EVALUACIÓN DE PREDICCIONES BASADA EN MEDIDAS DE INFORMACIÓN. NUEVAS ALTERNATIVAS

López Menéndez, Ana Jesús
Moreno Cuartas, Blanca
Universidad de Oviedo

Resumen

La teoría de la información proporciona un soporte adecuado para la evaluación de modelos econométricos, tanto en lo que se refiere a la bondad del ajuste como a la capacidad predictiva. En este segundo aspecto, el índice de desigualdad de Theil constituye una referencia habitual, debido en gran medida a su fácil interpretación y su propiedad de descomponibilidad.

En este trabajo examinamos nuevas alternativas que, basadas en medidas de información, resulten 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.

El análisis del comportamiento de estos indicadores en varias aplicaciones prácticas, y su comparación con las medidas habitualmente empleadas para evaluar las predicciones económicas, permite extraer algunas conclusiones de interés.

Introducción

La predicción se ha convertido en una práctica imprescindible en el mundo actual, resultando de gran interés disponer de instrumentos capaces de anticipar el comportamiento futuro de una magnitud.

El enorme esfuerzo metodológico realizado en el ámbito de las técnicas de predicción no se ha visto acompañado de nuevas propuestas para la evaluación de resultados.

Así, 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, basado en la comparación entre tasas de variación previstas y efectivas.

En este trabajo exploramos nuevas alternativas para evaluar la bondad de las predicciones económicas partiendo de la teoría de la información. Más concretamente, analizamos varias posibilidades basadas en las medidas de inquietud cuadráticas analizando su comportamiento en distintas aplicaciones prácticas.

Antecedentes. 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 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 $p_1, ..., p_n$ ($\Sigma_i p_i=1$) 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 $p_i = \frac{1}{n} \forall i$.

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 β definidas por J. Havrda y F. Charvat, (1967) cuyo caso particular β=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[1].

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

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

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

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

y la ***inquietud cuadrática*** viene dada por:

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

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.

Evaluación de Predicciones. Medidas habituales

Consideremos una magnitud Y para la cual realizamos predicciones en cierto horizonte temporal de amplitud T. Si denotamos dichas predicciones $\hat{Y}$, el error de predicción vendrá dado en cada período por $e_t = Y_t - \hat{Y}_t$.

Las medidas más directas de la bondad de las predicciones son la raíz del error cuadrático medio (ECM):

$$\sqrt{ECM} = \frac{1}{T}\sqrt{\sum_{t=1}^{T} e_t^2} = \frac{1}{T}\sqrt{\sum_{t=1}^{T}\left(Y_t - \hat{Y}_t\right)^2}$$

y el error absoluto medio (EAM):

[1] En R. Pérez (1985) se definen estas medidas, analizando exhaustivamente sus propiedades y su estimación en poblaciones finitas.

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

Estas medidas presentan varias limitaciones: dependen de las unidades de medida de la variable investigada, no se encuentran acotadas y no tienen en cuenta la dificultad inherente a cada predicción. Precisamente este último inconveniente ha servido de motivación a H. Theil (1966) quien propone la medida de desigualdad:

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

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

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: $\hat{Y}_t = Y_{t-1} \Rightarrow A_t = 0\ \forall t$.[2]

Una ventaja adicional es la *descomponibilidad*, que permite distinguir las contribuciones o pesos de los factores de sesgo, varianza y covarianza:

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

cuya suma es unitaria[3].

---

[2] 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 inicialmente la expresión $U = \frac{\sqrt{\frac{1}{T-1}\sum_t (P_t - A_t)^2}}{\sqrt{\frac{1}{T-1}\sum_t A_t^2} + \sqrt{\frac{1}{T-1}\sum_t P_t^2}}$ 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.

[3] Teniendo en cuenta que el numerador de U^2 puede ser expresado como:

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

se cumple $U^S + U^V + U^C = 1$

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.

Además de la medida de desigualdad U generalmente empleada, H. Theil (1966) 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, 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 $A_i>0$, $\Sigma_i A_i=1$ y consideramos las predicciones $P_1,\dots, P_n$ con $P_i>0$, $\Sigma_i P_i=1$, 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[4]: $I(A:P)=\sum_{i=1}^{n} A_i \log \frac{A_i}{P_i}$, 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.

