

HISPALINK-ASTURIAS

---

Documentos de trabajo

**COMBINACIÓN DE PREDICCIONES Y MÉTODOS DE EVALUACIÓN**

**DOCUMENTO DE TRABAJO 2/03**

Han participado en la elaboración de este documento de trabajo:

Blanca Moreno Cuartas

Ana Jesús López Menéndez

**Depósito Legal : AS-5022-2003**

## COMBINACIÓN DE PREDICCIONES y MÉTODOS DE EVALUACIÓN

### INDICE DE CONTENIDOS

<b>INTRODUCCIÓN .....</b>	<b>5</b>
<b>1 TÉCNICAS Y FUENTES DE PREDICCIÓN ECONÓMICA .....</b>	<b>7</b>
<b>1.1 Tipología de las técnicas de predicción económica.....</b>	<b>8</b>
<b>1.2 Técnicas cuantitativas de predicción .....</b>	<b>13</b>
1.2.1 Análisis clásico de series temporales univariantes.....	13
1.2.1.1 Métodos de alisado .....	15
1.2.1.2 Métodos de descomposición.....	20
1.2.2 Análisis estocástico de series temporales .....	22
1.2.2.1 Modelos ARIMA .....	23
1.2.2.2 Modelos ARCH y GARCH .....	27
1.2.2.3 Representación en el espacio de los estados.....	29
1.2.2.4 Nuevas tendencias .....	31
1.2.3 Modelos multivariantes de series temporales.....	31
1.2.3.1 Análisis de intervención y de función de transferencia.....	32
1.2.3.2 Modelo estocástico multivariante.....	34
1.2.3.3 Modelos VAR y MCE .....	35
<b>1.3 Técnicas bayesianas de predicción: los puntos de giro. ....</b>	<b>36</b>
<b>1.4 Técnicas cualitativas de predicción: las Encuestas de Opinión.....</b>	<b>41</b>
1.4.1 Encuestas de Actitudes y Expectativas de Empresarios y Consumidores.....	42
1.4.2 La Encuesta de Coyuntura Industrial .....	44
<b>1.5 Algunas fuentes de predicción sobre crecimiento económico en España....</b>	<b>47</b>
1.5.1 Fuentes de predicción cuantitativa .....	48
1.5.2 Fuentes de predicción cualitativa .....	49
<b>2 MEDIDAS DE EVALUACIÓN DE PREDICCIONES .....</b>	<b>53</b>
<b>2.1 Tipificación de la precisión y de las medidas de evaluación.....</b>	<b>53</b>
<b>2.2 Contraste de insesgamiento de las predicciones individuales.....</b>	<b>56</b>
<b>2.3 Medidas basadas en el error de predicción.....</b>	<b>62</b>
2.3.1 Predicciones cuantitativas.....	63
2.3.2 Predicciones cualitativas.....	67
2.3.3 Predicciones probabilísticas: medidas de Calibración .....	69
<b>2.4 Evaluación de los puntos de giro .....</b>	<b>69</b>
2.4.1 Predicciones cuantitativas.....	70
2.4.2 Predicciones cualitativas.....	71
2.4.3 Predicciones probabilísticas .....	73

<b>2.5</b>	<b>Nuevas alternativas basadas en medidas de información.....</b>	<b>73</b>
2.5.1	Medida de Entropía de Theil .....	75
2.5.2	Medida cuadrática de imprecisión.....	76
2.5.3	Información cuadrática asociada a las predicciones cuantitativas.....	78
2.5.4	Imprecisión cuadrática basada en errores relativos .....	81
<b>3</b>	<b>COMBINACIÓN DE PREDICCIONES CUANTITATIVAS .....</b>	<b>83</b>
<b>3.1</b>	<b>Métodos cuantitativos de combinación.....</b>	<b>83</b>
<b>3.2</b>	<b>Enfoque Clásico .....</b>	<b>84</b>
3.2.1	Técnicas basadas en la varianza de los errores de predicción individual.....	85
3.2.2	Técnicas basadas en la regresión de las predicciones individuales.....	87
<b>3.3</b>	<b>Técnicas basadas en metodología bayesiana.....</b>	<b>89</b>
<b>3.4</b>	<b>Problemas en la combinación y técnicas desarrolladas .....</b>	<b>93</b>
3.4.1	Alteración de las hipótesis estructurales.....	95
3.4.1.1	Muestras pequeñas.....	95
3.4.1.2	Relaciones de colinealidad de predicciones .....	97
3.4.1.3	Predicciones atípicas.....	98
3.4.1.4	Cambio estructural: Coeficientes cambiantes en el tiempo.....	99
3.4.1.5	Procedimientos de combinación para series no estacionarias .....	105
3.4.1.6	Regresores estocásticos y variables desplazadas.....	106
3.4.1.7	Especificación errónea del modelo.....	107
3.4.2	Incumplimiento de Hipótesis sobre la perturbación aleatoria .....	107
3.4.2.1	Media no nula .....	108
3.4.2.2	Homocedasticidad .....	108
3.4.2.3	Errores autocorrelacionados .....	108
3.4.2.4	No normalidad y Matriz de covarianzas no escalar.....	110
<b>4</b>	<b>COMBINACIÓN DE PREDICCIONES CUALITATIVAS.....</b>	<b>112</b>
<b>4.1</b>	<b>Técnicas cualitativas de combinación.....</b>	<b>113</b>
4.1.1	Tipificación de las técnicas.....	113
4.1.2	Técnicas de consenso bayesianas .....	114
4.1.3	La predicción individual según el estadio de predicción.....	116
<b>4.2</b>	<b>Nuevas alternativas para las Encuestas de Coyuntura Industrial.....</b>	<b>117</b>
4.2.1	Técnicas cuantitativas de combinación .....	118
4.2.2	Técnicas bayesianas de combinación .....	119
4.2.2.1	Distribución Uniforme-Multinomial .....	121
4.2.2.2	Dirichlet-Multinomial.....	122
<b>4.3</b>	<b>Cuantificación de predicciones combinadas cualitativas.....</b>	<b>124</b>
4.3.1	Métodos de regresión .....	124
4.3.2	Métodos probabilísticos.....	127
4.3.2.1	Distribución Normal.....	129
4.3.2.2	Distribución Uniforme.....	131
4.3.2.3	Distribución Logística .....	132
4.3.2.4	Intervalo de indiferencia variable en el tiempo .....	133

<b>5</b>	<b>EVALUACIÓN DE LA PREDICCIÓN COMBINADA .....</b>	<b>136</b>
5.1	Experiencias .....	136
5.2	Contrastes sobre la ganancia de información y de precisión .....	139
5.2.1	Contraste de abarcamiento.....	139
5.2.2	Test de Stekler .....	140
5.2.3	Contraste de reducción significativa del error de predicción .....	140
5.3	Propiedades de las ponderaciones estimadas.....	142
5.3.1	Estabilidad en las ponderaciones.....	142
5.3.2	La distribución de los pesos de la combinación .....	143
5.4	Características de los errores de la predicción combinada .....	144
5.4.1	Sesgo y la Eficiencia de las predicciones combinadas .....	144
5.4.2	Distribución de los errores de la predicción combinada.....	145
5.4.2.1	Simetría de las distribución .....	145
5.4.2.2	Métodos para aproximar la distribución: Regresión Cuantílica .....	146
5.5	Frontera de eficiencia.....	150
	<b>BIBLIOGRAFÍA .....</b>	<b>152</b>

## **INTRODUCCIÓN**

La predicción es una actividad esencial cuya importancia ha crecido en las últimas décadas, convirtiéndose en imprescindible en la gran mayoría de los procesos de toma de decisiones. Planificadores y decisores disponen de una gran variedad de técnicas para realizar sus predicciones, que van desde las más subjetivas e intuitivas hasta los métodos cuantitativos más complejos. Entre ambos extremos existen innumerables posibilidades que difieren en su filosofía, su coste, su complejidad y su precisión.

Teniendo en cuenta que las predicciones relativas a una misma magnitud económica pueden ser realizadas por diferentes agentes y a partir de distintos métodos, puede ser conveniente combinar varias predicciones que mejoren la precisión de las predicciones individuales, puesto que cada método empleado y cada agente implicado pueden capturar diferentes aspectos de la información.

La idea de la combinación de predicciones asume que no es posible identificar mediante un único modelo el proceso subyacente que explica un fenómeno económico, y que cada modelo de predicción es capaz de aprovechar distinto tipo de información disponible para la predicción, por lo que es posible que una combinación de las predicciones efectuadas según distintas técnicas sea la predicción más precisa. Si consideramos además que existe información contextual que las técnicas de predicción no pueden captar y la percepción de los individuos sí, sería también conveniente examinar tanto sus predicciones subjetivas, como la combinación de éstas, puesto que si los individuos no comparten la misma información o valoración sobre el contexto económico, una integración de sus predicciones aprovecharía más la información disponible. Así pues, la combinación puede estudiarse tanto para las predicciones subjetivas como para las predicciones efectuadas por las técnicas cuantitativas, además de poder contemplar una combinación de ambos tipos de predicciones.

Además de existir muchas técnicas para la elaboración de predicciones individuales, los métodos empleados para la obtención de predicciones combinadas son también muy diversos, por lo que la elección del método de combinación y de las predicciones individuales que se combinan dependerá, además de la preferencia del investigador, de la complejidad y coste de la técnica empleada y del grado de precisión aportada por la predicción. Dado que las dos primeras implican una cierta subjetividad y las diferencias entre los distintos métodos son a veces imprecisas, los indicadores que se emplean habitualmente son los basados en la precisión de las predicciones que tienen

como objetivo la medición de la desutilidad o el coste asociado a los pares de predicciones y realizaciones.

Como consecuencia de estas consideraciones, en el capítulo 1 del documento clasificamos las técnicas de predicción económica según distintos criterios y describimos las más apropiadas para la predicción a corto plazo y el análisis de los cambios de fase en la coyuntura económica. Se indican además algunas de las fuentes y organismos que elaboran predicciones sobre el crecimiento económico en España.

En el capítulo 2 se realizan algunos análisis para evaluar las características de las predicciones que van a ser combinadas, utilizando medidas tradicionales y aportando nuevas alternativas, que adoptan como referencia la teoría de la información.

En el capítulo 3 examinamos las posibilidades para combinar varias predicciones cuantitativas a partir de los enfoques clásico y bayesiano.

Posteriormente, en el capítulo 4 estudiamos la combinación de predicciones cualitativas y desarrollamos nuevos métodos bayesianos a partir de la información suministrada por las encuestas de opinión. Además, analizamos métodos de cuantificación para las predicciones combinadas cualitativas.

Finalmente, en el capítulo 5 describimos herramientas para contrastar si todos los modelos de combinación son igualmente adecuados y verificar si realmente existe una ganancia en precisión con respecto a las predicciones individuales usadas.

## 1 TÉCNICAS Y FUENTES DE PREDICCIÓN ECONÓMICA

Consideremos una magnitud  $Y$  para la cual, en un momento  $t$ , deseamos conocer el valor que tendrá en cierto horizonte temporal  $h$  de amplitud  $T$  ( $h=1, \dots, T$ ):  $y_{t+h}$ . Las predicciones pueden realizarse empleando tanto métodos cuantitativos (técnicas estadístico-económicas) como cualitativos (encuestas de opinión empresarial, panel de expertos) por lo que para pronosticar  $y_{t+h}$  existirá diversidad de resultados según el método empleado. La idoneidad de la técnica empleada no es siempre la misma, puesto que ésta puede variar en base a las situaciones de predicción en que se aplica, esto es, la información disponible, los medios técnicos y humanos empleados y el horizonte temporal de predicción.

La delimitación de la situación de predicción y la tipología de las técnicas en base a esta situación puede ayudarnos a decidir qué método puede ser el más apropiado en cada situación.

En este capítulo presentamos una breve tipificación de las técnicas de predicción, describiendo los métodos de predicción a corto plazo que de acuerdo con la experiencia<sup>1</sup> resultan más apropiados: los basados en la extrapolación de series temporales y en las encuestas de opinión sobre expectativas a empresarios y consumidores.

En ocasiones puede resultar de crucial interés predecir los puntos de giro de una serie económica dado que su presencia señala cambios de fase en la coyuntura económica general. Por este motivo, en este capítulo analizamos cómo las series temporales de algunos indicadores pueden convertirse en predictores de esos puntos de inflexión.

Si bien la metodología empleada para las técnicas de series temporales es universal, en el caso de las encuestas esta metodología puede tener peculiaridades dependiendo de las características propias de cada encuesta. Puesto que en España la encuesta de expectativas más relevante es la Encuesta de Coyuntura Industrial, nos centraremos en la descripción de sus principales rasgos metodológicos<sup>2</sup>.

---

<sup>1</sup> Diversos autores como Newbold y Granger (1974), Armstrong (1978), Makridakis y Wheelwright (1978), Makridakis (1984) han puesto de manifiesto que los métodos de extrapolación de series temporales son mejores (o no peores) que los modelos econométricos para predicciones a corto plazo, si bien para el largo plazo los modelos econométricos pueden resultar más adecuados.

<sup>2</sup> Agradecemos al Jefe de Servicio de Estudios y responsable técnico de Encuestas de Coyuntura de la Subdirección General de Estudios de la Secretaría General Técnica del Ministerio de Ciencia y Tecnología la información suministrada referente a la metodología más reciente empleada en su elaboración.

Finalmente hacemos referencia a algunas de las fuentes y organismos dedicados a la elaboración de predicciones sobre el crecimiento económico en España y a otras encuestas de opinión españolas que se emplean como referencia para la elaboración de predicciones cualitativas sobre el ciclo económico.

### 1.1 Tipología de las técnicas de predicción económica

Los métodos para predecir una variable económica  $Y$  pueden ser muy numerosos y su clasificación depende de los criterios considerados.

Llamaremos  $Y$  a la variable para la cual, en el momento actual  $t$ , queremos anticipar su comportamiento futuro en un horizonte temporal  $h$  de amplitud  $T$  ( $h=1, \dots, T$ ). Denotaremos la predicción efectuada por  $\hat{y}_{t+h,t}$  y los valores verdaderos (desconocidos) por  $y_{t+h}$ .

Dependiendo de la base informativa considerada en el momento de predicción  $t$  ( $I_t$ ) y del procedimiento seguido para procesarla ( $g(I_t)$ ), es posible obtener distintas predicciones para  $y_{t+h}$ , que denotaremos con el vector  $\hat{Y}_{t+h,t} = (\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N)$ , donde cada elemento  $\hat{y}_{t+h,t}^i$ , es la predicción efectuada por cada “técnica de predicción”<sup>3</sup>  $i$  ( $i=1, \dots, N$ ).

Las distintas herramientas de predicción pueden ser clasificadas según la forma que adopten tanto  $I_t$  como  $g(I_t)$ , o bien según las características de la situación que se va a predecir. Si bien las clasificaciones implican un cierto grado de arbitrariedad, existiendo por tanto una amplia variedad de propuestas<sup>4</sup>, la tabla 1 resume algunas de tipologías más habitualmente empleadas.

---

<sup>3</sup> Empleamos aquí el término “técnicas de predicción” en sentido genérico. A lo largo de este capítulo iremos describiendo el tipo de “herramientas” y por tanto el contenido del vector de predicciones (y su notación) irá modificándose en la medida que vayan apareciendo los diferentes tipos de predicción individual.

<sup>4</sup> En Aznar y Trivez (1993) y Pulido (1989) se recopilan otras clasificaciones aportadas a lo largo de la historia.

**Tabla 1: Tipología de las técnicas de predicción**

Criterio	Clasificación	
Tipo de información	Subjetivas Series temporales Causales	Cualitativas  Cuantitativas
Grado de formalización del procedimiento	Objetivos (Causales y series temporales) Subjetivos	
Disponibilidad de información sobre la variable	Situaciones sin historia Situaciones con historia	
Horizonte de predicción	Corto Plazo Medio Plazo Largo Plazo	
Medios empleados	Escasos Medios Altos	

Si consideramos el tipo de información que se utiliza en la predicción, podemos considerar tres categorías:

1. Técnicas subjetivas, si emplean información sobre experiencias previas, opiniones y expectativas de sujetos o instituciones. En este caso  $I_t = O_t = (o_t^1, o_t^2, \dots, o_t^k)$ , donde cada elemento  $o_t^j$  hace referencia a la opinión de cada individuo  $j$  en el instante  $t$  ( $j=1, \dots, k$ ) sobre el comportamiento futuro que tendrá la variable en  $t+h$ .
2. Técnicas de series temporales, si la información que se emplea es la evolución temporal de la variable que se desea predecir. Es decir  $I_t = Y_t = (y_t, y_{t-1}, \dots, y_{t-n})$  o  $I_t = Y_t = (y_1, y_2, \dots, y_n)$ , siendo  $y_t$  el valor de  $Y$  en el momento  $t$  ( $t=1, \dots, n$ ).
3. Técnicas causales, si la información que se requiere va referida a la relación de la variable de interés con otras variables. En este caso  $I_t = x_t = (x_t^1, \dots, x_t^r)$ , siendo  $x_t^d$  la información de la variable  $d$  ( $d=1, \dots, r$ ) en el instante  $t$ . Podría ocurrir que además se considere información sobre la evolución temporal de dichas variables, con lo que  $x_t$  sería ahora una matriz  $X_t$  ( $n \times r$ ) donde cada elemento es la serie temporal de cada variable  $d$ :  $X_t^{dT} = (x_1^d, x_2^d, \dots, x_n^d)$ .

Esta clasificación no es exhaustiva, puesto que las técnicas pueden combinar todos o algunos de los tipos de información<sup>5</sup>, y por tanto  $I_t \subseteq (Y_t, X_t, O_t)$

Existen también clasificaciones que, teniendo en cuenta el tipo de información, distinguen entre las técnicas cualitativas (categoría equivalente a las técnicas denominadas subjetivas según la clasificación anterior) y cuantitativas (donde se incluyen las técnicas que requieren la información cuantificada, esto es, las clasificadas con anterioridad como técnicas de series temporales y causales).

Las técnicas que emplean información cuantitativa suelen generar predicciones cuantitativas mientras que las que emplean información cualitativa dan lugar generalmente a predicciones cualitativas. No obstante, las opiniones o expectativas pueden revelarse con valores puntuales (cuantitativos) o cuantificarse (como veremos en el capítulo 4) y también las predicciones cuantitativas pueden convertirse en predicciones cualitativas cuando se transformen en niveles relativos de variación o se considere únicamente el sentido del cambio de la variable de interés.

En cualquier caso, cuando las predicciones se efectúen en base a las opiniones de individuos, y por tanto respondan básicamente a sus valoraciones sobre el futuro (habitualmente cualitativas) las indexaremos con  $e$ , indicando que la predicción de la técnica  $j$  es en realidad una expectativa ( ${}^e\hat{y}_{t+h,t}^j$ ); en caso contrario las predicciones responderán a técnicas cuantitativas ( $\hat{y}_{t+h,t}^i$ ). Así pues, el vector  $\hat{Y}_{t+h,t}$  puede estar formado por ambos tipos de predicciones:  $\hat{Y}_{t+h,t} = (\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, {}^e\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2, \dots)$  ( $i+j=1, \dots, N$ ).

Dependiendo principalmente del grado de formalización del procedimiento  $g(I_t)$ , las técnicas se pueden clasificar en:

1. Métodos objetivos o cuantitativos: las predicciones se basan en modelos matemáticos y estadísticos, y por tanto son técnicas con un alto grado de formalización del proceso. Para todas las técnicas recogidas en esta categoría, la base informativa está disponible al comienzo del proceso de predicción en forma

---

<sup>5</sup> Como veremos más adelante, una predicción combinada, puede generarse a partir de predicciones individuales efectuadas por distintas técnicas y por tanto, indirectamente a través de ellas, generarse a partir de los tres tipos de información.

de valores por lo que se incluyen aquí las técnicas clasificadas en series temporales y las causales, es decir las cuantitativas. No obstante haciendo referencia al procedimiento éstas pueden definirse también como:

- Análisis causal: se basan en información externa a la variable. En este caso se elabora un modelo que sistematiza las leyes de comportamiento que rigen la relaciones de la variable de interés con las demás variables y esas mismas leyes son las que se emplean en la predicción.
  - Análisis de series temporales: se toma como base informativa la evolución temporal de la variable (serie temporal) a partir de la cual se elabora un modelo que describe el patrón subyacente en la serie que se empleará en la predicción. Si únicamente se consideran los valores presentes y pasados de la variable, el modelo temporal será univariante. Si se tiene en cuenta además las relaciones causales entre la serie temporal y otras variables y/o series temporales surgen los modelos multivariantes o multiecuacionales<sup>6</sup>.
2. Métodos subjetivos o cualitativos<sup>7</sup>: las predicciones se basan en opiniones que sobre el futuro de la variable de interés tienen los agentes. A diferencia de los métodos cuantitativos, las predicciones no pueden ser reproducidas por otro agente ya que no se especifica explícitamente la forma en que éste procesa la información. Como indican Aznar y Trivez (1993), en estos métodos se suele asumir que al comienzo del ejercicio de predicción no está disponible la base informativa en forma objetiva y se supone que toda la información relevante la poseen los agentes (experiencias, conjeturas, intuiciones) que efectúan la predicción. No obstante, cabe la posibilidad de que los agentes predictores puedan incluir en su base informativa valores cuantitativos y objetivos disponibles, y que además puedan compartir su propia información. En este caso, lo único que exclusivamente conoce cada agente es  $g^j(I_t)$ .

---

<sup>6</sup> Pulido (1989) realiza una breve descripción de algunas de las técnicas objetivas. Así, entre los métodos causales incluye las basadas en el establecimiento de analogías, el análisis de regularidades estadísticas, los modelos de simulación deterministas y los modelos econométricos y dentro las técnicas de series temporales incluye las fórmulas “ad hoc” de predicción, análisis de descomposición de series temporales, análisis frecuencial de procesos estocásticos, análisis temporal de procesos estocásticos y modelos probabilísticos. En el epígrafe 1.2 del documento analizamos en detalle las técnicas de series temporales.

<sup>7</sup> Entre estos procedimientos se encuentran las encuestas de intenciones o expectativas, encuestas de actitudes o sentimientos, agregación de predicciones individuales, consenso de grupo en interacción abierta, panel de expertos con retroacción controlada de la información, técnica Delphi, jurado de opinión y método de escenarios, cuya descripción puede verse en Pulido (1989).

Las características de la situación a predecir abarcan diversos aspectos referidos al horizonte de predicción, la información disponible y los medios disponibles bien sea en términos humanos, técnicos o financieros.

Según la disponibilidad de esta información y el conocimiento que se tenga sobre la variable de interés, las situaciones pueden clasificarse en:

1. Situaciones sin historia: La variable va referida a una situación nueva y por tanto  $Y = Y_t$  (entonces  $t=1$ ) o la situación aparece en el momento en el que se predice y por tanto  $Y = Y_{t+h}$  siendo  $t=0$ .
2. Situaciones con historia: Existe información pasada sobre la variable de interés,  $Y_t = (y_t, y_{t-1}, \dots, y_{t-n})$ , siendo posible por tanto tener conocimiento sobre cómo se relacionaron otras variables con ella, predicciones del pasado, etc. En este caso denotaremos por  $\hat{Y}_t = (\hat{Y}_t^1, \hat{Y}_t^2, \dots, \hat{Y}_t^N)$  a la matriz (nxN) del conjunto de predicciones pasadas que se realizaron sobre Y por las diferentes técnicas y/o individuos (expectativas) ( $i=1, \dots, N$ ), y donde cada elemento del vector es  $\hat{Y}_t^{iT} = (\hat{y}_{1,0}^i, \hat{y}_{2,1}^i, \dots, \hat{y}_{n,n-1}^i)$ <sup>8</sup>.

Por otra parte, según el horizonte de predicción h, las situaciones pueden clasificarse en:

1. Corto plazo: si se quiere predecir la variable para un periodo de menos de un año.
2. Medio plazo: si el período abarca de uno a tres o cinco años.
3. Largo plazo: si las predicciones contemplan más de tres o cinco años.

La consideración de los medios empleados, humanos, técnicos y financieros introduce bastante arbitrariedad en la tipificación de las técnicas. No obstante, los medios empleados pueden ser desde escasos (bien porque no se requiere un predictor especializado o porque la aplicación es rápida), a altos (si se necesitan profesionales y aplicaciones informáticas especializados).

La determinación de la situación de predicción en la que nos encontremos puede ayudarnos a seleccionar, a priori, aquel conjunto de técnicas que en principio parecen

---

<sup>8</sup> El caso  $\hat{Y}_t^{iT} = (\hat{y}_{1,0}^i)$  se correspondería con la predicción de una variable sin historia.

más adecuadas, si bien no resuelve la elección entre variantes de una misma técnica y entre técnicas distintas igualmente óptimas. En el capítulo 4 establecemos criterios objetivos de selección basados en la precisión de las técnicas, que nos ayudarán a escoger las más adecuadas.

Incluso una vez resuelta la decisión de qué técnica emplear, en el campo de la predicción siempre han existido discrepancias entre los análisis teóricos y los resultados empíricos. La principal razón de estas diferencias es que algunos de los supuestos teóricos sobre los que se basan las técnicas no se mantienen en el mundo real.

Como consecuencia, la implementación de cualquier método de predicción consiste en estimar un modelo  $g$  a partir del cual se procesa el conjunto de información que ayuda a explicar  $Y$  para después emplearlo en la predicción. Sin embargo modelizar el proceso y predecir, bajo la información para la que se desarrolla el modelo, no es lo mismo, al menos que se mantenga la hipótesis de constancia. Es decir, la información en  $t+h$  (por tanto fuera del período muestral) y aquella a partir del cual se estima el proceso, pueden ser diferentes estructuralmente y por tanto el proceso estimado no tiene porqué ser el mejor para la predicción. De ahí que uno de los aspectos más importantes en la predicción consista en saber qué métodos pueden ser mejores para los períodos en los que se va a realizar el pronóstico.

## **1.2 Técnicas cuantitativas de predicción**

Los métodos de predicción en series temporales tienen como objetivo modelizar el patrón subyacente en dichas series a partir de los valores pasados de la variable. Si se considera que ese patrón sigue un modelo determinista el enfoque de análisis de series temporales será clásico, mientras que si se consideran las series como realizaciones de procesos aleatorios, el enfoque será estocástico.

En este apartado nos centraremos en el estudio de los modelos de series temporales según éstos sean univariantes o multivariantes, clasificándolos según el enfoque clásico o determinista.

### **1.2.1 Análisis clásico de series temporales univariantes**

La modelización clásica de series temporales univariantes considera la existencia de un proceso determinista, dependiente del tiempo, que rige el comportamiento de la variable en la serie. Este proceso debe determinar de forma única los valores observados en la variable en cada momento del tiempo, siendo posible encontrar expresiones

cuantitativas para describir este proceso que reciben el nombre de "modelo de la serie" y que son función del tiempo.

Así pues, se persigue la construcción de un modelo econométrico dependiente del tiempo ( $t=1, \dots, n$ )  $y_t = f(t) + \varepsilon_t$  y/o del pasado de la serie  $y_t = f(t, y_{t-1}, y_{t-2}, \dots) + \varepsilon_t$ , donde  $\varepsilon_t$  es el componente de error, producto de la existencia en la serie de una perturbación aleatoria.

Partiendo de esta expresión, la *predicción de  $y_t$  para un horizonte  $h$*  vendrá dada por:  $\hat{y}_{t+h,t} = f(t+h, y_{t+h-1}, y_{t+h-2}, \dots)$  con  $h=1 \dots T$ .

Según el tipo de descripción utilizada para el modelo, el análisis clásico da lugar a dos métodos en el estudio de series temporales, cuyo tratamiento es objeto de este epígrafe: los métodos de Alisado y los métodos de descomposición de la serie.

Si bien estos dos procedimientos se diferencian en la forma de construir la función  $f$ , ambos tienen en cuenta en su construcción los elementos o *componentes* que se consideran patrones básicos para determinar las regularidades de comportamiento de una serie temporal: la tendencia<sup>9</sup>  $T_t$  o movimiento de la variable en períodos largos de tiempo, el ciclo<sup>10</sup>  $C_t$  o fluctuaciones periódicas de la variable más o menos regulares, en períodos relativamente prolongados de tiempo, la estacionalidad  $E_t$  o movimientos similares a los del ciclo pero relativos a períodos de tiempo más cortos y la aleatoriedad  $a_t$  o movimiento irregular de la variable, determinado por el azar e impredecible determinísticamente.

Trataremos en primer lugar las técnicas de alisado que se emplean habitualmente, clasificándolas según el tipo de series que traten de modelizar, es decir según los componentes que determinen o constituyan la serie temporal, y más adelante estudiaremos los métodos de descomposición de la serie, que se basan en la cuantificación de los componentes y en la construcción de una función  $f$ , cuya forma dependerá del esquema de composición de dichos componentes para determinar la serie temporal.

---

<sup>9</sup> Al hablar de tendencia debemos de tener en cuenta la existencia de tendencia determinista, y estocástica. En este caso nos referimos a la determinista que es aquella posible de modelizar.

<sup>10</sup> Cuando la tendencia y el ciclo se estudian conjuntamente hablaremos del componente extraestacional o tendencia-ciclo (TC).

### 1.2.1.1 Métodos de alisado

Los métodos de series temporales basados en las técnicas de alisado adoptan como premisa básica la existencia de algún patrón subyacente en los valores de la variable que va a ser predicha. Al promediar o alisar los valores de la serie, estas técnicas eliminan los valores extremos de la secuencia histórica, ayudando por tanto a encontrar ese patrón subyacente.

Según el tipo de series analizadas resultan adecuados distintos métodos de alisado que llevan asociadas las correspondientes ecuaciones de predicción.

**Series sin tendencia:** El modelo más sencillo es el llamado *naïve* o *ingenuo*, en el que el procedimiento de predicción repite de forma mecánica un comportamiento del pasado. Es decir, la predicción en cada período coincide con el último valor observado:

$$\hat{y}_{t+h,t} = y_t$$

Si se asume la igualdad de los incrementos se obtiene  $\hat{y}_{t+h,t}^i - y_t = y_t - y_{t-h}$ , mientras si se utiliza como predicción el valor medio de un período tenemos:  $\hat{y}_{t+h,t} = \bar{y}_t$ . La utilización de estas técnicas en aplicaciones empíricas es muy reducida. No obstante, conviene tener presente que estos modelos son utilizados como referencia para valorar la capacidad predictiva de otras alternativas más complejas, pues dada su sencillez cualquier técnica debería de producir unos errores de predicción sistemáticamente inferiores a los obtenidos por estas fórmulas ingenuas.

Otro método para reducir los impactos de las fluctuaciones aleatorias es computar una *Media Móvil de período s* ( $M_t^s$ ) donde, a efectos de predicción, se acepta la última

media móvil calculada<sup>11</sup>: 
$$\hat{y}_{t+h,t} = M_t^s = \frac{y_t + y_{t-1} + \dots + y_{t-s+1}}{s} .$$

Una de las limitaciones de las medias móviles simples es que asignan idéntico peso a todos los períodos, incluso si la información más reciente es considerada más importante. Para solventar este problema aparecen las medias móviles ponderadas, en las que los valores integrantes de la fórmula pueden tener diferentes pesos.

<sup>11</sup> Para algunas aplicaciones de las series temporales que no tienen como objetivo central la predicción también suele emplearse una media móvil que asigna la media de los valores al período medio, y debería de centrarse si el número de datos fuera par (doble media móvil). Así la predicción sería:

$$\hat{y}_{t+h,t} = M_t = \frac{y_{t+r} + y_{t+r-1} + \dots + y_t + y_{t-1} + \dots + y_{t-r}}{2r + 1} .$$

Así, cuando los pesos en la media móvil ponderada son decrecientes en forma de progresión geométrica nos encontramos ante un *Alisado Exponencial Simple* cuya expresión es:  $M_t^\alpha = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \dots$

Como fórmula de predicción, la media deberá calcularse iniciándola en el último dato disponible<sup>12</sup>:  $\hat{y}_{t+h,t} = M_t^\alpha$ .

Tal y como señalan Makridakis y Wheelwright (1978), una de las dificultades en la aplicación del alisado exponencial es la selección de la constante  $\alpha$  más adecuada. En el *Método Adaptativo del Alisado Exponencial Simple*<sup>13</sup> la ecuación básica de predicción es similar a la del alisado anterior pero con  $\alpha$  en función de t:

$$\hat{y}_{t+h,t} = \alpha_t y_t + (1 - \alpha_t) \hat{y}_t$$

donde:

$$\alpha_t = \left| \frac{E_t}{M_t} \right|$$

$$E_t = \beta e_t + (1 - \beta) E_{t-1}$$

$$M_t = \beta e_t + (1 - \beta) M_{t-1}$$

$$e_t = y_t - \hat{y}_t$$

$$\beta = 0.2$$

**Series con tendencia:** Los métodos anteriormente descritos sólo son aplicables a series sin tendencia determinista. Si la serie presenta tendencia, al adoptar como valor de la predicción la última media móvil se produce un sesgo positivo o negativo según la tendencia sea creciente o decreciente.

Cuando se contempla la existencia de tendencia, existen dos opciones inmediatas para adoptar procedimientos *ingenuos*: el primero consiste en añadir un elemento de tendencia; por ejemplo una tasa de crecimiento constante c quedando entonces la predicción  $\hat{y}_{t+h,t} = (1 + c)y_t$

La segunda opción, más general, es eliminar la tendencia de la serie, predecir la nueva serie y posteriormente añadir la tendencia.

<sup>12</sup> Una expresión alternativa sería  $\hat{y}_{t+h,t} = \alpha y_t + (1 - \alpha) \hat{y}_t$  o también, si  $e_t = y_t - \hat{y}_t$

$\hat{y}_{t+h,t} = \alpha y_t + (1 - \alpha) \hat{y}_t = \alpha(\hat{y}_t + e_t) + (1 - \alpha) \hat{y}_t = \hat{y}_t + \alpha e_t$

<sup>13</sup> Generalmente este método se suele denominar ARRSES de acuerdo con las iniciales de su nombre en inglés: *Adaptive-response-rate single exponential smoothing*.

Si trabajamos con la media como ecuación de predicción, nos llevaría en el caso de aplicarlo a la serie con incrementos  $\Delta y_t = y_t - y_{t-1}$ <sup>14</sup> a una predicción del tipo:

$$\Delta \hat{y}_{t+h,t} = \bar{\Delta y}_t = \frac{1}{n}(y_t - y_0)$$

$$\hat{y}_{t+h,t} = y_t + \Delta \hat{y}_{t+h,t} = y_t + \frac{1}{n}(y_t - y_0)$$

Si tomamos una *Media Móvil de período s*, la predicción<sup>15</sup> sería:

$$\hat{y}_{t+h,t} = M_t + \frac{y_t - y_{t-s}}{s}$$

Kauffman y Grobolloit (1967) demostraron que en el caso de tendencia lineal se obtienen predicciones óptimas:

$$\hat{y}_{t+h,t} = a_t + b_t h$$

con

$$a_t = M_t^s + (M_t^s + M_t^{s.s})$$

$$b_t = \left( \frac{2}{s-1} \right) (M_t^s + M_t^{s.s})$$

Este planteamiento general de Kauffman y Grobolloit fue adaptado por Brown al caso del alisado exponencial, llamado *Doble Alisado Exponencial de Brown o alisado exponencial lineal con parámetro único*, cuyas ecuaciones son:

$$M_t^\alpha = \alpha y_t + (1 - \alpha) M_t^\alpha$$

$$M_t^{\alpha,\alpha} = \alpha M_t^\alpha + (1 - \alpha) M_{t-1}^{\alpha,\alpha}$$

y donde la ecuación de predicción sería:

$$\hat{y}_{t+h,t} = a_t + b_t h$$

con

$$a_t = M_t^\alpha + (M_t^\alpha + M_t^{\alpha,\alpha})$$

$$b_t = \left( \frac{\alpha}{1 - \alpha} \right) (M_t^\alpha + M_t^{\alpha,\alpha})$$

<sup>14</sup> En caso de trabajar con la serie en incrementos, el cálculo de la media es igual al recorrido de la serie ( $y_t - y_0$ ) dividido por el número de períodos:

$$\bar{\Delta y}_t = \frac{1}{n} [(y_t - y_{t-1}) + (y_{t-1} - y_{t-2}) + \dots + (y_1 - y_0)] = \frac{1}{n} (y_t - y_0)$$

<sup>15</sup> Este ajuste es necesario para compensar el error sistemático en la predicción que se produce al utilizar medias móviles en series con tendencia.

Una técnica similar a la propuesta por Brown es la de *Holt-Winters o alisado exponencial con doble parámetro*, que no aplica la fórmula del doble alisamiento, sino que la propia media móvil incorpora ya un componente de tendencia:

$$M_t = \alpha y_t + (1 - \alpha)(M_{t-1} + b_{t-1})$$

donde  $b$  es un factor de variación definido a partir de otra nueva constante de alisado para la tendencia,  $\beta$ :

$$b_t = \beta(M_t - M_{t-1}) + (1 - \beta)b_{t-1}$$

y donde la ecuación de predicción será:

$$\hat{y}_{t+h,t} = M_t + b_t h$$

Este método de Holt y Winters ha sido diseñado para tendencias aditivas y no es aplicable cuando el incremento en la tendencia no adopta un valor constante, sino un porcentaje constante a los datos actuales. Cuando éste es el caso se introduce una modificación que expresa la tendencia como un ratio ( $b_t$ ) que ajusta el valor del alisamiento teniendo en cuenta el porcentaje constante. Esta técnica recibe el nombre de *Método Multiplicativo de Winters*, y sus ecuaciones son:

$$M_t = \alpha y_t + (1 - \alpha)(1 - b_{t-1})M_{t-1}$$

$$b_t = \beta \frac{M_t - M_{t-1}}{M_{t-1}} + (1 - \beta)b_{t-1}$$

donde la ecuación de predicción será:

$$\hat{y}_{t+h,t} = M_t + b_t h$$

En 1865 W.M. Chow diseñó un procedimiento llamado "*Procedimiento de control adaptativo*" que permite la autocorrección de la constante  $\alpha$  según las características de la serie. Las ecuaciones del alisamiento adaptativo de Chow son:

$$M_t = \alpha_t y_t + (1 - \alpha_t)M_{t-1}$$

$$b_t = \alpha_t (M_t - M_{t-1}) + (1 - \alpha_t)b_{t-1}$$

con la ecuación de predicción:

$$\hat{y}_{t+h,t} = M_t + \frac{(1 - \alpha_t)}{\alpha_t} b_t$$

Del mismo modo que el alisado exponencial lineal se podía emplear para series con tendencia lineal, técnicas mas complicadas de alisamiento pueden emplearse cuando el patrón subyacente de los datos es cuadrático, cúbico o de orden superior. Para ir de un alisamiento lineal a un cuadrático, la forma básica es incorporar un nivel adicional de alisamiento (triple alisamiento) y estimar un parámetro adicional para la ecuación cuadrática. Del modo análogo podemos pasar de formas cuadráticas a cúbicas, etc.

Así, las ecuaciones para el alisamiento cuadrático o *Alisado Exponencial Cuadrático de Brown* son:

$$M_t^\alpha = \alpha y_t + (1 - \alpha)M_{t-1}^\alpha$$

$$M_t^{\alpha,\alpha} = \alpha M_t^\alpha + (1 - \alpha)M_{t-1}^{\alpha,\alpha}$$

$$M_t^{\alpha,\alpha,\alpha} = \alpha M_t^{\alpha,\alpha} + (1 - \alpha)M_{t-1}^{\alpha,\alpha,\alpha}$$

$$a_t = 3M_t^\alpha - 3M_t^{\alpha,\alpha} + M_t^{\alpha,\alpha,\alpha}$$

$$b_t = \frac{\alpha}{2(1 - \alpha)^2} [(6 - 5\alpha)M_t^\alpha - (10 - 8\alpha)M_t^{\alpha,\alpha} + (4 - 3\alpha)M_t^{\alpha,\alpha,\alpha}]$$

$$c_t = \frac{\alpha^2}{(1 - \alpha)^2} (M_t^\alpha - 2M_t^{\alpha,\alpha} + M_t^{\alpha,\alpha,\alpha})$$

donde la ecuación de predicción será:

$$\hat{y}_{t+h,t} = a_t + b_t h + \frac{1}{2} c_t h^2$$

**Series con estacionalidad:** Si bien todos los métodos anteriores y otros no tratados<sup>16</sup> son válidos una vez desestacionalizada la serie, para predecir es necesario incorporar después el correspondiente índice estacional, que recoge el patrón de oscilación estacional más común para el conjunto de los datos. No obstante, Hibon (1984) en un estudio empírico en el que compara las predicciones realizadas con métodos de alisamiento sobre series con estacionalidad, comprueba que en cualquier caso la predicción es más precisa cuando se emplean métodos adecuados de alisamiento que tienen en cuenta en su formulación dicho componente.

Uno de estos métodos lo proporcionan Holt y Winters en su variante con el tratamiento de la estacionalidad llamada técnica de *Alisado Exponencial con Triple Parámetro*. Según este procedimiento, se calcula un índice de estacionalidad para cada

<sup>16</sup> Este es el caso del alisamiento de tres parámetros de Box-Jenkins, método de alisamiento armónico de Harrison, sistema monotorizado de Trigg, etc que pueden verse en Makridakis y Hibon (1979), Makridakis y Wheelwright (1978), Firth (1977), etc.

dato por relación a la media móvil ( $Y_t/M_t$ ), pero no se intenta establecer un coeficiente medio significativo sino un procedimiento de actualización de esa estacionalidad estimada:

$$s_t = \delta \frac{y_t}{M_t} + (1 - \delta)s_{t-s}$$

que se añade a las ecuaciones anteriormente utilizadas en la variante de alisado de Holt-Winters para series con tendencia:

$$M_t = \alpha \frac{y_t}{s_{t-s}} + (1 - \alpha)(M_{t-1} + b_{t-1})$$

$$b_t = \beta(M_t - M_{t-1}) + (1 - \beta)b_{t-1}$$

La fórmula de predicción en este caso será la de Holt-Winters pero multiplicada por el factor de estacionalidad:

$$\hat{y}_{t+h,t} = (M_t + b_t h)s_{t-s+h}$$

### 1.2.1.2 Métodos de descomposición

Tal y como hemos señalado con anterioridad, la modelización clásica intenta encontrar expresiones para describir el proceso determinista que gobierna el comportamiento temporal de una variable a través de un modelo función del tiempo,  $y_t = f(t) + \varepsilon_t$  donde  $t=1...n$ .

Según el método de descomposición de series temporales, esta función puede construirse por descomposición de los cuatro elementos (llamados *componentes*) que pueden aparecer en la serie temporal: la tendencia ( $T_t$ ), la estacionalidad ( $E_t$ ), el ciclo ( $C_t$ ), y el componente aleatorio ( $a_t$ ). Cuando en una serie se detectan los tres primeros componentes, en ésta sólo debe quedar la componente aleatoria, cuyo comportamiento según el enfoque clásico es impredecible, y por tanto, no modelable con los instrumentos disponibles.

En la mayor parte de las aplicaciones, los cuatro componentes anteriores quedan reducidos realmente a dos: uno que recoge los movimientos de más larga duración (tendencia o bien tendencia-ciclo, también llamado componente extraestacional) y otro que trata de cuantificar las variaciones por razones de estacionalidad.

Cada componente detectado en la serie es modelado y una vez estimadas sus expresiones analíticas se genera la función o modelo general  $f$  que se construye como una composición de los modelos particulares para cada serie.

La composición suele llevarse a cabo según dos tipos de esquema:

- Composición aditiva: los modelos de cada componente se suman entre sí:

$$Y_t = T_t + C_t + E_t + a_t$$

- Composición multiplicativa: los modelos de cada componente se multiplican entre sí<sup>17</sup>.  $Y_t = T_t * C_t * E_t + a_t$

Una vez obtenida la función  $f$ , mediante alguno de los esquemas anteriores o bien combinaciones de ambos (en caso de los denominados esquemas de composición mixta), los valores de la serie se estiman con propósitos descriptivos o predictivos.

A continuación describimos brevemente el análisis de cada componente.

**Análisis de tendencia:** Una vez detectada la presencia de la tendencia con los métodos de alisamiento ya estudiados, es necesario obtener un modelo, expresión matemática ( $T_t$ ) para representar este componente.

El método más habitual para la modelización de la tendencia consiste en ajustar a la serie una determinada curva geométrica cuya expresión algebraica sea conocida. Este modelo es una relación teórica entre la tendencia y el tiempo, o sea, una función cuya variable dependiente es el valor de la tendencia<sup>18</sup> y su variable independiente es el tiempo  $T_t=f(t)$  donde  $t=1\dots n$ .

**Análisis de la estacionalidad:** Una vez detectada y modelizada la tendencia  $T_t$  (o el componente extraestacional  $TC_t$  si incluimos el ciclo) es posible eliminar este movimiento regular a largo plazo del conjunto de oscilaciones, con el propósito de obtener el patrón de oscilación estacional más común para el conjunto de los datos analizados. Así las variaciones estacionales son aisladas y su incidencia puede ser medida mediante un promedio normalizado o estandarizado.

---

<sup>17</sup> En el esquema de composición multiplicativo el componente estacional y el cíclico suelen ser índices que recogen las oscilaciones respecto a la tendencia.

<sup>18</sup> Los modelos más comunes para representar la tendencia son el modelo lineal, el cuadrático, el exponencial, el exponencial inverso, el potencial, el logarítmico, el logístico, el de Gompertz, el autorregresivo y otros. Firth (1977) los describe gráficamente y Otero (1993) describe la expresión funcional que las origina.

Existen muchas variantes para realizar este proceso, aunque la línea más utilizada es la técnica de relación de la media móvil, que siguiendo las propuestas de Shinkin de finales de los años 50 en el *Bureau of the Census* de los Estados Unidos, se concretó en 1965 en un método muy elaborado, aún ampliamente utilizado a nivel internacional: X-11. Otros métodos utilizados son el FORAN, desarrollado por McLaughlin en los 60, y el "*Link relative model*"<sup>19</sup>.

Una vez identificado el componente estacional resulta a veces conveniente su eliminación de la serie inicial, procedimiento que se denomina desestacionalización y que consiste en contrarrestar el efecto que ejerce la estacionalidad sobre la serie temporal estudiada.

**Análisis del ciclo:** Existen distintas vías para la aproximación del componente cíclico. La más sencilla es el *método de los residuos*, según la cual el ciclo se obtiene tras proceder primero a una desestacionalización de la serie, después a la eliminación de la tendencia y por último a la eliminación de los componentes residuales.

Otro método es el del *análisis armónico*, en el que es necesario admitir que los ciclos presentan oscilaciones periódicas<sup>20</sup>.

Finalmente, si la serie no presenta componente estacional la predicción se efectúa a partir de la función de ajuste de la tendencia que en ocasiones se hace sobre la serie suavizada. En cambio, si la serie presenta componente estacional, es necesario incorporar dicho componente a la predicción efectuada en la serie desestacionalizada, correspondiente al período que se predice. La aplicación de este procedimiento asume que el patrón de estacionalidad en el período de predicción no es significativamente distinto del detectado en el comportamiento pasado de la serie estudiada.

### 1.2.2 Análisis estocástico de series temporales

Este tipo de análisis supone que el patrón subyacente en las series temporales no sigue un modelo determinista sino estocástico.

---

<sup>19</sup> Lewandowsky (1982) describe el método empleado para predecir las ventas en campañas europeas durante diez años. Se trata del método FORSYS, basado en la descomposición clásica pero en el que incorpora más componentes, tales como la influencia de días especiales, de acciones especiales y la influencia de variables exógenas que pueden afectar a las ventas. En Lewandowsky (1984) se realiza una comparación de este método con otros empleados en la predicción, obteniendo resultados que elevan su método como uno de los que origina una relación error de predicción-complejidad más favorable.

<sup>20</sup> Una descripción más amplia de los métodos clásicos de tratamiento de series temporales puede verse en Pérez y otros (1993), Makridakis y Wheelwright (1978), etc.

Siguiendo este enfoque, una serie  $y_t$  puede considerarse generada por una sucesión de variables aleatorias que se denomina proceso estocástico y a la que exigiremos un cierto equilibrio estadístico o estacionariedad<sup>21</sup>.

Las técnicas de predicción univariantes más difundidas son las basadas en modelos *Autorregresivos Integrados de Medias Móviles* (ARIMA). No obstante, cuando no se cumple la hipótesis de homocedasticidad en los residuos es aconsejable utilizar una variante: los modelos *Autorregresivos de Heterocedasticidad Condicionada Generalizados* (GARCH-ARCH).

Entre los métodos de análisis de series univariantes basados en la utilización de procesos estocásticos destaca también la representación en el espacio de los estados de series temporales univariantes. Asimismo las situaciones de no linealidad han sido abordadas a través de diversas alternativas como la teoría del caos, las redes neuronales y los modelos neuroborrosos.

### 1.2.2.1 Modelos ARIMA

El teorema de H. Wold (1938) establece que cualquier proceso estocástico  $y_t$  estacionario puede ser aproximado con precisión predeterminada mediante *un Modelo Autorregresivo de Medias Móviles* ARMA ( $p, q$ )

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

en el que  $y_t$  es el proceso estocástico estacionario,  $a_t$  es ruido blanco<sup>22</sup> y  $\phi$  y  $\theta$  son parámetros constantes (coeficientes de la regresión a estimar).

En la práctica la mayoría de las series económicas no son estacionarias, es decir, tienen tendencia en media y en varianza, luego la modelización de series de este tipo con modelos ARMA requiere previamente la conversión de una serie no estacionaria en

<sup>21</sup> Se dice que un proceso es *estacionario* si sus distribuciones finito-dimensionales son invariantes a cualquier retardo en el tiempo. No obstante, en la práctica resulta habitual trabajar con el concepto de *estacionariedad débil* que exige que para todo  $t$  se verifique:

$$E(y_t) = \mu_y$$

$$V(y_t) = \sigma_y^2$$

$$Cov(y_t, y_{t-k}) = \gamma_k \text{ para } k=0, 1, 2, \dots$$

es decir su valor medio es constante (estacionariedad en media), su varianza también es constante (estacionariedad en varianza) y la autocorrelación se mantiene constante a lo largo del tiempo, dependiendo sólo del retardo  $k$  y no del período de referencia  $t$ . Como veremos más adelante, si la serie original no es estacionaria será necesario realizar determinadas transformaciones hasta que cumpla este requisito.

<sup>22</sup> Un ruido blanco  $a_t$  con  $t=1 \dots n$  es un caso de serie estacionaria donde

$$E(a_t) = 0 \text{ para todo } t, V(a_t) = \sigma_a^2 \text{ para todo } t, \text{ y } Cov(a_t, a_{t-k}) = \gamma_k = 0$$

otra estacionaria. Así, será necesario la eliminación de la tendencia en varianza (a partir de transformaciones de Box-Cox) y la eliminación de la tendencia (y/o componente estacional) en media, ya sea ésta determinista o estocástica. En el segundo caso se utilizan los operadores diferencias<sup>23</sup>:

$$(1-L)y_t = \Delta y_t = y_t - y_{t-1}$$

$$(1-L^s)y_t \Delta_s y_t = y_t - y_{t-s}$$

Por último, cabe destacar que si la amplitud de la oscilación cíclica estacional va aumentando o disminuyendo, la serie presenta un aspecto de tendencia en variabilidad, que se eliminará aplicando varias veces el operador de diferencias estacionales anterior, antes de aplicar la transformación de Box-Cox.

Como ya hemos comentado, el teorema de Wold establece la modelización de procesos estocásticos estacionarios a partir de modelos ARMA  $(p, q)$ . En el caso de modelizar la serie  $y_t$ , no estacionaria será necesario proceder a calcular las diferencias sucesivas  $(d)$  de la serie con el propósito de convertirla en estacionaria, y su modelización se llamará ahora *Modelo Autorregresivo Integrado de Medias Móviles* ARIMA  $(p, d, q)$  donde  $d$  hace referencia al número de diferencias o integraciones.

Un modelo ARIMA  $(p, d, q)$  viene dado por la expresión:

$$\Delta^d y_t = \phi_1 \Delta^d y_{t-1} + \dots + \phi_p \Delta^d y_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$$

en la que se ha eliminado el término independiente, que será frecuentemente nulo al trabajar con series en diferencias.

Una expresión alternativa se obtiene al expresar las variables desplazadas en el tiempo con el operador retardo,  $L$ , tal que aplicado a un valor de la variable, desfase éste un período  $Ly_t = y_{t-1}$ . Si este operador es aplicado en forma sucesiva en  $s$  períodos se tendría  $L^s y_t = y_{t-s}$ , quedando la nueva expresión para ARIMA  $(p, d, q)$  como:

$$(1 - \phi_1 L - \dots - \phi_p L^p)(1-L)^d y_t = (1 + \theta_1 L + \dots + \theta_q L^q) a_t$$

$$\phi_p(L) \Delta^d y_t = \theta_q(L) a_t$$

---

<sup>23</sup> La no estacionariedad en media o lo que es lo mismo, la existencia de una tendencia determinista, puede ser eliminada mediante la especificación de dicha tendencia en la estimación del modelo, no siendo por tanto necesario diferenciar la serie para llegar a la estacionariedad. Sin embargo cuando la serie es estacionaria en varianza (esto es, tiene tendencia estocástica) la aplicación del operador diferencia  $\Delta$  es el método adecuado para transformar la serie en estacionaria.

Si la serie original presenta una tendencia en la estacionalidad, entonces puede exigir una o varias diferencias de orden estacional. Tenemos así los modelos SARIMA  $(p, d, q)$  cuya expresión sería:

$$\Phi(L^s)(1-L^s)^D y_t = \Theta(L^s)a_t$$

Todos los modelos estacionales pueden combinarse en uno general denominado ARIMA estacional multiplicativo cuya notación es:  $ARIMA(p, d, q) \times (P, D, Q)_s$

$$\phi(L)(1-L)^d \Phi(L^s)(1-L^s)^D y_t = \theta(L)\Theta(L^s)a_t$$

y cuya expresión resumida sería:

$$\phi_p(L)\Delta^d \Phi_p(L^s)\Delta_s^D y_t = \theta_q(L)\Theta_q(L^s)a_t$$

Una vez que el modelo ha sido estimado y sometido a diversas comprobaciones para su validación, se convierte en un instrumento útil para las predicciones. El uso más importante de los modelos autorregresivos de medias móviles es predecir los valores futuros de  $y_t$ .

Así, para un modelo AR(1)  $y_t = \phi_0 + \phi_1 y_{t-1} + a_t$  si se quiere predecir la variable en un horizonte  $h=1$  :

$$y_{t+1} = \phi_0 + \phi_1 y_t + a_{t+1}$$

si el coeficiente  $\phi_1$  es conocido se puede predecir  $y_{t+1}$  condicionada a la información disponible en el período t como<sup>24</sup>

$$E_t y_{t+1} = \phi_0 + \phi_1 y_t$$

donde  $E_t y_{t+h}$  es la esperanza condicional de  $y_{t+h}$  dada la información disponible en t, formalmente  $E_t y_{t+h} = E(y_{t+h} / y_t, y_{t-1}, y_{t-2}, \dots, a_t, a_{t-1}, \dots)$

De la misma manera si  $y_{t+2} = \phi_1 y_{t+1} + a_{t+2}$ , la predicción para  $y_{t+2}$  condicionada a la información disponible en t es:  $E_t y_{t+2} = \phi_0 + \phi_1 E_t y_{t+1}$ , donde sustituyendo

$$E_t y_{t+1} = \phi_0 + \phi_1 y_t \text{ obtenemos: } E_t y_{t+2} = \phi_0 + \phi_0 \phi_1 + \phi_1^2 y_t$$

$$\text{Procediendo de la misma manera: } E_t y_{t+3} = \phi_0 + \phi_0 \phi_1 + \phi_0 \phi_1^2 + \phi_1^3 y_t$$

$$\text{y en general: } E_t y_{t+h} = \phi_0 (1 + \phi_1 + \phi_1^2 + \dots + \phi_1^{h-1}) + \phi_1^h y_t$$

---

<sup>24</sup> Según el criterio del mínimo error cuadrático medio (MECM) para la elección de predictores la mejor predicción por punto para un proceso ARIMA es aquella que se obtiene mediante la esperanza matemática condicional del proceso.

ecuación que es llamada *función de predicción*.

