

MODELO ECONOMÉTRICO

A) Modelos Admisibles

B) Modelos Esféricos

C) Modelo Óptimo

Valores de los coeficientes de regresión lineales para las funciones:

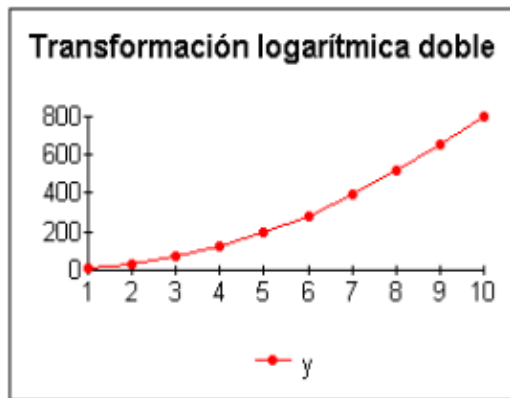
- $Y = A + B * X1$
- $Y = A + B * X1 + C * X2$

- El coeficiente de determinación **R²**, porcentaje de la varianza de la variable **Y** que es explicada por las variables **X1** y **X2**.
- El coeficiente de determinación que mide la asociación estadística entre las variables **X1** y **X2**, indicador de la existencia de *Multicolinealidad* en el modelo.
- Estimación de la desviación típica de la variable endógena.
- Valores de los estadísticos **D(A)** y **D(B)** tales que **(A-a) / D(A)** y **(B-b) / D(B)** siguen distribuciones t-Student de n-3 grados de libertad, (n-2 para el caso del modelo con una sola variable explicativa).
El cociente entre el valor del coeficiente de correlación y el estadístico **D()** permite contrastar la hipótesis de igualdad del parámetro del modelo a cero.
- Valor del contraste de **Durbin-Watson**, indicador de la existencia de *autocorrelación*, incumplimiento de la hipótesis inicial de independencia en las perturbaciones **E UtUt-i = 0**, es decir, no nulidad de las covarianzas de las perturbaciones.
El estadístico de **Durbin-Watson** detecta la autocorrelación del tipo **Ut = k*Ut-1 + Et** donde **Et** cumple las hipótesis clásicas. Se presume que cada observación es influida por la perturbación de la observación anterior, situación frecuente en los modelos de series temporales.
- Valores del contraste de **Goldfeld-Quandt** tanto si se clasifican las ternas de datos de acuerdo a **X1** como a **X2**.
Este contraste está diseñado para detectar la existencia de *heterocedasticidad*, desigualdad de las varianzas de las perturbaciones.
El applet clasifica las ternas o pares de datos aplicando regresiones lineales a los subconjuntos de datos de tamaño **(n*2)/5**. El primer subconjunto se forma con los valores más bajos de la variable independiente y el segundo con los más altos.
El cociente de las sumas de los cuadrados de los residuos sigue una distribución F de Snedecor con los mismos grados de libertad para el numerador y denominador : **((n+2)/5)-3**. La clasificación se efectúa para cada una de las dos variables explicativas.

El applet permite elegir entre diversas transformaciones de los datos, las dos últimas posibilidades de transformación aplican el procedimiento de mínimos cuadrados ponderados. Deflactando los datos por una de las variables explicativas intentaremos corregir la existencia de *heterocedasticidad*.

$$y = a + b * \ln(x)$$

$$\ln(y) = a + b * \ln(x)$$



$$y = a - b/x$$

$$\ln(y) = a - b/x$$

I MODELOS FINANCIERO-ECONOMÉTRICOS EMPLEADOS

1) Modelos regresivos

Un modelo regresivo (Aznar, 1989) trata de **explicar el comportamiento de una variable exógena en función de diversas variables explicativas** (exógenas) **o de valores anteriores de ella misma** (endógena). En modelos más complejos se pueden determinar conjuntamente varias variables endógenas a través de un sistema de ecuaciones resuelto de forma simultánea.

A lo largo de estos capítulos se emplearán modelos más o menos complejos pero en cualquier caso uniecuacionales, en los que sólo existirá una variable endógena, pudiendo existir varias explicativas. Todo este tipo de modelización parte a priori con un handicap la **existencia de factores reales explicativos** de la variable endógena, es bien conocido que en estadística siempre es posible incrementar el grado de explicación por el aumento de variables exógenas en el modelo.

