# Modelos de precios y backtesting

Gabriel Vergara Schifferli

19 Octubre 2021

Modelos de precios para considerar en estructura dinámica en el sistema de control aplicado a la gestión de riesgo.

# Agenda

- Modelos de series de tiempo
  - Procesos de memoria larga
  - Modelo de volatilidad estocástica
  - Markov-Switching GARCH
  - Modelos de mezclas de distribución
- Backtesting en métricas de riesgo
  - Test para Value at Risk
  - Test para Expected Shortfall y VaR conjunto
  - Expectiles*
- Comparación de modelos: *Model Confidence Set*

# ARIMA

Considerando la serie de precios logarítmica $p_t = ln(P_t)$
y el operador de rezago $B: B(p_t) = p_{t-1},\ B^n(p_t) = p_{t-n}$

**ARMA(p,q)**

considerando el proceso $\{p_t\}$

$$p_t - \phi_1 p_{t-1} - \cdots - \phi_p p_{t-p} = \varepsilon_t + \theta_1 \varepsilon_t + \cdots + \theta_q \varepsilon_{t-q},$$

donde $\varepsilon \sim WN(0, \sigma^2)$.

Considerando

$$\Phi(B) = (1 - \phi_1 B - \cdots - \phi_p B^p)$$

$$\Theta(B) = (1 + \theta_1 B + \cdots + \theta_q B^q)$$

Se escribe de forma compacta el proceso ARMA(p,q)

$$\Phi(B) p_t = \Theta(B) \varepsilon_t$$

# ARIMA

Considerando el operador de diferenciación:

$$\nabla x_t = x_t - x_{t-1} \Longrightarrow \nabla^k x_t = \nabla(\nabla^{k-1} x_{t-1})$$

$$\nabla^k = (1-B)^k = \sum_{j=1}^{k} \binom{n}{k} B^k$$

**ARIMA(p,d,q)**

el proceso $p_t$ sigue un modelo ARIMA(p,d,q) si el proceso diferenciado $\nabla^d p_t \sim ARMA(p,q)$

$$\Phi(B)\nabla^d p_t = \Theta(B)\varepsilon_t$$

# ARFIMA

Son procesos de memoria larga, donde $\gamma(h) = cov(p_t, p_{t+h})$ no es absolutamente sumable.

Otra definición: que tiene un decaimiento hiperbólico de la autocovarianza : $\gamma(h) \sim h^{2d-1}\ell_1(h)$ según $h \to \infty$, con $d$ el parámetro de memoria larga y $\ell(\cdot)$ una función de variación lenta.

**ARFIMA(p,d,q)**

un proceso $p_t$ sigue un modelo ARFIMA si:

$$\Phi(B)p_t = \Theta(B)\nabla^{-d}\varepsilon_t$$

donde $\Phi$ y $\Theta$ no tienen raíces comunes y $\nabla^{-d}$ es un operador de diferenciación fraccionario:

$$\nabla^{-d} = \sum_{j\geq 0} \eta_j B^j = \eta(B)$$

con

$$\eta_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)},$$

para $0 < d < 0.5$, y $\varepsilon_t \sim WN(0, \sigma^2)$

# Modelos de volatilidad

**EGARCH (Nelson (1991))**

*exponential GARCH*

$$ln(\sigma_t^2) = w + \sum_{i=1}^{p} \left\{\alpha_i \left(|\varepsilon_{t-i}| - E\left[|\varepsilon_{t-i}|\right]\right) + \gamma_i \varepsilon_{t-i}\right\} + \sum_{j=1}^{q} \beta_j ln(\sigma_{t-j}^2)$$

**GJRGARCH (Glosten et al. (1993))**

$$\sigma_t^2 = w + \sum_{i=1}^{p} [\alpha_i + \gamma_i I_{\varepsilon_{t-i}>0}]\varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

**APARCH (Ding, Granger y Engle 1993)**

*asymmetric power ARCH*

$$\sigma_t^\delta = \quad \alpha_0 \sum_{i=1}^{r} \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^{s} \beta_j \sigma_{t-i}^\delta$$

**TGARCH (Zakoian (1994))**

*Threshold GARCH*

$$\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i \left[(1-\gamma_i)\varepsilon_{t-i}^+ + (1+\gamma_i)\varepsilon_{t-i}^-\right] + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

**RCAR**

*random coefficient autoregressive model*

**CHARMA**

*conditional heteroscedastic autorregressive moving average model*

**GARCH(1,1)-M**

*garch in mean*

$$r_t = \mu + c\sigma_t^2 + a_t, \quad a_t = \sigma_t \varepsilon_t$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

$c$ se le llama *risk premium parameter*, el retorno se relaciona positivamente con su volatilidad pasada.