Como  $|\phi_1| < 1$  la función de predicción es convergente. En el límite si  $h \rightarrow \infty$

entonces 
$$E_t y_{t+h} \rightarrow \frac{\phi_0}{(1-\phi_1)}$$

Procediendo de la misma manera para los modelos ARMA ( $p, q$ )

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$$

expresión que adelantada un período proporciona:

$$y_{t+1} = \phi_1 y_t + \dots + \phi_p y_{t-p+1} + a_{t+1} + \theta_1 a_t + \dots + \theta_q a_{t-q+1}$$

y tomando esperanzas matemáticas condicionadas<sup>25</sup>:

$$\begin{aligned} \hat{y}_{t+h,t} &\equiv E_t y_{t+h} = \phi_1 y_t + \dots + \phi_p y_{t-p+1} + \theta_1 a_t + \dots + \theta_q a_{t-q+1} \\ &\equiv E(y_{t+h} / y_t, y_{t-1}, y_{t-p+1}, a_t, a_{t-1}, a_{t-q+1}) \end{aligned}$$

El error de predicción será por tanto:  $y_{t+1} - E_t y_{t+1} = a_{t+1}$ , que para ser calculado requiere antes haber obtenido la predicción para  $y_{t+1}$ , cuyo cálculo necesita tener calculados los  $a_t \dots a_{t+1-q}$ .

De la misma manera, para estimar  $a_t$  es preciso haber calculado la predicción para  $y_t$ , que obliga a haber obtenido antes  $a_{t-1}, \dots, a_{t-q}$ , y así sucesivamente.

Estimados así los  $a_t$  para el período muestral se obtienen los  $\hat{a}_t$ , que sustituidos dan lugar a la predicción de origen  $t$  con un período de adelanto ( $h=1$ ):

$$\hat{y}_{t+1,t} = \phi_1 y_t + \dots + \phi_p y_{t-p+1} + \theta_1 \hat{a}_t + \dots + \theta_q \hat{a}_{t-q+1}$$

y procediendo de la misma manera para la predicción de  $y_{t+2}$  se obtiene:

$$\hat{y}_{t+2,t} = \phi_1 \hat{y}_{t+1,t} + \dots + \phi_p y_{t-p+1,t} + \theta_1 \hat{a}_{t+1} + \dots + \theta_q \hat{a}_{t-q+2}$$

La predicción para un horizonte  $h$ , donde  $l < p$  y  $l < q$  será:

$$\hat{y}_{t+h,t} = \phi_1 \hat{y}_{t+h-1,t} + \dots + \phi_p y_{t-p+h,t} + \theta_1 \hat{a}_t + \dots + \theta_q \hat{a}_{t-q+h}$$

y cuando  $l > p$  y  $l > q$  será

<sup>25</sup>Al igual que ocurría con el modelo AR(1) las variables con subíndices inferiores a  $t+1$  han dejado de ser aleatorias y por tanto sus esperanzas matemáticas coinciden con sus realizaciones:

$$E(a_{t+1} / y_t, y_{t-1}, y_{t-p+1}, a_t, a_{t-1}, a_{t-q+1}) = 0$$

$$\hat{y}_{t+h,t} = \phi_1 \hat{y}_{t+h-1,t} + \dots + \phi_p y_{t-p+h}$$

En el límite, si  $h \rightarrow \infty$   $E_t y_{t+h}$  tiende a la media. Escribiendo el modelo ARMA como un proceso de media  $\mu$ :

$$y_t = \mu + \phi^{-1}(L)\theta(L)a_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots$$

adelantando la ecuación  $h$  períodos para obtener  $y_{t+h}$  y tomando esperanzas se obtiene:

$$\hat{y}_{t+h,t} = \mu + a_t + \psi_1 a_{t+h-1} + \psi_2 a_{t+h-2} + \dots$$

ecuación donde, cuando el horizonte temporal es infinito se anulan todos los términos  $a_t$ .

Si consideramos el error de predicción  $e_{t+h} = y_{t+h} - \hat{y}_{t+h,t}$  y sustituimos  $y_{t+h}$  e  $\hat{y}_{t+h,t}$  por las ecuaciones respectivas, obtenemos (teniendo en cuenta que  $E(e_{t+h})=0$ ) una varianza del error constante igual a  $E(e_{t+h})^2 = (1 + \psi_1^2 + \dots + \psi_{h-1}^2)\sigma_a^2$ , donde a medida que  $h$  crece aumenta la varianza del error de predicción.

### 1.2.2.2 Modelos ARCH y GARCH

En muchas series económicas la varianza residual no permanece constante, pudiendo estar relacionada con el tamaño de los residuos anteriores<sup>26</sup>, de modo que un gran valor de la perturbación del período actual incrementa la varianza condicional para períodos siguientes. Los modelos *autorregresivos de heterocedasticidad condicionada* ARCH y GARCH permiten a la varianza condicional de las series depender de las realizaciones pasadas del error.

Una aproximación para predecir la varianza en una serie es introducir explícitamente una variable independiente que ayude a predecir esta volatilidad. Si consideramos el caso más simple:  $y_{t+h} = a_{t+h}x_t^d$  donde  $a_{t+h}$  es un ruido blanco con varianza  $\sigma^2$ , la varianza de  $y_{t+h}$  condicionada a valores observados de  $x_t^d$  es:

$$Var(y_{t+h} / X_t^d) = (X_t^d)^2 \sigma^2$$

donde, dada la relación entre  $X_t^d$  y la varianza de  $y_{t+h}$ , si los valores sucesivos de  $X_t^d$  tienen una correlación positiva, la varianza condicional de  $y_{t+h}$  también la tendrá,

<sup>26</sup> En los modelos autorregresivos de medias móviles  $E(a_t a_{t-j}) = 0$

por lo que la introducción de la serie  $X_t^d$  ayudará a explicar la volatilidad de la serie  $Y_t$ <sup>27</sup>.

Utilizando elecciones *ad hoc* para  $X_t^d$  y/o transformaciones en los datos, en 1982 Engle muestra que es posible modelizar simultáneamente la media y la varianza de las series, introduciendo el concepto de modelos ARCH, que en 1986 fueron generalizados por Bollerslev dando lugar a los modelos GARCH. La importancia y reconocimiento de estos modelos fueron confirmados recientemente con la concesión a Engel del premio Nobel de Economía 2003<sup>28</sup>.

La especificación general de un modelo GARCH ( $p, q$ ) es

$$y_t = \gamma X_t^d + a_t$$

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i a_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$$

donde la primera ecuación es escrita en función de variables exógenas y del término de error y la segunda es la ecuación de predicción (con horizonte temporal de un período) de la varianza basada en información pasada, llamada varianza condicional. La ecuación de la varianza condicional es función de tres términos:

1. La media  $\omega$
2. Información sobre la volatilidad de períodos pasados, medidos como retardos del cuadrado del residuo de la ecuación principal  $a_{t-q}^2$  (término ARCH)
3. Predicciones de la varianza de períodos previos  $\sigma_{t-p}^2$  (término GARCH)

Un modelo ARCH es un caso especial de modelo GARCH donde no hay predicciones de la varianza retardadas en la ecuación de la varianza condicional.

Los modelos GARCH son muy utilizados en las series económicas del ámbito financiero, donde un agente predice la varianza de un período mediante una media ponderada de la media a largo plazo (la constante), la varianza predicha a partir de períodos pasados (término GARCH) y la información acerca de la volatilidad observada en períodos previos (término ARCH).

<sup>27</sup> En la práctica se suele modificar el modelo básico transformado en logaritmos  $\ln(y_t) = \phi_0 + \phi_1 \ln(x_{t-1}) + \ln(a_t)$

<sup>28</sup> Este galardón fue concedido conjuntamente a Engle y Granger, cuyas aportaciones aparecen resumidas en <http://www.nobel.se/economics/laureates/2003/>.

### 1.2.2.3 Representación en el espacio de los estados

Los modelos de predicción examinados hasta ahora pueden ser clasificados en dos categorías: modelos con parámetros y varianzas fijas, que exigen estacionariedad en media y varianza para todo el rango de datos y modelos con parámetros y varianzas variables, que se actualizan teniendo en cuenta los cambios en el patrón temporal.

Formalmente la representación en el espacio de los estados de un modelo temporal univariante consta de dos ecuaciones:

- a) La ecuación de transición o evolución que determina las variables estado en función de las entradas,  $S_t = CS_{t-1} + RV_t$ , donde:

$C$  es una matriz conocida,  $r \times r$ , denominada matriz de transición,

$S_t$  es un vector  $r \times 1$ , de variables estado

$R$  es una matriz conocida  $r \times r$  cuya inclusión es arbitraria

$V_t$  es un término de inputs o innovaciones  $r \times 1$  tal que  $V_t \approx NI(0, v_t)$  con  $v_t$  conocida

- b) La ecuación de medida que determina las salidas en función de las variables de estado y de un término de error de medida. En el caso univariante:

$$Y_t = Z^T S_t + u_t \text{ donde:}$$

$Z$  es un vector conocido  $r \times 1$

$u_t$  es un término de error tal que  $u_t \approx NI(0, \sigma_t^2)$  con  $\sigma_t^2$  conocida para  $t=1, 2, \dots, n$

$\sigma_t^2$  y  $v_t$  son constantes.

Tenemos así un sistema tempoinvariante que incluye todos los modelos estacionarios de series temporales univariantes. Por lo tanto cualquier modelo ARMA admite una representación tempoinvariante en el espacio de los estados.

Si consideramos la especificación general en el espacio de los estados de series temporales univariantes y tomamos  $Z^T = (1, 1)$  y  $S_t^T = (\mu_t, \gamma_t)$  que representan la componente de tendencia y la componente estacional respectivamente, tenemos el *modelo estructural básico de series temporales*:

$$Y_t = \mu_t + \gamma_t + u_t$$

y según las especificaciones que se hagan de las ecuaciones de transición (tendencia y componente estacional) da lugar a diferentes representaciones de series temporales<sup>29</sup>.

La ventaja de especificar un modelo en el espacio de los estados es que nos permite utilizar el filtro de Kalman, que permite obtener predicciones óptimas del vector de estado  $S_t$  para cada  $t$  con información disponible en cada momento  $t$ , combinando la predicción previa de  $S_t$  basada en informaciones pasadas y una nueva predicción del vector en base a nueva información. Se obtiene así una notable mejora de la predicción, período a período, a medida que se va añadiendo cada vez una nueva observación  $y_{t+h}$  ( $h=1\dots T$ ).

Aunque el filtro de Kalman se suele desarrollar con una formulación encuadrada dentro del espacio de los estados, con el propósito de ser más generales presentamos a continuación la notación de acuerdo con los desarrollos anteriores y posteriores del documento.

Si la predicción de la variable  $Y$  en  $t+h$ , puede realizarse en diferentes momentos o estadios ( $t+s$ ,  $s=0,..h$ ) siendo  $s=0$  el estadio donde se realiza la primera predicción, (esto es en  $t$ ) la predicción en cada estadio  $s$  varía de acuerdo con la última información disponible en el estadio anterior y la predicción en el estadio anterior ( $s-1$ ).

Así, si  $\hat{y}_{t+h,t+s}$  es la predicción de  $Y$ , para  $t+h$  en un momento  $t$  y en un estadio  $s$  y  $X_{t+s}$  es la última información disponible, la predicción para  $t+h$  en el siguiente estadio ( $t+s+1$ )  $\hat{y}_{t+h,t+s+1}$  puede ser expresada como una suma ponderada de  $X_{t+s}$  y  $\hat{y}_{t+h,t+s}$  obteniendo:  $\hat{y}_{t+h,t+s+1} = \omega X_{t+s} + (1-\omega)\hat{y}_{t+h,t+s}$ . Además la varianza se obtiene a partir de la expresión:  $\sigma^2 = (1-\omega)^2 \sigma_y^2 + \omega^2 \sigma_x^2$

Diferenciando respecto a  $\omega$  e igualando a cero para obtener un mínimo obtenemos:

$$\omega^* = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_x^2}$$

por lo que los pesos para combinar nueva y vieja información son función de las varianzas y la predicción cambia de acuerdo con la nueva información.

---

<sup>29</sup> También es posible representar modelos cíclicos  $Y_t = \mu + \Psi_t + u_t$  donde  $\Psi_t = A \cos(\lambda t + \phi)$  o modelos de tendencia y ciclo  $Y_t = \mu_t + \Psi_t + u_t$  y otros.

#### 1.2.2.4 Nuevas tendencias

Hasta ahora hemos asumido que el proceso que generaba los datos era lineal. Sin embargo, cuando esta hipótesis es falsa todavía muchos de los resultados básicos se mantienen<sup>30</sup>, luego los análisis precedentes siguen teniendo valor incluso en el caso de no linealidad.

Si el verdadero PGD no es lineal y estamos empleando métodos de predicción que suponen linealidad, se podrán esperar fallos en estimaciones y predicciones realizadas bajo ese patrón erróneo. Por estas razones, las tendencias más actuales en los métodos de predicción en series temporales están tomando nuevas direcciones:

- Un interés teórico en los sistemas no lineales y la posibilidad de que la no linealidad pueda generar diferencias en los comportamientos dinámicos. Esta es la literatura sobre la teoría del caos que en los últimos años ha proporcionado estudios que mejoran la capacidad para entender la presencia de sistemas caóticos (Fernández (1995)).
- El desarrollo de técnicas que permitan estimar sistemas no lineales desconocidos. Este es el caso de las redes neuronales<sup>31</sup>, técnica en la que no es preciso especificar la forma funcional entre inputs y outputs (estimación *free-model*) por lo que permite plantear una gran diversidad de modelizaciones (o arquitecturas) aprovechando su gran flexibilidad. Otra metodología alternativa que ha alcanzado auge en los últimos años es la ligada a sistemas borrosos, de especial interés cuando se desea procesar información cualitativa aportada por expertos. La conexión entre sistemas borrosos y neuronales es muy estrecha, dando lugar a los modelos neuroborrosos<sup>32</sup>.

#### 1.2.3 Modelos multivariantes de series temporales

Los métodos uniecuacionales de series temporales permiten capturar interesantes relaciones dinámicas con la única información proporcionada por los valores presentes y pasados de una variable. Sin embargo, no tienen en cuenta las relaciones causales entre variables ignorando por tanto información de gran importancia para la predicción ofrecida por otras características que puedan tener conexiones con la variable de

---

<sup>30</sup> Hall (1994) señala que el teorema de Wold se mantiene y que podemos pensar en los *ad hoc* de tendencia local como aproximaciones locales al verdadero proceso.

<sup>31</sup> Una descripción más detallada puede verse en Delgado (1997), donde además se realiza un estudio comparativo de este método con los modelos ARIMA y los métodos de alisado.

<sup>32</sup> Landajo, M. (1999) contiene un extenso análisis teórico de los modelos neuroborrosos, y aplicaciones de esta metodología a la predicción de series económicas.

interés<sup>33</sup>. Con el objetivo de tener en cuenta este tipo de situaciones, en los epígrafes que siguen desarrollamos brevemente el análisis de intervención y de función de transferencia, los modelos estocásticos multivariantes, los modelos VAR (vectores autorregresivos) y modelos de corrección de error (MCE).

### 1.2.3.1 Análisis de intervención y de función de transferencia

En los modelos de series temporales univariantes de una variable  $Y_t$  sólo se consideran como explicativos los valores pasados de la variable. En la práctica será frecuente localizar una o varias variables exógenas parcialmente explicativas de las variaciones de  $Y_t$ , mejorando las predicciones de  $Y_t$  al anticipar los cambios originados por las modificaciones de las variables  $X_t$ .

Así la serie  $Y_t$  quedaría explicada con el modelo

$$Y_t = a_0 + A(L)Y_{t-1} + C(L)X_t + B(L)a_t$$

donde la variable endógena  $Y_t$  depende linealmente de una serie de  $r$  variables exógenas  $X_t^d$  ( $d=1, \dots, r$  y que son variables con historia) y donde el polinomio  $C(L)$  es llamado función de transferencia.

La variable  $Y_t$  puede venir influida por cambios externos más o menos excepcionales de naturaleza cualitativa. En tal caso se define una variable ficticia que se denomina variable de intervención, que toma sólo dos valores:

$I_t=0$  antes y después del suceso

$I_t=1$  durante el suceso

con lo que el modelo anterior quedaría expresado como:

$$Y_t = a_0 + A(L)Y_{t-1} + c_0 I_t + B(L)a_t$$

---

<sup>33</sup> Ashley (1983) defiende la superioridad para la predicción de los modelos multivariantes frente a los univariantes siempre que las variables exógenas sean adecuadamente escogidas y analiza el efecto del error de predicción en las variables exógenas sobre el error de predicción en la predicción de la variable endógena. En una aplicación efectuada -Ashley (1988)- concluye que las predicciones a más de seis meses contienen errores de tal significación, que es mejor prescindir de ellas como fuente de información cuantitativa en la obtención de otras predicciones y utilizar modelos autoproyectivos que no precisan de dicha información.

Así por ejemplo, el modelo ARMA univariante se basa en que una serie temporal estacionaria  $Y_t$  puede ser generada aplicando una transformación a un ruido blanco<sup>34</sup>:

$$Y_t = \frac{\theta(L)}{\delta(L)} a_t$$

donde  $Y_t$  es considerado como un output producido al aplicar un filtro lineal a un input que en este caso es un ruido blanco, el cociente  $\frac{\theta(L)}{\phi(L)} = \Psi(L)$  es llamado función de transferencia del filtro.

Sin embargo aquí sólo se considera como variable explicativa de las variaciones de  $Y_t$  el ruido blanco  $a_t$ , junto a los propios valores retardados de  $Y$ . Si admitimos que la variable  $Y_t$  puede ser explicada además por variables exógenas y sus retardos, entonces puede descomponerse en dos términos aditivos:

$$Y_t = \hat{Y}_t + n_t$$

donde el término  $\hat{Y}_t$  puede explicarse exactamente mediante una o más variables explicativas  $X_t^d$  ( $d=1, \dots, r$ ):

$$\hat{Y}_t = \frac{\omega(L)}{\delta(L)} X_t^d = \sum_{d=1}^r \frac{\omega_d(L)}{\delta_d(L)} X_t^d = \sum_{d=1}^r v_d(L) X_t^d$$

siendo  $\omega(L)$  y  $\delta(L)$  polinomios de grados  $s$  y  $r$  respectivamente

y donde  $n_t$  es un término de error, que si es estacionario se representa mediante un modelo ARMA ( $p, q$ ), y en otro caso como un ARIMA:

$$n_t = \frac{\theta(L)}{\phi(L)} a_t$$

Por lo tanto el modelo general de función de transferencia se formula<sup>35</sup>:

<sup>34</sup> Una serie no estacionaria y por tanto modelizada como un ARIMA( $p, d, q$ ),  $\phi(L)\Delta^d y_t^{(\lambda)} = \theta(L)a_t$  permite realizar previsiones de  $\omega_t = \Delta^d Y_t^{(\lambda)}$  sin más que estimar los pesos de  $\omega_i$  del polinomio  $\Psi(L)$ ,  $\omega_t = \frac{\theta(L)}{\phi(L)} a_t = \Psi(L)a_t = \sum_{i=0}^{\infty} \psi_i a_{t-i}$ .

<sup>35</sup> Si las series  $Y_t$  y  $X_t^d$  no son estacionarias el modelo sería  $Y_t = \sum_{d=1}^r \frac{\omega_d(L)}{\delta_d(L)} X_t^d + \frac{\theta(L)}{\phi(L)} a_t$  con  $Y_t = (1-l)^d Y_t$  y  $X_t^d = (1-l)^d X_t^d$ . Si fuese necesario también se debería de hacer la diferenciación estacional.

$$Y_t = \sum_{d=1}^r \frac{\omega_d(L)}{\delta_d(L)} X_t^d + \frac{\theta(L)}{\phi(L)} a_t$$

Para el caso del análisis de intervención, donde la variable  $Y_t$  está influenciada exógenamente por una única variable que queda descrita por una dummy, el modelo se obtendría únicamente sustituyendo la variable  $X_t^d$  por  $I_t$ .

### 1.2.3.2 Modelo estocástico multivariante

Una generalización del modelo ARMA  $\phi(L)y_t = \theta(L)a_t$  es el modelo estocástico multivariante o MARMA, en el que se dispone de  $m$  series interrelacionadas:

$$Y_t^M = (Y_t^1, Y_t^2, \dots, Y_t^m)^T \text{ (mxn)} \quad \text{con} \quad t = 1 \dots n$$

y se constituye el modelo dinámico:

$$H(L)Y_t^M = F(L)A_t$$

en el que  $H(L)$  y  $F(L)$  son matrices de dimensión  $m \times m$

$$H(L) = \|h_{ij}(L)\| \quad F(L) = \|f_{ij}(L)\|$$

representando  $m$  ecuaciones en retardos sobre las  $m$  series temporales y siendo  $A_t$  un ruido blanco  $g$ -dimensional.

El modelo MARMA aprovecha las interrelaciones entre las variables que componen el vector  $Y_t^M$ , mejorando las predicciones obtenidas con un modelo ARMA uniecuacional.

Al igual que ocurría con el modelo estocástico univariante de series temporales, el modelo estocástico multivariante podría representar la solución del sistema en términos de errores. Así, la serie  $Y_t^M$  se puede obtener a partir del ruido blanco  $A_t$  mediante el filtro multivariante:

$$Y_t^M = H(L)^{-1} F(L)A_t = \Psi(L)A_t$$

La generalización de un modelo de función de transferencia al caso multivariante da lugar a la formulación de un modelo econométrico multiecuacional dinámico en el que existen  $r$  variables exógenas y  $g$  variables endógenas y a partir del cual y procediendo de igual manera que en el caso de series univariantes (pero con matrices) se obtiene la función de transferencia<sup>36</sup>:

<sup>36</sup> La deducción puede verse en Pulido (1993) y en Zellner y Palm (1974) siguiendo a Quenouille, M. H. (1957): *The analysis of Multiple Time Series*. Por otra parte, Charles Griffith emplea la ecuación

$$Y_t^M = D(L)^{-1}W(L)X_t + H(L)^{-1}F(L)A_t$$

donde las series estacionarias contenidas en  $X_t$ ,  $X_t = (X_t^1, X_t^2, \dots, X_t^r)$  pueden ser representadas mediante un modelo MARMA a efectos de predicción de las variables endógenas.

### 1.2.3.3 Modelos VAR y MCE

Otro modelo multivariante útil para predecir varias variables  $Y_t^M$  que están correlacionadas entre sí es el modelo de tipo VAR, propuesto inicialmente por Sims (1980), en el que no se distingue entre variables endógenas y exógenas.

Siguiendo el planteamiento de Sims cada variable influye en las demás y es influida por las otras, por lo que en cada ecuación aparece despejada una variable diferente en función de todas y cada una de las demás variables del sistema, siempre retardadas y de sus propios retardos. De este modo, las  $g$  variables consideradas anteriormente como endógenas serán función tanto de las  $r$  variables exógenas  $X_t = (X_t^1, X_t^2, \dots, X_t^r)$  como de las variables endógenas retardadas.

$$Y_t^M = A_{t-1}Y_{t-1}^M + \dots + A_{t-p}Y_{t-p}^M + BX_t + E_t$$

Las matrices  $A_1, A_2, \dots, A_r$  y  $B$  se estiman por mínimos cuadrados, bajo la hipótesis de no existencia de autocorrelación (supuesto que resulta muy restrictivo debido a la estructura de retardos).

Con el objetivo de evitar el riesgo de modelizar correlaciones espúreas entre las variables (es decir, situaciones en las que la causa que origina la correlación entre variables es que éstas no son estacionarias y presentan tendencias similares) surge el análisis de cointegración que conduce a los *modelos de corrección de error* (MCE o VEC).

Una relación de cointegración entre variables puede interpretarse como la existencia de una relación lineal de equilibrio entre ellas, dada por un vector de cointegración. Más concretamente, Engle y Granger –premios Nobel de Economía 2003- definen el concepto de cointegración de la siguiente forma:

Los componentes de un vector  $Y_t(mx1)$  se dice que están cointegrados de órdenes  $d$  y  $b$  y se denota por  $Y_t \approx CI(d,b)$  si:

---


$$Y_t = V(L)X_t + \phi(L)^{-1}\theta A_t$$

definiendo previamente  $V(L)$  como una matriz de  $g \times k$  polinomiales del tipo  $\omega_{ij}(L)/\delta_{ij}(L)$ .

1. Todos los componentes de  $Y_t^M$  son integrables del mismo orden<sup>37</sup>  $d$ ,  $I(d)$
2. Existe un vector  $\alpha$ , no nulo tal que  $\alpha_t^T Y_t^M = z_t \approx I(d-b)$ , con  $b > 0$ . A este vector se le denomina *vector de cointegración*.

Así, si admitimos esa relación de cointegración, trabajamos con un VAR restringido, que incorpora restricciones de cointegración dentro de su especificación y que se transforma en un modelo de corrección de error. En Granger (1981) se muestra la equivalencia entre MCE y cointegración demostrando que un conjunto de variables cointegradas pueden modelizarse mediante un MCE y, a la inversa, si la especificación de un MCE es correcta, entonces existe una relación de cointegración entre las variables implicadas. Esta equivalencia sería más tarde formalizada por Engle y Granger (1987) y se la conoce como *Teorema de Representación de Granger*. Así, si la ecuación de cointegración es:  $Y_t^1 = \beta Y_t^2$ , entonces el modelo de corrección de error (MCE)<sup>38</sup> vendría dado por:

$$\begin{aligned}\Delta Y_t^1 &= \gamma_1 (Y_{t-1}^2 - \beta Y_{t-1}^1) + a_t^1 \\ \Delta Y_t^2 &= \gamma_2 (Y_{t-1}^2 - \beta Y_{t-1}^1) + a_t^2\end{aligned}$$

Intuitivamente, la relación entre los modelos VAR y MCE es inmediata: Si las variables  $Y_t^M$  no son estacionarias, entonces el VAR deberá ser reescrito en diferencias. Si además se verifica una relación de cointegración en el sistema, entonces una combinación lineal de las variables deberá de ser añadida en cada ecuación del VAR, resultando por tanto adecuado un modelo MCE.

### 1.3 Técnicas bayesianas de predicción: los puntos de giro.

Los puntos de giro constituyen acontecimientos especialmente relevantes en la evolución económica, dado que su presencia señala cambios de fase en la coyuntura económica cuya predicción resulta de crucial interés. En este sentido, el uso de algún indicador adelantado ( $I_t^A$ ) puede ser de gran ayuda, puesto que éste presenta sus puntos de giro, picos (máximos relativos) y valles (mínimos relativos) antes que el nivel

---

<sup>37</sup> En el caso más simple si dos variables  $I(1)$  están cointegradas, entonces una combinación de ellas es  $I(0)$ . Estos autores ponen de manifiesto que una combinación lineal de dos o más series no estacionarias puede ser estacionaria.

<sup>38</sup> Existen otras especificaciones de los MCE, dependiendo de que las variables endógenas tengan o no tendencia, exista una constante en la ecuación de integración, etc. Granger y Swanson (1995) extienden el estudio de los MCE para series económicas no lineales, obteniendo MCE no lineales.

general de actividad económica. La última observación del indicador adelantado  $I_t^A$ , siendo  $I_t^A = (i_t^A, i_{t-1}^A, \dots, i_1^A)$ , se convierte en una predicción de los puntos de inflexión que  $Y_t$  tendrá en  $t+h$   $y_{t+h}^p$ , esto es:  $i_t^A = \hat{y}_{t+h,t}^p$  (entonces  $I_t^A = \hat{Y}_t^p$ ).

El objetivo de este apartado es estudiar cómo estos indicadores son capaces de anticipar los puntos de giro en la actividad económica, tratando de calcular cuál es la probabilidad de que ocurra un punto de giro. En concreto se presentan las probabilidades mensuales secuenciales recursivas de pico y valle, determinando previamente la densidad probabilística que ajusta los datos observados.

No obstante, antes de explicar cómo se obtienen dichas probabilidades, es necesario describir la metodología para determinar los puntos de giro de una serie económica mediante el empleo de técnicas empíricas o no paramétricas:

Estos métodos se basan en el empleo de filtros lineales y la aplicación de reglas de decisión que los analistas han venido utilizando para detectar los puntos de inflexión en la economía y que reflejan su experiencia práctica acumulada. Básicamente, estos procedimientos utilizan en grados diversos una combinación de técnicas de filtrado lineal, para obtener una señal cíclica exenta de irregularidad, y de algoritmos de búsqueda para identificar los máximos y los mínimos cíclicos. Así la determinación de los puntos de giro se inicia en una serie suavizada y se aproxima, en sucesivas etapas, a fechar los picos y valles en la serie original ajustada por estacionalidad cumpliendo siempre:

1. Que la distancia entre dos puntos de giro del mismo tipo sea como mínimo de  $L$  meses (duración del ciclo).
2. Que la distancia entre dos puntos de giro de distinto tipo sea como mínimo de  $M$  meses (duración de una fase de crecimiento o decrecimiento).

Si bien existe una amplia variedad de procedimientos no paramétricos, los planteamientos en los que éstos se basan son similares, diferenciándose en cambio las distintas técnicas en los filtrados empleados y en las restricciones propuestas<sup>39</sup>.

Una vez que se han detectado los puntos de giro del indicador adelantado, es posible analizar en qué medida éste es capaz de anticipar dichos puntos de giro en la

---

<sup>39</sup> Los métodos empíricos son subjetivos y en consecuencia la determinación de puntos de giro es una forma útil para cada analista de clasificar sus observaciones. Existen también procedimientos objetivos de carácter paramétrico, de los que puede verse una recopilación en Abad et al. (2000).

economía. El análisis teórico se basa en el modelo utilizado por Neftci (1982) y Diebold y Rudebusch (1989).

El método consiste en observar un proceso, cuya estructura cambia de forma aleatoria en algún momento del tiempo y predecir el instante en el que se presenta el cambio súbito. Interesa establecer el cambio de régimen del indicador para predecir los cambios en la actividad económica real  $Y_{t+h}$ . Así, las observaciones de  $I_t^A$  serán usadas para realizar la inferencia que permita predecir el momento en que la economía cambia de régimen.

Se supone que la actividad económica real  $Y_{t+h}$  tiene dos distribuciones distintas  $f_{t+h}^u(Y)$  y  $f_{t+h}^d(Y)$  correspondientes a las distribuciones en el régimen de expansión y de recesión respectivamente. Dichas distribuciones se estiman a partir de las observaciones pasadas de  $Y_t$  que denominaremos respectivamente  $f_t^u(Y)$  y  $f_t^d(Y)$ . Un punto de giro en la economía se define como el momento del tiempo en el que cambia la distribución de probabilidad de  $Y_{t+h}$  y el objetivo es predecir los puntos de giro de  $I_t^A$  con la idea de que los hechos que producen un cambio de giro en esa serie están presentes antes que en  $Y_{t+h}$ .

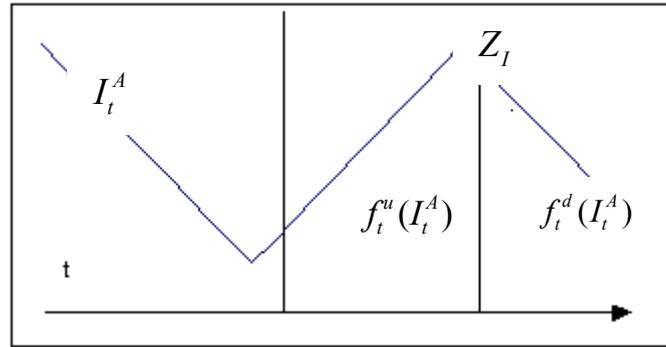
Denotaremos por  $Z_t$  al número entero que representa el primer periodo posterior a un punto de giro (pico) del indicador adelantado. Entonces el indicador adelantado  $I_t^A$  tiene dos distribuciones correspondientes a recuperaciones y contracciones que denotamos por  $f_t^u(I_t^A)$  y  $f_t^d(I_t^A)$  respectivamente:

$$f_t^u(I_t^A) \text{ cuando } t < Z_t$$

$$f_t^d(I_t^A) \text{ cuando } t \geq Z_t$$

y que serán estimadas a partir de las observaciones pasadas en el indicador ( $I_{t-1}^A$ ).

**Figura 1: Puntos de giro**



Conocida la última información disponible hasta  $t$  del indicador adelantado,  $I_{t-1}^A$ , la probabilidad de observar un pico en  $t$ , correspondiente a un anticipo del pico en la actividad económica futura  $Y_{t+h}$ ,  $Z_I \leq t$ , puede ser expresada como:

$$P_t = P(Z_I \leq t | I_{t-1}^A) = \frac{P(\hat{y}_{t+h,t}^p / Z_I \leq t) P(Z_I \leq t)}{P(I_{t-1}^A)}$$

que es la probabilidad a posteriori de un punto de giro dados los datos disponibles del indicador adelantado hasta el momento  $t$ .

Es posible obtener una fórmula recursiva de esta probabilidad a posteriori que para el caso de un pico viene dada por:

$$P_t = \frac{\left[ (P_{t-1} + \Gamma_t^u (1 - P_{t-1})) \right] f_t^d(\hat{y}_{t+h,t}^p | I_{t-1}^A)}{\left[ (P_{t-1} + \Gamma_t^u (1 - P_{t-1})) \right] f_t^d(\hat{y}_{t+h,t}^p | I_{t-1}^A) + (1 - P_{t-1}) f_t^u(\hat{y}_{t+h,t}^p | I_{t-1}^A) (1 - \Gamma_t^u)}$$

Para el cálculo de la probabilidad secuencial recursiva en los valles la fórmula será:

$$V_t = \frac{\left[ (V_{t-1} + \Gamma_t^d (1 - V_{t-1})) \right] f_t^u(\hat{y}_{t+h,t}^p | I_{t-1}^A)}{\left[ (V_{t-1} + \Gamma_t^d (1 - V_{t-1})) \right] f_t^u(\hat{y}_{t+h,t}^p | I_{t-1}^A) + (1 - V_{t-1}) f_t^d(\hat{y}_{t+h,t}^p | I_{t-1}^A) (1 - \Gamma_t^d)}$$

donde  $\Gamma_t^u = P(Z_I = t | Z_I > t)$  y  $\Gamma_t^d = P(Z_I = t | Z_I \leq t)$  son las probabilidades de que un punto de giro se presente en el momento  $t$  sabiendo que no se presentó previamente, denominadas probabilidades de transición de un régimen de expansión a uno de contracción en el primer caso (pico) y de una contracción a una expansión en el segundo (valle),  $f_t^u(\hat{y}_{t+h,t}^p | I_{t-1}^A)$  y  $f_t^d(\hat{y}_{t+h,t}^p | I_{t-1}^A)$  son las densidades de probabilidad del

indicador adelantado según la última observación provenga de una situación de recuperación o de contracción y  $P_{t-1}$  y  $V_{t-1}$  son las probabilidades calculadas para la observación anterior del indicador adelantado.

Los valores  $f^u(\hat{y}_{t+h,t}^p | I_{t-1}^A)$  y  $f^d(\hat{y}_{t+h,t}^p | I_{t-1}^A)$  se obtienen a partir de la estimación de la distribución de las observaciones pasadas  $f_t^u(I_{t-1}^A)$  y  $f_t^d(I_{t-1}^A)$  que son nuevamente estimadas al finalizar un ciclo.

De acuerdo con la evidencia empírica,  $\Gamma_t^u$  y  $\Gamma_t^d$  no dependen del tiempo (Diebold y Rudebusch (1989)) y por ello se les suele asignar como valores originales la razón del número de puntos de giro sobre la suma de la duración de todas las fases correspondientes.

Cuando nos encontramos en un punto de crecimiento en la economía, se calcula la probabilidad mensual secuencial recursiva de un pico (fin de la expansión)  $P_t$  y cuando supera un valor crítico establecido previamente concluimos que se presenta un pico en el momento  $t=Z_t$ , y se puede relacionar con un inminente punto de giro (en  $t+h$ ) en el crecimiento económico<sup>40</sup>, es decir  $\hat{y}_{t+h,t}^p$  es un predictor en  $t$  de los puntos de giro en  $y_{t+h}^p$ . Del mismo modo cuando estemos en un periodo de decrecimiento se calcula la probabilidad mensual secuencial recursiva de un valle (inicio de la expansión).

Cuando la probabilidad mensual es mayor que una probabilidad crítica fijada por el analista, se produce una señal de pico o valle, según la información contenida en la probabilidad secuencial recursiva. Cada vez que la probabilidad supera la probabilidad crítica y no acontece el punto de giro, se produce una falsa señal.

Se trata entonces de una predicción cuantitativa probabilística que adopta como referencia la metodología bayesiana<sup>41</sup>, al transformar distribuciones de densidad a priori de una variable aleatoria  $Y_{t+h}$ ,  $f_{Y_{t+h}}(y_{t+h})$  si es continua o  $P_{Y_{t+h}}(y_{t+h})$  si es discreta, en otras a posteriori ( $f_{Y_{t+h}}(y_{t+h} / I_t)$  o  $P_{Y_{t+h}}(y_{t+h} / I_t)$ ) a partir de la información suministrada en  $t$  por las distribuciones de probabilidad de otra variable aleatoria  $I$  (funciones de verosimilitud,  $f_{I_t}(i_t)$  o  $P_{I_t}(i_t)$ ):

<sup>40</sup> El valor  $P_{t-1}$  es nulo cuando se empieza a predecir un pico. Cuando se ha predicho y se busca otro punto de giro (pico) vuelve a valer cero y lo mismo ocurre para  $V_{t-1}$ .

<sup>41</sup> En el apartado 3 de este trabajo se explica con más detalle la metodología bayesiana aplicada a la combinación de predicciones.

$$f_{Y_{t+h}}(y_{t+h} | I_t) = \frac{f_{Y_{t+h}}(y_{t+h})f_{I_t}(i_t)}{\int_{Y_{t+h}} f_{Y_{t+h}}(y_{t+h})f_{I_t}(i_t)d(y_{t+h})} \text{ o } P_{Y_{t+h}}(y_{t+h} | I_t) = \frac{P_{Y_{t+h}}(y_{t+h})P_{I_t}(i_t)}{\sum_{Y_{t+h}} P_{Y_{t+h}}(y_{t+h})P_{I_t}(i_t)}$$

En este caso las distribuciones a priori son  $f_t^k(I_t^A)$ , las funciones de verosimilitud  $f_t^k(I_{t-1}^A)$ , y las distribuciones a posteriori  $f_t^k(I_t^A | I_{t-1}^A) = f_t^k(\hat{y}_{t+h,t}^p | I_{t-1}^A)$ , ( $k=u, d$ ).

#### 1.4 Técnicas cualitativas de predicción: las Encuestas de Opinión

Además de las predicciones de tipo cuantitativo, existen también técnicas de predicción basadas en las valoraciones sobre el futuro que de su entorno realizan los distintos agentes económicos. Estas técnicas son cualitativas, puesto que únicamente cada agente conoce el proceso a partir del cual elabora su opinión, si bien la información que emplea cada agente puede ser cuantitativa y/o cualitativa (y compartirla o no) y la valoración revelada cualitativa o cuantitativa.

Dentro de los métodos cualitativos de expectativas cualitativas están las encuestas de opinión, si bien el empleo de estas encuestas para la predicción está limitado por las cuestiones concretas que se plantean en cada una, por las posibilidades de respuesta, por la periodicidad con la que se realizan y por el horizonte de predicción en el cual los individuos revelan sus opiniones sobre la variable a predecir.

Las encuestas de opinión se pueden clasificar en tres tipos siendo distinta además su utilidad para la predicción: encuestas de intenciones, de actitudes y de expectativas.

Las encuestas de intención suministran información sobre las intenciones de los individuos. Sin embargo, la revelación de una intención puede no verse realizada efectivamente o aplazarse, y de ahí su limitación como herramienta para la predicción.

Las encuestas de actitudes y expectativas suelen referirse a una valoración *ex ante* que los individuos efectúan sobre variables generales de su entorno o sobre variables de su propio dominio. En el primer caso, las encuestas suministrarían valoraciones de los individuos acerca del entorno económico futuro, y al conocerse de alguna forma las opiniones generales de una colectividad éstas terminarían condicionando actitudes o intenciones más concretas, con lo que su utilidad en la predicción es importante. En el segundo caso, parece evidente la importancia de estas encuestas a efectos de predicción,

al suministrar información inmediata sobre expectativas de compra o inversión, de consumidores y empresarios respectivamente<sup>42</sup>.

En concreto, en todas las encuestas las cuestiones que se plantean suelen referirse a situaciones en las que se requiere conocer las expectativas para un periodo del tiempo muy reducido (máximo un año), lo que las convierte en herramientas idóneas para las predicciones a corto plazo.

Parece entonces que, a efectos de la predicción, las encuestas sobre actitudes y expectativas sean las más relevantes, por lo que en esta parte del trabajo el análisis quedará limitado a este tipo de encuestas. En los apartados que siguen describiremos sus características más generales, intentando tipificar sus contenidos y posteriormente nos centraremos en la Encuesta de Coyuntura Industrial que es la de mayor tradición y uso más generalizado en España.

#### **1.4.1 Encuestas de Actitudes y Expectativas de Empresarios y Consumidores**

La realización de las Encuestas de Actitudes y Expectativas de Empresarios y Consumidores (EAEEC), incluidas en las generalmente denominadas Encuestas Cualitativas de Empresarios y Consumidores (ECEC)<sup>43</sup> constituye un complemento muy útil de las encuestas estadísticas cuantitativas, al proporcionar información complementaria sobre aspectos no cubiertos por alguna de ellas.

En las EAEEC se pide a los administradores de empresas y a los consumidores, en tanto que agentes económicos que operan en el mercado, que den su opinión sobre ciertas variables económicas clave como, por ejemplo, decisiones de compra, nivel de inventarios, evolución de la producción, empleo, utilización de capacidades, inversión, innovación, precios, etc.

Por lo general, las variables son de tres tipos diferentes:

- realizaciones (datos *ex post*)

---

<sup>42</sup> Sin embargo, a pesar de la utilidad de este tipo de encuestas, la experiencia indica que éstas no están exentas de limitaciones, ya que los individuos (consumidores o empresas) pueden no tener incentivos para desvelar sus propias creencias. El hecho de que los individuos no sean “expertos predictores”, en la materia de que se trate no debería de considerarse un problema, ya que lo que influye en sus decisiones futuras son sus expectativas y no lo correcto o incorrecto de las mismas, tal y como señalan Estrada y Urtasum (1998).

<sup>43</sup> El vigente programa de ECEC armonizadas para toda la Unión Europea comprende una encuesta mensual sobre la industria, una encuesta mensual sobre la construcción, una encuesta mensual sobre el comercio minorista, desglosada en seis ramas, una encuesta semestral de la actividad inversora en la industria y una encuesta mensual de consumidores. En Schönborn (1997) se puede encontrar una descripción detallada del conjunto de encuestas enmarcadas dentro del programa de Encuestas Cualitativas de Empresarios y Consumidores armonizadas para toda la Unión Europea.

- valoraciones o estimaciones de la situación vigente
- planes o expectativas (datos *ex ante*)

Las respuestas más comunes son: *baja, normal o alta* si reflejan el nivel actual, *disminuyó, se mantuvo o aumentó*, si reflejan valoraciones sobre la evolución percibida y *disminuir, mantenerse o aumentar* si prevén la tendencia inmediata.

Cuando un individuo contesta a preguntas referentes a planes o expectativas sobre una variable Y (producción, empleo u otra), su respuesta puede considerarse una predicción para la variable Y sobre la que se les pregunta. Esto es, su opinión  $o_t^j$  se convierte en una predicción  ${}^e \hat{y}_{t+h,t}^j$ , y cada individuo (que genera sus expectativas en base a sus experiencias, conjeturas e intuiciones y en información disponible exclusiva o no) se convierte entonces en una herramienta cualitativa de predicción al ser el único conocedor del proceso a partir del cual procesa la información para generar su expectativa:  $g^j(I_t)$ .

En este caso y debido a las limitaciones impuestas por las tres únicas posibles respuestas en las encuestas,  $Y_{t+h}$  puede tomar 3 valores  $Y_{t+h} = (y_{t+h}^1, y_{t+h}^2, y_{t+h}^3)$ , disminuir, mantenerse y aumentar respectivamente, de acuerdo con la tendencia inmediata. Por consiguiente, los individuos manifestarán sus expectativas de acuerdo con esos posibles valores, es decir, los individuos escogen su predicción dentro de la triada posible de predicciones para Y:  $\hat{Y}_{t+h,t} = (\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \hat{y}_{t+h,t}^3)$ , disminuir, mantenerse y aumentar respectivamente y su elección puede ser considerada ya como una predicción cualitativa  ${}^e \hat{y}_{t+h,t}^j$ . Más concretamente denotaremos la opción escogida como  ${}^e \hat{y}_{t+h,t}^j(r)$ , siendo  $r=1,2,3$  según las posibles respuestas.

Normalmente, los resultados de las encuestas suelen difundirse como series en las que se agregan (ponderadas o no) las respuestas individuales  ${}^e \hat{y}_{t+h,t}^j$  en cada categoría (r). Dichas series pueden emplearse, por separado, combinadas<sup>44</sup> o resumidas en forma de indicadores compuestos.

Los indicadores de síntesis más destacados, que sirven de guía (pronóstico) de la actividad empresarial real, son los siguientes:

---

<sup>44</sup> En este caso, consideraremos el indicador como una predicción combinada obtenida a partir de predicciones cualitativas o expectativas individuales. En el capítulo 4 estudiaremos con detalle estos aspectos de la predicción.

- El indicador de confianza o clima de la industria, obtenido a partir de las expectativas de producción, las carteras de pedidos y los inventarios.
- El indicador de confianza o clima de la construcción, obtenido a partir de las carteras de pedidos y las expectativas de empleo.
- El indicador de confianza de los consumidores, obtenido a partir de la situación financiera de los hogares y la valoración de la situación económica general.
- El indicador de percepción de la economía, obtenido a partir de los indicadores de confianza de la industria, la construcción, los consumidores y el índice de cotización de las acciones.

#### **1.4.2 La Encuesta de Coyuntura Industrial**

Dentro de este tipo de encuestas en España hay que hacer referencia a la Encuesta de Coyuntura Industrial (ECI), que elabora con carácter mensual y trimestral el Ministerio de Ciencia y Tecnología (MCYT) desde 1963, con el objetivo de conocer las opiniones de los empresarios del sector industrial sobre la situación económico-industrial reciente y sus perspectivas a corto plazo<sup>45</sup>.

La muestra mensual de la ECI es un panel (muestra fija) de establecimientos, obtenido mediante la técnica de muestreo aleatorio estratificado. Actualmente, la muestra de esta encuesta es de alrededor de 2.700 empresas<sup>46</sup>, de las cuales alrededor del 90% corresponden a la encuesta en la industria, en sentido estricto, y el 10% restante a la encuesta en el sector de la construcción.

En concreto, el formulario va dirigido al personal gerencial de las empresas industriales y recopila básicamente información de dos tipos:

- Cualitativa: opiniones sobre los niveles actuales de la cartera de pedidos y de la producción y sobre la tendencia de los precios de venta, el empleo y la producción para los próximos meses.
- Cuantitativa: opiniones relativas al período de trabajo asegurado con su cartera de pedidos (en días) y el empleo actual (en personas).

---

<sup>45</sup> Las variables a las que los empresarios deben responder cada mes son, en términos de situación –nivel- y de expectativas de evolución a corto plazo (no más de tres meses) –tendencia.

<sup>46</sup> La muestra de la *Encuesta de Coyuntura Industrial (ECI)* es un panel de empresas. No obstante, cuando una empresa tiene varias líneas de producción en establecimientos diferenciados, éstos se consideran unidades informantes independientes

La ECI agrupa a las empresas industriales por subsectores en función de la actividad productiva (única o principal) de acuerdo con la clasificación de actividades CNAE-93.

Normalmente, las series cronológicas se obtienen agregando ponderadamente las respuestas individuales  ${}^e \hat{y}_{i+h,t}^i$  en cada categoría.

Con el incremento del nivel de agregación se obtienen los datos de los sectores industriales, clasificados por destino económico en bienes de consumo, intermedios, de inversión y total de la industria.

Una vez ponderadas adecuadamente las respuestas de las empresas que comprenden un subsector o subsectores que forman un sector industrial debe cumplirse que la suma del valor adjudicado a cada una de las tres opciones de respuesta a una variable o pregunta cualitativa debe ser 100.

Los resultados para cada variable pueden utilizarse directamente o bien resumidos en forma de indicadores compuestos<sup>47</sup>.

Generalmente, se suele emplear como indicador de cada variable su saldo, obtenido como diferencia entre los porcentajes de respuesta optimistas y pesimistas (Anderson (1952)). A partir de estos indicadores se construyen unos índices o niveles que en función de su magnitud y signo aportan una valiosa información sobre la evolución coyuntural de cada sector industrial y, por extensión, del conjunto de la industria española.

Uno de los indicadores de síntesis más destacados<sup>48</sup> es el *Indicador de Clima Industrial (ICI)* - o Indicador de Confianza Industrial, como se le denomina en el ámbito comunitario- que como ya hemos señalado representa el mayor o menor optimismo de los empresarios del sector industrial respecto a la situación actual de la actividad industrial y sus perspectivas de evolución futura. En concreto, el ICI se obtiene como media aritmética de los saldos netos obtenidos de los indicadores del nivel actual de la cartera de pedidos total, las expectativas de producción corregidas de variaciones estacionales y el nivel de existencias de productos terminados (cambiado de signo).

---

<sup>47</sup> La difusión de los resultados de la ECI se realiza en el sitio web [www.mcyt.es/coyuntura\\_eci/](http://www.mcyt.es/coyuntura_eci/). Además de los datos agregados para el total industrial y su comparación con los principales datos de la UE, también se ofrecen resultados por sectores clasificados por el destino económico de los bienes producidos (bienes de consumo, inversión e intermedios).

<sup>48</sup> Puesto que la ECI agrupa las empresas industriales por subsectores en función de la actividad productiva, es posible elaborar otros indicadores climáticos por sectores. Así, a partir de la Encuesta de Coyuntura para la Construcción se elabora el Indicador Climático de Construcción (ICC).

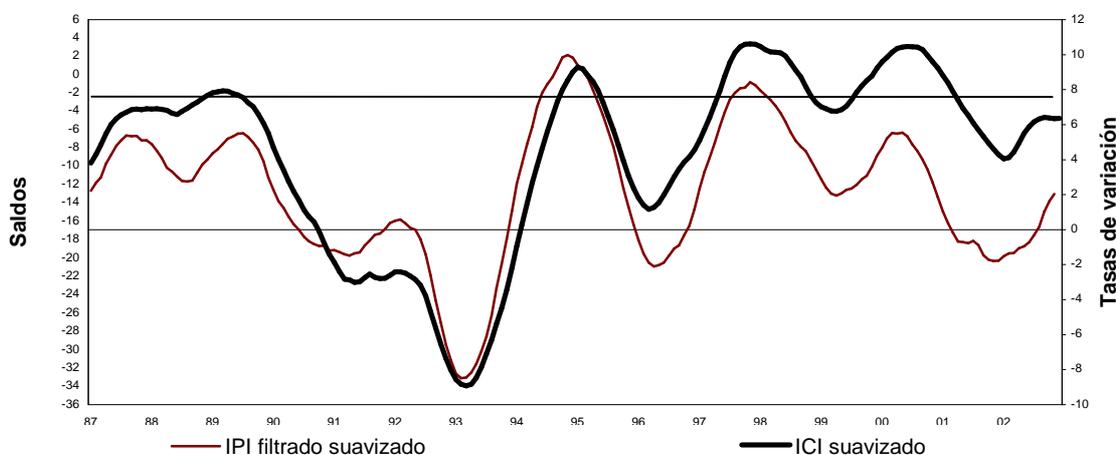
Los análisis disponibles indican que el ICI es un indicador útil para el diagnóstico de la coyuntura futura del sector industrial, dado que las variables que contiene son indicadores adelantados de la actividad empresarial (cuando el saldo de la cartera de pedidos y la producción previstas sean elevados el indicador reflejará una previsión positiva de la actividad industrial y cuando el ICI tome valores negativos reflejará una tendencia depresiva en la economía).

El ICI muestra una elevada correlación con el *Índice de Producción Industrial (IPI)*, lo que le configura como indicador adelantado de la actividad industrial.

En concreto, la figura 2 muestra las series suavizadas del ICI<sup>49</sup> y las tasas interanuales del IPI<sup>50</sup>, apreciándose que el ICI de un mes puede ser considerado como una predicción de la tasa del IPI del mes siguiente, y también que los puntos de inflexión en el ICI anticipan los del IPI.

**Figura 2: Evolución comparada del ICI y el IPI (filtrado). 1987-2002**

**Saldos y tasas de variación interanual suavizados**



**Fuente: Ministerio de Ciencia y Tecnología e INE**

<sup>49</sup> Las series utilizadas en el gráfico son los saldos originales del ICI, y las tasas de variación interanuales del IPI corregido de efectos calendario, lo que se denomina IPI filtrado, y ambas suavizadas mediante un regresor de orden 2 (AR-2)

<sup>50</sup> Gómez y Bengoechea (2000) indican que en la estimación del componente cíclico de una serie temporal no deben utilizarse las tasas de variación interanuales debido a la aparición de ciclos espúreos. Así, si se pretende realizar un análisis cíclico más heterodoxo para ambas series sería conveniente utilizar, en el caso del IPI filtrado, otro filtro diferente al de la tasa de variación anual. Los autores proponen varios filtros para captar la señal cíclica de las series, entre los que destaca un filtro *pasabanda* que logra estimar unos ciclos más suaves y potencia el estudio de máximos, mínimos y puntos de giro de las series.

Otro indicador adelantado que se obtiene de los datos de la ECI es el referido a la tendencia de los precios de venta, que se publica unos 25 días antes que otro de los indicadores coyunturales industriales de mayor interés para los analistas económicos por sus repercusiones sobre los precios de consumo, el *Índice de Precios Industriales (IPRI)*.

Así, la estrecha relación existente entre los indicadores sintéticos obtenidos de la ECI y el IPI y el IPRI que elabora el INE, permite la obtención con un mes de adelanto, de una serie de indicadores que anticipen señales sobre la marcha a corto plazo de la industria española.

### **1.5 Algunas fuentes de predicción sobre crecimiento económico en España**

El indudable interés que tiene para los distintos agentes económicos disponer de predicciones sobre el crecimiento económico futuro ha impulsado a numerosos organismos e instituciones a elaborar periódicamente informes de prospectiva. La difusión de estos resultados ha mejorado sustancialmente gracias al desarrollo de las TIC, encontrándose en la actualidad muchas de estas predicciones disponibles en Internet<sup>51</sup>.

Hasta ahora nos hemos centrado en describir las técnicas y encuestas más apropiadas para la predicción a corto plazo. No obstante, debido a la característica de la variable a predecir, generalmente el crecimiento económico, la mayoría de los organismos elaboran predicciones a medio plazo, con lo que las herramientas más apropiadas de predicción serían los modelos econométricos. La información disponible, las variables consideradas y la forma de modelizar su relación con la variable de interés puede ser diferente para cada organismo, pudiendo existir cierta heterogeneidad entre las predicciones económicas elaboradas.

En los apartados que siguen hacemos referencia a algunos de los organismos que elaboran predicciones cuantitativas para el crecimiento económico a medio plazo en España y a otras encuestas de opinión españolas cuya periodicidad las hace apropiadas en la elaboración de predicciones cualitativas.

---

<sup>51</sup> Una recopilación de instituciones que publican sus predicciones en Internet junto con las correspondientes direcciones puede verse en Pulido (2001a).

### 1.5.1 Fuentes de predicción cuantitativa

En España existen numerosos organismos tanto públicos como privados que realizan regularmente predicciones cuantitativas sobre variables macroeconómicas. Así, entre las instituciones públicas oficiales se encuentran el Ministerio de Economía, la Dirección General de Previsión y Coyuntura o el Banco de España, mientras los centros privados abarcan tanto entidades financieras nacionales (FUNCAS<sup>52</sup>, BBVA, Banesto, La Caixa, Caja Madrid, etc) como otras vinculadas a organizaciones financieras internacionales (J.P.Morgan, Morgan-Stanley, Goldman-Sachs) o empresariales (CEOE, Consejo Superior de Cámaras de Comercio, IEE, AFI, ...). A ambos tipos de organismos debemos añadir también algunas iniciativas ligadas a centros universitarios o de investigación como CEPREDE, Hispalink, ICAE o el Instituto Flores de Lemus.

