	Girona	Lleida	Tar.
$m_1$	$P_{1,B,p_1}, \dots, P_{1,B,p_n}$	$P_{1,G,p_1}, \dots, P_{1,G,p_n}$	$P_{1,L,p_1}, \dots, P_{1,L,p_n}$	$P_{1,T,p_1}, \dots, P_{1,T,p_n}$
$m_2$	$P_{2,B,p_1}, \dots, P_{2,B,p_n}$	$P_{2,G,p_1}, \dots, P_{2,G,p_n}$	$P_{2,L,p_1}, \dots, P_{2,L,p_n}$	$P_{2,T,p_1}, \dots, P_{2,T,p_n}$
...	...	...	...	...
$m_{10.000}$	$P_{N,B,p_1}, \dots, P_{N,B,p_n}$	$P_{N,G,p_1}, \dots, P_{N,G,p_n}$	$P_{N,L,p_1}, \dots, P_{N,L,p_n}$	$P_{N,T,p_1}, \dots, P_{N,T,p_n}$

S'aconsegueixen totes aquestes proporcions aplicant la creació de mostres en base la distribució multinomial. Es repeteix el procés 10.000 vegades. D'aquesta forma s'obtenen múltiples possibilitats d'enquestes perfectes. Ara només queda programar el pas de proporcions a escons segons la llei de Hondt.

## 5.3 Metodologia aplicada per l'obtenció d'escons

Barcelona, Girona, Lleida i Tarragona són les quatre províncies de Catalunya. Com ja s'havia comentat en el tema 4, amb els vots de tots els individus empadronats a algun dels municipis de Barcelona s'assignen 85 diputats. En el cas dels vots dels Gironins en són 17 diputats. A Lleida se'n adjudiquen 15. Per últim, a Tarragona en són 18. Espanya, a cada una de les circumscripcions, també té establerts uns certs diputats a repartir. Es troben en l'apartat 4.2.2 (taula 4.2.2.1).



Per passar les proporcions a escons doncs, també s'han de diferenciar els diputats que s'atribueixen a cada circumscripció. Es crea una matriu per cada una d'elles  $N \times (M-2)$  on hi consten els escons que s'adjudicarien a cada partit. Al haver-hi 10.000 mostres constaran 10.000 possibles combinacions de repartició d'escons. Amb unes instruccions amb R fàcils d'implementar, que es basen en el que s'ha estudiat en l'apartat 3.2 ( la Llei de Hondt), s'aconsegueix passar a diputats tots els percentatges de vot.

Ara ja només queda sumar els escons ( $e_{ij}$ ) de totes les províncies. Això es fa per cada mostra. En el cas de Catalunya es realitzaria de la forma que es veu en la taula 5.2.

**taula 5.2** Taula d'escons obtinguts en les diferents circumscripcions en les eleccions al Parlament i el total en les 10.000 mostres.

Mostres	Barna.	Girona	Lleida	Tar.	escons totals
$m_1$	$e_{11}$	$e_{12}$	$e_{13}$	$e_{14}$	$\sum_{j=1}^4 e_{1j} = e_{1t}$
$m_2$	$e_{21}$	$e_{22}$	$e_{23}$	$e_{24}$	$\sum_{j=1}^4 e_{2j} = e_{2t}$
...					
$m_{10000}$	$e_{100001}$	$e_{100002}$	$e_{100003}$	$e_{100004}$	$\sum_{j=1}^4 e_{10000j} = e_{10000t}$

Un cop sumats en totes les mostres ja es tenen milers de sondejos que estimen els escons per cada partit i serà possible fer un interval de confiança del nombre de diputats que pot obtenir cada partit amb un nivell de risc determinat. S'utilitzarà, com sol ser habitual, el risc del 5%.

El programa en R comentat el podeu veure en l'annex 2.

## 6 DEFINICIÓ DE LES DISTÀNCIES

---

En aquest apartat es pretén quantificar de forma numèrica les diferències entre els sondejos dels diaris, els sondejos perfectes i els resultats reals. Seria molt fàcil fer els càlculs si els pronòstics dels sondejos fossin puntuals. Només caldria preocupar-se d'establir una distància concreta. No és així, són intervals. Es discutirà com podria fer-se per aconseguir una distància justa i interpretable.

### 6.1 Selecció de les distàncies

Es poden pensar en diferent vies per expressar els errors amb un sol valor. Aquestes són totes les proposes :

1. La suma de residus al quadrat:

$$D1 = \sum_{i=1}^k (R_i - S_i)^2$$

2. La suma de residus amb valor absolut:

$$D2 = \sum_{i=1}^k |(R_i - S_i)|$$

3. La suma de residus al quadrat dividit pels escons reals :

$$D3 = \sum_{i=1}^k \frac{(R_i - S_i)^2}{R_i}$$

4. La suma de residus al quadrat dividit per les proporcions de vot reals:

$$D4 = \sum_{i=1}^k \frac{(R_i - S_i)^2}{P_i}$$

5. La suma de residus al quadrat dividit per la variància teòrica dels escons de cada partit :

$$D5 = \sum_{i=1}^k \frac{(R_i - S_i)^2}{V_i}$$

*S'utilitza la R per designar els escons que es varen donar realment el dia de les eleccions, S fa referència als escons publicats pels sondejos, P són les proporcions de vot reals, V estableix la variància teòrica extreta a partir de la realització dels sondejos ideals que es presenta en els apartats anteriors. Per últim, el subíndex "i" distingeix cada una de les candidatures amb presència al govern*

6. La suma de residus amb valor absolut dividit per la variància teòrica dels escons de cada partit :

$$D6 = \sum_{i=1}^k \frac{|(R_i - S_i)|}{V_i}$$

➤ **Proposta escollida**

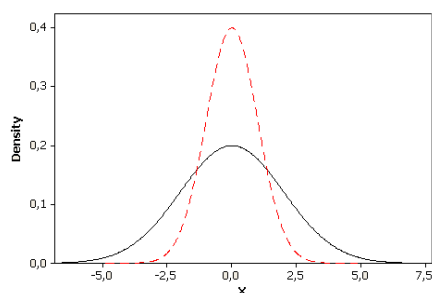
En totes s'aconsegueix clarificar quins sondejors i mostres són millors i pitjors. En aquest estudi s'agafarà una sola proposta: la 2 (suma de residus amb valor absolut).

S'ha cregut oportú agafar la suma de residus amb valor absolut perquè és molt senzilla d'entendre, de calcular i d'interpretar. Comptes quants escons en total ha fallat. Té el mateix sentit que la suma de residus al quadrat l'únic que amb aquesta última perds les unitats.

➤ **Proposta alternativa**

Una altra de les distàncies que es podia haver agafat és la que fa referència a dividir la suma de residus amb valor absolut (o la suma de residus al quadrat) per la variància teòrica d'escons de cada partit. Aquesta té un sentit més estadístic.

Al fer una enquesta a una petita part de la població, els escons que estimes de cada partit tenen una variància. És a dir, si s'agafa una altra mostra, independent de l'anterior, el més segur és que no s'obtinguin els mateixos escons. Es obvi pensar que és més fàcil obtenir el valor poblacional, o diferències molt petites respecte aquest, si la variància és petita que no si és gran.



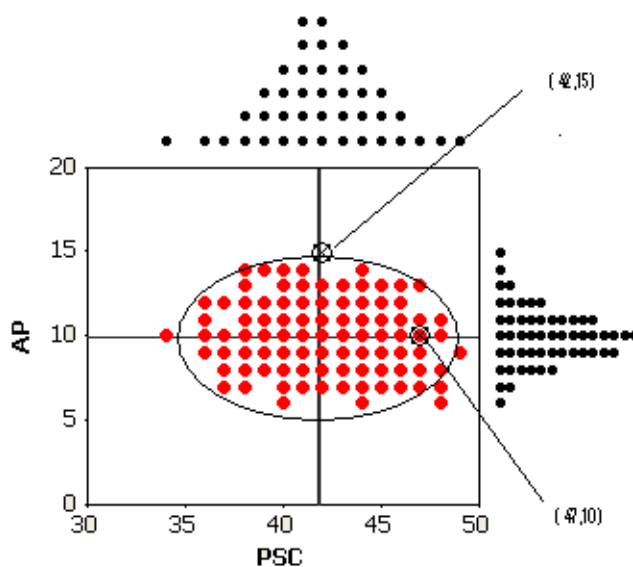
**figura 6.1.1** Dos distribucions normals. La negra amb mitjana 0 i desviació 2. i la vermella amb mitjan 0 i desviació 1.

Si es mira la figura 6.1.1 mentre que en la distribució vermella serà fàcil obtenir una mostra propera al valor central, en la negra és més probable que la mostra s'allunyi d'aquest mateix valor.

Això és el que passa en els escons a pronosticar de cada partit. Ni ha que són més fàcils que d'altres de predir ja que el rang de valors que poden prendre en la mostra és més petita que el d'altres candidatures.

Una altra manera d'entendre-ho és amb un exemple concret:

En la figura 6.1.2 hi ha els 10.000 escons del PSC i el PP (AP) que s'han obtingut en els sondejos ideals de les eleccions del parlament de Catalunya del 1984. La representació gràfica és mitjançant un diagrama Bivariant.



**figura 6.1.2** Diagrama bivariant dels resultats obtinguts en els sondejos ideals de 1984 en el PP i el PSC.

Els dos punts estan a la mateixa distància absoluta si s'usa la suma de residus amb valor absolut ( $|15 - 10| = 5$  i  $|47 - 42| = 5$ ). En canvi i pel fet que les distribucions dels escons dels dos partits són diferents, el primer punt queda fora del 97% de les 10.000 mostres fetes a l'atzar, mentre el segon entra a dins de les possibilitats normals de ser una més. D'aquesta manera sembla lògic potenciar de forma diferent segons el partit polític que s'està analitzant un error de la mateixa magnitud.

Tot i això, ja hi ha hagut altres projectes que han fet servir aquesta proposta de distància, també coneguda com a "distància de Mahalanobis". La suma d'errors amb valor absolut no és tan precisa però sí molt més fàcil de interpretar. Farem de la senzillesa una virtut i la prendrem com la nostra distància. A més cal remarcar que el final el que veu i l'interessa a l'immensa població són com de grans són els errors en els partits més importants. La distància en valor absolut possiblement és la que capta més aquest fet i la fa comprensible a la major part de la població.

## **6.2 Una única distància per sondeig: justificació**

Els sondejors que fan els diaris, tot i que al realitzar la mostra únicament obtenen una repartició d'escons dels diferents partits, donen un rang de possibilitats per cada un d'ells. N'hi ha que aquest interval és tant gran que no és gens informatiu. D'altres només publiquen el valor puntual que els hi ha sortit. En l'estudi es farà servir un sol valor que quantifiqui que lluny estan de la realitat per cada un dels sondejors.

D'aquesta manera es castiga una mica els sondejors amb intervals. I no és que estigui malament que els facilitin per fer prediccions. És just el contrari. És bo que sàpiguen que estan fent inferència d'una mostra petita sobre una població molt gran, i que ha estat per atzar que s'han donat aquests valors. Per tant, i per tenir una confiança més gran d'encertar el valor poblacional, és necessari prendre més d'una combinació.

Tot i això, per quantificar quant de ben fet estar el sondeig, és convenient valorar quant s'allunya amb un sol valor. I és que no oblidem que el que estem valorant és com està fet el sondeig. Per tant, totes les distàncies han de ser comparables independentment del diari i l'any de celebració.

## **6.3 Càlcul de les distàncies**

Els diaris donen uns intervals de confiança (no entrem en quin nivell de risc) dels escons que pronostiquen per cada candidatura. Per calcular les distàncies que s'han proposat és necessari interpretar aquestes múltiples combinacions que ens proporcionen. Ho podem plantejar per diferents vies.

Una opció que s'havia valorat era la de fer el promig dels escons de cada partit en els sondejors, i comparar directament aquests respecte la realitat. El que passa és que molts cops es tindrien escons amb decimals i no sempre sumarien el número d'escons totals establerts. Per aquest motiu s'ha descartat aquesta opció.

Finalment ens hem decantat per realitzar la interpretació amb un altre mètode. El que es fa és aconseguir tots els possibles resultats que ens permetin interpretar dels intervals. Es realitza per medi de permutacions. Es consideren únicament els que sumin el número d'escons totals a repartir. Serà a partir de totes aquestes opcions que es calcularan les distàncies.

Tot seguit es presenta un exemple per tal que es vegi com s'ha realitzat.

Data de l'any 1992 en les eleccions de la Generalitat. El diari és El País.

Sondeig :

**taula 6.3. 1** Sondeig d'El País en les eleccions de 1992.

CIU	PSC	PP	ERC	IC-IV
71-72	39-40	8	8	7-8

Possibles interpretacions:

**taula 6.3. 2** Possibles interpretacions dels intervals proposats pel sondeig d'El País l'any 1992.

CIU	PSC	PP	ERC	IC-IV	Suma
71	39	8	8	7	133
71	39	8	8	8	134
71	40	8	8	7	134
71	40	8	8	8	135
72	39	8	8	7	134
72	39	8	8	8	135
72	40	8	8	7	135
72	40	8	8	8	136

Les úniques que sumen 135 són les 3 que estan marcades. Seran les úniques combinacions que tindrem en compte per trobar les distàncies d'El País l'any 1992 respecte els sondejos teòrics i la realitat.

Un cop conegudes totes les possibilitats que presenten els diferents sondejos, s'ha realitzat un programa en R per establir les distàncies. El que si es desenvolupa és el càlcul de l'error en valor absolut dels sondejos dels diaris respecte els resultats reals. De fet, el que s'aconsegueix és el promig de totes les distàncies de cada un dels diaris (tindrem tants valors per diari com combinacions podem interpretar).

En aquest mateix programa també s'hi troben les diferències que prenen els diferents sondejos teòrics respecte els resultats. D'aquesta manera, es poden veure com de lluny estan els sondejos de les mostres teòriques respecte el que va succeir el dia de la celebració i comparar-ho amb les distàncies que presenten els diaris.

El procediment per aconseguir tots aquests valors és el següent :

Es guarden totes les possibilitats de tots els diaris en un mateix fitxer de text. Això s'ha fet per totes les convocatòries. En un altre s'hi guarden els resultats reals. I en el darrer totes les mostres dels sondejos perfectes. S'aconsegueixen unes taules com la taula 6.3.3, la taula 6.3.4 i la taula 6.3.5.

**taula 6.3.3** Possibles combinacions de repartició d'escons dels sondejos dels diaris a Catalunya.

Diaris	CIU	PSC	PP	ERC	IC-IV	Altres
LV <sub>1</sub>	$e_{1,CIU,LV}$	$e_{1,PSC,LV}$	$e_{1,PP,LV}$	$e_{1,ERC,LV}$	$e_{1,IV,LV}$	$e_{1,ALT,LV}$
...	...	...	...	...	...	...
LV <sub>n</sub>	$e_{n,CIU,LV}$	$e_{n,PSC,LV}$	$e_{n,PP,LV}$	$e_{n,ERC,LV}$	$e_{n,IV,LV}$	$e_{n,ALT,LV}$
...	...	...	...	...	...	...
EM <sub>1</sub>	$e_{1,CIU,EM}$	$e_{1,PSC,EM}$	$e_{1,PP,EM}$	$e_{1,ERC,EM}$	$e_{1,IV,EM}$	$e_{1,ALT,EM}$
...	...	...	...	...	...	...
EM <sub>n</sub>	$e_{n,CIU,EM}$	$e_{n,PSC,EM}$	$e_{n,PP,EM}$	$e_{n,ERC,EM}$	$e_{n,IV,EM}$	$e_{n,ALT,EM}$

**taula 6.3.4** Els 10.000(N) sondejos ideals aconseguits en una de les convocatòries a Catalunya.

Mostres	CIU	PSC	PP	ERC	IC-IV	Altres
M <sub>1</sub>	$e_{1,CIU}$	$e_{1,PSC}$	$e_{1,PP}$	$e_{1,ERC}$	$e_{1,IV}$	$e_{1,ALT}$
...	...	...	...	...	...	...
M <sub>N</sub>	$e_{N,CIU}$	$e_{N,PSC}$	$e_{N,PP}$	$e_{N,ERC}$	$e_{N,IV}$	$e_{N,ALT}$

**taula 6.3.5** Resultats electorals d'una convocatòria a Catalunya

CIU	PSC	PP	ERC	IC-IV	Altres
$e_{CIU}$	$e_{PSC}$	$e_{PP}$	$e_{ERC}$	$e_{IV}$	$e_{ALT}$

La distància entre els diaris i els resultats reals es troba fent el promig de totes les diferències entre possibles combinacions del sondeig i resultat real.

$$D_{diari\_reals} = \frac{\sum_{i=1}^n \sum_{p=CIU}^{altres} |Sd_{p,i} - R_p|}{n}$$

$Sd_{p,i}$  fa referència als escons pronosticats pels sondejos dels diaris,  $R_p$  als escons que es varen donar el dia de l'escrutini. La constant  $n$  és el número de combinacions que es poden interpretar del sondeig i es distingeixen per el subíndex "i". Per últim,  $P$  diferencia entre partits.

Continuació del cas d'El País a l'any 1992. Ara es tindrà la matriu de resultats del diari i l'assignació d'escons que es va donar el dia clau.

CIU	PSC	PP	ERC	IC-IV
71	40	8	8	8
72	39	8	8	8
72	40	8	8	7

CIU	PSC	PP	ERC	IC-IV
$e_{CIU}$ (70)	$e_{PSC}$ (40)	$e_{PP}$ (7)	$e_{ERC}$ (11)	$e_{IV}$ (7)

- $S1 = |71 - 70| + |40 - 40| + |8 - 7| + |8 - 11| + |8 - 7| = 6$
- $S2 = |72 - 70| + |39 - 40| + |8 - 7| + |8 - 11| + |8 - 7| = 8$
- $S3 = |72 - 70| + |40 - 40| + |8 - 7| + |8 - 11| + |7 - 7| = 6$

$$DPaís\_reals92 = \frac{\sum_{i=1}^3 S_i}{3} = \frac{6+8+6}{3} = 6,667$$

L'altra distància que cal expressar és la que fa referència als sondejos teòrics respecte els resultats reals. També es pot trobar el promig, si bé ens interessa més obtenir totes les diferències i estimar un interval al 95% d'aquestes. Així se sabrà en quines distàncies ens podem moure amb únicament el factor aleatori i comparar-les amb les dels sondejos dels diaris.

$$Dteòrics\_reals = \frac{\sum_{i=1}^{10.000} \sum_{p=CIU}^{altres} |St_{p,i} - R_p|}{10.000}$$

$St_{p,i}$  representen els escons dels sondejos teòrics i  $R_p$  els escons reals.  $P$  distingeix entre partits polítics.

Continuació del cas d'El País a l'any 1992. Tot i ser un cas comú per tots anys anem a veure com s'aconseguiria en les eleccions del 92. En aquesta situació de sumands n'hi ha 10.000, tants com sondejos ideals hi ha. L'estratègia a seguir és la mateixa al cas anterior.

- $S1 = |e_{1,CIU} - e_{CIU}| + |e_{1,PSC} - e_{PSC}| + |e_{1,PP} - e_{PP}| + |e_{1,ERC} - e_{ERC}| + |e_{1,IV} - e_{IV}|$
- ...
- $S10.000 = |e_{10.000,CIU} - e_{CIU}| + |e_{10.000,PSC} - e_{PSC}| + |e_{10.000,PP} - e_{PP}| + |e_{10.000,ERC} - e_{ERC}| + |e_{10.000,IV} - e_{IV}|$

$$Dteo\_reals92 = \frac{\sum_{i=1}^{10.000} S_i}{10.000}$$



El programa en R corresponent al càlcul de les distàncies es troba en l'annex 3.

#### 6.4 Distàncies entre els sondejos i els resultats reals

taula 6.4.1 *Distàncies entre els Sondejos (perfectes i de diaris) en totes les eleccions del Parlament de Catalunya.*

ANY	Sondejos/ resultats	Distància
1984	Perfectes	(2-10)
	La Vanguardia	13
	El Periódico	*
	El País	29,07692
	ABC	*
	El Mundo	*
1988	Perfectes	(2-14)
	La Vanguardia	14
	El Periódico	18
	El País	17,6667
	ABC	*
	El Mundo	*
1992	Perfectes	(2-12)
	La Vanguardia	6
	El Periódico	*
	El País	6,66667
	ABC	*
	El Mundo	*
1995	Perfectes	(2-10)
	La Vanguardia	21,52632
	El Periódico	20,33333
	El País	18,06452
	ABC	8,88889
	El Mundo	13,26667

ANY	Sondejos/ Resultats	distància
1999	Perfectes	(2-12)
	La Vanguardia	6
	El Periódico	19,23256
	El País	9,6
	ABC	16,66667
	El Mundo	10
2003	Mostres	(2-10)
	La Vanguardia	20
	El Periódico	16,90909
	El País	18
	ABC	19
	El mundo	15
2006	Perfectes	(2-12)
	La Vanguardia	14
	El Periódico	9,333
	El País	14
	ABC	10
	El Mundo	10

S'observen ja a simple vista diferents anys on els pronòstics dels diaris és van allunyar molt de les distàncies lògiques per culpa del factor aleatori. Mentre que els sondejos perfectes tenen uns intervals de distàncies respecte els resultats reals d'entre 2 i 12 (alguns arriben a 10 o 14), gran part de sondejos de diaris presenten distàncies superiors.