Dadas las características de esta medida, su aplicación resulta indicada para predicciones input-ouptut u otras situaciones donde el objetivo sea anticipar cuotas o participaciones sobre un total.

[4] 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: $-\sum_{i=1}^{n} A_i \log P_i + \sum_{i=1}^{n} A_i \log A_i = \sum_{i=1}^{n} A_i \log \frac{A_i}{P_i}$.

Nuevas alternativas

En este apartado esbozamos algunas posibilidades para la evaluación de predicciones, adoptando como referencia las medidas cuadráticas de entropía, incertidumbre útil e inquietud.

✓ **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: $A_1,..., A_n$ tales que $A_i>0$, $\Sigma_i A_i=1$ y predicciones $P_1,..., P_n$ con $P_i>0$, $\Sigma_i P_i=1$, podríamos construir las incertidumbres puntuales cuadráticas definidas por R. Pérez (1985)[5] 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:

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

que denominamos *medida cuadrática de imprecisión.*

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

- 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.
- No se garantiza la implicación en sentido contrario, pudiendo obtenerse resultados nulos en situaciones distintas a las predicciones perfectas[6].
- La medida cuadrática de imprecisión no se encuentra acotada inferiormente.

---

[5] 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 ser interpretada como el doble de lo que dista x_i de ser con total seguridad el verdadero valor poblacional.

[6] 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

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[7].

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.

Si nuestro objetivo es llegar a determinar los valores de una magnitud Y a lo largo del período temporal t=1,..., T, la inquietud cuadrática asociada viene dada por la expresión: $H(Y)=\frac{2}{T}\sum_t\left(\frac{E(Y)}{Y_t}-1\right)$[8].

La calidad de un conjunto de predicciones $\hat{Y}_1,...,\hat{Y}_T$ podría entonces ser analizada a partir de la inquietud cuadrática condicionada $H\left(Y/\hat{Y}\right)$, 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 que nos ocupa.

---

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.

[7] 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.

[8] R. Pérez define la medida inquietud cuadrática asociada a un sistema de probabilidades $\{p_1,...,p_n\}$ y utilidades $\{u_1,...,u_n\}$ como $HU^{*2}(p_1,...,p_n,u_1,...,u_n)=2\sum_{i=1}^{n}p_i\left(\frac{E(u)}{u_i}-1\right)$. 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).

✓ **Información cuadrática asociada a las predicciones**

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:

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

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

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

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 $\hat{Y}_1,...,\hat{Y}_T$, las definiciones anteriores no podrían ser aplicadas de forma directa puesto que, al llevar cada realización asociada una predicción, se tendría $E_{\hat{Y}_t}(Y) = Y_t \ \forall t = 1,...,T \Rightarrow H(Y/\hat{Y}_t) = 0 \ \forall t = 1,...,T \Rightarrow H(Y/\hat{Y}) = 0$.

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[9].

Por consiguiente, planteamos una solución intermedia consistente en agrupar la información según los intervalos de predicción, que denotamos por $[\hat{Y}_j]$, 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:

---

[9] La definición sería en este caso $H(Y/\hat{Y}) = 2\int_0^{+\infty}\left(\frac{E_{\hat{y}}(Y)}{y} - 1\right) f(y/\hat{y})dy$ donde $E_{\hat{y}}(Y) = \int_0^{+\infty} y f(y/\hat{y})dy$, obteniéndose la inquietud cuadrática $H(Y/\hat{Y}) = \int_0^{+\infty}\frac{E(Y)}{E_{\hat{y}}(Y)} H(Y/\hat{y}) f(\hat{y})d\hat{y}$.

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

que presenta los siguientes rasgos:

- 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.
- El resultado de la medida viene condicionado por el criterio de agrupación adoptado, aumentando su capacidad descriptiva con la precisión de dicha agrupación.
- 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:*

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

Como puede apreciarse, a diferencia de la información cuadrática introducida en trabajos previos[10] 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 $\left(1-r_{Y,\hat{Y}}\right)<1$ 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 $r_{Y,\hat{Y}}=1\Rightarrow IC\left(Y,\left[\hat{Y}\right]\right)=H(Y)$ con lo cual la información aportada por las predicciones sería del 100% de la inquietud inicial).

[10] 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 $HU^{*2}(A)$.

- Cuando se observan correlaciones negativas entre predicción y realidad, se tiene $\left(1-r_{Y,\hat{Y}}\right)>1$, 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.