**ZD-GARCH (Li, Zhang, Zhu and Ling (2018))**

*Zero drift GRACH*

$$\sigma_t^2 = \alpha a_{t-1}^2 + \beta \sigma_{t-1}^2$$

- Extender la propiedad de memoria larga del ARFIMA a la volatilidad.

**FIGARCH(p,d,q)**

$$
\begin{aligned}
a_t = &\quad \sigma_t \varepsilon_t \\
\sigma_t^2 = &\quad \omega + \beta(L)\sigma_t^2 + [1 - \beta(L) - [1 - \phi(L)](1 - L)^d]a_t^2
\end{aligned}
$$

donde $L$ es el operador de rezago, $\beta(L) = \sum_k^q \beta_k L^k$, $\phi(L) = \sum_k^q \phi_k L^k$ y $[1 - \beta(L)], [1 - \phi(L)]$ tienen raíces fuera del círculo unitario con $(1 - L)^d$ el operador de diferenciación fraccionario para el parámetro de memoria larga $0 < d < 1$.

# Time-Varying FIGARCH Boutahary y Belkhouja 2010

- Incorporar la propiedad de memoria larga a la volatilidad
- Considerar cambios en la dinámica de la volatilidad

**TV-FIGARCH**

$$\begin{aligned} a_t = &\quad \sigma_t \varepsilon_t \\ \sigma_t^2 = &\quad \omega + \beta(L)\sigma_t^2 + [1 - \beta(L) - [1 - \phi(L)](1 - L)^d] a_t^2 + f_t \\ f_t = &\quad \sum_{r=1}^{R} \omega_r F_r(s_t, \gamma_r, c_r) \end{aligned}$$

Donde $F_r(s_t, \gamma_r, c_r)$ son funciones de transición determinadas por cambios de régimen.

Una elección apropiada para sería una función de transición logística

$$F_r(s_t, \gamma_r, c_r) = (1 + exp\{-\gamma_r(s_t - c_r)\})^{-1}$$

con $\gamma_r > 0$ parámetro que controla el grado de suavidad, $c_r$ parámetro de umbral y $s_t = t/T$ la variable de transición.

- Surge al introducir un término estocástico en la ecuación de volatilidad

**Stochastic Volatility**

$$a_t = \sigma_t \varepsilon_t \quad ln(\sigma_t^2) = \alpha_0 + \sum_i^m \alpha_i ln(\sigma_{t-i}^2) + \nu_t$$

con $\varepsilon_t \sim^{iid} N(0,1)$, $\nu_t \sim^{iid} N(0,\sigma_\nu^2)$, $\nu_t \perp \varepsilon_t$

**Long memmory stochastic volatility**

$$a_t = \sigma_t \varepsilon_t, \quad \sigma_t = \sigma exp(u_t/2), \quad (1-L)^d u_t = \eta_t$$

$\sigma > 0$, $\varepsilon_t \sim^{iid} N(0,1)$, $\eta_t \sim^{iid} N(0,\sigma_\eta^2)$ $\varepsilon_t \perp \eta_t$, $0 < d < 0.5$

# Markov Switching GARCHs

GARCH con cambio estructural regido por una cadena de Markov no observable.

$$r_t|(S_t = k, \mathcal{F}_{t-1}) \sim F(0, \sigma^2_{k,t}, \theta_k)$$

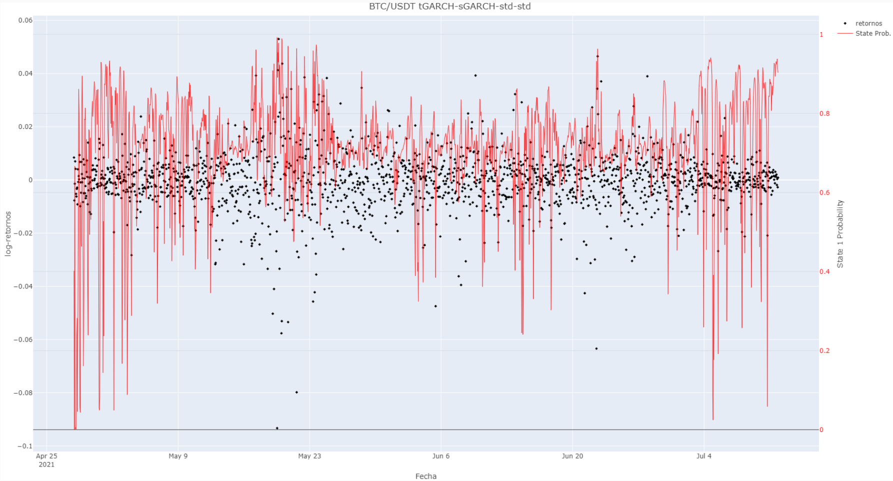
con $F(0, \sigma^2_{k,t}, \theta_k)$ distribución con media cero, varianza variando en el tiempo $\sigma^2_{k,t}$ y parámetros $\theta_k$