**taula 6.4.2** Distàncies entre els sondejos (perfectes i de diaris) en totes les eleccions del Congrés dels Diputats.

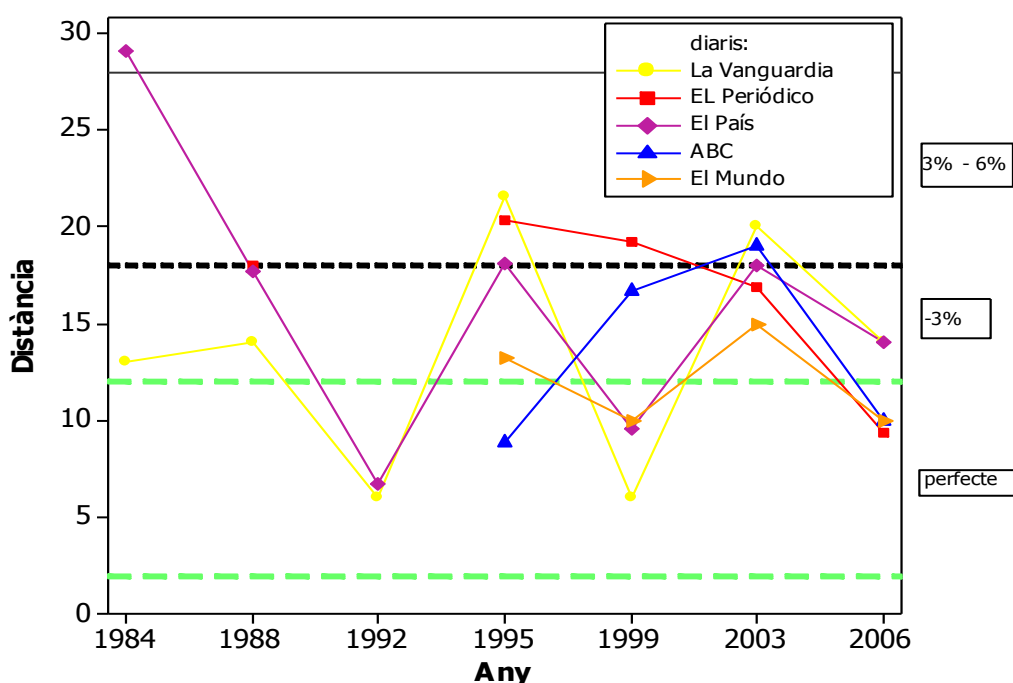
ANY	Sondejos/ resultats	Distància	ANY	Sondejos/ Resultats	distància
1986	Perfectes	4-20	2000	Perfectes	2-18
	La Vanguardia	-		La Vanguardia	36
	El Periódico	28		El Periódico	34
	El País	-		El País	30,91176
	ABC	64		ABC	41,31579
	El Mundo	-		El Mundo	51,42857
1989	Perfectes	4-22	2004	Perfectes	2-16
	La Vanguardia	12,96216		La Vanguardia	43
	El Periódico	-		El Periódico	60,13953
	El País	18,48508		El País	58,22857
	ABC	-		ABC	66,5
	El Mundo	26		El mundo	51,60825
1993	Perfectes	4-22	2008	Perfectes	2-16
	La Vanguardia	30,20833		La Vanguardia	13,77778
	El Periódico	31,65757		El Periódico	26,38545
	El País	34,49186		El País	12
	ABC	41		ABC	-
	El Mundo	45,56522		El Mundo	16,37310
1996	Perfectes	2-16			
	La Vanguardia	17,21569			
	El Periódico	37,48242			
	El País	44,70395			
	ABC	54,84536			
	El Mundo	49,88722			

En aquestes eleccions, de la mateixa manera que succeïa a Catalunya, ja es poden intuir els anys on els pronòstics van ser molt erronis. Sobretot se'n destaca l'any 2004. També s'ha de dir que no tots estan tant malament i atesa la dificultat de realitzar un bon sondeig n'hi ha alguns que aconseguen apropar-se força a la realitat. En l'última de les convocatòries, per exemple, sembla que varen estar bastant encertats.

## 7 REPRODUCCIÓ DE LES DISTÀNCIES OBTINGUDES

Un cop assolides les distàncies dels sondejos respecte els resultats del dia de l'escrutini es poden representar gràficament. En la figura 7.1 i la figura 7.2 s'hi presenten els diferents valors separats per any i diari. Les dues primeres línies discontinúes marquen l'interval al 95% de les distàncies que s'obtenen de realitzar un sondeig perfecte. La següent línia discontinúea és considerada la distància límit que poden arribar a presentar els sondejos (canvi en la intenció de vot d'un sondeig perfecte de com a màxim el 3%). Si el valor és més alt, el sondeig es podria qualificar d'incorrecte.

Fins a la línia contínua es tractaria d'un sondeig perfecte amb un canvi en la intenció de vot de com a màxim el 6%.



**figura 7.1** Distàncies entre els sondejos i els resultats separats pels diferents diaris a les eleccions del Parlament de Catalunya.

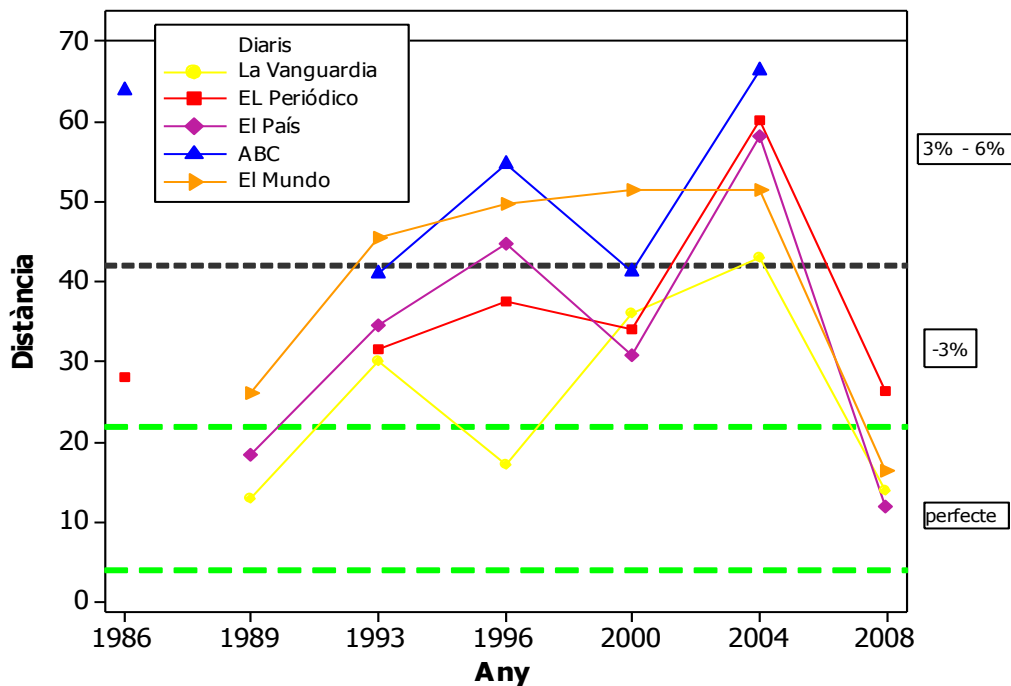


figura 7.2 Distàncies entre els sondejos i els resultats separats pels diferents diaris en les eleccions Generals.

### 7.1 Una altra via: Anàlisi de components principals

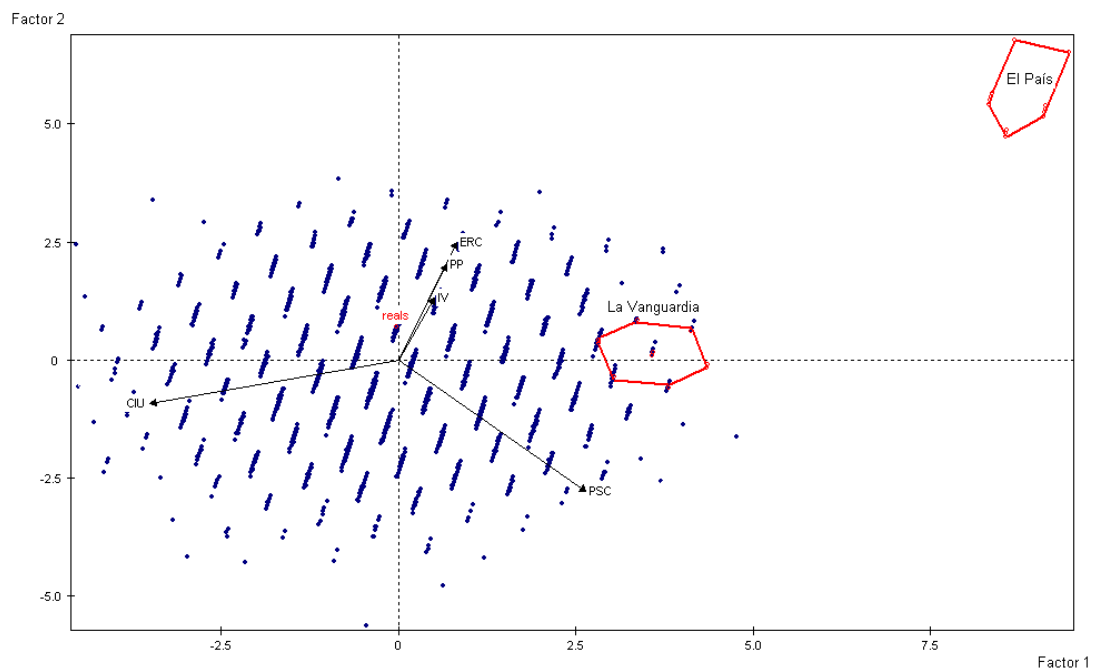
Una manera d'interpretar les diferències entre els sondejos teòrics i els dels diaris és mitjançant tècniques d'anàlisi multivariant de dades. El que es tracta és de reduir dimensions. Inicialment partim d'unes dades (tots els sondejos perfectes) amb 5 o 6 variables, tantes com partits polítics han aconseguit escons. Interpretar-ho gràficament és molt complicat. Per això, es talla el conjunt de les dades en forma de pla. Es realitza de tal forma que es maximitza la variabilitat total de les dades. No s'aconsegueixen percentatges molt elevats (vora el 50%), però si els suficients perquè es pugui entendre un gràfic amb dues dimensions. En definitiva, es perd informació per guanyar amb interpretació. La metodologia que s'aplica és la d'aconseguir els vectors i valors propis de la matriu inicial. En el cas que ens ocupa es tracta de matrius de dimensions 5 (o 6) x 10.000 i 6 (o 7) x 10.000. De valors propis n'hi ha  $n-1$  (sigui  $n$  el nombre de columnes o variables). I és per medi dels vectors obtinguts que es determinen els nous eixos.

A continuació es presenten les gràfiques que s'extreuen de realitzar l'anàlisi de components principals en totes les eleccions. S'agafen únicament les dues primeres components. És d'interès fixar-se que els resultats reals no corresponen a la mitjana de les mostres (ja que el punt central o origen correspon a la mitjana aritmètica de cada una

de les variables). Com ja s'ha comentat en apartats anteriors, existeix un biaix. La realització de l'anàlisi i els gràfics ha estat mitjançant el software estadístic SPAD.

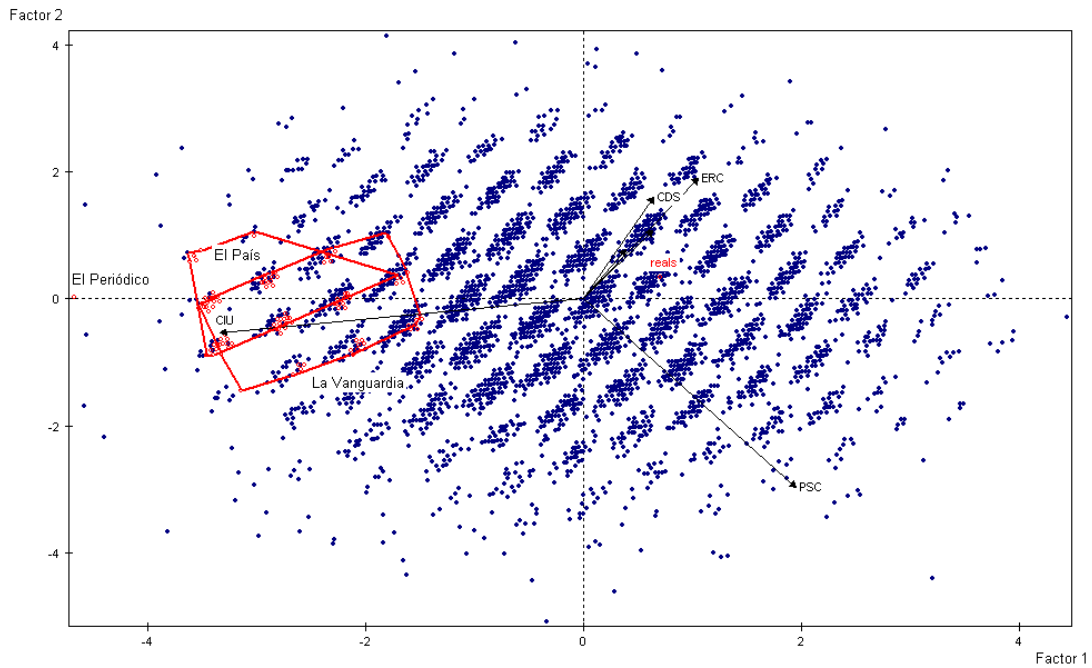
### 7.1.1 Gràfiques d'ACP en el Parlament de Catalunya

En les eleccions de 1984 el diari El País va errar totalment els pronòstics. Es comprova en la figura 7.1.1 que lluny està dels sondejos teòrics. Si es presta atenció en les direccions de creixement (les fletxes negres) s'indica com els sondejos creien que el PSC aconseguiria més escons en detriment de CIU.



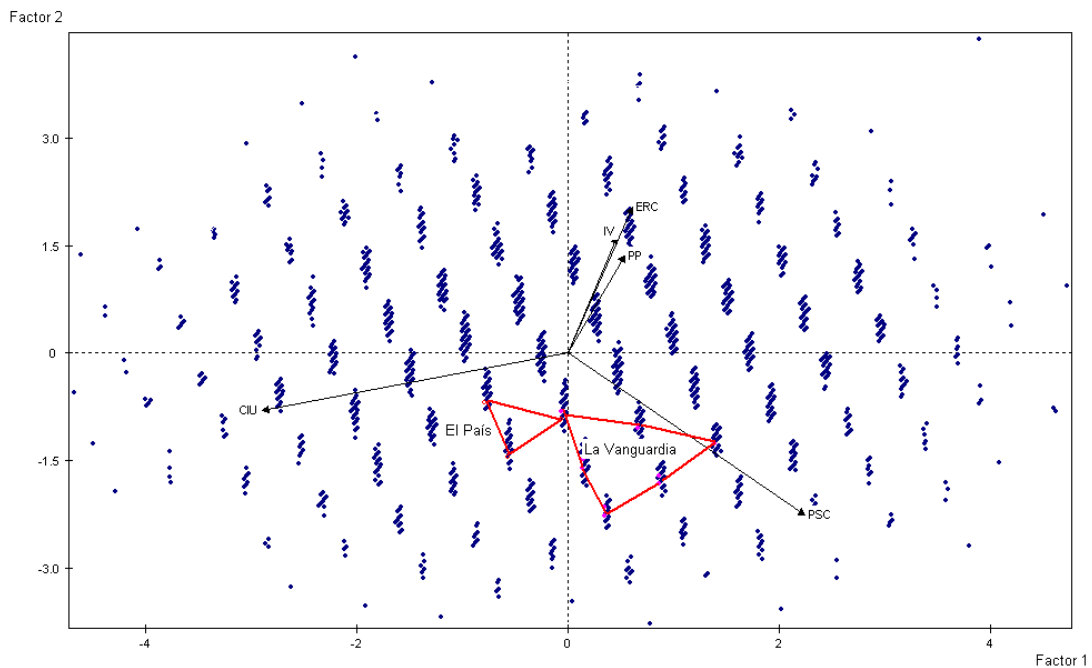
**figura 7.1.1** Anàlisi de les dues primeres components de la PCA de les eleccions del Parlament de 1984.

En el 1988 els sondejos no s'allunyen dels perfectes. Els tres estan encaminats en la mateixa direcció. Fan pronòstics a l'alça pels Convergents. El que pitjor ho fa sembla que és El Periódico.



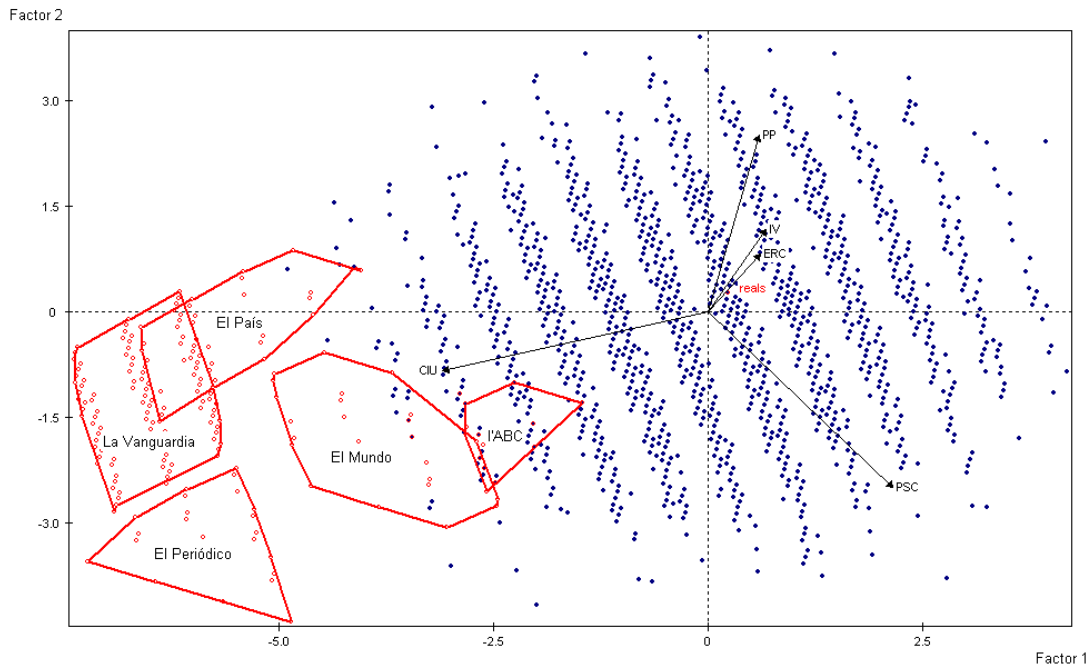
**figura 7.1.2** Anàlisi de les dues primeres components de la PCA de les eleccions del Parlament de 1988

En la tercera convocatòria que s'estudia del Parlament, els dos sondejos són ideals. Estan totalment a dins al interval que presenten els perfectes.



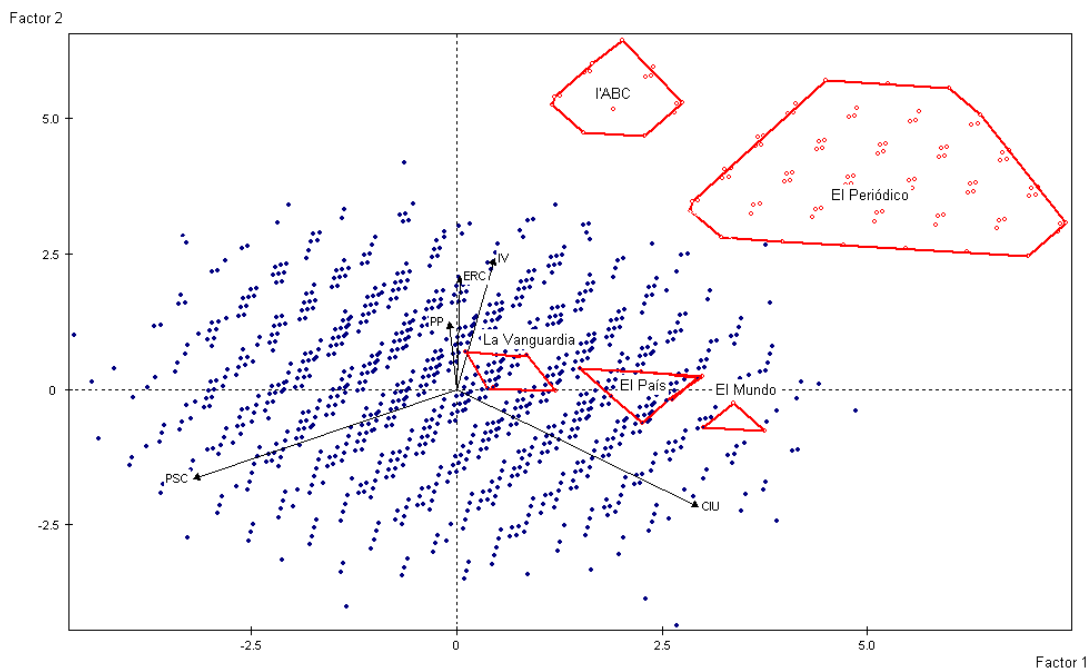
**figura 7.1.3** Anàlisi de les dues primeres components de la PCA de les eleccions del Parlament de 1992.

El 1995, contràriament a l'últim any, la majoria de sondejos són erronis. Tots creuen que CIU aconseguirà més escons dels que realment materialitza. El Mundo i l'ABC encara s'apropen als perfectes, els altres tres no.