Conviene aclarar, que no se ha tomado ningún partido en cuanto a la mayor o menor validez de los modelos regresivos o de cualquier otro, antes al contrario, se ha tratado de testear todos y cada uno de los modelos que los autores creían podían aportar buenos resultados. Las conclusiones son en muchos casos sorprendentes, y avalan que en muchas ocasiones ideas apriorísticas como por ejemplo la necesidad de modelizaciones distintas de la lineal para la volatilidad, no son ciertas en todas las ocasiones, sino más bien lo contrario.

En este estudio se ha pretendido en todo momento un **exhaustivo seguimiento de las posibles correlaciones cruzadas entre variables explicativas** para evitar redundancias y por supuesto de rechazar, mediante el empleo sistemático de test estadísticos de no significatividad, las variables que no fueran claramente relevantes en la explicación de la volatilidad.

Para la utilización de estos modelos debemos partir de una idea ex-ante sobre la relación entre una o varias variables y tratar de verificar que ésta realmente exista, pero el problema no sólo reside en la existencia de relaciones perdurables en el tiempo sino en la forma funcional que éstas adopten, por ello en este trabajo se emplearán **diversos modelos**, que a priori eran susceptibles de dar buenos resultados:

– **Relaciones lineales univariantes.**

$$y_t = \alpha_0 + \alpha_1 \beta_{1t} + u_t$$

Donde:

y es la variable dependiente. **beta** es la variable explicativa. Las **alphas** los parámetros especificados por el modelo es el término de error.

– **Relaciones lineales multivariantes.**

$$y_t = \alpha_0 + \alpha_1 \beta_{1t} + \alpha_2 \beta_{2t} + \dots + u_t$$

Donde:

y es la variable dependiente. **betas** son las variables explicativas. Las **alphas** los parámetros especificados por el modelo es el término de error.

– **Modelos de crecimiento exponencial.**

$$y_t = e^{\alpha_0 + \alpha_1 x_{1t} + \alpha_2 x_{2t} + \dots + u_t}$$

Donde:

y es la variable dependiente. x son las variables explicativas. Las **alphas** los parámetros especificados por el modelo. es el término de error

– **Modelos con punto de ruptura.**

$$y_t = \alpha_0^1 + \alpha_1^1 x_{1t}^1 + \alpha_2^1 x_{2t}^1 + \dots + u_t^1; \text{ si } x > \beta$$

Donde:

y es la variable dependiente. x son las variables explicativas. Las **alphas** los parámetros especificados por el modelo. es el término de error. el número e. Siendo **beta** el punto de ruptura.

Otros modelos como el estudio de eventuales relaciones logarítmicas entre precio y volatilidad, dieron escasos o nulos resultados por los que no se reseñan en el apartado anterior. Los ya mencionados modelos con punto de ruptura ("**piecewise models**") **si se muestran muy útiles** en la superación de algunos problemas ya reseñados de la base de datos, especialmente el inusitado incremento sufrido por la volatilidad implícita en los peores momentos de la crisis monetaria y **apuntar a un cambio de la relación precio– volatilidad implícita en diversos momentos del ciclo bursátil.**

2) Modelos SARIMA

Este tipo de modelización **permite soslayar en parte dos problemas** de los modelos regresivos, por un lado **no es necesaria la identificación de variables susceptibles de explicar** la variable endógena al aplicarse sobre una serie de datos correlativos de la propia variable, y además toda una serie de pautas de actuación preestablecidas hacen **más sencillo obtener la forma funcional adecuada del modelo.**

Hay que explicar que **el gran problema** de estos modelos, lo constituye precisamente el **cálculo del citado punto de ruptura**, en estos capítulos sólo se especificará uno de los puntos más relevantes de ruptura, que se produjo en la ruptura del 3000 por parte del Ibex, comienza a existir evidencia empírica de que puede haber un nuevo punto de ruptura en el año 99 cuando se produce una evidente falta de capacidad explicativa de los modelos regresivos empleados hasta ese momento.

El **paso previo** al empleo de la metodología SARIMA es la **obtención de series estacionarias**, es decir, bases de datos que cumplan los siguientes **criterios**

i) **Media constante** en el tiempo.

ii) **Varianza constante** en el tiempo.

Para ello **se procede a una serie de diferenciaciones** y ajustes que quedan explicados paso a paso en el momento en que se utilicen a lo largo del trabajo.