Cabe además señalar que, además de estos organismos nacionales, existen numerosas instituciones internacionales que realizan predicciones sobre variables macroeconómicas para un conjunto de países entre los que está España. Entre ellas se encuentran organismos oficiales (FMI, Comisión Europea, OCDE), centros universitarios o de investigación (Proyecto Link de Naciones Unidas<sup>53</sup>) y recopilaciones difundidas en publicaciones de ámbito internacional (The Economist, Consensus Economics, ..).

En ocasiones las predicciones disponibles son agrupadas en forma de panel, siendo posible elaborar comparaciones entre organismos, predicciones de consenso e incluso estudiar su comportamiento de acuerdo con el estadio en el que se realiza la predicción. A modo de ilustración, reproducimos el último panel que para el crecimiento económico de España (tasas interanuales) publica el instituto L.R. Klein en su boletín *Tendencias* (Junio 2003).

---

<sup>52</sup> La Fundación de Cajas de Ahorros Confederadas para la Investigación Económica y Social (FUNCAS) publica un panel de previsiones sobre la economía española que unifica y consensúa las cifras de trece instituciones españolas, tanto públicas como privadas, incluyendo ocho financieras, tres centros universitarios y dos centros empresariales. Esta información se encuentra actualmente disponible en <http://www.funcas.ceca.es>.

<sup>53</sup> El Proyecto Link (<http://www.chass.utoronto.ca/link> y <http://www.un.org/es/analysis.link>) proporciona predicciones macroeconómicas a partir de un panel de 80 centros mundiales de predicción.

Tabla2

FUENTES	CRECIMIENTO				
	2003			2004	
	Hace 6 meses	Hace 3 meses	Actual	Hace 3 meses	Actual
The Economist (Jun. 03)	2,4	2,2	2,0	2,9	2,7
Instituto L.R. Klein - Gauss (Jun. 03)	2,5	2,5	2,1	3,0	3,2
La Caixa (May. 03)	2,5	2,3	2,3	2,9	2,9
Caja Madrid (May. 03)	2,4	2,4	2,4	-	-
FUNCAS (May. 03)	2,5	2,3	2,3	3,1	3,1
Morgan Stanley (May. 03)	2,1	1,7	1,8	2,6	2,7
Santander Central Hispano (May. 03)	2,5	2,5	2,1	2,7	2,7
Consensus Economics (May. 03)	2,5	2,3	2,2	2,9	2,9
Instituto de Crédito Oficial (May. 03)	2,7	2,4	2,5	3,0	3,1
Insti. de Estudios Económicos (May. 03)	2,6	2,6	2,6	-	-
Instituto Flores de Lemus (May. 03)	2,6	2,4	2,4	3,0	3,0
Dresdner Bank (May. 03)	2,4	2,2	2,2	2,7	2,7
UBS Warburg (May. 03)	2,2	2,2	2,0	2,5	2,5
AFI (May. 03)	2,4	2,2	2,2	-	-
BNP Paribas (May. 03)	2,3	2,1	2,1	2,8	2,8
ICAE (May. 03)	2,5	2,3	2,3	-	-
Goldman Sachs (May. 03)	2,2	2,1	2,1	3,3	3,3
Intermoney (Abr. 03)	2,5	2,0	2,0	2,2	2,2
BBVA (Abr. 03)	2,5	2,5	2,5	-	-
ONU - Proyecto Link (Abr. 03) (1)	2,8	-	2,0	-	2,6
FMI (Abr. 03)	2,7	-	2,2	-	3,1
OCDE (Abr. 03) (1)	2,5	-	2,1	-	3,1
Comisión Europea (Abr. 03)	2,6	-	2,0	-	3,0
Gobierno (Ene. 03) (1)	3,0	3,0	3,0	-	-
PROMEDIO (Jun. 03)	2,4	2,3	2,2	2,9	2,9

### 1.5.2 Fuentes de predicción cualitativa

Si bien la Encuesta de Coyuntura Industrial es la principal fuente de predicción cualitativa para el crecimiento económico en España a corto plazo, existen otras estadísticas de interés cuya periodicidad y contenido las convierte en herramientas útiles para elaborar predicciones a medio plazo sobre aspectos relacionados con el crecimiento económico (inversión, empleo, consumo). Estas son por ejemplo la Encuesta de Inversiones Industriales que elabora el Ministerio de Ciencia y Tecnología, la Encuesta de Empresas Industriales del INE y la Encuesta de Coyuntura Laboral elaborada por el Ministerio de Trabajo y Asuntos Sociales. Por su parte, el Centro de Investigaciones Sociológicas (CIS) elabora encuestas sobre diversas cuestiones de la esfera económica, que pueden ser de interés para el seguimiento del ciclo económico.

La Encuesta de Inversiones en la Industria (EII) recoge cada semestre información cuantitativa sobre la inversión anual realizada y planeada por las empresas industriales, y cualitativa sobre la finalidad de la misma, así como los factores que están incidiendo en ella. La muestra a la que se dirige la encuesta se compone de alrededor de 2.200

empresas, sensiblemente inferior a la muestra a la que va dirigida la Encuesta de Coyuntura Industrial básicamente porque la unidad muestral es la empresa industrial<sup>54</sup> (aunque tenga varios establecimientos diferenciados) y porque los sectores energético-extractivos, actualmente no se investigan en la encuesta<sup>55</sup>.

La encuesta consiste en dos cuestionarios que se envían a las empresas cada año, en primavera y en otoño. El cuestionario de primavera permite obtener la tasa de variación de la inversión en el año anterior y una estimación con respecto a la tasa de variación de la inversión en el año en curso. Por su parte, el cuestionario de otoño permite obtener una estimación relativamente segura con respecto a la tasa de variación de la inversión en el año en curso y, en segundo lugar, una primera estimación sobre los planes de inversión para el año siguiente.

La Encuesta Industrial de Empresas, implantada a partir de 1993 con periodicidad anual, está orientada al estudio de las principales variables estructurales de los distintos sectores industriales, lo que permite disponer de una información básica para el conocimiento de la realidad industrial y para el análisis de sus principales características estructurales. La población objeto de estudio de la encuesta es el conjunto de empresas industriales con una o más personas ocupadas remuneradas, y cuya actividad principal figura incluida en las Secciones C a E de la CNAE-93. Abarca 100 sectores industriales dentro de todo el territorio nacional, a excepción de Ceuta y Melilla.

La Encuesta de Coyuntura Laboral, de periodicidad trimestral y anual<sup>56</sup>, tiene como objetivo obtener información sobre el mercado de trabajo desde la óptica de las empresas<sup>57</sup>. Sus fines son múltiples y variables dependiendo de la periodicidad de la

---

<sup>54</sup> Se entiende por empresa toda unidad jurídica que constituye una unidad organizativa de producción de bienes y servicios, y que disfruta de una cierta autonomía de decisión, principalmente a la hora de emplear los recursos corrientes de que dispone. La empresa puede ejercer una o más actividades en uno o varios lugares. Se considera que una empresa es industrial, a efectos de la encuesta, si su actividad principal está incluida dentro de las Secciones C a E de la Clasificación Nacional de Actividad Económicas (CNAE-93).

<sup>55</sup> No obstante, y como menciona López Serrano (2003), se tiene previsto la implementación de una nueva aplicación informática que ampliará tanto la investigación a los sectores mencionados como la información ofrecida, para el total industrial, en términos de dimensión empresarial.

<sup>56</sup> La ECL fue implantada en el segundo trimestre de 1990, experimentando una profunda modificación metodológica en el primer trimestre de 1997. Adicionalmente, desde 1998, la Encuesta de Coyuntura Laboral investiga, con periodicidad anual, en un módulo cuyo contenido es variable, otros aspectos complementarios a los recogidos trimestralmente.

<sup>57</sup> El directorio utilizado en la Encuesta es el Fichero de Cuentas de Cotización a la Seguridad Social, el marco lo constituyen las unidades de este directorio que pertenecen al ámbito investigado por la Encuesta y la unidad a investigar es la cuenta de cotización, que es, asimismo, la unidad elemental del citado fichero; dicha unidad está constituida por un colectivo de trabajadores de una misma empresa homogéneo frente a la cotización y cuya actividad laboral se realiza en una misma provincia. Aunque

encuesta, pudiéndose destacar como más importantes los siguientes: proporcionar información sobre los efectivos laborales y su composición según diversas características, evaluar la movilidad laboral, altas, bajas y modificaciones de contrato, de los efectivos laborales, cuantificar la jornada laboral efectivamente realizada, las horas no trabajadas y sus motivaciones y las horas extraordinarias, investigar determinados aspectos de las relaciones laborales y conocer las opiniones de los empresarios respecto a la evolución de sus plantillas.

En concreto, en el informe metodológico de la encuesta publicado por el Ministerio de Trabajo y Asuntos Sociales se definen las expectativas sobre la variación de efectivos laborales, como las previsiones del empresario sobre la evolución de los efectivos laborales de su centro con referencia al final del trimestre siguiente y con referencia al final de igual trimestre del año siguiente. Dichas previsiones se expresan como porcentaje esperado de incremento o disminución de los efectivos laborales con respecto a su plantilla. Con esta información se obtienen previsiones sobre la variación de efectivos laborales en cifras absolutas y en porcentaje<sup>58</sup> útiles para predicciones económicas al proporcionar información de los cambios previstos de tendencia en la evolución del empleo en sus empresas y de la intensidad de los mismos<sup>59</sup>.

Como complemento a las previsiones empresariales respecto a la evolución futura del empleo, desde 1999 se solicita, anualmente, la opinión de los empresarios relativa a determinados aspectos que permitan anticipar su comportamiento ante diversas situaciones o expectativas relacionadas con el mercado de trabajo, como las posibles medidas que se adoptarían en una empresa ante un hipotético aumento de la demanda o los factores que pueden influir en la decisión de realizar nuevas contrataciones.

El Centro de Investigaciones Sociológicas (CIS) elabora encuestas sobre diversas cuestiones de la esfera económica (mayoritariamente son encuestas dirigidas a la población de 18 años y más, realizadas mediante entrevista personal en los domicilios, con un tamaño muestral de 2.500 unidades).

---

conceptualmente no existe identidad entre establecimiento o centro de trabajo y cuenta de cotización, en la práctica ambos conceptos coinciden en un alto porcentaje de casos.

<sup>58</sup> En concreto la *Previsión del empresario sobre la variación en valores absolutos de los efectivos laborales* se calcula multiplicando el porcentaje de crecimiento o disminución de los efectivos laborales consignado por el empresario por el número de trabajadores del centro y la *Previsión del empresario sobre la variación porcentual de efectivos laborales* se calcula multiplicando el porcentaje de crecimiento o disminución de los efectivos laborales consignado por el empresario por la proporción que representan los trabajadores del centro respecto al total de efectivos laborales

<sup>59</sup> Pérez Corrales (2001) menciona que las tendencias obtenidas de la evolución prevista por los empresarios muestran una gran sintonía con la realidad observada posteriormente aunque las previsiones ponen de manifiesto la cautela de los empresarios respecto a la cuantía de estas estimaciones.

La información disponible es muy variada al contar con datos de variables relacionadas con el consumo, el desarrollo económico, las desigualdades, la economía sumergida, el papel del Estado en la economía, la valoración de la situación económica, etc.

En el ámbito del seguimiento del ciclo económico, Martínez (2001) indica que se ha comprobado que medidas del sentimiento o la opinión pública pueden predecir comportamientos económicos futuros y su impacto, posiblemente más rápido y de forma más precisa que otros tipos de datos. Una de las series disponibles en el CIS es por ejemplo, la valoración prospectiva de la situación económica de España (próximo año), contemplando como categorías de respuesta mejor, igual y peor.

Existen otras encuestas de opinión en España tales como la de Comercio al por Menor, la encuesta a los consumidores y la encuesta de opinión entre empresas exportadoras.

La Encuesta de Comercio al por Menor es llevada a cabo desde 1986 por la Comisión de las Comunidades Europeas, e incluye preguntas sobre la evolución pasada y esperada de ventas, existencias, compras a proveedores y empleo.

Por su parte, la Encuesta a los Consumidores también elaborada desde 1986 por la Comisión de las Comunidades Europeas, recoge información sobre su situación financiera, la situación económica del país, el coste de la vida, la realización de compras de bienes duraderos, el ahorro, desempleo, etc.

La Encuesta de Opinión entre Empresas Exportadoras es elaborada con carácter trimestral desde 1995 por el Ministerio de Economía con el objeto de obtener información sobre la coyuntura del comercio exterior. Las variables analizadas son cartera de pedidos, precios de exportación, precios internacionales, margen comercial exterior y su relación con el interior y actividad exportadora. En esta encuesta, únicamente en el caso de la cartera de pedidos existe una pregunta relativa a la evolución esperada de la misma con un horizonte de tres y doce meses.

## 2 MEDIDAS DE EVALUACIÓN DE PREDICCIONES

Como ya hemos visto, existe una amplia diversidad de técnicas de predicción, cuya idoneidad varía dependiendo de la situación a predecir, pudiendo además existir técnicas igualmente óptimas para una misma situación.

La evaluación de las predicciones obtenidas puede ayudarnos a decidir qué método individual utilizar o cuáles pueden emplearse en la combinación, siendo habitual el uso de indicadores basados en la precisión de las predicciones que tienen como objetivo la medición de la desutilidad o el coste asociado a los pares de predicciones y realizaciones.

Dependiendo del método empleado y de las características concretas que queramos predecir de una variable económica (su valor, la probabilidad de ese valor, los puntos de giro, etc), los pronósticos pueden tener diversa naturaleza (cuantitativa, cualitativa y probabilística), y entonces el par realización-predicción y los indicadores de precisión podrán adoptar formas diferentes que será necesario tipificar.

En este capítulo, una vez realizada esa tipificación, estudiamos las medidas tradicionales de evaluación de predicciones y examinamos nuevas alternativas basadas en medidas de información, que resultan adecuadas para evaluar las predicciones. Para ello adoptamos como referencia las medidas de incertidumbre e inquietud cuadráticas, proponiendo a partir de ellas algunos indicadores de imprecisión e información.

### 2.1 Tipificación de la precisión y de las medidas de evaluación

Dada una magnitud  $Y$  para la cual realizamos predicciones en cierto horizonte temporal  $h$  de amplitud  $T$  ( $h=1...T$ ), el análisis de su precisión será llevado a cabo según el tipo de predicción considerada:

- Cuantitativas:  $\hat{y}_{t+h,t}$ , tendremos un vector de  $T$  predicciones sobre las que “*evaluar la capacidad de la técnica para la predicción*”
- Cualitativas:  $\hat{y}_{t+h,t}^m$ , donde  $m$  hace referencia a la categoría nominal ( $M=1,...,m$ ) predicha para  $Y$  (clasificada también en las mismas categorías), dispondremos de una muestra ( $T$ ) de predicciones y realizaciones para “*evaluar la capacidad de una técnica para la predicción*”.

La clasificación conjunta se puede recoger en una tabla de contingencia como la siguiente:

**Tabla 3**

	PREDICCIONES				
REALIZACIONES	$\hat{Y}^1$	$\hat{Y}^2$	... $\hat{Y}^j$ ...	$\hat{Y}^m$	Total
$Y^1$	$n_{11}$	$n_{12}$	...	$n_{1m}$	$n_{1.}$
$Y^2$	$n_{21}$	$n_{22}$	...	$n_{2m}$	$n_{2.}$
$Y^i$	...	...	$n_{ij}$	...	$n_{i.}$
$Y^m$	$n_{m1}$	...	...	$n_{mm}$	$n_{m.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.j}$	$n_{.m}$	T

donde  $n_{ij}$  representa el número de las T observaciones clasificadas para la predicción en la categoría  $\hat{Y}^j$  y para la realización en la categoría  $Y^i$  ( $i, j = 1, \dots, m$ ).

Desde un punto de vista típicamente estadístico es más adecuado pensar en términos de la población de la que se supone que ha sido extraída dicha muestra y por tanto los valores  $n_{ij}$  de cada celda se modifican por las probabilidades conjuntas, que denotaremos por  $p_{ij}$ .

**Tabla 4**

	PREDICCIONES				
REALIZACIONES	$\hat{Y}^1$	$\hat{Y}^2$	.. $\hat{Y}^j$ ..	$\hat{Y}^m$	Total
$Y^1$	$p_{11}$	$p_{12}$	...	$p_{1m}$	$p_{1.}$
$Y^2$	$p_{21}$	$p_{22}$	...	$p_{2m}$	$p_{2.}$
$\vdots$ $Y^i$ $\vdots$	...	...	$n_{ij}$	...	$p_{i.}$
$Y^m$	$p_{m1}$	...	...	$p_{mm}$	$p_{m.}$
Total	$p_{.1}$	$p_{.2}$	$p_{.j}$	$p_{.m}$	1

- Probabilísticas:  $f_{Y_{t+h}}(y_{t+h} / I_t)$  (si Y es continua) o  $P_{Y_{t+h}}(y_{t+h} / I_t)$  (si es discreta), cuya esperanza es la predicción cuantitativa (que llamaremos  ${}^P E_{t+h,t}$ ), tendremos T distribuciones de probabilidad a posteriori o esperanzas para “*evaluar la capacidad de la técnica bayesiana para la predicción*”

La *evaluación de la predicción generada por una técnica* suele realizarse a partir del grado de precisión aportado por la predicción y los indicadores que se emplean habitualmente se basan en esa precisión, con el objetivo de medir la desutilidad o el coste asociado a los pares de predicciones y realizaciones  $(P_{t+h,t}, R_{t+h})$  que serán:

$(y_{t+h}, \hat{y}_{t+h,t})$  para predicciones cuantitativas

$(p_i, p_j)$  para predicciones cualitativas

$(y_{t+h}, {}^P E_{t+h,t})$  para predicciones probabilísticas

La relación existente entre el par realización-predicción, y la función (L) para medir el coste asociado a cada par pueden adoptar formas estadísticas sofisticadas, si bien la forma más habitual es definirla a partir del error de predicción:

$$L(P_{t+h,t}, R_{t+h}) = L(e_{t+h}).$$

Dependiendo entonces de cómo se defina el error de predicción y la función L que procesa la información, tendremos diferentes medidas de precisión para *evaluar la capacidad de predicción*. Así pues, la tipificación de ambos puede ayudarnos a escoger la medida de precisión que queramos emplear.

La tabla 5 resume las expresiones de los errores para las predicciones cuantitativas, cualitativas y probabilísticas:

**Tabla 5: Tipificación del error de predicción**

	<b>Predicciones cuantitativas</b>	<b>Predicciones cualitativas</b>	<b>Predicciones probabilísticas</b>
<b>Error</b>	$e_{t+h} = y_{t+h} - \hat{y}_{t+h,t}$	$e_{ij(i=j)} = p_i - p_j$	$e_{t+h} = y_{t+h} - {}^P E_{t+h,t}$
<b>Error absoluto</b>	$e_{t+h}^A =  y_{t+h} - \hat{y}_{t+h,t} $	$e_{ij(i=j)}^A =  p_i - p_j $	$e_{t+h}^A =  y_{t+h} - {}^P E_{t+h,t} $
<b>Error relativo</b>	$e_{t+h}^R = \frac{y_{t+h} - \hat{y}_{t+h,t}}{y_{t+h}}$	$e_{ij(i=j)}^R = \frac{p_i - p_j}{p_i}$	$e_{t+h}^R = \frac{y_{t+h} - {}^P E_{t+h,t}}{y_{t+h}}$

A la hora de realizar la evaluación no se suele disponer de las observaciones en  $t+h$ , por lo que generalmente la evaluación se realiza a partir de los errores pasados ( $t=1, \dots, n$ ) o de forma general ( $t=1, \dots, T$ ).

- Tipificación de L

Dependiendo de la expresión de la función L podemos tener una amplia batería de medidas que nos permiten contrastar hipótesis sobre los errores (tales como ausencia de sesgo o correlación), analizar la precisión, ver de qué depende esa precisión (si disponemos por ejemplo de información referente a los estadios de predicción o sobre las características de variables que se relacionan con la variable de interés a través del modelo predicción, ...) etc.

Sin embargo, mientras en los últimos años han aparecido numerosas contribuciones teóricas en cuanto a técnicas de predicción (análisis de cointegración, redes neuronales, sistemas neuro-borrosos, metodología input-output), en los análisis aplicados sigue resultando habitual la utilización de las medidas convencionales basadas en los errores de predicción (Error Cuadrático Medio, Error Absoluto Medio) y del índice de desigualdad de Theil.

## 2.2 Contraste de insesgamiento de las predicciones individuales

Si las predicciones son óptimas, entonces deberían cumplir las siguientes propiedades:

1. Los errores de predicción deben tener media cero (insesgadas).
2. Los errores de predicción un paso adelante deben ser ruido blanco (incorrelados).
3. Los errores de predicción se deben distribuir normalmente.

Generalmente, las propiedades 1 y 2 se suelen contrastar conjuntamente (y entonces nos referiremos a la *eficiencia* de las predicciones), para verificar la propiedad 3, se suelen realizar estudios teóricos o contrastar únicamente la simetría de la distribución. No obstante en el capítulo 5 describimos la regresión por cuantiles para ajustar los errores empíricos de predicción que emplearon Taylor y Bunn (1999) con el objetivo de producir modelos del error de predicción por cuantiles que son funciones de  $h$  (horizonte temporal de la predicción).

En este epígrafe, nos centraremos en estudiar la hipótesis de insesgadez, y/o autocorrelación de los errores según las predicciones sean cuantitativas, cualitativas o probabilísticas.

### **Predicciones cuantitativas**

Uno de los análisis más comúnmente empleados para el estudio del sesgo<sup>60</sup> y su corrección en futuras predicciones es el propuesto por Theil (1971) a partir de una regresión de las realizaciones ( $Y_t$ ) sobre las predicciones efectuadas en los periodos pasados  $\hat{Y}_t$ ,  $t=1, \dots, n$ .

$$Y_t = \alpha + \beta \hat{Y}_t + v_t$$

Los coeficientes estimados  $\hat{\alpha}$  y  $\hat{\beta}$  permiten analizar el sentido y la magnitud del sesgo, y efectuar correcciones en las predicciones futuras, tal y como veremos más adelante. Si las predicciones fuesen insesgadas se debería de cumplir  $\hat{\alpha} = 0$ ,  $\hat{\beta} = 1$ , requisito que Holden y Peel (1990) han demostrado que es condición suficiente pero no necesaria de insesgadez.

Estos autores proponen estimar el modelo  $e_t = v + \varepsilon_t$ , donde  $e_t = y_t - \hat{y}_t$  es el error de predicción y  $v$  es su valor medio. El contraste del supuesto de insesgadez equivale entonces a contrastar la nulidad del valor medio  $v$ , por lo que un rechazo de esta última hipótesis nos llevaría a concluir que las predicciones son sesgadas, indicando el parámetro estimado el signo y la magnitud del sesgo<sup>61</sup>.

Para series no estacionarias, Holden y Peel (1989) proponen un contraste alternativo que se basa en ajustar la ecuación:

$$Y_t - \hat{Y}_t = \alpha + v_t \quad \text{o} \quad Y_t - \hat{Y}_t = \gamma + \lambda X_{t-1} + u_t$$

donde  $X_{t-1}$  es el conjunto de información rezagada y  $u_t$  y  $v_t$  son los errores del proceso.

---

<sup>60</sup> Inicialmente se pueden considerar gráficos de los errores de predicciones los cuales informan sobre sus magnitudes, y posibles sesgos. Además se pueden detectar posibles errores atípicos, señalados estos por una alta curtosis.

<sup>61</sup> Si por ejemplo el valor estimado es positivo y significativamente diferente de cero, entonces las previsiones efectuadas son inferiores a los valores finales de la magnitud, y por tanto habría que hacer una revisión al alza en posteriores previsiones.

Si  $\alpha$  es nulo, entonces la predicción es insesgada. Si  $v_t$  no está serialmente correlacionado<sup>62</sup> (dado el horizonte de predicción), la predicción es *débilmente eficiente*. Si  $\gamma$  y  $\lambda$  son cero entonces la predicción es *fuertemente eficiente*.

Existen también algunos procedimientos no paramétricos para contrastar la hipótesis de insesgadez de las previsiones a través de una función que indica si el error de predicción es positivo:

$$\begin{aligned} f(e_t) &= 1 & e_t &\geq 0 \\ f(e_t) &= 0 & e_t &< 0 \end{aligned}$$

La hipótesis de insesgadez supondría que los errores se distribuyen simétricamente con media nula y bajo este supuesto el estadístico definido como

$S = \sum_{t=1}^T f(e_t)$  sigue una distribución binomial  $B(T, 0,5)$ , que para muestras grandes

puede ser aproximada a un modelo Normal:  $\frac{S - \frac{T}{2}}{\sqrt{\frac{T}{4}}} \approx N(0,1)$ .

Este test conducirá al rechazo de la hipótesis de insesgadez de las predicciones si el número de observaciones con error de predicción positivo es significativamente distinto al número de observaciones con error de predicción negativo.

Si las predicciones son sesgadas, se pueden emplear diferentes métodos estadísticos para corregir los errores sistemáticos en las predicciones futuras. En concreto, los métodos de corrección más empleados son los desarrollados por Theil (1971): El modelo de Corrección Lineal y el modelo de Corrección Adaptativo.

En la corrección lineal la predicción es corregida extrayendo su error esperado. Sin embargo, bajo condiciones de cambio, este método puede reducir la precisión de la predicción, por lo que puede ser conveniente aplicar una variante del modelo de Theil, el modelo de corrección adaptativo que pondera más las observaciones recientes<sup>63</sup>.

Theil muestra que una corrección lineal óptima puede ser empleada para extraer ambos sesgos a partir de un conjunto de predicciones pasadas (y por tanto con

---

<sup>62</sup> El análisis de la correlación serial de los errores puede realizarse a través del correlograma y un contraste de hipótesis sobre la significación de las autocorrelaciones puede ser realizado usando el contraste de Ljung y Box, si bien este contraste exige normalidad en los errores de predicción.

<sup>63</sup> Goodwin (1997) analiza los dos métodos en un experimento sobre predicciones cuantitativas generadas por cien expertos.

realizaciones conocidas) adoptando como base los parámetros estimados en la regresión.

$$Y_t = \hat{\alpha} + \hat{\beta} \hat{Y}_t$$

La predicción corregida para el periodo t+h sería entonces:

$$\hat{y}_{t+h,t}(c) = \hat{\alpha} + \hat{\beta} \hat{y}_{t+h,t}$$

pudiendo ser empleada para la corrección de T predicciones futuras, lo que implícitamente supone predecir el error de la propia predicción:

$$\hat{e}_{t+h} = \hat{\alpha} + (\hat{\beta} - 1) Y_{t+h}$$

Si los residuos del modelo de regresión tienen una distribución conocida, entonces es posible conocer la probabilidad de que la precisión de una predicción se incremente con la corrección<sup>64</sup>:

Un incremento de la precisión ocurrirá cuando:

$$y_{t+h} > 0,5(\hat{y}_{t+h,t}(c) + \hat{y}_{t+h,t}) \text{ cuando } \hat{y}_{t+h,t}(c) > \hat{y}_{t+h,t}$$

o

$$y_{t+h} < 0,5(\hat{y}_{t+h,t}(c) + \hat{y}_{t+h,t}) \text{ cuando } \hat{y}_{t+h,t}(c) < \hat{y}_{t+h,t}$$

En otro caso el error de predicción aumentará o no cambiará con la corrección.

En el método de corrección adaptativo se emplea el filtro de Kalman para obtener estimaciones recursivas de los parámetros del modelo  $Y_t = \hat{\alpha}_t + \hat{\beta}_t \hat{Y}_t$ , permitiendo aplicar diferentes factores de descuento a cada parámetro que ponderen más las observaciones más recientes. Los factores de descuento que han de ser aplicados a los parámetros son especificados en una matriz W diagonal. De este modo, en t la matriz de varianzas covarianzas de los parámetros de la regresión en t-1 ( $Q_{t-1}$ ) se calcula como:  $WQ_{t-1}W$ .

Así por ejemplo un valor de  $w=1+\delta$  sobre la estimación del  $\alpha_t$  en la regresión  $Y_t = \hat{\alpha}_t + \hat{\beta}_t \hat{Y}_t$  tendría el efecto de aumentar la varianza asociada a ese parámetro por un factor  $(1+\delta)^2$ , lo que significa reducir el valor de la información disponible en t-1. Si  $\delta$  aumenta la información pasada se descuenta aún mas. Así lo más usual sería que el

---

<sup>64</sup> Así por ejemplo si los residuos sigan un modelo normal,  $y_{t+h}$  seguirá una distribución t-student con media  $\hat{y}_{t+h,t}(c)$  por tanto hay una probabilidad del 50% de que mejore la predicción. Esta mejoría puede disminuir en el caso de que el predictor “coja experiencia” y la distancia entre la predicción original y la corregida disminuya.

valor de  $w$  estuviese comprendido entre 1 (no descuento) y 1,2 (porque si es mayor solo dependería de 2 o 3 periodos recientes).

Este método requiere la especificación inicial de los parámetros  $\alpha$  y  $\beta$  y de sus varianzas. Si se asume desconocimiento total se puede suponer que los parámetros son nulos y sus varianzas muy grandes. En el caso de que no haya descuento, el método sería igual al de MCO con actualización temporal.

### **Predicciones cualitativas**

Generalmente las predicciones cualitativas son valoraciones efectuadas por agentes que participan en la economía, por lo que la hipótesis de insesgadez se suele llamar “*Hipótesis de Expectativas Racionales (REH)*”

Para contrastar esta hipótesis es posible cuantificar las expectativas y aplicar los métodos del anterior epígrafe, o bien aplicar directamente contrastes<sup>65</sup> sobre las expectativas cualitativas.

Estos tests requieren un panel de información con al menos dos periodos para contrastar la eficiencia y tres periodos para contrastar la ausencia de correlación.

Gourieroux-Pradel (1986) proponen llevar a cabo un test directo sobre las tablas de contingencia: “*La hipótesis de expectativas racionales es satisfecha si y solo si:*  $P_{jj} \succ P_{ij}$ ”

Sin embargo este test no detecta el signo del sesgo (optimismo o pesimismo), existiendo otros indicadores que sí pueden ayudarnos a detectar si los individuos tienden a sobrevalorar o infravalorar las realizaciones.

Cohen (1968) y Light (1971) definen coeficientes de acuerdo entre dos clasificaciones como:

$$\kappa_{ii} = \frac{P_{ii} - P_i.P_i}{P_{.j} - P_i.P_i}$$

A partir de estas expresiones, Pichery (1989) propone una extensión para contrastar REH y el sentido del sesgo. Más concretamente, calcula los coeficientes de desacuerdo condicional correspondientes a pares con  $i$  en expectativas y  $j$  para realizaciones:

$$\kappa_{ij} = \frac{P_{ij} - P_i.P_j}{P_{.j} - P_i.P_j}$$

---

<sup>65</sup> Entre los posibles contrastes se encuentran los de Nerlove (1981); Gourieroux y Pradel (1986); Ivaldi (1990).

Cada coeficiente varía entre un límite inferior,  $k_{ij} \text{ inf} = k_{ii} \text{ inf} = -p_i / (1 - p_i)$ , que es siempre negativo pero no constante, y un límite superior igual a 1. Con el objetivo de tener coeficientes acotados entre 0 y 1 se realiza una normalización:

$$k_{ij}^* = \frac{k_{ij} + |k_{ij} \text{ inf}|}{1 + |k_{ij} \text{ inf}|} = \frac{p_{ij}}{p_{.j}}$$

cumpléndose entonces que la *hipótesis de expectativas racionales es satisfecha* si para cada elemento j, el elemento de la diagonal principal es mayor o igual que cada elemento de la misma fila asociada con la expectativa del elemento j.

El resultado sería el mismo si contrastamos

“*La hipótesis de expectativas racionales es satisfecha si y solo si:  $k_{jj}^* > \max k_{ij}^*$ ”*

Cuando hay información disponible, basta con ver que el gráfico de  $k_{jj}^*$  esté por encima de lo otros. En caso afirmativo, entonces es posible aceptar la hipótesis de REH<sup>66</sup>.

A partir de este coeficiente se pueden obtener coeficientes sintéticos tales como el coeficiente de acuerdo  $I_a = \sum_{i=1}^m k_{ii}^*$  o dos coeficientes de desacuerdo, uno cuando es positivo, es decir cuando la realización es mejor que la expectativa y otro negativo para el caso opuesto. Para calcularlos es necesario que las m categorías sean clasificadas en orden decreciente (de crecimiento a decrecimiento) y entonces

$$I_p = \sum_{i=1}^m \sum_{j < i}^m k_{ij}^*$$

$$I_n = \sum_{i=1}^m \sum_{j > i}^m k_{ij}^*$$

$I_p$  es calculada sobre los elementos encima de la diagonal principal (las realizaciones son mejores que las expectativas, en términos de utilidad) y  $I_n$  sobre los elementos debajo de la diagonal principal.

Un desacuerdo positivo aparece asociado con el caso en el que la realización es mejor que la expectativa que se tenía (por ejemplo crecimiento superior al esperado), lo que puede ser interpretado como una actitud pesimista a priori de los empresarios en

---

<sup>66</sup> Se puede hacer otro test empleando esta información que corresponde a un juicio a posteriori y es condición necesaria y suficiente para expectativas correctas:  $k_{ii}^* > \max k_{ij}^*$

torno a la variable a predecir. En caso de un desacuerdo negativo, el coeficiente revela una actitud optimista a priori.

A nivel agregado tres indicadores son los más habitualmente empleados para resumir la situación de optimismo (IO), pesimismo (IP) o expectativas correctas (CC) y son calculados a partir de la suma (manteniendo el orden anterior en la tabla de contingencia) de la suma del triángulo por debajo de la matriz, por encima y la diagonal respectivamente

$$IO = \sum_{j=1}^m \sum_{i<j}^m p_{ij}$$

$$IP = \sum_{i=1}^m \sum_{i>j}^m p_{ij}$$

$$CC = \sum_{i=1}^m p_{ii}$$

Finalmente los coeficientes de desacuerdo, respectivamente positivos o negativos, corresponderán a las diferencias de 1 categoría ( $I_{dn}(1)$  y  $I_{dp}(1)$ ), o dos o más ( $I_{dn}$  e  $I_{dp}$ ). Cuyos valores están entre 0 y 1

$$I_a = \frac{1}{m} \sum_{i,j=1}^m k_{ij}^*, \quad I_{dp(1)} = \frac{1}{m-1} \sum_{i=2}^m k_{ii-1}^*, \quad I_{dn(1)} = \frac{1}{m-1} \sum_{i=1}^{m-1} k_{ii+1}^*$$

$$I_{dp} = \frac{1}{(m-2)(m-3)} \sum_{i=3}^m \sum_{j=1}^{i-2} k_{ij}^* \quad I_{dn} = \frac{1}{(m-2)(m-3)} \sum_{i=1}^{m-2} \sum_{j=i+1}^k k_{ij}^*$$

$$I_d = \frac{1}{m(m-1)} \sum_{i=2}^m \sum_{j=1}^{i-1} k_{ij}^* + \frac{1}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=1+1}^m k_{ij}^*$$

### Predicciones probabilísticas

Cabe por último señalar que cuando las predicciones son probabilísticas, se pueden aplicar los mismos test que para las predicciones cuantitativas, puesto que

$${}^P E_{t+h,t} = \hat{y}_{t+h,t}$$

### 2.3 Medidas basadas en el error de predicción.

En este apartado estudiamos las medidas tradicionalmente empleadas en la evaluación de la precisión de las predicciones, según su naturaleza sea cuantitativa, cualitativa o probabilística.

### 2.3.1 Predicciones cuantitativas

Las medidas más directas de la bondad de las predicciones se basan en los errores de las predicciones y vienen dadas por las siguientes expresiones:

*Raíz del error cuadrático medio (ECM)<sup>67</sup>:*

$$\sqrt{ECM} = \frac{1}{T} \sqrt{\sum_{t=1}^T e_t^2} = \frac{1}{T} \sqrt{\sum_{t=1}^T (y_t - \hat{y}_t)^2}$$

*Error absoluto medio (EAM):*

$$EAM = \frac{1}{T} \sum_{t=1}^T |e_t| = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|$$

*Error absoluto porcentual medio (EAPM):*

$$EAPM = \frac{\sum_{t=1}^T \frac{|y_t - \hat{y}_t|}{y_t}}{T} 100$$

Estas medidas presentan varias limitaciones: así, las dos primeras dependen de las unidades de medida de la variable investigada, y además ninguna de ellas se encuentra acotada ni tiene en cuenta la dificultad inherente a cada predicción.

Con el objetivo de elaborar una medida invariante con respecto a las unidades usadas y que además tenga en cuenta los problemas que rodean a la predicción, H. Theil<sup>68</sup> (1966) propone una medida de desigualdad dada por la expresión:

$$U = \sqrt{\frac{\sum_t (P_t - A_t)^2}{\sum_t A_t^2}}$$

donde  $P_t$  y  $A_t$  representan respectivamente las tasas de variación interanual pronosticadas y efectivas:  $P_t = \frac{\hat{y}_t - y_{t-1}}{y_{t-1}}$ ,  $A_t = \frac{y_t - y_{t-1}}{y_{t-1}}$ .

y donde el numerador puede ser considerado como un indicador de la manera en la que los errores de predicción se dispersan en torno a cero, y el denominador como un indicador de la dificultad de la predicción<sup>69</sup>.

<sup>67</sup> Si bien  $t=1, \dots, T$ , generalmente la evaluación suele llevarse a cabo a partir de las observaciones disponibles ( $t=1 \dots n$ ), lo cual implica evaluar el modelo estimado más que la predicción, que se evaluaría para ( $t=h, \dots, T$ ).

<sup>68</sup> En 1955 Theil expone en *International Economic Papers*, 5 194-199 las primeras ideas de una medida que tenga en cuenta ambos inconvenientes.

El índice de Theil es utilizado con generalidad como medida de la bondad de las predicciones debido a su sencillez de cálculo e interpretación. Así, el índice U adoptará valores nulos únicamente en el caso de coincidencia entre tasas pronosticadas y reales, mientras el resultado  $U=1$  se corresponde con las predicciones ingenuas:  $\hat{y}_t = y_{t-1} \Rightarrow P_t = 0 \forall t$ . Este valor representa un límite superior que si se superase implicaría que el modelo considerado predice peor que el modelo ingenuo<sup>70</sup>.

Una ventaja adicional es la *descomponibilidad*, que se basa en la posibilidad de expresar el numerador de  $U^2$  como:

$$\sum_t (P_t - A_t)^2 = (\bar{P} - \bar{A})^2 + (S_p - S_A)^2 + 2(1 - r_{PA}) S_p S_A$$

que permite distinguir las contribuciones o pesos de los factores de sesgo, varianza y covarianza:

$$U^S = \frac{(\bar{P} - \bar{A})^2}{\sum_t (P_t - A_t)^2} \quad U^V = \frac{(S_p - S_A)^2}{\sum_t (P_t - A_t)^2} \quad U^C = \frac{2(1 - r_{PA}) S_p S_A}{\sum_t (P_t - A_t)^2}$$

cuya suma es unitaria.

Aunque el objetivo perseguido sería llegar a predicciones coincidentes con la realidad ( $U=0$ ), que conllevarían valores nulos para los tres componentes del error, en general esto no es posible. Por tanto, la recomendación más general consiste en evitar sesgos o errores sistemáticos, concentrando la mayor participación en el componente de covarianza, cuya corrección no es posible dada su naturaleza asistemática.

Theil (1971) elabora una medida capaz de tener en cuenta la dificultad del año de la predicción ( $t=1...T$ ). Es más, si consideramos series temporales multivariantes, entonces se añadirá a su predicción un problema más: la dificultad de predecir  $Y_t$  en función de la dificultad para predecir las variables  $X_t^d$  ( $d=1...r$ ) que forman parte del modelo. Si además se tiene en cuenta que las predicciones para un periodo se renuevan

<sup>69</sup> Dado que las predicciones son hechas año a año, los mayores errores pueden esperarse cuando las diferencias entre los valores reales de años sucesivos son mayores.

<sup>70</sup> Obsérvese que el índice considerado puede sin embargo adoptar valores superiores a la unidad al no encontrarse acotado superiormente. En trabajos previos, Theil (1958) había propuesto la expresión

$$U = \frac{\sqrt{\frac{1}{T-1} \sum_t (P_t - A_t)^2}}{\sqrt{\frac{1}{T-1} \sum_t A_t^2 + \frac{1}{T-1} \sum_t P_t^2}} \quad \text{que sí se encuentra acotada entre 0 y 1, pero en cambio presenta la}$$

limitación de incluir las tasas previstas  $P_t$  como referencia para la comparación.

de acuerdo con los diferentes estadios de la información disponible ( $s=0\dots h$ ), donde 0 corresponde a la información más reciente sobre los valores de las variables y h al estadio donde se dispone de las informaciones de las variables cerrado el año), es lógico pensar que el error cuadrático medio variará también de acuerdo con el estadio de la información en el que nos encontremos.

Así, si  $A_{tdh}$  es el valor real de la tasa de la variable i en el año t y  $P_{tds}$  la predicción en los diferentes estadios s, el error de predicción será  $A_{tdh} - P_{tds}$ , Theil y Scholes (1967) proponen un patrón para determinar el error estándar de acuerdo con las fuentes de error en la predicción:

$$\sum_{t,d,s} (A_{tdh} - P_{tds})^2 = \alpha_t \beta_d \gamma_s$$

donde  $\alpha_t$  mide la calidad de la predicción de acuerdo con el año de predicción,  $\beta_d$  mide la calidad en los sucesivos estadios y  $\gamma_s$  con respecto a la variable d.

A partir de la ecuación anterior se obtienen:

$$\frac{1}{T} \sum_{d=1}^r \sum_{s=1}^h \frac{(A_{tdh} - P_{tds})^2}{\beta_d \gamma_s} = \alpha_t$$

$$\frac{1}{r} \sum_{s=1}^h \sum_{t=1}^T \frac{(A_{tdh} - P_{tds})^2}{\alpha_t \gamma_s} = \beta_d$$

$$\frac{1}{h} \sum_{d=1}^r \sum_{t=1}^T \frac{(A_{tdh} - P_{tds})^2}{\alpha_t \beta_d} = \gamma_s$$

y procediendo de manera iterativa (debido a la no linealidad), es posible estimar  $\alpha_t, \beta_d, \gamma_s$ . La mejor fiabilidad del modelo corresponderá, de acuerdo con el criterio del error estándar, con valores bajos de estos tres componentes.

Makridakis y Hibon (1979) desarrollan una metodología para evaluar, en series temporales, la capacidad de factores temporales para explicar las diferencias de precisión (EAPM) de diferentes métodos de predicción.

Según estos autores, dos conjuntos de factores producen las mejores regresiones (esto es, coeficientes estadísticamente significativos y mayores  $R^2$ ). Ambos tienen al EAPM como variable dependiente y dos combinaciones de factores como variables independientes.

Una de las ecuaciones de regresión ajusta el EAPM de los modelos de predicción de acuerdo con el cambio absoluto porcentual medio de la tendencia, de la aleatoriedad,

de la estacionalidad y del número de datos utilizados para estimar los modelos de predicción.

La segunda regresión ajusta el EAPM de las predicciones de diferentes horizontes temporales  $(h_1, h_2, \dots, h_k)$ <sup>71</sup> con el cambio absoluto porcentual medio de la tendencia, de la aleatoriedad, de la estacionalidad y con el horizonte temporal de la predicción.

El método que emplean para cuantificar los factores se basa en la descomposición clásica de las series<sup>72</sup>, obteniendo posteriormente el cambio absoluto porcentual medio de cada uno de los componentes de las series temporales como:

$$EAPMc = 100 \left( \frac{1}{n-12} \sum_{t=2}^{n-h_k} \frac{|C_t - C_{t-1}|}{C_{t-1}} \right)$$

donde  $C_t$  es cualquiera de los componentes de las series temporales.

Los EAPM de los componentes son utilizados como variables independientes de las regresiones.

Puesto que el objetivo de nuestro estudio es la evaluación de las técnicas de predicción en función de su capacidad para predecir, más que en la bondad de los modelos, nos centramos en el análisis del segundo tipo de regresiones, es decir aquellas que tienen como variable dependiente al EAPM de las predicciones de cada método  $i$

Las ecuaciones consideradas son de la forma:

$$EAPM_i = a + b_2 X_2 + b_3 X_3 + b_4 X_4 + b_5 X_5 + u$$

donde  $a$  es la constante,  $b_2$ ,  $b_3$ ,  $b_4$ , y  $b_5$  son los coeficientes de la regresión, y  $X_2$ ,  $X_3$  y  $X_4$  son respectivamente el cambio absoluto porcentual medio en la tendencia de las series para el período  $(n - h_k)$ , en la aleatoriedad de las series y en la estacionalidad, y  $X_5$  es el horizonte de predicción  $(h_1, h_2, \dots, h_k)$ .

Los parámetros de las ecuaciones pueden aportarnos información relevante sobre la relativa actuación en la predicción de las diferentes técnicas. Así por ejemplo a partir

<sup>71</sup> En concreto emplean los horizontes 1, 6, 12

<sup>72</sup> Calculando la media móvil centrada de la serie original se logra capturar el ciclo-tendencia de ésta. Posteriormente, mediante resta o división (según la hipótesis sea aditiva o multiplicativa) de la serie original, se aíslan el componente aleatorio y el componente estacional conjuntamente, a partir del cual se calcula la serie desestacionalizada. Efectuando una media móvil de la serie desestacionalizada y dividiendo esta última serie con la serie desestacionalizada original se logra extraer el componente aleatorio separadamente.

$\hat{b}_2$  puede verse el efecto que tiene sobre EAPM de cada método una variación en el cambio porcentual medio en la tendencia<sup>73</sup>.

El intento de especificar y medir la relación entre la precisión de las predicciones y los factores que la afectan es válido para ilustrar la comparación entre métodos, aunque tal y como indican Makridakis y Hibon (1979) aún son necesarias notables mejorías en esta metodología como por ejemplo introducir alguna medida cuadrática de precisión como variable dependiente, o introducir más variables independientes.

### 2.3.2 Predicciones cualitativas

A partir de las diferencias entre frecuencias (probabilidades) marginales de la tabla de contingencia, es posible definir una amplia batería de indicadores de precisión para variables cualitativas, similares a las que se han descrito para variables cuantitativas, por ejemplo a partir de  $e_{ij(i=j)}^A = |p_i - p_j|$ , el Error Absoluto Medio se calcula como  $E_a = \frac{1}{m} \sum_{i,j=1}^m |p_i - p_j|$ , si bien existen otras expresiones que se han desarrollado con el propósito de explotar más la información contenida en las tablas.

La  $\gamma$  de Goodman-Kruskal, se emplea para estudiar la precisión empleando el número de pares realización-predicción coincidentes y no coincidentes y se utiliza para ver si existe o no asociación entre idénticas categorías ( $M=1, \dots, m$ ). Un valor positivo ( $\gamma=1$  máximo) indica predicciones perfectas.

Se pueden calcular también índices desagregados asociados con cada posible par de respuestas (para cada celda de la matriz de contingencia) y calculados a partir de las probabilidades condicionales. Así, el índice kappa, propuesto inicialmente por Cohen (1960), se define como:

$$\kappa = \frac{\sum_{i,j=1}^m p_{ij} - \sum_{i,j=1}^m p_i \cdot p_j}{1 - \sum_{i,j=1}^m p_i \cdot p_j}$$

expresión para la que se obtiene el valor  $\kappa=1$  en el caso de máxima concordancia entre predicciones y realizaciones.

---

<sup>73</sup> Makridakis y Hibon (1979), en su estudio empírico extraen algunas conclusiones tales como que un aumento en los cambios porcentuales medios de la tendencia o un aumento en el horizonte temporal, provoca un mayor incremento en el EAPM de los modelos ARIMA.

A la hora de interpretar el valor de  $\kappa$  es útil disponer -a pesar de su arbitrariedad- de una escala como la siguiente:

**Tabla 6**

<b>Valoración del Índice Kappa</b>	
<b>Valor de k</b>	<b>Fuerza de la concordancia</b>
< 0,20	Pobre
0,21 – 0,40	Débil
0,41 – 0,60	Moderada
0,61 – 0,80	Buena
0,81 – 1,00	Muy buena

Por su parte, la variabilidad puede ser relevante tanto en la formulación de contrastes de hipótesis como en la construcción de intervalos de confianza. La distribución asintótica del estimador cuando el verdadero valor de  $\kappa$  es nulo viene dada por la expresión:

$$\sigma_0^2(\kappa) = \frac{\sum_{i,j=1}^m p_i p_j + \left( \sum_{i,j=1}^m p_i p_j \right)^2 - \sum_{i,j=1}^m p_i p_j (p_i p_j)}{\left( 1 - \sum_{i,j=1}^m p_i p_j \right)^2 T}$$

cuyo valor muestral que denotaremos por  $S_0^2(\kappa)$  se obtiene reemplazando las probabilidades teóricas, que desconocemos, por las proporciones muestrales.

Estos resultados se emplean para contrastar:  $H_0 : \kappa = 0$   
 $H_1 : \kappa \neq 0$ , utilizando como

estadístico del contraste la expresión  $\frac{|\hat{\kappa}|}{S_0^2(\kappa)}$  con distribución normal estándar.

Una variante de este planteamiento consiste en asignar un peso a las diferentes posibilidades de desacuerdo, de tal manera que se considere como más importante un desacuerdo entre categorías alejadas que entre las próximas. Este peso variará entre 1 (acuerdo, misma categoría) y 0 (desacuerdo con categorías extremas) y la idea del índice ponderado es asignar a cada celda de la tabla un peso  $w_{ij}$  comprendido entre 0 y 1 que represente la importancia del desacuerdo. El peso máximo corresponde al acuerdo perfecto, y se obtienen pesos proporcionalmente menores según la importancia del

desacuerdo ( $w_{ij} = 1$  y  $0 \leq w_{ij} < 1$ ), siendo los pesos más comúnmente utilizados los lineales y los bicuadrados<sup>74</sup>.

### 2.3.3 Predicciones probabilísticas: medidas de Calibración

Los tests denominados de “calibración de probabilidad” evalúan las predicciones probabilísticas teniendo en cuenta aspectos como la exactitud (la cercanía, en promedio, de las probabilidades con las realizaciones observadas) y su calibración (la cercanía de la probabilidad con la frecuencia relativa observada). Ambos conceptos se miden respectivamente con el error cuadrático medio (PS) de Brier (1950) y el sesgo cuadrático global (GSB).

$$PS = \frac{1}{T} \sum_{t=1}^T 2(y_{t+h} - P(y_{t+h} / I_t))^2 \text{ donde } PS \in [0,2] \text{ indicando valores más pequeños}$$

predicciones más precisas<sup>75</sup>.

$$GSB = 2(\overline{{}^p E_{t+h,t}} - \bar{y}_{t+h})^2 \text{ medida comprendida en los mismos límites que PS y con}$$

la misma orientación.

En el caso de que las predicciones sean cualitativas y provengan por ejemplo de encuestas de opinión, se puede tomar la mayor frecuencia relativa<sup>76</sup> de cada categoría  $r$  ( $r$ =disminuir, mantenerse o aumentar) como una predicción probabilística o grado de creencia sobre la verdadera tendencia que tendrá la variable en  $t+h$ . En ese caso las realizaciones serán 1 ó 0 si la tendencia finalmente observada, coincide o no con la de mayor porcentaje.

## 2.4 Evaluación de los puntos de giro

En este apartado estudiamos medidas para evaluar si la información proporcionada por las predicciones es adecuada para seguir los ciclos económicos y determinar puntos de inflexión.

---

<sup>74</sup> Los pesos lineales serían del tipo  $w_{ij} = 1 - \frac{|i-j|}{m-1}$  y los pesos bicuadrados  $w_{ij} = 1 - \frac{(i-j)^2}{(m-1)^2}$ .

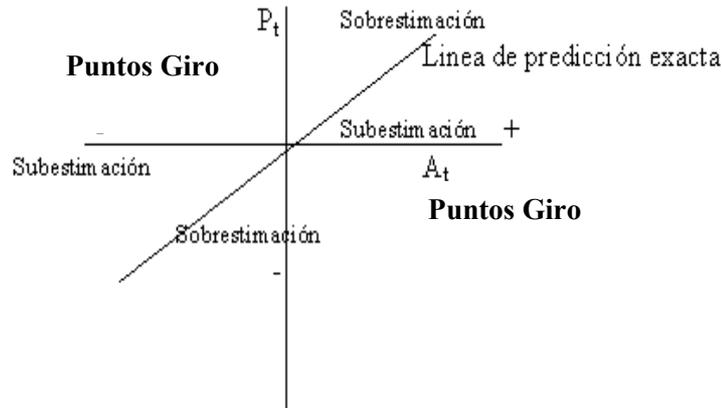
<sup>75</sup> Dado que las predicciones y realizaciones no se conocen, el análisis se realiza a partir de los valores pasados de  $R_t$  y  $C_t$ .

<sup>76</sup> Siguiendo a Savage (1971), cuando los sucesos son excluyentes la creencia puede ser tomada como una probabilidad y en concreto como una probabilidad frecuencalista.

### 2.4.1 Predicciones cuantitativas

Mediante la representación de la nube de puntos, Theil (1961, 1966) propone el *diagrama de realización- predicción* en el que se representan las tasas de crecimiento reales  $A_t$  y previstas  $P_t$ .

**Figura 3**



En el caso ideal la recta de regresión debería de coincidir con la bisectriz del primer y tercer cuadrantes del diagrama, sobre los que se delimitan también zonas de infra y sobrevaloración de  $A_t$  frente a  $P_t$ , es decir, errores de cuantía pero no de signo (sistemáticos).

En el segundo y cuarto cuadrantes se recogen los errores que han sido tanto de cuantía como de signo, es decir, donde el modelo no ha sido capaz de predecir los cambios en la tendencia de la serie o *puntos de giro*. La visión conjunta de los cuadrantes puede así darnos una idea de la fiabilidad de un modelo, tanto de los errores sistemáticos como de los cambios de signo. A partir de aquí es posible elaborar diversos indicadores atendiendo a las frecuencias de los puntos en los diversos cuadrantes. Así por ejemplo llamando  $f_i$ ,  $f_{ii}$ ,  $f_{iii}$  y  $f_{iv}$  a las frecuencias de los puntos de los cuadrantes I, II, III y IV respectivamente podemos obtener una medida sobre la capacidad del modelo para predecir el cambio en la tendencia o puntos de giro, medida acotada entre 0 y 1:

$$F = \frac{f_{ii} + f_{iv}}{f_i + f_{ii} + f_{iii} + f_{iv}} = \frac{f_{ii} + f_{iv}}{T}$$

### 2.4.2 Predicciones cualitativas

Existe un conjunto de test diseñados para evaluar la calidad de las predicciones únicamente en la medida en que éstas sirven para indicar la dirección futura de la economía. Merton (1981) y Henriksson y Merton (1981)<sup>77</sup> consideran dos categorías correspondientes a una dirección negativa y positiva de la economía (1 y 2 respectivamente), obteniendo así la clasificación:

**Tabla 7**

REALIZACIONES	PREDICCIONES		
	$\hat{Y}^1$	$\hat{Y}^2$	Total
$Y^1$	$n_{11}$	$n_{12}$	$n_{1.}$
$Y^2$	$n_{21}$	$n_{22}$	$n_{2.}$
Total	$n_{.1}$	$n_{.2}$	T

Al igual que antes, los valores  $n_{ij}$  de cada celda se modifican por las probabilidades conjuntas, que denotaremos por  $p_{ij}$

Las probabilidades de una predicción correcta en la dirección corresponderían con  $p_{11}$  y  $p_{22}$  y la estimación del total de éstas ( $p=p_{11}+p_{22}$ ) se obtendrían a partir de las frecuencias relativas, es decir:  $p_{11} + p_{22} = \frac{n_{11}}{n_{.1}} + \frac{n_{22}}{n_{.2}}$ .