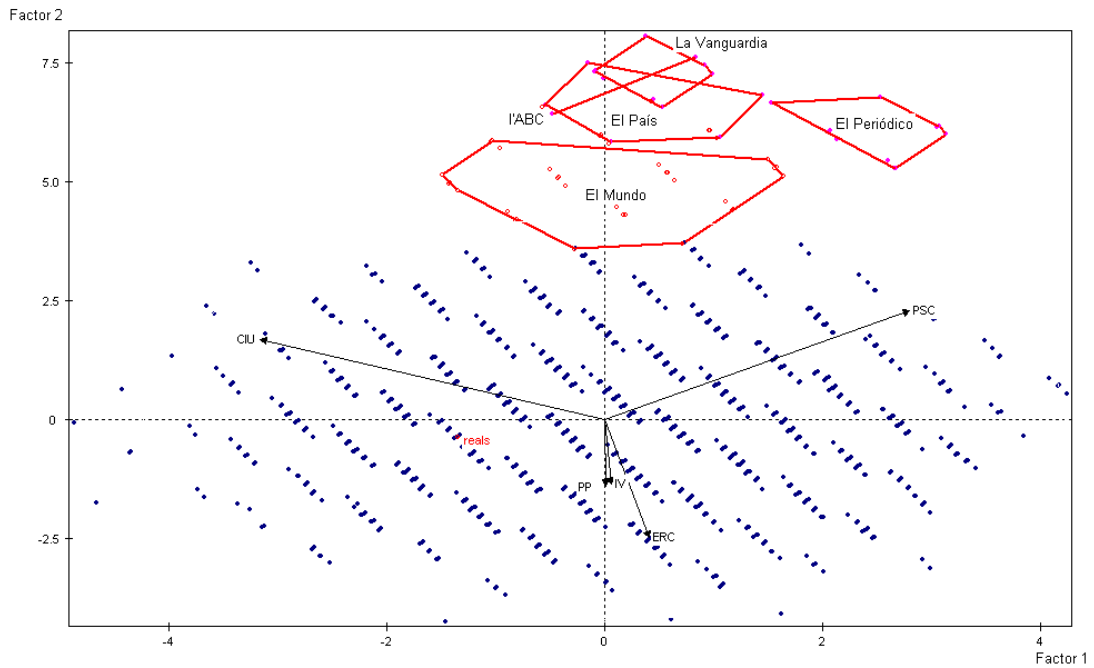
**figura 7.1.4** Anàlisi de les dues primeres components de la PCA a les eleccions del Parlament de Catalunya 1995.

En 1999 clarament s'aprecien dos sondejos fora dels perfectes i tres a dins. També sembla que estan encaminats a sobreestimar a CIU.



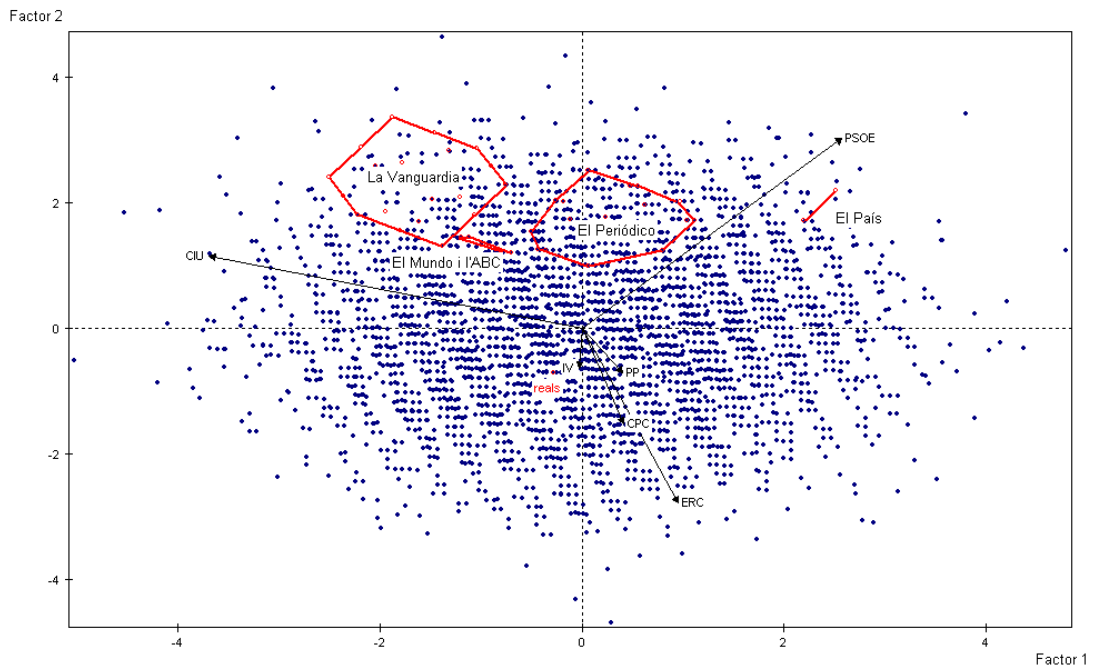
**figura 7.1.5** Anàlisi de les dues primeres components de la PCA de les eleccions del Parlament de Catalunya al 1999.

En el 2003, els sondejos fallen sobretot en els partits més minoritaris. El que millor s'aproxima és El Mundo.



**figura 7.1.6** Anàlisi de les dues primeres components de la PCA de les eleccions del Parlament de Catalunya al 2003.

En la darrera de les convocatòries al Parlament, tots els sondejos aconseguen reflexar bé el que va passar el dia de les eleccions.

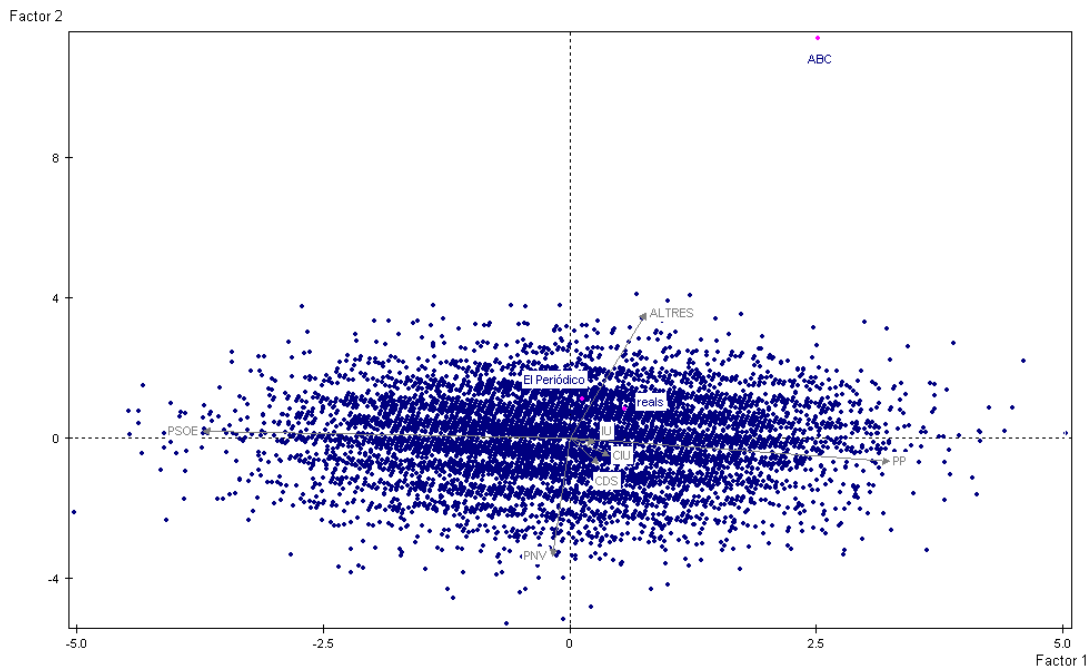


**figura 7.1.7** Anàlisi de les dues primeres components de la PCA de les eleccions del Parlament de Catalunya 2006.



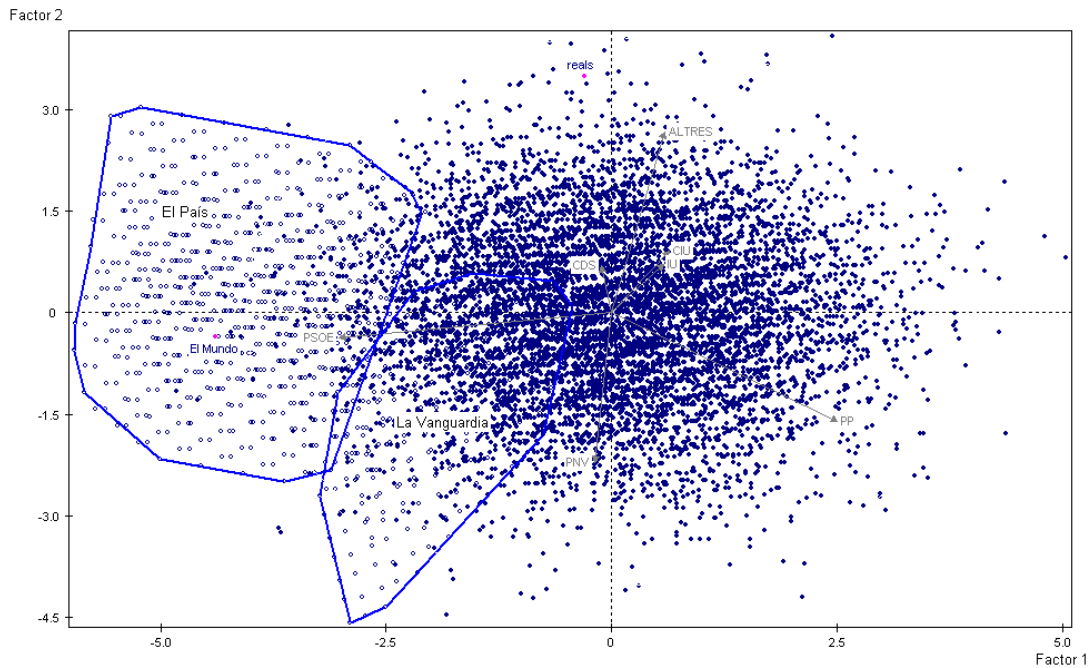
### 7.1.2 Gràfiques d'ACP en el Congrés dels Diputats

En la primera de les convocatòries a la cambra espanyola, hi ha dos sondejors. El d'El Periódico que s'adequa bastant bé a la realitat i l'ABC que falla a l'alça en els escons a partits més minoritaris (Altres).



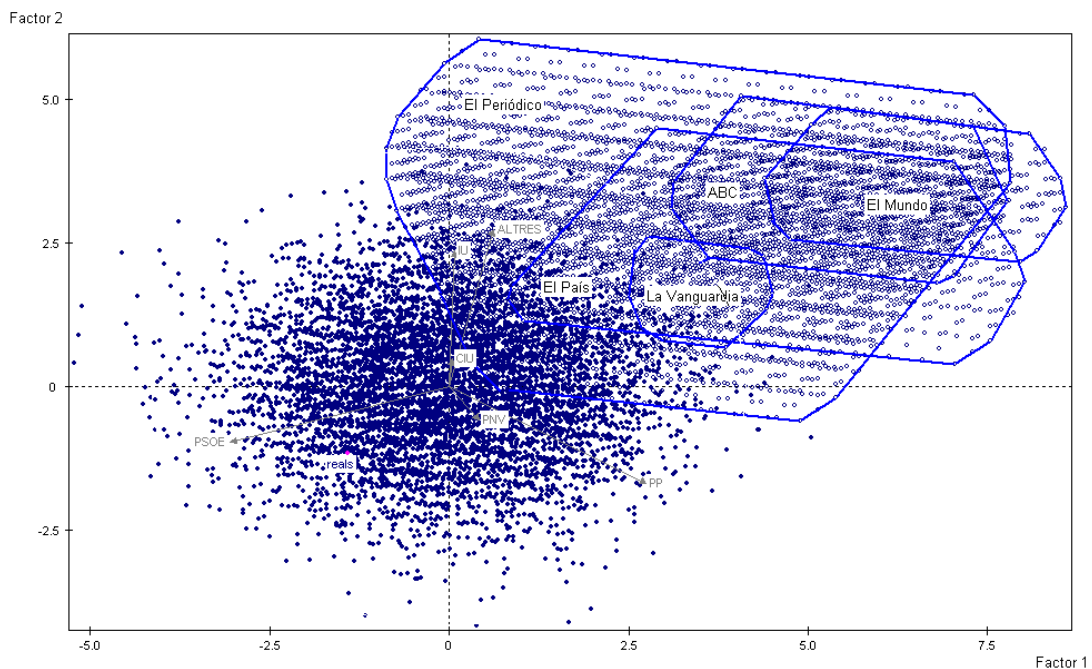
**figura 7.1.8** Anàlisi de les dues primeres components de la PCA de les eleccions del Congrés dels Diputats al 1986.

El 1989, els tres diaris pronostiquen a favor del Partido Socialista. No s'allunyen molt dels sondejos perfectes. Cal destacar el fet que els intervals de La Vanguardia i El País siguin tant grans.



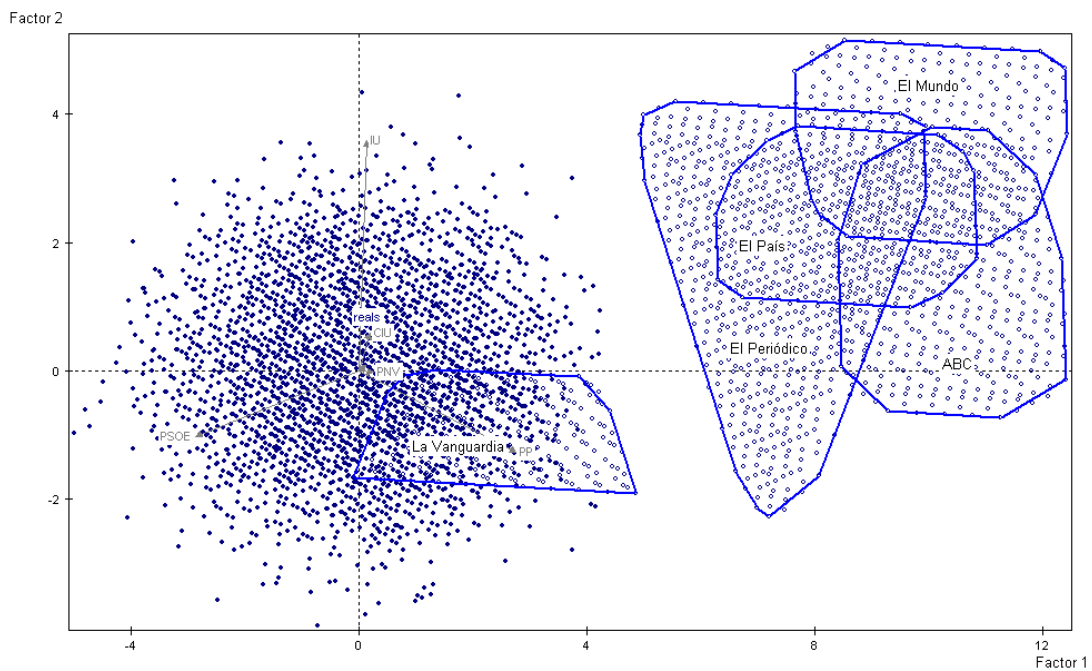
**figura 7.1.9** Anàlisi de les dues primeres components de la PCA de les eleccions del Congrés dels Diputats al 1989.

El 1993, destaca el interval tant gran que proporciona El Periódico. No és gens informatiu. El Mundo i l'ABC són els que erren més. Tots solen desfavorir el PSOE.



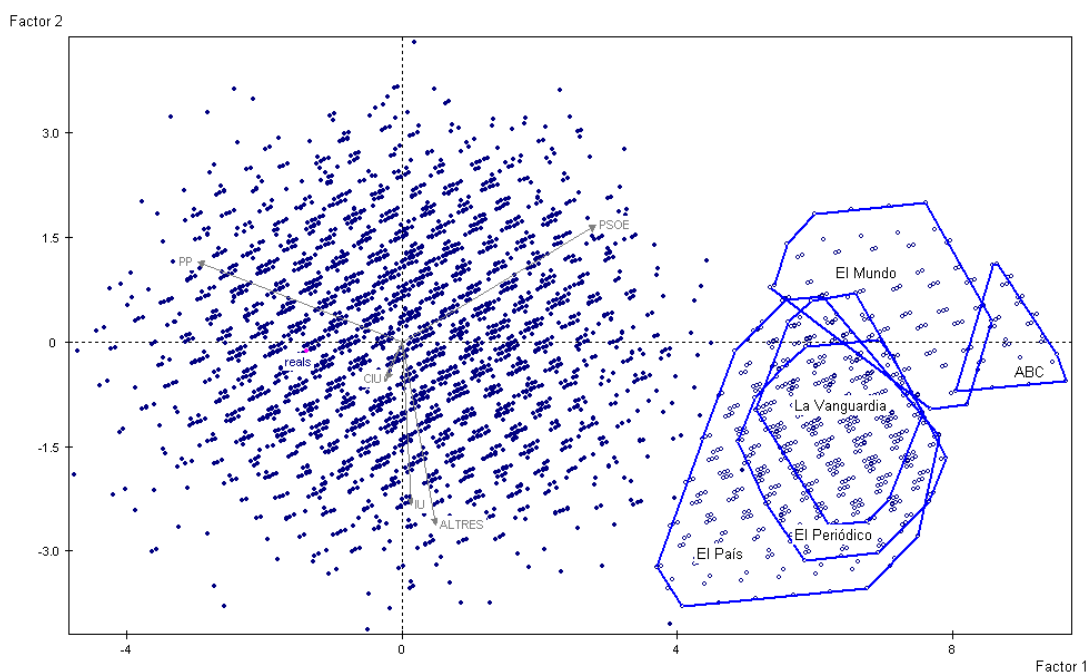
**figura 7.1.10** Anàlisi de les dues primeres components de la PCA de les eleccions del Congrés dels Diputats al 1993.

En 1996, La Vanguardia és l'únic diari que encerta el resultat del dia de l'escrutini. Tots els altres s'allunyen molt dels sondejos ideals. Igual que en 1993 pronostiquen a la baixa pel PSOE.



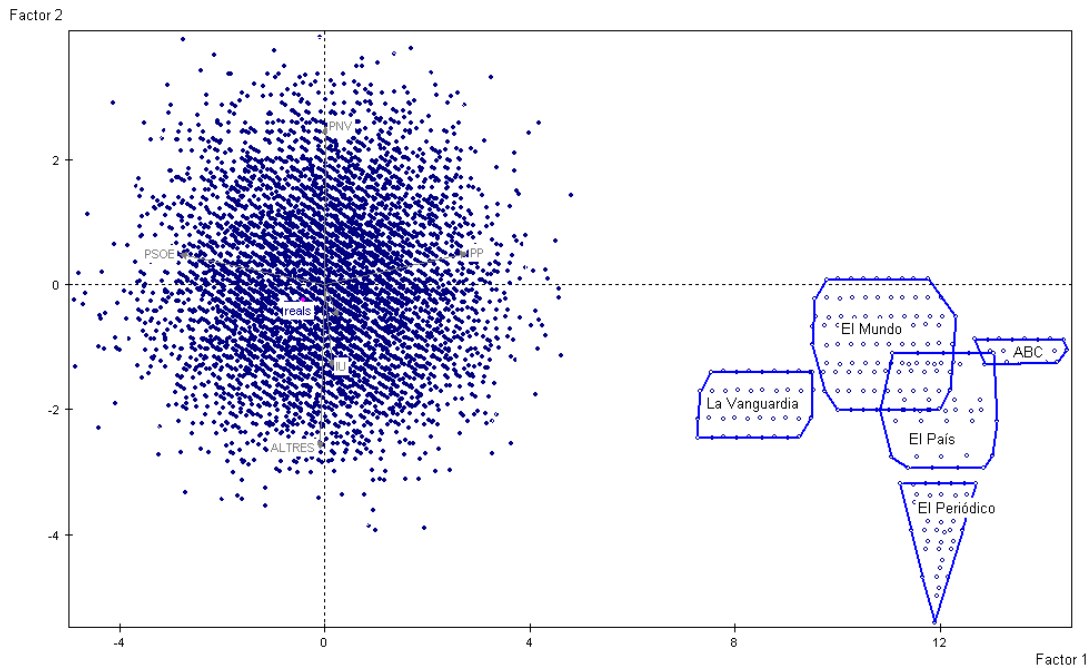
**figura 7.1.11** Anàlisi de les dues primeres components de la PCA de les eleccions del Congrés dels Diputats al 1996.

Al 2000 tots els diaris s'equivoquen considerablement. Potencien més del compte els escos al PSOE.