**MS-GARCH**

$$\begin{aligned} a_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_{0,s_t} + \sum_{i=1}^{u} \alpha_{i,s_t} a^2_{t-i} + \sum_{j=1}^{v} \beta_{j,s_t} \sigma^2_{t-j} \end{aligned}$$

done $s_t$ es una cadena de Markov ergódica en un conjunto finito $S = \{1, 2, \ldots, n\}$ con una probabilidad de transición $\{\eta_{ij} = \mathbf{P}(s_t = i | s_{t-1} = j)\}$ y una medida de probabilidad invariante $\{\pi_i\}$:
$P(s_{t+1} = i | s_t = j, s_{t-1} = j_1, s_{t-2} = j_2 ...) = P(s_{t+1} = i | s_t = j) = \eta_{ij}$

- Se tienen $n$ regímenes distintos
- Cada régimen tiene estructura distinta
  *e.g.* tGARCH-GARCH


BTC/USDT tGARCH-sGARCH-std-std
retornos
State Prob.
log-retornos
State 1 Probability
Fecha
0.06
0.04
0.02
0
−0.02
−0.04
−0.06
−0.08
−0.1
1
0.8
0.6
0.4
0.2
0
Apr 25
2021
May 9
May 23
Jun 6
Jun 20
Jul 4

$$r_t \sim DM(p1, \ldots, pk; \mu_1, \ldots, \mu_k; h_1, \ldots, h_k)$$

Donde *DM* es una mezcla de densidades con la siguiente forma:

$$\vartheta(x) = \sum_i^k p_i f_i(x), \quad \sum p_i = 1, \quad f_i(x) = f(y; \mu_i; h_i)$$

con $[p_1, \ldots, p_k]$ la regla de mezcla de densidades.

Se puede ver como un modelo de Markov-Switching GARCH, en el cual las probabilidades de transición son independientes del estado anterior.

**K-GARCH(1,1)**

Considerando mezclas de normales,

$$\begin{aligned} \sigma_{i,t}^2 = &\quad w_i + \alpha_i \varepsilon_{t-1}^2 + \beta_i \sigma_{i,t-1}^2, \quad \forall\ i = 1, \ldots, K \\ \sigma_t^2 = &\quad \sum p_i \sigma_{i,t}^2 + \sum p_i \mu_i^2 \end{aligned}$$

$\Rightarrow$ se mezclan distintos tipos de garch, con distintas distribuciones
*e.g. eGARCH-tGARCH-ged-std*

# Ejemplo modelo de precios

considerando $r_t$ la serie de retornos logarítmicos

**$r_t \sim AR(2) - GARCH(1,1)$**

$$\begin{aligned} r_t = &\quad \mu + \phi_1 r_{t-1} + \phi_2 r_{t-2} + a_t \\ a_t = &\quad \sigma_t \varepsilon_t \\ \sigma_t^2 = &\quad \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \beta_1 a_{t-1}^2 \\ &\quad \varepsilon_t \sim N(0,1) \end{aligned}$$

Entonces, los predictores quedan :

$$E(r_{t+1}|\mathcal{F}_{t-1}) = r_t(1) = \mu + \phi r_t + \phi_2 r_{t-1}$$

$$E(\sigma_{t+1}^2|\mathcal{F}_{t-1}) = \sigma_t^2(1) = \alpha_0 + \alpha_1 a_t^2 + \beta_1 \sigma_t^2$$

Luego,

$$\hat{r}_{t+1} \sim N(r_t(1), \sigma_t^2(1))$$

$\Longrightarrow$ se tiene una estructura para estimar predictores de VaR y ES por ejemplo.

# Backtesting

*Para evaluar el desempeño de los modelos de riesgo, las medidas de riesgo proyectadas son comparadas con las pérdidas realizadas sobre un periodo de tiempo y su correcta especificación es testeada estadísticamente.*

### Value at Risk

Pérdida máxima esperada con una probabilidad $\alpha$ sobre un periodo:

$$VaR_\alpha(X) = \inf\{x : F_X(x) \geq \alpha\} = F_X^{-1}(\alpha)$$

Para realizar un test sobre las predicciones de 1 paso sobre el VaR se construye una sucesión de aciertos para las violaciones del VaR:

$$\begin{aligned} I_{t+1} &= 1, \quad si \quad r_{t+1} < VaR_{t+1}^{\alpha} \\ I_{t+1} &= 0, \quad si \quad r_{t+1} > VaR_{t+1}^{\alpha} \end{aligned}$$

donde $VaR_{t+1}^{\alpha}$ es el predictor del VaR para el tiempo $t+1$ a un nivel $\alpha$.