Henriksson y Merton (1981) demostraron que la distribución de  $n_{11}$  condicionada a  $n_{.1}$ , T y  $n_{.2}$  viene dada por una distribución hipergeométrica:

$$p(n_{11} = x | n_{.1}, T, n_{.2}) = \frac{\binom{n_{.1}}{x} \binom{T - n_{.1}}{n_{.1} - x}}{\binom{T}{n_{.1}}}$$

donde  $n_{11}$  puede tomar valores dentro del rango:

$$n_{11} = \max\{0, n_{.1} - n_{.2}\} \leq n_{11} \leq \min\{n_{.1}, n_{.1}\} = \overline{n_{11}}$$

Merton (1981) indica que una condición necesaria y suficiente para que las predicciones tengan valor para predecir la dirección es que:  $p_{11} + p_{22} \geq 1$ . Por tanto para examinar la calidad de las predicciones es necesario contrastar:

<sup>77</sup> Estos autores desarrollan su trabajo en el ámbito financiero, si bien Ash (1998) lo aplica en la evaluación de las predicciones que sobre crecimiento económico elabora la OCDE.

$$H_0 : p_{11} + p_{22} \geq 1$$

$$H_1 : p_{11} + p_{22} < 1$$

donde el valor crítico observado a partir de la información muestral será

$$\sum_{x=n_{11}}^{n_{11}} \binom{n_{.1}}{x} \binom{T-n_{.1}}{n_{1.}-x} / \binom{T}{n_{1.}}$$

Si la hipótesis de calidad en la dirección se rechaza es posible contrastar:

$$H_0 : p_{11} + p_{22} = 1$$

$$H_1 : p_{11} + p_{22} < 1$$

donde la hipótesis nula equivale a la independencia entre realizaciones y predicciones<sup>78</sup>.

En este caso el valor crítico observado sería:  $\sum_{x=n_{11}}^{\bar{n}_{11}} \binom{n_{.1}}{x} \binom{T-n_{.1}}{n_{1.}-x} / \binom{T}{n_{1.}}$

Pesaran y Timmerman (1992) desarrollan bajo este enfoque un test no paramétrico para realizar el mismo contraste, considerando:

$y_{t+h} = 1$  cuando la economía crece, con probabilidad  $p_{y_{t+h}} = p(y_{t+h} > 0)$ , siendo  $y_{t+h} = 0$  en otro caso.

$\hat{y}_{t+h,t} = 1$  cuando la predicción revela crecimiento en la economía, con probabilidad  $p_{\hat{y}_{t+h,t}} = p(\hat{y}_{t+h,t} > 0)$ , siendo  $\hat{y}_{t+h,t} = 0$  en otro caso.

$B = 1$  cuando las predicciones y realizaciones coincidan ( $y_{t+h} \hat{y}_{t+h,t} > 0$ ), con probabilidad  $p = p(y_{t+h} \hat{y}_{t+h,t} > 0)$  siendo  $B=0$  en otro caso.

Puesto que sólo se dispone de los pares predicción-realización con información pasada ( $Y_t$  e  $\hat{Y}_t$  respectivamente), es necesario estimar dichas probabilidades a partir de las n observaciones. Sustituyendo las probabilidades anteriores por las correspondientes frecuencias observadas ( $\hat{p}_{Y_t}$ ,  $\hat{p}_{\hat{Y}_t}$ ,  $\hat{p}$  para  $p_{\hat{Y}_{t+h}}$ ,  $p_{\hat{Y}_{t+h}}$  y  $p$  respectivamente), se define el estadístico:

$$S^2 = \frac{\hat{p} - \hat{p}^*}{\text{var}(\hat{p}) - \text{var}(\hat{p}^*)} \approx \chi_1^2$$

donde:

$$\hat{p}^* = \hat{p}_{Y_t} \hat{p}_{\hat{Y}_t} + (1 - \hat{p}_{Y_t})(1 - \hat{p}_{\hat{Y}_t})$$

<sup>78</sup> Bajo esta hipótesis los modelos ingenuos no tendrían valor puesto que o bien  $p_{11}=1$  y  $p_{22}=0$  o viceversa.

$$\text{var}(\hat{p}) = \frac{1}{n}(\hat{p}^*(1 - \hat{p}^*))$$

$$\text{var}(\hat{p}^*) = \frac{1}{n} \left[ \begin{aligned} & (2\hat{p}_{Y_i} - 1)^2 \hat{p}_{\hat{Y}_i} (1 - \hat{p}_{\hat{Y}_i}) + (2\hat{p}_{\hat{Y}_i} - 1)^2 \hat{p}_{Y_i} (1 - \hat{p}_{Y_i}) + \\ & + 4\hat{p}_{Y_i} \hat{p}_{\hat{Y}_i} (1 - \hat{p}_{Y_i})(1 - \hat{p}_{\hat{Y}_i}) \end{aligned} \right] / n$$

### 2.4.3 Predicciones probabilísticas

Las predicciones generadas por la probabilidad secuencial recursiva (PSR) denominadas  $P_t$  y  $V_t$  para pico o valle respectivamente y estudiadas en 1.3., se evalúan a partir del error cuadrático medio (PS) y el sesgo cuadrático global (GBS) de Brier (1950).

En este caso la realización  $y_{t+h}^p = 1$  si se produce un punto giro cuando el PSR lo indica (esto es, cuando es mayor que una probabilidad crítica fijada por el analista) o  $y_{t+h}^p = 0$  en caso contrario.

En caso de las evaluaciones de las probabilidades secuenciales recursivas de pico o valle, el error cuadrático medio y el sesgo global tendrán como expresiones:

$$PS = \frac{1}{T} \sum_{t=1}^T 2(y_{t+h}^p - P_t)^2, \quad GSB = 2(\bar{y}^p - \bar{P})^2$$

$$PS = \frac{1}{T} \sum_{t=1}^T 2(y_{t+h}^p - V_t)^2, \quad GSB = 2(\bar{y}^p - \bar{V})^2$$

donde PS y GBS  $\in [0, 2]$  indicando valores más pequeños predicciones más precisas.

### 2.5 Nuevas alternativas basadas en medidas de información

Desde sus inicios en los años 40, la teoría de la información se ha revelado como una herramienta de gran versatilidad, siendo habituales sus aplicaciones a ámbitos como la medición de la concentración industrial, la diversidad, la desigualdad de renta, ...

Las aportaciones pioneras en este ámbito son debidas a C.E. Shannon (1948) autor de la medida de información de uso más generalizado. Dada una variable aleatoria X con sistema de probabilidades  $p_1, \dots, p_n$  ( $\sum_{i=1}^n p_i = 1$ ) la entropía poblacional de Shannon

viene dada por la expresión  $H(X) = -\sum_{i=1}^n p_i \log p_i$ , que adopta valor nulo cuando  $p_i = 1$

para algún  $i$  y  $p_j=0$  para  $j \neq i$  y alcanza su valor máximo en el caso de resultados equiprobables  $p_i = \frac{1}{n} \forall i$ .

Pese a sus indudables ventajas esta medida presenta también limitaciones, asociadas en gran medida a su estimación. Dichas limitaciones son solucionadas por las medidas de incertidumbre tipo  $\beta$  definidas por J. Havrda y F. Charvat, (1967) cuyo caso particular  $\beta=2$  da lugar a la incertidumbre cuadrática propuesta por R. Pérez (1985). Este autor define también la incertidumbre útil cuadrática y la inquietud cuadrática, poniendo de manifiesto algunas ventajas de estos indicadores con respecto a las medidas tipo Shannon<sup>79</sup>.

Dada una población sobre la que se define la variable aleatoria  $X$  con sistema de probabilidades  $\{p_1, \dots, p_n\}$  se define la *incertidumbre o entropía cuadrática*:

$$H^2(X) = H^2(p_1, \dots, p_n) = 2 \sum_{i=1}^n p_i (1 - p_i) = 2 \left( 1 - \sum_{i=1}^n p_i^2 \right)$$

Si asociado a la variable  $X$  definimos además un sistema de utilidades positivas  $\{u_1, \dots, u_n\}$ ,  $u_i > 0$   $i=1, \dots, n$ , la *incertidumbre útil cuadrática* viene dada por la expresión:

$$HU^2(X) = HU^2(p_1, \dots, p_n, u_1, \dots, u_n) = 2 \sum_{i=1}^n p_i \left( \frac{E(u)}{u_i} - p_i \right)$$

y la *inquietud cuadrática* viene dada por:

$$HU^{*2}(X) = HU^{*2}(p_1, \dots, p_n, u_1, \dots, u_n) = 2 \sum_{i=1}^n p_i \left( \frac{E(u)}{u_i} - 1 \right)$$

Esta segunda medida se obtiene al eliminar de la anterior la incertidumbre debida exclusivamente a las probabilidades, por lo cual aísla el nivel de incertidumbre asignable directamente a las utilidades.

---

<sup>79</sup> En R. Pérez (1985) se definen estas medidas, analizando exhaustivamente sus propiedades y su estimación en poblaciones finitas.

### 2.5.1 Medida de Entropía de Theil

Basándose en *la teoría de la información*, H. Theil ha desarrollado otras propuestas para la evaluación de predicciones. En concreto, este autor propone la utilización de medidas de información basadas en la incertidumbre de Shannon<sup>80</sup> (1948), para evaluar previsiones que cumplan los requisitos de no negatividad y suma unitaria, siendo así susceptibles de ser interpretadas como probabilidades.

Así, si el objetivo es anticipar los valores:  $A_1, \dots, A_n$  tales que  $A_i > 0$ ,  $\sum_i A_i = 1$  y consideramos las predicciones  $P_1, \dots, P_n$  con  $P_i > 0$ ,  $\sum_i P_i = 1$ , éstas serán tanto más precisas cuanto menor sea la información que aportan los verdaderos valores, dada la predicción. De ahí que Theil proponga la utilización de la medida<sup>81</sup>:

$I(A : P) = \sum_{i=1}^n A_i \log \frac{A_i}{P_i}$ , que puede ser interpretada como el *valor de la información aportada por A, dados los P* y también como la *imprecisión o pérdida de información de P con respecto a A*. Como consecuencia, se trata de una medida de imprecisión o inexactitud de las predicciones, que adoptaría valores nulos en el caso de predicciones perfectas.

Para el caso de predicciones cuantitativas, cualitativas y probabilísticas las expresiones serían las siguientes:

- Para las predicciones cuantitativas, donde cada observación  $y_{t+h}$  y  $\hat{y}_{t+h,t}$  es única y por tanto de probabilidades unitarias, es necesario agrupar la información intervalos de predicción, que denotamos por  $[\hat{Y}_j]$  y de realizaciones  $[Y_i]$  en k

<sup>80</sup> Dada una variable aleatoria X con sistema de probabilidades  $p_1, \dots, p_n$  ( $\sum_i p_i = 1$ ) la entropía poblacional de Shannon viene dada por la expresión  $H(X) = -\sum_{i=1}^n p_i \log p_i$ , que adopta valor nulo cuando  $p_i = 1$  para algún i y  $p_j = 0$  para  $j \neq i$  y alcanza su valor máximo en el caso de resultados equiprobables  $p_i = \frac{1}{n} \forall i$ .

<sup>81</sup> La construcción de esta medida se lleva a cabo identificando las predicciones  $P_i$  con probabilidades *a priori* y las realizaciones  $A_i$  con probabilidades *a posteriori*. De este modo, antes de conocer el valor de una observación su incertidumbre puntual (o *autoinformación*) vendría dada por  $-\log P_i$ , mientras que una vez conocida la observación se obtendría la incertidumbre a posteriori  $-\log A_i$ . Si calculamos las esperanzas de las incertidumbres a priori y a posteriori (usando como ponderaciones en ambos casos las probabilidades a posteriori  $A_i$ ) la diferencia entre ambas cuantifica la información que la realización aporta sobre la predicción:  $-\sum_{i=1}^n A_i \log P_i + \sum_{i=1}^n A_i \log A_i = \sum_{i=1}^n A_i \log \frac{A_i}{P_i}$ .

categorías, de tal manera que se puedan clasificar las observaciones T en una

tabla de contingencia. En ese caso:  $I([Y_i], [\hat{Y}_j]) = \sum_{i=j=1}^k p_i \log \frac{p_i}{p_i}$

- Para las predicciones cualitativas se tendría  $I(Y^M : \hat{Y}^M) = \sum_{i=j=1}^m p_i \log \frac{p_i}{p_i}$
- Para las predicciones probabilísticas la expresión, que se obtiene al relacionar la funciones de probabilidad a priori y a posteriori, sería:

$$I(P_Y(Y) : P_Y(Y/I)) = \sum_{t=1}^T P_{Y_t}(y_t) \log \frac{P_{Y_t}(y_t)}{P_{Y_t}(y_t/I_t)}$$

### 2.5.2 Medida cuadrática de imprecisión

La propuesta de Theil basada en la entropía de tipo Shannon podría ser extrapolada a otras medidas de incertidumbre. Así, si consideramos valores efectivos:  $A_1, \dots, A_n$  tales que  $A_i > 0$ ,  $\sum_i A_i = 1$  y predicciones  $P_1, \dots, P_n$  con  $P_i > 0$ ,  $\sum_i P_i = 1$ , podríamos construir las incertidumbres puntuales cuadráticas definidas por R. Pérez (1985)<sup>82</sup> que, antes de conocer las observaciones, vendrían dadas por la expresión  $2(1 - P_i)$  pasando posteriormente a  $2(1 - A_i)$ .

Siguiendo un procedimiento análogo al propuesto por Theil, la diferencia de incertidumbres esperadas antes y después de conocer las observaciones (ponderadas por las probabilidades a posteriori) cuantificaría la ganancia informativa de la realidad con respecto a las predicciones, es decir, la imprecisión de éstas.

Llegaremos así a una expresión:

$$I(A : P) = 2 \sum_{i=1}^n A_i (1 - P_i) - 2 \sum_{i=1}^n A_i (1 - A_i) = 2 \sum_{i=1}^n A_i (A_i - P_i)$$

que denominamos *medida cuadrática de imprecisión*.

Para el caso de predicciones cuantitativas, cualitativas y probabilísticas y las expresiones serían:

$$I([Y_i], [\hat{Y}_j]) = 2 \sum_{i=j=1}^k p_i (p_i - p_j)$$

<sup>82</sup> Según este autor la incertidumbre puntual cuadrática para un valor  $x_i$  de la variable X se define como  $H^2(x_i) = 2(1 - p_i)$ , que puede ser interpretada como el doble de lo que dista  $x_i$  de ser con total seguridad el verdadero valor poblacional.

$$I(Y^M : \hat{Y}^M) = 2 \sum_{i=j=1}^m p_i (p_i - p_j)$$

$$I(P_Y(Y) : P_Y(Y/I)) = 2 \sum_{t=1}^T P_{Y_t}(y_t) (P_{Y_t}(y_t) - P_{Y_t}(y_t/I_t))$$

Sin entrar aquí en un análisis detallado del comportamiento de esta medida, señalamos algunas de sus ventajas y limitaciones:

1. Cuando las predicciones y las realizaciones son coincidentes, la medida de imprecisión adopta valor nulo, mostrando un comportamiento análogo al de la medida propuesta por Theil.
2. No se garantiza la implicación en sentido contrario, pudiendo obtenerse resultados nulos en situaciones distintas a las predicciones perfectas<sup>83</sup>.
3. La medida cuadrática de imprecisión no se encuentra acotada inferiormente.

Además de las limitaciones anteriores, debemos tener en cuenta que la medida cuadrática ha sido derivada partiendo de indicadores de incertidumbre, basados directamente en probabilidades. No obstante, teniendo en cuenta que en general los resultados posibles de una experiencia llevan asociada una utilidad, y que dicha utilidad afectará al nivel de incertidumbre percibido, resulta conveniente proponer medidas de incertidumbre asociadas a un sistema de probabilidades y utilidades<sup>84</sup>.

Con este nuevo planteamiento, sería posible considerar tres tipos de medidas, según nos interese únicamente la incertidumbre probabilística (*entropía o incertidumbre*), la incertidumbre en probabilidades y utilidades (*incertidumbre útil*) o bien la incertidumbre asociada sólo a las utilidades (*inquietud*). Este último parece ser el enfoque más adecuado en el ámbito de la predicción económica, donde la incertidumbre futura va directamente asociada a la utilidad asociada a la magnitud objeto de estudio.

---

<sup>83</sup> De hecho, se comprueba fácilmente que la medida es nula cuando las observaciones reales son constantes, ya que en este caso  $A_i = \frac{1}{n} \forall i \Rightarrow I(A : P) = 2 \sum_{i=1}^n \frac{1}{n} \left( \frac{1}{n} - P_i \right) = 2 \left( \frac{1}{n} - \frac{1}{n} \right) = 0$ . Si bien este

hecho no resulta deseable desde el punto de vista operativo, podría ser justificado conceptualmente teniendo en cuenta que en esta situación la realidad llevaría asociada incertidumbre máxima o información mínima, por lo cual su aportación con respecto a la predicción es nula.

<sup>84</sup> Los trabajos pioneros en este sentido se deben a M. Belis y S. Guiasu (1968), quienes proponen una medida *cuantitativo-cualitativa de información*. Posteriormente, P. Gil (1975) y A. Gil (1981) definen las medidas de incertidumbre útil y de inquietud respectivamente (asociadas ambas a entropías de Shannon) y R. Pérez (1985) define las medidas de incertidumbre e inquietud cuadráticas.

R. Pérez (1985) define la medida de inquietud cuadrática asociada a un sistema de probabilidades  $\{p_1, \dots, p_n\}$  y utilidades  $\{u_1, \dots, u_n\}$  como:

$$HU^{*2}(p_1, \dots, p_n, u_1, \dots, u_n) = 2 \sum_{i=1}^n p_i \left( \frac{E(u)}{u_i} - 1 \right).$$

Si nuestro objetivo es llegar a determinar los valores de una magnitud  $Y$  a lo largo del período temporal  $t=1, \dots, T$ , la inquietud cuadrática asociada (para predicciones cuantitativas) viene dada por la expresión:  $H(Y) = \frac{2}{T} \sum_t \left( \frac{E(Y)}{y_t} - 1 \right)^{85}$ .

La calidad de un conjunto de predicciones  $\hat{y}_1, \dots, \hat{y}_T$  podría entonces ser analizada a partir de la inquietud cuadrática condicionada  $H(Y/\hat{Y})$ , conduciendo la comparación de esta expresión con  $H(Y)$  a una medida de precisión de las predicciones

Desde el marco de la teoría de la información, P. Gil (1981) sugiere dos posibilidades para la consecución de una incertidumbre útil condicionada: condicionar mediante una variable aleatoria, cuyos resultados no intervengan en las utilidades iniciales o bien condicionar mediante un nuevo campo de probabilidad y utilidad. A continuación exploramos las posibilidades que ambas alternativas ofrecen en el caso de predicciones cuantitativas.

### 2.5.3 Información cuadrática asociada a las predicciones cuantitativas

La definición de la inquietud cuadrática condicionada a una variable aleatoria  $X$  aparece en M. Alvargonzález y N. Muñoz (1992). Dada una v.a.  $X$  con valores  $x_j$ , que modifica las probabilidades de un campo pero no sus utilidades, se define la inquietud del campo  $A$  condicionada al valor  $x_j$ :

$$H(A/x_j) = 2 \sum_i p_{i/j} \left( \frac{E_{x_j}(u)}{u_i} - 1 \right) \text{ donde } p_{i/j} = p(A/x_j) \quad E_{x_j}(u) = \sum_i u_i p_{i/j}$$

A partir de estas medidas, la inquietud cuadrática de  $A$  condicionada por  $X$  se define entonces como:

---

<sup>85</sup> R. Pérez define la medida inquietud cuadrática asociada a un sistema de probabilidades  $\{p_1, \dots, p_n\}$  y utilidades  $\{u_1, \dots, u_n\}$  como  $HU^{*2}(p_1, \dots, p_n, u_1, \dots, u_n) = 2 \sum_{i=1}^n p_i \left( \frac{E(u)}{u_i} - 1 \right)$ . No obstante, en este caso consideramos como utilidades las observaciones  $Y_t$ , con probabilidades  $p_t = \frac{1}{T}$  utilizando por comodidad la notación  $H(Y)$ .

$$H(A/X) = \sum_j p_j \frac{E(u)}{E_{x_j}(u)} H(A/x_j)$$

Teniendo en cuenta que en el caso que nos ocupa la inquietud de referencia es

$$H(Y) = \frac{2}{T} \sum_t \left( \frac{E(Y)}{y_t} - 1 \right),$$

y la variable aleatoria que condiciona son las predicciones  $\hat{y}_1, \dots, \hat{y}_T$ , las definiciones anteriores no podrían ser aplicadas de forma directa puesto que, al llevar cada realización asociada una predicción, se tendría

$$E_{\hat{y}}(Y) = y_t \quad \forall t = 1, \dots, T \Rightarrow H(Y/\hat{Y}_t) = 0 \quad \forall t = 1, \dots, T \Rightarrow H(Y/\hat{Y}) = 0.$$

Una posible solución sería la extensión de las medidas al caso continuo, si bien esta alternativa exigiría conocer las funciones de densidad marginales y condicionadas<sup>86</sup>. Por consiguiente, planteamos una solución intermedia de agrupar la información según los intervalos de predicción,  $[\hat{Y}_j]$ , definiendo la inquietud condicionada por cada intervalo:

$$H(Y/[\hat{Y}_j]) = 2 \sum_t p_{Y/[\hat{Y}_j]} \left( \frac{E_{[\hat{Y}_j]}(Y)}{y_t} - 1 \right) \text{ con}$$

$$p_{Y/[\hat{Y}_j]} = p(Y/[\hat{Y}_j]); E_{[\hat{Y}_j]}(Y) = \sum_t y_t p_{Y/[\hat{Y}_j]}$$

y la *inquietud de la realización condicionada a la predicción* vendrá dada por la expresión:

$$H(Y/[\hat{Y}]) = \sum_j p_{[\hat{Y}_j]} \frac{E(Y)}{E_{[\hat{Y}_j]}(Y)} H(Y/[\hat{Y}_j])$$

que presenta los siguientes rasgos:

1. Esta medida adopta valor nulo cuando existe máxima homogeneidad en la variable Y según los intervalos de predicción, aumentando su valor a medida que crece la heterogeneidad.

---

<sup>86</sup> La definición sería en este caso  $H(Y/\hat{Y}) = 2 \int_0^{+\infty} \left( \frac{E_y(Y)}{y} - 1 \right) f(y/\hat{y}) dy$  donde  $E_{\hat{y}}(Y) = \int_0^{+\infty} y f(y/\hat{y}) dy$ ,

obteniéndose la inquietud cuadrática  $H(Y/\hat{Y}) = \int_0^{+\infty} \frac{E(Y)}{E_{\hat{y}}(Y)} H(Y/\hat{y}) f(\hat{y}) d\hat{y}$ .

2. El resultado de la medida viene condicionado por el criterio de agrupación adoptado, aumentando su capacidad descriptiva con la precisión de dicha agrupación.
3. La inquietud de la realización condicionada a la predicción depende de la inquietud asociada a las observaciones dentro de los intervalos de predicción, ignorando la relación existente entre realización y predicción.

Esta última característica constituye una importante limitación de la medida, ya que podrían obtenerse valores bajos para las inquietudes condicionadas en situaciones donde las observaciones disten mucho de las predicciones. Para solucionar al menos parcialmente este inconveniente planteamos la *medida cuadrática de información asociada a las predicciones*:

$$IC(Y, [\hat{Y}]) = H(Y) - H(Y | [\hat{Y}]) (1 - r_{Y, \hat{Y}})$$

Como puede apreciarse, a diferencia de la información cuadrática introducida en trabajos previos<sup>87</sup> esta medida incorpora un factor de corrección según la correlación existente entre observaciones y predicciones que intenta atenuar la ambigüedad de la inquietud condicionada a los intervalos de predicción:

Para correlaciones predicción- realidad de signo positivo, se tiene  $(1 - r_{Y, \hat{Y}}) < 1$  con lo cual la inquietud condicionada *efectiva* es menor que la cuantificada inicialmente mediante los intervalos (obsérvese que en el caso límite de correlación perfecta se obtendría  $r_{Y, \hat{Y}} = 1 \Rightarrow IC(Y, [\hat{Y}]) = H(Y)$  con lo cual la información aportada por las predicciones sería del 100% de la inquietud inicial).

Cuando se observan correlaciones negativas entre predicción y realidad, se tiene  $(1 - r_{Y, \hat{Y}}) > 1$ , factor que corrige al alza la inquietud condicionada y por tanto a la baja la cantidad de información cuadrática (que podría llegar incluso a ser negativa).

La corrección no tendrá efecto en los casos de predicciones incorreladas con las observaciones.

---

<sup>87</sup> M. Alvargonzález y N. Muñoz (1992) definen la cantidad de información útil cuadrática que una variable X contiene sobre un campo como la diferencia entre la inquietud cuadrática del campo y la inquietud del campo condicionada por X:  $IU^{*2} = HU^{*2}(A) - HU^{*2}(A/X)$ , expresión que se encuentra acotada entre 0 y  $HU^{*2}(A)$ .

### 2.5.4 Imprecisión cuadrática basada en errores relativos

La consideración de un sistema de utilidades conjuntas  $u_{ij}$  con probabilidades  $p_{ij}$  permite definir la inquietud de un campo B condicionada por el campo A:

$$H(B/A) = \sum_i p_i \frac{E_i(u)}{E(u)} H(B/A_i) \text{ con } H(B/A_i) = 2 \sum_j \left( \frac{E_i(u)}{u_j} - 1 \right) p_{j/i}$$

obteniéndose la inquietud conjunta como la suma de la inquietud condicionada anterior más la inquietud de A.

Si consideramos ahora los campos de valores de realización y predicción podríamos adoptar como utilidades conjuntas los errores relativos de predicción:

$$u_{ij} = \frac{\hat{y}_j}{y_i} \text{ (que serían en realidad "desutilidades"). No obstante, la utilización directa de}$$

la información disponible nos conduciría a inquietudes nulas al trabajar con observaciones realización-predicción de frecuencias unitarias.

Una posible solución podría ser la clasificación de los datos en intervalos de predicción y realización, que denotamos respectivamente por  $[\hat{Y}_j], [Y_i]$ , cuyas

$$\text{utilidades conjuntas se obtienen como } u_{ij} = \sum_{t \in [\hat{Y}_j] \times [Y_i]} \frac{\hat{y}_t}{y_t} \frac{1}{T_{ij}} \text{ con probabilidades } p_{ij} = \frac{T_{ij}}{T}.$$

En este contexto, la inquietud condicionada a  $[\hat{Y}_j]$  vendría dada por la expresión<sup>88</sup>:

$$H(Y/[\hat{Y}_j]) = 2 \sum_i \left( \frac{u_{[\hat{Y}_j]}}{u_{ij}} - 1 \right) \frac{p_{ij}}{p_j}, \quad \text{con} \quad u_{[\hat{Y}_j]} = \sum_i u_{ij} \frac{p_{ij}}{p_j} = \sum_i \frac{\hat{Y}_j}{Y_i} \frac{p_{ij}}{p_j},$$

$$p_j = \sum_i p_{ij}$$

obteniéndose como resumen la inquietud condicionada a la predicción:

$$H(Y/\hat{Y}) = \sum_j p_j \left( \frac{E(u)}{u_{[\hat{Y}_j]}} \right) H(Y/[\hat{Y}_j])$$

<sup>88</sup> Estas definiciones siguen la metodología propuesta por P. Gil (1981) para su medida de incertidumbre útil. Obsérvese que la definición de las utilidades marginales  $u_{[\hat{Y}_j]} = \sum_i u_{ij} \frac{p_{ij}}{p_j}$  y  $u_{[Y_i]} = \sum_j u_{ij} \frac{p_{ij}}{p_i}$

garantiza  $E(u) = E(u_{[\hat{Y}_j]}) = E(u_{[Y_i]})$ .

Por su parte, la incertidumbre cuadrática conjunta se obtendría como:

$$H(Y, \hat{Y}) = 2 \sum_i \sum_j \left( \frac{E(u)}{u_{ij}} - 1 \right) p_{ij}$$

verificándose  $H(Y, \hat{Y}) = H(Y / \hat{Y}) + H(\hat{Y}) = H(\hat{Y} / Y) + H(Y)$ .

Partiendo de las inquietudes cuadráticas anteriores, la expresión:

$$I(\hat{Y}, Y) = H(Y) + H(\hat{Y}) - H(Y, \hat{Y})$$

puede ser interpretada como una medida de la imprecisión de las predicciones respecto a las realizaciones o equivalentemente, de la cantidad de información que la realidad incorpora con respecto a la predicción.

Este indicador, que denominamos *Imprecisión cuadrática basada en errores relativos*, presenta los siguientes rasgos:

- Se trata de una medida simétrica, ya que se cumple:

$$I(\hat{Y}, Y) = H(\hat{Y}) - H(\hat{Y} / Y) = H(Y) - H(Y / \hat{Y})$$

- El indicador de imprecisión no se encuentra acotado, aumentando su valor con los errores de predicción.

### 3 COMBINACIÓN DE PREDICCIONES CUANTITATIVAS

El elevado número de fuentes y métodos para la obtención de predicciones, hace posible disponer de un amplio abanico de predicciones ( $\hat{y}_{t+h,t}^i$ ) o expectativas ( ${}^e\hat{y}_{t+h,t}^i$ ) tanto cuantitativas como cualitativas para  $y_{t+h}$ ,  $\hat{Y}_{t+h,t} = (\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N)$ .

Debido a la heterogeneidad de las mismas, cada predicción captura diferentes aspectos de la información disponible y en consecuencia una combinación de todas ellas podría mejorar la precisión, por lo que puede resultar útil el cálculo de predicciones de consenso mediante síntesis de predicciones individuales.

Al igual que ocurre con las técnicas individuales, los métodos para obtener la predicción combinada son muy numerosos, dependiendo de la base informativa que consideran, predicciones individuales, ( $I_t = \hat{Y}_{t+h,t}$ ) y del procedimiento que se siga para procesarla ( $g(\hat{Y}_{t+h,t})$ ).

En este capítulo estudiamos métodos cuantitativos de predicciones cuantitativas por lo tanto la base informativa contendrá únicamente predicciones individuales cuantitativas  $\hat{Y}_{t+h,t} = (\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N)$ <sup>89</sup>.

Además de estudiar técnicas convencionales, analizamos otras que surgen con el propósito de intentar solucionar los problemas derivados del incumplimiento de los supuestos teóricos sobre los que se basan las técnicas convencionales originalmente propuestas.

#### 3.1 Métodos cuantitativos de combinación

Todos los métodos de combinación, asumen la idea de que a partir de las predicciones individuales se puede obtener una predicción de consenso  $y_{t+h,t}^c$  que combina toda la información contenida en las predicciones individuales. Es decir,  $y_{t+h,t}^c = f(\hat{Y}_{t+h,t}(\alpha)) + \mu_{t+h}$ , donde  $\alpha$  es un vector de ponderaciones de las predicciones individuales y  $\mu_{t+h}$  un término de error que incluye factores no captados por las predicciones individuales. Dependiendo de cómo se formalice la relación causal,

---

<sup>89</sup> Las predicciones cualitativas pudieran aparecer cuantificadas de alguna manera o corresponder directamente a valoraciones cuantitativas (y en ese caso también podrían formar parte del vector de predicciones individuales al que nos referimos a lo largo de este capítulo). En el capítulo 3.2. veremos métodos de cuantificación de predicciones cualitativas, que pueden incluirse dentro de la base informativa de éstas técnicas cuantitativas de combinación.

y se obtengan las estimaciones de  $\alpha$ , surgen una serie de métodos que tienen como objetivo generar una predicción combinada  $\hat{y}_{t+h,t}^c$ , esto es  $\hat{y}_{t+h,t}^c = f(\hat{Y}_{t+h,t}(\hat{\alpha}))$ .

Así, estos pesos pueden obtenerse en función de la varianza de error de predicción de cada método (y de las covarianzas entre ellos) o mediante técnicas de regresión. Todos estos métodos están basados en el enfoque clásico, al resultar los primeros casos particulares de los segundos.

La metodología Bayesiana ha sido empleada también para la elaboración de predicciones combinadas puesto que ésta ofrece consistentes y fuertes herramientas para desarrollar mecanismos de predicción. En este enfoque la variable a predecir y el conjunto de predicciones han de considerarse variables aleatorias sobre las que es necesario conocer sus distribuciones de probabilidad si bien bajo algunos supuestos los resultados obtenidos coinciden con los del enfoque clásico.

La teoría de la información constituye también una herramienta importante para explicar relaciones entre variables económicas, por lo que basándonos en este enfoque desarrollaremos nuevas técnicas para la combinación.

### 3.2 Enfoque Clásico

Los pioneros en la combinación de predicciones fueron Bates y Granger (1969)<sup>90</sup>, desarrollando técnicas para obtener la predicción combinada a partir de combinaciones lineales de predicciones individuales  $y_{t+h,t}^c = \hat{Y}_{t+h,t}(\alpha)$ , y donde los pesos se obtienen a partir de la varianza de las predicciones individuales. Con esta idea se desarrollaron numerosos métodos, hasta que Granger y Ramanathan (1984) ponen de manifiesto que los convencionales métodos de combinación son equivalentes a mínimos cuadrados ordinarios restringidos donde la predicción combinada es explicada por las predicciones individuales. Su trabajo ofrece un importante ímpetu para el uso de métodos econométricos más sofisticados en la combinación de predicciones y solucionan la limitación de los métodos de MCO.

---

<sup>90</sup> Realmente, la primera referencia sobre el uso de predicciones combinadas se encuentra en el trabajo de Barnard (1963), donde compara las predicciones basadas en modelos ARIMA con las obtenidas usando procedimientos de suavización exponencial clásicos, y observa que un promedio simple de las dos predicciones resulta mejor (Melo y Castaño, 1998).

### 3.2.1 Técnicas basadas en la varianza de los errores de predicción individual

Bates y Granger (1969) analizan la combinación para el caso de dos predicciones. Dadas dos predicciones insesgadas<sup>91</sup>, su combinación lineal será:

$$y_{t+h,t}^c = \hat{Y}_{t+h,t} \alpha \text{ donde } \alpha^T = (\alpha_1, \alpha_2) \text{ son los pesos dados a las predicciones, } l \text{ es}$$

un vector de unos (Nx1), en este caso (2x1) y cumplen  $l^T \alpha = 1$  y por lo tanto  $\alpha_2 = 1 - \alpha_1$

Si denotamos por  $\sigma_1^2$  y  $\sigma_2^2$  a las varianzas de los errores de las dos predicciones, la varianza de los errores de la predicción combinada será:

$$\sigma_c^2 = \alpha_1^2 \sigma_1^2 + (1 - \alpha_1)^2 \sigma_2^2 + 2\rho k \sigma_1 (1 - \alpha_1) \sigma_2$$

donde  $\rho$  es el coeficiente de correlación entre los errores de predicción de las predicciones individuales.

La elección de  $\alpha_1$  se debe realizar de tal manera que los errores de las predicciones combinadas sean pequeños, es decir minimizando  $\sigma_c^2$ : obtenemos que el mínimo valor de  $\sigma_c^2$  se obtiene cuando:

$$\hat{\alpha}_1 = \frac{\sigma_2^2 - \rho \sigma_1 \sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2}$$

que en el caso de  $\rho = 0$  es:

$$\hat{\alpha}_1 = \frac{\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)}$$

donde  $\sigma_c^2$  no es mayor que la más pequeña de las dos varianzas individuales.

Si bien queda así justificado el uso de una combinación de predicciones, en la práctica no es posible calcular  $\sigma_c^2$  al desconocerse las varianzas de los errores de predicción de t+h, y por lo tanto el valor óptimo de  $\alpha$  se desconoce al comienzo de la combinación.

Bates y Granger (1969) utilizan la expresión primera de  $\alpha$  para obtener estimaciones maximoverosímiles de  $\alpha$  partiendo de las predicciones individuales de un conjunto de períodos previos a aquél para el cual se va a efectuar la combinación de

---

<sup>91</sup> Dicks y Burrell (1994) ponen de manifiesto que cuando la combinación se realiza para predicciones dadas por diferentes agentes, la condición de predicciones insesgadas no suele verificarse debido a la aversión al riesgo de desviarse de la visión convencional. En este caso puede ser conveniente aplicar la corrección óptima de Theil (1971).

predicciones. Los autores sugieren la elección de los pesos basados en dar mayor importancia a la predicción individual que se haya comportado mejor en el pasado.

Así, entre las alternativas que sugieren la más empleada consiste en sustituir las varianzas de los errores de predicción por sus estimaciones de los t períodos pasados (t=1...n):

$$\hat{\alpha}_1 = \frac{\sum_{t=1}^n (e_t^2)^2}{\sum_{t=1}^n [(e_t^1)^2 + (e_t^2)^2]}$$

donde  $e_t^1 = y_t - \hat{y}_t^1$  es el error cometido por la técnica 1 al predecir para  $y_t$ . El conjunto de todos los errores pasados será  $e_t^{1T} = (e_1^1, e_2^1, \dots, e_n^1)$ . Para la técnica 2, los errores se obtendrían de forma similar. Los pesos estimados se emplearían para obtener la predicción combinada:

$$\hat{y}_{t+h,t}^c = \hat{Y}_{t+h,t} \hat{\alpha}$$

La metodología de Bates y Granger (1969) es ampliada por Newbold y Granger (1974) para el caso de más de 2 predicciones<sup>92</sup>. En ese caso,  $y_{t+h,t}^c = \hat{Y}_{t+h,t} \alpha$  donde  $\alpha^T = (\alpha_1, \alpha_2, \dots, \alpha_N)$ ,  $1$  es un vector (Nx1) de unos  $1^T \alpha = 1$  y  $0 \leq \alpha_i \leq 1$  para todo  $i$ , la varianza del error de la predicción combinada es minimizada tomando<sup>93</sup>:

$$\hat{\alpha} = \frac{(\sum^{-1} l)}{(l^T \sum^{-1} l)} \text{ donde } \sum = E(e_{t+h} e_{t+h}^T) \text{ y } e_{t+h} = y_{t+h} l^T - \hat{Y}_{t+h,t}$$

Al desconocerse las realizaciones de  $y_{t+h}$  también será necesario sustituir las varianzas de los errores de predicción por sus estimaciones es periodos previos.<sup>94</sup>

<sup>92</sup> Makridakis y Winkler (1983) discuten sobre el impacto del número de predicciones incluidas en la combinación.

<sup>93</sup> Además, el resultado del trabajo de Newbold y Granger (1974) demuestra que en la práctica se pueden ignorar las correlaciones en la estimación de la combinación ponderada, es decir, la estimación ponderada que depende de las correlaciones estimadas actúa peor que aquella que la ignora. No obstante al referirnos a predicciones subjetivas la correlación entre expertos es posible, por lo que no se debería de ignorar este efecto. Winkler y Makridakis (1983) estudiaron el problema de la obtención de pesos “óptimos” para las combinaciones ponderadas usando las técnicas de Newbold y Granger (1974) confirmando que generalmente es mejor ignorar los efectos de las correlaciones en el cálculo de los ponderaciones en las combinaciones, si bien este supuesto no tiene porque mantenerse en realidad. En el siguiente epígrafe trataremos los métodos propuestos para su solución.

<sup>94</sup> En ese caso se emplea:  $e_t^i = y_t - \hat{y}_t^i$ , es el error cometido por cada técnica individual (i=1,...N) al predecir  $y_t$ , en el periodo anterior. El conjunto de todos los errores cometidos en el pasado (n periodos)

### 3.2.2 Técnicas basadas en la regresión de las predicciones individuales

El método de regresión para la combinación de predicciones fue propuesto por Granger y Ramathan (1975) que mostraron que el vector óptimo de pesos basados en varianza-covarianza de los errores de predicción tiene una interpretación como vector de coeficientes de la proyección lineal de  $y_{t+h}$  a partir de las predicciones de los N distintos métodos.

$$y_{t+h} = \alpha_1 \hat{y}_{t+h,t}^1 + \dots + \alpha_N \hat{y}_{t+h,t}^N + u_{t+h} = \hat{Y}_{t+h,t} \alpha + u_{t+h}$$

Al desconocer el verdadero valor de  $y_{t+h}$  los pesos se obtienen a partir de las n predicciones pasadas. Recordando que  $Y_t = (y_1, y_2, \dots, y_n)$  es la evolución pasada de la serie y  $\hat{Y}_t = (\hat{Y}_t^1, \hat{Y}_t^2, \dots, \hat{Y}_t^N)$  la matriz nxN de predicciones pasadas que se realizaron sobre Y por las diferentes técnicas y/o individuos, donde cada elemento  $\hat{Y}_t^{iT} = (\hat{y}_{1,0}^i, \hat{y}_{2,1}^i, \dots, \hat{y}_{n,n-1}^i)$ , el objetivo será estimar  $\alpha$  en la regresión  $Y_t = \hat{Y}_t \alpha + u_t$  que estime  $Y_t$  como proyección lineal de  $\hat{Y}_t$ , para emplearla después en la predicción.

Según el modelo se estime o no con constante o con alguna restricción sobre los pesos de las predicciones Granger y Ramanathan estudian tres modelos de MCO:

Modelo A: No restringido sobre los pesos y sin constante

Modelo B: Restringido  $l^T \hat{\alpha} = 1$  y sin constante

Modelo C: No restringido y con constante

Llamaremos  ${}_A \hat{y}_{t+h,t}^c$  a la predicción combinada obtenida por el modelo A:

$${}_A \hat{y}_{t+h,t}^c = \hat{Y}_{t+h,t} \hat{\alpha}, \quad {}_B \hat{y}_{t+h,t}^c \text{ a la obtenida por el modelo B}^{95}: {}_B \hat{y}_{t+h,t}^c = \hat{Y}_{t+h,t} \hat{\alpha} \text{ y}$$

${}_C \hat{y}_{t+h,t}^c$  a la obtenida por el modelo C.

El objetivo es analizar qué modelo obtiene una varianza de la predicción combinada menor.

por todas las técnicas será:  $e_t = Y_t^T l^T - \hat{Y}_t$ , que es una matriz nXN, donde cada columna está formada por los errores cometidos por cada técnica en el pasado:  $e_t^{iT} = (e_{1,0}^i, e_{2,1}^i, \dots, e_{n,n-1}^i)$

<sup>95</sup> Para mayor comodidad en el método B denotaremos a  $\alpha$  como  $\beta$  donde  $\beta^T = (\beta_1, \beta_2, \dots, \beta_N)$   
 $l^T \beta = 1$

En el modelo A, el error de predicción será  $e_t^A = Y_t^T - \hat{Y}_t \alpha$ , y utilizando MCO la función a minimizar será  $(Y_t^T - \hat{Y}_t \alpha)^T (Y_t^T - \hat{Y}_t \alpha)$  con respecto a  $\alpha$ , cuya solución será:  $\hat{Y}_t^T (Y_t^T - \hat{Y}_t \alpha) = 0$ ,  $\hat{\alpha} = (\hat{Y}_t^T \hat{Y}_t)^{-1} \hat{Y}_t^T Y_t^T$

La suma cuadrática de los errores de predicción obtenida en la optimización es:

$$e_{CA}^2 = (Y_t^T - \hat{Y}_t \hat{\alpha})^T (Y_t^T - \hat{Y}_t \hat{\alpha}) = Y_t^T Y_t^T - Y_t^T \hat{Y}_t \hat{\alpha}$$

Por tanto una vez estimados los pesos la regresión estimada, la predicción combinada para t+h con este método sería:

$${}_A \hat{y}_{t+h,t}^c = \hat{Y}_{t+h,t} \hat{\alpha}$$

Para el modelo B el error de predicción será:  $e_t^B = Y_t^T - \hat{Y}_t \beta$  y entonces la función a minimizar será  $(Y_t^T - \hat{Y}_t \beta)^T (Y_t^T - \hat{Y}_t \beta)$  s.a  $l^T \beta = 1$

Es decir la función objetivo es  $\min_{\beta} \left( (Y_t^T - \hat{Y}_t \beta)^T (Y_t^T - \hat{Y}_t \beta) + 2\lambda_{\beta} (l^T \beta - 1) \right)$

donde  $\lambda_{\beta}$  es el multiplicador de Lagrange.

La condición de primer orden es

$$Y_t^T (Y_t - \hat{Y}_t \beta) - \lambda_{\beta} l = 0$$

$$\hat{\beta} = (\hat{Y}_t^T \hat{Y}_t)^{-1} \hat{Y}_t^T Y_t - \lambda_{\beta} (\hat{Y}_t^T \hat{Y}_t)^{-1} l = \hat{\alpha} - \lambda_{\beta} (\hat{Y}_t^T \hat{Y}_t)^{-1} l$$

donde  $\hat{\alpha}$  es la expresión obtenida mediante el método A y el multiplicador se obtiene teniendo en cuenta la restricción  $l^T \beta = 1$

$$\hat{\lambda}_{\beta} = \frac{(l^T \hat{\alpha} - 1)}{\left[ l^T (\hat{Y}_t^T \hat{Y}_t)^{-1} l \right]}$$

Ahora la suma del error cuadrático de la predicción es:

$$e_{CB}^2 = e_{CA}^2 + \lambda_{\beta}^2 \left[ l^T (\hat{Y}_t^T \hat{Y}_t)^{-1} l \right]$$

tal que

$$e_{CB}^2 \geq e_{CA}^2$$

y la predicción  ${}_B \hat{y}_{t+h,t}^c = \hat{Y}_{t+h,t} \hat{\beta}$

Así pues, el método de la regresión obtiene mejores resultados mediante la proyección de los valores pasados de la variable sobre las predicciones realizadas por distintos métodos sin imponer ninguna restricción. No obstante la combinación del modelo A sería insesgada si las predicciones individuales los fuesen, pero no hay ninguna garantía de que con predicciones individuales insesgadas, la predicción combinada también lo sea.

Analizando el modelo C de forma similar que en A y en B se obtiene que

$$e_{CB}^2 \geq e_{CA}^2 \geq e_{CC}^2$$

Granger y Ramanathan (1984) han comprobado que la ponderación obtenida bajo el enfoque de minimizar la varianza de los errores de predicción es igual que estimar los coeficientes de una regresión restringida a que estos sumen la unidad, con la misma condición de insesgadez de las predicciones individuales que imponen Bates y Granger (1969). Es más, si no se somete a ninguna restricción se puede producir predicciones combinadas más precisas<sup>96</sup>, y si se deja la posibilidad de incluir un término independiente en la regresión, no será preciso imponer que las predicciones de cada método individual sean insesgadas.

### 3.3 Técnicas basadas en metodología bayesiana

El desarrollo de los métodos de combinación bajo este enfoque requiere que tanto la variable a predecir como las predicciones individuales sean consideradas variables aleatorias. Así, la variable a predecir  $Y_{t+h}$  es una variable aleatoria definida de tal manera que sus valores representan las posibles opciones en que puede el fenómeno aleatorio antes de llevar a cabo un experimento, y donde la probabilidad de los diferentes valores se encuentra expresado por una función de probabilidad  $p(y_{t+h})$  que

---

<sup>96</sup> Si las predicciones individuales están fuertemente correlacionadas positivamente, algunos de los coeficientes estimados en la regresión que las combina pueden ser negativos y entonces, es posible, que la predicción combinada falle fuera del rango de las predicciones individuales. Y que por tanto actúe mal en el rango de predicción. En este caso habría que preguntarse si los modelos con ponderación deberían seguir usándose o no. No obstante, la superioridad de estos modelos frente aquellos cuya restricción es  $l^T \hat{\alpha} = 1$ , ha sido estudiada por numerosos autores. Por ejemplo Thomson (1982) estudia las propiedades de los estimadores de MCO restringidos a la no negatividad bajo diferentes asunciones y Peña (1997) considera la combinación de información estudiando la estructura y las propiedades de los estimadores que forman parte de la combinación.

recibe el nombre de función de probabilidad a priori de  $Y_{t+h}$  (si  $Y_{t+h}$  es discreta) o función de densidad de probabilidad a priori,  $f(y_{t+h})$  si la variable es continua.

El objetivo será entonces modificar esa función a priori a partir de la información suministrada por la muestra de predicciones individuales. La función modificada será una función a posteriori que llamaremos “función a posteriori (combinada)”, para recalcar que es el resultado de una combinación de predicciones individuales. De la esperanza de esa “función a posteriori (combinada)” se obtendrá un valor concreto de  $Y_{t+h}$ , que será  $\hat{y}_{t+h,t}^c$ .

La “función a posteriori (combinada)” de  $Y_{t+h}$ , obtenida partir de la información suministrada por las predicciones individuales, para cualquier horizonte  $h$  se calcula empleando la metodología bayesiana como:

$$f(y_{t+h} / \hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N) = \frac{f_{Y_{t+h}}(\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N / y_{t+h}) f_{Y_{t+h}}(y_{t+h})}{\int f_{Y_{t+h}}(\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N / y_{t+h}) f_{t+h}(y_{t+h}) d(y_{t+h})} \text{ si } Y_{t+h}$$

continua o

$$f(y_{t+h} / \hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N) = \frac{f_{Y_{t+h}}(\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N / y_{t+h}) p_{Y_{t+h}}(y_{t+h})}{\sum_{Y_{t+h}} f_{Y_{t+h}}(\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N / Y_{t+h}) p_{Y_{t+h}}(y_{t+h})} \text{ si } Y_{t+h} \text{ es}$$

discreta.

La especificación de  $p(y_{t+h})$  o  $f(y_{t+h})$  depende de la convicción del investigador-predictor antes de que la información muestral se encuentre disponible. Esta convicción se puede basar en cualquier tipo de información, y en concreto en nuestro caso se tomará como información la muestra predicciones individuales  $\hat{Y}_{t+h,t} = (\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N)$ . Si  $f_{Y_{t+h}}(\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N / y_{t+h})$  representa la función de densidad de probabilidad condicional de variables aleatorias, la cuales representan evidencia muestral en función de una alternativa fija de  $y_{t+h}$  en  $Y_{t+h}$ , esta función recibe el nombre de función de verosimilitud puesto que representa el grado de concordancia del resultado muestral  $\hat{Y}_{t+h,t} = (\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N)$ , dado el valor  $y_{t+h}$  de  $Y_{t+h}$ .

Cuando la información a priori con respecto a los valores de  $Y_{t+h}$  se combina con la información muestral de predicciones individuales, el resultado es un conjunto de información modificada con respecto a la variable aleatoria  $Y_{t+h}$ . Es decir, la

combinación de la distribución a priori y de la función de verosimilitud origina una distribución condicional para  $Y_{t+h}$ , dado el resultado muestral de predicciones individuales que se conoce como la distribución a posteriori de  $Y_{t+h}$  y que refleja la probabilidad modificada del respecto a la variable aleatoria  $Y_{t+h}$  después de obtener información muestral:  $f(y_{t+h} / \hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N)$ .

Según las diferentes formas que se den tanto a las funciones de verosimilitud como a la función a priori del investigador<sup>97</sup>, obtendremos distintas expresiones para la distribución a posteriori (combinada) de  $Y_{t+h}$ .

La solución del problema bayesiano, bajo determinadas condiciones es la misma que la por el enfoque clásico: Así por ejemplo Dickinson (1972) comprueba que si todos las distribuciones de las predicciones individuales son normales, y entonces  $f(\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N / y_{t+h})$  es una distribución multinormal, la mejor combinación será la misma que la obtenida por Bates y Granger (1969) y Newbold y Granger (1974)<sup>98</sup>.

Por otra parte, si  $f(\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N / y_{t+h})$  es una distribución log-normal<sup>99</sup>, derivando la expresión para este caso y calculando la esperanza, la mejor estimación para calcular la combinación será:

$$\hat{y}_{t+h,t}^c = \prod_{i=1}^N (\hat{y}_{t+h,t}^i)^{w_i} e^{(\sigma^2/2)} \text{ que es la media geométrica.}$$

Otro de los pioneros en emplear el enfoque bayesiano en la combinación de predicciones fue Bunn (1975), quien para una combinación de dos predicciones,

<sup>97</sup> Una de las mayores críticas a esta metodología, es la dificultad para conocer esas distribuciones de probabilidad así como la dificultad para calcular las distribuciones a posteriori (Winkler (1981))

<sup>98</sup> Si las predicciones individuales están distribuidas normalmente entorno al verdadero valor de  $y_{t+h}$  y entonces  $f(\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N / y_{t+h})$  es una distribución multinormal:

$$= \frac{1}{\sqrt{(2\pi)^N |\Sigma|}} \exp \left[ -\frac{1}{2} (\hat{Y}_{t+h,t} - y_{t+h} I^T)^T \Sigma^{-1} (\hat{Y}_{t+h,t} - y_{t+h} I^T) \right]$$

donde  $\Sigma^{-1}$  es la matriz de varianzas covarianzas.

Desarrollando el término

$\left[ -\frac{1}{2} (\hat{Y}_{t+h,t} - y_{t+h} I^T)^T \Sigma^{-1} (\hat{Y}_{t+h,t} - y_{t+h} I^T) \right]$  se obtiene que la mejor estimación, si los errores de la predicción de las predicciones individuales tienen una distribución normal, de la predicción combinada (obtenida a partir de la esperanza de  $f(y_{t+h} / \hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N)$ ) es la misma que la que obtienen Bates y Granger (1969) y Newbold y Granger (1974).

<sup>99</sup> Por ejemplo este puede ser el caso que ocurre cuando las técnicas individuales tienden a infravalorar las predicciones y cometen errores muy pequeños y sin embargo cuando comenten una sobrevaloración el error es muy grande.

interpreta los pesos de la combinación como variables aleatorias que indican la actuación relativa<sup>100</sup> del método  $i$  con respecto a  $j$ .

Así, considera los pesos como v.a. que representan el número relativo de veces que un método actúa mejor que otro, y por tanto cada elemento de  $\alpha$  puede ser representado como una fracción  $a_1 = \frac{a_1}{a_1 + a_2}$  donde  $a_1$  es el número de veces que el método 1 actuó mejor que 2 en  $k$  momentos o situaciones de predicción  $\left( k = \sum_{i=1}^2 a_i \right)$ .

Para dos métodos de predicción individuales se considera que esta fracción puede ser convenientemente ajustada a una distribución beta  $\alpha \approx \beta(k|a_1, a_2)$ . La distribución a priori de la misma puede ser  $\alpha \approx \beta(k|1/2, 1/2)$ , que correspondería a una situación en la que se daría igual ponderación a los métodos de predicción, porque no hay información suficiente de que una técnica sea superior a otra, y modificarla a partir de las predicciones pasadas  $\hat{Y}_t = (\hat{Y}_t^1, \hat{Y}_t^2, \dots, \hat{Y}_t^N)$ , y entonces  $k=n$ , obteniéndose así distribución a posteriori (combinada) de las ponderaciones.

Esta distribución puede ir modificándose a medida que se vaya disponiendo de nuevas experiencias de predicción. Con cada nueva predicción de  $Y_{t+h}$  se puede representar si el método 1 ha actuado mejor que 2 ( $\delta = 1$ ) o no ( $\delta = 0$ ) y entonces  $\delta$  puede ser considerada como una variable de Bernoulli. Después de  $z$  predicciones realizadas la distribución posterior de  $\alpha$  será  $\alpha \approx \beta(k|a_1 + s_z, a_2 + z - s_z)$  con  $s_z = \sum_{i=1}^z \delta_i$ . Asumiendo

una forma de coste o desutilidad del error cuadrática, los pesos óptimos para 1 y 2 son  $\hat{\alpha}_1$  (la media posterior de  $\alpha$ ) y  $1 - \hat{\alpha}_1$  respectivamente. Después de  $z$  realizaciones,

$$\hat{\alpha}_1 = \frac{(a_1 + s_z)}{(a_1 + a_2 + z)}.$$

El caso estudiado para dos predicciones individuales es extendido para el caso de  $N$  predicciones. La distribución análoga a la distribución beta para el caso multivariante es conocida como la distribución Dirichlet y en este caso puede usarse para conjugar las probabilidades en el proceso multinomial.

---

<sup>100</sup> La manera de evaluar la calidad de actuación de un método con respecto a otro puede ser muy arbitraria, de ahí que es conveniente el uso de indicadores objetivos, como por ejemplo los basados en errores de predicción estudiados anteriormente.