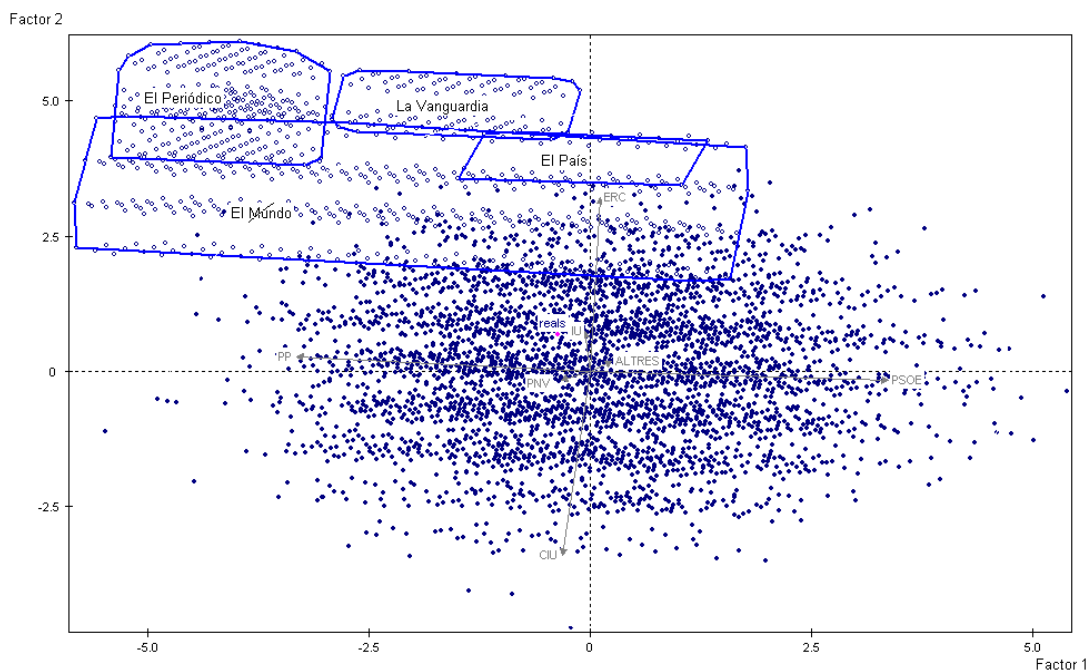
**figura 7.1.12** Anàlisi de les dues primeres components de la PCA de les eleccions del Congrés dels Diputats al 2000.

En l'any 2004 es veuen les diferències més grans. Tots els diaris fallen en la mateixa direcció. Pronostiquen més escons del compte pel PP en detriment del PSOE.



**figura 7.1.13** Anàlisi de les dues primeres components de la PCA de les eleccions del Congrés dels Diputats al 2004.

En l'última convocatòria al Congrés, els diaris no s'allunyen molt de la realitat. Tendeixen a afavorir al Partido Popular si bé és situen força a prop dels sondejos perfectes.



**figura 7.1.14** Anàlisi de les dues primeres components de la PCA de les eleccions del Congrés dels Diputats al 2008.

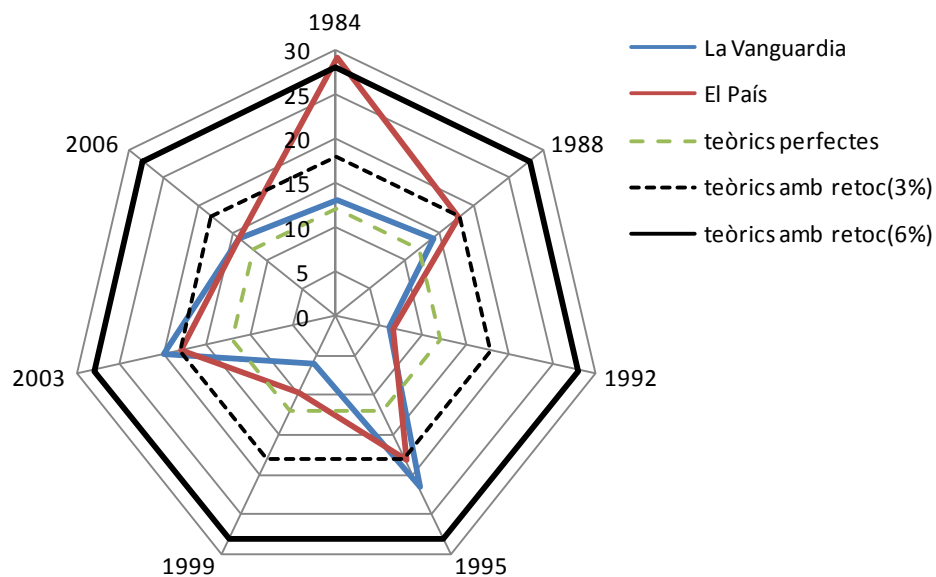
## 8 ANÀLISI DE LES DISTÀNCIES I COMPARACIONS

### 8.1 Evolució dels sondejos al llarg dels anys

Pot ser interessant estudiar com han evolucionat els sondejos al llarg dels anys. Es vol constatar en quins anys els sondejos s'han aproximat més a la realitat i s'hi ha existit una millora a mesura que s'han succeït les celebracions.

#### 8.1.1 Evolució dels sondejos al Parlament de Catalunya

No queda clar que els pronòstics hagin estat més encertats al llarg del temps. S'ha d'aclarir que no és fàcil establir tendències tenint tant poques observacions. Al Parlament de Catalunya, els dos diaris que han publicat més sondejos són La Vanguardia i El País, concretament 7 de les 8 convocatòries que s'han produït. Es pot comprovar com s'apropen a la realitat.



**figura 8.1.1** Distàncies individuals en el Parlament dels diaris La Vanguardia i El País ordenades per l'any. La traça verda discontinua és el límit superior dels sondejos perfectes i la negra discontinua el límit superior dels sondejos teòrics amb un retoc a les proporcions de vot del 3%. Per últim la línia contínua de color negre és la d'un canvi en la intenció de vot del 6%.

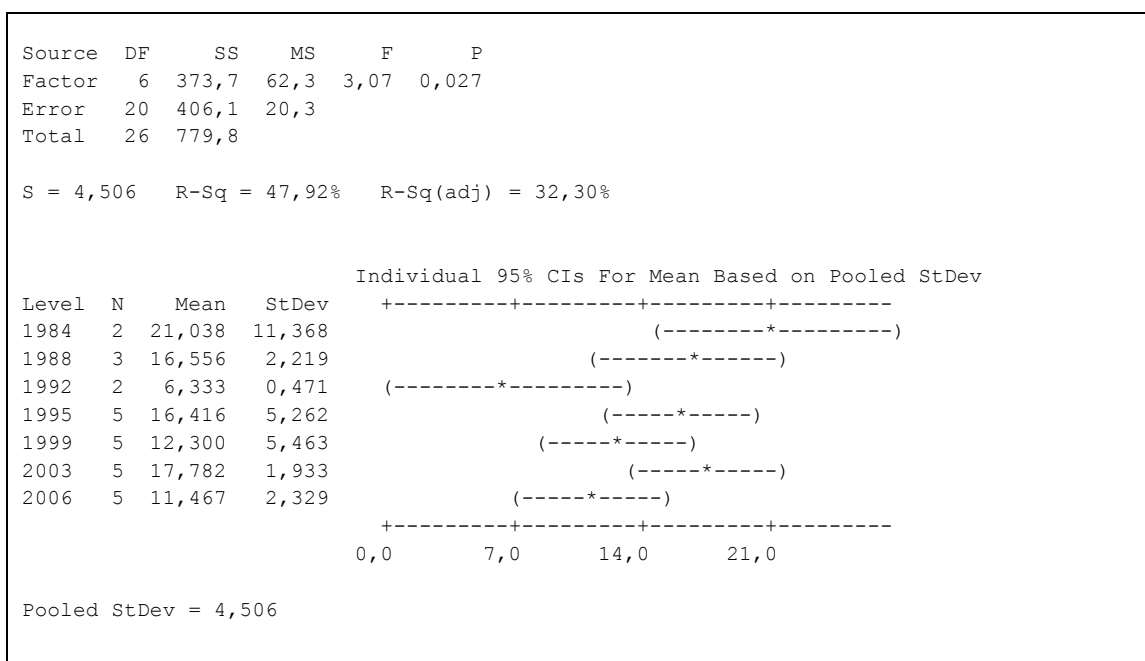
Mirant la figura 8.1.1 no s'aprecia la millora en cap dels dos diaris. Això sí, el comportament de les distàncies no és molt normal. Llevat el primer any, presenten un patró molt semblant d'encert de forma intermitentment. Un any sí i l'altre no. Es un fet

curiós, i més si no tenim constància de cap motiu en especial que afectés de forma considerable la intenció de vot en els anys més desafortunats. El més lògic serien una sèrie de dades independents de l'any i del diari i pel contrari no és així.

Malgrat que no s'hi veu una millora, sí que hi ha anys que han estat mes bons que d'altres. Encara que només consten els pronòstics d'El País i La Vanguardia per establir diferències, el millor sembla que és el 1992 (mirar la figura 8.1.3).

La realització d'una comprovació formal per veure si hi ha diferències significatives en els múltiples anys és mitjançant un Anàlisi de la Variància i està en la figura 8.1.2. Amb un p-valor del 0,027 el factor any és significatiu. Aquesta probabilitat indica que hi ha al menys alguna diferència entre dos anys. Si es mira el gràfic de confiança que ens proporciona l'ANOVA es pot apreciar que l'any 1984 i el 2003 els pronòstics han estat pitjors que el 1992. Entre els altres anys no s'hi veuen diferències significatives. Malgrat tot, aquesta prova no és gens informativa pel fet que hi ha molt poques observacions en algun dels anys. Hi ha casos on únicament es varen presentar dos pronòstics, pel qual deduir diferències és molt complicat.

En definitiva, tot i haver-hi anys on els sondejos van estar més encertats que d'altres s'aprecien poques diferències i ni molt menys una tendència a millorar.



**figura 8.1.2** Sortida del Minitab de l'ANOVA amb factor a analitzar l'any (Parlament de Catalunya).

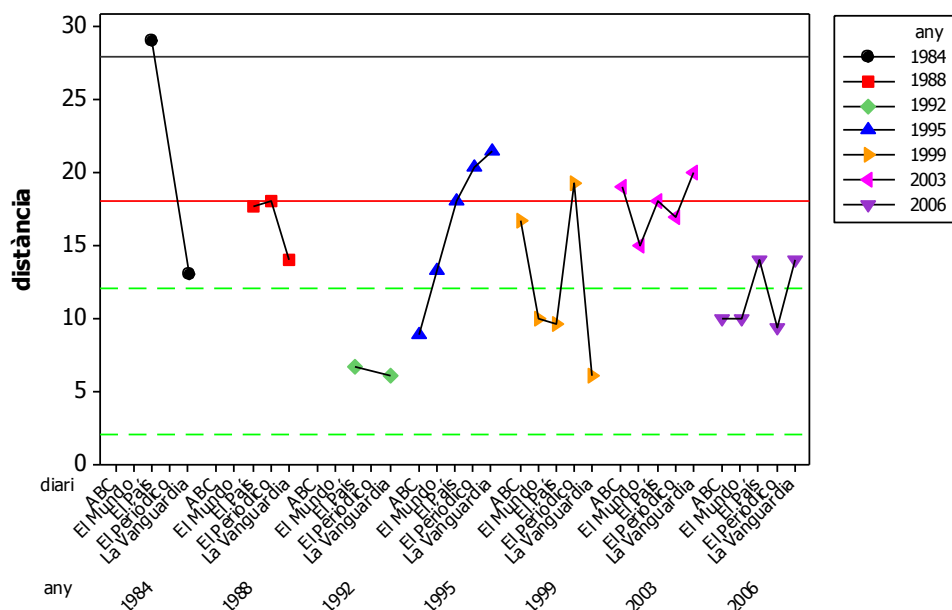
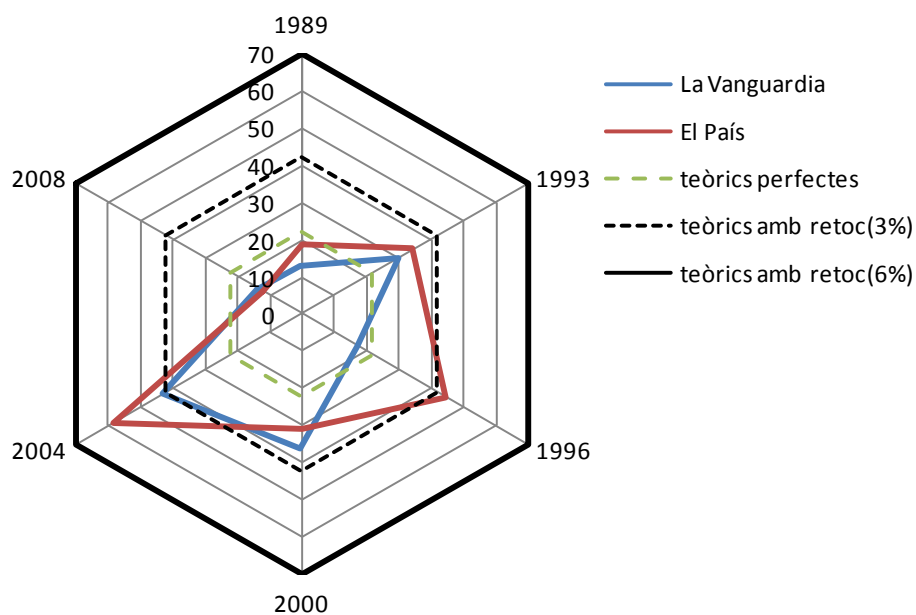


figura 8.1.3 Gràfic de les distàncies que presenten els diaris a mesura que s'han succeït les celebracions al Parlament de Catalunya.

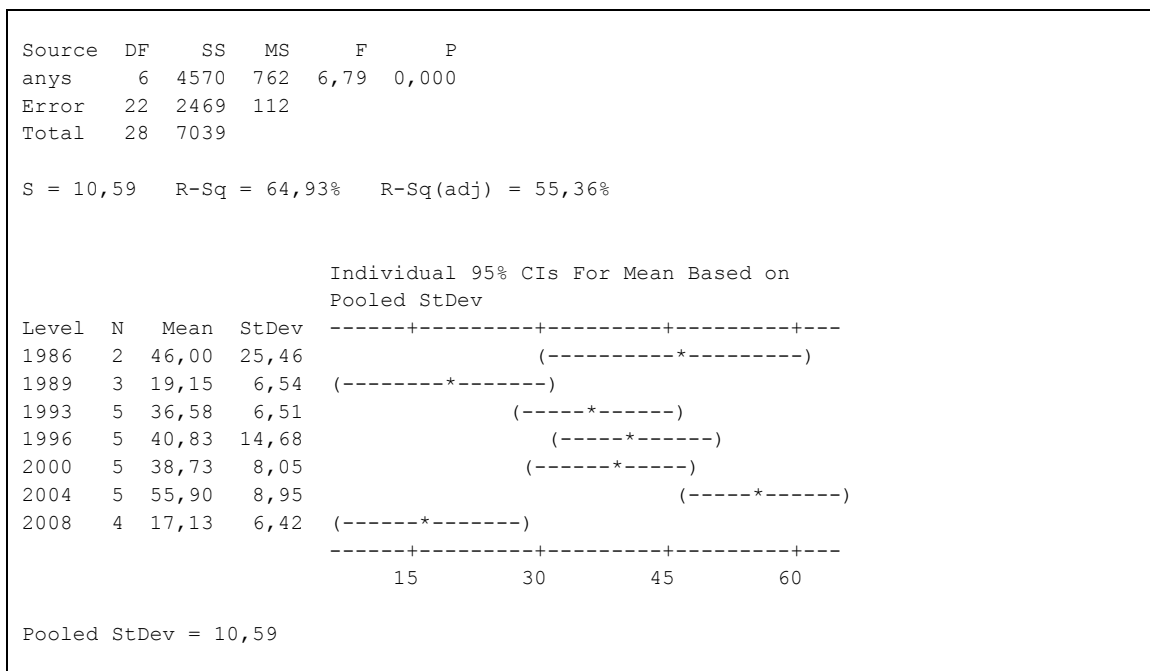
## 8.1.2 Evolució dels sondejos al Congrés dels Diputats

En les eleccions al Congrés tampoc s'hi destaca cap millora. Si es miren les distàncies que presenten els diaris El País i La Vanguardia, a mesura que s'han succeït les celebracions (en la figura 8.1.4), no es veu cap mena d'evolució positiva. En el primer i últim any que varen publicar els pronòstics, els dos estan encertats. En els altres anys els errors són més grans. De la mateixa manera que a Catalunya les distàncies que presenten segueixen un patró bastant similar (potser no tant clar com en el Parlament). Si s'obvia l'any 1996, on La Vanguardia es va aproximar molt més a la realitat que El País, s'hi aprecia un comportament comú. L'any que es van equivocar més els dos diaris, va ser el 2004 i el que menys el 2008.

Tot i que tampoc es pot concloure una millora en els pronòstics al Congrés si que s'intueixen anys millors que d'altres. En la figura 8.1.5 hi ha la sortida de minitab de l'Anàlisi de la Variança amb variable explicativa l'any. Com ja s'intuïa, únicament mirant la gràfica de distàncies de La Vanguardia i El País, els anys que varen aproximar-se més als resultats reals són el 1989 i el 2008. Hi ha diferències significatives respecte totes les altres convocatòries, on els pronòstics es van allunyar molt més. Destaquen negativament les eleccions del 2004. Cal recordar que a pocs dies de la celebració d'aquestes eleccions es van produir els atemptats a l'estació d'Atocha de Madrid. Sembla que una de les conseqüències va ser un canvi en quan a la intenció de vot del ciutadà. Igual que succeïa amb la prova ANOVA al Parlament, es tracta d'un test molt poc informatiu al tenir molt poques observacions per any.



**figura 8.1.4** Distàncies individuals en el Congrés del Diputats dels diaris La Vanguardia i El País ordenades per l'any.



**figura 8.1.5** Sortida del Minitab de l'ANOVA amb factor a analitzar l'any ( Eleccions Generals). Al tractar-se de tant poques observacions la prova és bastant irrellevant.

Hi ha una alta consideració que cal esmentar. En la figura 8.1.6 hi ha els errors absoluts del tots els diaris al transcórrer les convocatòries. Podem intuir l'existència de dos pics. Un a 1996 i un altre a 2004. Sobretot es donen en El Periódico, El País i L'ABC (mirar



també la figura 7.2 del tema 7). Curiosament, aquests dos anys són els que es va produir un canvi de govern. El 1996, el Partido Popular va ser per primera vegada el màxim representant en el Congrés. Tot i això, i contràriament al que es podia pensar, les prediccions encara donaven una diferència major a favor del partit liderat per Aznar. El 2004 i després de 8 anys de lideratge popular, són els Socialistes els que varen tornar entrar al govern. Aquest cop sí que els sondejos apostaven pel PP com a partit més votat i va succeir just el contrari. Pot ser d'interès estudiar si aquesta hipòtesi es reafirma en propers canvis de governs.

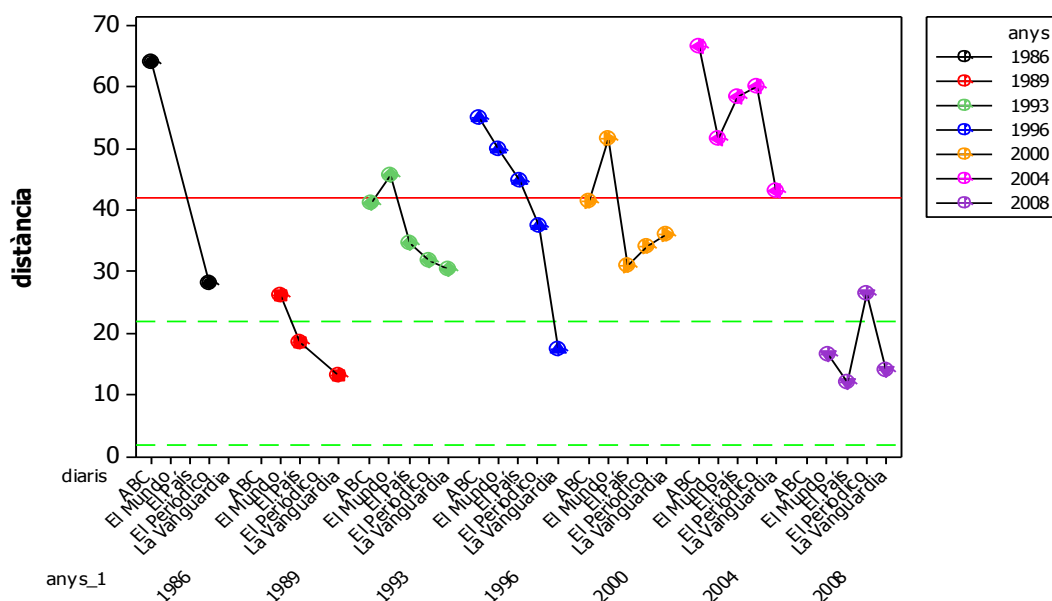


figura 8.1.6 Gràfic de les distàncies que presenten els diaris a mesura que s'han succeït les celebracions al Congrés dels Diputats.

## 8.2 Exactitud dels sondejos publicats

Ara es pretén jutjar els sondejos dels diaris respecte els teòrics. Ho farem any per any i determinarem com i quants cops els diaris han fallat en les prediccions. Al tractar-se de sondejos molt complicats, es valoraren positivament prediccions que s'apropen a les distàncies que presenten els sondejos teòrics.

Introduint el retoc en les proporcions de vot del 3% entre els dos partits principals (es comenta la seva obtenció en l'apartat 4.3) les distàncies respecte els resultats reals que presenten els sondejos teòrics han augmentat. La distància límit, que és fixada com la màxima distància possible per considerar un sondeig vàlid, és la que presenta la fita superior del interval al 95% d'aquest nou sondeig teòric. Aquesta és de 22 en el Parlament de Catalunya i de 42 en el Congrés. Si el retoc és del 6% les distàncies màximes són de 28 a les catalanes i 70 a les espanyoles.