Bajo la hipótesis de una correcta especificación del modelo, la sucesión de aciertos debe ser independiente y seguir una distribución Bernoulli.

$$H_0 : I_{t+1} \sim Bernoulli(\alpha), \tag{1}$$

$$f(I_{t+1}, p) = (1-p)^{1-I_{t+1}} p^{I_{t+1}} \tag{2}$$

**Objetivo:**
Testear si la cantidad de retornos menores al $VaR_\alpha$ predicho corresponde al nivel $\alpha$ (*portion of failures POF*).

Para $\{I_t\}_t$ anterior,

$$L(\pi) = \prod (1-\pi)^{1-I_{t+1}} \pi^{i_{i+1}} = (1-\pi)^{T_0} \pi^{T_1}$$

se tiene que

$$\hat{\pi}_{MLE} = \frac{T_1}{T_0 + T_1}$$

Bajo $H0 : \pi = \alpha$

$$LR_{uc} = -2 \ln \left[ \frac{L(\alpha)}{L(\hat{\pi})} \right] \sim \chi^2_1, \tag{3}$$

## Conditional Coverage test (CC) Christoffsen (1998)

**Objetivo:**
Ver que la serie corresponde a la cantidad de aciertos y no se generan clusters de retornos menores al VaR predicho.

Considerando que $\{I_t\}_t$ presenta una dependencia temporal y sigue una sucesión de Markov con matriz de probabilidad de transición:

$$\Pi_1 = \begin{pmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{pmatrix}, \tag{4}$$

donde $\pi_{01}$ es la probabilidad de transición de 0 a 1, con verosimiltud:

$$L(\Pi_1) = (1-\pi_{01})^{T_{00}} \pi_{01}^{T_{01}} (1-\pi_{11})^{T_{10}} \pi_{11}^{T_{11}}, \tag{5}$$

Si la $\{I_t\}_t$ es independiente del tiempo, se tiene que $\pi_{01} = \pi_{11} = \pi$, entonces:

$$\hat{\Pi}_{1_{MLE}} = \begin{pmatrix} 1-\hat{\pi} & \hat{\pi} \\ 1-\hat{\pi} & \hat{\pi} \end{pmatrix}, \quad LR_{cc} = -2\ln\left[\frac{L(\alpha)}{L(\hat{\Pi}_1)}\right] \sim \chi^2_1$$

**Objetivo:**

Testear que no exista capacidad predictiva sobre las excepciones del VaR predicho.

Consideramos la suseción $Hit_t(\alpha) = I_t(\alpha) - \alpha$, bajo la correcta especificación del modelo se tienen las condiciones de momentos: $E[Hit_t(\alpha)] = 0$, $E[Hit_t(\alpha)I_{t-1}] = 0$, $E[Hit_t(\alpha)Hit_{t'}(\alpha)] = 0 \ \forall t \neq t'$ por lo tanto, se tiene el siguiente modelo de regresión:

$$Hit_t(\alpha) = \delta + \sum_{k=1}^{K} \beta_k Hit_{t-k}(\alpha) + \sum_{k=1}^{K} \gamma_k g[\mathcal{F}_{t-k}] + \varepsilon_t, \tag{6}$$

Entonces, bajo la hipótesis nula de la eficiencia condicional es equivalente a testear si los coeficientes son conjuntamente nulos:

$$H_0 : \delta = \beta_1 = \cdots = \beta_k = \gamma_1 = \cdots = \gamma_k = 0, \ \forall k = 1, \ldots, K \tag{7}$$

- VaR no es capaz de medir pérdidas más allá de un nivel $\alpha$.
- ES mide el riesgo de cola que no captura el VaR.

**Expected Shortfall**

ES corresponde a la pérdida esperada dado que la pérdida excede el $\alpha-$cuantil del VaR:

$$ES_{\gamma}(X) = \frac{1}{\gamma}\int_{0}^{\gamma} VaR_{\alpha}(X)d\alpha \tag{8}$$

donde $ES_{\gamma}(X)$ es el *Expected Shortfall* de la v.a. $X$ a un nivel $\gamma$, esto es, dada una pérdida $r$ se tiene que $ES = E[r|r < VaR]$.