✓ 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: $u_{ij}=\frac{\hat{Y}_i}{Y_j}$ (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 $\left[\hat{Y}_i\right],\left[Y_j\right]$, cuyas utilidades conjuntas se obtienen como $u_{ij}=\sum_{t\in\left[\hat{Y}_i\right]x\left[Y_j\right]}\frac{\hat{Y}_t}{Y_t}\frac{1}{T_{ij}}$ con probabilidades $p_{ij}=\frac{T_{ij}}{T}$.

En este contexto, la inquietud condicionada a $\left[\hat{Y}_i\right]$ vendría dada por la expresión[11]:

$$H\left(Y/\left[\hat{Y}_i\right]\right)=2\sum_j\left(\frac{u_{\left[\hat{Y}_i\right]}}{u_{ij}}-1\right)\frac{p_{ij}}{p_i} \text{ , con } u_{\left[\hat{Y}_i\right]}=\sum_j u_{ij}\frac{p_{ij}}{p_i}=\sum_j\frac{\hat{Y}_i}{Y_j}\frac{p_{ij}}{p_i}, \qquad p_i=\sum_j p_{ij}$$

obteniéndose como resumen la inquietud condicionada a la predicción:

[11] Estas definiciones siguen la metodología propuesta por P. Gil (1981) para su medida de incertidumbre útil. Obsérvese que la definición de las utilidades marginales $u_{\left[\hat{Y}_i\right]}=\sum_j u_{ij}\frac{p_{ij}}{p_i}$ y $u_{\left[Y_j\right]}=\sum_i u_{ij}\frac{p_{ij}}{p_j}$ garantiza $E(u)=E\left(u_{\left[\hat{Y}_i\right]}\right)=E\left(u_{\left[Y_j\right]}\right)$.

$$H(Y/\hat{Y}) = \sum_i p_i \left(\frac{E(u)}{u_{[\hat{Y}_i]}} \right) H(Y/[\hat{Y}_i])$$

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.

Algunos resultados

Presentamos a continuación un primer análisis de las medidas anteriores en diferentes situaciones. Para ello consideramos dos series (A y B) con diferente nivel de dificultad, analizando sobre cada una de ellas seis predicciones alternativas cuya descripción aparece recogida en la siguiente tabla:

Caso 1	Predicciones efectuadas por método ingenuo $\hat{Y}_t = Y_{t-1}$
Caso 2	Predicciones aproximadamente constantes
Caso 3	Predicciones con sesgo absoluto constante $\hat{Y}_t = Y_t + B$
Caso 4	Predicciones lineales con pendiente positiva $\hat{Y}_t = a + bt$
Caso 5	Predicciones lineales con pendiente negativa $\hat{Y}_t = a - bt$
Caso 6	Predicciones con error aleatorio $\hat{Y}_t = Y_t + u_t \quad u_t \approx N(0,\sigma)$

Tabla 1

Para cada uno de los casos descritos hemos calculado la batería de indicadores estudiados en los epígrafes anteriores. El tratamiento de la información muestral, consistente en 16 observaciones de realización-predicción, ha sido llevado a cabo según los siguientes criterios:

- Las medidas de imprecisión basadas en sistemas de probabilidades A y P han sido obtenidas expresando observaciones y predicciones como peso sobre su valor total, esto es: $A_t = \frac{Y_t}{\sum_t Y_t}$, $P_t = \frac{\hat{Y}_t}{\sum_t \hat{Y}_t}$. Con esta información hemos calculado tanto la medida de imprecisión de Theil como la cuadrática.
- Para la obtención de la inquietud cuadrática condicionada hemos efectuado una clasificación de las observaciones según cuartiles de predicción.
- De modo análogo, la cuantificación de la imprecisión cuadrática basada en errores relativos ha sido llevada a cabo mediante la consideración de los cuartiles de observaciones y predicciones, obteniéndose la utilidad conjunta de cada intersección de cuartiles como la correspondiente media de error relativos de la predicción.

A grandes rasgos, las medidas presentadas en este trabajo muestran un comportamiento coherente con los indicadores habitualmente empleados, como puede apreciarse en las tablas 2 y 3, conduciendo a la selección de los casos 6 (con errores aleatorios) o 3 (con sesgo moderado).