En la taula 8.2.1 i la taula 8.2.2 hi ha tots els sondejos considerats correctes i tots els que no ho són. Els que contenen la qualificació d'excel·lent, signifiquen sondejos que van ser molt encertats i és troben dins de les distàncies que presenten els teòrics sense cap imperfecció. Els aprovats ja no són tant precisos. Malgrat tot, atès que la seva realització és molt complicada de realitzar, estarien dins les distàncies dels teòrics que s'han retocat les proporcions de vot. Aquests es podrien considerar uns sondejos perfectes on hi ha hagut un canvi de com a molt el 3% en la intenció de vot entre els dos partits principals. Els sondejos marcats com a suspesos són totalment erronis i serien uns sondejos perfectes amb un canvi en la intenció de vot dels dos partits principals de entre el 3% i el 6% (el sondeig d'El País al Parlament de Catalunya l'any 1984 seria un canvi en la intenció de vot de més del 6%).

**taula 8.2.1** Taula resum de l'encert dels sondejos en els diferents anys al Parlament de Catalunya.

Diaris	1984	1988	1992	1995	1999	2003	2006
LV	Aprovat	Excel·lent	Excel·lent	Suspès	Excel·lent	Suspès	Aprovat
EPE	*	Aprovat	*	Suspès	Suspès	Aprovat	Excel·lent
EPA	Suspès	Aprovat	Excel·lent	Suspès	Excel·lent	Aprovat	Aprovat
ABC	*	*	*	Excel·lent	Excel·lent	Suspès	Excel·lent
EMU	*	*	*	Aprovat	Excel·lent	Aprovat	Excel·lent

**taula 8.2.2** Taula resum de l'encert dels sondejos en els diferents anys al Congrés dels Diputats.

Diaris	1986	1989	1993	1996	2000	2004	2008
LV	*	Excel·lent	Aprovat	Excel·lent	Aprovat	Aprovat	Excel·lent
EPE	Aprovat	*	Aprovat	Aprovat	Aprovat	Suspès	Aprovat
EPA	*	Excel·lent	Aprovat	Suspès	Aprovat	Suspès	Excel·lent
ABC	Suspès	*	Aprovat	Suspès	Aprovat	Suspès	*
EMU	*	Aprovat	Suspès	Suspès	Suspès	Suspès	Excel·lent

### 8.2.1 Exactitud dels sondejos al Parlament de Catalunya

Estudiem detalladament les distàncies al Parlament. Si es mira la figura 8.2.1 no sembla que hi hagin diferències entre els diferents diaris.

Per quantificar numèricament les distàncies respecte els resultats reals en els diferents diaris s'hauria de realitzar una ANOVA a dos factors. No és possible atès que falten algunes dades. Aleshores la matriu del disseny no està balancejada. L'estudi més vàlid llavors, és fer una comparació de mitjanes de dades aparellades. No hi ha cap prova significativa, si bé El Mundo és el que presenta unes distàncies més baixes. La prova que resulta més a prop de la significació és la comparació entre El País i El Mundo. Amb un P-valor de 0,088 no refusem la hipòtesi nul·la d'igualtat de distàncies.

Si tornem a fixant-se amb la figura 8.2.1, s'aprecia com el primer sondeig d'El País va ser exageradament erroni. Se'n va amb molta diferència per sobre les distàncies que presenten tots els altres sondejos. Donada aquesta circumstància, s'ha cercat quin va ser el institut encarregat de realitzar tal enquesta. Es tracta del grup Metra-Seis. Com podia semblar lògic aquesta va ser l'última contractació de l'empresa en sondejos d'El País. Posteriorment, els encarregats a realitzar els pronòstics van ser Demoscopia (els següents 4 anys) i Opina (els 2 últims).

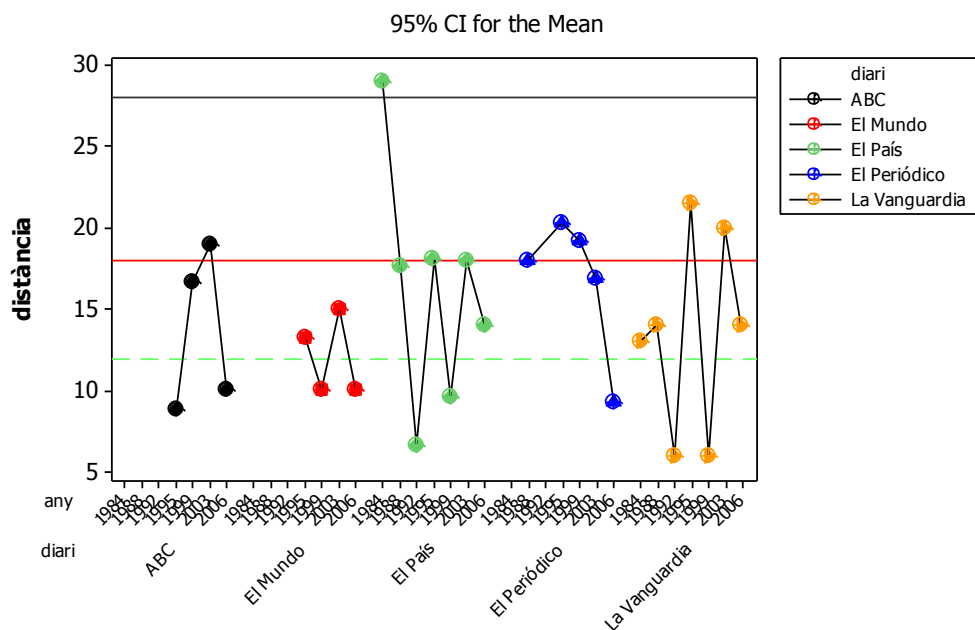


figura 8.2.1 distàncies respecte els resultats reals al llarg dels anys depenent del diari al Parlament de Catalunya.

## 8.2.2 Exactitud dels sondejos al Congrés dels Diputats

Ara anem a estudiar el que succeeix en el Congrés dels Diputats. En la següent figura (8.2.2) ja es pot intuir com l'ABC presenta distàncies molt més elevades. En El Mundo s'observa com falla en molts dels sondejos. Com s'ha indicat en la taula 8.2.2 suspèn en 4 dels 6 que presenta. El salven els altres dos, on si que es donen bons pronòstics. La Vanguardia sembla el diari amb millors prediccions. Erra únicament en un i en presenta 3 de perfectes. El País és el que presenta més variabilitat. En fa dos molt bé i dos totalment erronis. El Periódico, llevat del 2004, sempre ho fa correctament. No s'equivoca en excés, però tampoc fa sondejos perfectes.

S'ha analitzat tot de forma numèrica i la conclusió és que hi ha diaris més encertats que d'altres de forma significativa. S'identifica amb els diaris La Vanguardia i El Mundo. El primer s'aproxima molt més al que realment va succeir el dia de la celebració que no pas el segon. La prova que ho demostra es troba en la figura 8.2.3. Es pot concloure, que en mitjana, La Vanguardia encerta més els pronòstics que El Mundo, amb una diferència puntual de 14,62 diputats. Una altra prova significativa correspon a estudiar

les distàncies d'El País i les de l'ABC. El País presenta valors molt més baixos que l'ABC i amb un p-valor del 0,002 (figura 8.2.4) sembla que aconsegueix fer prediccions més bones que el diari de caràcter popular. Si es compara La Vanguardia amb l'ABC també s'intueix un grau d'encert més alt en el primer. La prova no és significativa per molt poc (és en la figura 8.2.5). L'últim test significatiu compara El Periódico i l'ABC. De nou l'ABC ha publicat els pitjors sondejos. La prova es localitza en la figura 8.2.6.

Sembla doncs que L'ABC ha estat el pitjor de tots els diaris. El Mundo també ha errat considerablement i ha estat menys encertat que La Vanguardia.

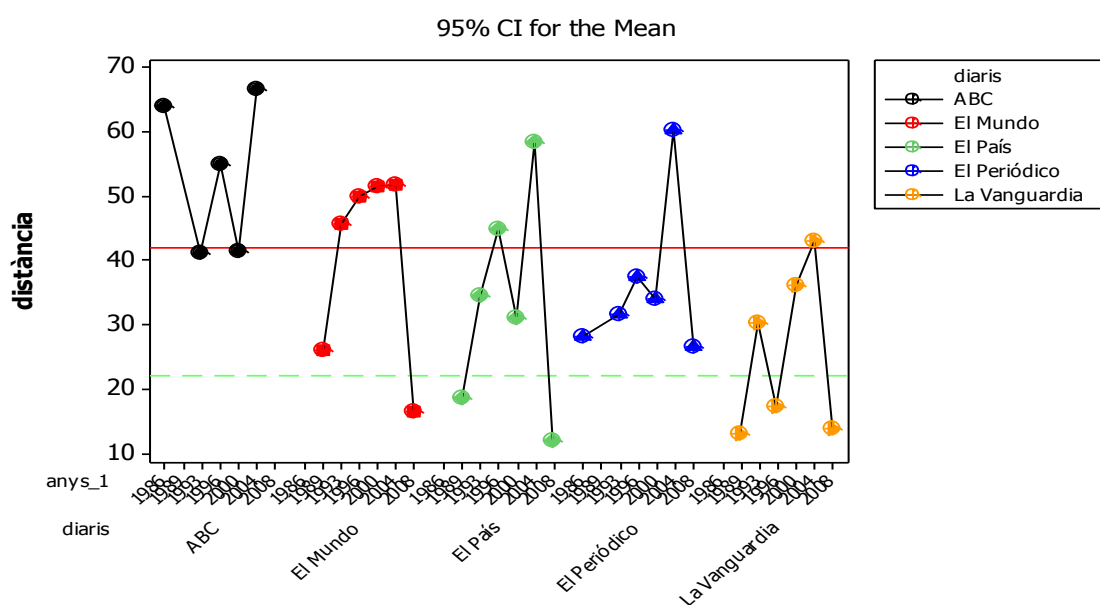


figura 8.2.2 Distàncies respecte els resultats reals al llarg dels anys depenen del diari en les eleccions generals.

Paired T for La Vanguardia - El Mundo				
	N	Mean	StDev	SE Mean
La Vanguardia	6	25,53	12,66	5,17
El Mundo	6	40,14	15,15	6,19
Difference	6	-14,62	10,10	4,12

95% CI for mean difference: (-25,22; -4,01)  
T-Test of mean difference = 0 (vs not = 0): T-Value = -3,54 P-Value = 0,016

figura 8.2.3 Sortida de minitab amb la prova de comparació de mitjanes de dades aparellades de les distàncies dels sondejos de El Mundo i La Vanguardia respecte els resultats reals en el Congrés.

Paired T for ABC - El País

	N	Mean	StDev	SE Mean
ABC	4	50,92	12,23	6,12
El País	4	42,08	12,25	6,12
Difference	4	8,831	1,817	0,908

95% CI for mean difference: (5,941; 11,722)

T-Test of mean difference = 0 (vs not = 0): T-Value = 9,72 P-Value = 0,002

**figura 8.2.4** Sortida de minitab amb la prova de comparació de mitjanes de dades aparellades de les distàncies dels sondejos de L'ABC i El País respecte els resultats reals en el Congrés.

Paired T for ABC - La Vanguardia

	N	Mean	StDev	SE Mean
ABC	4	50,92	12,23	6,12
La Vanguardia	4	31,61	10,93	5,46
Difference	4	19,31	14,39	7,20

95% CI for mean difference: (-3,59; 42,21)

T-Test of mean difference = 0 (vs not = 0): T-Value = 2,68 P-Value = 0,075

**figura 8.2.5** Sortida de minitab amb la prova de comparació de mitjanes de dades aparellades de les distàncies dels sondejos de L'ABC i La Vanguardia respecte els resultats reals en el Congrés.

Paired T for ABC - EL Periódico

	N	Mean	StDev	SE Mean
ABC	5	53,53	12,10	5,41
EL Periódico	5	38,26	12,71	5,68
Difference	5	15,28	12,37	5,53

95% CI for mean difference: (-0,08; 30,63)

T-Test of mean difference = 0 (vs not = 0): T-Value = 2,76 P-Value = 0,051

**figura 8.2.6** Sortida de minitab amb la prova de comparació de mitjanes de dades aparellades de les distàncies dels sondejos de L'ABC i El Periódico respecte els resultats reals en el Congrés.

### 8.3 Tendències polítiques dels diaris

A tots els diaris se'ls associa amb una tendència política. Tot i això, el més normal seria que en fer l'enquesta per saber la intenció de vot dels ciutadans no esbiaixessin els resultats afavorint al seu partit polític preferit. Es pot analitzar per assegurar-nos que no es produeix tal incongruència.

Una primera comprovació que es constata és com s'equivoquen els diaris en els diferents partits al llarg dels anys. Se separen els partits en dues tendències diferents. Probablement no és una separació que agradi a tothom, però el cert és que és utilitzada i coneguda per gran part de la població. Es diferencia entre dretes i esquerres. A Catalunya, el primer bloc el formen Convergència i Unió, el Partit Popular i el Centre Democràtic Social. En el segon hi tindrem el Partit Socialista, Esquerra Republicana de Catalunya, Iniciativa Per Catalunya- Els Verds i Ciutadans Per la Ciutadania. A Espanya, a la dreta hi haurà Convergència i Unió, El Partido Popular, el Partido Nacionalista Vasco i el Centro Democrático y Social. A les esquerres El Partido Socialista Obrero Español, Izquierda Unida i Esquerra Republicana de Catalunya

Error en negatiu signifiquen una infravaloració d'escons de la candidatura. Els positius indiquen sobrevaloració. Aquests valors s'aconsegueixen fent el promig de les diferències entre els escons reals i els escons que es poden deduir de totes les combinacions possibles de cada un dels sondejos. S'identifiquen com a errors per no confondre amb els valors que s'anomenen distància. El programa en R es pot consultar en l'annex 4.

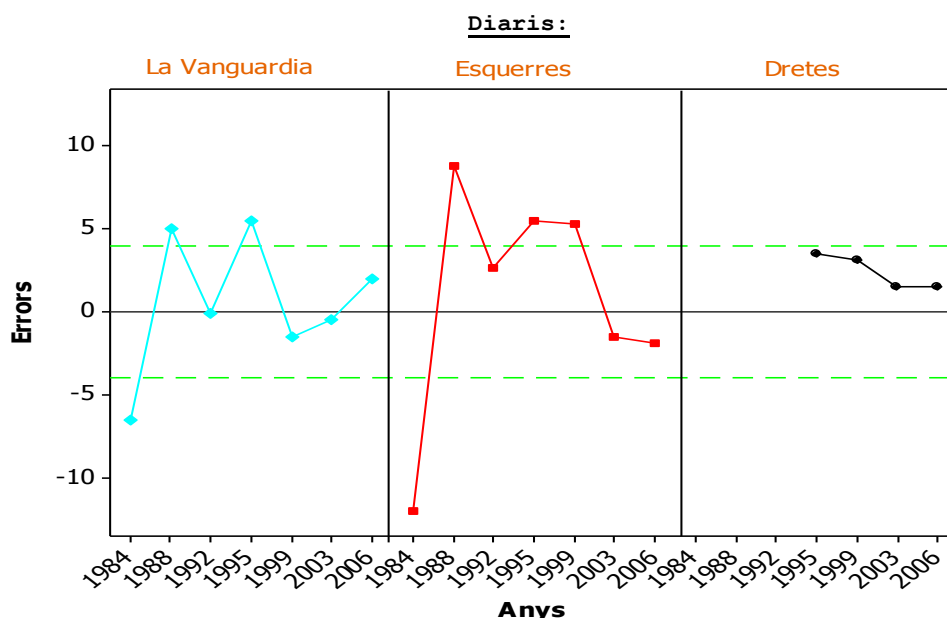
En el Parlament de Catalunya els errors es complementen entre dretes i esquerres. Se sap perquè la suma dels diputats previstos ha de ser igual al total a repartir (135). Així, si s'han pronosticat uns certs escons més del compte  $n$  per un dels dos grups, per compensar i obtenir els 135 escons desitjats, en l'altre grup ha de haver-hi una subestimació de  $-n$ . Per tant presentant les diferències d'una de les dues tendències ja n'hi ha prou. S'estudia el comportament a la dreta. En el Congrés dels Diputats, en canvi, hi ha una petita diferència. És deguda per el bloc d'escons destinats a candidatures més minoritàries que s'engloben com a *Altres*. Aquests no s'identifiquen en cap de les dues tendències i pot fer variar aquesta relació perfecte entre dretes i esquerres. Per aquest motiu es presenten els errors en les dues ideologies.

Pel que fa als diaris també es poden separar segons aquests mateixos grups. És sabut que els diaris El País i El Periódico recolzen més aviat el cantó de les esquerres. En canvi, El Mundo i l'ABC s'orienten més cap a les dretes. La Vanguardia no el sabem col·locar en cap de les dues tendències i anirà per lliure.

### 8.3.1 Tendències polítiques al Parlament de Catalunya

En la figura 8.3.1 hi ha el gràfic dels errors a mesura que s'han succeït les convocatòries en el Parlament de Catalunya separats per cada una de les tendències dels diaris. El primer any, els partits d'esquerres van ser molt sobrevalorats (tant per La Vanguardia com pels diaris d'esquerres). En els següents anys, els diaris s'inclinen més per pronosticar a l'alça pels partits de dretes. Tot i això, no es pot determinar en cap cas que els diferents grups de diaris sobreestimin alguna de les dues inclinacions de forma significativa i ni molt menys que presentin estimacions diferents entre ells.

Les línies discontinües determinen l'interval d'errors que presentarien els sondejos perfectes. Es poden apreciar anys on es va tendir a afavorir en excés a una de les dues tendències. Se'n destaquen el 1984 als partits d'esquerres i el 1988 i 1995 als de dretes.

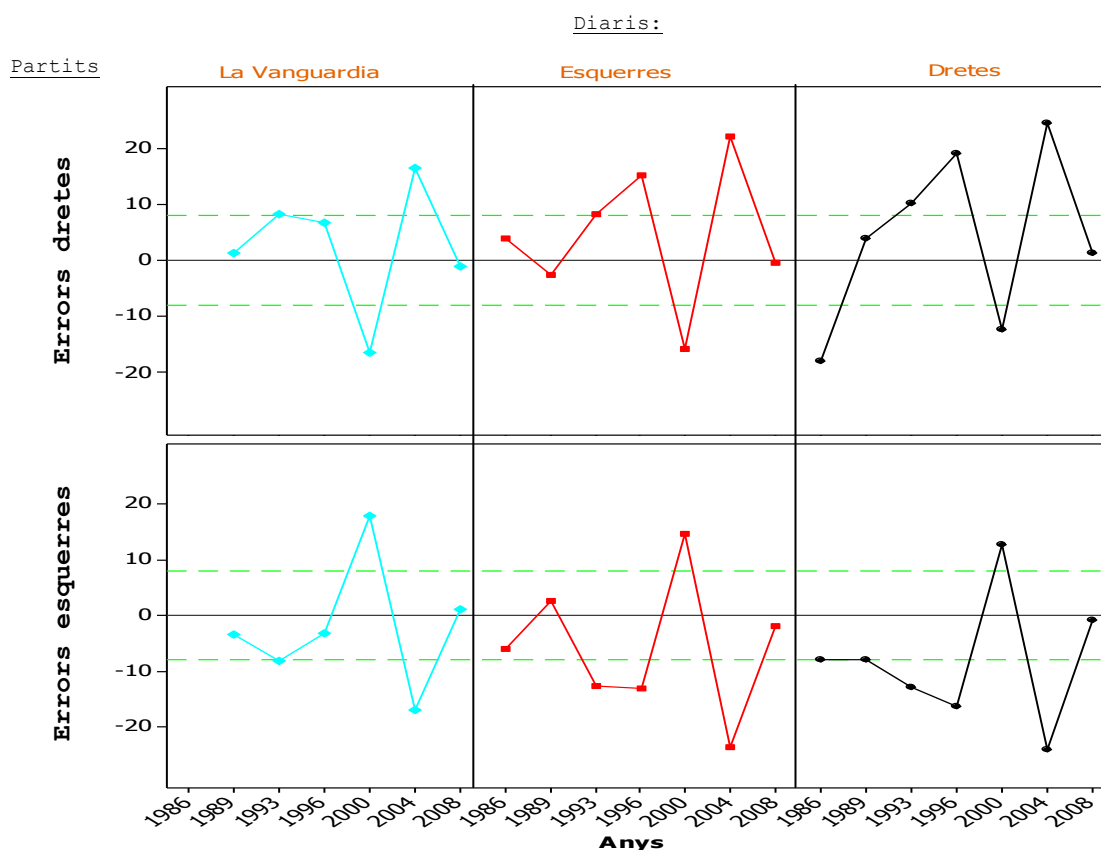


**figura 8.3.1** Errors dels sondejos en totes les convocatòries en estimar vots per els partits de dretes en el Parlament.

### 8.3.2 Tendències polítiques al Congrés dels Diputats

En la figura 8.3.2 es contempen els errors en el cas del Congrés dels Diputats. Es torna a detectar un comportament similar dels diferents diaris (els errors són principalment induïts pel PP i pel PSOE i això s'analitza amb detall en el proper apartat). A més, si s'obvia el primer any, on els diaris de dretes van subestimar molt més els partits de dretes que no pas els diaris més socialistes, sembla que l'ABC i El Mundo pronostiquin una mica més a favor als partits de dretes que no La Vanguardia i els dos diaris d'esquerres. De forma similar deduïm el mateix si es miren els errors en els partits d'esquerres.