# Exceedance Residuals test (ER) McNeil y Frey (2000)

**Objetivo:**

Se testea si el valor esperado de ER es cero contra la hipótesis unilateral de que son mayores que cero, esto es, que el ES esta sistemáticamente siendo subestimado

Considerando

$$\left\{ R_{t+1} : t \in T, r_{t+1} < \hat{VaR}_{\alpha}^{t+1}, \quad R_{t+1} = \frac{r_{t+1} - \hat{ES}_{\alpha}^{t+1}}{\hat{\sigma}_{t+1}} \right\}$$

las excedencias de residual.

Bajo $H0$ se estima correctamente la dinámica de $(\mu_{t+1} \ y \ \sigma_{t+1})$ y el $ES$.

Por argumentos asintóticos de *Teoría de valores extremos (EVT)* estos residuos deberían comportarse como iid con media cero.

## Conditional Calibration test Nolde y Ziegel (2017)

Aplicado a un par $(VaR, ES)$ a un nivel $\alpha$, considerando la función de identificación estricta:

$$V(Y, \nu, e) = \begin{pmatrix} \alpha - \mathbb{1}_{\{Y \leq \nu\}} \\ e - \nu + \mathbb{1}_{\{Y \leq \nu\}}(\nu - Y)/\alpha \end{pmatrix}$$

donde $E(V(Y, \nu, e)) = 0 \iff \nu$ y $e$ son los verdaderos VaR y ES de $Y$.

Entonces se tienen los test:

$$H_0^{2s} : E(V(Y_t, \hat{\nu}_t, \hat{e}_t)|\mathcal{F}_{t-1}) = 0 \quad v/s \quad H_1^{2s} : E(V(Y_t, \hat{\nu}_t, \hat{e}_t)|\mathcal{F}_{t-1}) \neq 0$$

$$H_0^{1s} : E(V(Y_t, \hat{\nu}_t, \hat{e}_t)|\mathcal{F}_{t-1}) \geq 0 \quad v/s \quad H_1^{1s} : E(V(Y_t, \hat{\nu}_t, \hat{e}_t)|\mathcal{F}_{t-1}) < 0$$

con el estadístico tipo Wald

$$T_{CC} = T \left( \frac{1}{T} \sum_{t=1}^{T} \mathbf{h}_t V(Y_t, \hat{\nu}_t, \hat{e}_t) \right)^T \hat{\Delta}_T^{-1} \left( \frac{1}{T} \sum_{t=1}^{T} \mathbf{h}_t V(Y_t, \hat{\nu}_t, \hat{e}_t) \right) \sim \chi_q^2,$$

donde $\hat{\Delta}_T^{-1} = \frac{1}{T} \sum_{t=1}^{T} (\mathbf{h}_t V(Y_t, \hat{\nu}_t, \hat{e}_t))(\mathbf{h}_t V(Y_t, \hat{\nu}_t, \hat{e}_t))^T$.

# Conditional Calibration test Nolde y Ziegel (2017)

**Objetivo:**

Se testea conjuntamente (VaR, ES) la correcta especificación del modelo y en cierto sentido la volatilidad $\sigma$.

***simple* CC test**

Utiliza solamente los predictores de VaR y ES.

$$\mathbf{h}_t = I_2$$

para el test unilateral y bilateral.

***general* CC test**

Utiliza los predictores de VaR, ES y la volatilidad $\hat{\sigma}_t$

$$\mathbf{h}_t = \hat{\sigma}_t((\hat{e}_t - \hat{\nu}_t)/\alpha, 1), \quad y \quad \mathbf{h}_t = \begin{pmatrix} 1 & |\hat{\nu}_t| & 0 & 0 \\ 0 & 0 & 1 & \hat{\sigma}_t^{-1} \end{pmatrix}$$

para el test unilateral y bilateral respectivamente.

**Objetivo:**

Testear la correcta especificación del modelo en cuanto a los predictores del ES.

Considerando los retornos $\{r_t, t = 1, \ldots, T\}$ y la serie de predictores a 1 paso de ES $\{\hat{e}_t, t = 1, \ldots T\}$ se tiene un esquema general de regresión de los retornos con el ES y un intercepto,

$$r_t = \gamma_1 + \gamma_2 \hat{e}_t + u_t^e,$$

donde $ES_\alpha(u_t^e | \mathcal{F}_{t-1}) = 0$ $c.t.p.$ o equivalentemente

$$ES_\alpha(r_t | \mathcal{F}_{t-1}) = \gamma_1 + \gamma_2 \hat{e}_t.$$

Considerando el test :

$$H_0 : (\gamma_1, \gamma_2) = (0, 1) \quad v/s \quad H_1 : (\gamma_1, \gamma_2) \neq (0, 1).$$

Bajo $H_0$, los predictores $\hat{e}_t$ están correctamente especificados y se tiene que $\hat{e}_t = ES_\alpha(r_t | \mathcal{F}_{t-1})$ $c.t.p.$

# Pero por problemas técnicos

Considerando una formulación más general $Y_t = W_t^T \gamma + u_t^e$, estimar el parámetro $\gamma$ no es posible pues no existe función de pérdida estrictamente consistente y funciones de identificación para el ES, pero sí se tiene conjuntamente para (VaR, ES).

Entonces, se consideran una regresión conjunta:

$$Y_t = V_t^T \beta + u_t^q, \quad Y_t = W_t^T \gamma + u_t^e,$$

donde $V_t, W_t$ son $\mathcal{F}_{t-1}-$medibles, y $VaR_\alpha(u_t^q|\mathcal{F}_{t-1}) = 0$ y $ES_\alpha(u_t^e|\mathcal{F}_{t-1}) = 0$ c.t.p.

Por lo que se proponen los siguientes tests considerando ($\hat{\nu} = VaR, \hat{e} = ES$):

# ESR Backtest Bayer y Dimitriadis (2018)

**Auxiliary ESR**

Se testea conjuntamente $(VaR, ES)$, considerando $V_t = (1, \hat{\nu}_t)$ y $W_t = (1, \hat{e}_t)$,

$$r_t = \beta_1 + \beta_2 \hat{\nu}_t + u_t^q, \quad r_t = \gamma_1 + \gamma_2 \hat{e}_t + u_t^e,$$

$$H_0 : (\gamma_1, \gamma_2) = (0, 1) \quad v/s \quad H_1 : (\gamma_1, \gamma_2) \neq (0, 1)$$

**Strict ESR**

Se testea solamente $ES$, considerando $V_t = W_t = (1, \hat{e}_t)$,

$$r_t = \beta_1 + \beta_2 \hat{e}_t + u_t^q, \quad r_t = \gamma_1 + \gamma_2 \hat{e}_t + u_t^e,$$

$$H_0 : (\gamma_1, \gamma_2) = (0, 1) \quad v/s \quad H_1 : (\gamma_1, \gamma_2) \neq (0, 1)$$

**Intercept ESR**

Se testea solamente ES, considerando el error de predicicón $r_t - \hat{e}_t$

$$r_t - \hat{e}_t = \beta_1 + u_t^q, \quad r_t - \hat{e}_t = \gamma_1 + u_t^e,$$

$$H_0^{2s} : \gamma_1 = 0 v/s \quad H_1^{2s} : \gamma_1 \neq 0$$
$$H_0^{1s} : \gamma_1 \geq 0 v/s \quad H_1^{1s} : \gamma_1 < 0$$

# Expectiles

Otra métrica de riesgo con buenas propiedades estadísticas y similar información al ES.

$\tau-$expectile $e_\tau(X)$ de $X$ con media finita, es la única solución $x = e_\tau(X)$ de la ecuación

$$\tau \int_x^\infty (y-x)dF_X(y) = (1-\tau)\int_{-\infty}^x (x-y)dF_X(y)$$

es una medida de riesgo obtenible y coherente para $\tau \in [1/2, 1)$, es una generalización de la esperanza.

Un articulo sobre métricas de riesgo, en cuanto a obtenibilidad y coherencia entre otras propiedades de las medidas VaR, ES y Expectiles:


Chen, J.M. (2018).
**On Exactitude in Financial Regulation: Value-at-Risk, Expected Shortfall, and Expectiles.**
*Risks 2018, 6, 61. https://doi.org/10.3390/risks6020061.*

# Model Confidence Set Hansen et al. (2011)

- Test anteriores no proporciona comparación entre modelos.
- Se busca construir un conjunto de modelos superiores que tienen misma capacidad predictiva.

**Objetivo:**

Determinar los modelos que tienen la misma capacidad predictiva dado una función de pérdida.

