

Evaluación de pronósticos para la volatilidad empleando modelos autoregresivos con heterocedasticidad condicional generalizados aplicados a precios diarios de acciones en Argentina

UNR - FCEYE - Maestría en Estadística Aplicada

27 de octubre de 2017

Autor: Luis Damiano

Directora: Mg. María Teresa Blaconá

Hechos Estilizados: Retorno y Volatilidad

Estimadores de la Volatilidad

Heterocedasticidad Condicional

Aplicación al Mercado de Acciones en Argentina

Bibliografía

Problemática muy desafiante desde los puntos de vista estadístico y financiero.

- **Estadística:** el proceso ocurre una sola vez, fenómeno de gran complejidad, errores de medición, datos faltantes, efectos no observables, parámetros inestables a lo largo del tiempo, variables integradas, heterocedasticidad condicional, densidad no gaussiana, modelos no lineales.
- **Finanzas:** medición, gestión y predicción del riesgo para finanzas personales, finanzas corporativas, toma de decisiones de política económica, ahorro público.

- Doce acciones más líquida del BYMA (84% del volumen negociado) e Índice Merval.
- Desde 1-ene-2009 hasta 31-may-2016.
- Serie diaria de precios homogéneos.
- Apertura, máximo, mínimo y cierre (*OHLC*).
- Plazo de negociación 72hs.
- Aprox. 1760 observaciones por serie.
- Datos faltantes eliminados.
- Retornos logarítmicos truncados al 1%.

Hechos Estilizados: Retorno y Volatilidad

Sea P_t el precio de un activo al cierre del día de negociación t .

- **Retorno logarítmico:** tasa de variación compuesta en tiempo continuo

$$r_t = \ln(P_t/P_{t-1}) \quad (1)$$

- **Información:** conjunto de información disponible al momento $t - 1$

$$\mathcal{F}_{t-1} \quad (2)$$

- **Media condicional**

$$E \langle r_t | \mathcal{F}_{t-1} \rangle \equiv E_{t-1} \langle r_t \rangle \equiv \mu_t \quad (3)$$

- **Variancia condicional**

$$V \langle r_t | \mathcal{F}_{t-1} \rangle \equiv V_{t-1} \langle r_t \rangle \equiv \sigma_t^2 \quad (4)$$

Características suficientemente consistentes para ser generalmente aceptadas como verdaderas.

- Ciertas **regularidades** en las series de precios.
- **Propiedades estadísticas** no triviales.
- Generalidades que permiten **explotar las similitudes** a cambio de perder cierta habilidad para hacer observaciones precisas.

Estudiar el cambio en los precios a través de la función de autocorrelación (τ representa el rezago).

$$\rho(\tau) = \text{corr}(r_t, r_{t+\tau}) \quad (5)$$

- **Teoría:** hipótesis de los mercados eficientes.
- **Evidencia:** en mercados líquidos, los retornos no observan autocorrelación lineal significativa para series de tiempo muestreadas con una frecuencia mayor a los 20 minutos.

El arbitraje blanquea el espectro de la variación en los precios (Mandelbrot 1971).

No confundir falta de correlación lineal con independencia!

- **Generalización:**

$$\rho_{\alpha}(\tau) = \text{corr}(|r_t|^{\alpha}, |r_{t+\tau}|^{\alpha}), \quad (6)$$

- **Efecto de Taylor:** Para un rezago dado, la correlación se maximiza en $\alpha = 1$ (Ding, Granger, and Engle 1993; Granger and Ding 1994; Taylor 1986).

Los retornos absolutos son los más predecibles.

La volatilidad es **persistente**:

- Volatilidad de un momento influencia la volatilidad esperada para varios períodos futuros.
- Manifiesta la existencia de bloques o conglomerados (*clustering*).
- Proceso de memoria larga.

El **efecto apalancamiento** es otra forma de dependencia no lineal en los retornos.

- Correlación entre un retorno y el cuadrado del subsiguiente

$$L_{\tau} = \text{corr} (r_t, |r_{t+\tau}|^2). \quad (7)$$

- Usualmente comienza con un valor negativo y decae hacia cero, sugiriendo que una **variación negativa en el precio incrementa la volatilidad**.
- Pruebas de hipótesis (Engle and Ng 1993);
 - **Sesgo en el signo**: perturbaciones positivas y negativas presentan un impacto diferente en la volatilidad condicional.
 - **Sesgo en el tamaño**: la magnitud de la perturbación de un signo dado está asociada a la asimetría en la volatilidad.
 - **Prueba conjunta**.

▪ Correlación

- En la mitad de las series de retornos se rechaza la hipótesis nula de independencia hasta el décimo rezago en los retornos.
- No se observan mayores discrepancias entre los resultados con datos originales y valores truncados.
- Se manifiesta dependencia no lineal en la forma de autocorrelación lineal en el cuadrado de los retornos.

▪ Efecto apalancamiento

- Empleando los datos originales, cinco series muestran un efecto conjunto significativo.
- Muchos de estos efectos desaparecen al emplear retornos truncados.
- **TECO2** presenta un efecto conjunto significativo por el sesgo en el signo y en el tamaño de los retornos negativos.
- **MERVAL** sólo muestra un efecto conjunto significativo.