Así si cada  $a_i$  (donde  $a_i = \frac{a_i}{\sum_{i=1}^N a_i}$ ) es el número de veces que el método  $i$  actuó

mejor que los  $N-1$  métodos restantes en  $k$  momentos o situaciones de predicción siendo  $\left(k = \sum_{i=1}^N a_i\right)$ ,  $\alpha$  (siendo  $\sum_{i=1}^N \alpha_i = 1$ ) sigue una distribución de Dirichlet. Después de  $z$  realizaciones la media posterior de  $a_i$  (que se ha comparado con un método  $j$ ) será:

$$\bar{\alpha}_{i,j} = \frac{(a_i + s_{i,z})}{\left(\sum_{i=1}^N a_i + z\right)} \text{ (siendo } i \neq j \text{) donde } s_{i,z} \text{ es el número de veces que la técnica } i$$

actuó mejor las nuevas  $z$  realizaciones.

Cada método  $i$  se revisa con los  $j$  métodos obteniendo  $\binom{N}{2}$  ponderaciones  $\bar{\alpha}_{i,j}$ .

Si llamamos  $\bar{\alpha}_i$  a los pesos normalizados, el ajuste sobre todo el conjunto se obtiene

como  $\hat{\alpha}_i = \frac{\bar{\alpha}_i}{S} \prod_{j=1}^{i-1} ((1 - \bar{\alpha}_j) / \bar{\alpha}_j)$  con  $S = \sum_{i=1}^N \bar{\alpha}_i$ . Se trata por tanto de descomposiciones

de beta para cada par de métodos dos a dos.

Bordley (1982) subraya que el trabajo de Bunn (1975) es contrario a la metodología Bayesiana, la cual implica el uso de toda la información relevante, pues únicamente emplea la actuación relativa de un método con todos los demás para la actualización de las ponderaciones, ignorando la actuación relativa entre todos los métodos.

Sin embargo, estos métodos, al no tener en cuenta la correlación entre predicciones, son menos sensibles a los cambios de la matriz de covarianzas del error que la metodología de la regresión o de mínima varianza, y por lo tanto pueden evitar problemas que ello ocasiona en la estimación de las ponderaciones.

A nuestro criterio, si bien estos métodos no requieren más que saber que método actuó mejor que otro, no tienen en cuenta “cuánto” actuaron mejor.

### 3.4 Problemas en la combinación y técnicas desarrolladas

La estimación eficiente de los métodos de combinación hasta ahora desarrollados, dependerá del cumplimiento de los supuestos teóricos sobre los que se basan.

En los métodos basadas en Mínimos Cuadrados Ordinarios, las hipótesis pueden clasificarse según que afecten a la propia estructura del modelo (hipótesis estructurales) o a la perturbación aleatoria. Entre las alteraciones de tipo estructural podemos considerar:

- Muestras pequeñas
- Relaciones de colinealidad entre predicciones individuales
- Predicciones individuales atípicas
- Series no estacionarias
- Cambio estructural
- Especificación errónea
- Regresores estocásticos

y por lo que se refiere a la perturbación aleatoria consideramos los siguientes problemas:

- Media no nula
- Heterocedasticidad
- Autocorrelación
- No Normalidad

De la violación de algunos de los supuestos o la presencia de problemas muestrales (que no sólo afectan a las técnicas basadas en MCO), se derivan una serie de modelos que permiten mejorar la estimación de las ponderaciones<sup>101</sup>. En este epígrafe hacemos mención a todas las hipótesis, aunque revisaremos con más detalle aquellas que han generado el desarrollo de otros métodos para la combinación de predicciones.

Conviene tener presente que en las técnicas bayesianas, los problemas surgen sobre todo por la dificultad de conocer las funciones de distribución de las variables y de la dificultad operativa para obtener la función a posteriori (combinada), por lo que sus soluciones estarán sujetas a la capacidad del investigador o al desarrollo de aplicaciones informáticas para estimarlas.

---

<sup>101</sup> Si se usa un método de optimización inestable, los resultados podrían ser mucho peores que el método simple de usar ponderaciones iguales.

### 3.4.1 Alteración de las hipótesis estructurales

A continuación estudiamos el incumplimiento sobre las hipótesis estructurales, su origen, sus efectos sobre la predicción y las soluciones desarrolladas.

#### 3.4.1.1 Muestras pequeñas

El cálculo eficiente de las ponderaciones  $\alpha$  exige, como requisito mínimo, que el número de observaciones ( $n$ ) sea superior al número de parámetros en el modelo (que al menos son  $N$ ), es decir ( $n > N$ ), aumentando la garantía de la estimación ha medida que esta diferencia aumenta (grados de libertad).

Sin embargo, en ocasiones, bien porque la situación a predecir es nueva o se trata de una situación con poca historia, no es posible disponer de suficiente información para el cálculo eficiente de las ponderaciones.

La existencia de muestras pequeñas produce estimaciones que, aunque insesgadas y eficientes con respecto a cualquier otro método de estimación lineal, presentan varianzas comparativamente elevadas respecto de las que se obtendrían con tamaños muestrales relativamente superiores. Como consecuencia, los intervalos de predicción resultan demasiado amplios y los contrastes de significación estadística no son conclusivos al respecto (las varianzas elevadas de los estimadores hacen difícil rechazar cualquier contraste de no significación de los parámetros -contrastos  $t$  de nulidad).

Una de las soluciones pasaría por disminuir el número de predicciones individuales que se incluyen en el modelo, de tal manera que al ir elevándose los grados de libertad por encima del mínimo, el modelo pueda ir cobrando significación en las variables que se mantengan.

Como alternativa a las técnicas de regresión podríamos emplear las alternativas derivadas bajo otros enfoques (bayesiano, teoría de la información), si bien todos ellos también requieren de información muestral suficiente.

Si no es posible disponer de información objetiva suficiente, será entonces conveniente el empleo de información subjetiva u otra referente a la precisión con que los  $N$  modelos individuales predicen otra variable  $X$ , relacionada con la variable que nos interesa predecir,  $Y$ , y para la que existan observaciones suficientes tal que  $n > N$ . Es decir, los pesos se obtendrían a partir de la información proporcionada por la estimación de  $\beta$  en  $X_t = \beta \hat{X}_t + u_t$ , donde  $\hat{X}_t$  es la matriz de las  $N$  predicciones individuales de  $X$  para  $n$  periodos.

Gupta y Wilton (1987) proponen un método para la inclusión de información relevante e información empírica en la predicción proporcionándonos un sistema de obtención de ponderaciones independientes del número de observaciones previas a la realización de la predicción.

El método que proponen Gupta y Wilton (1987) es el llamado Matriz de ventajas (Odds-Matrix, OM) donde para estimar las ponderaciones de la predicción combinada se emplea una matriz de ventajas relativas por pares de métodos sobre la actuación de éstos.

Si se tiene una matriz de pesos  $\beta$  la matriz de comparación entre métodos dos a dos (OM) se construye como:

$$OM = \begin{pmatrix} \frac{\beta_1}{\beta_1} & \frac{\beta_1}{\beta_2} & \dots & \dots & \frac{\beta_1}{\beta_N} \\ \frac{\beta_2}{\beta_1} & \frac{\beta_2}{\beta_2} & \dots & \dots & \frac{\beta_2}{\beta_N} \\ \frac{\beta_3}{\beta_1} & \frac{\beta_3}{\beta_2} & \dots & \dots & \frac{\beta_3}{\beta_N} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\beta_N}{\beta_1} & \frac{\beta_N}{\beta_2} & \dots & \dots & \frac{\beta_N}{\beta_N} \end{pmatrix}$$

donde cada  $om_{ij}$  puede ser interpretada<sup>102</sup> como la ventaja que el modelo  $i$  tendrá sobre el  $j$ . Esta OM se puede aproximar una vez que se estime  $X_t = \beta \hat{X}_t + u_t$ , y emplear la información suministrada para la estimación de  $\alpha$ .

Realmente, Gupta y Wilton (1987) estiman únicamente la matriz de ventajas de cada método comparado con los otros uno a uno. Es preciso entonces normalizar los pesos de OM (que llamaremos  $\overline{om}_i$ ) y el ajuste sobre todo el conjunto se obtiene como

$$\hat{\beta}_i = \frac{\overline{om}_i}{S} \prod_{j=1}^{i-1} ((1 - \overline{om}_j) / \overline{om}_j) \text{ con } S = \sum_{i=1}^N \overline{om}_i .$$

Las ventajas que presenta este método, además de resolver el problema para predicciones con poca historia y refiriéndonos a la falta de información, es que en ocasiones no es posible obtener las desviaciones del error de predicción de cada modelo y sin embargo sí es posible obtener el número de veces que un modelo funciona mejor que otro.

---

<sup>102</sup> Obsérvese que  $om_{ij} = 1 / om_{ji}$

### 3.4.1.2 Relaciones de colinealidad de predicciones

En ocasiones las predicciones individuales están generadas a partir de una base informativa muy homogénea (por ejemplo las instituciones que realizan predicciones pueden compartir la misma información, hecho que cada vez es más frecuente debido al fuerte desarrollo de las TIC), o presentan una tendencia común (este caso podría ocurrir si las predicciones para  $t+h$  pueden realizarse en diferentes estadios del tiempo, y entonces en las predicciones que elaboran distintos organismos exista una tendencia a converger a la predicción media<sup>103</sup>). Estos hechos, pueden ocasionar la existencia de una cierta dependencia entre las predicciones individuales:

Esta dependencia afecta a la matriz de varianzas-covarianzas de los estimadores ya que ésta depende muy directamente de la mayor o menor relación de las variables explicativas  $\hat{Y}_{t+h,t}$  del modelo. Concretamente, las varianzas de los estimadores crecen según aumenta la correlación entre las predicciones individuales.

Una elevada colinealidad produce una sobrestimación de los errores estándar de los coeficientes de regresión con lo que los estimadores, aunque insesgados, pueden sufrir errores considerables y también sufrir cambios sensibles al añadir nueva información, es decir, en el proceso de actualización del modelo poniendo nuevamente de manifiesto la inseguridad en la estimación. Además las varianzas elevadas de los estimadores hacen difícil rechazar cualquier contraste de no significación de los parámetros (contrastes t de nulidad)<sup>104</sup>.

La presencia de la multicolinealidad conduce al empleo de técnicas alternativas de estimación entre las cuales las más empleadas son las Componentes Principales y la Regresión *Ridge* o Cresta.

El análisis de Componentes Principales busca identificar singularidades cercanas de las variables independientes (es decir, de las predicciones individuales<sup>105</sup>) y

---

<sup>103</sup> Moreno y López (2002) muestran una evidencia empírica de cómo este fenómeno efectivamente ocurre en predicciones sobre el crecimiento económico para España realizadas por distintas instituciones tanto públicas como privadas.

<sup>104</sup> Si bien la multicolinealidad puede afectar a la significación de cada parámetro individual, no tiene porqué incidir en la validez del modelo en su conjunto, introduciendo elementos de duda sobre la influencia de cada variable en particular. Como consecuencia, un modelo con problemas de colinealidad puede resultar discutible a efectos de análisis estructural, pero puede emplearse como herramienta, siempre que para el periodo extramuestral se admita una correlación entre variables similar a la observada en el período muestral.

<sup>105</sup> Podría ocurrir que las predicciones efectuadas por instituciones públicas tuviesen características comunes frente a las efectuadas por las instituciones privadas, y en ese caso el análisis de componentes principales lo detectaría

determinar su valor predictivo. Su objetivo es eliminar las singularidades cercanas no predictivas de las variables dependientes e independientes estandarizadas.

Por su parte, la Regresión *Ridge* o Cresta supone cambiar un estimador insesgado por un estimador al cual se le agrega un sesgo.

Así si mediante MCO obtenemos los estimadores de  $Y_t = \alpha \hat{Y}_t + u_t$  como  $\hat{\alpha} = (\hat{Y}_t^T \hat{Y}_t)^{-1} \hat{Y}_t^T Y_t$ , el estimador de la regresión Cresta será  $\hat{\alpha}_r = (\hat{Y}_t^T \hat{Y}_t + rD)^{-1} \hat{Y}_t^T Y_t$ , donde D es una matriz diagonal cuyos elementos diagonales coinciden con los de la matriz  $\hat{Y}_t^T \hat{Y}_t$ , y el escalar r se escoge arbitrariamente. En la práctica este r (que puede ser encontrado por inspección o estimado) debe escogerse de manera que los estimadores resultantes sean “estables” con respecto a pequeñas variaciones de r<sup>106</sup>. Si bien es fácil demostrar que el estimador cresta es sesgado  $E(\hat{\alpha}_r) = \sigma^2 (\hat{Y}_t^T \hat{Y}_t + rD)^{-1} \hat{Y}_t^T Y_t \alpha$ , a su favor puede decirse que su matriz de covarianzas, que viene dada por la expresión  $Var(\hat{\alpha}_r) = \sigma^2 (\hat{Y}_t^T \hat{Y}_t + rD)^{-1} \hat{Y}_t^T Y_t (\hat{Y}_t^T \hat{Y}_t + rD)^{-1}$  es inferior a la del estimador de MCO.

Además de emplear técnicas alternativas ya mencionadas (Gupta y Wilton (1987)<sup>107</sup>) a nuestro criterio, el empleo de componentes principales podría ser muy útil en la combinación de predicciones. El estimador cresta puede tener menor error cuadrático medio que el de mínimos cuadrados ordinarios, pero este error es función de los parámetros desconocidos que estamos intentando estimar (en este caso, con un sesgo evidente) y por tanto es poco probable que este criterio resulte útil.

### 3.4.1.3 Predicciones atípicas

Si existe alguna predicción atípica, entonces los mínimos cuadrados ordinarios pueden producir estimadores con sesgos arbitrariamente grandes.

Si las condiciones del teorema de Gauss Markov son satisfechas por la población, pero la muestra contiene observaciones atípicas, entonces los mínimos cuadrados ordinarios pueden producir estimadores con sesgos arbitrariamente grandes, mientras

<sup>106</sup> Los autores de esta técnica sugieren tomar en principio un valor pequeño, por ejemplo 0,01, y escoger después valores cada vez mayores hasta que los coeficientes se estabilicen.

<sup>107</sup> Este método, al no tener en cuenta la correlación entre los métodos, es menos sensible a los cambios de la matriz de covarianzas que la metodología de la regresión o de mínima varianza

que un procedimiento robusto, por construcción, limita la influencia de cualquier observación<sup>108</sup>.

Hallman y Kamstra (1989) sugieren que el uso de estimaciones robustas en la regresión que combina las predicciones individuales puede ayudar a mejorar la predicción. Sin embargo, el estimador de mínimos cuadrados sigue siendo asintóticamente el mejor estimador lineal insesgado, de modo que en este contexto las técnicas robustas son importantes en muestras pequeñas.

#### 3.4.1.4 Cambio estructural: Coeficientes cambiantes en el tiempo

Una de las hipótesis básicas es la de la constancia de los parámetros de los modelos de regresión, es decir, la existencia de una estructura única, válida para todo el período de observación, y que suponemos mantenida para el horizonte de predicción.

Cuando se trabaja con series temporales, o situaciones con historia, es habitual encontrarse con cambios institucionales y del contexto económico que transforma la estructura interna de funcionamiento de los fenómenos, y que deberían de considerarse en los modelos. Si no es así puede producirse el incumplimiento de la hipótesis de constancia estructural y en resumen, la mezcla de dos estructuras diferentes dará lugar, en relación con cada una de las estructuras componentes a estimadores sesgados, inconsistentes<sup>109</sup>, ineficientes y con errores en la aplicación de los contrastes de significación.

Además cada investigador puede tratar de mejorar su método de predicción con el tiempo, lo cual sugiere que el comportamiento de las predicciones podría estar cambiando en el tiempo y por tanto sus ponderaciones en la combinación probablemente también lo deberían hacer (a medida que el modelo individual mejora la predicción, se le debería de dar mayor ponderación en la combinación).

La solución más inmediata es la de permitir cambiar las ponderaciones en el tiempo  $y_{t+h,t}^c = f\left(\hat{Y}_{t+h,t}(\alpha_{t+h})\right) + \mu_{t+h}$ , que puedan recoger ese cambio de estructura.

---

<sup>108</sup> Las técnicas robustas proveen resultados similares a MCO cuando se tienen unos errores que provienen de una distribución normal, sin embargo las técnicas robustas presentan la ventaja de limitar la influencia de valores atípicos. En Greene (1999) pueden verse las propiedades estadísticas del estimador de MCO en muestras finitas, y la construcción de estimadores “robustos”.

<sup>109</sup> Al mezclar indebidamente dos estructuras diferentes, los parámetros estimados para el conjunto serán una cierta media ponderada de los correspondientes a cada submuestra, por lo que la estimación conjunta es necesariamente sesgada e inconsistente con respecto al modelo adecuado para explicar el comportamiento de cada submuestra. Por otra parte, como recoge Pulido (1989), la propia heterogeneidad de los datos indebidamente mezclados tendrá un efecto de incrementar la varianza de los residuos, y por tanto de los estimadores de los residuos.

Sea cual sea el método empleado, la estimación de los parámetros se realizará a partir de la información pasada.

Debido a esto, Bates y Granger (1969) sugieren usar subconjuntos móviles según el retardo temporal considerado y la posibilidad de ir actualizando la información. Así la estimación de  $\hat{\alpha}_t^i$  puede obtenerse para el caso de 2 técnicas como:

$$\hat{\alpha}_t^1 = \frac{\sum_{t=n-v}^{n-1} (e_t^2)^2}{\sum_{t=n-v}^{n-1} \left( (e_t^1)^2 + (e_t^2)^2 \right)}$$

donde las ponderaciones se calculan utilizando las V

observaciones más recientes.

$$\hat{\alpha}_t^1 = \frac{\sum_{t=n-v}^{n-1} \left( (e_t^2)^2 - e_t^1 e_t^2 \right)}{\sum_{t=n-v}^{n-1} \left( (e_t^1)^2 + (e_t^2)^2 - 2e_t^1 e_t^2 \right)}$$

s.a.  $0 \leq \hat{\alpha}_t^1 \leq 1$  cuya filosofía es igual a la

anterior aunque no ignora la covarianza entre los errores de predicción de las diferentes predicciones individuales

$$\hat{\alpha}_t^1 = x\hat{\alpha}_{n-1}^1 + (1-x) \frac{\sum_{t=n-v}^{n-1} (e_t^2)^2}{\sum_{t=n-v}^{n-1} \left( (e_t^1)^2 + (e_t^2)^2 \right)}$$

con  $0 < x < 1$ , es un método adaptativo

que ignora la covarianza de los errores de predicción de las diferentes predicciones individuales.

Finalmente hay otros dos métodos (según se ignore o no la covarianza entre los errores) en los que se pondera la información más reciente

$$\hat{\alpha}_t^1 = \frac{\sum_{t=1}^{n-1} W^t (e_t^2)^2}{\sum_{t=1}^{n-1} W^t \left( (e_t^1)^2 + (e_t^2)^2 \right)}$$

con  $W \geq 1$

$$\hat{\alpha}_t^1 = \frac{\sum_{t=1}^{n-1} W^t \left( (e_t^2)^2 - e_t^1 e_t^2 \right)}{\sum_{t=1}^{n-1} W^t \left( (e_t^1)^2 + (e_t^2)^2 - 2e_t^1 e_t^2 \right)}$$

s.a.  $0 \leq \hat{\alpha}_t^1 \leq 1$

donde en todos los casos será importante la elección de  $v$ ,  $x$  y  $W$ .

El cálculo de estos pesos puede ser extendido a la combinación de N predicciones.

Sessions y Chatterjee (1989), emplean un modelo de espacio de estados para obtener ponderaciones cambiantes en el tiempo. Más concretamente, estos autores sugieren emplear el siguiente modelo y variaciones de él :

$$Y_t = \hat{Y}_{t+h,t} \alpha_t + u_t$$

$$\alpha_t = R \alpha_{t-1} + w_t$$

siendo R es una matriz conocida cuya inclusión es arbitraria.

En los modelos de regresión lo más habitual es emplear modelos de transición, de tal manera que los pesos puedan incluir el cambio de estado en t+h.

Si denotamos por  $I_{t+h}^1$  e  $I_{t+h}^2$  a los 2 diferentes estados o regímenes del estado de la economía en t+h, la forma general de un modelo de este tipo aplicado a la combinación sería:

$$y_{t+h} = \hat{Y}_{t+h,t} (\hat{\alpha} + \lambda_{t+h} \gamma) + \varepsilon_{t+h} \quad 0 \leq \lambda_{t+h} \leq 1,$$

donde  $\alpha_{t+h} = \hat{\alpha} + \lambda_{t+h} \gamma$ , se calcula a partir de los pesos ya estimados e incluye la posibilidad de incluir el cambio de régimen que se prevé en t+h a través de  $\lambda_{t+h}$ .

$\{\lambda_{t+h} | h \in [1, \dots, T]\}$  que recogería el cambio estructural en t+h de la siguiente manera:

$$\lambda_{t+h} = 1 \text{ cuando } t+h \in I_{t+h}^1$$

$$\lambda_{t+h} = 0 \text{ cuando } t+h \in I_{t+h}^2.$$

Si  $\lambda_{t+h}$  es una función continua entre 0 y 1, entonces los pesos cambiarán suavemente de un régimen a otro y en este caso el modelo se llamará modelo de transición suave.

Para estos modelos el principal problema práctico es la estimación de los regímenes asociados con el modelo de regresión cuando dichos regímenes son desconocidos<sup>110</sup>. Para resolver el problema se han considerado varias alternativas; la primera de ellas asume que el error retardado de las predicciones individuales puede ser

---

<sup>110</sup> Es decir, se desea conocer qué periodos del tiempo pertenecen al subconjunto  $I^1$  y cuáles al subconjunto  $I^2$ .

útil para conocer el régimen<sup>111</sup> y la segunda asume que alguna variable económica relevante puede ser útil para aproximarlos.

No obstante, los errores de predicción para  $t+h$ , cuando se deban calcular las ponderaciones, son desconocidos, por lo que hay que emplear información sobre los errores pasados. Por lo tanto en la regresión que se emplea para la predicción en  $t+h$  han de utilizarse los parámetros estimados a partir de:

$$Y_t = \hat{Y}_t(\hat{\alpha} + \lambda_t \gamma) + \varepsilon_t \quad 0 \leq \lambda_t \leq 1, \text{ entonces } \{\lambda_t | t \in [1, \dots, n]\} \text{ es el patrón de transición de la forma}$$

$$\lambda_t = 1 \text{ cuando } t \in I_t^1 \text{ y}$$

$$\lambda_t = 0 \text{ cuando } t \in I_t^2$$

donde  $I_t^1$  e  $I_t^2$  son los diferentes estados o regímenes del estado de la economía en el pasado.

Deutsch, Granger y Teräsvirta (1994) para decidir el estado de la naturaleza se fijan los signos de los errores retardados, su magnitud o su tamaño relativo en el pasado. Es decir decide cuándo ese estado de la naturaleza será 1 o 2,

1.  $I_t^1$  es  $e_t^1 \geq 0$
2.  $I_t^1$  es  $e_t^2 \geq 0$
3.  $I_t^1$  es  $|e_t^1| \geq c \sum_{t=1}^n |e_t^1| / n$
4.  $I_t^1$  es  $|e_t^2| \geq c \sum_{t=1}^n |e_t^2| / n$
5.  $I_t^1$  es  $e_t^1$  es un proceso AR (1)
6.  $I_t^1$  es  $e_t^2$  es un proceso AR (1)

Otros modelos de este tipo serían aquellos en los que el estado viene determinado por una variable económica  $X_t$ . Los dos modelos propuestos serían

1.  $I_t^1$  es  $X_t > 0$

---

<sup>111</sup> Si el error cometido por una predicción individual en  $t+h$  es inusualmente grande o si es negativo, podría indicar que la economía está en estado 1.

$$2. I_t^1 \text{ es } |X_t| \frac{\sum_{t=1}^n |W_t|}{n}$$

Puede ocurrir que el paso de un estado a otro se produzca de manera gradual, esto es  $\lambda$  es una función continua entre 0 y 1, por lo que el modelo general sería:

$$Y_{t+h} = \hat{Y}_{t+h} \left( \hat{\alpha} + \lambda_{t+h} (1 - e^{\lambda_{t+h-1}})^{-1} \right) + \varepsilon_{t+h} \quad 0 \leq \lambda_{t+h} \leq 1 \text{ que implica estimar:}$$

$$Y_t = \hat{Y}_t \left( \hat{\alpha} + \lambda_t (1 - e^{\lambda_{t-1}})^{-1} \right) + \varepsilon_t \quad 0 \leq \lambda_t \leq 1$$

donde  $V_{t-1}$  es alguna función del error de predicción retardado. Algunas posibilidades son:

1.  $V_{t-1}$  es  $e_{t-1}^1 \geq 0$
2.  $V_{t-1}$  es  $e_{t-1}^2 \geq 0$
3.  $V_{t-1}$  es un proceso AR de  $e_t^1$
4.  $V_{t-1}$  es un proceso AR de  $e_t^2$

Otro método de obtener ponderaciones cambiantes en el tiempo sería resolver la regresión con las  $v$  observaciones más recientes. Es más, podemos iterar sobre  $V$  para producir  $V^*$  que minimice el error cuadrático medio del error de predicción combinado.

Otra posibilidad es emplear mínimos cuadrados ordinarios ponderados que permitan ponderar más las observaciones más recientes.

Se trata de obtener  $\alpha$  que minimice:

$e_t^{cT} W e_t^c$  donde  $W$  es una matriz ( $n \times n$ ) de ponderaciones :

$$e_t^{cT} e_t^c = \sum_{t=1}^n \left( y_t - \sum_{i=1}^N \hat{y}_{t+h,t}^i \alpha_i \right)^2$$

$$\text{en lugar de } e_t^{cT} e_t^c = \sum_{t=1}^n \left( y_t - \sum_{i=1}^N \hat{y}_{t+h,t}^i \alpha_i \right)^2$$

En la mayor parte de las aplicaciones se suele suponer que la matriz de ponderaciones es diagonal, es decir  $W = \text{diagonal}(w_{11}, w_{22}, \dots, w_{nn})$  lo que implica entonces minimizar:

$$e_t^{cT} W e_t^c = \sum_{t=1}^n w_{tt} \left( y_t - \sum_{i=1}^N \hat{y}_{t+h,t}^i \alpha_i \right)^2$$

Una manera sencilla de dar menor relevancia a la información más alejada en el tiempo es especificar  $W = diagonal(w_1, w_2, \dots, w_n)$  donde  $w_t \geq w_{t-1} \quad \forall t = 1 \dots n$

Así por ejemplo Diebold y Pauly (1987) proponen:

$$W = N \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & n \end{pmatrix}$$

$$W = N \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda^{n-1} & 0 \\ 0 & 0 & \dots & 0 & \lambda^n \end{pmatrix} \quad \text{para } \lambda \geq 1$$

$$W = N \begin{pmatrix} 1^\lambda & 0 & \dots & 0 \\ 0 & 2^\lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & n^\lambda \end{pmatrix} \quad W = diag[w_t] = [Nt^\lambda] \quad \text{donde } N, \lambda > 0 \quad \text{y}$$

$$\frac{d^2 w_t}{dt^2} = N \lambda (1 - \lambda) t^{(\lambda-2)} \begin{cases} > 0 & \text{si } \lambda > 1 \\ < 0 & \text{si } \lambda < 1 \end{cases}$$

$$w_t^*(\lambda) = \begin{cases} (t^\lambda - 1) / \lambda & \text{si } 0 < \lambda \leq 1 \\ \ln t & \text{si } \lambda = 0 \end{cases}$$

Es decir:

1. Pesos iguales  $w_t = 1$  para todo t
2. Lineal  $w_t = t$  para todo t
3. Geométrico  $w_t = \lambda^{n-t}$ ,  $1 < \lambda \leq 1$ , o  $w_t = \lambda^t$ ,  $\lambda \geq 1 \Rightarrow \lambda > 1$
4.  $t^\lambda$  (t-landa):  $w_t = t^\lambda$ ,  $\lambda > 0$
5. Box-Cox:  $w_t^*(\lambda) = \begin{cases} (t^\lambda - 1) / \lambda & \text{si } 0 < \lambda \leq 1 \\ \ln t & \text{si } \lambda = 0 \end{cases}$

Entonces según el  $\lambda$  escogido se debería obtener

$$\min_{\lambda, \alpha} e_i^{cT} W e_i^c$$

$$\min_{\lambda, \alpha} \sum_{i=1}^n w_t^*(\lambda) \left( y_t - \sum_{i=1}^N \hat{y}_t^i \alpha_i \right)^2$$

### 3.4.1.5 Procedimientos de combinación para series no estacionarias

Hasta ahora todos los procedimientos de combinación presentados exigen que la serie  $Y_t$  sobre la que se van a realizar predicciones sea estacionaria. Sin embargo, en la práctica, muchas de las series macroeconómicas no son estacionarias y por lo tanto es necesario incorporar esta característica en el procedimiento de combinación.

En el caso de series no estacionarias, si se supone que los predictores candidatos para la combinación son *insesgados* y débilmente eficientes, Holden y Peel (1989) sugieren que no se debe incluir el intercepto en la regresión que combina las predicciones individuales propuesta por Granger y Ramanathan. Ahora bien, si las predicciones no son insesgadas dichos autores sugieren introducir una constante en la regresión y restringir las ponderaciones de las predicciones a que sumen la unidad. En decir, Holden y Peel sugieren llevar a cabo la regresión:

$$y_{t+h} = c + \hat{Y}_{t+h,t} \alpha \text{ sujeta a } \sum_{i=1}^N \alpha_i = 1$$

Otros procedimientos incorporan directamente el concepto de cointegración en la obtención de la regresión que combina las predicciones individuales. Cuando la serie es integrada de orden 1 de manera que los cambios de la serie son estacionarios, Hallman y Kamstra (1989) discuten cómo obtener un predicción combinada para  $y_{t+h}$ . Si  $Y_t$  es  $I(1)$ , cualquier predicción razonable  $\hat{y}_{t+h,t}^i$  ( $i=1, \dots, N$ ) debería ser cointegrada con ella, a través del vector de cointegración (1, -1). Si esto no ocurre, el error del predicción no será estacionario.

En una serie no estacionaria, la regresión combinada para  $h=1$  será:

$$\Delta y_{t+1} = m_0 + (e_{t+1,t}) \alpha + v_{t+1}, \text{ siendo } e_{t+1,t} \text{ el error de predicción al predecir } y_{t+1} \text{ en } t$$

$$\Delta y_{t+1} = m_0 + (\hat{Y}_{t+1,t} - Y_t) \alpha + v_{t+1}$$

Estimando los parámetros a partir de:

$$\Delta Y_t = m_0 + (\hat{Y}_{t,t-1} - Y_{t-1}) \alpha + v_t$$

donde cada elemento del vector  $(\hat{Y}_{t,t-1} - Y_{t-1})$  puede interpretarse como la predicción que cada técnica  $i$  realiza del cambio que se produce en la variable de interés al pasar de  $t-1$  a  $t$ . Así,  $\Delta Y_t$  es explicado a partir de la combinación lineal de las predicciones individuales de este cambio. Esto no es igual a la combinación lineal de las predicciones en niveles, pues si rescribimos la última ecuación,

$$Y_t = m_0 + Y_{t-1}(1 - l^T \alpha) + \hat{Y}_{t,t-1} \alpha + v_t$$

se observan varias diferencias; primero, cuando se combinan predicciones de una serie integrada, se debe incluir un retardo de la serie pronosticada. Segundo, los coeficientes de la regresión exceptuando  $m_0$ , deben ser restringidos a que sumen uno. Tercero, si los coeficientes  $\alpha_i$  suman la unidad la variable dependiente rezagada en la ecuación desaparece.

Para el caso de ponderaciones cambiantes de Deutsch, Granger y Teräsvirta (1994) los autores sugieren, siguiendo a Hallman y Kamstra (1989), modificar el primer modelo de combinación como:

$$Y_t = \hat{Y}_t(\hat{\alpha} + \lambda_t \gamma) + Y_{t-1} \gamma + \varepsilon_t \quad 0 \leq \lambda_t \leq 1$$

### 3.4.1.6 Regresores estocásticos y variables desplazadas

Si en la metodología bayesiana es necesario considerar las variables (variable a predecir y predicciones individuales) como aleatorias, en la metodología clásica se requiere que las variables explicativas tengan carácter no estocástico. Para que esta hipótesis se mantuviese deberíamos trabajar con experimentos controlables, en los que podrían repetirse unos valores prefijados de las predicciones individuales y explicar su incidencia sobre la predicción combinada. Sin embargo, cuando el experimento es el que proporciona unos datos como consecuencia del propio paso del tiempo o del comportamiento observado de ciertos individuos, el mismo experimento es irrepetible, por tanto no controlable, tal y como ocurre en numerosos estudios empíricos dentro del ámbito económico.

No obstante, cuando las variables explicativas son las predicciones individuales efectuadas por técnicas, por ejemplo de series temporales, el experimento sí podría ser controlable. Se podría ver cómo afecta a la predicción combinada el cambio de técnicas empleadas, con lo que podríamos repetir observaciones para el mismo período de tiempo.

Cuando las predicciones están basadas en las encuestas de opinión, la hipótesis es de más difícil cumplimiento, puesto que las observaciones no pueden repetirse (se pregunta a individuos). No obstante, si en los modelos de series temporales incluyésemos variables endógenas desplazadas o fuese un modelo multiecuacional

(MARMA) la hipótesis no se mantiene al depender las observaciones del comportamiento de otras variables o de perturbaciones aleatorias.

Sin embargo, cuando se trabaja con series temporales, el empleo de variables desplazadas está justificado, ya que con datos temporales las relaciones entre variables solo son entendibles en un cierto contexto dinámico, pues pueden existir efectos desplazados que se presentan con intensidad creciente cuanto menor es el lapso del tiempo sobre el cual queda definido el modelo (mensual, trimestral, anual, etc).

### **3.4.1.7 Especificación errónea del modelo**

Cuando la variable a predecir tenga una relación lineal con las predicciones individuales que tienen un efecto destacable sobre ella, entonces la especificación del modelo será correcta. Sin embargo en la práctica de la combinación econométrica difícilmente puede tenerse la seguridad de haber conseguido una especificación correcta, bien sea porque la forma funcional no debiera ser lineal, bien porque se ha omitido alguna predicción individual significativa<sup>112</sup> o se ha incluido alguna irrelevante<sup>113</sup>.

Únicamente un proceso de iteración y perfeccionamiento en los modelos de combinación puede mejorar la especificación del modelo.

Por ejemplo, Donaldson y Kamstra (1996) describen un método de combinación basado en Redes Neuronales, el cual tiene en cuenta relaciones no lineales complejas que no son capturadas por los modelos lineales tradicionales.

### **3.4.2 Incumplimiento de Hipótesis sobre la perturbación aleatoria**

A continuación estudiamos las hipótesis sobre la perturbación aleatoria (necesarias para la predicción), el problema que puede originar su rechazo, los efectos sobre la predicción combinada y las soluciones desarrolladas.

Generalmente, el incumplimiento de alguna de estas hipótesis se encuentra relacionado con el incumplimiento de hipótesis estructurales, de ahí que la solución en la mayoría de las ocasiones pase por solucionar los problemas estructurales..

---

<sup>112</sup> La omisión de variables relevantes produce estimadores sesgados e inconsistentes, problema que no desaparece si aumenta el tamaño muestral.

<sup>113</sup> En este caso los estimadores de las variables correctamente incluidos son insesgados y consistentes. Sin embargo las varianzas estimadas son superiores a las que les correspondería con los consecuentes problemas de pruebas t-de significación. Además los grados de libertad pueden disminuir y aparecer multicolinealidad .

### 3.4.2.1 Media no nula

La aparición de una media no nula para la perturbación aleatoria puede admitirse como un efecto de la mala especificación del modelo. Si el problema se debe a la omisión de variables relevantes, además de no ser nula será variable al depender de las variaciones de las incidencias que hayan sido indebidamente excluidas en el modelo.

### 3.4.2.2 Homocedasticidad

Las causas que originan heterocedasticidad de las perturbaciones aleatorias son muy diversas. Por ejemplo es frecuente que a mayores valores de la variable pueda crecer la dispersión absoluta e incluso la relativa. Si las predicciones están basadas en encuestas de opinión a consumidores y empresarios (en datos transversales), la heterogeneidad de los datos puede ser más frecuente. Si las preguntas van referidas a evolución de la plantilla, expectativas de gasto, etc, puede ser habitual encontrarse oscilaciones directamente relacionadas con las oscilaciones en la renta de los individuos, tamaños de empresas, etc. Si las predicciones se basan en datos temporales, la heterocedasticidad puede ser menos habitual al ser los datos menos dispares entre sí.

Otras causas de la heterocedasticidad son la omisión de predicciones individuales relevantes (especificación errónea de modelo de combinación) y la existencia de un cambio estructural no considerado, ya que el término de error puede tener un comportamiento distinto de unos periodos a otros. La revisión del modelo de combinación y el empleo de aquellos que permitan cambiar las ponderaciones en el tiempo pueden ser las soluciones más inmediatas.

### 3.4.2.3 Errores autocorrelacionados

Otro aspecto relevante en la combinación de predicciones aparece cuando surgen errores correlacionados en la regresión que combina las predicciones.

La especificación errónea, posibles cambios de estructura, la correlación entre predicciones individuales, la posibilidad de realizar las predicciones en diferentes estadios  $s$  ( $s=1\dots h$ )  $\hat{y}_{t+h,t+s}^i$  de manera que se mantengan efectos de las predicciones en estadios pasados sobre las predicción en el estadio actual<sup>114</sup>, e incluso el propio manejo previo de datos (como el alisado de series si una predicción individual emplea esta

---

<sup>114</sup> Efectivamente esto ocurre si los individuos, organismos predictores manifiestan sus predicciones individuales al resto, esta información influye en la predicción en el siguiente estadio por motivos de aversión a dispersarse del resto, etc.

técnica) puede provocar autocorrelación en los residuos del modelo de regresión que combina las predicciones individuales.

En general puede ser una buena idea permitir que la correlación serial, en los métodos de regresión para la combinación, capture la dinámica en la variable que va a ser pronosticada y que no explican las distintas predicciones individuales. De otro modo se puede incorporar algún dinamismo en la variable a predecir que no pueda ser capturada por las predicciones individuales (retardos en la variables dependientes y en las predicciones).

Diebold (1988) muestra que si las predicciones individuales se encuentran correlacionadas<sup>115</sup>, una combinación dinámica de ellas puede generar mejores predicciones.

Coulson y Robins (1993), siguiendo a Hendry y Mizon (1978), señalan que una regresión combinante con errores serialmente correlacionados es un caso especial de una regresión de combinación la cual incluye rezagos de la variable dependiente y rezagos de los predicciones:

Uno de los métodos más difundidos para la eliminación de autocorrelación es el de Cochran-Orcutt<sup>116</sup>, en el que la estimación de

$$e_{t+h} = \rho e_{t+h-1} + v_{t+h}$$

se emplea para estimar

$$y_{t+h} = \alpha_0 + \hat{Y}_{t+h,t} \alpha + e_{t+h}$$

de tal manera que:

$$y_{t+h} = \alpha_0 (1 - \rho) + \hat{Y}_{t+h,t} \alpha - \hat{Y}_{t+h-1,t-1} \alpha \rho + y_{t+h-1} \rho + v_{t+h}$$

Reagrupando las ponderaciones esta expresión es similar a:

$$y_{t+h} = \beta_0 + \hat{Y}_{t+h,t} \gamma + \hat{Y}_{t+h-1,t-1} \psi + y_{t+h-1} \eta + v_{t+h}$$

donde  $\hat{Y}_{t+h-1,t-1}$  son los valores rezagados de los predicciones individuales

Como  $e_{t+h}$  se desconoce, para estimarla se emplea el pasado de las series de modo que las ecuaciones anteriores se transforman como:

$$Y_t = \alpha_0 + \hat{Y}_t \alpha + e_t$$

$$e_t = \rho e_{t-1} + v_t$$

<sup>115</sup> La correlación suele aparecer en modelos de regresión no restringidos cuando la suma de las ponderaciones son distintas a 1 (Holden y Peel, 1990).

<sup>116</sup> En Pulido (1993) se explica el método y el proceso iterativo que se requiere para estimar  $\rho$

$$Y_t = \alpha_0(1 - \rho) + \hat{Y}_t\alpha - \hat{Y}_{t-1}\alpha\rho + \rho Y_{t-1} + v_t$$

$$Y_t = \beta_0 + \hat{Y}_t\gamma + \hat{Y}_{t-1}\psi + Y_{t-1}\eta + v_t$$

Coulson y Robins (1993) sugieren que en el contexto de combinación de predicciones debería cumplirse que  $\psi = 0$  puesto que posiblemente valores rezagados de las predicciones individuales ( $\hat{Y}_{t-1}$ ) no serán útiles en el cálculo de la predicción de  $Y_t$ , mientras que  $\gamma$  y  $\eta$  probablemente serán significativamente diferentes de cero, Además, si las predicciones pasadas,  $\hat{Y}_{t-1}$ , son redundantes en la ecuación sin restricciones, con su eliminación se ganaría eficiencia en la estimación, y particularmente en el período postmuestreal. De acuerdo con la anterior, Coulson y Robins (1993) sugieren que la ecuación de regresión para la combinación de predicción debería ser de la forma:

$$Y_t = \beta_0 + \hat{Y}_t\gamma + Y_{t-1}\eta + v_t$$

En nuestra opinión, cuando las predicciones se realizan en diferentes estadios s las predicciones rezagadas sí tienen sentido, por lo que en ese caso el modelo debería de incluir esas predicciones rezagadas referidas a las predicciones en los estadios pasados en los que se realiza la predicción.

Las combinaciones más efectivas son logradas cuando no existe correlación positiva entre los errores de las predicciones individuales. La correlación negativa es bastante rara, pero cuando ocurre la ganancia en la combinación puede ser mayor.

#### 3.4.2.4 No normalidad y Matriz de covarianzas no escalar

El incumplimiento de las hipótesis de homocedasticidad y ausencia de correlación (nulidad de las covarianzas) implica que la matriz de varianzas-covarianzas de la perturbación aleatoria se reduzca a un escalar ( $I\sigma^2$ ). Se puede considerar que MCO es un caso particular de un modelo de regresión generalizado (MCG) que considerase la existencia de una matriz de covarianzas en las perturbaciones que pudiera tener tanto elementos distintos en la diagonal principal como valores no nulos fuera de ella.

En MCG se trata de obtener  $\alpha$  que minimicen:  $e_t^{cT} V e_t^c$ , es decir, la suma de los errores al cuadrado ponderados por V que es una matriz (nxn).

En la combinación de predicciones este tipo de modelos se ha empleado cuando se incumplía la hipótesis de constancia estructural. El incumplimiento de esta hipótesis

puede producir la autocorrelación serial de los residuos y por lo tanto que la matriz no sea escalar. En este caso, el empleo de los modelos que explicamos para resolver el problema estructural, podrían resolver también el de  $\sigma^2$ .

Finalmente la hipótesis de normalidad en la distribución de las perturbaciones aleatorias se basa en posibles propiedades asintóticas de los errores, y es la base de todo el proceso inferencial sobre el modelo lineal básico, por lo que el incumplimiento de esta hipótesis podría afectar a los contrastes de significación. No obstante, los estimadores mínimo cuadráticos son ELIO, es decir lineales, insesgados y óptimos, y estas propiedades se cumplen con independencia de la distribución probabilística de la perturbación aleatoria.

#### 4 COMBINACIÓN DE PREDICCIONES CUALITATIVAS

La predicción de la actividad industrial o económica a corto plazo no puede ser correcta si sólo tenemos en cuenta el resultado que arroja una extrapolación de unos datos ya pasados y que además suponemos que serán regulares en el futuro.

Precisamente porque la valoración o las decisiones que toman los agentes económicos (los empresarios industriales, por ejemplo) no siempre sigue una continuidad con datos anteriores, surge la necesidad de considerar sus expectativas.

Cada agente económico  $j$  puede emplear información exclusiva (referente a su empresa por ejemplo), para elaborar su predicción de  $y_{t+h}$  ( ${}^e\hat{y}_{t+h,t}^j$ ) así que una predicción combinada de todas ellas  $\hat{y}_{t+h,t}^c$  podría ser más precisa al aprovechar toda esa información contenida en las predicciones individuales.

Generalmente, las expectativas se obtienen a partir de las encuestas de opinión, con lo que las predicciones son de tipo cualitativo. Sin embargo los agentes pueden dar un valor concreto para  $y_{t+h}$ , o manifestar su grado de creencia respecto a los posibles valores que pueda tomar  $Y_{t+h}$ , y en ese caso las expectativas constituirán predicciones cuantitativas y probabilísticas respectivamente. No obstante, el método a partir del cual cada agente genera su expectativa es cualitativo, ya que éste no especifica explícitamente la forma en que realiza su predicción.

Independientemente del tipo de predicciones y métodos individuales que se consideren, las técnicas de combinación podrán ser cualitativas o cuantitativas, dependiendo del grado de formalización por el cual se procesen las expectativas individuales, y la predicción combinada cualitativa y cuantitativa.

En este capítulo después de tipificar las técnicas de combinación en el caso que nos ocupa, describimos aquellas técnicas cualitativas desarrolladas bajo el enfoque bayesiano. Adoptando como referente esa metodología, aportamos nuevos métodos que resultan adecuados para la combinación de las expectativas de la Encuesta de Coyuntura Industrial.

Finalmente, analizamos métodos empleados para cuantificar las predicciones combinadas cualitativas.

#### 4.1 Técnicas cualitativas de combinación

Si denotamos por  $\hat{Y}_{t+h,t} = (\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^N)$  al vector de expectativas de  $N$  individuos, el objetivo será obtener  $\hat{y}_{t+h,t}^c$  a partir de algún mecanismo que calibre la actuación de cada individuo :  $y_{t+h,t}^c = f(\hat{Y}_{t+h,t}(\alpha))$ . Dependiendo de qué manera se definen las ponderaciones, los métodos de combinación de predicciones pueden ser cualitativos o cuantitativos.

Cuando el proceso a partir del cual se generan las ponderaciones o la predicción combinada no puede reproducirse por otro agente al no especificarse explícitamente la forma en la que éste procesa las expectativas individuales, entonces estamos hablando de técnicas cualitativas de combinación.

##### 4.1.1 Tipificación de las técnicas

Los métodos de combinación de predicciones subjetivas se clasifican en:

1. Métodos de ponderación: las ponderaciones son asignadas a partir de métodos tanto cualitativos (evaluación subjetiva de cada experto) como cuantitativos (medias ponderadas asociadas a la posición del predictor en la organización y asociadas a la precisión relativa de las predicciones pasadas).
2. Métodos de consenso: Se trata de métodos cualitativos de combinación donde las ponderaciones individuales, varían en sucesivas etapas hasta que se llega a una previsión consensuada. Cada agente suministra sus previsiones y a partir de la opinión global del grupo, proceden a reconsiderarlas hasta que se logra un acuerdo aceptable por todos. Entonces, las ponderaciones individuales finales son iguales, puesto que los agentes tienden a homogeneizar sus previsiones con las del resto a lo largo de todo el proceso. Puesto que no se asume la presencia de un decisor fuerte, distintos métodos de consenso pueden derivarse como casos especiales de Bayes.
3. Métodos de calibración: A cada experto se le evalúa su capacidad para asignar probabilidades en función de actuaciones pasadas. La metodología bayesiana se aplica para obtener una función de distribución calibrada en la que el decisor tenga en cuenta alguna medida de evaluación de cada experto. Generalmente, el decisor no revela el sistema a partir del cual obtiene la distribución a posteriori, y por lo tanto se tratan de métodos cualitativos de combinación.

#### 4.1.2 Técnicas de consenso bayesianas

Si las predicciones a combinar son todas subjetivas, y éstas son probabilidades de ocurrencia de un suceso que puede presentar diferentes estados o valores, Stone (1961) propone un método en el que a partir de éstas se obtiene una probabilidad de consenso para cada valor posible del suceso. Si cada individuo  $j$  de un grupo  $N$  da una probabilidad sobre un posible valor ( $s$ ) de  $y_{t+h}^s$   $p_j(y_{t+h}^s)$  la probabilidad del consenso

sería la obtenida mediante ponderaciones:  $p_\lambda(y_{t+h}^s) = \sum_{j=1}^N \alpha_j p_j(y_{t+h}^s)$ . Si no hay un

decisor fuerte, la predicción correspondería con una media aritmética.

De Groot (1974) explica estos resultados: a la hora de llevar a cabo una opinión de consenso, los individuos revisan sus probabilidades asignadas inicialmente a la variable asignando también un peso a las predicciones (probabilidades) de los demás individuos. Mediante reiteradas revisiones el consenso sólo se obtendrá cuando todas las probabilidades revisadas de cada individuo coincidan. Los pesos entonces serán iguales para todos.

Si la toma de decisiones es considerada tarea de un solo individuo que pertenece al grupo, (decisor  $d \in N$ ), Stone (1961) indica que las ponderaciones se obtienen maximizando la utilidad de éste ante la probabilidad combinada,  $\max u[d | p_\alpha(y_{t+h}^s)]$ , y analiza bajo qué condiciones la utilidad de la predicción combinada es mayor que la utilidad de cada opinión por separado:

$$u[d | p_\alpha(y_{t+h}^s)] \geq \min_{j=1 \dots N(d \in N)} u[d_j | p_\alpha(y_{t+h}^s)] \text{ puesto que la probabilidad}$$

combinada final no es conocida por el decisor,  $p_\alpha(y_{t+h}^s)$ , éste ha de examinar diferentes ponderaciones y tomar aquélla que maximice su utilidad.

Si admitimos la posibilidad de la existencia de un individuo decisor ajeno al grupo, entonces sería él únicamente quien valoraría las probabilidades del resto de individuos. Morris (1974) estudia cómo el decisor obtiene una función de distribución a posteriori teniendo en cuenta las funciones de probabilidad de cada individuo (que realiza sus predicciones en base a información económica que puede solaparse o ser exclusiva para cada uno).

Los estudios de Morris fueron completados por Clemen (1987) que estudia el problema de la agregación de la información dependiente entre individuos y cómo afecta a la obtención de la función de probabilidad del decisor, analizando el problema

al que se enfrenta un agente decisor para agregarla y cómo afecta a la probabilidad a posteriori.

Si las funciones de probabilidad de cada individuo son normales la función del decisor seguiría una normal mientras si se parte de distribuciones de Bernoulli se llegaría a una distribución compuesta beta.

Sea  $X_t$  el conjunto de información disponible en el entorno económico, siendo  $x_t^d$  la información de la variable  $d$  ( $d=1,\dots,r$ ) en el instante  $t$ <sup>117</sup>, cada experto puede tomar una parte de ella ( $X_t^j$ ) a partir de la cual genera su distribución de probabilidad  $g_{t+h}^j(y_{t+h} / X_t^j)$  y compartir parte de esta información con el resto de individuos  $X_t^{ij} = X_t^i \cap X_t^j$  dos a dos, y en general un individuo puede emplear información que han empleado los demás  $X_t^j = X_t^1 \cap \dots \cap X_t^N$ .

En un modelo normal donde  $\sigma^2 = 1$  y cada individuo toma como media  $\bar{X}_t^j$  el problema del decisor será  $f(y_{t+h} | \bar{X}_t^1, \dots, \bar{X}_t^N) = f(y_{t+h} | \bar{X})$ .

Esta función de probabilidad se distribuirá también normalmente con  $\mu^+ = \sigma^{*2} l^T \Sigma^{-1} \bar{X}$  y  $\sigma^{*2} = 1/l^T \Sigma^{-1} l$  donde  $l = (1, \dots, 1)^T$  y  $\Sigma$  se obtienen a partir de la cantidad de información que comparten los individuos,  $1/n_i$  y  $n_{ij}/n_i n_j$ .

En un modelo de Bernoulli, cada agente elabora su  $g_{t+h}^j(y_{t+h} / X_t^j)$  como una probabilidad  $p_{t+h}^j = \binom{s_j}{n_j}$ , siendo  $s_j$  un estado de la economía y  $n_j$  el número de casos que cada uno analiza. Entonces el problema del decisor será generar  $f(p_{t+h} | p_{t+h}^1, \dots, p_{t+h}^N)$  en función de  $n_1, \dots, n_N$  y  $S = (s_1, \dots, s_N)$ . La expresión obtenida será

$$f(p_{t+h} | S) = \sum_{q=0}^n \alpha_q^* f_\beta(p | q, n)$$

En cualquier caso, cuando el decisor dispone de la información revelada por cada individuo y ésta se solapa:  $X_t^i \subset X_t^j$  entonces  $f(y_{t+h} | g_{t+h}^i, g_{t+h}^j) = f(y_{t+h} | g_{t+h}^j)$ , con

<sup>117</sup> Podría ocurrir que además los individuos considerasen información referente a la evolución temporal de dichas variables  $X_t^{dT} = (x_1^d, x_2^d, \dots, x_n^d)$  siendo  $x_t^d$  el valor de la serie temporal  $X^d$  en el momento  $t$  ( $t=1, \dots, n$ ).

tres individuos  $f(y_{t+h} | g_{t+h}^i, g_{t+h}^j, g_{t+h}^k) \propto f(g_{t+h}^k | y_{t+h}, g_{t+h}^i, g_{t+h}^j) f(y_{t+h} | g_{t+h}^j) \propto f(g_{t+h}^k | y_{t+h}, g_{t+h}^i, g_{t+h}^j) f(y_{t+h} | g_{t+h}^j) f(y_{t+h} | g_{t+h}^i, g_{t+h}^j, g_{t+h}^k) \propto f(g_{t+h}^k | y_{t+h}, g_{t+h}^j) f(y_{t+h} | g_{t+h}^i, g_{t+h}^j, g_{t+h}^k) \propto f(g_{t+h}^k | y_{t+h}, g_{t+h}^j) f(y_{t+h} | g_{t+h}^i, g_{t+h}^j) \propto f(y_{t+h} | g_{t+h}^i, g_{t+h}^j, g_{t+h}^k)$  con  $f(g_{t+h}^k | y_{t+h}, g_{t+h}^i, g_{t+h}^j) = f(g_{t+h}^k | y_{t+h}, g_{t+h}^j)$ , por lo que se tiene  $f(y_{t+h} | g_{t+h}^i, g_{t+h}^j, g_{t+h}^k) \propto f(g_{t+h}^k | y_{t+h}, g_{t+h}^j) f(y_{t+h} | g_{t+h}^i, g_{t+h}^j) \propto f(y_{t+h} | g_{t+h}^i, g_{t+h}^j, g_{t+h}^k)$

En muchos de los estudios mencionados hasta ahora la función de consenso es tomada como la combinación lineal de las funciones de probabilidad de los expertos [Stone (1961), De Groot (1974)]. En los estudios en los que el consenso se realiza a partir de distribuciones normales con varianzas conocidas se suele llegar a una distribución normal mezcla de las individuales [Morris (1977), Wrinkler (1981)] considera que ese modelo normal de consenso no puede ser tomado como una “mezcla” sino como un modelo individual en el que su media es una combinación lineal de la media de los expertos con los pesos referentes a su precisión y a la dependencia vía varianzas.

#### 4.1.3 La predicción individual según el estadio de predicción.

Como hemos visto, en el consenso y en ausencia de un decisor fuerte, las predicciones individuales tienden a converger según se avanza en los estadios temporales en los que se lleva a cabo éste.

Gallo, Granger y Leon (1999) señalan además que la correlación entre expertos hace que sea muy difícil apreciar las predicciones individuales. Para corroborar este hecho, proponen un modelo que explica cómo cada individuo u organismo genera su predicción cuando hay más expertos y canales de información entre ellos.

Si llamamos  $\hat{y}_{t+h,t-s}^i$  a la predicción realizada en t-s para  $y_{t+h}$ , por un organismo  $i$  ( $i=1, \dots, N$ ), y dicha predicción puede revisarse a medida que se dispone de nueva información, entonces para la predicción de  $y_{t+h}$  en t podremos disponer de distintas predicciones de acuerdo con los distintos estadios de la predicción que denominaremos  $s$  ( $s=0, \dots, k$ ), siendo  $k$ , el primer estadio del consenso y 0 el momento en el que se dispone de la predicción consensuada final.

Así pues, en t dispondremos de distintas predicciones, tanto individuales ( $\hat{y}_{t+h,t-s}^i$ ) como de consenso ( $\hat{y}_{t+h,t-s}^c$ ), de acuerdo con el instante  $s$  en el que nos situemos.