Es comprova en la figura 8.3.3 amb un anàlisi de comparació de mitjanes de dades aparellades. Sense prendre el primer any i amb un p-valor del 0,006 podem confirmar diferències significatives en quan a l'estimació d'escons de les dretes segons si els diaris són d'una tendència o de l'altra. En l'estimació d'escons a la banda esquerrana la prova no és significativa (figura 8.3. 5) si bé en 6 de les 7 observacions els diaris de dretes han realitzat prediccions menys favorables als partits d'esquerres. Si es compara La Vanguardia amb els diaris de dretes s'hi distingeix una diferència significativa en els dos blocs. La Vanguardia doncs, tendeix a situar-se cap a la banda esquerrana en quan als pronòstics de les eleccions generals.



**figura 8.3.2** Errors dels sondejos en totes les convocatòries en estimar vots per els partits de dretes i esquerres en el Congrés dels Diputats.



Paired T for esq - dret

	N	Mean	StDev	SE Mean
esq	6	4,47	13,72	5,60
dret	6	7,85	13,28	5,42
Difference	6	-3,377	1,838	0,750

95% CI for mean difference: (-5,305; -1,448)

T-Test of mean difference = 0 (vs not = 0): T-Value = -4,50 P-Value = 0,006

**figura 8.3.3** Sortida de minitab de la prova d'hipòtesis de comparació de mitjanes (errors en els partits de dretes) de dades aparellades entre els diaris de dretes i els d'esquerres

Paired T for LV - dret

	N	Mean	StDev	SE Mean
LV	6	2,54	11,15	4,55
dret	6	7,85	13,28	5,42
Difference	6	-5,31	4,18	1,71

95% CI for mean difference: (-9,70; -0,91)

T-Test of mean difference = 0 (vs not = 0): T-Value = -3,11 P-Value = 0,027

**figura 8.3.4** Sortida de minitab de la prova d'hipòtesis de comparació de mitjanes (errors en els partits de dretes) de dades aparellades entre els diaris de dretes i La Vanguardia

Paired T for Esquerres - Dretes

	N	Mean	StDev	SE Mean
Esquerres	7	-5,77	12,47	4,71
Dretes	7	-8,23	11,82	4,47
Difference	7	2,46	3,88	1,47

95% CI for mean difference: (-1,13; 6,05)

T-Test of mean difference = 0 (vs not = 0): T-Value = 1,67 P-Value = 0,145

**figura 8.3.5** Sortida de minitab de la prova d'hipòtesis de comparació de mitjanes (errors en els partits d'esquerres) de dades aparellades entre els diaris de dretes i els d'esquerres

Paired T for La Vanguardia - Dretes				
	N	Mean	StDev	SE Mean
La Vanguardia	6	-2,18	11,64	4,75
Dretes	6	-8,27	12,95	5,29
Difference	6	6,09	3,91	1,60

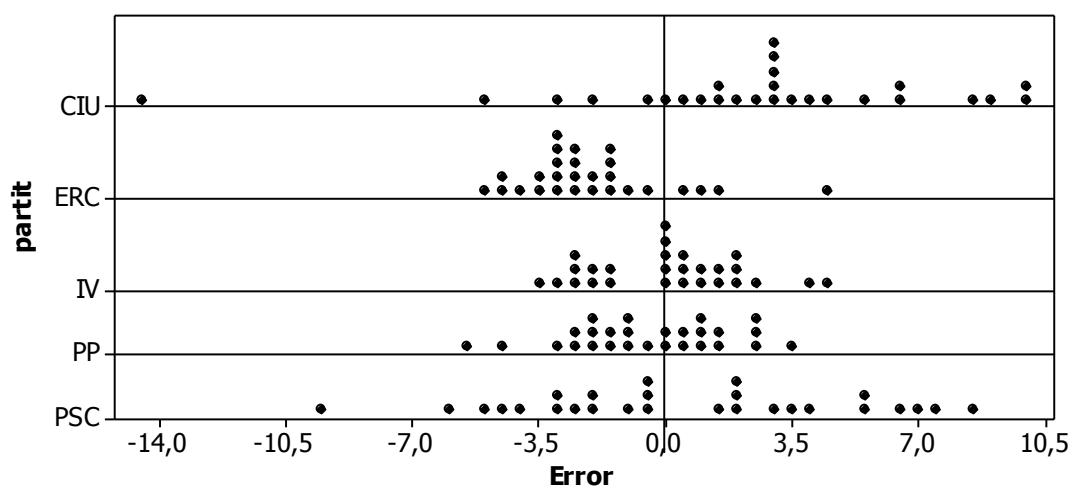
95% CI for mean difference: (1,99; 10,19)  
T-Test of mean difference = 0 (vs not = 0): T-Value = 3,81 P-Value = 0,012

**figura 8.3. 6** Sortida de minitab de la prova d'hipòtesis de comparació de mitjanes (errors en els partits d'esquerres) de dades aparellades entre els diaris de dretes i La Vanguardia.

#### 8.4 Anàlisi dels errors en les principals candidatures

En aquest apartat, es vol estudiar de quina forma han pronosticat els diaris els escons per cada partit. Es realitza l' anàlisi complet per descobrir com han errat en cada un d'ells.

Una primera comprovació és constatar si hi ha partits que són infravalorats o sobrevalorats de forma significativa. L'anàlisi gràfic és a la figura 8.4.1 (Catalunya) i a la figura 8.4.2 (Espanya). A Catalunya sembla que es tendeix a sobreestimar a CIU i infravalorar a ERC. A Espanya sobretot veiem com es sobreestima el PNV i IU. En aquest anàlisi es prenen totes les observacions, independentment del diari i la data de realització.



**figura 8.4.1** Dotplot de tots els errors que s'han produït en els diferents partits a Catalunya.

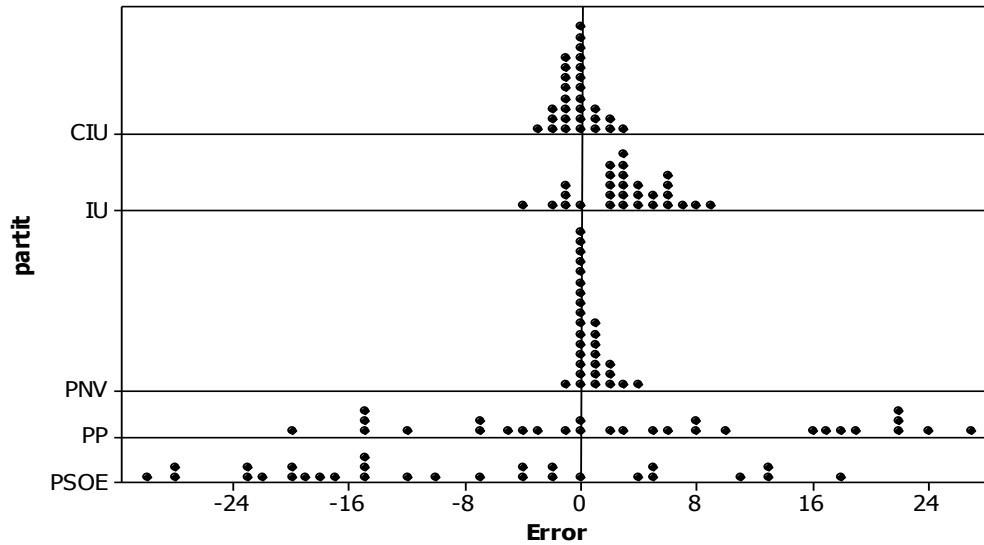


figura 8.4.2 Dotplot de tots els errors que s'han produït en els diferents partits a Espanya.

Per comprovar-ho de forma més formal es fa la prova d'hipòtesi mitjançant la t-student (estan en les següents figures, 8.4.3 a Catalunya i 8.4.4 a Espanya). Com ja s'intuïa surten dues proves significatives en el Parlament. La primera és una sobreestimació a Convergència i Unió. Té un p-valor de 0,017. La segona és una subestimació a Esquerra Republicana de Catalunya. El P-valor és de 0,000 i ens confirma la tendència a pronosticar a la baixa per part dels sondejos al Partit Independentista.

En el Congrés hi ha 3 proves significatives. La subestimació del Partido Socialista amb un p-valor del 0,017, la sobrevaloració envers Izquierda Unida amb un p-valor de 0,000 i la també sobrevaloració al PNV amb un p-valor de 0,003. La prova referent al Partido Popular ratlla la significació. El P-valor és de 0,098 i amb un risc del 10% es podria rebutjar la hipòtesis nul·la de mitjana d'errors igual a 0. El PP sol rebre una sobreestimació d'escons.

```

One-Sample T: CIU; PSC; PP; ERC; IC-IV

Test of mu = 0 vs not = 0

Variable   N    Mean   StDev   SE Mean   95% CI          T      P
CIU        27    2,480   5,034   0,969   ( 0,489;  4,472)  2,56  0,017
PSC        27    0,536   4,588   0,883   (-1,279;  2,351)  0,61  0,549
PP         27   -0,470   2,206   0,424   (-1,342;  0,403) -1,11  0,278
ERC        27   -2,003   2,051   0,395   (-2,815; -1,192) -5,08  0,000
IC-IV     27    0,099   2,086   0,401   (-0,726;  0,924)  0,25  0,808

```

figura 8.4.3 Sortida de minitab de la prova de la t-student per comprovar si la mitjana d'errors a de cada partits és igual o diferent de 0 en Catalunya.

One-Sample T: PSOE; PP; CIU; IU; PNV							
Test of mu = 0 vs not = 0							
Variable	N	Mean	StDev	SE Mean	95% CI	T	P
PSOE	29	-9,16	13,69	2,54	(-14,36; -3,95)	-3,60	0,001
PP	29	4,27	13,45	2,50	(-0,84; 9,39)	1,71	0,098
CIU	29	-0,293	1,265	0,235	(-0,774; 0,188)	-1,25	0,223
IU	29	2,962	3,026	0,562	( 1,811; 4,113)	5,27	0,000
PNV	29	0,587	0,982	0,182	( 0,214; 0,960)	3,22	0,003

**figura 8.4.4** Sortida de minitab de la prova de la t-student per comprovar si la mitjana d'errors a de cada partits és igual o diferent de 0 a Espanya

Procedim a analitzar si hi ha diferències entre els diferents diaris en els errors de cada un dels 5 partits principals, tant a Espanya com a Catalunya. No es pot incloure els casos de CDS i CPC en les eleccions catalanes i ERC i el CDS en les espanyoles per falta de dades.

Mitjançant un anàlisi gràfic ja es poden distingir patrons, diferències entre diaris i com s'allunyen del errors del sondeig perfecte (es marca el interval en les figures de forma discontinua).

A més, s'ha realitzat un anàlisi numèric per via comparació de dades aparellades al ser important la data de celebració. Únicament es presenten els tests on hi ha alguna diferència significativa. Les taules on hi consten tots els errors, tant del Parlament com del Congrés, dels diaris estan en l'annex 5.

#### ❖ Cas CIU:

No hi ha diferències significatives en els errors d'estimació d'escons dels diferents diaris en aquest partit en cap de les dos tipus d'eleccions. Mirant la figura 8.4.5, es poden deduir errors comuns en els partits en cada convocatòria. És un fet curiós. Ens indica que totes les enquestes (teòricament independents) estan encaminades en la mateixa direcció. És rellevant ja que no tenim constància de cap succés en la setmana d'impàs entre sondejos i celebracions que pogués justificar aquests errors comuns en el Parlament. Podria semblar que es copiessin els uns als altres. Els sondejos referents als anys 1992, 1999, 2003 i 2006 s'aproximen molt als resultats reals i l'interval del sondeig perfecte els engloba. En 1984 varen subestimar excessivament als Convergents, en canvi 1988 i 1995 va ser sobreestimat per sobre el que presenta el sondeig perfecte.

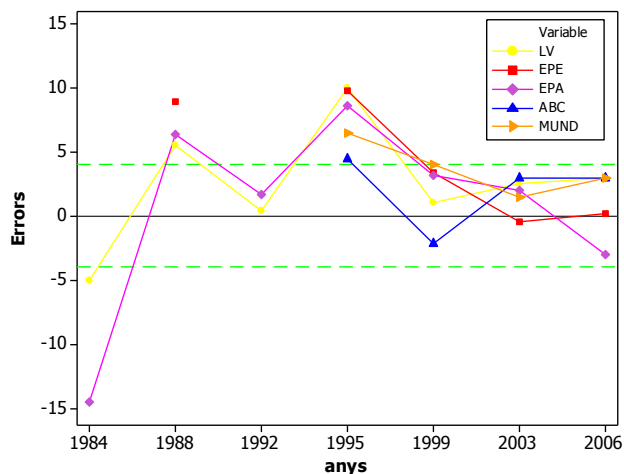


figura 8.4.5 Errors dels sondejos en l'estimació d'escons de Convergència I Unió a Catalunya.

A les eleccions generals també s'aprecia aquesta situació d'errors comuns. Als anys on menys bé es pronostiquen els escons per CIU són el 1993 i el 1996. Tot i això, en tots ells s'ajusten molt bé i tenen errors similars als sondejos perfectes.

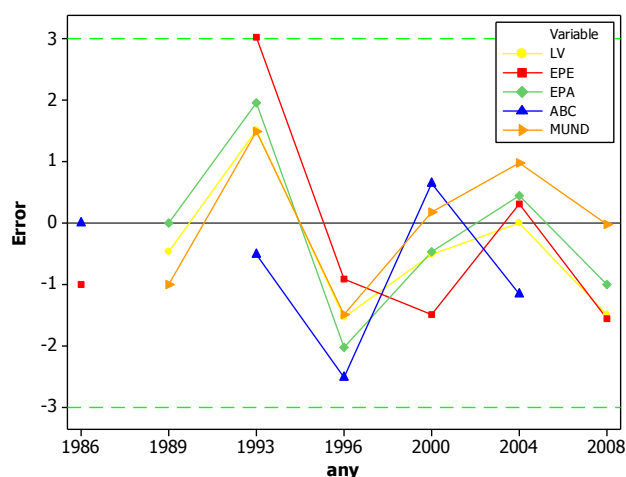


figura 8.4.6 Errors dels sondejos en l'estimació d'escons de Convergència I Unió a Espanya.

#### ❖ Cas PSC/PSOE:

Es confirma que cap diari pronostica diferent els escons pels socialistes en el Parlament de Catalunya. Amb el gràfic (figura 8.4.7) ja n'hi ha prou per veure com tots els sondejos tenen el mateix patró d'errors al llarg dels anys. En les 4 primeres convocatòries els errors en el Partit Socialista no van ser molt grans i es situen a dins l'interval dels perfectes. En canvi, les dues següents (sobretot la referent al 2003), hi ha diaris que fallen massa en fer els pronòstics al partit d'esquerres. En 1999 el varen subestimar (queden fora de l'interval El Periódico i l'ABC) i en 2003 tots els diaris el

van sobreestimar. En 2006 també s'inclinen a valorar més positivament del compte el partit. Tot i això, únicament El País surt de les bandes de confiança.

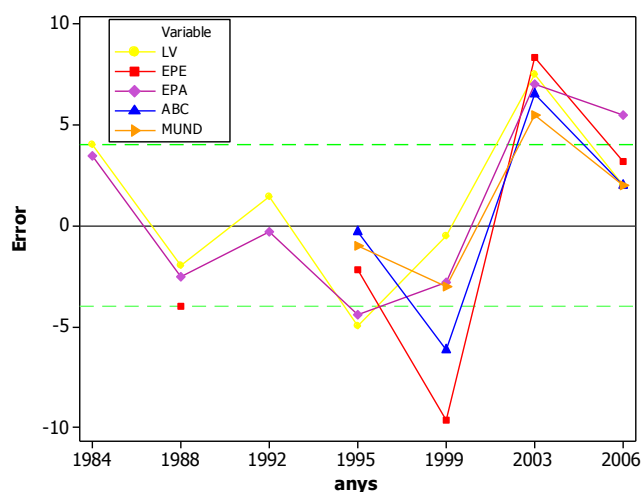


figura 8.4.7 Errors dels sondejos en l'estimació d'escons del Partit Socialista a Catalunya.

En les eleccions generals tampoc hi ha cap diari específic que potencí significativament més o menys als socialistes. En la figura 8.4.8 s'hi destaquen els errors (comuns segons l'any). El 1993, 1996 i 2004 hi va haver-hi una infravaloració excessiva i el 2000 es va sobrevalorar massa al partit d'esqueres.

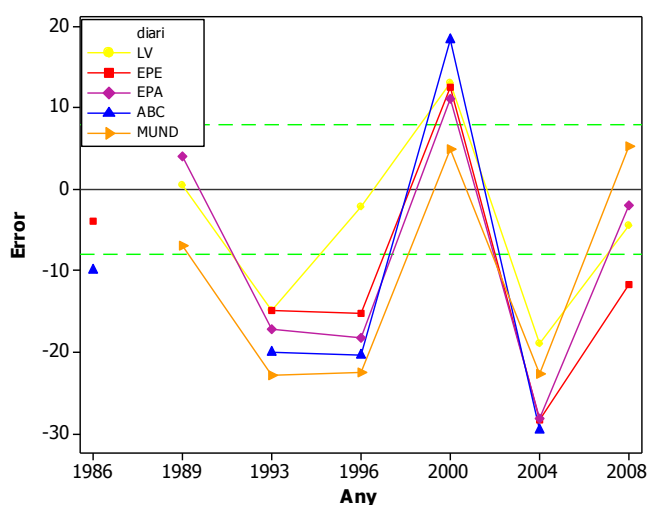
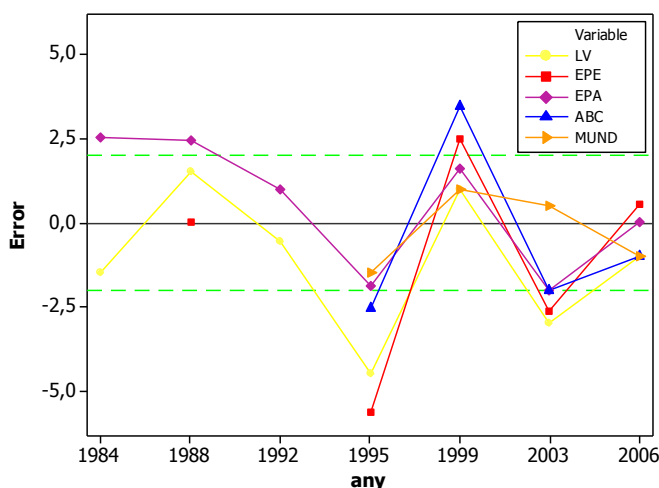


figura 8.4.8 Errors dels sondejos en l'estimació d'escons del Partido Socialista a Espanya.

❖ **Cas PP:**

En el Partit Popular tampoc hi ha grans diferències entre diaris en el Parlament. El que presenta errors més positius és El País (que és conegut per ser contrari a la ideologia del PP). Tot i això, no són suficients per considerar-lo significativament diferent.

Els dos únics anys que s'aproximen bé a la realitat tots els diaris són el 1992 i el 2006. En les altres convocatòries sempre, un o altre diari, ha sobrevalorat o infravalorat en excés als populars. Destaquen negativament, els errors de La Vanguardia i El Periódico en l'any 1995 on van subestimar molt a la candidatura. Mirar la figura 8.4.9.



**figura 8.4.9** Errors dels sondejos en l'estimació d'escons del Partit Popular en Catalunya.

En canvi en les eleccions generals és El País qui estima més a la baixa els escons del Partido Popular (mirar figura 8.4.10). Presenta diferències significatives respecte el diari El Mundo (possiblement el diari més partidari dels populars). La prova es troba en la figura 8.4.11. A més cal destacar, i com és lògic, que es va tendir a sobreestimar al partit els anys on es va subestimar el PSOE. Al 1996 i el 2004. També és dóna quan s'ha subestimat el partit de dretes (2000). Aleshores el Partido Socialista ha estat sobrevalorat.

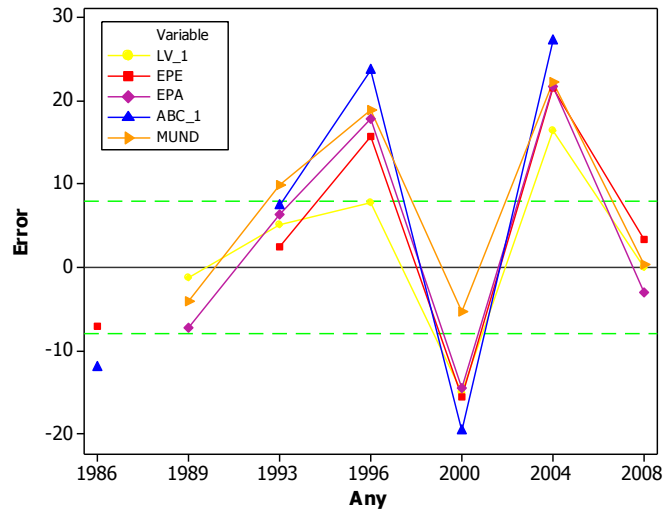


figura 8.4.10 Errors dels sondejos en l'estimació d'escons del Partido Popular a Espanya.

Paired T for MUND - EPA				
	N	Mean	StDev	SE Mean
MUND	6	6,99	11,82	4,83
EPA	6	3,56	14,35	5,86
Difference	6	3,43	3,09	1,26

95% CI for mean difference: (0,19; 6,67)  
 T-Test of mean difference = 0 (vs not = 0): T-Value = 2,72 P-Value = 0,042

figura 8.4.11 Prova de mitjanes amb dades aparellades entre El País i El Mundo en el Partit Popular



❖ **Cas ERC:**

A l' Esquerra Republicana de Catalunya tampoc s'hi mostren diferències en els diferents diaris. Les proves que hem realitzat, surten amb un p-valor molt elevat. Ens indica que els pronòstics dels diaris són molt semblants. En tots es sol subestimar els escons pel partit catalanista. Sobretot se'n destaca el 2003. Tots els diaris van errar excessivament a la baixa els escons del partit d'esqueres. En els altres anys els errors no s'allunyen dels que presenta el sondeig perfecte. Mirar la figura 8.4.12.