A la hora de **especificar un modelo ARIMA** son necesarios **3 parámetros**:

i) El **retardo de la parte autorregresiva (AR)** que viene explicitado por el subíndice p en la siguiente fórmula:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + u_t$$

Donde:

y es la variable dependiente Las **phi** representan en este caso, a los parámetros especificados por el modelo. es el término de error.

En realidad la relación matemática no hace sino poner en evidencia que la observación de la variable y en el período t depende de las misma variable y en t-1, t-2,... y t-p.

ii) El **retardo de la parte de medias móviles (MA)**, q en la fórmula siguiente:

$$y_t = \mu + u_t + \sigma_1 u_{t-1} + \sigma_2 u_{t-2} + \dots + \sigma_q u_{t-q}$$

Donde:

y es la variable dependiente Las **sigma** representan en este caso, a los parámetros especificados por el modelo. son los términos de error.

La significación de este tipo de modelos es también bastante clara, se trata en suma de explicar la variable y en t en función de una constante y de una corrección de q errores del modelo en los períodos anteriores.

iii) Número de **veces que hemos de diferenciar (I) la serie** para hacerla estacionaria.

Se entenderá por **diferencial de yt**, la diferencia entre la volatilidad en t y t-1, es decir:

$$\Delta y_t = y_t - y_{t-1}$$

Donde:

d indica que se trata de la diferencial de la variable de que se trate y es la variable dependiente En ocasiones se emplea la **diferenciación logarítmica para obtener un mayor alisado de la serie**, sin embargo, este tipo de diferenciación no dió excesivos buenos resultados a la hora de modelizar las series de datos empleadas en este estudio.

La utilización de estos modelos permite constatar un **hecho** que no deja de ser preocupante, la **volatilidad en t no es independiente de la manifestada en otros períodos inmediatamente anteriores**, poniendo en tela de juicio las propias fórmulas valorativas de opciones , empleadas por la práctica totalidad de los participantes en el mercado.

Antes de entrar en el siguiente apartado, sería conveniente explicar las **principales diferencias entre los modelos ARIMA y los de la familia ARCH**, mientras que **los ARIMA se aplican sobre series homocedásticas**, es decir, de varianza constante, **los ARCH se caracterizan por tratar de modelizar precisamente la heterocedasticidad de la serie**. Por lo tanto mientras que en los modelos ARIMA se debe de alisar la serie como paso previo a la modelización de la misma, ésto no es necesario en los ARCH, lo que no es por si mismo una ventaja al obligar a elegir una de las múltiples posibilidades de modelización ofrecidos por los ARCH que se tratarán en el apartado posterior.

Los resultados del trabajo de **Poon y Taylor** (1991) hacen intuir que los modelos ARIMA y cualquiera de los ARCH no son en principio excluyentes, mientras **los ARIMA son válidos para identificar los cambios de tendencia en volatilidad**, en el caso de los heterocedásticos se observa una mayor inestabilidad en el resultado, aunque la serie prevista se ajuste más a los frecuentes cambios en la volatilidad real. En resumen, la serie, en este caso generada por un GARCH (1,1), prevista por el modelo heterocedástico se ajustaba mejor, pero era más volátil en la cuantía de los errores.

Un tema que **no se trata en este trabajo es el análisis de la posible existencia de estacionalidad** (la S de los modelos SARIMA) en la series de volatilidad, se ha de reconocer que existen períodos concretos en los que se producen cambios de una manera casi sistemática en la volatilidad implícita. Se puede citar el fenómeno de reducción de la volatilidad los días inmediatamente anteriores a puentes festivos (en los días inmediatamente anteriores a la festividad de semana santa de 1996, la volatilidad implícita se redujo desde el 17% a niveles cercanos al 13% en las series "at-the-money") pero sería muy difícil llegar más lejos en el análisis por la relativa limitación de la base de datos.

Hay sin embargo **una excepción** a esta afirmación anterior, se podrá constatar a lo largo de la presentación de los resultados de los modelos que recurrentemente aparece un efecto estacional (que será denominado pseudoestacional en este trabajo) que coincide con el período de cálculo de la volatilidad histórica homocedástica que se esté modelizando.