La forma de la distribución marginal de los retornos difiere para distintas frecuencias de muestreo.

- **Leptokurtosis**
- **Asimetría**
- **Efecto apalancamiento**

- Picos más apuntalados y colas más pesadas.
- Más aún, tienen una distribución **condicional de colas pesadas** a pesar de ser filtrado por un modelo de heterocedasticidad condicional.
- Parte importante de la kurtosis está explicada por los retornos nocturnos (*overnight*), en especial producto de eventos, anuncios y noticias.

En la práctica, esto implica que la variación en los precios suele fluctuar en un rango más acotado que lo parametrizado por una normal, aunque en algunas ocasiones los retornos saltan hacia valores más extremos.

- Prevalece en activos de inversión (los agentes esperan obtener retornos positivos).
- No se observa en divisas.
- El efecto se manifiesta sin importar la medida de volatilidad seleccionada.

Observaciones empíricas en Argentina:

- **Normalidad:** la prueba de Jarque-Bera rechaza la hipótesis de simetría y exceso de kurtosis nulo (incluso retornos truncados).
- **Asimetría:** nueve series tienen una distribución asimétrica hacia la derecha (frecuentes pérdidas pequeñas y pocas grandes ganancias).
- **Leptokurtosis:** se destaca YPF por tener la mayor kurtosis con un coeficiente de 14.1385 que se reduce a 5.0516 al truncar los datos.

Distribución paramétrica - Aplicación YFPD

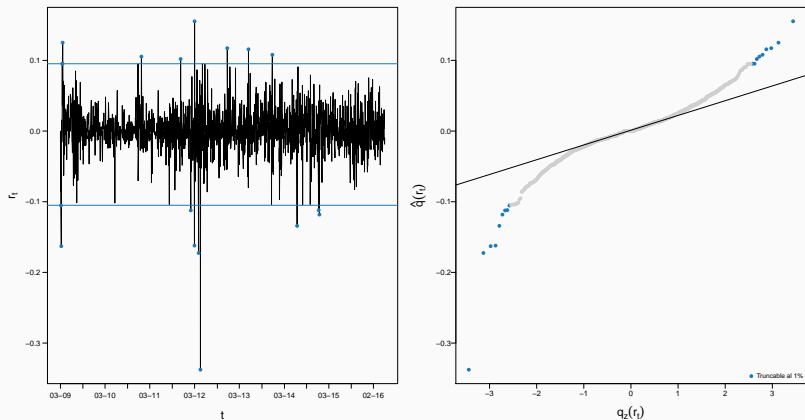


Figure 1: Valores extremos (YPFD). Izquierda: Serie histórica del retorno logarítmico. Los puntos azules destacan aquellas observaciones que serían truncadas al 1%. Las líneas marcan el valor del cuantil. Derecha: gráfico de cuantil-cuantil de los retornos logarítmicos para una distribución gaussiana. Elaboración propia.

Las características mencionadas no resultan suficientes para identificar una única distribución.

- **Literatura habitual:** t de Student, hiperbólica, normal inversa, estable y estable exponencialmente truncada.
- **Parametrizaciones:** los parámetros de locación, dispersión, simetría y peso en las colas son necesarios para reflejar las características empíricas.

Sinclair (2013) afirma que la distribución en particular no es importante, siempre que la misma sea fuertemente asimétrica hacia la derecha.

Debe modelarse explícitamente que la volatilidad transcurre mucho más tiempo en estados bajos que altos.

Crisis financiera global del año 2007:

- Subestimación del riesgo que derivaron en **pérdidas excesivamente mayores a las esperadas** (Daníelsson 2011).
- Movimientos en los precios equivalentes a **25 veces el desvío estándar** (Larsen 2007).
- Eventos modelados con una probabilidad de ocurrencia de una vez cada 10.000 años que finalmente se verificaron durante **tres días seguidos** (Whitehouse 2007).

No es solamente un problema académico!

El alto grado de variabilidad en los precios implica que los movimientos más bruscos tienen una **probabilidad de ocurrencia no despreciable**.

- Las grandes fluctuaciones no pueden ser excluidas ya que forman una fracción importante de los retornos agregados a lo largo del tiempo.
- Motivan desarrollos teóricos para replicar las colas pesadas.
- La teoría financiera define el Valor en Riesgo (*Value At Risk*) como el cuantil hacia la derecha de la función de pérdida para un horizonte de tiempo.

- Los valores extremos influyen fuertemente los estadísticos empleados en la **descripción de las series**.
- En 160 series de índices bursátiles y tipos de cambios, el efecto de Taylor desaparece si se emplean **estadísticos robustos** para la asimetría (Teräsvirta and Zhao 2011).
- Analistas financieros suelen **truncar datos** (*winsoring*):¹

$$\hat{r}_\alpha = \max \left[\min \left(r, r_{1-\alpha/2} \right), r_{\alpha/2} \right]. \quad (8)$$

¹Expresión que lleva el nombre del ingeniero y bioestadístico Charles P. Winsor (1895-1951).