Analizando la serie A, que presenta un nivel de dificultad "normal", se aprecia que los criterios de Theil (U e I(A:P)) y la información cuadrática aconsejarían el caso de predicciones sesgadas, frente a la alternativa con errores aleatorios (caso 6) que sería preferible según los criterios de ECM y de imprecisión cuadrática (tanto la medida basada en probabilidades como la asociada a errores relativos).

En el extremo opuesto, la alternativa 5 sería la menos deseable según los criterios cuadráticos examinados en este trabajo, mientras los criterios de Theil penalizan especialmente el método ingenuo (caso 1) y el ECM las predicciones constantes (caso 2).

Las conclusiones son similares para la serie B que, como puede apreciarse en la figura 2, presenta menor dificultad predictiva. Como consecuencia de esta menor dificultad, las medidas –aun manteniendo sus criterios de selección- conducen ahora a resultados más equilibrados, excepto el caso 5, calificado por unanimidad como el peor.

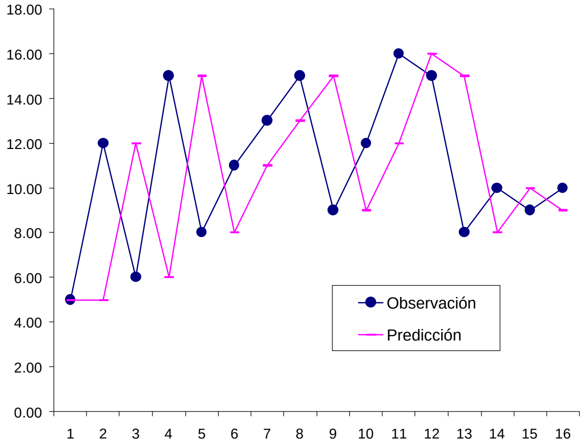
SERIE A	CASO 1	CASO 2	CASO 3	CASO 4	CASO 5	CASO 6
$\sqrt{\mathbf{ECM}}$	4.6704	6.6950	2.0000	5.7879	6.6144	0.8660
Indice de Theil (U_{66})	1.6973	1.0005	0.2435	0.9453	1.0345	0.3790
Componente de Sesgo	0.01%	8.32%	13.75%	2.37%	16.49%	14.57%
Componente de Varianza	0.00%	91.29%	80.32%	90.93%	77.58%	64.21%
Componente de Covarianza	99.99%	0.38%	5.93%	6.71%	5.93%	21.22%
Imprecisión $\mathbf{I(A:P)}$						
Medida de Theil	0.0439	0.0200	0.0006	0.0361	0.0527	0.0017
Medida Cuadrática	0.0104	0.0111	0.0017	0.0089	0.0133	-0.0013
Inquietud cuadrática H(Y)	0.2238	0.2238	0.2238	0.2238	0.2238	0.2238
Inquietud cuadrática condicionada $H\left(Y/\left[\hat{Y}\right]\right)$	0.1046	0.0913	0.0181	0.0913	0.0913	0.0277
Coeficiente de Correlación $r_{Y,\hat{Y}}$	0.0553	0.0173	1.0000	0.1717	-0.1717	0.9802
Información cuadrática asociada a predicciones $\mathbf{IC\left(Y,\left[\hat{Y}\right]\right)}$	0.1250 (55.85%)	0.1341 (59.92%)	0.2238 (100%)	0.1482 (66.22%)	0.1168 (52.21%)	0.2232 (99.75%)
Imprecisión cuadrática basada en errores relativos $\mathbf{I(\hat{Y},Y)}$	0.0557	0.0521	0.0059	0.0844	0.2048	0.0020