3) Modelos heterocedasticos

A modo de introducción sería conveniente recordar que estos modelos son relativamente novedosos, su desarrollo comienza con los trabajos de **Engle** (1982, op. citada) y se muestran especialmente **útiles** según Crouhy y Rockinger (op. citada) **para:**

- La **modelización de una volatilidad no constante** en el tiempo.
- **Recoger** en modelos teóricos la evidencia empírica de **que la volatilidad se manifiesta en olas**.
- Identificar la **existencia de una memoria importante** en el proceso.
- **Predicción** de la volatilidad futura

A) ARCH

El modelo ARCH fue desarrollado por Engle en 1982 y se basa en la **explicación de la volatilidad condicional como una función lineal de q errores pasados** de predicción elevados al cuadrado (los epsilon de la fórmula).

$$\sigma_t^2 = \alpha_0 + \sum_{k=1}^q \alpha_k \varepsilon_{t-k}^2, \alpha_0 > 0, \alpha_k \geq 0$$

Donde:

sigma es la variable condicional Los **alpha** son los parámetros especificados por el modelo **Epsilón** son los términos de error

Tanto este modelo como el resto de los que se tratará a continuación, que no son sino desarrollos más o menos complejos del mismo, definen la **volatilidad condicional**, que es la única que se puede predecir por las hipótesis del modelo, como la volatilidad condicionada **a la información existente** en ese periodo (denotada por Phi en la formula siguiente).

$$\sigma_t^2 = E(\varepsilon_t^2 / \phi_{t-1})$$

Donde:

Epsilón en t es el "shock" o error de predicción **Phi** representa la información existente en t y que no existía en t-1.

Los **modelos tradicionales** trataban sólo la volatilidad incondicional, es decir:

$$\sigma^2 = E(\varepsilon_t^2)$$

Si se piensa en el caso de **la volatilidad implícita**, ésta viene condicionada por toda la información existente hasta el período considerado, por lo que el concepto de "shock" de volatilidad podría ajustarse sin excesivos problemas. Es tal vez por esta razón por la que es posible llegar a una conexión de las volatilidades implícita y condicional sin excesivos problemas, a pesar de que no exista relación funcional alguna entre ellas

En cuanto a la **memoria del proceso**, se corresponde con la **suma de todos los coeficientes alfa**, en concreto:

$$mem = \sum_{k=1}^k \alpha_k$$

Siendo **r** el número de coeficientes alpha distintos de 0.

De acuerdo con un **modelo ARCH(1)** se puede prever la volatilidad s períodos más adelante como sigue:

$$\sigma_{t+s}^2 = \sigma^2 + \alpha_1^s (\sigma_t^2 - \sigma^2)$$

Siendo **sigma cuadrado** la volatilidad incondicional, es decir, aquella que no depende del período en que nos encontremos. En el largo plazo es destacable que la volatilidad condicional e incondicional tienden a aproximarse en cualquier caso, como veremos más tarde es muy raro que alpha1 sea mayor que la unidad, ésto no es así necesariamente en los modelos GARCH.

B) GARCH

Este modelo fue desarrollado por **Bollerslev** (1987), extendiendo el modelo ARCH para incluir retardos en la varianza condicional. En definitiva **un GARCH es un modelo ARCH infinito**, un GARCH (p,q) se define como:

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{k=1}^q \alpha_k \epsilon_{t-k}^2; \beta_j \geq 0, \alpha_k \geq 0, \alpha_0 > 0$$

Donde:

sigma es la variable condicional Los **alpha** y **beta** son los parámetros especificados por el modelo **Epsilon** son los términos de error

Si p es cero el proceso se reduce a un ARCH(q). Si tuviésemos por ejemplo un GARCH de reducida p, que son los más comunes en los estudios de mercado, sus propiedades vendrían a ser equivalentes a un ARCH con una q elevada, en general con q mayor o igual a 20.

Como quiera que **por las hipótesis propias al modelo:**

$$\epsilon_t^2 = \sigma_t^2 + \mu_t$$

Se puede **reformular un GARCH(p,q)** como sigue:

$$\epsilon_t^2 = \alpha_0 + \sum_{k=1}^m (\alpha_j + \beta_j) \epsilon_{t-k}^2 - \sum_{j=1}^p \beta_j \mu_{t-j} + \mu_t$$

Donde **m** es el máximo entre p y q, y **mu** no manifiesta correlación con las otras series.