Gallo, Granger y Leon (1999), proponen el siguiente modelo para explicar el efecto del consenso sobre la predicción individual :

$$\hat{y}_{t+h,t-s}^i = \alpha + w_1^i \hat{y}_{t+h,t-s+1}^i + w_2^i \hat{y}_{t+h,t-s+1}^c + w_3^i \sigma_{t+h,t-s+1} + u_{t+h,t-s}^i$$

donde la predicción del organismo  $j$  en un momento  $t-s$  dependerá de la convicción que éste tiene en su predicción previa ( $\hat{y}_{t+h,t-s+1}^i$ ), de la predicción de consenso anterior ( $\hat{y}_{t+h,t-s+1}^c$ ) y la dispersión entre el resto de organismos observada en el estadio anterior<sup>118</sup>  $\sigma_{t+h,t-s+1}$ . El último término de la regresión corresponde al error de predicción  $u_{t+h,t-s}^i$ , siendo de suponer que a medida que se avanza en el proceso de predicción, los organismos tiendan a converger en sus predicciones, de una parte por el efecto a imitar al resto y de otra porque al disponer de más información más precisa será su predicción.

#### 4.2 Nuevas alternativas para las Encuestas de Coyuntura Industrial

En apartados previos resaltamos la utilidad de la Encuesta de Coyuntura Industrial (ECI) para el seguimiento y análisis de las tendencias de la actividad económica al proporcionar una información contextual para seguir los ciclos económicos, determinar puntos de inflexión y realizar pronósticos.

En la ECI se pregunta a 2.700 gerentes empresariales (N) sobre la tendencia que tendrá una variable económica a corto plazo. Puesto que cada individuo tiene información exclusiva referente a su empresa, puede resultar conveniente una combinación de todas sus expectativas  $\hat{Y}_{t+h,t} = \left( {}^e \hat{y}_{t+h,t}^1, {}^e \hat{y}_{t+h,t}^2 \dots {}^e \hat{y}_{t+h,t}^N \right)$ .

Generalmente, la difusión de los resultados de la encuesta se realiza a partir de series cronológicas obtenidas a partir de la agregación ponderada de las respuestas individuales. En este sentido, se puede considerar que se ha efectuado una predicción combinada mediante técnicas cuantitativas.

En este apartado estudiamos cómo se realiza esta combinación y, adoptando como referencia la metodología bayesiana y la información suministrada por esas ponderaciones iniciales, desarrollamos un nuevo método cualitativo que resulta adecuado para la obtención de predicciones combinadas a partir de los datos publicados por la ECI.

---

<sup>118</sup> O'Connor, Remus y Griggs (2000) tratan de separar la reacción derivada de la información pasada de la derivada de la nueva información incorporada a la pasada y se preguntan en qué medida la reacción también puede ser debida a información contextual.

#### 4.2.1 Técnicas cuantitativas de combinación

Como ya hemos descrito anteriormente, en las encuestas de coyuntura cada gerente escoge su predicción dentro de la triada  $\hat{Y}_{t+h,t} = (\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \hat{y}_{t+h,t}^3)$ , cuyos valores corresponden a las tres posibilidades de respuesta para la tendencia  $Y_{t+h} = (y_{t+h}^1, y_{t+h}^2, y_{t+h}^3)$ , disminuir, mantenerse y aumentar respectivamente. Su elección puede considerarse como una predicción cualitativa  ${}^e\hat{y}_{t+h,t}^j$ . En concreto indicaremos con el (r) la opción escogida:  ${}^e\hat{y}_{t+h,t}^j(r)$ , siendo r=1,2,3 de acuerdo con las posibles respuestas.  $\hat{Y}_{t+h,t} = ({}^e\hat{y}_{t+h,t}^1(r), {}^e\hat{y}_{t+h,t}^2(r), \dots, {}^e\hat{y}_{t+h,t}^N(r))$

La ECI agrupa a las empresas industriales por subsectores en función de la actividad productiva (única o principal), utilizando como criterio de agrupación la clasificación de actividades CNAE-93. La respuesta de cada empresa del subsector  ${}^e\hat{y}_{t+h,t}^j(r)$ , se pondera de acuerdo con el porcentaje que representa el número de empleados de la empresa respecto al empleo total del subsector ( $w_j$ )<sup>119</sup>. Se obtiene así una respuesta porcentual:  ${}^e\hat{y}_{t+h,t}^j(r)w_j$ .

Las N respuestas porcentuales se agregan en cada categoría  $\sum_{j=1}^N {}^e\hat{y}_{t+h,t}^j(r)w_j$ , para obtener los tres valores adjudicados en las series cronológicas y que suman 100:

$$\sum_{r=1}^3 \sum_{j=1}^N {}^e\hat{y}_{t+h,t}^j(r)w_j = 100.$$

Cada  $\sum_{j=1}^N {}^e\hat{y}_{t+h,t}^j(r)w_j$  puede considerarse como una combinación de expectativas

${}^e\hat{y}_{t+h,t}^c(r) = \sum_{j=1}^N {}^e\hat{y}_{t+h,t}^j(r)w_j$ , por lo que el vector de predicciones individuales se

transforma en un vector de predicciones combinadas:

$$\hat{Y}_{t+h,t} = ({}^e\hat{y}_{t+h,t}^c(1), {}^e\hat{y}_{t+h,t}^c(2), {}^e\hat{y}_{t+h,t}^c(3))$$

Este proceso es el primer escalón de ponderación y se actualiza cada tres meses con las respuestas de las propias empresas al número de empleados en plantilla.

<sup>119</sup> En otros países se usa el valor de las ventas como factor de ponderación pero es una variable más difícil de conocer que el empleo y más complicada de obtener con la desagregación necesaria.

En una segunda fase, a cada subsector  ${}^e \hat{y}_{t+h,t}^j(r)$  se le adjudica un peso  $\beta_j$  dentro de los sectores agregados en función de su valor añadido bruto<sup>120</sup> respecto al del total de la industria, obtenido de los datos de la *Encuesta Industrial de Empresas* (EIE) del Instituto Nacional de Estadística (INE)

Así,  ${}^e \hat{y}_{t+h,t}^j(r)\beta_j$  se agrega como en el primera fase, obteniendo una predicción combinada para cada  $r$ :  ${}^e \hat{y}_{t+h,t}^c(r) = \sum_{j=1}^N {}^e \hat{y}_{t+h,t}^j(r)\beta_j$ , y cumpliendo:

$$\sum_{r=1}^3 \sum_{j=1}^N {}^e \hat{y}_{t+h,t}^j(r)\beta_j = 100.$$

Este proceso constituye el segundo escalón de ponderación y suele actualizarse cada cinco años para recoger los cambios estructurales que afectan al peso de los diferentes sectores que integran la industria. Con el incremento del nivel de agregación se obtienen los datos de los sectores industriales agregados por destino económico de los bienes (consumo, intermedios e inversión) y total industria.

#### 4.2.2 Técnicas bayesianas de combinación

Bajo el enfoque bayesiano vamos a suponer un decisor que ha de generar su distribución de probabilidad a posteriori de una variable del entorno económico en el futuro  $Y_{t+h}$ <sup>121</sup>. Dicha magnitud (producción, empleo u otra) puede tomar 3 valores  $Y_{t+h} = (y_{t+h}^1, y_{t+h}^2, y_{t+h}^3)$  que identificamos con las tres posibilidades para la tendencia inmediata (*aumentar, mantenerse y disminuir*).

En la encuesta de opiniones empresariales se pregunta a  $N$  empresarios su predicción para cada variable  $Y_{t+h}$  revelando cada uno de ellos un único valor. Por lo tanto, considerando  $N$  predicciones individuales subjetivas  $\hat{Y}_{t+h,t} = ({}^e \hat{y}_{t+h,t}^1, {}^e \hat{y}_{t+h,t}^2, \dots, {}^e \hat{y}_{t+h,t}^N)$  el decisor ha de transformar su distribución a priori de  $Y_{t+h}$  en una distribución a posteriori teniendo en cuenta la realización muestral u opiniones empresariales que dará lugar a la distribución condicional conjunta de verosimilitud  $f_{Y_{t+h}}({}^e \hat{y}_{t+h,t}^1, {}^e \hat{y}_{t+h,t}^2, \dots, {}^e \hat{y}_{t+h,t}^N / y_{t+h})$ .

<sup>120</sup> En ocasiones se emplea una variable equivalente, que se aproximaría al concepto de ingresos de explotación menos variaciones de existencias e impuestos.

<sup>121</sup> Las predicciones se hacen con horizonte trimestral y en concreto para la Encuesta de Opiniones empresariales publicada por SADEI para Asturias, se presentan para tres meses (el trimestre siguiente al que se efectúa la encuesta).

Teniendo en cuenta que  $Y_{t+h}$  es discreta, la distribución a posteriori se calculará:

$$f_{Y_{t+h}}(y_{t+h} / {}^e\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2 \dots {}^e\hat{y}_{t+h,t}^N) = \frac{f_{Y_{t+h}}({}^e\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2 \dots {}^e\hat{y}_{t+h,t}^N / y_{t+h}) p_{Y_{t+h}}(y_{t+h})}{\sum_{Y_{t+h}} f_{Y_{t+h}}({}^e\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2 \dots {}^e\hat{y}_{t+h,t}^N / y_{t+h}) p_{Y_{t+h}}(y_{t+h})}$$

y el decisor  $d$  deberá de revelar el estado en el que encontrará la economía en  $t+h$ , es decir, predecir si la variable  $Y$  aumentará, permanecerá igual o disminuirá en  $t+h$ .

A priori, el decisor establece una opinión en función de experiencias pasadas, información contextual, etc sobre  $Y_{t+h}$ , asignando probabilidades<sup>122</sup> iniciales  $P_{t+h,t} = (p_{t+h}^1, p_{t+h}^2, p_{t+h}^3)$  que reflejan su grado de creencia sobre  $Y_{t+h} = (y_{t+h}^1, y_{t+h}^2, y_{t+h}^3)$  respectivamente. El problema bayesiano implicará entonces, transformar esas probabilidades a priori en otras a posteriori, a partir de la información  $\hat{Y}_{t+h,t} = ({}^e\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2 \dots {}^e\hat{y}_{t+h,t}^N)$ .

Se tendrá así

$$f_{P_{t+h}}(p_{t+h} / {}^e\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2 \dots {}^e\hat{y}_{t+h,t}^N) = \frac{f_{Y_{t+h}}({}^e\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2 \dots {}^e\hat{y}_{t+h,t}^N / p_{t+h}) f_{P_{t+h}}(p_{t+h})}{\int f_{P_{t+h}}({}^e\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2 \dots {}^e\hat{y}_{t+h,t}^N / p_{t+h}) f_{P_{t+h}}(p_{t+h}) d(p_{t+h})}$$

donde  $P_{t+h}$  es continua al estar las probabilidades comprendidas entre 0 y 1.

En función de la forma asumida para  $f_{Y_{t+h}}({}^e\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2 \dots {}^e\hat{y}_{t+h,t}^N / p_{t+h})$  y  $f_{P_{t+h}}(p_{t+h})$  llegaremos a diferentes resultados para la expresión de la distribución a posteriori. Así, podemos decir que  $f_{Y_{t+h}}({}^e\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2 \dots {}^e\hat{y}_{t+h,t}^N / p_{t+h})$  sigue una distribución multinomial o polinomial, ya que la encuesta puede ser considerada como un experimento aleatorio que repetimos  $N$  veces<sup>123</sup> y en cada una de esas repeticiones (que se asumen independientes entre sí) el resultado será una y sólo una de las modalidades de  $Y_{t+h}$  ( $Y_{t+h} = (y_{t+h}^1, y_{t+h}^2, y_{t+h}^3)$ ).

Designando por  $p_{t+h}^r$  ( $r=1, 2, 3$ ) la probabilidad de que en una de las realizaciones independientes ocurra  $y_{t+h}^r$ , y asumiendo que dichas probabilidades se mantienen constantes en las  $N$  pruebas realizadas, el vector aleatorio  $(A_{t+h}^1, A_{t+h}^2, A_{t+h}^3)$  cuyas

<sup>122</sup> Cuando los sucesos son excluyentes, la creencia puede ser tomada como una probabilidad y en concreto como una probabilidad frecuencialista (Shafer (1974)).

<sup>123</sup> En concreto, el vector de expectativas disponible  $\hat{Y}_{t+h,t} = ({}^e\hat{y}_{t+h,t}^1, {}^e\hat{y}_{t+h,t}^2 \dots {}^e\hat{y}_{t+h,t}^N)$  es

$$\hat{Y}_{t+h,t} = ({}^e\hat{y}_{t+h,t}^c(1), {}^e\hat{y}_{t+h,t}^c(2), {}^e\hat{y}_{t+h,t}^c(3)), \text{ donde cada } {}^e\hat{y}_{t+h,t}^c(r) \text{ se obtiene como } \sum_{j=1}^N {}^e\hat{y}_{t+h,t}^c(r) w_j,$$

por lo que el único componente aleatorio es la respuesta de cada empresa, al mantenerse fijas las ponderaciones.

componentes recogen el número de empresas (de las  $N$ ) que han escogido cada modalidad para  $t+h$ , seguirá una distribución multinomial de parámetros  $(N, p_{t+h}^1, p_{t+h}^2, p_{t+h}^3)$ . A pesar de que existen tres resultados, al ser éstos mutuamente excluyentes sólo es necesario definir dos variables aleatorias, dado que para cualquier número específico de cada una la suma de las tres es  $N$  y se cumple  $\sum_{r=1}^3 p_{t+h}^r = 1$ .

Así pues, si dada una realización muestral  $A_{t+h}^1 = a_{t+h}^1$  y  $A_{t+h}^2 = a_{t+h}^2$ , el número de empresas que opinan que la variable  $Y_{t+h} = y_{t+h}^3$  será  $N - a_{t+h}^1 - a_{t+h}^2$ ; por lo tanto

$(A_{t+h}^1, A_{t+h}^2) \approx M(N, p_{t+h}^1, p_{t+h}^2)$  y la distribución de probabilidad será:

$$p(a_{t+h}^1, a_{t+h}^2; N | p_{t+h}^1, p_{t+h}^2) = \frac{N!}{a_{t+h}^1! a_{t+h}^2! (N - a_{t+h}^1 - a_{t+h}^2)!} p_{t+h}^1{}^{a_{t+h}^1} p_{t+h}^2{}^{a_{t+h}^2} (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2}$$

Una de las mayores dificultades que implica la aplicación de la metodología bayesiana es la de necesitar el conocimiento de la distribución a priori del decisor, ya que éste no necesariamente dispondrá de información inicial.

En concreto, en nuestro caso son las empresas las que tienen percepción del estado en el que se encontrará la economía y el decisor toma un papel neutral en la combinación.

#### 4.2.2.1 Distribución Uniforme-Multinomial

Si asumimos que el único conocimiento del que dispone el decisor es:

$$\left. \begin{array}{l} 0 \leq p_{t+h}^1 \leq 1 \\ 0 \leq p_{t+h}^2 \leq 1 \\ p_{t+h}^1 + p_{t+h}^2 \leq 1 \end{array} \right\} \mathbf{R}$$

su distribución a priori  $f_{p_{t+h}}(p_{t+h})$  será uniforme, y la expresión de la función de densidad será:

$$f_{p_{t+h}}(p_{t+h}^1, p_{t+h}^2) = \begin{cases} 2 & (p_{t+h}^1, p_{t+h}^2) \in \mathbf{R} \\ 0 & (p_{t+h}^1, p_{t+h}^2) \notin \mathbf{R} \end{cases}$$

En esta situación, la distribución a posteriori del decisor puede ser obtenida como:

$$\begin{aligned}
 f\left(p_{t+h}^1, p_{t+h}^2 \mid a_{t+h}^1, a_{t+h}^2 : N\right) &= \frac{p\left(a_{t+h}^1, a_{t+h}^2 : N \mid p_{t+h}^1, p_{t+h}^2\right) f_{P_{t+h}}\left(p_{t+h}^1, p_{t+h}^2\right)}{\int_0^1 \int_0^1 p\left(a_{t+h}^1, a_{t+h}^2 : N \mid p_{t+h}^1, p_{t+h}^2\right) f_{P_{t+h}}\left(p_{t+h}^1, p_{t+h}^2\right) d p_{t+h}^1 d p_{t+h}^2} = \\
 &= \frac{\frac{N!}{a_{t+h}^1! a_{t+h}^2! (N - a_{t+h}^1 - a_{t+h}^2)!} p_{t+h}^1 a_{t+h}^1 p_{t+h}^2 a_{t+h}^2 (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2} 2}{\int_0^1 \int_0^1 \frac{N!}{a_{t+h}^1! a_{t+h}^2! (N - a_{t+h}^1 - a_{t+h}^2)!} p_{t+h}^1 a_{t+h}^1 p_{t+h}^2 a_{t+h}^2 (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2} 2 d p_{t+h}^1 d p_{t+h}^2} \\
 &= \frac{p_{t+h}^1 a_{t+h}^1 p_{t+h}^2 a_{t+h}^2 (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2}}{\int_0^1 \int_0^1 p_{t+h}^1 a_{t+h}^1 p_{t+h}^2 a_{t+h}^2 (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2} d p_{t+h}^1 d p_{t+h}^2}
 \end{aligned}$$

y teniendo en cuenta que la integral del denominador es una Beta

$$\int_0^1 \int_0^1 p_{t+h}^1 a_{t+h}^1 p_{t+h}^2 a_{t+h}^2 (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2} d p_{t+h}^1 d p_{t+h}^2 = B\left(a_{t+h}^1 + 1, a_{t+h}^2 + 1, N - a_{t+h}^1 - a_{t+h}^2 + 1\right)$$

se obtiene una distribución a posteriori:

$$f\left(p_{t+h}^1, p_{t+h}^2 \mid a_{t+h}^1, a_{t+h}^2 : N\right) = \frac{\Gamma(N + 3)}{\Gamma\left(a_{t+h}^1 + 1\right) \Gamma\left(a_{t+h}^2 + 1\right) \Gamma\left(N - a_{t+h}^1 - a_{t+h}^2 + 1\right)} p_{t+h}^1 a_{t+h}^1 p_{t+h}^2 a_{t+h}^2 (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2}$$

que es una densidad beta  $B\left(a_{t+h}^1 + 1, a_{t+h}^2 + 1, a_{t+h}^3 + 1\right)$ .

Las estimaciones del decisor se corresponderán en este caso con

$$E\left(p_{t+h}^r \mid a_{t+h}^1, a_{t+h}^2 : N\right) = \frac{a_{t+h}^r + 1}{N + 3}.$$

#### 4.2.2.2 Dirichlet-Multinomial

Teniendo en cuenta que el decisor está interesado en estimar proporciones, la distribución a priori puede ser también una beta  $B\left(P_{t+h} : \alpha, \beta, \gamma\right)$ , ya que este modelo resulta adecuado para representar variables (habitualmente proporciones), cuyos valores se encuentran restringidos a un intervalo de longitud finita ( $p_{t+h}$  deben estar entre 0 y 1).

Más concretamente, puesto que se cumple  $\sum_{r=1}^3 p_{t+h}^r = 1$ , se tendrá en este caso como distribución a priori una Beta restringida, denominada distribución de Dirichlet  $D\left(P_{t+h} : \alpha, \beta, \gamma\right)$ , donde las cantidades de  $\alpha, \beta, \gamma$  son parámetros de perfil que indican la forma de la distribución.

La función de densidad a priori del decisor tendrá en este caso una expresión:

$$f_{P_{t+h}}\left(p_{t+h}^1, p_{t+h}^2 ; \alpha, \beta, \gamma\right) \begin{cases} \frac{\Gamma(\alpha + \beta + \gamma)}{\Gamma(\alpha) + \Gamma(\beta) + \Gamma(\gamma)} p_{t+h}^1 \alpha^{-1} p_{t+h}^2 \beta^{-1} (1 - p_{t+h}^1 - p_{t+h}^2)^{\gamma-1} & \left(p_{t+h}^1, p_{t+h}^2\right) \in \mathbb{R} \\ & \alpha, \beta, \gamma > 0 \\ 0 & \left(p_{t+h}^1, p_{t+h}^2\right) \notin \mathbb{R} \end{cases}$$

donde  $\alpha, \beta$  y  $\gamma$  dependen del criterio del decisor. Si a priori éste no tiene ninguna

preferencia sobre algún estado de la economía la distribución será simétrica y por tanto los tres parámetros serán 1; si el decisor es optimista sesgará la forma de la distribución con  $\alpha > 1$  y  $\beta < 1$  y  $\gamma < 1$ . Los tres parámetros también pueden ser dispuestos por el decisor para contrarrestar los sesgos de los expertos en experiencias pasadas.

La distribución a posteriori del decisor será en este caso:

$$f(p_{t+h}^1, p_{t+h}^2 | a_{t+h}^1, a_{t+h}^2 : N) = \frac{p(a_{t+h}^1, a_{t+h}^2 : N | p_{t+h}^1, p_{t+h}^2) f_{P_{t+h}}(p_{t+h}^1, p_{t+h}^2 : \alpha, \beta, \gamma)}{\int_0^1 \int_0^1 p(a_{t+h}^1, a_{t+h}^2 : N | p_{t+h}^1, p_{t+h}^2) f_{P_{t+h}}(p_{t+h}^1, p_{t+h}^2 : \alpha, \beta, \gamma) dp_{t+h}^1 dp_{t+h}^2} =$$

$$= \frac{\frac{N!}{a_{t+h}^1! a_{t+h}^2! (N - a_{t+h}^1 - a_{t+h}^2)!} p_{t+h}^{a_{t+h}^1} p_{t+h}^{a_{t+h}^2} (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2} \frac{\Gamma(\alpha + \beta + \gamma)}{\Gamma(\alpha) + \Gamma(\beta) + \Gamma(\gamma)} p_{t+h}^{\alpha-1} p_{t+h}^{\beta-1} (1 - p_{t+h}^1 - p_{t+h}^2)^{\gamma-1}}{\int_0^1 \int_0^1 \frac{N!}{a_{t+h}^1! a_{t+h}^2! (N - a_{t+h}^1 - a_{t+h}^2)!} p_{t+h}^{a_{t+h}^1} p_{t+h}^{a_{t+h}^2} (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2} \frac{\Gamma(\alpha + \beta + \gamma)}{\Gamma(\alpha) + \Gamma(\beta) + \Gamma(\gamma)} p_{t+h}^{\alpha-1} p_{t+h}^{\beta-1} (1 - p_{t+h}^1 - p_{t+h}^2)^{\gamma-1} dp_{t+h}^1 dp_{t+h}^2} =$$

$$= \frac{p_{t+h}^{a_{t+h}^1 + \alpha - 1} p_{t+h}^{a_{t+h}^2 + \beta - 1} (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2 + \gamma - 1}}{\int_0^1 \int_0^1 p_{t+h}^{a_{t+h}^1 + \alpha - 1} p_{t+h}^{a_{t+h}^2 + \beta - 1} (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2 + \gamma - 1} dp_{t+h}^1, dp_{t+h}^2}, \text{ donde se}$$

cumple:

$$\int_0^1 \int_0^1 p_{t+h}^{a_{t+h}^1 + \alpha - 1} p_{t+h}^{a_{t+h}^2 + \beta - 1} (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2 + \gamma - 1} dp_{t+h}^1, dp_{t+h}^2 = B(a_{t+h}^1 + \alpha, a_{t+h}^2 + \beta, N - a_{t+h}^1 - a_{t+h}^2 + \gamma)$$

Por lo tanto, se obtiene la expresión:

$$f(p_{t+h}^1, p_{t+h}^2 | a_{t+h}^1, a_{t+h}^2 : N) = \frac{p_{t+h}^{a_{t+h}^1 + \alpha - 1} p_{t+h}^{a_{t+h}^2 + \beta - 1} (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2 + \gamma - 1}}{B(a_{t+h}^1 + \alpha, a_{t+h}^2 + \beta, N - a_{t+h}^1 - a_{t+h}^2 + \gamma)} =$$

$$= \frac{\Gamma(\alpha + \beta + \gamma + N)}{\Gamma(a_{t+h}^1 + \alpha) \Gamma(a_{t+h}^2 + \beta) \Gamma(N - a_{t+h}^1 - a_{t+h}^2 + \gamma)} p_{t+h}^{a_{t+h}^1 + \alpha - 1} p_{t+h}^{a_{t+h}^2 + \beta - 1} (1 - p_{t+h}^1 - p_{t+h}^2)^{N - a_{t+h}^1 - a_{t+h}^2 + \gamma - 1}$$

que es una densidad beta =  $B(a_{t+h}^1 + \alpha, a_{t+h}^2 + \beta, a_{t+h}^3 + \gamma)$

donde las estimaciones del decisor corresponderán con:

$$E(p_{t+h}^1 | a_{t+h}^1, a_{t+h}^2 : N) = \frac{a_{t+h}^1 + \alpha}{N + \alpha + \beta + \gamma}, \quad E(p_{t+h}^2 | a_{t+h}^1, a_{t+h}^2 : N) = \frac{a_{t+h}^2 + \beta}{N + \alpha + \beta + \gamma} \quad y$$

$$\text{por tanto } E(p_{t+h}^3 | a_{t+h}^1, a_{t+h}^2 : N) = \frac{a_{t+h}^3 + \gamma}{N + \alpha + \beta + \gamma}$$

$$E(p_{t+h}^3 | a_{t+h}^1, a_{t+h}^2 : N) = \frac{a_{t+h}^3 + \gamma}{N + \alpha + \beta + \gamma}.$$

### 4.3 Cuantificación de predicciones combinadas cualitativas

La información contenida en las encuestas de opinión permite estimar de forma directa las expectativas de los agentes económicos. Si bien las respuestas en estas encuestas son de tipo cualitativo, se ha intentado transformar las respuestas sobre la dirección del cambio en medidas cuantitativas que pueda emplearse en aplicaciones empíricas junto con otro tipo de información cuantitativa

Uno de los pioneros en la cuantificación fue Anderson (1951, 1952) con la elaboración de un estadístico de síntesis elaborado a partir de la sustracción de las respuestas ponderadas de las categorías extremas y se puede calcular indistintamente para la pregunta referente a la evolución percibida ( ${}^v\hat{y}_{t-h,t}^c(r)$ ) y para la esperada ( ${}^e\hat{y}_{t+h,t}^c(r)$ ):

$$S_t = {}^v\hat{y}_{t-h,t}^c(3) - {}^v\hat{y}_{t-h,t}^c(1) \text{ y}$$

$$S_{t+h} = {}^e\hat{y}_{t+h,t}^c(3) - {}^e\hat{y}_{t+h,t}^c(1) \text{ respectivamente.}$$

A partir de esta idea se han desarrollado numerosos métodos estadísticos entre los que destacan el método de regresión y el método probabilístico, y que desarrollamos en este apartado.

Para la cuantificación por cualquiera de los métodos que empleemos, es necesario disponer de un indicador cuantitativo de referencia (su tasa de variación  $y$ ), a partir del cual es posible dar valores a las encuestas de opinión<sup>124</sup>.

#### 4.3.1 Métodos de regresión

El método de regresión fue desarrollado por Pesaran (1984) con el propósito de justificar la relación entre el saldo  $S_t$  propuesto por Anderson (1952) y las series cuantitativas oficiales ( $y_t$ ).

Si consideramos que  $y_t$  (cambio porcentual) está relacionado con las expectativas que los agentes tuvieron respecto a esa variable en el pasado para el momento  $t$ , entonces el cambio actual de  $y$ , es resultado de los cambios anunciados por las empresas en el pasado  $t-h$  ( ${}^e\hat{y}_{t,t-h}^i$ ):

---

<sup>124</sup> El indicador de referencia es necesario para “anclar” los valores de las encuestas.

$$y_t = \sum_{i=1}^N w_t^i e \hat{y}_{t,t-h}^i, \text{ donde } w_t^i \text{ es la ponderación de la empresa sobre el total de la}$$

magnitud del indicador de referencia.

Si categorizamos las empresas de acuerdo con las respuestas “aumentar” o “disminuir”, entonces:

$$y_t = \sum_{i=1}^p w_t^{i+} e \hat{y}_{t,t-h}^i(3) + \sum_{i=1}^m w_t^{i-} e \hat{y}_{t,t-h}^i(1) \text{ }^{125}$$

Puesto que las respuestas individuales  $e \hat{y}_{t,t-h}^i(r)$  no aparecen en las encuestas, es necesario realizar algún supuesto sobre las variaciones de las respuestas entre firmas y en el tiempo.

Anderson asumió implícitamente que

$$e \hat{y}_{t,t-h}^i(3) = \alpha + v_t^{i+}$$

$$e \hat{y}_{t,t-h}^i(1) = \beta + v_t^{i-}$$

y entonces:

$$y_t = \alpha \sum_{i=1}^p w_t^{i+} + \beta \sum_{i=1}^m w_t^{i-} + v_t$$

donde  $\sum_{i=1}^p w_t^{i+}$  y  $\sum_{i=1}^m w_t^{i-}$  son el porcentaje de firmas (ponderadas) que han respondido “aumentar” y “disminuir”, esto es  $e \hat{y}_{t,t-h}^i(3)$  y  $e \hat{y}_{t,t-h}^i(1)$  respectivamente<sup>126</sup> y

$$v_t = \sum_{i=1}^p w_t^{i+} v_t^{i+} + \sum_{i=1}^m w_t^{i-} v_t^{i-}$$

Así, la ecuación que relaciona el indicador de referencia con las expectativas pasadas puede ser reescrito como:

$$y_t = \alpha e \hat{y}_{t,t-h}^c(3) + \beta e \hat{y}_{t,t-h}^c(1) + v_t.$$

Suponiendo que  $v_t$  no presenta autocorrelación, la mejor predicción lineal de  $y_{t+h}$  vendría dada por:

<sup>125</sup> Los agentes que responden “igual” no se tienen en cuenta.

<sup>126</sup> En 4.2.1. ya vimos que  $\sum_{i=1}^N e \hat{y}_{t+h,t}^i(r) w_i$  era la información proporcionada para cada posible respuesta en la encuesta y que llamamos  $e \hat{y}_{t+h,t}^c(r)$  para indicar que se trataba de una combinación de expectativas. En este caso  $w_{t+h}^{i+}$  representaría la respuesta ya ponderada de cada empresa, esto es  $e \hat{y}_{t+h,t}^i(3) w_i$  (similar para  $w_{t+h}^{i-}$ ).

$$\hat{y}_{t+h}^e = \hat{\alpha}^e \hat{y}_{t+h,t}^c (3) + \hat{\beta}^e \hat{y}_{t+h,t}^c (1)$$

donde  $\hat{\alpha}$  y  $\hat{\beta}$  son las estimaciones MCO de la expresión  $y_t = \hat{\alpha}^e \hat{y}_{t,t-h}^c (3) + \hat{\beta}^e \hat{y}_{t,t-h}^c (1)$ .

A partir de este modelo, es posible contrastar si es necesario distinguir entre las respuestas al alza y a la baja o si los saldos son suficientes ( $\alpha = -\beta$ )<sup>127</sup>.

Theil (1952) propone especificar la siguiente expresión general  $y_{t+h}^e = \hat{\alpha} + \hat{\beta}_1^e \hat{y}_{t+h,t}^c (3) + \hat{\beta}_2^e \hat{y}_{t+h,t}^c (1)$

Los parámetros de esta expresión se estiman o bien a partir de las expectativas pasadas o a partir de la regresión de las valoraciones sobre el comportamiento pasado de la variable<sup>128</sup>:

$$y_t = \alpha + \beta_1^v \hat{y}_{t-h,t}^c (3) + \beta_2^v \hat{y}_{t-h,t}^c (1) + u_t.$$

Pesaran (1984, 1987) y Smith y McAleer (1995) han realizado extensiones adicionales de esta metodología, que permiten contrastar si existen asimetrías en las respuestas al alza y a la baja y tener en cuenta la autocorrelación residual.

Las estimaciones de los parámetros pueden ser obtenidos observando las tasas de variación individuales, pero esta información es difícil de obtener, por lo que hay que estimarlos a partir de la regresión:

$${}^e \hat{y}_{t,t-h}^{i+} (3) = \alpha + \lambda_1 y_t + v_t^{i+}$$

$${}^e \hat{y}_{t,t-h}^{i-} (1) = \beta + \lambda_2 y_t + v_t^{i-}$$

sustituyendo en:

$$y_t = \sum_{i=1}^p w_t^{i+} y_t^{i+} + \sum_{i=1}^m w_t^{i-} y_t^{i-}, \text{ se obtiene:}$$

$$(1 - \lambda_1 {}^e \hat{y}_{t,t-h}^c (3) - \lambda_2 {}^e \hat{y}_{t,t-h}^c (1)) y_t = \alpha {}^e \hat{y}_{t,t-h}^c (3) + \beta {}^e \hat{y}_{t,t-h}^c (1) + v_t$$

$$y_t = \frac{\alpha {}^e \hat{y}_{t,t-h}^c (3) + \beta {}^e \hat{y}_{t,t-h}^c (1)}{1 - \lambda_1 {}^e \hat{y}_{t,t-h}^c (3) - \lambda_2 {}^e \hat{y}_{t,t-h}^c (1)} + u_t^{129}$$

<sup>127</sup> Bajo los supuestos de distribución uniforme, simetría e idéntico impacto de las respuestas al alza y a la baja, es posible cuantificar las expectativas a partir de los saldos.

<sup>128</sup> Pesaran (1985) indica la importancia de contrastar la racionalidad en la formación de expectativas de los agentes, porque sino pueden existir problemas econométricos, justificando así el empleo de valoraciones sobre la evolución pasada.

que permite contrastar la asimetría entre categorías extremas.

La ecuación de cuantificación será:

$$y_{t+h}^e = \frac{\alpha^e \hat{y}_{t+h,t}^c(3) + \beta^e \hat{y}_{t+h,t}^c(1)}{1 - \lambda_1^e \hat{y}_{t+h,t}^c(3) - \lambda_2^e \hat{y}_{t+h,t}^c(1)} + u_{t+h},$$

Si las perturbaciones  $v_t$  fuesen autocorreladas, por ejemplo si siguen un esquema autoregresivo de orden 1 con parámetro  $\rho$ , entonces la expresión correcta será

$$y_{t+h}^e = \frac{\alpha^e \hat{y}_{t+h,t}^c(3) + \beta^e \hat{y}_{t+h,t}^c(1) + \rho \left( (1 - \lambda_1^e \hat{y}_{t-h,t}^c(3) - \lambda_2^e \hat{y}_{t-h,t}^c(1)) y_t - \alpha \hat{y}_{t-h,t}^c(3) - \beta \hat{y}_{t-h,t}^c(1) + v_t \right)}{1 - \lambda_1^e \hat{y}_{t+h,t}^c(3) - \lambda_2^e \hat{y}_{t+h,t}^c(1)}$$

Pesaran (1987) estima la ecuación con AR(2):

$$y_{t+h}^e = \frac{\alpha^e \hat{y}_{t+h,t}^c(3) + \beta^e \hat{y}_{t+h,t}^c(1) + \phi_1 \hat{v}_t + \phi_2 \hat{v}_{t-1}}{1 - \lambda_1^e \hat{y}_{t+h,t}^c(3) - \lambda_2^e \hat{y}_{t+h,t}^c(1)}$$

Si consideramos que los parámetros de la regresión varían a lo largo del tiempo, y que por ejemplo hacemos que dependan de los valores pasados de la serie de referencia (Smith-Mcleer (1995))

$$\alpha_t = 1 + \sum_{j=0}^m \delta_j y_{t-j}$$

$$\beta_t = 1 + \sum_{j=0}^m \gamma_j y_{t-j}$$

se llega a un modelo de regresión dinámico no lineal:

$$y_{t+h}^e = \frac{\alpha^e \hat{y}_{t+h,t}^c(3) + \beta^e \hat{y}_{t+h,t}^c(1) + \hat{y}_{t+h,t}^c(3) \sum_{j=0}^m \lambda_{1j} y_{t-j} + \hat{y}_{t+h,t}^c(1) \sum_{j=0}^m \lambda_{2j} y_{t-j}}{1 - \lambda_{10}^e \hat{y}_{t+h,t}^c(3) - \lambda_{20}^e \hat{y}_{t+h,t}^c(1)}$$

### 4.3.2 Métodos probabilísticos

El método probabilístico fue desarrollado por Theil (1952) para dar justificación teórica al saldo de Anderson (1952) y posteriormente fue ampliado y aplicado por Carlson y Parkin (1975), Pesaran (1987), etc. Estos métodos se basan en las siguientes hipótesis:

---

<sup>129</sup> Con  $u_t = \frac{v_t}{1 - \lambda_1^e \hat{y}_{t,t-h}^c(3) - \lambda_2^e \hat{y}_{t,t-h}^c(1)}$ .

$u_t$  (y  $v_t$ ) tendrá 0 de media pero no tendrá que ser necesariamente homocedástica e incorrelada.

1. Cada firma basa su respuesta en una distribución de probabilidad subjetiva  $g^i(y_{t+h}^i / I_t^i)$  definida sobre el cambio futuro de la variable en su empresa condicionada al conjunto de información de que dispone la empresa en el momento  $t$ .
2. Existe un intervalo de indiferencia  $(-a_{t+h}^i, b_{t+h}^i)$  en el que los encuestados contestarán que esperan un aumento de la variable en  $t+h$ ,  ${}^e\hat{y}_{t+h,t}^i(3)$ , si  ${}^e\hat{y}_{t+h,t}^i \geq b_{t+h}^i$ , y que esperan una disminución,  ${}^e\hat{y}_{t+h,t}^i(1)$ , si  ${}^e\hat{y}_{t+h,t}^i \leq -a_{t+h}^i$ .

Es decir,  $(-a_{t+h}^i, b_{t+h}^i)$  determinan el rango de tasas de variación de la variable alrededor de cero que los encuestados no distinguen de cero.

3. Los intervalos de indiferencia son iguales para todas las firmas  $(-a_{t+h}^i, b_{t+h}^i) = (-a_{t+h}, b_{t+h})$ , pudiendo ser constantes a lo largo del tiempo  $(-a, b)$  y/o simétricos  $a=b=c$ .
4. Las distribuciones de probabilidad subjetivas son independientes y tienen la misma forma conocida para todas las empresas.
5. La probabilidades subjetivas, pueden ser empleadas para derivar una distribución de probabilidad agregada  $g(y_{t+h} / I_t)$  con momentos finitos de primer y segundo orden, donde  $I_t$  es la unión de los conjuntos de información de las empresas individuales  $I_t = \bigcup_{i=1}^N I_t^i$  y  $y_{t+h}$  el porcentaje de cambio real en la variable media de la industria  $(y_{t+h} = \sum_{i=1}^N w_i y_{t+h}^i)$ .

Bajo estas hipótesis la formulación general de los métodos probabilísticos es:

$$\begin{cases} \text{prob}\{y_{t+h} \leq -a_{t+h} / I_t\} = G_t(-a_{t+h}) = {}^e\hat{y}_{t+h,t}^c(1) \\ \text{prob}\{y_{t+h} \geq b_{t+h} / I_t\} = 1 - G_t(b_{t+h}) = {}^e\hat{y}_{t+h,t}^c(3) \end{cases}$$

donde  $G(\cdot)$  denota la función de densidad acumulativa de  $g(y_{t+h} / I_t)$

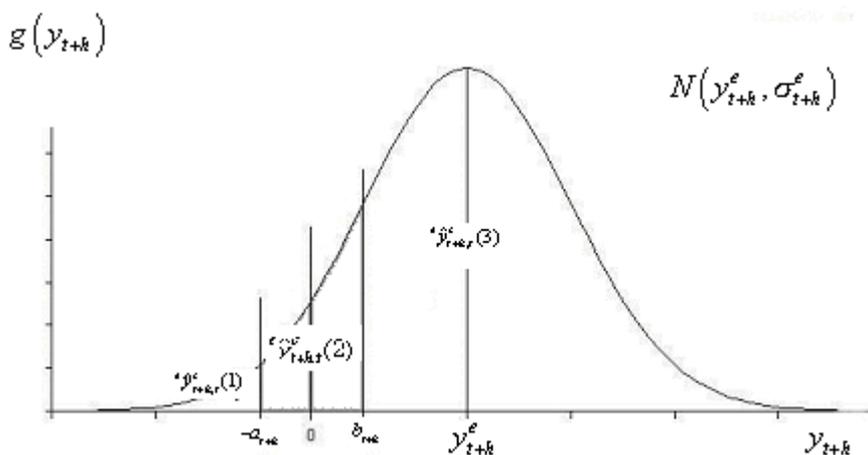
La expectativa agregada puede obtenerse como  $y_{t+h}^e = E(y_{t+h} / I_t)$ , suponiendo que el intervalo de indiferencia es conocido o que puede estimarse independientemente del sistema anterior, y que excepto para su media y/o varianza, la forma de la función de densidad agregada  $g(y_{t+h} / I_t)$  es completamente conocida.

La cuantificación de las expectativas a partir de esa formalización general dependerá de la forma que adopte  $g(y_{t+h}/I_t)$ , y del intervalo de indiferencia en cuanto a simetría y variación en el tiempo.

### 4.3.2.1 Distribución Normal

La distribución  $g(y_{t+h}/I_t)$  habitualmente considerada en la literatura es la normal<sup>130</sup>. En la figura 8 se representa el marco teórico desarrollado por Theil (1952).

Figura 8. Distribución agregada normal



Bajo esta premisa, si se supone que la distribución de frecuencias de los encuestados se comporta como una normal, el sistema general (en términos de normal estandarizada) será:

$$\begin{cases} \text{prob} \left\{ \frac{y_{t+h} - y_{t+h}^e}{\sigma_{t+h}^e} \leq \frac{-a_{t+h} - y_{t+h}^e}{\sigma_{t+h}^e} / I_t \right\} = {}^e \hat{y}_{t+h,t}^c (1) \\ \text{prob} \left\{ \frac{y_{t+h} - y_{t+h}^e}{\sigma_{t+h}^e} \leq \frac{b_{t+h} - y_{t+h}^e}{\sigma_{t+h}^e} / I_t \right\} = 1 - {}^e \hat{y}_{t+h,t}^c (3) \end{cases}$$

<sup>130</sup> El Teorema Central del Límite justifica el empleo de la distribución normal desde un punto de vista teórico. Muchos autores no han hallado diferencias significativas entre las series de expectativas estimadas mediante las distribuciones uniforme, logística,  $t$  de Student con respecto a las obtenidas a partir de la normal en las series de expectativas estimadas, y recalando la importancia de la simetría en la distribución. Berk (1999), por ejemplo, realiza un estudio empírico para encuestas expectativas sobre precios en Alemania, para ver la los efectos de relajar el supuesto de normalidad de las expectativas (curtosis y asimetría) comprobando que las predicciones más precisas son las de basadas en distribuciones simétricas y que existen pocas diferencias entre los resultados obtenidos con la t-student y la normal. Además, la evidencia empririca indica que la distribución de las expectativas es simétrica, pero no estrictamente normal (Batchelor y Dua (1987)). Dasgupta y Lahiri (1992) llegan a conclusiones similares a sobre la simetría de las distribución

Si llamamos

$$r_{t+h} = \frac{-a_{t+h} - y_{t+h}^e}{\sigma_{t+h}^e}$$

$$s_{t+h} = \frac{b_{t+h} - y_{t+h}^e}{\sigma_{t+h}^e}, \text{ entonces mediante este sistema obtenemos la ecuación de}$$

cuantificación<sup>131</sup>:

$$y_{t+h}^e = \frac{r_{t+h}b_{t+h} + s_{t+h}a_{t+h}}{r_{t+h} - s_{t+h}}$$

Si definimos la variable:

$$d_{t+h} = \frac{r_{t+h} + s_{t+h}}{r_{t+h} - s_{t+h}}$$

entonces podemos obtener una expresión equivalente:  $y_{t+h}^e = \alpha_{t+h} + \beta_{t+h}d_{t+h}$  con

$$\alpha_{t+h} = \frac{b_{t+h} - a_{t+h}}{2} \text{ y } \beta_{t+h} = \frac{b_{t+h} + a_{t+h}}{2}$$

Puesto que no se conoce  $y_{t+h}$  en el momento  $t$ , para cuantificar las expectativas, es necesario estimar la expectativa agregada a partir de las realizaciones del indicador de referencia y de las expectativas pasadas. Así los parámetros estimados en

$$y_t = \frac{r_t b_t + s_t a_t}{r_t - s_t} + u_t;$$

serán empleados para calcularla la expectativa agregada estimada<sup>132</sup>:

$$\hat{y}_{t+h}^e = \frac{r_{t+h} \hat{b}_t + s_{t+h} \hat{a}_t}{r_{t+h} - s_{t+h}}$$

Además, Theil (1952) supone que el intervalo de indiferencia es simétrico y permanece fijo tanto entre las empresas como a lo largo del tiempo  $a_{t+h} = b_{t+h} = c$   $\forall t = 1, \dots, T$ .

Carlson y Parkin (1975) realizan el mismo supuesto sobre el intervalo añadiendo además, que en ausencia de evidencia contraria,  $c$  es proporcional como media al nivel

<sup>131</sup> La varianza será:  $\sigma_{t+h}^e = \frac{a_{t+h} + b_{t+h}}{s_{t+h} - r_{t+h}}$

<sup>132</sup> La sustitución de los parámetros estimados implica asumir que la distribución agregada y el intervalo de indiferencia es igual para expectativas que para las realizaciones, Pesaran (1987).

de la variable durante un período de tiempo específico<sup>133</sup>. Así introducen el supuesto de insesgadez y estiman el parámetro de indiferencia como el valor necesario para igualar la media de la tasa de inflación esperada a la media de la tasa de inflación real:

$$y_{t+h}^e = \hat{c}d_{t+h} \quad \text{donde} \quad \hat{c} = \frac{\sum_{t=1}^n y_t}{\sum_{t=1}^n d_t}$$

Danes (1975) con el objetivo de relajar el supuesto de insesgadez de las expectativas, propone sustituir  $d_{t+h}$  por  $d_{t-h}$ , es decir utilizar las valoraciones sobre la evolución pasada de la variable ( ${}^v\hat{y}_{t-h,t}^c(r)$ ) en lugar de las expectativas sobre la tendencia futura:

$$y_{t+h} = \hat{c}d_{t+h} \quad \text{donde} \quad \hat{c} = \frac{\sum_{t=1}^n y_t}{\sum_{t=1}^n d_{t-h}}$$

Con el mismo objetivo Bennett (1984) propone obtener  $c$  a partir de la estimación MCO de la regresión  $y_{t-h} = \gamma + cd_{t-h} + u_{t-h}$ . Al suponerse ambas series independientes, la estimación mínimo-cuadrática del intervalo de indiferencia,  $\hat{c}$ , debe interpretarse como el coeficiente de correlación entre ambas series. La expresión resultante a partir de este método es:

$$y_{t+h}^e = \hat{c}d_{t+h} \quad \text{donde} \quad \hat{c} = \frac{\sum_{t=1}^n (y_{t-h} - \bar{y}_{t-h})(d_{t-h} - \bar{d}_{t-h})}{\sum_{t=1}^n (y_{t-h} - \bar{y}_{t-h})^2}$$

#### 4.3.2.2 Distribución Uniforme

En el caso de que  $g(y_{t+h} / I_t)$  sea uniforme  $(-q, q)$ :

$$g(y_{t+h} / I_t) = \frac{1}{2q} \quad \text{para} \quad y_{t+h}^e - q \leq y_{t+h} \leq y_{t+h}^e + q$$

$$g(y_{t+h} / I_t) = 0 \quad \text{en otro caso.}$$

<sup>133</sup> Otros autores asumen arbitrariamente otros valores para el intervalo de indiferencia. Knöbl (1974), por ejemplo, asume el valor de  $c$  arbitrariamente igual a dos y supone que el intervalo de indiferencia es independiente de la variable de referencia

y generalizando para un intervalo asimétrico de indiferencia  $(-a_{t+h}, b_{t+h})$ , las respuestas de las encuestas pueden escribirse como:

$${}^e\hat{y}_{t+h,t}^c(1) = \frac{-y_{t+h}^e + q - a_{t+h}}{2q}$$

$${}^e\hat{y}_{t+h,t}^c(3) = \frac{y_{t+h}^e + q - b_{t+h}}{2q}$$

y

$$y_{t+h}^e = \frac{2q(q - a_{t+h})}{2q - a_{t+h} - b_{t+h}} {}^e\hat{y}_{t+h,t}^c(3) + \frac{2q(q - b_{t+h})}{2q - a_{t+h} - b_{t+h}} {}^e\hat{y}_{t+h,t}^c(1)$$

$$\sigma_{t+h}^e = \frac{a_{t+h} + b_{t+h}}{2\sqrt{3}S_{t+h}}$$

La expectativa estimada agregada se realiza a partir de la información pasada, si bien es necesario determinar primero  $q$ . Carlson y Parkin (1975) asumen que las

expectativas son insesgadas, con lo que estiman  $\hat{q} = \frac{\sum_{i=1}^n y_t}{\sum_{i=1}^n S_t}$

#### 4.3.2.3 Distribución Logística

Si  $g(y_{t+h}/I_t)$  es logística, las expresiones que se emplearían serían:

$$r_{t+h} = \log \left[ \frac{1}{{}^e\hat{y}_{t+h,t}^c(1)} - 1 \right]$$

$$s_{t+h} = -\log \left[ \frac{1}{{}^e\hat{y}_{t+h,t}^c(3)} - 1 \right]$$

La expectativa agregada se obtendría como:

$$y_{t+h}^e = \frac{r_{t+h}b_{t+h} + s_{t+h}a_{t+h}}{r_{t+h} - s_{t+h}}$$

#### 4.3.2.4 Intervalo de indiferencia variable en el tiempo

Hasta ahora hemos supuesto que los intervalos de indiferencia son constantes, si bien los hemos indexado en todo momento con  $t+h$  para evitar perder generalidad. Si se prescinde de esta hipótesis, y se supone que estos evolucionan con el tiempo, surgen nuevos métodos de cuantificación, algunos de los cuales estudiamos a continuación.

Uno de los métodos [Pesaran (1984) y Batchelor (1986)], hace que los parámetros del intervalo dependan de la tasa de la variable de referencia y de alguna medida de incertidumbre sobre esa variable ( $\sigma_t$ ):

$$a_{t+h} = a^0 + a^1 y_{t+h} + a^3 \sigma_{t+h}$$

$$b_{t+h} = b^0 + b^1 y_{t+h} + b^3 \sigma_{t+h}$$

Los parámetros estimados con la última información disponible serán los empleados en  $t+h$ .

Clavería y Pons (2001) indican que, puesto que las encuestas son mensuales, los individuos encuestados no disponen de la tasa de variación de la variable del período en curso por lo que sólo podrán incorporar a su conjunto de información la tasa de variación del período anterior, la cual sí que conocen. Estos autores desarrollan un método de estimación en dos etapas que parte del marco teórico desarrollado por Theil con intervalo de indiferencia simétrico (1952).

Así, introducen el supuesto de que los individuos únicamente considerarán que la variable analizada ha variado cuando la tasa de variación esperada para el siguiente período sea mayor en valor absoluto que la tasa de variación del indicador cuantitativo de referencia observada en el período anterior.

De esta forma, el parámetro  $c$  se corresponderá con la tasa de variación del indicador cuantitativo de referencia en el período  $t-1$ . Así, esta metodología de estimación del parámetro de indiferencia da como resultado una primera serie de expectativas de tasas de variación de la variable estudiada que denominan como *input*:

$$input y_{t+h}^e = \hat{c}^* d_{t+h}$$

donde  $\hat{c}^* = |y_{t-1}|$ , y  $y_{t-1}$  recoge la tasa de la tasa de crecimiento del indicador cuantitativo de referencia del período anterior.

Los autores justifican la necesidad de un escalamiento en dos etapas porque el supuesto realizado respecto al parámetro de indiferencia no es suficiente para acabar de escalar la serie de expectativas resultante. Con este propósito, una vez derivada la serie de expectativas input con parámetros cambiantes en el tiempo, se especifica la siguiente ecuación de regresión:

$$y_{t+h} = \alpha + \beta^{input} y_{t+h}^e + u_{t+h}$$

en la que se introduce la serie input como variable exógena. A partir de la estimación MCO de los parámetros de esta regresión se construye la siguiente ecuación de cuantificación:

$$y_{t+h}^e = \hat{\alpha} + \hat{\beta}^{input} y_{t+h}^e \quad \text{donde } {}^{input}y_{t+h}^e = \hat{c}^* d_{t+h} \text{ y } \hat{c}^* = |y_{t-1}|$$

donde  $y_{t+h}^e$  es la serie de expectativas estimadas de la tasa de variación de la variable analizada que se puede utilizar como *proxy* de la tasa de variación de la variable en el futuro.

Otros métodos alternativos se basan en la aplicación de las técnicas de estimación de parámetros cambiantes en el tiempo:

El modelo más general fue propuesto por Cooley y Prescott (1973, 1976) que especifica la siguiente estructura para los parámetros de  $y_{t+h}^e = \alpha_{t+h} + \beta_{t+h} d_{t+h}$  :

$$\alpha_{t+h} = \alpha_{t+h}^p + u_{t+h}^\alpha$$

$$\alpha_{t+h}^p = \alpha_t^p + v_{t+h}^\alpha$$

y

$$\beta_{t+h} = \beta_{t+h}^p + u_{t+h}^\beta$$

$$\beta_{t+h}^p = \beta_t^p + v_{t+h}^\beta$$

donde  $\alpha_{t+h}^p$  y  $\beta_{t+h}^p$  representan los componentes permanentes de los parámetros y  $u$  y  $v$ , los errores aleatorios.

Sustituyendo en la ecuación de  $\alpha_{t+h}$  la expresión de  $\alpha_{t+h}^p$ , los sistemas anteriores pueden describirse como:

$$\alpha_{t+h} = \alpha_0^p + u_{t+h}^\alpha + \sum_{s=1}^{t+h} v_s^\alpha$$

$$\beta_{t+h} = \beta_0^p + u_{t+h}^\beta + \sum_{s=1}^{t+h} v_s^\beta$$

Estimando las expresiones  $y_t = \alpha \hat{y}_t^c(3) + \beta \hat{y}_t^c(1) + v_t$ ,  $y_t = \alpha_t + \beta_t d_t + u_t$ , y empleando el filtro de Kalman y suponiendo que  $\alpha_{t+h} = \alpha_t$  y  $\beta_{t+h} = \beta_t$ , pueden obtenerse la serie de expectativas cuantificadas.

De forma similar, Seitz (1988), propone el siguiente sistema::

$$y_{t+h} = \alpha_{t+h} + \beta_{t+h} d_{t+h}$$

$$\alpha_{t+h} = \alpha_t + \xi_{t+h}^\alpha$$

$$\beta_{t+h} = \beta_t + \xi_{t+h}^\beta$$

Sobre esta metodología Clavería, Pons y Suriñach (2002), proponen un esquema más general, que parte del método probabilístico con límites de indiferencia asimétricos y cambiantes en el tiempo propuesto por Seitz (1988) y suponen que:

$$\begin{cases} a_t = \alpha a_{t-1} + w_t \\ b_t = \beta b_{t-1} + w_t \end{cases}$$

donde  $v_t$  y  $w_t$  son dos términos de perturbación independientes e idénticamente distribuidos como una normal con media cero<sup>134</sup>. Se especifica la siguiente ecuación de medida:  $y_t = \alpha_t + \beta_t d_t + u_t$  donde  $u_t \sim N(0, \sigma_u^2)$

La representación en el espacio de estados del modelo

$$y_t = \alpha_t + \beta_t d_t + u_t$$

$$\begin{cases} \alpha_t = \gamma \alpha_{t-1} + w_t \\ \beta_t = \lambda \beta_{t-1} + v_t \end{cases}$$

permite obtener las estimaciones de  $\sigma_u$ ,  $\sigma_v$ ,  $\sigma_w$ ,  $\gamma$ ,  $\lambda$  y los vectores de estado estimados, equivalentes a los límites de indiferencia para cada momento en el tiempo, mediante la aplicación del filtro de Kalman.

---

<sup>134</sup> Con el objetivo de estimar las varianzas de ambos términos de perturbación y los parámetros del modelo utilizan las expectativas sobre la evolución futura de la variable analizada en lugar de las percepciones sobre la evolución pasada que emplean otros autores.

## **5 EVALUACIÓN DE LA PREDICCIÓN COMBINADA**

El objetivo de la combinación de predicciones es aprovechar la información contenida en las predicciones individuales, que produciría una alta ganancia en precisión.

Cuando se tienen diferentes métodos de combinación y se obtienen predicciones combinadas, es necesario contrastar si todos los modelos de combinación son igualmente adecuados y verificar si realmente existe una ganancia en precisión con respecto a cada una de las predicciones individuales empleadas.