Tabla 2: Evaluación de predicciones para la serie A

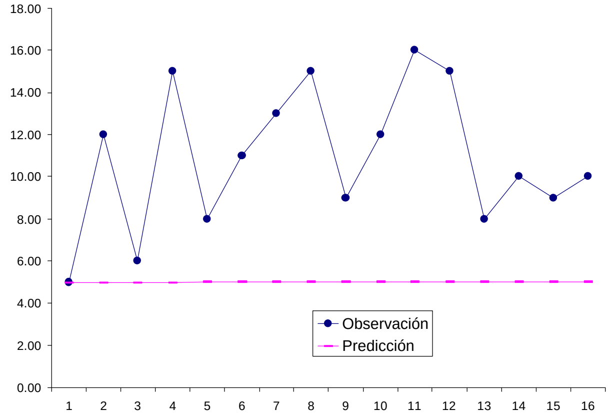
SERIE B	CASO 1	CASO 2	CASO 3	CASO 4	CASO 5	CASO 6
$\sqrt{\mathbf{ECM}}$	0.6999	3.8498	2.0000	0.7569	5.7642	0.4919
Indice de Theil (U_{66})	1.2215	0.9967	0.2650	0.7901	1.7746	1.1432
Componente de Sesgo	0.12%	52.97%	44.18%	13.46%	83.75%	0.02%
Componente de Varianza	0.05%	46.18%	49.77%	4.41%	2.63%	39.57%
Componente de Covarianza	99.84%	0.85%	6.05%	82.13%	13.62%	60.41%
Imprecisión $\mathbf{I(A:P)}$						
Medida de Theil	0.0009	0.0227	0.0010	0.0028	0.1299	0.0008
Medida Cuadrática	-0.0001	0.0129	0.0026	-0.0033	0.0292	-0.0002
Inquietud cuadrática H(Y)	0.2459	0.2459	0.2459	0.2459	0.2459	0.2459
Inquietud cuadrática condicionada $H\left(Y/\left[\hat{Y}\right]\right)$	0.0080	0.0080	0.0080	0.0080	0.0080	0.0153
Coeficiente de Correlación $r_{Y,\hat{Y}}$	0.9831	0.9730	1.0000	0.9934	-0.9934	0.9825
Información cuadrática asociada a predicciones $\mathbf{IC\left(Y,\left[\hat{Y}\right]\right)}$	0.2458 (99.95%)	0.2457 (99.91%)	0.2459 (100%)	0.2459 (100%)	0.2301 (93.55%)	0.2457 (99.89%)
Imprecisión cuadrática basada en errores relativos $\mathbf{I(\hat{Y},Y)}$	0.0002	0.2305	0.0111	0.0266	1.4697	0.0003