La volatilidad sigue ciertos **patrones que permiten la predicción**.

- **Varía** a lo largo del tiempo (Akgiray 1989; Turner and Weigel 1992).
- **Persistente** (autocorrelación) y se manifiesta en **bloques** (*clustering*).
 - El efecto tiende a ser mayor en los mercados desarrollados (Aggarwal, Inclan, and Leal 1999).
 - Los bloques suelen resultar más pronunciados cuando los precios sufren variaciones negativas.
 - Las bajas de precios más pronunciadas presentan caídas más rápidas en la correlación.
- **Reversión a la media**: la volatilidad de corto plazo vuelve hacia la media de largo plazo (Engle and Patton 2001).

Estos hechos estilizados no dependen del activo bajo análisis y son observados para índices, acciones, commodities y monedas.

Estimadores de la Volatilidad

Concepto central en las finanzas, en especial para las teorías de valuación de activos, selección de portafolios y administración del riesgo.

- Medida de la **variabilidad en el precio** de un activo.
- Habitualmente asociada al **desvío estándar** de los retornos.
- **Múltiples definiciones** que responden a diferentes problemas del conocimiento científico.
- **Diferencias en las estimaciones** en función de cada modelo subyacente y los supuestos empleados.

- No es observable **instantáneamente** sino que requiere del paso del tiempo para manifestarse.
- No es observable **directamente**: autores han diseñado un amplio menú de estimadores estadísticos con ventajas y desventajas.
 - **Sesgo**: error sistemático de estimación.
 - **Eficiencia**: velocidad de convergencia al verdadero valor, usualmente en forma relativa a la variancia del estimador del desvío estándar.

La **volatilidad** es la suma de infinitos retornos instantáneos tomados en períodos infinitesimales y no solapados que componen una unidad de tiempo.

En tiempo discreto, se traduce en la raíz cuadrada de la variancia de los retornos

$$\hat{\sigma} = \sqrt{\frac{1}{T} \sum_{t=1}^T (r_t - \bar{r})^2} \quad (9)$$

- Sólo requiere convergencia en la suma.
- No realiza supuestos sobre la distribución del proceso generador de datos.

En los procesos financieros resulta difícil **diferenciar la media**, llamada desplazamiento o *drift*, de la variancia de los retornos.

- Los estimadores para la media tienen un **gran componente de ruido**, en especial para muestras de tamaño pequeño.
- Habitualmente se **fija el retorno medio en cero** a fin de eliminar una fuente de ruido e incrementar la precisión de la medición.

Aunque es un estimador insesgado para la variancia, su raíz cuadrada subestima el desvío estándar producto de la **desigualdad de Jensen**.

- La transformación convexa de la media es siempre menor o igual a la media del valor transformado

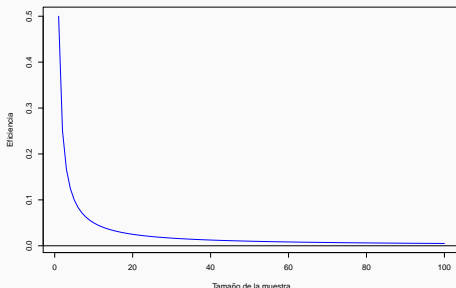
$$E \langle \hat{\sigma} \rangle = E \langle \sqrt{\hat{\sigma}^2} \rangle \leq \sqrt{E \langle \hat{\sigma}^2 \rangle} = \sqrt{\sigma^2} = \sigma. \quad (10)$$

Estimadores - Cierre a Cierre: eficiencia

Asumiendo ruidos Gaussianos, la variancia del estimador es

$$V \langle \hat{\sigma}_{cc} \rangle = \frac{1}{T} \left(T - 1 - \frac{2\Gamma^2 \left(\frac{T}{2} \right)}{\Gamma^2 \left(\frac{T-1}{2} \right)} \right) \sigma^2 = \frac{1}{T} (T - 1 - \sqrt{T}b^2) \sigma^2 \approx \frac{\sigma^2}{2T}. \quad (11)$$

El estimador es **ineficiente**: converge lentamente al verdadero valor de la volatilidad.



Dinámica poco favorable:

- Muestras pequeñas dan lugar a estimaciones con mucho ruido.
- Incrementar el número de observaciones ampliando el ancho de la ventana incorpora información irrelevante para el estado actual del mercado.

Tomando $T = 22$ días, el intervalo de confianza del 95% es tan amplio como 29.55% de la verdadera volatilidad.

$$100 \times \frac{z_{1-\alpha} \sigma / \sqrt{2T}}{\sigma} = 100 \times \frac{1.96}{\sqrt{44}} \approx 29.55 \quad (12)$$

Parkinson (1980) toma herramientas de la física para modelar el cambio en los precios como el movimiento de una partícula.

Sean h_t y l_t los valores máximo y mínimo que un precio toma dentro de un intervalo de tiempo dado t ,

$$\hat{\sigma}_{park} = \sqrt{\frac{1}{4T \ln 2} \sum_{t=1}^T \left(\ln \frac{h_t}{l_t} \right)^2}. \quad (13)$$

- Intuitivamente, la diferencia entre los valores mínimos y máximos de un precio en un intervalo de tiempo arroja un buen estimador de la variancia.
- Esta diferencia tiende a ser menor o mayor cuando la constante de difusión también lo es.