La **previsión** en un modelo GARCH(1,1) se efectúa como sigue:

$$\sigma_{t+1}^2 = \sigma^2 + (\alpha_1 + \beta_1)^2 (\sigma_t^2 - \sigma^2)$$

Sólo en el caso de que $\alpha_1 + \beta_1$ sea inferior a 1 la volatilidad prevista decrecerá hacia la incondicional, en ese caso se dice que el modelo es integrado.

C) EGARCH o "exponential GARCH"

Este modelo surge **para paliar un problema** de los modelos GARCH consistente en que **los efectos de una "sorpresa"**, entendida como información inesperada por el mercado, son los mismos se trate de una noticia negativa o positiva. Sin embargo parece empíricamente demostrado que los picos en volatilidad coinciden con caídas de mercado, por lo que una mayor asimetría de la distribución de volatilidad sería conveniente. Este modelo a pesar de su complejidad teórica **permite estudiar desde otro punto de vista la asimetría precio-volatilidad** que ya se estudia en la parte regresiva, pues no es necesario acudir a la serie de precios para la modelización.

En un modelo **EGARCH(1,1)** la **varianza condicional** se escribiría como sigue:

$$\ln \sigma_t^2 = \alpha_0 + \beta_1 \ln \sigma_{t-1}^2 + \delta \frac{\epsilon_{t-1}}{\sigma_{t-1}^{0.5}} + \alpha_1 \left[\frac{|\epsilon_{t-1}|}{\sigma_{t-1}^{0.5}} - \left(\frac{2}{\pi} \right)^{0.5} \right]$$

Donde:

sigma es la variable condicional Los **alpha** son los parámetros especificados por el modelo **Epsilon** son los términos de error **Delta** el parámetro de asimetría **Pi** es el número pi

Un coeficiente **delta** estimado negativo confirmaría la presencia de un sobreimpacto de las noticias negativas, la contribución de un "**shock**" **positivo** en la volatilidad condicionada es:

$$(\alpha_1 + \delta) \frac{\varepsilon_{t-1}}{\sigma_{t-1}^2}$$

Por el contrario un "**shock**" **negativo**:

$$(\alpha_1 - \delta) \frac{\varepsilon_{t-1}}{\sigma_{t-1}^2}$$

La **asimetría** de la volatilidad condicionada la mediría el siguiente ratio:

$$\rho = \frac{\alpha_1 + \delta}{\alpha_1 - \delta}$$

D) Non-Linear Assymmetric ARCH (AGARCH)

Desarrollado por **Engle**(1990) y **Sentana** (1991). Un **AGARCH (1,1)** :

$$\sigma_t^2 = w + \alpha(\varepsilon_{t-1} - \lambda)^2 + \beta \sigma_{t-1}^2; \text{si } \lambda > 0$$

Donde:

sigma es la variable condicional **w** es la variable incondicional **alpha y beta** son los parámetros especificados por el modelo **Epsilon** son los términos de error **Lambda** el coeficiente de asimetría

Se formulan a continuación **algunos modelos susceptibles de ser probados en nuestro mercado**, que no han sido probados en el presente trabajo.

E) Otros modelos Heterocedásticos

La lista de modelos heterocedásticos es muy amplia, de hecho una gran parte de la innovación de la ciencia econométrica se ha centrado en los últimos años en el desarrollo de modelos que se ajusten cada vez más a las series de volatilidades.

– Glosten, Jagannathan y Runkle desarrollan el **GJR-ARCH** en 1989. Un GJR-ARCH(1,1) se expresaría como sigue:

$$\sigma_t^2 = w + \alpha \varepsilon_{t-1}^2 + \delta D \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

Donde:

sigma es la variable condicional Los **alpha, beta, delta** son los parámetros especificados por el modelo **D** es una variable binaria que toma el valor 0 o 1. **Epsilon** son los términos de error

D es una variable que toma el valor 1 en t, si el error en t-1 es negativo y 0 si el error en t-1 es positivo.

– **TGARCH**

$$\sigma_t = w + \sum_{i=1}^p (\alpha_i^+ \varepsilon_{t-i}^+ - \alpha_i^- \varepsilon_{t-i}^-) + \sum_{j=1}^q \beta_j \sigma_{t-j}; \varepsilon_t^+ = \max(\varepsilon_t, 0); \varepsilon_t^- = \min(\varepsilon_t, 0)$$

Donde:

sigma es la variable condicional Los **alpha y beta** son los parámetros especificados por el modelo **Epsilon** son los términos de error que se definen como doble variable tal y como se expresa en la fórmula