Tabla 3: Evaluación de predicciones para la serie B

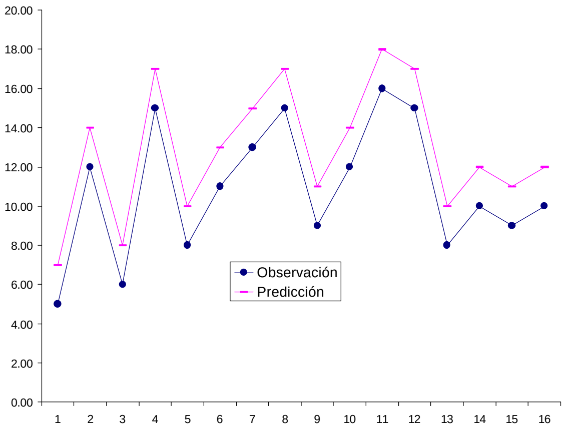
CASO 1



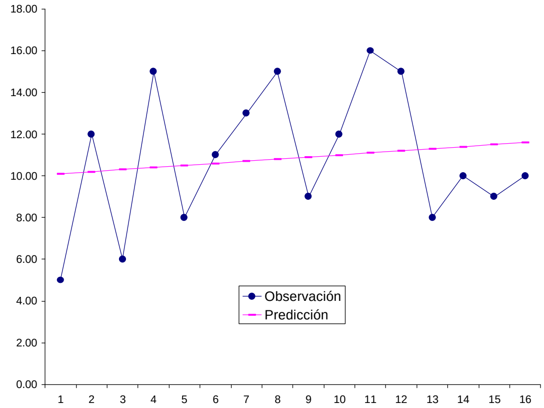
CASO 2



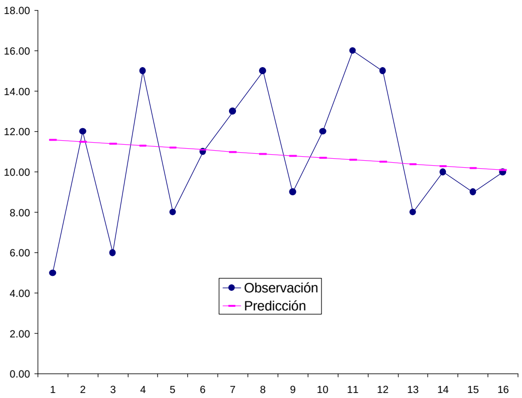
CASO 3



CASO 4



CASO 5



CASO 6

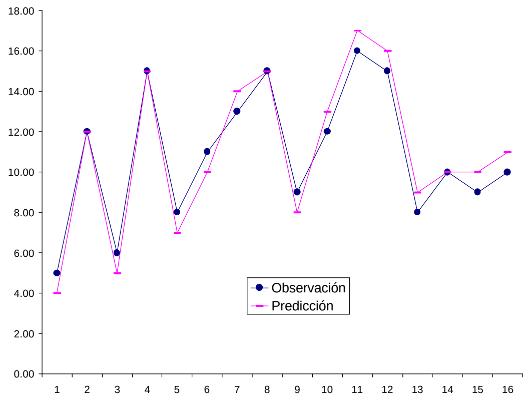
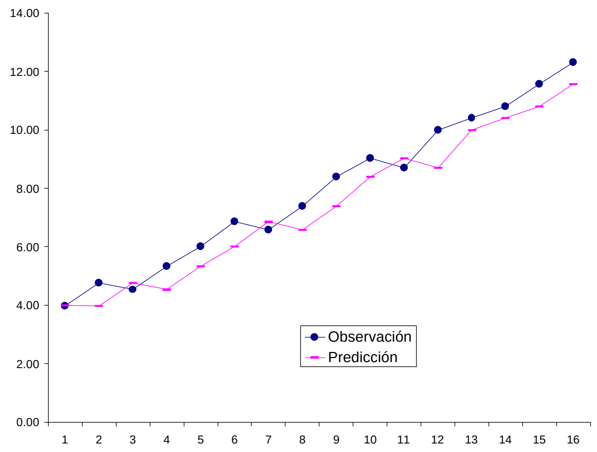
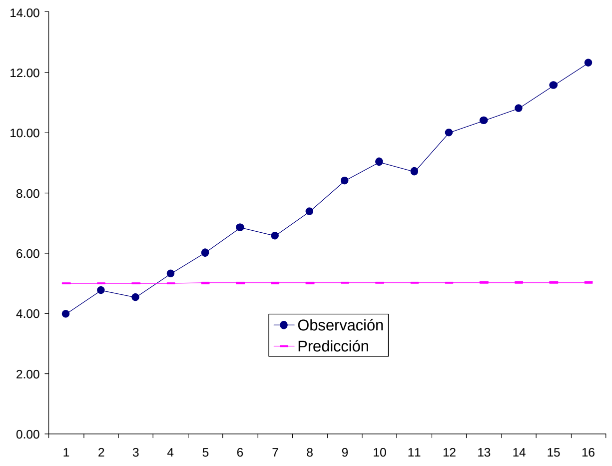