- **Sesgo:**
 - Los precios son registrados en tiempo discreto. Factor de corrección cuantificado mediante simulaciones (Garman and Klass 1980).
 - No considera los saltos en la apertura.
 - Desigualdad de Jensen.
- **Robustos** al ruido generado por la microestructura (*bid-ask spread* y *bid-ask bounce*) (Brandt and Diebold 2003).

- **Eficiencia relativa:** estimador de distancias es aproximadamente cinco veces más eficiente que la variancia muestral

$$V \langle \hat{\sigma}_{park}^2 \rangle = E \langle (\hat{\sigma}_{park}^2 - \sigma^2)^2 \rangle = \left[\frac{E \langle L^4 \rangle}{E \langle L^2 \rangle} - 1 \right] \frac{\sigma^2}{T} \approx \frac{0.41\sigma^2}{T}, \quad (14)$$

donde $L_t = h_t - l_t$ es la diferencia entre el mayor y el menor precio en el i -ésimo período de tiempo (rango).

- **Convergencia** más rápida porque emplea la información proveniente de dos precios para cada período de tiempo.

El analista financiero obtiene una estimación cinco veces más precisa, o bien emplear sólo el 20% de los datos más actuales.

- **Garman-Klass:** usa precios de **apertura y cierre**, pero asume desplazamiento nulo (Garman and Klass 1980).
- **Rogers-Satchell:** generalizan GK para **cualquier desplazamiento**, pero mejora el rendimiento sólo si el verdadero valor del parámetro es distinto de cero (L. C. G. Rogers and Satchell 1991; L. C. Rogers, Satchell, and Yoon 1994).
- **Yang-Zhang:** además de desplazamientos, admite **saltos** pero supone falta de correlación entre la volatilidad diaria y la nocturna (Yang and Zhang 2000).

Estimador	Datos	Desplazamiento	Saltos	Eficiencia máxima
Cierre a cierre	Cierre	No	No	1.0
Parkinson	Máximo y mínimo	No	No	5.2
Garman-Klass	Apertura, máximo, mínimo y cierre	No	No	7.4
Rogers-Satchell	Apertura, máximo, mínimo y cierre	Sí	No	8.0
Yang-Zhang	Apertura, máximo, mínimo y cierre, apertura y cierre del período anterior	Sí	Sí	14.0

Table 1: Resumen de los estimadores seleccionados para la volatilidad.

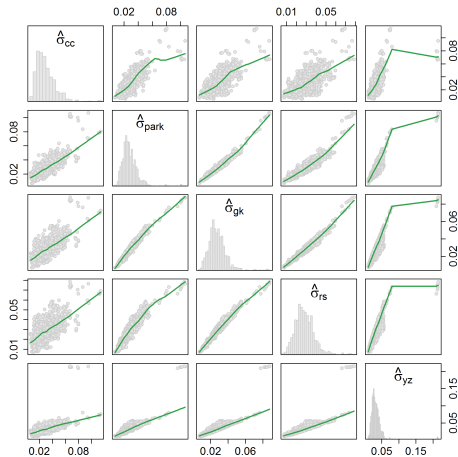


Figure 2: Volatilidad de COME (transversal).

Al igual que Brandt and Kinlay (2005), se observa una fuerte correlación lineal para estimaciones contemporáneas.

- **Cierre a cierre:** mayor grado de variabilidad. Difiere un tanto de los restantes con una correlación estimada superior al 0.60.
- **Basados en distancias:** muy estrechamente relacionados entre sí con coeficientes de correlación estimados mayores a 0.90.
- **Valores extremos:** el proceso de limpieza de los datos resulta clave para COME.

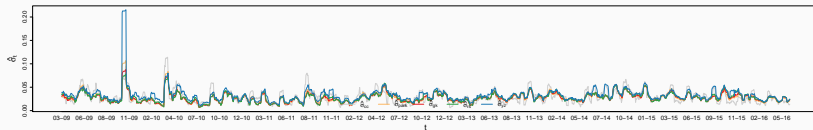


Figure 3: Volatilidad de COME.

Observaciones empíricas:

- **Asimetría** hacia la derecha, es decir el mercado se encuentra con mayor frecuencia en un estado de baja volatilidad.
- **Similitud entre estimadores**: siguen un mismo patrón y parecen congruentes entre sí.
- **Estimador de Cierre a Cierre**: tiende a mostrar valores más extremos en ambos sentidos.
- **Estimador de Yang-Zhang** se aleja de otras medidas para mostrar valores mayores al resto.

Heterocedasticidad Condicional

Se asume que el retorno logarítmico sigue un modelo simple para series de tiempo como ser un proceso estacionario autoregresivo de media móvil $ARMA(p, q)$ con variables explicativas

$$r_t = \mu_t + a_t, \quad (15)$$

$$\mu_t = \sum_{i=1}^p \phi_i \left(r_{t-i} - \phi_0 - \sum_{j=1}^k \beta_j x_{j,t-i} \right) - \sum_{i=1}^q \theta_i a_{t-i}, \quad (16)$$

donde k es la cantidad de covariables, p es el orden del proceso autoregresivo, q es el orden del proceso de media móvil, todos ellos enteros no negativos, y finalmente x_j son covariables.