Estudios empíricos, no han dado conclusiones sobre la superioridad de un método frente a otros, si bien la precisión en la predicción ha sido tomada generalmente como criterio para la comparación entre métodos más que la estabilidad de ponderaciones y la distribución del error de predicción. Analizar la estabilidad de éstas puede ayudar a contrastar si la estructura de las varianza-covarianza de los errores de predicción combinada es estable y conocer la calidad del error en términos de simetría y normalidad ayudarnos a decidir entre distintos métodos de combinación.

Sobre la base de estas consideraciones, en este capítulo, después de repasar algunos trabajos empíricos, describimos herramientas desarrolladas que permiten contrastar la ganancia en información (contraste de abarcamiento) y ,precisión (test de Stekler y test de reducción significativa del error cuadrático medio), la estabilidad de los pesos y la forma de la distribución del error de la predicción combinada.

Finalmente, estudiamos el grado de complejidad de cada método individual y combinado que puede ayudar a decidir entre técnicas, cuando otros criterios no permiten discriminar entre ellas.

### **5.1 Experiencias**

A la hora de escoger un método de combinación de predicciones nos deberíamos preguntar si hay alguno más adecuado que otro. Lo cierto es que tanto los trabajos teóricos que hemos visto y trabajos empíricos no concluyen con unanimidad la superioridad de un métodos frente a otros.

Algunos estudios empíricos concluyen que aquel basado en una media aritmética da mejores resultados que otras reglas más complicadas (Granger y Newbold (1975),

Makridakis y Hibon (1979), Moreno y López (1999), Moreno, López y Landajo (2000) etc)<sup>135</sup>.

Granger y Newbold (1975) indican además la combinación es más beneficiosa cuanto más difiera la naturaleza de los procedimientos que se combinan, idea confirmada en estudios posteriores.

Winkler y Makridakis (1983) y Winkler (1984)<sup>136</sup> obtienen resultados muy buenos con medias ponderadas y modelos de ajuste, si bien resaltan que el incremento de la precisión de la predicción combinada en relación a la complicación del método es muy pequeña.

Tal y como menciona Otero (1993), sigue siendo una incógnita el saber por qué la media aritmética simple funciona mejor que otros métodos de ponderación más complicados. Algunas respuestas se basan en el incumplimiento de las hipótesis básicas de modelización cuando éstas se aplican a las series de predicciones individuales, lo que explica que métodos más simples resultan más robustos y proporcionan mejores resultados. Por otra parte la media aritmética supone restringir las ponderaciones a una suma unitaria que como ya hemos visto ofrece errores cuadráticos de predicción mayores, con lo que los resultados deberían de ser generalizadamente peores.

Algunos autores han especulado sobre la inestabilidad de combinaciones ponderadas (Kang 1986) y sobre la no estacionariedad del proceso que se predice (Diebold y Pauly (1986, 1987).

Además dar pesos iguales a métodos menos precisos implica disminuir la precisión de la combinación como han puesto de manifiesto Asthon y Asthon (1985)<sup>137</sup> y Clemen y Wrinkler (1986), Ahlesrs y Lakonishok (1983)<sup>138</sup>, etc.

---

<sup>135</sup> Si bien Makridakis y Hibon (1979) no examinan modelos combinados, su trabajo es importante porque es el inicio de la "M-competition" realizada por Makridakis et al. (1982, 1983) en la que se aplicaron varios métodos de predicción a 1001 series temporales económicas. En el estudio se efectuaron predicciones ex ante y su predicción se analizó a partir de un conjunto de medidas del error de predicción. Aunque el propósito del trabajo era comparar la precisión de las técnicas empleadas, también se aplicó dos modelos de combinación obteniendo resultados relativamente mejores que los que se obtenían con las técnicas empleadas individualmente, siendo la combinación aritmética la más precisa. Moreno y López (1999), Moreno, López y Landajo (2000) efectúan un trabajo empírico sobre series de la economía asturiana.

<sup>136</sup> Winkler (1984) en su estudio comprobó que el número de veces que una combinación fue mejor que la de la media aritmética estaba entre el 52% y el 66%.

<sup>137</sup> Asthon y Asthon (1985) analizan predicciones de ventas en publicidad efectuadas por ejecutivos empleando como método de combinación una media ponderada asociada con la posición del ejecutivo en la organización y otra basada en pesos basados en el criterio de la CEO sobre la relativa precisión de las predicciones.

<sup>138</sup> Ahlesrs y Lakonishok (1983) muestran que la predicción obtenida por el consenso de expertos (con igual ponderación) no es eficiente, es decir no incorpora toda la información disponible y no supera a los

Una alternativa a la media aritmética es calcular los pesos minimizando el error de la predicción combinada (Bates y Granger (1969)) que llega a los mismos resultados que en el enfoque bayesiano realizado por Morris (1977)<sup>139</sup>.

No obstante si la correlación entre modelos es grande y positiva, los pesos asignado a uno o varios modelos puede ser negativa (la suma de todas las ponderaciones suma la unidad) y entonces es posible que la predicción combinada falle fuera del rango de las predicciones individuales.

Granger y Ramanathan (1984) señalan que si no se somete a ninguna restricción al modelo de regresión éste puede producir predicciones combinadas más precisas, aumentando más si incluso se deja la posibilidad de incluir un término independiente en la regresión, no será preciso que las predicciones de cada método individual sean insesgadas. No obstante los problemas de negatividad y multicolinealidad pueden persistir.

Newbold y Granger (1974) demuestran que en la práctica se pueden ignorar las correlaciones en la estimación de la combinación ponderada, es decir, la estimación ponderada que depende de las correlaciones estimadas actúa peor que aquella que la ignora. Winkler y Makridakis (1983) estudiaron el problema de la obtención de pesos “óptimos”.

No obstante, en lo que si parece que hay unanimidad en la literatura es en la idoneidad de realizar predicciones combinadas<sup>140</sup>, y en las propiedades que deberían de tener los métodos de combinación:

---

modelos simples. No obstante el estudio realizado para diez series económicas muestran que los resultados en la eficiencia pueden variar según la serie.

<sup>139</sup> Morris (1974) propone un método bayesiano para la agregación de información basándose en la idea de que el decisor bayesiano debe considerar su opinión de experto como información en la actualización bayesiana. El problema radica en establecer la función de verosimilitud. Morris (1977) extiende su trabajo de 1974 y presenta resultados con distribuciones normales. Su análisis demuestra que la agregación de opinión experta es similar a un problema de calibración experta.

<sup>140</sup> Bunn y Mustafaoglu (1978) realizan un trabajo empírico en el que agrega utilizando la metodología bayesiana, juicios realizados por un grupo de expertos. Este juicio combinado es utilizado con posterioridad para generar probabilidades acerca de riesgos de algunos eventos. Vélez y Melo (1998) concluyen que para diferentes horizontes de predicción y en todos los casos que analizan el procedimiento de la mezcla de predicciones genera predicciones más precisas que los predicciones individuales de los modelos considerados. Conroy y Harris (1987) Estudian el Intitutional Brokers Estimate System (IBES) sugiriendo que sus predicciones se pueden mejorar empleando análisis combinados y predicciones realizadas con series temporales. Lawrence, Edmunson y O'Connor (1986) basándose en el experimento de Makridakis con 111 series, se combinan predicciones estadísticas y subjetivas mostrando un incremento en la precisión. La combinación mecánica es mejor que la combinación subjetiva. Estos autores consideran dos tipos de predicciones subjetivas:

Una consiste en ofrecer a un grupo de expertos un gráfico con los valores pasados de la serie a predecir: Consideran estas predicciones con las obtenidas por un modelo exponencial simple sobre las series

1. El procedimiento para la obtención de pesos no debe depender de un gran número de información. Si no hay mucha información disponible los juicios subjetivos pueden combinarse con la información disponible.
2. El método de combinación debe de distinguir la actuación relativa de los métodos individuales. Además ha de proporcionar resultados robustos.
3. Los pesos han de tener significado intuitivo.

## 5.2 Contrastes sobre la ganancia de información y de precisión

### 5.2.1 Contraste de abarcamiento

Los contrastes de abarcamiento de predicciones permiten determinar una predicción  $i$  incorpora la misma información que otras  $j \neq i = 1, \dots, N$ .

En el caso de dos predicciones el contraste sería estimar:

$$y_{t+h} = \alpha_0 + \alpha_1 \hat{y}_{t+h,t}^i + \alpha_2 \hat{y}_{t+h,t}^j + u_{t+h}$$

Si  $(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2) = (0, 1, 0)$  se dice que el modelo de predicción  $i$  incorpora la información del modelo  $j$ , y si  $(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2) = (0, 0, 1)$ , entonces el modelo  $j$  incorpora al modelo  $i$ . Para cualquier otro valor de  $(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2)$  se dice que ningún modelo incorpora al otro, y por tanto los dos predicciones contienen información útil sobre  $y_{t+h}$  y sería conveniente una combinación.

Cuando las series son integradas de orden uno, Fair y Shiller (1989, 1990) se basan en la regresión

$$y_{t+h} - y_t = \alpha_0 + \alpha_1 (\hat{y}_{t+h,t}^i - y_t) + \alpha_2 (\hat{y}_{t+h,t}^j - y_t) + \varepsilon_{t+h}$$

$(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2) = (0, 1, 0)$  o  $(0, 0, 1)$  indican la existencia de abarcamiento.

---

desestacionalizadas y proponen métodos de combinación de media aritmética y ponderadas por el error de predicciones pasadas.

y otra un gráfico, de tal manera que éstos han de dar predicciones para la variable en un horizonte  $h$ . Sus resultados muestran que la inclusión de predicciones subjetivas mejora la precisión de las predicciones y que el método de gráficos es mejor para predicciones a corto plazo mientras que el de la tabla lo es para predicciones a largo plazo. Tuljapurkar y Boe (1999) en un estudio para la predicción de la fertilidad en USA obtienen que las predicciones en las que se combina información pasada e información subjetiva mejoran la precisión en la predicción. En Clemen (1989), pueden verse una extensa referencia bibliográfica sobre la combinación de predicciones.

### 5.2.2 Test de Stekler

Si tomamos como criterio cuantificable para comparar entre los métodos individuales o de combinación su precisión, podemos estar interesados en contrastar si todos los métodos tendrán la misma precisión esperada.

Es decir si tenemos  $j$  métodos ( $j=1\dots N$ ) y una vez que hemos decidido el indicador de precisión que vamos a tomar como criterio de comparación podemos contrastar la hipótesis de que desutilidad esperada de la predicción calculada por el método  $i$  es igual que la desutilidad esperada de la predicción calculada por el método  $j$ :

$$\text{Contrastar } E[L(y_{t+h}, \hat{y}_{t+h,t}^i)] = E[L(y_{t+h}, \hat{y}_{t+h,t}^j)]$$

contra la alternativa de que un método,  $i$  ó  $j$ , es mejor.

Stekler (1987) propone<sup>141</sup> un test basado en el ranking que ocupan el conjunto de  $h$  predicciones ( $h=1\dots T$ ) obtenidas por  $N$  métodos, contrastando que cada conjunto de predicciones tienen igual desutilidad esperada en cada método.

Así a cada predicción en cada momento  $h$  le asigna un rango de acuerdo con su precisión (el mejor recibe el rango de  $N$ , el segundo mejor  $N-1$ , y así sucesivamente). Entonces agregando periodo a periodo los rangos para cada método utilizado:

$$H^j = \sum_{h=1}^T \text{Rank}(L(y_{t+h}, \hat{y}_{t+h,t}^j)) \text{ donde } j=1\dots N$$

y la forma que adopta el estadístico del test de bondad sigue, bajo la hipótesis nula, una  $\chi_{N-1}^2$  de la forma:

$$H = \sum_{j=1}^N \frac{\left( H^j - \frac{NT}{2} \right)^2}{\frac{NT}{2}}$$

### 5.2.3 Contraste de reducción significativa del error de predicción

Cuando las predicciones que son combinadas son insesgadas, el error cuadrático medio de la predicción es igual a la varianza de los errores de predicción (Lawrence, Edmunson y O'Connor, 1986), es decir:

$$\sigma_c^2 = 0,25(\sigma_i^2 + \sigma_j^2 + 2\rho\sigma_i\sigma_j)$$

con lo que ECM de la combinación solo será menor que una predicción individual cuando:

---

<sup>141</sup> Al no disponer del artículo original hemos tomado como fuente secundaria Diebold y López (1995)

$$\frac{\sigma_j}{\sigma_i} > \frac{\rho + (\rho^2 + 3)^{0,5}}{3} = \Phi \text{ que es también el caso que } \frac{\sigma_j}{\sigma_i} < 1/\Phi$$

Cuando las predicciones que se van a combinar son sesgadas, la precisión de la predicción combinada dependerá del tamaño relativo y del signo de los errores medios de la predicción ( $\bar{Y}_t - \hat{Y}_t$ )

Si la media de los errores de las predicciones individuales i y j son respectivamente: u y v, entonces el error cuadrático medio de la predicción combinada será:

$$\sigma_c^2 = 0,25(\sigma_i^2 + \sigma_j^2 + 2\rho\sigma_i\sigma_j + (v + w)^2)$$

SI  $u=-w$ , el sesgo de una predicción anula a la otra. Si la predicción i es insesgada, pero la predicción j no ( $v^2$ ), entonces la combinación podría solo remover el 75% de este sesgo medio.

Sería entonces conveniente aplicar la corrección de Theil. Sin embargo, efectuar la corrección de las predicciones individuales antes de la combinación podría disminuir las ganancias potenciales de la combinación:

1. Si la corrección de Theil elimina el sesgo de predicciones futuras, entonces también puede eliminar los beneficios potenciales de errores medios de signo opuesto que se podrían cancelar con la combinación.
2. Es posible que los beneficios de la corrección sobre predicciones subjetivas (Goodwin, 1997)<sup>142</sup> pueda incrementar la correlación de sus errores con los de otras predicciones.

Cuando se comparan los Errores Cuadráticos Medios de dos predicciones es importante contrastar si la reducción en el error de uno de ellos es significativa. Para contrastar esta hipótesis, Granger y Newbold (1977) construyeron el siguiente contraste:

Si  $e_{t+h}^i$  y  $e_{t+h}^j$  son los errores de predicción de dos modelos<sup>143</sup> i y j, y si, son insesgados, incorrelados y proceden de una distribución normal, el contraste consta de dos etapas:

---

<sup>142</sup> No obstante la efectividad de la aplicación de la corrección depende de la validez de la hipótesis de que el patrón de los errores es constante en el tiempo. En muchas situaciones, los errores de las predicciones subjetivas no suelen ser estacionarias.

En la corrección Goodwin (1997) encontró que la corrección es más conveniente cuando las series tienen altos niveles de ruido. En particular en series de ruido blanco la corrección afecta alisando la variación en las predicciones subjetivas causadas por la reacción excesiva a movimientos aleatorios de las series.

1.  $S_{t+h} = e_{t+h}^1 + e_{t+h}^2$  y  $D_{t+h} = e_{t+h}^1 - e_{t+h}^2$
2. Se estima  $S_{t+h} = \alpha + \beta D_{t+h} + u_{t+h}$

Si el coeficiente de  $D_{t+h}$  es significativamente diferente de cero, entonces se puede concluir que el Error Cuadrático Medio del modelo  $i$  es significativamente distinto al del modelo  $j$ . Además, si el término constante de la regresión estimada es significativamente diferente de cero entonces  $e_{t+h}^i$  y/o  $e_{t+h}^j$  son sesgados.

Por tanto, ya que un término constante significativo indica la violación de uno de los supuestos necesarios para la validez del contraste, esta condición debe verificarse.

### 5.3 Propiedades de las ponderaciones estimadas

Si en la combinación de predicciones las ponderaciones permanecen estables a lo largo del tiempo, la estructura de las varianzas-covarianzas de los errores de predicción serían estables. Sin embargo si esta condición no es satisfecha, las ponderaciones pueden ser inestables a lo largo del tiempo y por tanto no hay garantía de que la predicción combinada tenga la varianza del error de predicción menor que la de alguno de los métodos de predicción.

#### 5.3.1 Estabilidad en las ponderaciones

Kang (1986) compara a partir de serie que genera la actuación de la predicción combinada y de la predicción compuesta<sup>144</sup>, y señala que cuando los pesos no son estable, la predicción compuesta puede ser una alternativa a la combinación de predicciones si los modelos individuales son conocidos.

Kang (1986) estudia la estabilidad de las ponderaciones de diferentes métodos de combinación (fijos, cambiantes a lo largo del tiempo empleando medias móviles de diferentes periodos, etc) a partir de una media porcentual de las fluctuaciones de éstas como  $\left[ \frac{(|\text{máximo}| - |\text{mínimo}|)}{|\text{mínimo}|} \right] * 100$  donde el máximo y el mínimo es obtenido a partir de los diferentes esquemas que emplea para la obtención de las ponderaciones.

---

<sup>143</sup> Deutsch, Granger y Teräsvirta (1994) Compara la reducción de la varianza de los errores obtenidos por métodos de combinación basados en modelos de transición con los obtenidos mediante la simple combinación lineal

<sup>144</sup> Cuando se tienen dos modelos  $i$  y  $j$ , una predicción compuesta es obtenida a partir de un modelo compuesto, en el cual los modelos  $i$  y  $j$  son combinados previamente. Es decir, si  $i$  emplea  $X_1$  y  $X_2$  para predecir  $Y$  en una regresión lineal múltiple y  $j$  emplea  $X_1$  y  $X_3$  para predecir  $Y$  de forma similar, un modelo compuesto emplea  $X_1$ ,  $X_2$  y  $X_3$  en la regresión. Una predicción compuesta sería la que se obtiene a partir de la estimación este último modelo.

Contrasta la hipótesis de que los pesos permanecen constantes empleando el test MANOVA (Multivariante análisis of variance) del SPSS.

Los procedimientos bayesianos óptimos implican el uso de una matriz de ventajas o ayudas posteriores sobre los modelos alternativos, derivada y calculada para modelos de parámetros fijos y cambiantes. Estas “ventajas” pueden facilitarnos el test sobre la constancia de los parámetros<sup>145</sup>.

Algunos trabajos han discutido la posibilidad de que las predicciones a combinar sean insesgadas (Clemen (1989), Diebold y Pauly (1987), Palm y Zellner (1990, Winkler (1981) en el análisis bayesiano, etc). Si las predicciones son insesgadas y los pesos de la combinación se calculan con el objetivo de minimizar el error cuadrático medio, se puede comprobar que el ECM de la combinada es menor que el ECM de cada predicción individual. Sin embargo si se considera que las predicciones son insesgadas cuando en realidad no lo son, el uso de pesos apropiados para predicciones insesgadas puede ser subóptima cuando estas son sesgadas, es decir, la presencia de sesgo en la predicción puede originar peores resultados en términos del ECM.

### 5.3.2 La distribución de los pesos de la combinación

Dickinson (1972) estudia la distribución de los pesos de la combinación. Así por ejemplo para el caso de Bates y Granger (1969) donde los pesos serían<sup>146</sup>:

$$\alpha_1 = \frac{\sum_{t=1}^n e_{2,t}^2}{\sum_{t=1}^n (e_{1,t}^2 + e_{2,t}^2)} \quad \alpha_2 = \frac{\sum_{t=1}^n e_{1,t}^2}{\sum_{t=1}^n (e_{1,t}^2 + e_{2,t}^2)}, \quad \text{estudia la distribución y genera}$$

intervalos de confianza para los pesos bajo diferentes longitudes de n y diferentes ratios

de  $\frac{\sum_{t=1}^n e_{1,t}^2}{\sum_{t=1}^n e_{2,t}^2}$  concluyendo que ha medida que el ratio aumenta la estimación puntual de  $\alpha_1$ ,

y los intervalos, reflejan mayor peso a la predicción de menor varianza en el error. Sin

<sup>145</sup> Min y Zellner (1993) estudian diversas técnicas de combinación bayesiana y desarrollan una regla bayesiana de selección de las predicciones basada en un criterio de error de predicción que indica cual de los modelos, incluido el combinado es óptimo.

<sup>146</sup> Cabría preguntarse que ocurriría si fuese posible una autocorrelación entre los errores de las predicciones individuales.

embargo a medida que incrementa  $n$ , manteniendo constante el ratio, los resultados dejan de ser robustos<sup>147</sup>.

#### 5.4 Características de los errores de la predicción combinada

En la metodología tradicional se suele resaltar únicamente los dos momentos de los errores de predicción y como consecuencia los intervalos de predicción suelen realizarse suponiendo normalidad. En los métodos individuales es habitual chequear la distribución de los errores, pero en la combinación no es habitual diagnosticar la distribución de éstos como un aspecto inherente a la predicción.

##### 5.4.1 Sesgo y eficiencia de las predicciones combinadas

Holden y Peel (1990) Analizan el sesgo y la eficiencia de las predicciones combinadas. El test más usual para comprobar la insesgadez de las predicciones es realizar  $Y_t = \alpha + \beta \hat{Y}_t^c + v_t$  si las predicciones combinadas fuesen insesgadas se debería de cumplir  $\hat{\alpha} = 0$ ,  $\hat{\beta} = 1$ , requisito que Holden y Peel (1990) han demostrado que es condición suficiente pero no necesaria de insesgadez.

Holden y Peel (1990) comprueban que esta es una condición suficiente pero no necesaria. En particular comprueban que las predicciones combinadas pueden ser insesgadas (o eficientes) y no pasar el test, por lo que proponen un test alternativo:

Bates y Granger (1969) combinan predicciones insesgadas para producir una combinación insesgada. Empleando la metodología de estos autores Holden y Peel (1990) combinan una predicción sesgada y otra insesgada y comprueban que la condición necesaria y suficiente de insesgadez es  $\alpha = (1 - \mu)\beta$ , donde  $\mu$  es la esperanza poblacional ( $E(y_t)$ ). Puesto que esta esperanza poblacional puede ser desconocida, la condición que se obtiene cuando se trabaja con una muestra se obtendría estimando

$Y_t = \alpha + \beta \hat{Y}_t^c + v_t$  a partir de MCO y la condición resultaría  $\hat{\alpha} = (1 - \hat{\beta})\bar{Y}$ , donde

$$\hat{\beta} = 1 - \frac{\sum (\hat{Y}_t^c - \bar{\hat{Y}}^c) u_t}{\sum (\hat{Y}_t^c - \bar{\hat{Y}}^c)}$$

distinto a uno, en tanto en cuanto haya correlación entre las predicciones y  $u_t$ . Es más, la predicción puede ser sesgada (y puede ocurrir cuando la predicción se deriva de una función no-simétrica) y si no hay correlación  $\hat{\beta} = 1$  (por tanto hablaríamos de eficiencia débil)

<sup>147</sup> Además un incremento en  $n$  aumenta el riesgo de dependencia entre los errores

### **5.4.2 Distribución de los errores de la predicción combinada**

La precisión en la predicción ha sido tomada como criterio para la decisión de un método de predicción, más que la especificación de la distribución del error de predicción. Así, aunque la precisión ha sido habitualmente el principal criterio para decidir entre métodos de predicción, cuando se hacen análisis de riesgo e incertidumbre, la calidad de la distribución del error de predicción en términos de normalidad, simetría y la correlación serial<sup>148</sup> puede ser tan importante como la insesgadez y la varianza del valor esperado.

Si diferentes reglas de combinación conducen a diferentes formas en la distribución del error combinado, la actitud del decisor respecto al riesgo puede ser un factor adicional in seleccionar la reglas de combinación y si la combinación de diferentes métodos conducen a diferentes formas en la distribución, entonces la elección de que técnicas incluir puede afectar también a este criterio.

No obstante, los estudios sobre la normalidad habitualmente han sido de carácter teórico y asintótico, y la normalidad en muestras pequeñas se ha estudiado a partir con test de los momentos de la muestra.

#### **5.4.2.1 Simetría de la distribución**

Meneces y Bunn (1998) estudian la forma de la distribución del error de la predicción combinada considerando la no normalidad de éstos y justifican la simetría de la distribución como un relevante criterio para decidir entre el método de combinación más apropiado ya que la distribución del error combinado puede no funcionar tan bien con respecto a la normalidad como la distribución de los errores de cada método de predicción.

En concreto el trabajo empírico que efectúan es el siguiente:

A partir de 800 combinaciones de dos métodos de predicción y seis tamaños muestrales donde los errores son especificados con diferentes distribuciones ( $N(0,1)$ ,  $y(a, b)$ ), contrastan el sesgo y la curtosis de la muestra de los errores de distribución a partir del test de Bowman-Shenton y coeficientes de cusrtosis.

Como resultados obtienen que:

---

<sup>148</sup> La correlación serial entre los errores de la predicción combinada ha sido estudiada por Diebold (1988) que mostró que la combinación basada en la regresión sin restringir puede tener los errores correlacionados.

El sesgo de la distribución de los errores persiste en las combinaciones si las distribuciones individuales son sesgadas, aunque con combinaciones más sofisticadas parece suavizarse. Y entonces en distribuciones individuales asimétricas es un error asumir normalidad en la distribución del error de la predicción combinada<sup>149</sup>.

#### 5.4.2.2 Métodos para aproximar la distribución: Regresión Cuantílica

Además de ser importante chequear la simetría de los errores de la predicción combinada puede ser interesante intentar conocer su distribución probabilística<sup>150</sup>.

Puesto que los modelos de combinación no imponen un procedimiento formal para identificar el proceso generador de datos, la expresión de la varianza teórica no es fácilmente derivada o hay dudas sobre las hipótesis asumidas en el modelo: Aunque algunos autores realizan estudios analíticos sobre la incertidumbre de los pesos de la combinación no suelen analizar explícitamente o derivar la varianza del error de predicción combinada<sup>151</sup>. En los métodos de combinación teóricos (MCO, varianza-covarianza) el método para estimar la distribución de los errores de predicción consiste en derivar una expresión para la varianza del error que puede utilizarse como una asunción de la distribución para aproximar la distribución de los errores.

El uso de métodos empíricos puede ser una alternativa. Taylor y Bunn (1999) proponen aplicar la regresión por cuantiles (Koenker y Basset (1978, 1982))<sup>152</sup> para

<sup>149</sup> Además, Menezes y Bunn (1998) ofrecen algunos resultados sobre las propiedades de la distribución de los errores combinados en relación del número de métodos empleados en la combinación.

<sup>150</sup> En ocasiones puede ser necesario para realizar intervalos de confianza de la predicción combinada, y si no se conoce la distribución del error, la predicción por intervalos se basa entonces en las varianzas de los errores y en la hipótesis de normalidad (algunos autores proponen el uso de la desigualdad de Chebyshev pero los intervalos no sería muy precisos y por tanto poco prácticos).

<sup>151</sup> Dickinson (1972) investiga la distribución teórica de los pesos (como vimoa) y del la varianza del error. En concreto, estudia la distribución probabilística para la varianza del error combinado, y compara los intervalos de confianza de ésta varianza con los intervalos de las varianzas de los errores de predicción individual. Así cuando el intervalo de la varianza de una predicción  $i$  es mucho más preciso que otra  $j$  (con el mismo nivel de confianza) no parece que el de la predicción combinada gane más en precisión. Sin embargo, para la expresión para la varianza del error no tiene en cuenta la incertidumbre en los pesos

<sup>152</sup> Koenker y Basset (1978, 1982) desarrollan la técnica de la regresión por cuantiles de una variable  $y_t$  que es función de otras variables. El  $\theta$ th cuantil de una variable  $y$  es el valor  $Q(\theta)$  para el cual  $P(y < Q(\theta)) = \theta$  una aproximación de la distribución completa puede realizarse a partir de las estimaciones de los cuantiles correspondientes a rango de valores de  $\theta$  ( $0 < \theta < 1$ ).

Así de modo intuitivo si consideramos una regresión solo con constante  $y_t = \beta_0 + e_t$  siendo  $e_t$  un vector de variables aleatorias i.i.d. a partir de la muestra se puede obtener la solución  $\beta_0(\theta)$  minimizando:

$$\sum_{t|y_t \geq \beta_0} \theta |y_t - \beta_0| + \sum_{t|y_t < \beta_0} (1 - \theta) |y_t - \beta_0|$$

ajustar los errores empíricos de predicción con el objetivo de producir modelos del error de predicción por cuantiles que son funciones de  $h$  (horizonte temporal de la predicción)<sup>153</sup>.

La serie de errores a partir de la cual efectúan la regresión cuantílica la obtienen de la siguiente manera:

A partir del pasado de la serie, obtienen los errores de predicción para el método de combinación considerado, comenzando por el principio y reestimando la varianza de esos errores cada vez.

El modelo de combinación no es reestimado cada vez y las varianzas de los errores a lo largo del tiempo se basan en el modelo generado en las observaciones de la muestra.

La regresión cuantílica:

Taylor y Bunn (1999) consideran el caso general de un modelo de heterocedasticidad sistemática:

$$y_t = \mu_t(x_t) + \sigma_t(x_t)e_t^{154}$$

donde  $x_t$  es una columna de variables explicativas,  $\mu_t(x_t)$  puede considerarse como la media condicional del proceso de regresión,  $\sigma_t(x_t)$  como la escala condicional y  $e_t$  es un error independiente del vector  $x_t$ .

El  $\theta$ th cuantil de  $e_t$  es el valor  $Q_e(\theta)$  para el cual  $P(e_t < Q_e(\theta)) = \theta$ .

Las funciones cuantiles condicionadas de  $y_t$  serán entonces:

$$Q_{y_t}(\theta|x_t) = \mu_t(x_t) + \sigma_t(x_t)Q_e(\theta)$$

Si se considera que  $\mu_t$  y  $\sigma_t$  son funciones lineales de  $x_t$ , el cual tiene 1 como primer elemento:

---

<sup>153</sup> Además de proponer el método, estudian y comparan los problemas derivados en la estimación de la distribución de predicción para tres métodos de combinación teóricos ( mínimos cuadrados ordinarios (con término independiente), mínima varianza y la media aritmética) viendo su potencial para ver en que medida puede ser considerado en aproximaciones más empíricas.

<sup>154</sup>  $\mu_t$  y  $\sigma_t$  no tienen porque depender del mismo vector  $x_t$

$$Q_{y_t}(\theta|x_t) = x_t\beta + (1 + x_t\gamma)Q_e(\theta)$$

donde  $\beta$  y  $\gamma$  son vectores de parámetros.

Si todos los elementos de  $\gamma$  son cero es lo mismo que asumir que el error de  $y_t$  es i.i.d. y entonces la ecuación anterior puede escribirse como:

$$Q_{y_t}(\theta|x_t) = x_t\beta(\theta) \text{ donde } \beta(\theta) \text{ es un vector de parámetros dependientes de } \theta.$$

Se define entonces como la  $\theta$ th regresión cuantil ( $0 < \theta < 1$ ) como una solución  $\hat{\beta}(\theta)$  al problema de minimizar la regresión cuantil:

$$\sum_{t|y_t \geq x_t\beta} \theta |y_t - x_t\beta| + \sum_{t|y_t < x_t\beta} (1 - \theta) |y_t - x_t\beta|.$$

Esa expresión puede ser esta  $\sum_{t|y_t \geq \hat{Y}_t\alpha} \theta |y_t - \hat{Y}_t\alpha| + \sum_{t|y_t < \hat{Y}_t\alpha} (1 - \theta) |y_t - \hat{Y}_t\alpha|$ , puesto que

$y_t = x_t\beta + e_t$ , en el caso de combinación de predicciones es  $y_t = \hat{Y}_t\alpha + e_t$ <sup>155</sup>.

Taylor y Bunn (1999) aplican ya esta metodología a los errores pasados de la serie<sup>156</sup>, como describimos a continuación:

Para predicciones de series temporales, la varianza del error de predicción no es constante y es función del horizonte temporal  $h$ <sup>157</sup>.

Si consideramos la variación en las series de los errores de predicción desde varios horizontes temporales como una sistemática heterocedasticidad, se puede usar la regresión cuantil para regresar el error de predicción a partir de variables explicativas escogidas para dar un modelo estimado para el cuantil seleccionado del error de predicción.

$$\sum_{t|y_t \geq \hat{Y}_t\alpha} \theta |y_t - \hat{Y}_t\alpha| + \sum_{t|y_t < \hat{Y}_t\alpha} (1 - \theta) |y_t - \hat{Y}_t\alpha| \Rightarrow \sum_{t|y_t \geq \hat{y}_t^c} \theta |y_t - \hat{y}_t^c| + \sum_{t|y_t < \hat{y}_t^c} (1 - \theta) |y_t - \hat{y}_t^c|$$

<sup>155</sup> Si bien hemos empleado la notación usada por Koenker y Basset (1978, 1982) y Taylor y Bunn (1999), a partir de aquí, empleamos esta expresión, con el propósito de mantener la coherencia en la notación empleada en el resto del documento.

<sup>156</sup> Es necesario que haya bastantes errores de predicción pasados.

<sup>157</sup> Si se genera una distribución de los errores para una amplitud temporal puede que esta distribución no fuese constante en el tiempo.

$$\Rightarrow \sum_{t|y_t \geq \hat{y}_t \alpha} \theta |e(\hat{y}_t^i)| + \sum_{t|y_t < \hat{y}_t \alpha} (1-\theta) |e(\hat{y}_t^i)| \Rightarrow \sum_{t|y_t \geq \hat{y}_t \alpha} \theta |e(\hat{y}_t^i(x_t))| + \sum_{t|y_t < \hat{y}_t \alpha} (1-\theta) |e(\hat{y}_t^i(x_t))|$$

En este caso las variables explicativas que toman será h o una función de h<sup>158</sup>. Así, emplean los errores de predicción obtenidos sobre diferentes longitudes h<sup>159</sup>, y entonces

$$\Rightarrow \sum_{t|y_t \geq \hat{y}_t \alpha} \theta |e(h)| + \sum_{t|y_t < \hat{y}_t \alpha} (1-\theta) |e(h)|$$

La regresión por cuantiles del vector de los errores de predicción se regresa sobre simples funciones de h (h, h<sup>2</sup>, etc).

Por ejemplo el 95 cuantil de los errores de predicción Q<sub>fe</sub>(0,95) debería estimarse usando  $\theta=0,95$  en

$$\sum_{t|y_t \geq \hat{y}_t \alpha} \theta |e(h)| + \sum_{t|y_t < \hat{y}_t \alpha} (1-\theta) |e(h)|$$

y si por ejemplo el error es función de h, tal que  $a+bh+ch^2$ , entonces

$$\sum_{t|y_t \geq \hat{y}_t \alpha} \theta |a+bh+ch^2| + \sum_{t|y_t < \hat{y}_t \alpha} (1-\theta) |a+bh+ch^2|$$

Con el vector de errores de predicción como variable dependiente y vectores como h y h<sup>2</sup> como regresores. El resultado podría ser una ecuación de la forma:  $\hat{Q}_e(\theta) = \hat{a} + \hat{b}h + \hat{c}h^2$  donde a, b y c son parámetros estimados por la regresión cuantil.

A partir de las estimaciones que queramos realizar, podemos aproximar las distribución de los errores de cualquier método de predicción combinada.

<sup>158</sup> Lo justifican diciendo que si los errores eran función de otras variables que forman parte de los modelos individuales, podía ocurrir que no se puedan hacer inferencias para la distribución de los errores. De ahí que empleen como variables explicativa h

<sup>159</sup> Por ejemplo si se quieren construir distribuciones de predicción para diferentes adelantos temporales h (1,3, 6, 9,12, 15 y 18), las series individuales se construirían con los errores de predicción efectuados para un periodo de adelanto, otra para 3, etc. Se toman los elementos de las h series como cogiendo un valor de h cuando el correspondiente elemento de las series de los errores de predicción es el error para un determinado h. Es decir, si se construyen distribuciones de predicción para 1, 3, etc las series h estarían constituidas por 1's, seguidas por 3's, etc.

Taylor (1999) estudia el sesgo y la varianza del estimador de la regresión por cuantiles. Así, el estimador  $\hat{Q}_e(\theta)$  del  $\theta$ th cuantil será insesgado si la proporción de valores de  $e_t$  que son menores  $\hat{Q}_e(\theta)$  tiende a  $\theta$  cuando la muestra se hace mayor.

Taylor y Bunn (1998) dan una condición necesaria y suficiente para probar la condición de insesgadez: Un estimador  $\hat{Q}_e(\theta)$  del  $\theta$ th cuantil de  $e_t$  será insesgado si y solo si la regresión del  $\theta$ th cuantil de la variable  $(e_t - \hat{Q}_t(\theta))$  sobre una constante y sin variables explicativas resulta en una constante cero cuando el número de observaciones aumenta.

### 5.5 Frontera de eficiencia

Hasta ahora sólo nos hemos fijado en la precisión de los métodos de predicción como criterio para clasificarlos y evaluarlos. Sin embargo no hay que olvidar que tanto la complejidad como el grado de dificultad para entender los resultados obtenidos en cada método son también importantes a la hora de optar por una técnica u otra.

Teniendo en cuenta el grado de complejidad de los métodos considerados y los resultados, se puede elaborar una frontera de eficiencia con el propósito de ilustrar el *trade-off* existente entre la complejidad y la precisión de las técnicas.

Según Lewandowski (1984) el grado de complejidad es el tiempo, medido en horas, que se requiere para enseñar a una persona, sin conocimientos especiales en predicción o estadística los principios de un método. El índice que elabora, que además tiene en cuenta el grado de comprensión de los resultados obtenidos, coincide básicamente con el utilizado por Makridakis (1983) que lo elabora en función de la complejidad percibida por él. Por tanto el criterio adoptado para elaborar el índice es subjetivo.

Por ejemplo, en técnicas de series temporales, parece existir acuerdo entre varios autores respecto al hecho de que la metodología de Box-Jenkins es la que requiere el mayor tiempo, pues es necesario observar el gráfico de cada serie, las funciones de autocorrelación, identificar el modelo, estimar sus parámetros, chequear las autocorrelaciones de los residuos, etc. Otros métodos tienen una base prácticamente automática, si bien es necesario detenerse en detectar la estacionalidad, la existencia de tendencia etc.

De acuerdo con los criterios de cualquiera de los autores anteriores, los índices de complejidad para predicciones basadas en análisis de series temporales sería:

Métodos ingenuos: 1

Métodos de alisado exponencial: 5

Métodos ARIMA: 10

En la combinación de predicciones se promedia el índice de complejidad de los métodos que se combinan, en caso de emplear un método de media aritmética. Sin embargo, si se emplean técnicas más complicadas, el índice de complejidad se debería de obtener sumando los índices individuales con el índice de complejidad (arbitrario) del método de combinación que empleemos.

Teniendo en cuenta la clasificación de los diferentes métodos de acuerdo con su precisión y con su complejidad es posible elaborar un de frontera de eficiencia, que ilustre el *trade-off* entre mayor precisión y mayor complejidad.

Makridakis (1983) construye la frontera teniendo en cuenta el número de veces que cada método es el mejor o el segundo mejor y el índice de complejidad y Lewandowski (1984) relaciona la complejidad con los EAPM.

La elección entre cada método dependerá entonces de cada situación considerada y de las preferencias del agente predictor<sup>160</sup>.

---

<sup>160</sup> Moreno y López (1999) elaboran fronteras de eficiencia para métodos de predicción de series temporales (para variables de la economía asturiana) y métodos de combinación lineal, concluyendo que la calidad de las predicciones no se incrementa cuando la complejidad del método aumenta, y que a pesar de la complejidad del método de Box-Jenkins, ofrece peores resultados que otros más sencillos. La metodología ARIMA es muy compleja, sus resultados son difíciles de interpretar y la calidad de sus predicciones es baja. Además la combinación de las predicciones obtenidas por el método más sencillo (ingenuo) con el más complejo (ARIMA) es la que mejores resultados ofrece.

## BIBLIOGRAFÍA

- ABAD, A.M.; CRISTÓBAL, A. y QUILIS, E.M. (2000): “Fluctuaciones económicas, puntos de giro y clasificación cíclica”, *Boletín Trimestral de Coyuntura*, nº 78, 1-31.
- AGNEW, C.E. (1986): “Bayesian consensus forecast of macroeconomic variables”, *Journal of Forecasting*, Vol. 3, nº 3, 363-376.
- AGRESTI, A. (1990): *Categorical Data Analysis*, John Wiley & Sons, Inc. Wiley InterScience.
- AGUIRRE, A. (1994): *Introducción al tratamiento de series temporales*, Ed. Díaz de Santos, Madrid.
- AHLERS, D. Y LAKONISHOK, J. (1983): “A study of economist’ consensus forecast”, *Management Science*, Vol. 29 nº 10, 113-1125.
- ALCAIDE, J. (2001): “Expectativas económicas 2001-2002: Fuentes privadas”, *Revista Fuentes Estadísticas*, nº 60. [www.fuentesestadisticas.com](http://www.fuentesestadisticas.com).
- ALVARGONZÁLEZ, M. (2003): *Medidas doble cuadráticas de información. Algunas aplicaciones económicas*. Tesis doctoral. Universidad de Oviedo.
- ANDERSEN, A. y WEISS, A. (1984): “Forecasting: The Box-Jenkins Approach”, en *The Forecasting Accuracy of Major Time Series Methods*, John Wiley & Sons, 167-199.
- ANDERSON, O. (1952): “The business test of the IFO-institute for economic research, Munich, and its theoretical model”, *Revue de l’Institut International de Statistique*, nº 20, 1-17.
- ARANDA, D., GONZÁLEZ, A. PETITBÓ, A. (1994): “Las encuestas de opiniones empresariales: Un instrumento útil para conocer la coyuntura industrial”, *Economía Industrial*, nº 299, 137-156.
- ARMSTRONG, J.S. (1978): “Forecasting with Econometric Methods: Folklore versus Fact”, *Journal of Business*, Vol. 51, nº 4, 549-564.
- ARMSTRONG, J.S. (2001): *Principles of Forecasting*, Kluwer Academic Publishers, Massachusetts.
- ARMSTRONG, J.S. y COLLOPY, F. (1993): “Causal forces: structuring knowledge for time series extrapolation”, *Journal of Forecasting*, Vol. 12, nº1, 103-115.
- ARMSTRONG, J.S. Y FILDES, R. (1995): "Correspondence on the selection of error measures for comparisons among forecasting methods", *Journal of Forecasting*, Vol. 14, nº 1, 67-71.
- ARMSTRONG, J.S., MORWITZ, V.G. y KUMAR, V. (2000): “Sales forecasting for existing consumer products and services: Do purchase intentions contribute to accuracy?”, *International Journal of Forecasting*, Vol. 16, 383-397.
- ARTIS, M.J.; BLANDEN-HOVELL y ZHANG, W. (1995): “Turning points in the international business cycle: “An analysis of the OCDE leading indicators for the G-7 countries”, *OCDE Economic Studies*, nº 24, 125-165.
- ASH, J.C.K.; SMYTH, D.J. y HERAVI S.M. (1998): “Are ODCE forecast rational and useful?: A directional analisis”, *International Journal of Forecasting*, Vol.14, 381-391.
- ASHTON, A.H. y ASHTON, R.H. (1985): “Aggregating subjective forecasts: Some empirical results”, *Management Science*, Vol. 31, nº12, 1499-1508.
- AUKUTSIONEK, S.P. y BELIANIN, A.V. (2001): “Quality of forecasts and business performance: A survey study of Russian managers”, *Journal of Economic Psychology*, nº 22, 661-692.

- AZNAR, A. y TRÍVEZ, F.J. (1993): *Métodos de predicción en Economía, Tomo I: Fundamentos, Input-Output, Modelos econométricos y métodos no paramétricos de series temporales*, Ed. Ariel, Barcelona.
- AZNAR, A. y TRÍVEZ, F.J. (1993): *Métodos de predicción en Economía, Tomo II: Análisis de series temporales*, Ed. Ariel, Barcelona.
- BALCOMBE, K. (1996): "The Carlson-Parkin method applied to NZ price expectations using QSBO survey data", *Economic Letters*, nº 51, 51-57.
- BATCHELOR, R. (2001): "How useful are the forecast of intergovernmental agencies? The IMF and OCDE versus the consensus", *Applied Economics*, nº 33, 225-235.
- BATCHELOR, R. y DUA, P. (1987): "The accuracy and rationality of UK inflation expectations: some quantitative evidence", *Applied Economics*, nº 19, 819-828.
- BATCHELOR, R. y DUA, P. (1992): "Survey expectations in time series consumption function", *The Review of Economics and Statistics*, 598-606.
- BATCHELOR, R. y DUA, P. (1995): "Forecaster diversity and the benefits of combining forecasts", *Management Science*, Vol. 41, nº1, 68-75.
- BATCHELOR, R.A. (1981): "Aggregate expectations under stable laws", *Journal of Econometrics*, nº 16, 199-210.
- BATCHELOR, R.A. (1986): "Quantitative vs. qualitative measures of inflation expectations", *Oxford Bulletin of Economics and Statistics*, Vol 48, nº2, 99-120.
- BATCHELOR, R.A. (1990): "All forecasters are equal", *Journal of Business and Economics Statistics*, Vol. 8, nº 1, 143-144.
- BATCHELOR, R.A. y ORR. A.B. (1987): "Inflation expectations revisited", *Economica*, nº 55, 317-331.
- BATES, J.M. y GRANGER, C.W.J. (1969): "The Combination of Forecast", *Operational Research Quarterly* Vol. 20, nº 4, 451-468.
- BELIS, M. y GUIASU, S. (1968): "A quantitative-qualitative Measure of Information in Cibernetic System", *I.E.E.E. Trans. Inf. Th*, nº 14, 593-594.
- BENNETT, A. (1984), "Output expectations of manufacturing industry", *Applied Economics*, nº 16, 869-879.
- BERK, J. M. (1999), "Measuring inflation expectations: a survey data approach", *Applied Economics*, nº 31, 1467-1480.
- BERK, J.A. (2000): "Consumers' inflation expectations and monetary policy in Europe", *Contemporary Economic Policy*, nº 20, 122-132.
- BESSLER, D.W. y CHAMBEWRLAIN, J. (1987): "On bayesian composite forecasting", *Omega*, nº 15, 43-48.
- BOHARA, A., McNOWN, R. y BATTS, J.T. (1987): "A re-evaluation of the combination and adjustment of forecasts", *Applied Economics*, nº 19, 437-445.
- BOHARA, A.K. (1991): "Testing the Rational-Expectations Hypothesis: Further evidende", *Journal of Bussines and Economic Statistics*, Vol 9, nº3. 337-340.
- BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heterocedasticity", *Journal of Econometrics*, nº 31, 307-327.
- BONHAM, C.S. y COHEN, R.H. (2001): "To aggregate, pool, or neither: Testing the Rational-Expectations Hypothesis using survey data", *Journal of Bussines and Economic Statistics*, Vol 19, nº3. 278-291.
- BOPP. A.E. (1985): "On combining forecasts: Some extensions and results", *Management Science*, Vol. 31, nº 12, 1492-1498.
- BORDLEY, R.F. (1982): "The combination of forecast: a Bayesian approach", *Journal of Operational Research Society*, Vol. 33, 171-174.
- BOX, G.E.P. y G.M. JENKINS (1976): *Time Series Analysis: Forecasting and Control*, (2ª ed.), Ed. Holden-Day, San Francisco.

- BOX, G.E.P. y TIAO, G.C. (1976): "Comparison of Forecast and Actuality", *Applied Statistics*, Vol. 25, nº 3, 195-200.
- BOX, G.E.P., G.M. JENKINS y REINSEL, G.C. (1994): *Time Series Analysis: Forecasting and Control*, (3ª ed.), Prentice-Hall, Inc, New Jersey.
- BRIER, G.W. (1950): "Verification of forecast expressed in terms of probability", *Monthly Weather Review*, Vol. 78 nº 1, 1-3.
- BROWN, R.G. (1959): *Statistical Forecasting for Inventory Control*, McGraw-Hill Book Company.
- BRY, G. y BOSCHAN, C. (1971): "Cyclical Analysis of Time Series: Selected Procedures and Computer Programs", *National Bureau of Economic Research, Technical Papers*, nº 20. New York: Columbia University Press.
- BUNN, D. W. (1975): "A bayesian approach to the linear combination of forecast", *Operational Research Quarterly*, Vol. 26, nº 2, 325-329.
- BUNN, D. y WRIGHT, G. (1991): "Interaction of judgmental and statistical forecasting methods: Issues and analysis", *Management Science*, Vol. 37, nº 5, 501-518.
- BUNN, D.W. y MUSTAFAOGLU, M.M. (1978): "Forecasting political risk", *Management Science*, Vol. 24, nº 15, 1557-1567.
- BUNN, D.W. y TAYLOR, J.W. (2001): "Setting accuracy targets for short-term judgemental sales forecasting", *International Journal of Forecasting*, nº 17, 159-169.
- CARBONE, R.; ANDERSEN, A.; CORRIVEAU, Y. y CORSON, P.P. (1983): "Comparing for different time series methods the value of technical expertise individualized analysis, and judgemental adjustment", *Management Science*, Vol. 29, nº 5.
- CARIDAD y OCERIN, J. M. (1998): *Modelos econométricos y series temporales*, Ed. Reverté, Barcelona.
- CARLSON, J.A. y PARKIN, M. (1975): "Inflation Expectations", *Economica*, nº 42, 123-138.
- CASKEY, J. (1985): "Formation a price expecations: A bayesian approach", *The American Economic Review*, Vol. 75, nº 4, 768-776.
- CASTAÑO, E. y MELO, L.F. (2000): "Métodos de combinación de pronósticos: Una aplicación a la inflación colombiana", *Lecturas de Economía*, nº 52, 1-58.
- CHO, D.W. (2002): "Do revisions improve forecasts?", *International Journal of Forecasting*, nº 18, 107-115.
- CLARK, T.E. y McCRACKEN, M.W. (2001): "Test of equal forecast accuracy and encompassing for nested models", *Journal of Econometrics*, Vol. 105, nº 1, 85-111.
- CLAVERÍA, O. Y PONS, E. (2001): "Una estimación de las expectativas sobre precios industriales en Cataluña a partir de encuestas de opinión", *Actas IV Encuentro de Economía Aplicada* (Reus).
- CLAVERIA, O., PONS, E. Y SURIÑACH, J. (2002): "Predicción de los precios de manufacturas en España a partir de encuestas de opinión: Un nuevo método", *Actas V Encuentro de Economía Aplicada* (Oviedo).
- CLEMEN, R.T. (1987): "Combining overlapping information", *Management Science*, Vol. 33, nº 3, 373-380.
- CLEMEN, R.T. y WINKLER, R.L. (1985): "Limits for the precision and value of information from dependent sources", *Operations Research*, Vol. 33, nº 2, 427-442.
- CLEMEN, R.T. y WINKLER, R.L. (1986): "Combining economic forecasts", *Journal of Business and Economic Statistics*, Vol. 4, nº 1, 39-46.

- CLEMEN, R.T. y WINKLER, R.L. (1990): "Unanimity and compromise among probability forecasters", *Management Science*, Vol. 36, nº 7, 767-779.
- CLEMEN, R.T. y WINKLER, R.L. (1993): "Aggregating point estimates: A flexible modeling approach", *Management Science*, Vol. 39, nº.4, 501-515.
- CLEMEN, T. (1989): "Combining forecast: A review and annotated bibliography", *International Journal of Forecasting*, Vol. 5, 559-583.
- CLEMENT, M.B. (1999): "Analyst forecast accuracy", *Journal Accounting and Economics*, nº 27, 285-303.
- CLEMENTS, M.P. (2003): "Some possible directions for future research", *International Journal of Forecasting*, 19, 1-3.
- CLEMENTS, M.P. y HENDRY, D.F. (1993): "On the limitations of comparing mean square forecast errors", *Journal of Forecasting*, Vol. 12, 617-637.
- CLEMENTS, M.P. y HENDRY, D.F. (2002): *A companion Economic Forecasting*, Blackwell Publishers.
- COHEN, J. (1960): "A coefficient of agreement for nominal scales", *Educational Psychology Measures*, nº 20, 37-47.
- COHEN, J. (1968): "Weighted Kappa: nominal scale agreement with provision for scaled and disagreement or partial credit", *Psychological Bulletin*, nº 70, 213-220
- CONROY, R. y HARRIS, R. (1987): "Consensus forecast of corporate earnings: Analysts' forecast and time series methods", *Management Science*, Vol. 33 nº 6, 725-738.
- COOLEY, T.F., PRESCOTT, E.C. (1976): "Estimation in the presence of stochastic parameter variation", *Econometrica*, Vol. 44, nº 1, 167-184.
- COULSON, N.E. Y ROBINS, R.P. (1993): "Forecast Combination in a Dinamic Setting", *Journal of Forecasting*, nº 12, 63-67
- COX, D. R. (1961): "Prediction by Exponentially Weighted Moving Averages and Related Methods", *Journal of the Royal Statistical Society*, Serie B, 414-422.
- CROUSHORE, D. (1998): *Evaluating inflation forecasts*, Federal Reserve Bank of Philadelphia, Documento de trabajo, nº 98-14.
- CURRIE, D. (1994): "Economic forecast: their relevance and use", en *Applied Economic Forecasting Techniques*, Stephen Hall, London.
- D'ESOPPO, D.A. (1961): "A note on forecasting by the exponential smoothing operator", *Operations Research*, Vol. 5, nº 5, 667-668.
- DANES, M. (1975): "The measurement and explanation of inflationary expectations in Australia", *Australian Economic Papers*, 19, 75-87.
- DASGUPTA, S. y LAHIRI, K. (1992): "A comparative study of alternative methods of quantifying qualitative survey responses using NAPM data", *Journal of Business and Economic Statistics*, Vol. 10, nº 4, 391-400.
- DAUTEN, C. A. y VALENTINE, L. M. (1968): *Business Cycles and Forecasting*, South-Western Publishing Company, Ohio.
- DAWID, A. P. (1984): "Statistical Theory: The precuential approach (with discussion)", *Journal of the Royal Statistical Society*, Serie A, 147 (part 2), 278-292.
- DeGROOT, M.H. (1962): " Uncertainty, information and sequential experiments", *Annals of Mathematical Statistics*, Vol. 33, nº 2, 404-419.
- DeGROOT, M.H. (1970): *Optimal Statistical Decisions*, McGraw-Hill, Inc.
- DeGROOT, M.H. (1974): "Reaching Consensus", *Journal of the American Statistical Association*, Vol. 69, nº 345, 118-121.

- DELGADO, F.J. (1997): *Análisis de las técnicas de redes neuronales artificiales para la predicción univariante de series temporales*. Universidad de Oviedo. trabajo de Investigación.
- DEUTSCH, M.; GRANGER, C.W.J. Y TERÄSVIRTA, T. (1994): "The combination of forecast using changing weights", *International Journal of Forecasting*, nº 10, 47-57.
- DHRYMES, J. y PERISTIANI, C. (1988): "A comparison of the forecasting performance of WEFA and ARIMA time series methods", *International Journal of Forecasting*, Vol. 4, 81-101.
- DICKINSON, J.P. (1972): "Some statistical results in the combination of forecast", *Operational Research Quarterly*, Vol. 24, nº 2, 253-260.
- DICKS, G. y BURRELL, A. (1994): "Forecasting in practice", en *Applied Economic Forecasting Techniques*, Stephen Hall, London.
- DIEBOLD, F.X. (2001): *Elements of Forecasting*, South-Western.
- DIEBOLD, F.X. (1988): "Serial correlation and the combination of forecast", *Journal of Business and Economic Statistics*, nº 6, 105-111.
- DIEBOLD, F.X. y KILIAN, L. (2000): "Unit-Root test are useful for selecting forecasting models", *Journal of Business and Economic Statistics*, Vol. 18, nº 3. 265-273.
- DIEBOLD, F.X. y LÓPEZ, J.A. (1995): *Forecast Evaluation and Combination*, Working Paper, University of Pennsylvania.
- DIEBOLD, F.X. y MARIANO, R.S. (1995): "Comparing predictive accuracy", *Journal of Business & Economic Statistics*, Vol. 13, nº 3, 253-263.
- DIEBOLD, F.X. y PAULY, P. (1987): "Structural change and the combination of forecast", *Journal of Forecasting*, Vol. 6, 21-40.
- DIEBOLD, F.X. y PAULY, P. (1990): "The use of prior information in forecast combination", *International Journal of Forecasting*, nº 6, 503-508.
- DIEBOLD, F.X. y RUDEBUSCH, G. D. (1991): "Turning point prediction with the composite leading index: An ex ante analysis", en *Leading economic indicators: New approaches and forecasting records*, Lahiri, K. y Moore G. H. (eds), Cambridge University Press, 231-256.
- DIEBOLD, F.X. y RUDEBUSCH, G.D. (1989): "Scoring the leading indicators", *Journal of Business*, nº 62, 369-391.
- DIEBOLD, F.X. y RUDEBUSCH, G.D. (1990): "A Nonparametric Investigation of Duration Dependence in the American Business Cycle", *Journal of Political Economy*, nº 98, 595-616.
- DIEBOLD, F.X. y RUDEBUSCH, G.D. (1991): "Forecasting output with the composite leading index: A Real-Time analysis", *Journal of the American Statistical Association*, Vol. 86, nº 415, 603-610
- DIEBOLD, F.X., GUNTHER, T.A., y TAY, A.S. (1997): *Evaluating density forecasts*, Research Department of the Federal Reserve Bank of Philadelphia, Working paper 97-6.
- DIEBOLD, F.X., GUNTHER, T.A., y TAY, A.S. (1998): "Evaluating density forecasts with applications to financial risk management", *International Economic Review*, Vol. 39, nº 4, 863-883.
- DOMINITZ, J. y MANSKI, C.F. (1997): "Using expectations data to study subjective income expectations", *American Statistical Association*, Vol. 92, nº 439, 855-867.
- DONALDSON, R.G. Y KAMSTRA, M. (1996): "Forecast combining with Neuronal Networks", *Journal of Forecasting*, nº 15, 59-61.