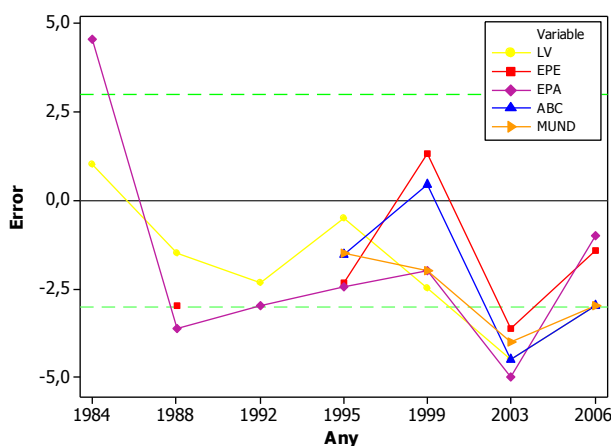


figura 8.4.12 Errors dels sondejos en l'estimació d'escons de Esquerra Republicana de Catalunya.

❖ **Cas IC-IV (PSUC)/IU:**

A Catalunya tampoc s'hi aprecia cap diferència en els diaris. La figura 8.4.13 ens descobreix com el comportament dels errors de predicció pels Verds és comú a tots els diaris. El que fa pronòstics més a la baixa és El Mundo. Tot i això no són significativament inferiors als altres. La majoria de les prediccions encerten força els escons que van aconseguir, si bé en 1999 i 2003 hi va haver-hi diaris amb errors excessivament grans (sobretot l'ABC en el 99).

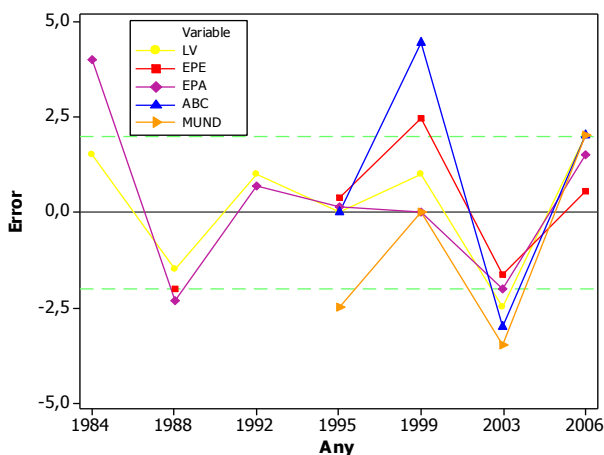


figura 8.4.13 Errors dels sondejos en l'estimació d'escons de Iniciativa Per Catalunya- Els Veds a Catalunya.

A Espanya succeeix el mateix. Tots els diaris tendeixen a sobreestimar al partit d'esquerres. Són molts els anys on aquesta valoració a l'alça és excessiva. Sobretot se'n destaquen els sondejos de 1993.

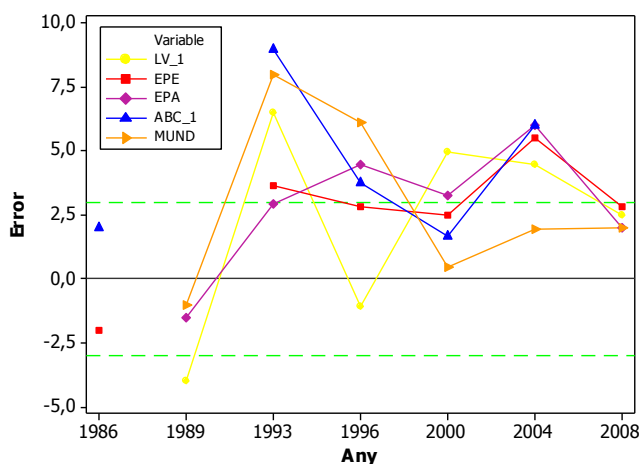


figura 8.4.14 Errors dels sondejos en l'estimació d'escons de Izquierda Unida a Espanya .

#### ❖ Cas PNV:

L'últim dels partits polítics que estudiem, el PNV, no presenta diferències entre els diferents diaris. En tots s'intueix una tendència a sobreestimar el partit. Tot i això, els errors no són gaire grans i podem considerar-los fruit de l'atzar.

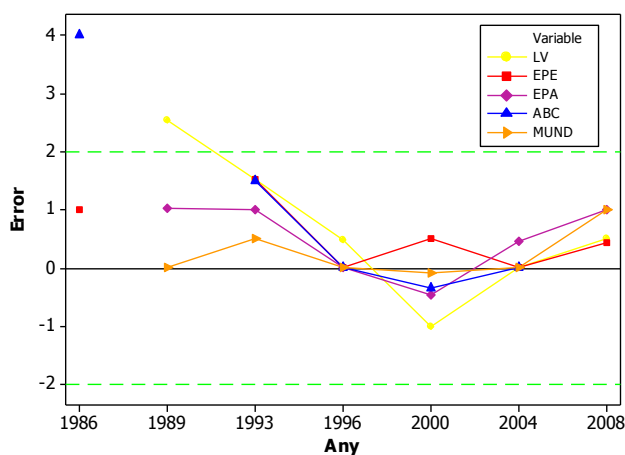


figura 8.4.15 Errors dels sondejos en l'estimació d'escons del Partido Nacionalista Vasco.

## 9 CONCLUSIONS

---

Després d'haver estat analitzant a fons els sondejos que presenten els diaris a una setmana vista de les eleccions, tant del Parlament de Catalunya com del Congrés dels Diputats, es poden donar unes conclusions del que s'ha estudiat.

Un primer aspecte que ha quedat clar és la dificultat de realitzar un bon sondeig. Els recursos econòmics que tenen els diaris són limitats i fan inviable l'elaboració d'una enquesta que representi de forma adequada la intenció de vot dels ciutadans. Fan prediccions de poblacions de l'ordre de milions de persones amb únicament mostres de l'ordre de mil persones. A més no distribueixen aquests individus enquestats de forma aleatòria al 100% .

En les dues eleccions que s'han estudiat es presenten sondejos bons i altres de no tant. Tot i això no es produeixen distàncies massa grans de forma sistemàtica. En les del Parlament de Catalunya sobretot hi ha dues convocatòries on els sondejos s'equivoquen considerablement. La de 1995 i 2003 (erren significativament 3 i 2 dels 5 sondejos respectivament). Les millors són les de 1992 i 2006. En el Congrés dels Diputats, curiosament, s'han produït els pitjor pronòstics en els anys posteriors als pitjors de les eleccions catalanes. El 1996 i el 2004 (erren significativament 3 i 4 dels 5 sondejos respectivament). Els millors són el 1989 i el 2008. Ni en les eleccions generals ni en les catalanes s'aprecia una tendència a fer pronòstics més encertats a mesura que s'han succeït les celebracions. Es podria continuar analitzant en propers convocatòries, però tot sembla indicar que és complicat millorar en aquest aspecte.

Els diaris, quan fallen, ho fan més o menys tots i en la mateixa direcció. D'aquí es poden deduir uns patrons comuns que no serien lògics de no ser pel fet que la intenció de vot ha canviat en una setmana. Aquest sembla el motiu més indicat si bé també es podria plantejar l'opció que els sondejos que realitzen les empreses subcontractades es copiïn entre elles. Tot i això sembla més raonable pensar que es dona la primera circumstància.

Mentre al Parlament de Catalunya, no es troben diferències significatives de distàncies entre els diferents diaris, en el Congrés dels Diputats si n'hi ha que ho fan millor que d'altres. El millor en aquestes eleccions és La Vanguardia i presenta distàncies significativament menors a El Mundo i l'ABC. Aquest últim és el pitjor de tots i també s'equivoca significativament més que El País i El Periódico. Entre El País, El Periódico i La Vanguardia les diferències són menyspreables, al igual que els dos primers i El Mundo.

Una menció especial al primer pronòstic d'El País en El Parlament de Catalunya on va presentar un sondeig molt per sobre de les distàncies habituals. En aquest la empresa encarregada que varen subcontractar era Metra-Seis. En la següent convocatòria el diari

ja va prescindir del servei d'aquesta empresa i va adjudicar la feina a una altra (Demoscopia).

Al comparar les distàncies de les dues eleccions intuïm el fet que a Catalunya (11/27) es solen fer bastants més pronòstics perfectes que Espanya (6/29). La relació dels totalment erronis amb els correctes no és tant clara. 7/27 a Catalunya i 10/29 a Espanya. En definitiva, sembla que s'ha aconseguit descobrir amb més exactitud el que passarà en les eleccions del Parlament, que no a les del Congrés.

S'ha analitzat les tendències a afavorir a una certa ideologia política, separant en dos grups els diaris i els partits. Dretes (PP, CIU, PNV, CDS / El Mundo, ABC) i esquerres (PSC/PSOE, ERC, IU/IV, CPC / El País, El Periódico). La Vanguardia no s'ha sabut col·locar a cap dels dos llocs i va per lliure. En el Parlament no s'hi distingeix cap diferència per afavorir una de les dues tendències per part dels diaris. En el Congrés, si obviem el primer any (on passa el contrari) els diaris de dretes sobreestimen significativament més que els diaris d'esquerres i La Vanguardia el conjunt de partits de orientació de dretes. La Vanguardia doncs, sembla que està més a prop de les idees dels diaris de tendència socialista.

A partir de l'observació específica de cada un dels partits, a Catalunya s'identifica que CIU es sobreestima significativament en mitjana per tots els diaris entre 0,49 i 4,47 punts. En canvi ERC és subestimat entre -2,82 i -1,19. Pel que fa a Espanya, el PSOE és subestimat entre -14,36 i -3,95 punts en mitjana, IU és sobreestimat entre 1,81 i 4,11 i el PNV és també sobreestimat entre 0,21 i 0,96 punts. Entre els diaris només s'hi aprecia una diferència significativa. Fa referència als pronòstics del Partido Popular a les eleccions generals, El País tendeix a sobrevalorar menys el PP de forma significativa que no El Mundo. Fet que pot semblar lògic per les ideologies contràries que defensen, però incoherent atès que les mostres que realitzen haurien de ser aleatòries i independents de les candidatures.

## 10 BIBLIOGRAFIA

---

### Llibres o textos per a revistes:

- Delicado, Pedro & Udina, Federic (2001). *¿cómo y cuánto fallan los sondeos electorales?*. Reis. 96/01 pp:123-150.
- Box, George E.P. Hunter, William G. Hunter, Stuart J. (1999). *Estadística para investigadores*. Versió traduïda a l'Espanyol. Editorial Reverté, S.A.
- Escofiter B. & Pagès J. *Analyses factorielles simples et multiples* (1992). Versió traduïda al servei de la universitat del País Basc.

### Pàgines Web o textos electrònics:

- Preelectoral eleccions generales y autonòmiques de Andalusia 2004 [en línea]. <[http://www.cis.es/cis/opencm/CA/1\\_encuestas/estudios/listaTematico.jsp?tema=230&todos=no](http://www.cis.es/cis/opencm/CA/1_encuestas/estudios/listaTematico.jsp?tema=230&todos=no)>. [consulta: 6 de juny de 2010].
- Base històrica dels resultats electorals (Congrés) [en línea]. <<http://www.elecciones.mir.es/MIR/jsp/resultados/index.htm>>. [consulta: 6 de juny de 2010].
- Dades electorals de totes les convocatòries (Parlament) [en línea]. <<http://www20.gencat.cat/portal/site/governacio/...>>. [consulta: 4 de juny de 2010]
- El repartiment d'escons: aplicació de la llei de Hondt [en línea]. <<http://www20.gencat.cat/portal/site/governacio/...>>. [consulta: 20 de maig de 2010]

### Pàgines web dels diaris on s'han cercat els sondejos:

- [www.lavanguardia.es](http://www.lavanguardia.es)
- [www.Elperiodico.es](http://www.Elperiodico.es)
- [www.Elpais.com](http://www.Elpais.com)
- [www.abc.es](http://www.abc.es)
- [www.elmundo.es](http://www.elmundo.es)

## ANNEXES

---

### Annex 1: PROGRAMA PER ACONSEGUIR TOTES LES COMBINACIONS DE SONDEJOS POSSIBLES QUE PRESENTEN ELS DIARIS (CAS CONGRES DELS DIPUTATS)

```
## intervals que han proposat els diaris ##
```

```
epsoea =199
epsoeb = 178
eppa =91
eppb =71
eciua=21
eciub=20
eiua=12
eiub=7
epnva=11
epnvb=9
ealta=21
ealtb=13
eerca=28
eercb=17
```

```
## número total de permutacions que es donaran ##
```

```
numper=(epsoea-epsoeb+1)*(eppa-eppb+1)*(eciua-eciub+1)*(epnva-epnvb+1)*(ealta-
ealtb+1)*(eerca-eercb+1)
```

```
## aconseguiment de les permutacions ##
```

```
l=1
```

```
permu = matrix(ncol=8,nrow= numper)
```

```
for (i in epsoeb:epsoea){
  for (j in eppb:eppa){
    for (k in eciub:eciua){
      for (p in eiub:eiua){
        for (s in epnvb:epnva){
          for (r in ealtb:ealta){
            for (t in eercb:eerca){

              if (sum(i,j,k,p,s,r,t)== 350){
                permu[l,]=c(i,j,k,p,s,r,t,sum(i,j,k,p,s,r,t))
                l=l+1
              }

              else { l=l
                    }
            }
          }
        }
      }
    }
  }
}
permu[1:l,]
```

**Annex2: PROGRAMA PER ACONSEGUIR ELS 10.000 SONDEJOS TEÒRICS AL PARLAMENT DE CATALUNYA (es destaquen trossos utilitzats per trobar els tamany de mostra adequats a cada circumscripció(estan en cursiva))**

```

# diferents combinacions de mides mostrals #

inttotal=matrix(0,ncol=10,nrow=3)
int=0
mides=matrix(0,ncol=4,nrow=3)
mides[1,1]= 1650
mides[1,2]=450
mides[1,3]=450
mides[1,4]=450
mides[2,1]= 1800
mides[2,2]=400
mides[2,3]=400
mides[2,4]=400
mides[3,1]= 1950
mides[3,2]=350
mides[3,3]=350
mides[3,4]=350

for (proves in 1:10){
  for (mid in 1:3){

    #numero de enquestats: quantitat d'individus que formen la mostra en les
    diferents diputacions.
    nmost = 10000
    nmostb = mides[mid,1]
    nmostg = mides[mid,2]
    nmostl = mides[mid,3]
    nmostt = mides[mid,4]

    #numero de enquestats: quantitat d'individus que formen la mostra en les
    diferents diputacions.

    nmost = 10000
    nmostb = 1800
    nmostg = 400
    nmostl = 400
    nmostt = 400

    #proporcions de partits: proporcions de vots de cada partit el dia de les
    eleccions segons la provincia.

    absb = 0.4406
    PSCb = 0.279
    CIUb = 0.2989
    ERCb = 0.1261
    PPb = 0.1116
    IVb = 0.1038
    ALTRESb = 0.0353
    rest_blancsb = 1-(PSCb+CIUb+ERCb+PPb+IVb+ALTRESb)

    absg = 0.4287
    PSCg = 0.2213
    CIUg = 0.3819
    ERCg = 0.1923
    PPg = 0.0722
    IVg = 0.0764
    ALTRESg = 0.0094
    rest_blancsg = 1-(PSCg+CIUg+ERCg+PPg+IVg+ALTRESg)

    absl = 0.41
    PSCl = 0.2199
    CIUl = 0.3999
    ERCl = 0.1772
    PPl = 0.0911
    IVl = 0.0659
    ALTRESl = 0.0097
    rest_blancsl = 1-(PSCl+CIUl+ERCl+PPl+IVl+ALTRESl)
    abst = 0.4587
  }
}

```

```

PSCt = 0.2598
CIUt = 0.3243
ERct = 0.1762
PPt = 0.11
IVt = 0.0651
ALTRESt = 0.0242
rest_blancst = 1-(PSCt+CIUt+ERct+PPt+IVt+ALTRESt)

```

**##### creació de matrius de dades (prporcions) #####**

```

barcelona=matrix(ncol=nmost,nrow=8)
barcelona2=matrix(ncol=nmost,nrow=7)
girona=matrix(ncol=nmost,nrow=8)
girona2=matrix(ncol=nmost,nrow=7)
lleida=matrix(ncol=nmost,nrow=8)
lleida2=matrix(ncol=nmost,nrow=7)
tarragona=matrix(ncol=nmost,nrow=8)
tarragona2=matrix(ncol=nmost,nrow=7)
prop_total= matrix(ncol=nmost,nrow=8)
prop_total2= matrix(ncol=nmost,nrow=6)

```

**##### proporcions de vots amb mostres de 3000 individus: creació de les proporcions de vot dels diferents partits en totes les diputacions #####**

```

### (considerem el llindar establert per la llei de hondt per tal de poder optar a escons: és de 0,03) ###

```

**##Barcelona##**

```

for (i in 1:nmost) {
  barcelona[,i]= rmultinom(1, size = nmostb, prob=c((1-absb)*PSCb,(1-absb)*CIUb,(1-
absb)*ERCb,(1-absb)*PPb,(1-absb)*IVb,(1-absb)*ALTRESb,(1-absb)*rest_blancsb,absb))
  barcelona[,i]=barcelona[,i]/nmostb
  barcelona[,i]=barcelona[,i]/sum(barcelona[1:7,i])
  for (j in 1:6){
    if (barcelona[j,i]>(0.03)){
      barcelona2[j,i]=barcelona[j,i]
    }
    else {barcelona2[j,i]=0}
  }
}

```

**##Girona##**

```

for (i in 1:nmost) {
  girona[,i]= rmultinom(1, size = nmostg, prob=c((1-absg)*PSCg,(1-absg)*CIUg,(1-
absg)*ERCg,(1-absg)*PPg,(1-absg)*IVg,(1-absg)*ALTRESg,(1-absg)*rest_blancsg,absg))
  girona[,i]=girona[,i]/nmostg
  girona[,i]=girona[,i]/sum(girona[1:7,i])
  for (j in 1:6){
    if (girona[j,i]>(0.03)){
      girona2[j,i]=girona[j,i]
    }
    else {girona2[j,i]=0}
  }
}

```

**## Lleida##**

```

for (i in 1:nmost) {
  lleida[,i]= rmultinom(1, size = nmostl , prob=c((1-absl)*PSCl,(1-absl)*CIUl,(1-
absl)*ERCl,(1-absl)*PPl,(1-absl)*IVl,(1-absl)*ALTRESl,(1-absl)*rest_blancsl,absl))
  lleida[,i]=lleida[,i]/nmostl
  lleida[,i]=lleida[,i]/sum(lleida[1:7,i])
  for (j in 1:6){
    if (lleida[j,i]>(0.03)){
      lleida2[j,i]=lleida[j,i]
    }
    else {lleida2[j,i]=0}
  }
}

```

**## Tarragona##**



```

for (i in 1:nmost) {
  tarragona[,i]= rmultinom(1, size = nmostt, prob=c((1-abst)*PSct, (1-abst)*CIUt, (1-
abst)*ERCt, (1-abst)*PPt, (1-abst)*IVt, (1-abst)*ALTRESt, (1-abst)*rest_blancst, abst))
  tarragona[,i]=tarragona[,i]/nmostt
  tarragona[,i]=tarragona[,i]/sum(tarragona[1:7,i])
  for (j in 1:6){
    if (tarragona[j,i]>(0.03*(1-abst))){
      tarragona2[j,i]=tarragona[j,i]
    }
    else {tarragona2[j,i]=0}
  }
}

```

**### proporcions de vot totals: es multipliquen les proporcions aconseguides en cada diputació amb els vots vàlids totals que s'han donat a cada circumscripció. ###**

```

for (i in 1:nmost) {
  prop_total[,i]=
(barcelona[,i]*2178561+girona[,i]*268993+lleida[,i]*178157+tarragona[,i]*282579)/2908290
}
for (i in 1:6){
  prop_total2[i,]=sort(prop_total[i,]) ## ordenació de les proporcions
}

```

**##### repartició d'escons segons la mostra ( llei hondt) en cada província i el total###**

**#escons a Barcelona#**

```

escb=matrix(0,ncol=6,nrow=nmost)

partit=0
totesc=85

for ( i in 1:nmost){

  while (totesc > 0){
    partit=which.max(barcelona2[,i]) ##aconsegum la proporció més alta
    escb[i,partit]=escb[i,partit]+1
    barcelona2[partit,i]=barcelona[partit,i]/(escb[i,partit]+1)
    totesc=totesc-1
  }

  totesc=85
}

```

**#escons a Girona#**

```

escg=matrix(0,ncol=6,nrow=nmost)
partit=0
totesc=17

for ( i in 1:nmost){

  while (totesc > 0){
    partit=which.max(girona2[,i])
    escg[i,partit]=escg[i,partit]+1
    girona2[partit,i]=girona[partit,i]/(escg[i,partit]+1)
    totesc=totesc-1
  }

  totesc=17
}