La variancia condicional es

$$\sigma_t^2 = \mathbf{V}_{t-1} \langle r_t \rangle = \mathbf{V}_{t-1} \langle a_t \rangle, \quad (17)$$

donde a_t es el residuo de la ecuación para la media.

En el ámbito de la econometría, es habitualmente llamado **perturbación**, **innovación** o **shock** en el retorno de un activo al momento t .

Engle (1982) propone el primer **conjunto sistemático de herramientas** para modelar la volatilidad.

El modelo autoregresivo con heterocedasticidad condicional de orden m *ARCH*(m) descompone la perturbación en dos términos

$$a_t = \sqrt{h_t} \epsilon_t, \quad (18)$$

donde ϵ_t es una secuencia de variables aleatorias independientes e idénticamente distribuidas con $E \langle \epsilon_t \rangle = 0$ y $V \langle \epsilon_t \rangle = 1$.

La dinámica de la variancia condicional está dada por

$$h_t = \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2, \quad (19)$$

con $\alpha_0 > 0$ y $\alpha_i \geq 0 \forall i > 0$. Los coeficientes α_i deben respetar ciertas condiciones de regularidad para garantizar que la variancia marginal de las perturbaciones $V \langle a_t \rangle$ sea finita.

Logra reproducir parte de los hechos estilizados.

- **Clustering:** grandes perturbaciones pasadas implican una variancia condicional mayor para futuras observaciones.
- **Kurtosis condicional:** colas de los shocks son pesadas pues el exceso de kurtosis marginal es positivo incluso cuando ϵ_t sigue una distribución gaussiana. El exceso no es suficiente para reproducir la evidencia empírica.

Aun así, tiene ciertas limitaciones.

- **Simetría:** la reacción de la volatilidad a la perturbación es simétrica.
- **Lentitud:** cierta lentitud para responder a perturbaciones de magnitud aislada.

El modelo ARCH es simple y habitualmente requiere un orden m elevado para replicar el proceso de la volatilidad.

De forma independiente, Bollerslev (1986) y Taylor (1986) proponen un modelo más parsimonioso.

La dinámica de la volatilidad condicional está dada por

$$h_t = \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j}, \quad (20)$$

con $\alpha_0 > 0$, $\alpha_i \geq 0 \forall i > 0$, $\beta_j \geq 0 \forall j$ y $\sum_{i=1}^{\max(m,s)} \alpha_i + \beta_i < 1$ con $\alpha_i = 0 \forall i > m$ y $\beta_i = 0 \forall i > s$.

Equivalencias:

- $GARCH = ARCH(m)$ para $m = \infty$.
- $GARCH(m) = ARCH(m)$ cuando $s = 0$.
- ARMA para la serie del cuadrado de las perturbaciones al definir $\eta_t = a_t^2 - h_t$ y reemplazar h_t .

- **Decaimiento exponencial:** la función de autocorrelación del cuadrado de la perturbación, si existe, decae lenta y exponencialmente.
- **Parsimonia:** ARCH muestra una tasa de disminución muy rápida respecto de lo observado en las series de tiempo financieras, salvo que el rezago máximo m sea muy grande.
- **Hechos estilizados:** replica los bloques de volatilidad y el exceso de kurtosis.
- **Fácil de explicar:** provee una función determinística simple que describe la evolución de la volatilidad.

Desventajas:

- **Simetría:** la respuesta de la volatilidad a los shocks es simétrica.
- **Kurtosis:** el exceso de kurtosis cuando ϵ_t sigue una distribución gaussiana no refleja la leptokurtosis observada empíricamente.

Nelson (1991) propone GARCH Exponencial (EGARCH).

$$\log h_t = \alpha_0 + \sum_{i=1}^m \alpha_i \frac{|a_{t-i}| + \gamma_i a_{t-i}}{\sqrt{h_{t-i}}} + \sum_{j=1}^s \beta_j \log h_{t-j}. \quad (21)$$

Emplea el logaritmo natural de la variancia condicional:

- **Conveniencia matemática:** garantiza valores positivos de la variancia.
- **Asimetría:** función no lineal que permite responder en forma distinta según el signo de las perturbaciones rezagadas.

$$\begin{cases} \exp(\alpha_i(1 + \gamma_i)a_{t-i}h_{t-i}^{-0.5}) & \text{cuando } a_{t-i} \geq 0 \\ \exp(\alpha_i(1 - \gamma_i)a_{t-i}h_{t-i}^{-0.5}) & \text{cuando } a_{t-i} < 0. \end{cases} \quad (22)$$

El parámetro γ_i captura el **efecto apalancamiento**. En función de los hechos estilizados estudiados se espera que $\gamma_i < 0$, de forma que las **malas noticias incrementen la volatilidad**.