Se define la variable de rendimiento considerando la función de pérdida $L_{i,t} = L(Y_t, \hat{Y}_{i,t})$:

$$d_{ij,t} = L_{i,t} - L_{j,t} \quad \forall i, j \in M^0$$

El estadístico de pérdida relativa muestral:

$$\bar{d}_{ij} = \frac{1}{n}\sum_{t=1}^{n} d_{ij,t}, \quad \bar{d}_i = \frac{1}{m}\sum_{j \in M}^{n} \bar{d}_{ij}$$

## Model Confidence Set

Asumiendo que $\mu_{ij} = E(d_{ij,t}) < \infty$ e invariante en $t$ $\forall i, j \in \mathcal{M}^0$ se rankean en términos de la pérdida esperada, donde el modelo $i$ es preferido al $j$ si $\mu_{ij} < 0$

**Set of superior models (SSM)**

$$\mathcal{M}^* := \{i \in \mathcal{M}^0 : \mu_{ij} \leq 0 \forall j \in \mathcal{M}^0\}$$

Entonces, se busca determinar el conjunto $\mathcal{M}^*$ mediante una secuencia de tests de significancia , donde los conjuntos que son significativamente inferior a los demás elementos de $\mathcal{M}^0$ son eliminados. Esto es ,

**Equal predictive ability test (EPA)**

$$H_{0,\mathcal{M}} : \mu_{ij} = 0 \; \forall i, j \in \mathcal{M} \quad v/s \quad \mu_{ij} \neq 0$$

donde $\mathcal{M} \subset \mathcal{M}^0$, y el caso análogo para $\mu_i$

# Model Confidence Set

Luego se derivan los estadísticos t

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\hat{var}(\bar{d}_{ij})}}, \quad t_i = \frac{\bar{d}_i}{\sqrt{\hat{var}(\bar{d}_i)}} \; para \; i, j \in \mathcal{M}.$$

y las hipótesis EPA se resumen en los estadísticos de prueba:

$$T_{R,\mathcal{M}} = \max_{i,j \in \mathcal{M}} |t_{ij}|, \quad T_{max,\mathcal{M}} = \max_{i \in \mathcal{M}} t_i$$

con una regla de eliminación

$$e_{R,\mathcal{M}} = \arg\max_i \left\{ \sup_{j \in \mathcal{M}} \frac{\bar{d}_{ij}}{\sqrt{\hat{var}(\bar{d}_{ij})}} \right\}, \quad e_{max,\mathcal{M}} = \arg\max_{i \in \mathcal{M}} \frac{\bar{d}_i}{\sqrt{\hat{var}(\bar{d}_i)}}$$

**Algoritmo**

1. iniciar $\mathcal{M} = \mathcal{M}^0$.
2. Testear para EPA, si EPA no es rechazada terminar y dejar $\mathcal{M} = \mathcal{M}^*$, en otro caso, eliminar el peor modelo mediante la regla de eliminación y volver al paso 1.

# Funciones de pérdida

**Quantile Loss**

$$QL_t^{\alpha} = (\alpha - I_t(\alpha))(r_t - VaR_t^{\alpha})$$

**Bayer and Dimitriadis (2018)***

De una familia más general (Fissler y Ziegel 2016), una forma particular:
(funciones de scoring fuertemente consistentes para VaR y ES)

$$S(VaR_t^{\alpha}, ES_t^{\alpha}, r_t) = \frac{1}{-ES_t^{\alpha}} \left( ES_t^{\alpha} - VaR_t^{\alpha} + \frac{(VaR_t^{\alpha} - r_t) 1_{\{r_t < VaR_t^{\alpha}\}}}{\alpha} \right) + ln(-ES_t^{\alpha})$$

# Ejemplo

Se consideraron precios de cierre de Bitcoin valorizados en USDT a una frecuencia de 1 hora durante 5 meses.
Datos descargados a través de *python-Binance* API.

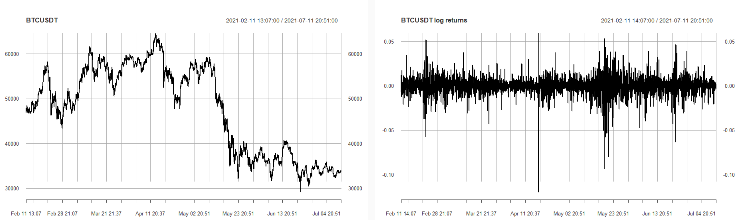
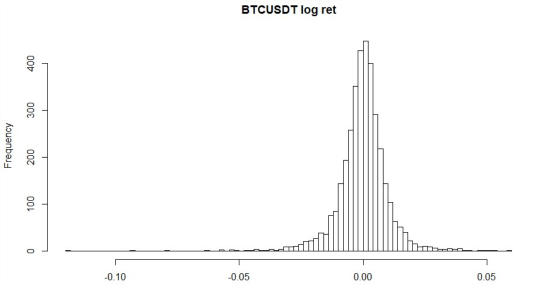