```

**#escons a lleida#**

```

escl=matrix(0,ncol=6,nrow=nmost)

partit=0
totesc=15

for ( i in 1:nmost){

```

```

while (totesc > 0){
  partit=which.max(lleida2[,i])
  escl[i,partit]=escl[i,partit]+1
  lleida2[partit,i]=lleida[partit,i]/(escl[i,partit]+1)
  totesc=totesc-1
}
totesc=15
}

#escons a tarragona#

esct=matrix(0,ncol=6,nrow=nmost)

partit=0
totesc=18

for ( i in 1:nmost){

  while (totesc > 0){
    partit=which.max(tarragona2[,i])
    esct[i,partit]=esct[i,partit]+1
    tarragona2[partit,i]=tarragona[partit,i]/(esct[i,partit]+1)
    totesc=totesc-1
  }
  totesc=18
}

## escons totals ##

esctotal=matrix(0,ncol=6,nrow=nmost)
for ( i in 1:nmost){
  for (j in 1:6){
    esctotal[i,j]=escl[i,j]+escl[i,j]+escl[i,j]+esct[i,j]
  }
}

## escons totals ordenats per cada partit ( llavors les mostres 250 i 9750 determinaran un interval al 95%)

esctotal2=matrix(0,ncol=6,nrow=nmost)

for (i in 1:6){
  esctotal2[,i]=sort(esctotal[,i])
}
mitjmed=matrix(ncol=2,nrow=6)
for (i in 1:6){
  mitjmed[i,1]=mean(esctotal[,i])
  mitjmed[i,2]=(esctotal2[499,i]+esctotal2[500,i])/2
}
desviacions=matrix(ncol=6,nrow=6)
desviacions=var(esctotal)

prop_total2[,250]
prop_total2[,9750]
esctotal2[250,]
esctotal2[9750,]

## suma per establir quina combinació de mides és la millor ##

for (i in 1:6){
  int = int+esctotal2[9750,i]-esctotal2[250,i]
}
(inttotal[mid,proves]= int )
int=0
}

```

### Annex 3: PROGRAMA QUE RECULL LES DISTÀNCIES ENTRE ELS SONDEJOS I ELS RESULTATS REALS EN EL CONGRÉS DELS DIPUTATS L'ANY 2004

```

#### distancia absoluta dels sondejors amb els resultats reals####

sondejors=read.table("sond04.txt") #obtenció dels resultats dels sondejors
reals= read.table("esp04res.txt") #obtenció dels resultats reals
mat_sr= matrix(0,nrow=38) #matriu on hi guardem les distàncies

## La Vanguardia
for (i in 1:38){
  mat_sr[i]=sum(abs(sondejors[i,]-reals[]))
}
(dist_lv=sum(mat_sr)/38) ## distància promig de La Vanguardia

## El Periódico
mat_sr= matrix(0,nrow=43)
for (i in 39:81){
  mat_sr[i-38]=sum(abs(sondejors[i,]-reals[]))
}
(dist_epe=sum(mat_sr)/43) ## distància promig de El Periódico

## El País
mat_sr= matrix(0,nrow=35)
for (i in 82:116){
  mat_sr[i-81]=sum(abs(sondejors[i,]-reals[]))
}
(dist_epa=sum(mat_sr)/35) ## distància promig de El País

## ABC
mat_sr= matrix(0,nrow=20)
for (i in 117:136){
  mat_sr[i-116]=sum(abs(sondejors[i,]-reals[]))
}
(dist_abc=sum(mat_sr)/20) ## distància promig de l'ABC

## El Mundo
mat_sr= matrix(0,nrow=98)
for (i in 137:234){
  mat_sr[i-137]=sum(abs(sondejors[i,]-reals[]))
}
(dist_emu=sum(mat_sr)/97) ## distància promig de El Mundo

#### distàncies absolutes de les mostres amb els resultats reals###

reals= read.table("esp04res.txt")
esctotal=read.table("sondperf04.txt") #obtenció dels resultats dels sondejors perfectes
mat_mr= matrix(0,nrow=10000) #matriu on hi guardem les distàncies

for (i in 1:10000){
  mat_mr[i]=sum(abs(esctotal[i,]-reals[]))
}
(dist_mr=sum(mat_mr)/10000) ## distància promig

```

#### Annex 4: PROGRAMA PER DETERMINAR ELS ERRORS DELS SONDEJOS DE CADA PARTIT, AIXÍ COM AJUNTATS PER DRETES I ESQUERRES, AL CONGRÉS DELS DIPUTATS L'ANY 2004:

```
#### errors absoluts dels sondejos amb els resultats reals####

p=7

sondejos=read.table("sond04.txt") #obtenció dels resultats dels sondejos
reals= read.table("esp04res.txt") #obtenció dels resultats reals
teorics = read.table("sondperf04.txt")
mat_sr= matrix(0,nrow=38,ncol=p) #matriu on hi guardem les distàncies
distpartits_lv= matrix(0,ncol=p) # distàncies de La Vanguardia en tots els partits
distpartits_epe= matrix(0,ncol=p) # distàncies d'El Periódico en tots els partits
distpartits_epa= matrix(0,ncol=p) # distàncies d'El País en tots els partits
distpartits_eabc= matrix(0,ncol=p) # distàncies de L'ABC en tots els partits
distpartits_eemu= matrix(0,ncol=p) # distàncies de EL Mundo en tots els partits

# errors ajuntant els partits en dretes i esquerres

de_elv=matrix(0,ncol=2)
de_epe=matrix(0,ncol=2)
de_epa=matrix(0,ncol=2)
de_abc = matrix(0,ncol=2)
de_emu=matrix(0,ncol=2)
dretes=matrix(0,ncol=2)
esquerres=matrix(0,ncol=2)

## La Vanguardia

for (i in 1:38){
  for (j in 1:p){
    mat_sr[i,j]=sum(sondejos[i,j]-reals[j])
  }
}

for (i in 1:p){
  distpartits_lv[i] = mean(mat_sr[,i])
}

de_elv[1]=distpartits_lv[2]+distpartits_lv[3]+distpartits_lv[5]
de_elv[2]= distpartits_lv[1]+distpartits_lv[4]+distpartits_lv[7]

## El Periódico

mat_sr= matrix(0,nrow=43,ncol=p)
for (i in 39:81){
  for (j in 1:p){
    mat_sr[i-38,j]=sum(sondejos[i,j]-reals[j])
  }
}

for (i in 1:p){
  distpartits_epe[i] = mean(mat_sr[,i])
}

de_epe[1]=distpartits_epe[2]+distpartits_epe[3]+distpartits_epe[5]
de_epe[2]= distpartits_epe[1]+distpartits_epe[4]+distpartits_epe[7]

## El País

mat_sr= matrix(0,nrow=35,ncol=p)
for (i in 82:116){
  for (j in 1:p){
    mat_sr[i-81,j]=sum(sondejos[i,j]-reals[j])
  }
}

for (i in 1:p){
  distpartits_epa[i] = mean(mat_sr[,i])
}

de_epa[1]=distpartits_epa[2]+distpartits_epa[3]+distpartits_epa[5]
de_epa[2]= distpartits_epa[1]+distpartits_epa[4]+distpartits_epa[7]
```

**## ABC**

```

mat_sr= matrix(0,nrow=20,ncol=p)
for (i in 117:136){
  for (j in 1:p){
    mat_sr[i-116,j]=sum(sondejors[i,j]-reals[j])
  }
}

for (i in 1:p){
  distpartits_eabc[i] = mean(mat_sr[,i])
}

de_abc[1]=distpartits_eabc[2]+distpartits_eabc[3]+distpartits_eabc[5]
de_abc[2]= distpartits_eabc[1]+distpartits_eabc[4]+distpartits_eabc[7]

```

**## EL MUNDO**

```

mat_sr= matrix(0,nrow=98,ncol=p)
for (i in 137:234){
  for (j in 1:p){
    mat_sr[i-137,j]=sum(sondejors[i,j]-reals[j])
  }
}

for (i in 1:p){
  distpartits_eemu[i] = mean(mat_sr[,i])
}

de_emu[1]=distpartits_eemu[2]+distpartits_eemu[3]+distpartits_eemu[5]
de_emu[2]= distpartits_eemu[1]+distpartits_eemu[4]+distpartits_eemu[7]

```

**## errors en dretes i esquerres**

```

(dretes[1]=(de_emu[1]+de_abc[1])/2)
(dretes[2]=(de_emu[2]+de_abc[2])/2)
(esquerres[1]=(de_epe[1]+de_epa[1])/2)
(esquerres[2]=(de_epe[2]+de_epa[2])/2)
de_elv

```

**## errors en cada un dels partits**

```

distpartits_lv
distpartits_epe
distpartits_epa
distpartits_eabc
distpartits_eemu

```

**## errors de cada partit en els teòrics**

```

mat_tr= matrix(0,nrow=10000,ncol=p)
for (i in 1:10000){
  for (j in 1:p){
    mat_tr[i,j]=sum(teorics[i,j]-reals[j])
  }
}

mat_tr2 = matrix(0,nrow=10000,ncol=p)

for (j in 1:p){
  mat_tr2[,j]=sort(mat_tr[,j])
}

mat_tr2[250,]
mat_tr2[9750,]

```

**## dretes i esquerres en els teòrics**

```
mat_de= matrix(0,nrow=10000,ncol=2)
for (i in 1:10000){
  mat_de[i,1] = mat_tr[i,2]+mat_tr[i,3]+mat_tr[i,5]
  mat_de[i,2] = mat_tr[i,1]+mat_tr[i,4]+ mat_tr[i,7]
}
mat_de2 = matrix(0,nrow=10000,ncol=2)
for (j in 1:2){
  mat_de2[,j]=sort(mat_de[,j])
}
mat_de2[250,]
mat_de2[9750,]
```

**Annex 5: TAULES DE ERRORS DELS SONDEJOS EN ELS DIFERENTS PARTITS**

Les dades referents als errors dels sondejos en cada un dels partits les cerquem a continuació en les següents taules. L'error és un sol valor ja que seguim amb la mateixa metodologia que hem aplicat per trobar les distàncies. Trobem els errors calculant el promig de totes les diferències que podem interpretar de les combinacions de l'interval que presenten els diaris. Es tracta d'una resta entre els escons pronosticats pels sondejos i els escons donats el dia clau.

**taula 1** *Errors dels sondejos en els diferents anys amb Convergència I Unió*

Anys	LV	EPE	EPA	ABC	MUND
1984	-5	-	-14,5385	-	-
1988	5,5	9	6,38095	-	-
1992	0,44444	-	1,66667	-	-
1995	10	9,79167	8,64516	4,444444	6,5
1999	1	3,38372	3,2	-2,16667	4
2003	2,5	-0,45455	2	3	1,5
2006	3	0,16667	-3	3	3

**taula 2** *Errors dels sondejos en els diferents anys amb el Partit Socialista de Catalunya.*

Anys	LV	EPE	EPA	ABC	MUND
1984	4	-	3,46154	-	-
1988	-2	-4	-2,54762	-	-
1992	1,44444	-	-0,33333	-	-
1995	-5	-2,20833	-4,45161	-0,33333	-1
1999	-0,5	-9,6163	-2,8	-6,16667	-3
2003	7,5	8,36364	7	6,5	5,5
2006	2	3,16667	5,5	2	2

**taula 3** *Errors dels sondejos en els diferents anys amb el Partit Popular a Catalunya .*

Anys	LV	EPE	EPA	ABC	MUND
1984	-1,5	-	2,53846	-	-
1988	1,5	0	2,45238	-	-
1992	-0,55556	-	1	-	-
1995	-4,5	-5,625	-1,87097	-2,55556	-1,5
1999	1	2,465116	1,6	3,4444	1
2003	-3	-2,63636	-2	-2	0,5
2006	-1	0,55556	0	-1	-1

**taula 4** Errors dels sondejos en els diferents anys amb Esquerra Republicana de Catalunya a Catalunya..

Anys	LV	EPE	EPA	ABC	MUND
1984	1	-	4,53846	-	-
1988	-1,5	-3	-3,61905	-	-
1992	-2,33333	-	-3	-	-
1995	-0,5	-2,33333	-2,45161	-1,55556	-1,5
1999	-2,5	1,30233	-2	0,44444	-2
2003	-4,5	-3,63636	-5	-4,5	-4
2006	-3	-1,44444	-1	-3	-3

**taula 5** Errors dels sondejos en els diferents anys amb Iniciativa per Catalunya- Els Verds a Catalunya.

Anys	LV	EPE	EPA	ABC	MUND
1984	1,5	-	4	-	-
1988	-1,5	-2	-2,33333	-	-
1992	1	-	0,66667	-	-
1995	0	0,375	0,12903	0	-2,5
1999	1	2,46512	0	4,4444	0
2003	-2,5	-1,63636	-2	-3	-3,5
2006	2	0,55556	1,5	2	2

**taula 6** Errors dels sondejos en els diferents anys amb el Partido Socialista Obrero Español a Espanya .

Anys	LV	EPE	EPA	ABC	MUND
1986	-	-4	-	-10	-
1989	0,464865	-	4,0869	-	-7
1993	-14,8125	-14,8569	-17,2459	-20	-22,7826
1996	-2,16667	-15,308	-18,3191	-20,3608	-22,485
2000	13	12,5	11,12059	18,34211	4,978571
2004	-19	-28,3954	-28,1143	-29,55	-22,7347
2008	-4,5	-11,6255	-2	-	5,255965

**taula 7** Errors dels sondejos en els diferents anys amb el Partido Popular a Espanya .

Anys	LV	EPE	EPA	ABC	MUND
1986	-	-7	-	-12	-
1989	-1,33514	-	-7,18807	-	-4
1993	5,1875	2,496294	6,41531	7,5	9,846154
1996	7,833333	15,69198	17,88487	23,63918	18,83459
2000	-15	-15,5	-14,5147	-19,6579	-5,33571
2004	16,5	21,6046	21,7429	27,25	22,2653
2008	0	3,3745	-3	-	0,3189



**taula 8** Errors dels sondejos en els diferents anys amb *Convergència i Unió a Espanya* .

Anys	LV	EPE	EPA	ABC	MUND
1986	-	-1	-	0	-
1989	-0,47027	-	0	-	-1
1993	1,520833	3,029134	1,967427	-0,5	1,491639
1996	-1,52941	-0,90014	-2,01645	-2,51031	-1,48872
2000	-0,5	-1,5	-0,47059	0,657895	0,178571
2004	0	0,32558	0,457143	-1,15	0,989796
2008	-1,5	-1,56727	-1	-	-0,00868

**taula 9** Errors dels sondejos en els diferents anys amb *Izquierda Unida a Espanya* .

Anys	LV	EPE	EPA	ABC	MUND
1986	-	-2	-	2	-
1989	-4	-	-1,49416	-	-1
1993	6,520833	3,633018	2,967427	9	7,976589
1996	-1,07843	2,825598	4,467105	3,783505	6,109023
2000	5	2,5	3,255882	1,657895	0,45
2004	4,5	5,53488	6	6	1,96939
2008	2,5	2,81454	2	-	2

**taula 10** Errors dels sondejos en els diferents anys amb el *Partido Nacionalista Vasco a Espanya*

Anys	LV	EPE	EPA	ABC	MUND
1986	-	1	-	4	-
1989	2,52973	-	1,015564	-	0
1993	1,520833	1,510861	1	1,5	0,491639
1996	0,470588	0	0	0	0
2000	-1	0,5	-0,47059	-0,34211	-0,09286
2004	0	0	0,45714	0	0
2008	0,5	0,43273	1	-	1

## Annex 6: EXEMPLE DE SONDEIG: LA VANGUARDIA A LES ELECCIONS DEL PARLAMENT DE CATALUNYA L'ANY 1999

16 LA VANGUARDIA

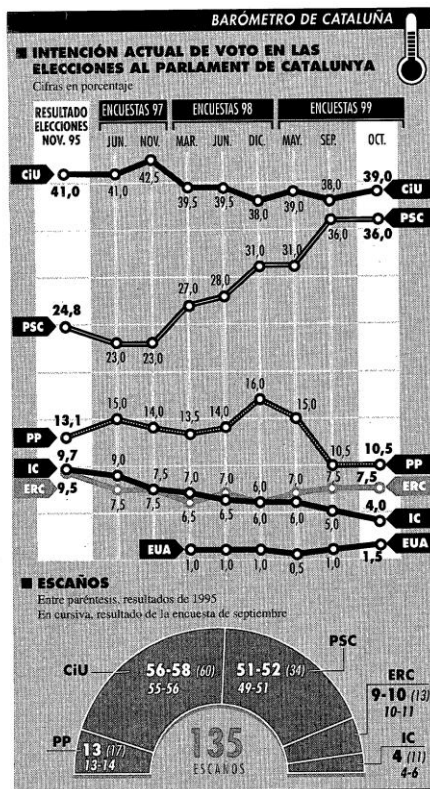
POLÍTICA

DOMINGO, 10 OCTUBRE 1999

## parlament

## Pujol afianza su ventaja sobre Maragall

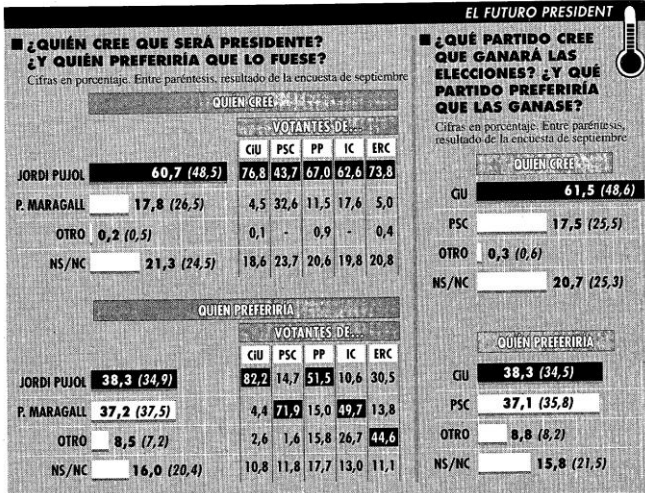
El candidato de CiU supera al del PSC en 3 puntos y sus expectativas de victoria son más elevadas que en septiembre



## ENCUESTA INSTITUTO OPINA

BARCELONA. (Redacción.) - Jordi Pujol parece afianzar su moderada ventaja sobre Pasqual Maragall cuando falta apenas una semana para la cita con las urnas. Al menos así se desprende del sondeo del Instituto Opina realizado para "La Vanguardia" durante los días 4, 5 y 6 de

octubre. Según esa encuesta, el candidato de CiU cuenta actualmente con una intención de voto del 39%, frente al 36% de su oponente socialista. Esa proyección debe aceptarse, sin embargo, con la máxima cautela, ya que sobre ella pesan dos datos de signo contrario. Por un lado, la intención de voto directa -la que expresan espontáneamente los ciudadanos consultados- ofrece una situación de empate técnico, con una



levisima ventaja de Pujol: ocho décimas, que se reducen a cinco en lo relativo a la simpatía de los electores hacia uno u otro partido. Por otra parte, sin embargo, los indicadores relativos a la evaluación que hacen los ciudadanos de la gestión de Pujol, o los parámetros respecto a las expectativas de triunfo de uno u otro candidato, son muy favorables al líder de CiU. Y queda, por último, la incógnita de si los electores socialistas -inclinados a la abstención en las autonómicas- acudirán a las urnas, tal como afirman.

El resultado en escaños de esta proyección daría a CiU una horquilla de entre 56 y 58 escaños, frente a 51 o 52 al PSC. Iniciativa, el socio electoral de los socialistas en Tarragona, Lleida y Girona, se quedaría con cuatro escaños en Barcelona -donde concurre en solitario-, aunque obtendría otros dos en las listas conjuntas PSC-IC de Tarragona y Girona. Globalmente, IC pierde un punto respecto al sondeo de septiembre -realizado antes de anunciarse su parcial coalición-, aunque el PSC no logra atraer esos votos.

Los restantes partidos mantendrían una intención de voto similar

CiU recupera un punto en un mes y se sitúa en el 39%, mientras el PSC sigue en el 36%

La intención de voto directa reduce a ocho décimas la ventaja de Pujol

PP y ERC pueden bajar 4 escaños

IC pierde un punto, que no va al PSC, y logra 4 escaños en Barcelona, uno en Tarragona y otro en Girona

a la del barómetro de septiembre, con una leve disminución en los escaños adjudicados. Así, con un 10,5% de los votos, el PP obtendría 13 diputados, y con un 7,5% Esquerra oscilaría alrededor de nueve.

La ventaja del candidato de CiU coincide con una leve mejora de la puntuación que le otorgan los ciudadanos, suficiente para ponerse por delante de Maragall, que mantiene casi intacta la calificación del mes pasado, aunque, eso sí, con su nota más baja de los últimos dos años. Entre los restantes líderes políticos se registra un estancamiento a la baja, salvo en el caso de Carod-Rovira y de Lucchetti, mientras que Albert Fernández continúa último.

Esta leve inflexión en favor de Pujol se aprecia especialmente en las expectativas de victoria de cada candidato y en las preferencias de los electores sobre el próximo presidente. Así, con relación a la primera cuestión, vuelven a ser una sólida mayoría quienes creen que Pujol será el próximo presidente, de manera que suman de nuevo más del 60%, después de que en la encuesta

Continúa en la página 18



Si quiere experimentar lo que se siente al volante del Audi A4, no pierda ni un segundo.

Acérquese a Grupo Letamendi y empiece a disfrutar hoy mismo de su Audi A4.



**Grupo Letamendi, S.A.**  
Muntaner, 81  
Tel. 934 543 500 - 08011 Barcelona  
Valencia, 177  
Tel. 934 538 800 - 08011 Barcelona