- **GJRGARCH** (Glosten, Jagannathan, and Runkle 1993)

$$h_t = \alpha_0 + \sum_{i=1}^m (\alpha_i + \gamma_i S_{t-i}^-) a_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j}, \quad (23)$$

- El parámetro γ_i captura el **efecto apalancamiento**.

- **TGARCH** (Zakoian 1994)

$$\sqrt{h_t} = \alpha_0 + \sum_{i=1}^m (\alpha_i + \gamma_i S_{t-i}^-) |a_{t-i}| + \sum_{j=1}^s \beta_j \sqrt{h_{t-j}}. \quad (24)$$

- **Desvío estándar** condicional.

- **NGARCH** (Higgins and Bera 1992)

$$h_t^\delta = \alpha_0 (\sigma^2)^\delta + \sum_{i=1}^m \alpha_i (a_{t-i}^2)^\delta, \quad (25)$$

- El parámetro δ funciona como la **transformación de potencias** de Box-Cox.

- **Pronóstico óptimo** bajo un modelo $ARMA(p, q) - GARCH(m, s)$:

$$\hat{h}_{t+j|t} = \hat{\alpha}_0 \sum_{i=0}^{j-2} (\hat{\alpha}_1 + \hat{\beta}_1)^i + (\hat{\alpha}_1 + \hat{\beta}_1)^{j-1} \hat{h}_{t+1}. \quad (26)$$

- **Volatilidad determinística:** Si los parámetros del modelo son conocidos, las perturbaciones a_t y las variancias condicionales h_t pueden ser reconstruidas perfectamente.
- **Invariante** al supuesto de homocedasticidad o heterocedasticidad condicional (Baillie and Bollerslev 1992).

La amplitud del intervalo de pronóstico para la media varía a lo largo del tiempo. Los períodos de **mayor incertidumbre** están asociados a intervalos de mayor apertura.

Para un coeficiente de $1 - \alpha$, el intervalo de pronóstico está dado por

$$\left(\hat{r}_{t+j} - z_{1-0.5\alpha} \sqrt{V \langle e_{t+j|t} \rangle}, \hat{r}_{t+j} + z_{1-0.5\alpha} \sqrt{V \langle e_{t+j|t} \rangle} \right) \quad (27)$$

donde z_p es el cuantil p de la función de densidad supuesta para ϵ_t y $e_{t+j|t} = r_{t+j} - \hat{r}_{t+j|t}$ es el error de pronóstico.

- **Convergencia no resulta monótona** a diferencia de lo que sucede al suponer homocedasticidad.

Los modelos presentados suponen que ϵ_t es una secuencia de variables aleatorias independiente e idénticamente distribuidas con $E \langle \epsilon_t \rangle = 0$ y $V \langle \epsilon_t \rangle = 1$.

- **Distribuciones habituales:** gaussiana, t de Student (Bollerslev 1987) y distribución generalizada del error (Nelson 1991).
- **Otras distribuciones:** mezcla de normal y poisson (Jorion 1988), mezcla de normal y lognormal (Hsieh 1989), gaussiana inversa (Venter and Jongh 2002).
- Parametrizaciones simétricas y asimétricas (Fernández and Steel 1998).

Son modelos **no lineales** en sus parámetros.

- **Desafíos:**

- Función de verosimilitud con máximos locales y zonas planas.
- Variantes asimétricas suman dimensiones al espacio paramétrico.

- **Técnicas:**

- Condicionar la primera observación (Zivot 2009).
- **Métodos numéricos** como Newton-Raphson (Berndt et al. 1974; Liu and Nocedal 1989).
- **Cuasi-verosimilitud** (Weiss 1984; Weiss 1986; Bollerslev and Wooldridge 1992; Glosten, Jagannathan, and Runkle 1993).²
- **Estimador de Huber** o sandwich (Bollerslev and Wooldridge 1992).
- Elección del algoritmo de optimización, los valores iniciales y los criterios de convergencia son muy influyentes (Fiorentini, Calzolari, and Panattoni 1996; McCullough and Renfro 1998; Brooks, Burke, and Persaud 2001).

²Bajo ciertos supuestos, la optimización de la log verosimilitud gaussiana permite arribar a una estimación consistente y asintóticamente normal para perturbaciones no gaussianas.

Aplicación al Mercado de Acciones en Argentina

¿Los diferentes modelos diseñados para dar respuesta a los hechos estilizados **mejoran la capacidad de predicción** respecto de un modelo base parametrizado como $ARMA(1,1)$ – $GARCH(1,1)$ con distribución gaussiana para el error aleatorio ?

Se filtra cada serie de tiempo a través de un modelo de heterocedasticidad condicional.

- **Media:** $ARMA(1, 1)$.
- **Volatilidad:** GARCH, EGARCH, GJRGARCH, TGARCH y NGARCH $m = 1$ y $s = 1$.
- **Densidad:** gaussiana, t de Student y generalizada del error, en sus versiones simétricas y asimétricas.
- **Estimación:** máxima verosimilitud empleando rugarch (Ghalanos 2015) en R 3.3.1 (R Core Team 2016).