- DRIVER, C. y URGA, G. (2002): "Transforming Qualitative Survey Data: Performance Comparisons for the UK", *Forthcoming in Oxford Bulletin of Economic and Statistics*.
- DRIVER, C., IMAI, K., TEMPLE, P. y URGA, G. (2003): "The Effects of Uncertainty on UK Investment Authorisation: Homogenous vs Heterogenous Estimators" *Forthcoming in Empirical Economics*.
- DUA, P. (1988): "Multiperiod Forecasts of Interest Rates", *Journal of Business and Economic Statistics*, Vol 6, nº 3. 381-384.
- DUDLEY, L. (1987): "Explaining forecasting bias: The case of real exchange rate variance", *Applied Economics*, nº 19, 1249-1260.
- DuMOUCHEL, W.H. y HARRIS, J.E. (1983): "Bayes methods for combining the results of cancer studies in humans and other species (with comments and rejoinder)", *Journal of the American Statistical Association*, Vol.78, nº 382, 293-315.
- EDMUNDSON, R.H. (1990): "Descomposition: A strategy for judgemental forecasting", *Journal of Forecasting*, Vol. 9, 305-314.
- ELLIOT, G., TIMMERMANN, A. (2002): *Optimal forecast combinations under general loss functions and forecast Error distributions*, University of California, San Diego, Working Paper 2002-08,
- ELTON, E.J.; GRUBER, M.J. y GULTEKIN, M. (1981): "Expectations and share prices", *Management Science*, Vol. 27, nº 9, 975-987.
- ENDERS, W. (1995): *Applied Econometric Time Series*, John Wiley & Sons.
- ENGLE, R. (1982): "Autoregressive Conditional Heterocedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, nº 50, 987-1008.
- ENGLE, R. F. y GRANGER, C. (1987): "Co-integration and Error Correction: Representation, Estimation and Testing", *Econometrica*, nº 55, 251-276.
- ESTRADA, A. Y URTASUN, A. (1998): *Cuantificación de expectativas a partir de las encuestas de opinión*, Servicio de Estudios del Banco de España, Documento de Trabajo nº 9803,
- EUROPEAN COMMISSION (2003): *Methodology of the country weights used for aggregation purpose*, Economic studies and research, Business Surveys, Documento ECOFIN.
- EUROPEAN COMMISSION (2003): *The Joint Harmonised EU Programme of Business and Consumer Surveys*. User Guide (2/06/2003), Economic studies and research, Business Surveys, Documento ECOFIN.
- FAIR, R.C. Y SHILLER, (1989): "The informational content of ex ante forecasts", *Review of Economics and Statistics*, nº 71, 3325-331.
- FAIR, R.C. Y SHILLER, (1990): "Comparing Information in forecast from econometrics models", *American Economic Review*, nº 80, 375-389.
- FAN, D.K., LAU, K-N., LEUNG, P-L. (1996): "Combining ordinal forecasts with an application in financial market", *Journal of Forecasting*, Vol. 15, 37-48.
- FANG, Y. (2003): "Forecasting combination and encompassing tests", *International Journal of Forecasting*, Vol. 19, 87-94.
- FELDMAN, M. (1987): "An example of convergence to rational expectations with heterogeneous beliefs", *International Economic Review*, Vol. 28, nº 3, 635-650.
- FERNANDEZ, C.; LEY, E. y STEEL M. (2001): "Benchmark priors for bayesian model averaging", *Journal of Econometrics*, Vol. 100, 381-427.
- FERNÁNDEZ, F. (1995): *El problema de la predicción en series temporales: Aplicaciones del caos determinista*, Servicio de publicaciones, Universidad de las Palmas de Gran Canaria.

- FILDES, R. (1984): "Bayesian Forecasting", en *The Forecasting Accuracy of Major Time Series Methods*, John Wiley & Sons, 221-237.
- FINETTI, B. (1964): "Foresight: Its logical laws, its subjective sources", en *Studies in Subjective Probability*, Kyburg, H.E. y Smokler H.E. editores, John Wiley, New York, 97-158. Originalmente publicado en los *Annales de l'institut Henri Poincaré*, Vol. 7, (1937).
- FIRTH, M. (1977): *Forecasting Methods in Business and Management*, Edward Arnold, London.
- FISCHER, I. Y HARVEY, N. (1999): "Combining forecasts: What information do judges need to outperform the simple average?", *International Journal of Forecasting*, Vol. 15, nº 3, 227-246.
- FISCHHOFF, B. (1988): "Judgemental aspects of forecasting: Needs and possible trends", *International Journal of Forecasting*, Vol. 4, 331-339.
- FOMBY, T.B. y SAMANTA, S.K. (1993): "Average of forecasts: an improvement", *Applied Economics*, nº 25, 689-694.
- FOSTER, D.P. y VOHRA, R.V. (1993): "A randomization rule for selecting forecasts", *Operations Research Society of America*, Vol. 41, nº 4, 704-803.
- FUHRER, J.C. (1992): "Inferring changes in expectation behaviour over time: An application of Nonlinear Time-Varying-Parameters Estimation", *Journal of Business and Economic Statistics*, Vol. 10, nº 2, 169-177.
- GALLO, G.M.; GRANJER, C.W.J. Y JEON, Y. (1999): *The impact of the use of forecast in information sets*, University of California, San Diego, Discussion paper 99-18.
- GARCIA-FERRER, A. Y BUJOSA-BRUN, M. (2000): "Forecasting OECD industrial turning points using unobserved components models with business survey data", *International Journal of Forecasting*, nº 16, 207-227.
- GARDES, F., GHABRI, S., MADRE, J.L. y PICHERY, M.C. (1998): *Rationality of price and unemployment expectations: test on French qualitative Micro-Data*, Working Paper 1' IME, nº98, Septiembre, <http://web.hhs.se/stat/workshops/ec2/ghabri.pdf>.
- GARRATT, A., LEE, K., PESARAN, M.H. y SHIN, Y. (1999): *Forecast Uncertainties in macroeconomic Modelling: An application to the UK economy*, ESCR Working Paper, <http://www.econ.cam.ac.uk/faculty/pesaran/prob.pdf>
- GAVER, D. P.; DRAPER, D.; GOEL, P. K.; GREENHOUSE, J. B.; HEDGES, L. V.; MORRIS, C. N. y WATERAUX, C. (1992): *Combining Information: Statistical Issues and Opportunities for Research*, National Academy Press, Washington, D.C.
- GEWEKE, J. (2001): "Bayesian econometrics and forecasting", *Journal of Econometrics*, 100, 11-15.
- GHOSH, D. y DUTT, S.D. (2000): "Macroeconomic forecasts: an empirical test of rationality and parameter stability", *Applied Economic Letters*, nº 7, 49-52.
- GIACOMINI, R. y KOMUNJER, I. (2002): *Evaluation and combination of conditional quantile forecasts*, Boston College Department of Economics, working paper in economics nº 571.
- GIL, M.A. (1979): *Incertidumbre y Utilidad*, Tesis Doctoral, Universidad de Oviedo.
- GIL, P. (1975): "Medidas de incertidumbre e información en problemas de decisión estadística", *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales*, Madrid, 549-610.
- GIL, P. (1981): *Teoría Matemática de la Información*, Ed. ICE, Madrid

- GÓMEZ, V. Y BENGOCHEA, P. (2000): "Estimación del components cíclico de las series económicas con filtros pasabanda: Una aplicación a los indicadores de clima e índices de producción industrial", *Investigaciones Económicas*, Vol. XXIV (2), 473-485.
- GOODWIN, P. (1997): "Adjusting judgemental extrapolations using Theil's Method and Discounted Weighted Regression", *Journal of Forecasting*, Vol. 16, 37-46.
- GOODWIN, P. (2000): "Correct or combine? Mechanically integrating judgemental forecast with statistical methods", *International Journal of Forecasting*, Vol. 16, 261-275.
- GOODWIN, P. y LAWTON, R. (1999): "On the asymmetry of the symmetric MAPE", *International Journal of Forecasting*, Vol. 15, 405-408.
- GOULD, J.P. (1986): "Is the rational expectations hypothesis enough?", *Journal of Business*, Vol. 59, nº 4, 371-377.
- GOURIEROUX, C. y PRADEL, J. (1986): "Direct test of the rational expectation hipótesis", *European Economic Review*, Vol. 30, nº 2, 265-284.
- GRANGER, C.W.J y NEWBOLD, P. (1980): *Forecasting Economic Time Series*, Academic Press, New York.
- GRANGER, C.W.J y RAMANATHAN, C. (1984): "Improved Methods of Combining Forecast", *Journal of Forecasting*, Vol. 3, 197-204.
- GRANGER, C.W.J. (1981): "Some properties of time series data and their use in econometric model specification", *Journal of Econometrics*, Vol. 16, 121-130.
- GRANGER, C.W.J. y NEWBOLD, P. (1975): "Economic Forecasting: The Atheist's Viewpoint", en *Modelling the Economy*, Ed. G.A. Renton, London. 131-147.
- GRANGER, C.W.J. y PESARAN, M.H. (1996): *A decision theoretic approach to forecast evaluations*, Manuscript, Departments of Economics, University of California, San Diego and Cambridge University. <http://citeseer.nj.nec.com/granger96decision.html>
- GRANGER, C.W.J. y PESARAN, M.H. (1999): *Economic and statistical measures of forecast accuracy*, Department of Applied Economics, University of Cambridge, working paper in economics nº 9910.
- GRANGER, C.W.J. y SWANSON, N. (1995): *Further developments in the study of cointegrated variables*, UCSD Working Papers.
- GRANGER, C.W.J. (1969): "Prediction with a Generalized Cost of Error Function", *Operational Research Quarterly*, Vol. 20, nº 2, 199-207.
- GREENE, W. H. (1999): *Análisis econométrico*. Tercera edición, Prentice Hall Iberia.
- GREER, M. (2003): "Directional accuracy test of long-term interest rate forecast", *International Journal of Forecasting*, Vol. 19, nº 2, 291-300.
- GREGORY, A.W. (1987): "Testing interest rate parity and rational expectations for Canada and the United States", *Canadian Journal of Economics*, nº 2, 289-305.
- GROFF, G.K. (1973): "Empirical comparison of models for short range forecasting", *Management Forecasting*, Vol. 20, nº 1, 22-31.
- GUERRERO, V.M. y PEÑA, D. (2000): "Linear combination of restrictions and forecast in time series analysis", *Journal of Forecasting*, Vol. 19, 103-122.
- GUNTER, S.I. y AKSU, C. (1997): "The usefulness of Heuristic N(E)RLS Algorithms for combining forecast", *Journal of Forecasting*, Vol. 16, 439-463.
- GUPTA, S. y WILTON, P.C. (1987): "Combination of forecasts: An extension", *Management Science*, Vol. 33, nº 3, 356-372.
- GUPTA, S. Y WILTON, P.C. (1988): "Combination of Economic Forecasts: An Odds-Matrix approach", *Journal of Business and Economic Statistics*, Vol. 8, nº 3, 373-379.

- HAFER, R.W. y HEIN, S.E. (1985): "On the accuracy of time-series, interest rate and survey forecasts of inflation", *Journal of Business*, Vol. 58, nº 4, 377-398.
- HALL, S. (1994): "Time series forecasting", en *Applied Economic Forecasting Techniques*, Stephen Hall, London.
- HALLMAN, J. Y KAMSTRA, M. (1989): "Combining Algorithms based on robust estimation techniques and Co-integration restrictions", *Journal of Forecasting*, Vol. 8, 189-198.
- HALTIWAGNER, J. y WALDMAN, M. (1985): "Rational expectations and the limits of rationality: An analysis of heterogeneity", *The American Economic Review*, Vol. 75, nº 3, 326-340.
- HARRIES, C. (1999): "Judgmental inputs to the forecasting process: research and practice", *Technological Forecasting and Social Change*, nº 61, 75-82.
- HARRISON, P. y STEVENS, H. (1971): "A Bayesian approach to Short-term forecasting", *Operational Research Quarterly*, Vol. 22, nº 4, 341-363.
- HARRISON, P. y STEVENS, H. (1976): "Bayesian forecasting", *Journal of the Royal Statistical Society*, A139, 205-228.
- HARVEY, D., LEYBOURNE, S. Y NEWBOLD, P. (2001): "Analysis of a panel of UK macroeconomic forecasts.", *Econometrics Journal*, Vol. 4, 37-55.
- HARVEY, D.I. y NEWBOLD, P. (2003): "The non-normality of some macroeconomic forecast errors", *International Journal of Forecasting*, Vol. 19, nº 4, 635-654.
- HAVRDA, J. y CHARVAT, F. (1967): "Quantification method of classification processes", *Kybernetika*, nº 3, 30-35.
- HEDGES, L.V. Y OLKIN, I. (1985): *Statistical Methods for Meta-Analysis*, Academic Press.
- HENDRY, D.F. Y MIZON, G. E. (1978): "Serial correlation as a convenient simplification, nor a nuisance: A comment of the demand for money by the Bank of England", *Economic Journal*, nº 88, 549-563.
- HENDRY, D.F. y CLEMENTS, M.P. (2002): "Pooling Forecasts", *Econometrics Journal*, Vol. 5, 1-26.
- HENDRY, D.J. (1978): "Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom", *The Economic Journal*, nº 88, 661-692.
- HENRIKSSON, R.D. y MERTON, R.C. (1981): "On market timing and investment performance 2: statistical procedures for evaluating forecasting skills", *Journal of Business*, nº 54, 513-553.
- HENRIKSSON, R.D. y MERTON, R.C. (1981): "On market timing and investment performance 2: statistical procedures for evaluating forecasting skills", *Journal of Business*, nº 54, 513-553.
- HERREA, G. (2001): "La Encuesta de Coyuntura Industrial, fuente adelantada de información sobre la evolución de la industria", *Revista Fuentes Estadísticas*, nº 60. [www.fuentesestadisticas.com](http://www.fuentesestadisticas.com).
- HERRERO, A. (1996): "Encuesta de Coyuntura Industrial", *Revista Fuentes Estadísticas*, Notas Metodológicas, Indicadores Económicos de Oferta, Encuesta de Coyuntura Industrial. [www.fuentesestadisticas.com](http://www.fuentesestadisticas.com).
- HEYDE, C.C. y JOHNSTONE, I.M. (1979): "On Asymptotic posterior normality of stochastic processes", *Journal of the Royal Statistical Society*, Serie B, Vol. 41, (2), 184-189.
- HIBON, M. (1984). "Naive, Moving Average, Exponential Smoothing, and Regression Methods", *The Forecasting Accuracy of Major Time Series Methods*, John Wiley & Sons, 239-245.

- HILLIER, B. (1987): "Forecasting and policy evaluation in economies with rational expectations: The discrete time case-A comment", *Bulletin of Economic Research* Vol. 39, nº 1, 71-78.
- HOETING, J. A.; MADIGAN, D.; REFTERY, A.E. y VOLINSKY, C. T. (1999): "Bayesian model averaging: A tutorial", *Statistical Science*, Vol 14, nº 4, 331-363.
- HOGARTH, R. M (1975): "Cognitive processes and the assessment of subjective probability distributions", *Journal of the American Statistical Association*, Vol. 70, nº 3, 271-294
- HOLDEN, K. y PEEL, D.A. (1989): "Unbiasedness, efficiency and the Combination of Forecasts", *Journal of Forecasting*, Vol. 8, 175-188.
- HOLDEN, K. y PEEL, D.A. (1990): "On testing for unbiasedness and efficiency of forecast", *The Manchester School*, Vol. LVIII, nº 2, 120-127.
- HUGHES HALLET, A.J. (1987): "Forecasting and policy evaluation in economies with rational expectations: The discrete time case", *Bulletin of Economic Research*, Vol. 39, nº 1, 49-70.
- INE (2001): *Revisión del sistema de indicadores cíclicos, Nota Metodológica*, S.G. Cuentas Nacionales.
- INE (Varios años): *Boletín Mensual de Estadística*, [www.ine.es](http://www.ine.es)
- IVALDI, M. (1992): "Survey evidence on rationality of expectations", *Journal of Applied Econometrics*, Vol. 7, 225-241.
- JACOBSON, T. y KARLSSON, S. (2002): *Finding good predictors for inflation: A bayesian model averaging approach*, Sveriges Riskbank Working Paper Series, nº 138, Central Bank of Sweden.
- JEONG, J. y MADDALA, G.S. (1996): "Testing the rationality of survey data using the weighed double-bootstrapped method of moments", *The Review of Economic and Statistics*, 296-302.
- JEX, C. F. (1994): "Recursive estimation as an aid to exploratory data analysis: an application to market share models", *International Journal of Forecasting*, nº 10, 445-453.
- JORRAT, J.M. y CERRO, A.M. (2000): "Computing turning point monthly probability of the Argentine economy according to the leading index: 1970-2000", *Estudios de Economía*, Vol. 27, nº 2, 279-295.
- JOUNI, M.H. y CLEMEN, R.T. (1996): "Copula models for aggregating expert opinions", *Operations Research*, Vol. 44, nº 3, 444-457.
- JUDD, K.L. (1998): *Numerical Methods in Economics*, Massachusetts Institute of Technology.
- JUDGE, G.G. (1978): *The Statistical Implications of Pre-test and Stein-Rule Estimators in Econometrics*, North-Holland Publishing Company.
- JUDGE, G.G. y BOCK, M. E. (1978): *The Statistical Implications of Pre-test and Stein-Rule Estimators in Econometrics*, North-Holland Publishing Company.
- KADANE, J.B. y DICKEY, J.M. (1980): "Bayesian decision theory and the simplification of models" en *Evaluation of Econometric models*, Kmenta, J. y Ramsey, J.B. editores, Academic Press, 245-268.
- KALMAN, R. (1960): "A New Approach to Linear Filtering and Prediction Problems", *Journal of Basic Engineering, Transactions ASME, Series D*, nº 82, 35-45.
- KAMSTRA, M. y KENNEDY, P. (1998): "Combining qualitative forecast using logit", *International Journal of Forecasting*, 14, 83-93.
- KANG, H. (1986): "Unstable weights in the combination of forecast", *Management Science*, Vol. 32, nº 6, 683-695.

- KANO, S. y LI, Z.D. (1990): "A method of exploring the mechanism of inflationary expectations based on qualitative survey data", *Journal of Business and Economic Statistics*, Vol. 8, nº 4, 395-403.
- KAUFFMAN Y GROBOLLIOT (1967): series temporales, alisado
- KAUFMANN, A. y GIL, J. (1990): *Las matemáticas del azar y de la incertidumbre*. Editorial Centro de Estudios Ramón Areces, Madrid.
- KAWASAKI, S. y ZIMMERMANN, K.F. (1986): "Testing the rationality of price expectations for manufacturing firms", *Applied Economics*, nº 18, 1335-1347.
- KIRBY, R. (1966): "A comparison of short and medium range statistical forecasting methods", *Management Science*, Vol. 13, nº 4, 202-210.
- KLAPPER, D. y HERWARTZ, H. (2000): "Forecasting market share using predicted values of competitive behaviour: further empirical results", *International Journal of Forecasting*, Vol. 16, 399-421.
- KLING, J.L. y BESSLER, D.A. (1989): "Calibration-based predictive distributions: An application of prequential analysis to interest rates, money, prices and output", *Journal of Business*, Vol. 62, nº4, 477-499.
- KNÖBL, A. (1974), "Price expectations and actual price behaviour in Germany", *International Monetary Fund Staff Papers*, nº 21, 83-100.
- KOENKER, R. y BASSET G. Jr. (1978): "Regression quantiles", *Econometrica*, Vol. 46, nº 1, 33-50.
- KOENKER, R. Y BASSET G. Jr. (1982): "Robust test for heterocedasticity based regression quantiles", *Econometrica*, nº 50, 43-62.
- KUAN, C.M. y WHITE, H. (1994): "Artificial Neural Networks: An Econometric Approach", *Econometric Reviews*, 13 (1), 1-91.
- LABORDA, A. (2001): "Panel FUNCAS de previsiones económicas para España", *Revista Fuentes Estadísticas*, nº 55. [www.fuentesestadisticas.com](http://www.fuentesestadisticas.com).
- LACKMAN, C. y BRANDON, C. (1994): "Consensus forecasting using relative error weights", *Marketing Intelligence & Planning*, Vol. 12, nº 1, 37-41.
- LAHIRI, K., ZAPOROWSKI, M. (1987): "More flexible use of survey data on expectations in macroeconomic models", *Journal of Business and Economic Statistics*, Vol. 5, nº 4, 69-76.
- LAMONT, O.W. (2002): "Macroeconomic forecasts and microeconomic forecasters", *Journal of Economic Behavior and Organization*, Vol. 48, 165-280.
- LANDAJO, M. (1999): *Modelos neuroborrosos para la predicción económica*. Tesis Doctoral, Universidad de Oviedo.
- LANDAJO, M. (2000): "Neural and Fuzzy Models for Economic Forecasting. An Econometric View and Some Practical Experience", *Fuzzy Economic Review*, Vol. 5, nº 1, 3-28.
- LAPEDES, A. y FARBER, R. (1987): "Nonlinear signal processing using neural networks", *Technical Report LA-UR-87-2662*, Los Alamos National Laboratory, Los Alamos, N.M.
- LATHA, E.M.M. (1998): "Direct measures of inflationary expectations: A review of selected surveys and related problems", *The Indian Economic Journal*, Vol. 45, nº 1, 80-92.
- LAWRENCE, M.J., EDMUNDSON, R.H. y O'CONNOR, M.J. (1986): "The accuracy of combining judgemental and statistical forecasts", *Management Science*, Vol. 32, nº 12, 1521-1532.
- LAYTON, A. P. (1998). "A further test of the influence of leading indicators on the probability of US business cycle phase shifts", *International Journal of Forecasting*, Vol. 14, 67-70.

- LEE, K.C. (1994): "Formation of price and cost inflation expectations in British manufacturing industries: A multi-sectorial analysis", *The Economic Journal*, nº 104, 372-385.
- LEWANDOWSKI, R. (1982): "Sales Forecasting by FORSYS", *Journal of Forecasting*, Vol. 1, 205-214.
- LEWANDOWSKI, R. (1984): "Lewandowski's FORSYS Method", en *The forecasting Accuracy of Major Time Series Methods*, John Wiley & Sons, 245-253.
- LI, F. y TKACZ, G. (2001): *Evaluating linear and non-linear time-varying forecast-combination methods*, Bank of Canada Working Paper 2001-12.
- LIGHT, R.J. (1971): "Measures of response agreement for qualitative data: some generalization and alternatives", *Psychological Bulletin*, nº 76, 365-377.
- LIM, G.C. y MCKENZIE, C.R. (1998): "Testing the rationality of expectations in the Australian foreign exchange market using survey data with missing observations", *Applied Financial Economics*, Vol. 8, 181-190.
- LINDLEY, D. V. (1985): "Reconciliation of discrete probability distributions" en *Bayesian Statistics 2*, J.M. Bernardo, M.H. DeGroot, D.V. Lindley y A.F.M. Smith editores, North-Holland, 375-390.
- LINDLEY, D. V.; TVERSKY, A. y BROWN, R. V. (1979): "On the reconciliation of probability assessments", *Journal of the Royal Statistical Society*, serie A, nº 142, part 2, 146-180.
- LINDLEY, D.V. (1956): "On a measure of the information provided by an experiment" *Annals of Mathematical Statistics*, Vol. 27, nº 4, 986-1005.
- LINDLEY, D.V. (1964): "The Bayesian analysis of contingency tables", *Annals of Mathematical Statistics*, Vol. 35, nº 4, 1622-1643.
- LINDLEY, D.V. y SINGPURWALLA, N.D. (1986): "Reliability and fault tree analysis using expert opinions", *Journal of American Statistical Association* (Theory and Methods), Vol. 81, nº 393, 87-90.
- LINDLEY, D.V. y SMITH, A.F.M. (1972): "Bayes estimates for the linear model", *Journal of the Royal Statistical Society*, Serie B, Vol. 34, (1), 1-41.
- LINDSTRÖM, T. (2000): *Qualitative Survey Responses and production over the business cycle*, Working Paper Series nº 116, Sveriges Riksbank, Central Bank of Sweden, <http://swopec.hhs.se/rbnkwp/abs/rbnkwp0116.htm>
- LJUNG, G. Y G. BOX (1979): "On a Measure of Lack of Fit in Time Series Models", *Biometrika*, nº 66, 265-270.
- LÖFFER, G. (1999): "Refining the Carlson-Parkin method", *Economics Letter*, nº 64, 167-171.
- LOPEZ SERRANO (2003): *Cambios recientes en la Encuesta de Coyuntura Industrial*, Documento interno Secretaria General Técnica MCYT. Texto actualizado respecto al publicado en *Economía Industrial*, nº 335-336/2000v.
- LÓPEZ, A.J. y MORENO, B. (1999): "Evaluación de predicciones basada en medidas de información. Nuevas Alternativas", *Anales de Economía Aplicada XIII Reunión Asepelt-España*, Burgos.
- LOPEZ, A.J., MORENO, B. y PÉREZ, R. (2003): "Forecasting Evaluation Based on Information measures", *Actas 54<sup>th</sup> session International Statistical Institute*, Berlin.
- LÓPEZ, A.J.; MUÑOZ, N. y PÉREZ, R. (1993): Base de datos Asturdat. *Documento de trabajo 2/93, Hispalink-Asturias*.
- LOVELL, M.C. (1986): "Test the rational expectations hypothesis", *The American Economic Review*, Vol. 76, nº 1, 110-124.

- LUBECKE, T.H., NAM, K.D., MARKLAND, R.E. y KWOK, C.C.Y. (1998): "Combining foreign exchange rate forecasts using neural networks", *Global Finance Journal*, 9 (1), 5-27.
- MAKRIDAKIS, S. (1984): "Forecasting: State of the Art", en *The Forecasting Accuracy of Major Time Series Methods*, John Wiley & Sons, 1-17.
- MAKRIDAKIS, S. y HIBON, M. (1979): "Accuracy of Forecasting: An Empirical Investigation (with Discussion)", *J. R. Statistical Society*, Series A, 142, Part. 2, 97-145.
- MAKRIDAKIS, S. y WHEELWRIGHT, S. (1978): *Forecasting: Methods & Applications*, John Wiley & Sons.
- MAKRIDAKIS, S. Y WINKLER, R.L. (1983): "Averages of forecasts: Some empirical results", *Management Science*, Vol 29, nº 9, 987-998.
- MAKRIDAKIS, S., ANDERSEN, A., CARBONE, R., FILDES, R., HIBON, M., LEWANDOWSKY, R., NEWTON, J., PARZEN, E. y WINKLER, R. (1982): "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition", *Journal of Forecasting*, Vol. 1, 111-153.
- MARCELLINO, M. (2002): *Forecasting pooling for short time series of macroeconomic variables*, Working Paper, nº 212, IEP-Bocconi University, IGER y CEPR, 1-31.
- MARITZ, J.S. (1995): *Distribution-free Statistical Methods, Monographs on Statistics and Applied Probability, 17*, Chapman & Hall.
- MARTINEZ, P. (2001): "Datos de opinión acerca de cuestiones económicas", *Revista Fuentes Estadísticas*, nº 60. [www.fuentesestadisticas.com](http://www.fuentesestadisticas.com).
- MARTZ, H.F. y BRYSON, M.C. (1984): "A statistical model for combining biased expert opinions", *IEEE Transactions on Reliability*, Vol. R-33, nº 3. 227-232
- MASCAREHNAS, B. y SAND, O.C. (1989): "Combination of forecasts in the international context: predicting debt reschedulings", *Journal of International Business Studies*, Fall, 539-552.
- MATTHEWS, K. (1985): "Forecasting with a rational expectations model of the UK", *Oxford Bulletin of Economics and Statistics*, Vol. 47, nº 4, 311-336.
- McKENZIE, C.R. y McALLER, M. (1994): "On the effects of misspecification errors in models with generated regressors", *Oxford Bulletin of Economics and Statistics*, Vol. 56, nº 4, 441-455.
- McNESS, S. K. (1991): "Forecasting cyclical turning points: The record in the past three recessions", en *Leading economic indicators: New approaches and forecasting records*, Lahiri, K. y Moore G. H. (eds), Cambridge University Press, 151-168.
- MEADE, N. y ISLAM, T. (1998): "Technological forecasting: Model selection, model stability, and combining models", *Management Science*, Vol. 44, nº 8. 1115-1130.
- MEADE, N. Y ISLAM, T. (1998): "Technological forecasting--model selection, model stability, an combining models", *Management Science*, Vol. 44, nº 8, 1115-1130.
- MENEZES, L.M. y BUNN, D.W. (1998): "The persistence of specification problems in the distribution of combined forecast errors", *International Journal of Forecast*, Vol. 14, 415-426.
- MENEZES, L.M., BUNN, D.W. y TAYLOR, J.W. (2000): "Review of guidelines for the use of combined forecasts", *European Journal of Operational Research*, nº 120, 190-204.

- MERTON, R.C. (1981): "On market timing and investment performance 1: an equilibrium theory of value for market forecasts", *Journal of Business*, nº 54, 363-406.
- MIN, C. y ZELLNER, A. (1993): "Bayesian and non-Bayesian methods for combining models and forecasts with applications to forecasting international growth rates", *Journal of Econometrics*, nº 56, 89-118.
- MINISTERIO DE CIENCIA Y TECNOLOGÍA (Varios años): *Encuesta de Coyuntura Industrial*. Secretaría General del Ministerio de Ciencia y Tecnología, [www.mcyt.es](http://www.mcyt.es)
- MITCHELL, J. (2002): "The use of non-normal distributions in quantifying qualitative survey data on expectations", *Economic Letters*, nº 76, 101-107.
- MOORE, G. H. (1980): *Business Cycles, Inflation and Forecasting*, National Bureau of economic Research Studies in Business Cycles, nº 24.
- MORENO, B. (1999): *Técnicas de predicción en series temporales y medidas de evaluación*, Memoria de Proyecto de Investigación. Departamento de Economía Aplicada, Universidad de Oviedo.
- MORENO, B. y LÓPEZ, A. J. (2002). "El tratamiento de la variabilidad en las predicciones sobre crecimiento económico", *Actas V Encuentro de Economía Aplicada*, Oviedo.
- MORENO, B. y LÓPEZ, A. J. (2002): "Evaluating the suitability of qualitative leading indicators as economic forecasting tools", *Proceedings of the II International Meeting on Economic Cycles*, Madrid.
- MORENO, B. y LOPEZ, A.J. (2001): "La combinación de predicciones subjetivas. Aplicación a las encuestas de opiniones empresariales", *Anales de Economía Aplicada, XV Reunión ASEPELT-España*, La Coruña.
- MORENO, B. y LOPEZ, A.J. (2002): "El tratamiento de la variabilidad en las predicciones sobre crecimiento económico", *Actas V Encuentro de Economía Aplicada*, Oviedo.
- MORENO, B. y LOPEZ, A.J. (2002): "Evaluación de las predicciones sobre crecimiento económico", *Anales de Economía Aplicada, XVI Reunión ASEPELT-España*, Madrid.
- MORENO, B., LÓPEZ, A.J. y LANDAJO, M (2000): "Comparación y combinación de predicciones. Aplicación a las series temporales", *Anales de Economía Aplicada XIV Reunión ASEPELT-España*, Oviedo.
- MORIARTY, M.M. (1990): "Boundary value models for the combination of forecasts", *Journal of Marketing Research*, Vol. XXVII, 402-417.
- MORRIS, P.A. (1974): "Decision analysis expert use", *Management Science*, Vol. 20, nº 9, 1233-1241.
- MORRIS, P.A. (1977): "Combining expert judgments: A Bayesian approach", *Management Science*, Vol. 23, nº 7, 679-693.
- MOSTAGHIMI, M. (1996): "combininig ranked mean value forecasts", *European Journal of Operational Research*, nº 94, 505-516.
- MUÑOZ, N. (1993): "Técnicas cuantitativas de predicción", Resumen de?
- MURPHY, A.H. (1973): "A new vector partition of the probability score", *Journal of Applied Meteorology*, Vol. 12, 595-600
- MURPHY, A.H. (1988): "Skill Scores based on the mean square error and the relationships to the correlation coefficient", *Monthly Weather Review*, Vol. 116, 2417-2424.
- MURPHY, A.H. y WRINKLER, R.L. (1987): A general framework for forecast verification, *Monthly Weather Review*, Vol. 115, 1330-1338.

- MURPHY, A.H. y WRINKLER, R.L. (1992): "Diagnostic verification of probability forecasts", *International Journal of Forecasting*, Vol. 7, 435-455.
- NEFTCI, S. (1982): "Optimal prediction of cyclical downturns", *Journal of Economic Dynamics and Control*, nº 4, 225-241.
- NELSON, R.G. y BESSLER, D.A. (1989): "Subjective probabilities and scoring rules: Experimental Evidence", *American Journal of Agricultural Economics*, 363-369.
- NERLOVE, M. (1983): "Expectations, plans and realisations in theory and practice", *Econometrica*, Vol. 51, 1251-1279
- NERLOVE, M., ROSS, D. y WILSON, D. (1993): "The importance of seasonality in inventory models. Evidence from business survey data", *Journal of Econometrics*, nº 55, 105-128.
- NEWBOLD, P. Y GRANGER, C.W.J. (1974): "Experience with Forecasting Univariate Time Series and the Combination of Forecast", *Journal of the Royal Statistical Society, Serie A*, 137. Part. 2, 131-165.
- NEWTON, H.J. y PARZEN, E. (1984): "Forecasting and Time Series Model Types of 111 Economic Time Series", en *The Forecasting Accuracy of Major Time Series Methods*, John Wiley & Sons, 267-287.
- NICKELL, S. (1980): *Error correction, partial adjustment and all that: an expository note*, Working paper nº255, London School of Economics, 1-8.
- NIEMIRA, M. P. (1991): "An international application of Neftci's probability approach for signaling growth recessions and recoveries using turning point indicators", en *Leading economic indicators: New approaches and forecasting records*, Lahiri, K. y Moore G. H. (eds), Cambridge University Press, 91-108.
- NILSON, R. (2000): "Confidence indicators and composite indicators", *Actas CIRET conference*, Paris, 1-31.
- NOVELL, J. y SOLE, A (2000): "Racionalidad de las previsiones presupuestarias y comportamiento estratégico. El caso español", *Actas III Encuentro de Economía Aplicada*, Valencia.
- O'CONNOR, M., REMUS, W., GRIGGS, K. (2001): "The asymmetry of judgemental confidence intervals in time series forecasting", *International Journal of forecasting*, Vol. 17, 623-633.
- O'CONNOR, M.; REMUS, W. y GRIGGS, K. (2000): "Does updating judgmental forecast improve forecast accuracy?", *International Journal of Forecasting*, Vol. 16, 101-109.
- OECD (1997): "Development of business and consumer surveys in Central and Eastern Europe. Summary of Workshops 1991-1996", Transition Economies Division, Statistics Directorate, Paris, 1-108.
- OECD (1997): *Cyclical indicators and business tendency surveys*, General Distributions.
- ORTEGA, E. (1998): *The Spanish business cycle and its relationship to Europe*, Documento de Trabajo nº 9819, Servicio de Estudios del Banco de España.
- OSTERBERG, W.P. (2000): "New results on the rationality of survey measures of exchange-rate expectations", *Economic Review*, nº 14, 14-21.
- OTERO, J. M. (1978): *Lógica y limitaciones de la econometría*. Ediciones ICE, Madrid.
- OTERO, J.M. (1989): *Modelos econométricos y predicción de series temporales*, Ed. AC, Madrid.
- OTERO, J.M. (1994): *La predicción en Economía: Perspectiva histórica y tendencias actuales*. Lección Inaugural, Universidad de Málaga.
- PALM, F.C. y ZELLNER, A. (1992): "To combine or not to combine? Issues of combining forecast", *Journal of Forecasting*, Vol. 11, 687-701.

- PEDERSEN, T.M. y ELMER, A.M. (2000): *Technical Appendix. Determination of Business Cycle Turning Points with Computer Algorithm*, Working Paper, Denmark
- PEEL, D.A. (1989): "On testing the properties of directly obtained expectations data", *Economic Letters*, nº 30, 137-139.
- PEÑA, D. (1997): "Combining information in statistical modeling", *The American Statistician*, Vol. 54, nº 4, 326-332.
- PÉREZ, M.A. (2001): "La opinión de los empresarios ante diversas expectativas de futuro", *Revista Fuentes Estadísticas*, nº 60, [www.fuentesestadisticas.com](http://www.fuentesestadisticas.com).
- PÉREZ, R. (1985): *Estimación de la incertidumbre, la incertidumbre útil y la inquietud en poblaciones finitas. Una aplicación a las medidas de desigualdad*, Tesis Doctoral, Universidad de Oviedo.
- PÉREZ, R. y LÓPEZ, A. J. (1997): *Análisis de datos económicos II. Métodos inferenciales*. Ed. Pirámide, Madrid.
- PÉREZ, R. y otros (1993): *Análisis de datos económicos I. Métodos descriptivos*. Ed. Pirámide, Madrid.
- PESARAN, M. H. (1985), "Formation of inflation expectations in British manufacturing industries", *Economic Journal*, nº 95, 948-975.
- PESARAN, M. H. (1987), *The Limits to Rational Expectations*, Basil Blackwell, Inc, Oxford.
- PESARAN, M.H. (1984): "Expectation formations and macroeconomic modelling", en *Contemporary Macroeconomic Modelling*, Malgrange P. y Muet, P-A, eds, , Basil, Blackwell, Oxford, 27-55.
- PESARAN, M.H. (1985): "Formation of inflation expectations in british manufacturing industries", *The Economic Journal*, nº 95, 948-975.
- PESARAN, M.H. y SKOURAS,S. (2000): "Decison-based methods for forecast evaluation", en *A Companion to Economic Forecasting*, M.P. Clements and D.F. Hendry (eds.), Oxford: Basil Blackwell, chapter 11, 241-267.
- PESARAN, M.H. Y TIMMERMANN, A. (1992): "A simple nonparametric test of predictive performance", *Journal of Business & Economic Statistics*, Vol. 10, nº 4, 461-465.
- PESARAN, M.H. y TIMMERMANN, A. (1992): "A simple nonparametric test of predictive performance", *Journal of Business & Economic Statistics*, Vol. 10, nº 4, 461-465.
- PICHERY, M.C. (1989): *Agreement and disagreement between expectations and realizations*, Document l'IME, N° 9001, September.
- POIRIER, D.J. (1995): *Intermediate statistics and econometrics: a comparative approach*. Massachusetts Institute of Technology Press.
- POLE, A.; WEST, M. Y HARRISON, J. (1994): *Applied bayesian forecasting and time series Analysis*, Chapman & Hall.
- POLLOCK, A.C., MACAULAY, A., ÖNKAL-ATAY, D. y WILKIE-THOMSON, M.E. (1999): "Evaluating predictive performance of judgemental extrapolations from simulated currency series", *European Journal of Operational Research*, nº 114, 281-293.
- PONS, J. (1999): "La fiabilidad de las predicciones económicas", *Anales de Economía Aplicada XIII Reunión Asepelt-España*, Burgos.
- PUCH, L. y LICANDRO, O. (1995): *Teoría de los ciclos reales y fluctuaciones agregadas de la economía española*, Documento de trabajo 95-24, FEDEA y Universidad Carlos III de Madrid.

- PUGH, L. y LICANDRO, O. (1997): *Are they any special features in the Spanish business cycle?*, Documento de trabajo 97-06, FEDEA y Universidad Carlos III de Madrid.
- PULIDO, A (1998): *Una apuesta por el futuro. Predicciones y profecías económicas*. Ed. Pirámide, Madrid.
- PULIDO, A. (1987): *Modelos econométricos*, Ed. Pirámide, Madrid. (4ª edición 1993)
- PULIDO, A. (1989): *Predicción económica y empresarial*, Ed. Pirámide, Madrid.
- PULIDO, A. (2001a): “Fuentes internacionales y nacionales en predicción económica”, *Revista Fuentes Estadísticas*, nº 55. [www.fuentesestadisticas.com](http://www.fuentesestadisticas.com).
- PULIDO, A. (2001b): “Consenso de predicciones”, *Revista Fuentes Estadísticas*, nº 60. [www.fuentesestadisticas.com](http://www.fuentesestadisticas.com).
- PULIDO, A. y LÓPEZ, A.M. (1999): *Predicción y Simulación aplicada a la economía y gestión de empresas*, Ed. Pirámide, Madrid.
- PULIDO, A (1999): *Predicción económica: Técnicas y aplicaciones*, Curso del Doctorado en Economía Aplicada, Universidad de Oviedo
- QUANTITATIVE MICRO SOFTWARE (1998): *EViews User's Guide, Command and Programming Reference*.
- RAIFFA, H y SCHLAFER, R. (1961): *Applied Statistical Decision Theory*, The M.I.T. press.
- RAMANATHAN, R. (1998): *Introductory econometrics*, Harcourt Brace College Publishers, Orlando, 4ª edición.
- RAMSEY, F. P. (1926): “Truth and Probability” en *The Foundation of Mathematics and Other Logical Essays*, Routledge y Kegan P. editores, London, 156-198.
- RAO, C.R. (1965): *Linear Statistical Inference and its Applications*, John Wiley & Sons, Inc.
- RASVIRTA, T.; TJOSTHEIM, D. y GRANGER, C.W. (1994): “Aspects of modelling nonlinear time series”, en *Handbook of Econometrics*, R.F. Engle & D.L. McFadden (eds.), vol. IV, cap. 48, 2917-2957. Elsevier Science, New York.
- REEVES, G.R. y LAWRENCE, K.D. (1983): “Combining multiple forecast given multiple objectives”, *Journal of Forecasting*, Vol. 1, 271-280.
- RIOS, S. (1976): *Análisis de decisiones*, Ediciones ICE, Madrid.
- ROBERT, C. P. (1994): *The Bayesian Choice: A decision-Theoretic motivation*, Springer, Berlin.
- ROBERTSON D. y WICKENS, M. (1994): “VAR modelling”, en *Applied Economic Forecasting Techniques*, Stephen Hall, London, 24-47.
- ROWE, G. Y WRIGHT, G, (1999): “ The Delphi technique as forecasting tool: Issues and analysis”, *International Journal of Forecasting*, Vol. 15, 353-375.
- SALMON, M. (1982): “Error correction mechanisms”, *The Economic Journal*, Vol. 92, 615-629.
- SAVAGE L. J. (1971): “Elicitacion of personal probabilities and Expectations”, *Journal of the American Statistical Association*, vol 66, nº 336. 783-801.
- SCHNADER, M.H. Y STEKLER, H.O.(1990): "Evaluating predictions of change", *Journal of Business*, Vol. 63, nº 1, 99-107.
- SCHÖNBORN, F. (1997): “Las encuestas cualitativas reflejan los planes y expectativas de empresarios y consumidores”, *Revista Fuentes Estadísticas*, nº 55. [www.fuentesestadisticas.com](http://www.fuentesestadisticas.com).
- SCHÖNBORN, F. (1997): “Las encuestas cualitativas reflejan los planes y expectativas de empresarios y consumidores”, *Revista Fuentes Estadísticas*, nº 55, [www.fuentesestadisticas.com](http://www.fuentesestadisticas.com).

- SEITZ, H. (1988), "The estimation of inflation forecasts from business survey data", *Applied Economics*, Vol. 20, 427-438.
- SENGUPTA, J.K. (1993): "Information theory and economic growth models", *International Journal Systems Science*, Vol. 24, nº 2, 403-418.
- SESSIONS, D.N, Y CHATTERJEE (1989): "The combining of Forecast using Recursive techniques with nonstationary weights", *Journal of Forecasting*, nº 8, 239-251.
- SHANNON, C.E. (1948): "A Mathematical Theory of Communication", *Bell System Technology Journal*, 27, 379-423.
- SHAW, G.K. (1987): "Rational expectations", *Bulletin of Economic Research*, Vol. 39, nº3, 187-209.
- SILVER, S. J. (1991): "Forecasting peaks and troughs in the business cycle: On the choice and use of appropriate leading indicator series", en *Leading economic indicators: New approaches and forecasting records*, Lahiri, K. y Moore G. H. (eds), Cambridge University Press, 183-195.
- SILVERMAN, B.W. (1986): *Density Estimation for Statistics and Data Analysis*, *Monographs on Statistics and Applied Probability*, 26, Chapman & Hall.
- SIMPSON, P.W., OSBORN, D.R y SENSIER, M. (2001): "Forecasting UK Industrial Production over business cycle", *Journal of Forecasting*, nº 20, 405-424.
- SIMS, C. (1988): "Uncertainty in macroeconomics: Uncertainty across models", *AEA Papers and Proceedings*, Vol. 78, nº2, 163-167.
- SIMS, C.A. (1980): "Macroeconomics and Reality", *Econometrica*, Vol. 48, nº 1, 1-48.
- SISA: Simple Interactive Statistical Analysis INTERNET (tablas 2\*2)
- SMITH, A.F.M. (1983): "Bayesian approaches to outliers and robustness" en *Specifying Statistical Models; Lecture Notes in Statistics*, 16, Florens, J. et al editores, Springer-Verlag, Berlin, 14-35.
- SMITH, J. Y MCALEER, M. (1995), "Alternative procedures for converting qualitative response data to quantitative expectations: an application to Australian manufacturing", *Journal of Applied Econometrics*, Vol. 10, 165-185..
- SPANOS, A. (1999): *Probability theory and statistical inference*, Cambridge University Press.
- STARR-McCLUER, M. (1999): *The measurement of Consumer expectations using survey data*, Finance and Economics Discussion Series, nº1999-17, Board of Governors of the Federal Reserve System (U.S.).
- STECKLER, H.O. (1987): "Who forecasts better?", *Journal of Business and Economic Statistics*, Vol. 5, 155-158.
- STECKLER, H.O. (1991): "Turning point predictions, errors, and forecasting procedures", en *Leading economic indicators: New approaches and forecasting records*, Lahiri, K. y Moore G. H. (eds), Cambridge University Press, 169-181.
- STECKLER, H.O. (1994): "Are economic forecasts valuable?", *Journal of Forecasting*, Vol. 13, 495-505.
- STEKLER, H.O. (1987): "Who forecast better?", *Journal of Business and Economic Statistics*, Vol. 5, nº 1, 155-158.
- STONE, M. (1959): "Application of a measure of information to the design and comparison of regression experiments", *Annals of Mathematical Statistics*, Vol. 30, nº 1, 55-70.
- STONE, M. (1961): "The opinion pool", *Annals of Mathematical Statistics*, nº 32, 1339-42.
- SURIÑACH, J. y otros (1995): *Análisis económico regional, nociones básicas de la teoría de la cointegración*, Ed. Boch, Barcelona.

- SWANSON, N.R. y ZENG, T. (2001): "Choosing among competing econometric forecasts: Regression-based forecast combination using model selection", *Journal of Forecasting*, 20,425-440.
- SYDSAETER, K. y HAMMOND, P. J. (1996): *Matemáticas para el análisis económico* Prentice Hall, Madrid.
- TANNER, M.A. y WONG, W.H. (1987): "The calculation of posterior distributions by data augmentation (with discussion)", *Journal of the American Statistical Association*, Vol. 82, nº 398, 528-550.
- TAY, A.S. y WALLIS, K.F. (2000): "Density Forecasting: A Survey", *Journal of Forecasting*, Vol. 19, 235-254
- TAYLOR, J.W. (1999): "Evaluating volatility and interval forecast", *Journal of Forecasting*, 18, 111-128.
- TAYLOR, J.W. y BUNN, D.W. (1998): "Combining forecast quantiles using quantile regression: Investigating the derived weights, estimator bias and imposing constraints", *Journal of Applied Statistics*, Vol 25, nº2, 193-206.
- TAYLOR, J.W. Y BUNN, D.W. (1999): "Investigating improvements in the accuracy of prediction intervals for combination of forecast: A simulation study." *International Journal of Forecasting*, 15, 325-339.
- TERREGROSSA, S.J. (1999): "Combining analysts' forecasts with causal model forecasts of earnings growth", *Applied Financial Economics*, 9, 143-153.
- TERUI, N. y van DIJK, H.K. (1999): "Combined forecasts from linear and nonlinear time series models", *Econometric Institute Report EI-9949/A*, 1-24.
- THEIL, H. (1952): "On the time shape of economic microvariables and the Munich Business test", *Review of the International Statistical Institute*, 20, 105-120.
- THEIL, H. (1955): "Recent experiences with the Munich business test: An expository article", *Econometrica*, 23, 184-192.
- THEIL, H. (1955): "Who Forecast Best?", *International Economic Papers*, Vol. 5, 194-199.
- THEIL, H. (1961): *Economic Forecast and Policy*, North Holland Publishing (First Edition 1958)
- THEIL, H. (1966): *Applied Economic Forecasting*, North Holland Publishing.
- THEIL, H. (1967): *Economics and Information Theory*, North Holland Publishing Company, Amsterdam.
- THEIL, H. (1971): *Principles of Econometrics*, John Wiley and Sons, New York.
- THEIL, H. y GOLDBERGER, A.S. (1961): "On pure and mixed statistical estimation in economics", *International Economic Review*, 2, 65-78.
- THEIL, H. y SCHOLLES, M. (1967): "Forecast Evaluation Based on a Multiplicative Decomposition of Mean Square Errors", *Econometrica*, 35, 70-88.
- THOMSON, M. (1982): "Some results on the statistical properties of an inequality constrained least squares estimator in a linear model with two regressors", *Journal of Econometrics*, 19, 215-231.
- TIERNEY, L. y KADANE, J.B. (1986): "Accurate approximations for posterior moments and marginal densities", *Journal of the American Statistical Association*, Vol.81, nº 393, 82-86.
- TRABELSI, A. y HILLMER, S.C. (1989): "A benchmarking approach to forecast combination", *Journal of Business and Economic Statistics*, Vol 7, nº3, 353-362.
- TROCONIZ, A. F. (1975): *Probabilidades, estadística, procesos aleatorios, ¿Ed. Grafor?*

- TULJAPURKAR, S. Y BOE, C. (1999): "Validation, probability-weighted priors, and information in stochastic forecast", *International Journal of Forecasting*, 15, 259-271.
- TVERSKY, A. (1974): "Assessing Uncertainty", *Journal of the Royal Statistical Society*, serie B, nº 2, 148-159.
- VERHOEF, P.C. y FRANSES, P.H. (2002): On combining revealed and stated preferences to forecast customer behaviour: Three case studies, *Econometric Institute Report 2002-04*, 1-12.
- VISCO, I. (1984): Price expectations in rising inflation, Elsevier Science Publishers. B.V., Amsterdam.
- VON MISES, R. (1964): Mathematical Theory of probability and statistics, Academic Press INC, London.
- WALLSTEN, T.S. y DIEDERICH, A. (2001): "Understanding pooled subjective probability estimates", *Mathematical Social Sciences*, 41, 1-18.
- WEBB, R. H. (1991): "On predicting the stage of the bussines cycle", en Lahiri, K. y Moore G. H. (eds), *Leading economic indicators: New approaches and forecasting records*, Cambridge University Press, 109-127.
- WELCH, E., BRETSCHEIDER, S. y ROHRBAUG, J. (1998): Accuracy of judgemental extrapolation of time series data: Characteristics, causes, and remediation strategies for forecasting", *International Journal of Forecasting*, Vol. 14, 95-110.
- WEST, K. (2001): "Test for forecast encompassing when forecasts depend on estimated regression parameters", *Journal of Bussines and Economic Statistics*, Vol. 19, nº 1, 29-33.
- WEST, K.D. (1994): *Asymptotic inference about predictive ability*. Working paper 9417, University of Wisconsin-Madison, 1-55.
- WEST, K.D. (1994): *Asymptotic inference about predictive ability: Additional appendix*. Working paper, University of Wisconsin-Madison.
- WEST, M y HARRISON, J. (1989): "Subjetive intervention in formal models", *Journal of Forecasting*, Vol. 8 nº 1, 33-54.
- WHITE, H. (1994): *Estimation, Inference and Specification Analysis*. Cambridge University Press, New York.
- WHITE, M., TANSEY, R., SMITH, M. y BARNETT, T. (1993): "Log-linear modeling in personnel research", *Personnel Psychology*, 46, 667-686.
- WINKLER, R.L. (1981): "Combining probability distributions from dependent information sources", *Management Science*, Vol. 27, nº4, 479-488.
- WINKLER, R.L. (1984): "Combining Forecasts", *The Forecasting Accuracy of Major Time Series Methods*, John Wiley & Sons, 289-295.
- WINKLER, R.L. y CLEMEN, R.T. (1992): "Sensitivity of weights in combining forecasts", *Operations Research*, Vol 40, nº3, 609-614.
- WINTERS, P. (1960): "Forecasting sales by exponentially weighted moving averages", *Management Science*, Vol. 6, 324-342.
- WOLFE, C. y FLORES, B. (1990): "Judgemental adjustment of earnings forecast", *Journal of Forecasting*, Vol. 9, nº4, 389-406.
- WREN-LEWIS, S. (1985): "The cuantification of survey data on expectations", *National Institute Economic Review*, 113, 39-49.
- WRINKLER, R. L. (1981): "Combining probability distribution from dependent information sources", *Management Science*, Vol. 27 nº 4, 479-488.
- WRINKLER, R.L. y MAKRIDAKIS, S. (1983): "The Combination of Forecast", *Journal of the Royal Statistical Society*, Serie A, 146. Part. 2, 150-157.

- YANG, R. y BERGER, J.O. (1998): *A catalog of Noninformative priors*. Working Paper Purdue University.
- ZAMAN, A. (1996): *Statistical foundations for econometric techniques*, Academic Press INC.
- ZARNOWITZ, V. (1985): "Rational Expectations and macroeconomic forecasts", *Journal of Business and Economic Statistics*, Vol 3, nº4, 293-311.
- ZELLNER, A. (1971): *An introduction to Bayesian Inference in econometrics*, John Wiley, New York.
- ZELLNER, A. (1985): "Bayesian econometrics", *Econometrica*, Vol. 53, nº 2, 253-269.
- ZELLNER, A. (1986); "On assessing prior distributions and bayesian regression analysis with g-prior distributions" en *Bayesian Inference and Decision Techniques*, P. Goel y A Zellner editores, 233-243.
- ZELLNER, A. (1989): *Bayesian and non-bayesian methods for combining models and forecast*, Manuscript, H.G.B. Alexander Research Foundation, Universidad de Chicago, I.L.
- ZELLNER, A. (2000): *Bayesian and non-bayesian approaches to scientific modelling and inference in economics and econometrics*, Lectura presentada como invitado en la Universidad Carlos III, Madrid, Junio 2000.
- ZELLNER, A. y PALM, F. (1974): "Time series analysis and simultaneous equation econometric models", *Journal of Econometrics*, nº 2, 17-54.
- ZIMMERMANN, K.F. (1997): "Analysis of Business Surveys", en *Handbook of Applied Econometrics-Microeconomics*, Pesaran, M.H. y Schmidt, P., eds., Blackwell Publishers, Oxford, 407-441.