**Figure 1:** Gráficos de precios de cierre y retornos logarítmicos con frecuencia horaria para el periodo establecido.

## Resumen

| Estadísticas | Bitcoin |
|---|---|
| Cantidad de obs. | 3599 |
| Mínimo | -0.1195336 |
| Q1 | -0.0043263 |
| Mediana | 0.0002772 |
| Media | -0.0000901 |
| Q3 | 0.0047216 |
| Máximo | 0.0588610 |
| Desv. std. | 0.01003132 |
| Asimetría | -1.027493 |
| Curtosis | 15.69013 |

**(a)** Tabla resumen de estadísticas.




**(b)** Histograma de retornos logarítmicos.

| N° | model | UC | CC | DQ | ER | ESR | MCS |
|---|---|---|---|---|---|---|---|
| 1 | AR-tGARCH-sGARCH-std-norm | 27,47% | 17,21% | 48,01% | 23,10% | 23,29% | 100,00% |
| 2 | AR-gjrGARCH-tGARCH-std-norm | 22,24% | 13,14% | 38,75% | 15,50% | 17,59% | 100,00% |
| 3 | AR-eGARCH-gjrGARCH-std-std | 42,16% | 17,26% | 29,65% | 35,80% | 33,36% | 100,00% |
| 4 | AR-eGARCH-tGARCH-std-norm | 34,76% | 43,37% | 28,44% | 27,70% | 21,66% | 100,00% |
| 5 | AR-tGARCH-tGARCH-std-norm | 17,21% | 10,34% | 44,64% | 12,90% | 14,57% | 100,00% |
| 6 | AR-tGARCH-sGARCH-std-std | 27,16% | 13,43% | 36,43% | 24,60% | 16,21% | 100,00% |
| 7 | AR-eGARCH-tGARCH-std-std | 34,76% | 58,25% | 28,57% | 50,10% | 42,63% | 100,00% |
| 8 | AR-eGARCH-sGARCH-std-std | 28,08% | 50,08% | 30,10% | 42,80% | 29,66% | 100,00% |
| 9 | ARMA-eGARCH-tGARCH-norm-std | 34,76% | 58,25% | 29,63% | 25,10% | 23,08% | 100,00% |
| 10 | AR-tGARCH-tGARCH-norm-std | 10,34% | 10,34% | 59,80% | 11,40% | 20,00% | 100,00% |
| 11 | AR-eGARCH-tGARCH-norm-std | 22,24% | 34,76% | 30,77% | 18,90% | 15,36% | 100,00% |
| 12 | AR-tGARCH-gjrGARCH-std-std | 21,76% | 40,61% | 27,24% | 17,70% | 13,57% | 100,00% |
| 13 | AR-tGARCH-tGARCH-std-std | 17,21% | 13,43% | 30,02% | 19,90% | 12,48% | 100,00% |
| 14 | AR-gjrGARCH-eGARCH-norm-std | 13,14% | 35,56% | 26,94% | 33,80% | 34,04% | 100,00% |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

**Table 1:** Valores P para los procedimientos de backtesting y MCS utilizando la función de pérdida QL VaR. Se muestra el conjunto de modelos superiores con su especificación para la media y sus regímenes para la varianza (MS-GARCH).
Se probaron 2 especificaciones para media, 2 distribuciones y 4 modelos de varianza, en total 128 modelos de los cuales 89 fueron validados y 68 seleccionados en MCS a un 70%.

# Gráfico de POF

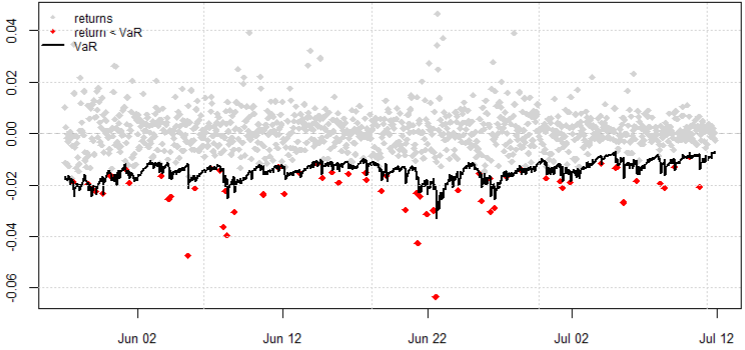



**Figure 3:** modelo AR-tGARCH-sGARCH-std-norm, excesos esperados 54, obtenidos 62 para VaR con $\alpha = 0.05$ (Pvalue = 27.47%)

# Modelos de precios y backtesting

Gabriel Vergara Schifferli

19 Octubre 2021

Modelos de precios para considerar en estructura dinámica en el sistema de control aplicado a la gestión de riesgo.