Se genera un pronóstico fuera de la muestra de la volatilidad $\sqrt{\hat{h}_{t+j|t}}$ para $j = 1$ paso hacia adelante.

- **Iteraciones:** 500 veces.
- **Ventana móvil:** aproximadamente 1250 observaciones de entrenamiento.
- **Estimación:** en cada iteración.

Los procesos de diagnóstico de los residuos y evaluación de pronósticos son realizados para cada serie.

- **Residuos estandarizados.**
 - **Distribución:** estimación puntuales de los primeros cuatro momentos, prueba de Jarque-Bera, bondad de ajuste χ^2 con $k = 50$ celdas, histograma, gráfico de cuantil-cuantil.
 - **Modelo de la media:** prueba de Ljung-Box aplicada a los residuos en los primeros 10 y 22 rezagos, correlogramas.
 - **Modelo de la volatilidad:** prueba de Ljung-Box aplicada al cuadrado de los residuos en los primeros 10 y 22 rezagos, correlogramas.
 - **Asimetría:** pruebas de Engle and Ng (1993), curva de impacto de noticias.
- **Pronósticos fuera de la muestra.**
 - **Direccional:** razón de acierto para movimientos positivos, negativos y medios.
 - **Error de predicción:** Error Cuadrático Medio, Cuasi-Verosimilitud y R^2 de Mincer-Zarnowitz.
 - **Pronóstico por intervalos:** cantidad de observaciones fuera de la muestra que exceden el VaR al 1% contra la cantidad esperada.

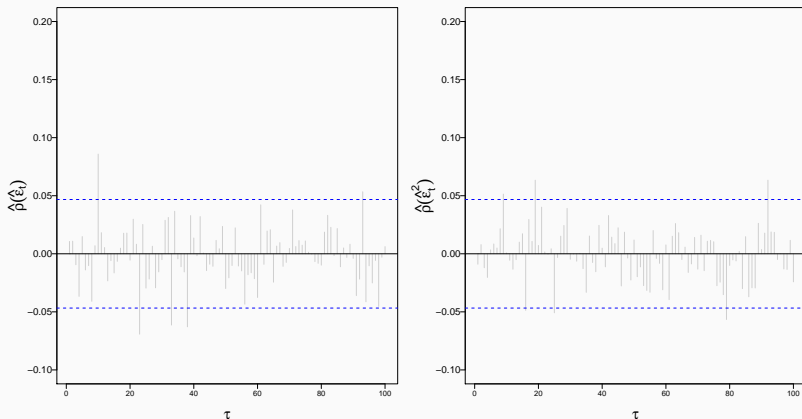


Figure 4: Autocorrelación de residuos (COME). Izquierda: Autocorrelograma para los residuos. Derecha: Autocorrelograma para el cuadrado de los residuos. Elaboración propia.

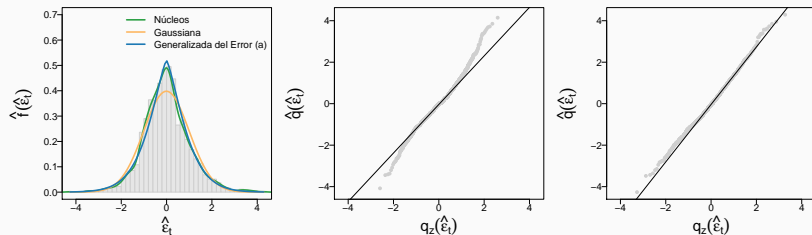


Figure 5: Distribución de residuos (ERAR). Izquierda: Histograma de los retornos logarítmicos. Centro: gráfico de cuantil-cuantil de los retornos logarítmicos para una distribución gaussiana. Derecha: gráfico de cuantil-cuantil de los retornos logarítmicos para una distribución generalizada del error asimétrica. Elaboración propia.

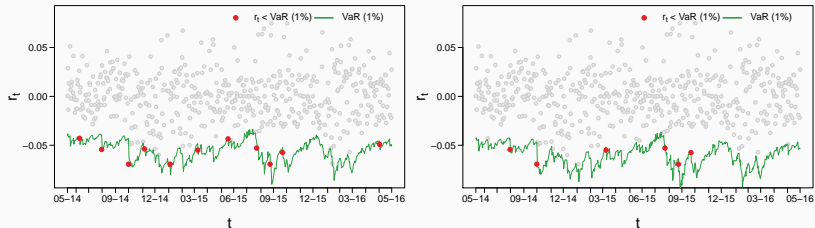


Figure 6: Retornos observados y *Value At Risk* al 1% predicho en el momento anterior (TS). En rojo, aquellos retornos que exceden el límite estimado. Izquierda: Modelo base con error gaussiano. Derecha: Modelo alternativo con Distribución Generalizada del Error. Elaboración propia.

- **Hipótesis de normalidad:** en los residuos del modelo base no se cumple.
- **Asimetría:** presente en los residuos de nueve de las doce series analizadas. Sólo tres casos con marcada asimetría.
- **Kurtosis:** todas la series son leptocúrticas. Once casos con marcada leptokurtosis.
- **Autocorrelación:** captura satisfactoriamente el efecto de la autocorrelación lineal en los retornos y el cuadrado de las perturbaciones. Sólo un caso con falla en el modelo de la media.
- **Efecto apalancamiento:** muy difundido en los trabajos de investigación para otros mercados, es observado sólo en circunstancias muy particulares. Tres casos sugieren la existencia de algún efecto individual, pero sólo uno presenta evidencia para un efecto conjunto.