Figura 1: Predicciones alternativas para la Serie A

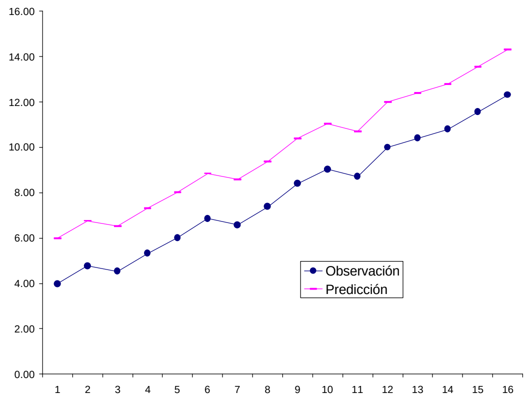
CASO 1



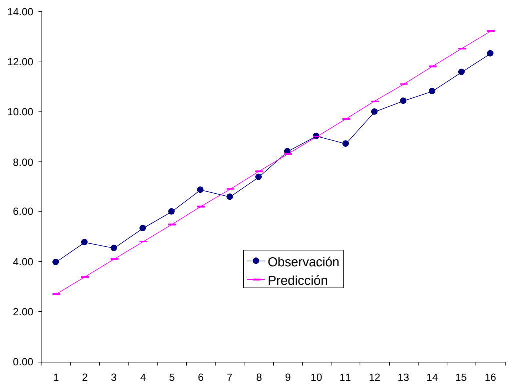
CASO 2



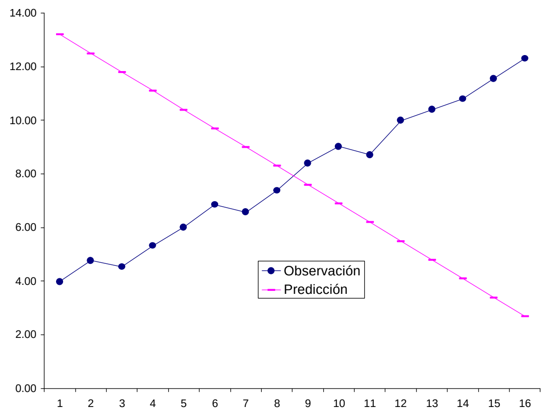
CASO 3



CASO 4



CASO 5



CASO 6

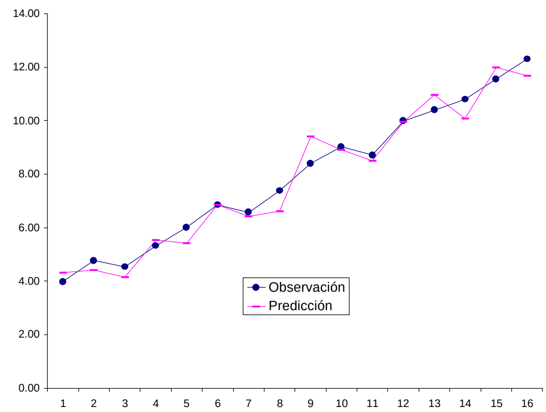


Figura 2: Predicciones alternativas para la Serie B

Reflexiones finales

Como ya hemos señalado, los contenidos de los apartados anteriores deben ser considerados como aproximaciones a la evaluación de predicciones desde la teoría de la información, sin pretender en esta etapa proponer una medida concreta. De ahí que para finalizar recojamos una visión sintética de las conclusiones extraídas hasta el momento y las necesidades percibidas como objeto de trabajos futuros:

La teoría de la información proporciona un marco de gran potencial para el diseño de indicadores de la bondad de predicciones, especialmente cuando -además de las probabilidades- se tienen en cuenta las utilidades asociadas a cada par realización-predicción.

La consideración de la inquietud cuadrática de las observaciones condicionadas a ciertos intervalos de predicción permite construir una medida cuadrática de información, que corrige dicha inquietud según la correlación entre observaciones y predicciones.

A partir de un sistema de utilidades conjuntas asociadas a las observaciones y las predicciones es posible definir una medida de imprecisión cuadrática cuyo comportamiento aplicado es similar al de los indicadores propuestos por Theil.

Dado que las expresiones de información e imprecisión cuadráticas son construídas con datos agrupados, resulta deseable analizar el coste informativo de dicha agrupación así como la posible extensión de los indicadores al caso continuo.

Las medidas cuadráticas de información en las que se basan los indicadores analizados cumplen la propiedad de ramificación, rasgo que invita a examinar su posible descomponibilidad según diferentes componentes del error de predicción.

Bibliografía

ALVARGONZÁLEZ, M.; MUÑOZ, N. (1992): “Medidas cuadráticas de información y dependencia estadística”, *Actas de la VI Reunión ASEPELT*, Granada.

ALVARGONZÁLEZ, M.; PÉREZ, R. (1989): Información cuadrática e independencia en información, *Actas de las XIV Jornadas Hispano-Lusas de Matemáticas*, Tenerife.

BELIS, M.; GUIASU, S. (1968): “A quantitative-qualitative Measure of Information in Cibernetic System”, *I.E.E.E. Trans. Inf. Th*, 14, 593-594.

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

HAVRDA, J.; CHARVAT, F. (1967): “Quantification method of classification processes”, *Kybernetica*, 3, 30-35.

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.

SHANNON, C.E. (1948): *A Mathematical Theory of Communication*, Bell System Tech. J., 27, 379-423.

THEIL, H. (1955):”Who Forecasts Best?”, *International Economic Papers*, 5, 194-199.

THEIL, H. (1958): *Economic Forecasts and Policy*, North Holland Publishing, Amsterdam.

THEIL, H. (1966): *Applied Economic Forecasting*, North Holland Publishing, Amsterdam.

THEIL, H. (1967): *Economics and Information Theory*, North Holland Publishing Company, Amsterdam.