La leptokurtosis se impone como la característica más determinante y se constituye como la principal falla al supuesto de densidad gaussiana.

- **Selección de modelos**

- Modelos GARCH en diez casos, EGARCH en dos casos (uno de ellos asociado al efecto apalancamiento).
- Distribución generalizada del error en sus versiones asimétrica y simétrica y la distribución t de Student en seis, cinco y una oportunidad respectivamente.
- Sólo en la mitad de los casos la asimetría resulta ser lo suficientemente marcada para compensar la la pérdida de grados de libertad.

- **Direccional:** sólo en un caso³.
- **Error medio:**
 - Ocho casos muestran un desempeño similar.
 - COME y PAMP incrementan su precisión,
 - Se destaca COME gracias a una función de densidad que modela en simultáneo leptokurtosis y asimetría.
 - Dos individuos con peor desempeño producto de la sobreparametrización del modelo seleccionado por los criterios de información.
 - **YPFD:** modelo eGARCH con distribución generalizada del error cuando los retornos no muestran asimetría.
 - **APBR:** modelo con distribución t de Student aunque los retornos no poseen colas pesadas.

³Caso que además muestra asimetría, un indicio que estos modelos mejorarían la predicción de la dirección para acciones con retornos asimétricos.

- **Amplitud del Intervalo:**

- Los modelos mejoran la cantidad de excesos al VaR en siete casos (BMA, COME, FRAN, GGAL, Merval, TECO2, TS).
- Uso de una función de densidad que se adapta a colas más pesadas y, en tres casos, asimetría.
- **YFPD** y **APBR** empeoran por sobreajuste.

- **BIC:** en los casos donde los criterios no seleccionan un mismo modelo, el BIC se orienta en favor del más natural para la práctica de las finanzas.
- **Sobreajuste:** la penalización observada no siempre se mostró suficiente para evitar el sobreajuste (YPFD y APBR).
- **Conclusiones:** no son unívocas!
 - Estadísticos pueden arrojar resultados mixtos sobre una hipótesis.
 - Preferencias a la hora de elegir la función de pérdida para evaluar los pronósticos.

- Especial atención a las decisiones que hacen al preprocesamiento de los datos (ej. truncado).
- No resulta necesario incrementar el orden más allá de $GARCH(1, 1)$.
- Foco selección de la función de densidad.
- Leptokurtosis tiene más peso que la asimetría.
- La elección de una ecuación no lineal para la volatilidad no es prioritario.
- El sobreajuste es el principal riesgo.
- La principal mejora se observa en la amplitud de los intervalos de pronóstico.
- Si el interés yace en el pronóstico puntual, debe evaluarse otras estrategias como el uso de regresores exógenos.

Bibliografía

- Aggarwal, Reena, Carla Inclan, and Ricardo Leal. 1999. "Volatility in Emerging Stock Markets." *Journal of Financial and Quantitative Analysis* 34 (01). Cambridge Univ Press: 33–55.
- Akgiray, Vedat. 1989. "Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts." *Journal of Business*. JSTOR, 55–80.
- Baillie, Richard T, and Tim Bollerslev. 1992. "Prediction in Dynamic Models with Time-Dependent Conditional Variances." *Journal of Econometrics* 52 (1-2). Elsevier: 91–113.
- Berndt, Ernst R, Bronwyn H Hall, Robert E Hall, and Jerry A Hausman. 1974. "Estimation and Inference in Nonlinear Structural Models." In *Annals of Economic and Social Measurement, Volume 3, Number 4*, 653–65. NBER.
- Bollerslev, Tim. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31 (3). Elsevier: 307–27.
- . 1987. "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return." *The Review of Economics and Statistics*. JSTOR, 542–47.
- Bollerslev, Tim, and Jeffrey M Wooldridge. 1992. "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances." *Econometric Reviews* 11 (2). Taylor & Francis: 143–72.
- Brandt, Michael W, and Francis X Diebold. 2003. "A No-Arbitrage Approach to Range-Based Estimation of Return Covariances and Correlations." National Bureau of Economic Research.
- Brandt, Michael W, and Jonathan Kinlay. 2005. "Estimating Historical Volatility." *Research Article, Investment Analytics*.
- Brooks, Chris, Simon P Burke, and Gita Persaud. 2001. "Benchmarks and the Accuracy of Garch Model Estimation." *International Journal of Forecasting* 17 (1). Elsevier: 45–56.
- Daniélsson, Jón. 2011. *Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk with Implementation in R and Matlab*. Vol. 588. John Wiley & Sons.
- Ding, Zhuixin, Clive WJ Granger, and Robert F Engle. 1993. "A Long Memory Property of Stock Market Returns and a New Model." *Journal of Empirical Finance* 1 (1). Elsevier: 83–106.
- Engle, Robert F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation."