

Modelando a Variância

Ricardo Ehlers

ehlers@icmc.usp.br

Departamento de Matemática Aplicada e Estatística

Universidade de São Paulo

Introdução

Nos modelos vistos até aqui a variância dos erros foi assumida constante ao longo do tempo, i.e. $Var(\epsilon_t) = E(\epsilon_t^2) = \sigma_\epsilon^2$.

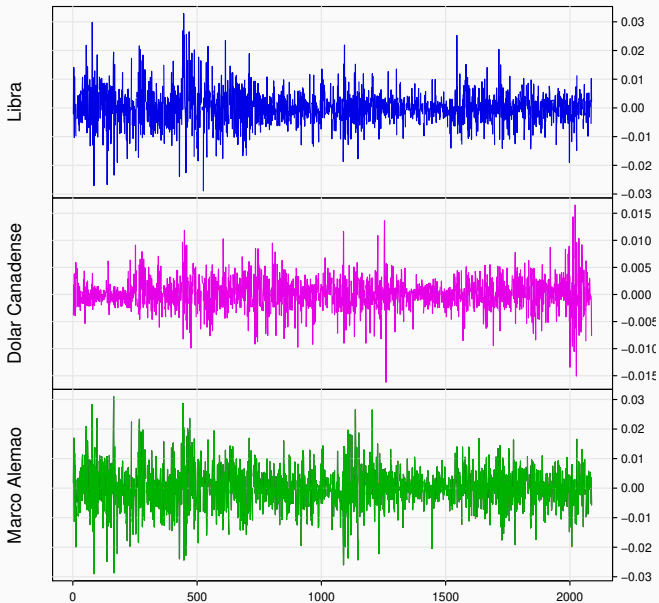
- Muitas séries temporais exibem períodos de grande volatilidade seguidos de períodos de relativa tranquilidade.
- Nestes casos, a suposição de variância constante pode não ser apropriada.
- Embora a variância incondicional dos erros ainda possa ser assumida constante, sua variância condicional pode estar mudando ao longo do tempo.

- Dados financeiros (preços de ações, taxas de cambio, etc.) podem ser observados com frequencia muito alta.
- Isto dá origem a séries temporais longas que costumam exhibir fortes efeitos não lineares e não normalidade.

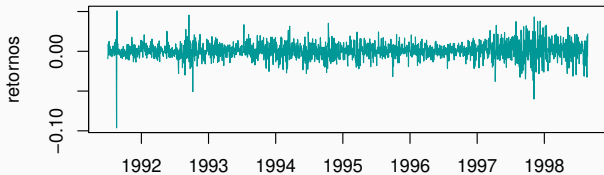
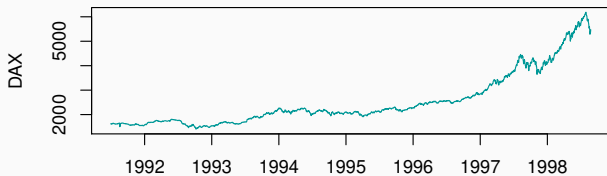
Taxas de câmbio diárias em relação ao Dolar Americano, de janeiro/1991 a dezembro/1998.

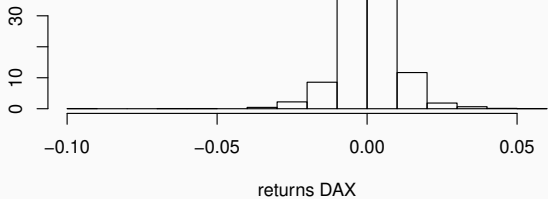


Retornos diários das taxas de câmbio em relação ao Dolar Americano, de janeiro/1991 a dezembro/1998.

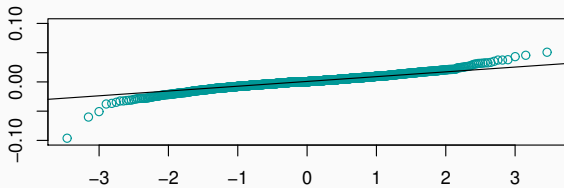


Preços diários no fechamento de um índice de mercado da Alemanha, 1991-1998 e respectivos retornos



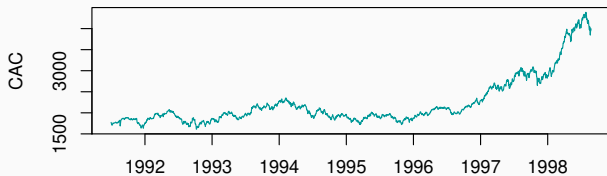


Q-Q norm DAX

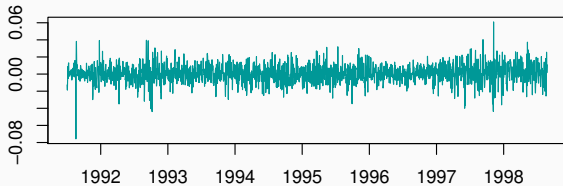


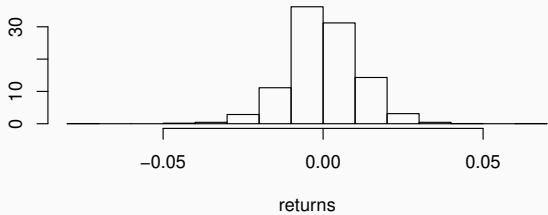
Preços diários no fechamento de um índice da Bolsa de Paris (CAC), 1991-1998 e respectivos retornos.

France CAC

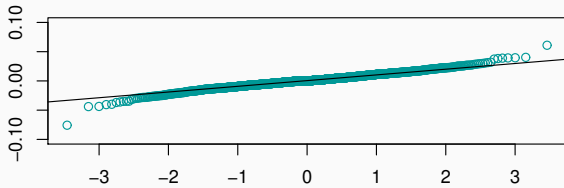


returns





Q-Q norm



Retornos

Seja P_t o preço de um ativo no tempo t .

- A variação de preços entre $t - 1$ e t (sem dividendos pagos) é, $P_t - P_{t-1}$.
- Variação relativa de preços,

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1.$$

- Retorno composto continuamente, ou log-retorno,

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(1 + R_t).$$

Alguns fatos estilizados sobre retornos

- Retornos tendem a apresentar caudas pesadas com pico mais alto em torno da média.
- Retornos tendem a apresentar variabilidades diferentes ao longo do tempo.
- Retornos tendem a apresentar variabilidades agrupadas. Retornos grandes tendem a ser seguidos por retornos grandes e retornos pequenos tendem a ser seguidos por retornos pequenos.
- A variabilidade tende a crescer mais seguindo uma queda de preço (retorno negativo) do que após um aumento de preço de mesma magnitude (efeito alavancagem).

Objetivo: modelar a variancia condicional do *retorno*

$$x_t = \log \left(\frac{P_t}{P_{t-1}} \right)$$

$Var(x_t | \mathcal{F}_{t-1})$: medida de volatilidade dos retornos.

\mathcal{F}_{t-1} : informação até o tempo $t - 1$.

Por exemplo,

$$\begin{aligned}x_t &= \epsilon_t \sqrt{h_t} \\E(\epsilon_t | \mathcal{F}_{t-1}) &= 0 \\Var(\epsilon_t | \mathcal{F}_{t-1}) &= 1 \\Var(x_t | \mathcal{F}_{t-1}) &= h_t\end{aligned}$$

Como a variância condicional h_t varia no tempo?

Modelos ARCH (*Autoregressive conditionally heteroscedastic*)

Definição

O modelo ARCH de ordem 1 é dado por,

$$x_t = \epsilon_t \sqrt{c + \alpha x_{t-1}^2}$$

ou equivalentemente,

$$\begin{aligned}x_t &= \epsilon_t \sqrt{h_t} \\h_t &= c + \alpha x_{t-1}^2\end{aligned}$$

sendo ϵ_t independentes com média 0 e variância 1.

Restrições no modelo,

- Restrição de positividade: $c > 0$ e $\alpha \geq 0$.
- Condição de estacionariedade: $0 < \alpha < 1$.

Segue então que,

$$\begin{aligned}E(x_t | \mathcal{F}_{t-1}) &= E(\epsilon_t) \sqrt{c + \alpha x_{t-1}^2} = 0 \\E(x_t^2 | \mathcal{F}_{t-1}) &= c + \alpha x_{t-1}^2\end{aligned}$$

$$E(x_t) = E[E(x_t|\mathcal{F}_{t-1})] = 0$$

$$\begin{aligned} \text{Var}(x_t) &= E(x_t^2) = E[E(x_t^2|\mathcal{F}_{t-1})] \\ &= E[c + \alpha x_{t-1}^2] = c + \alpha E(x_{t-1}^2). \end{aligned}$$

Se o processo é estacionário,

$$E(x_t^2) = E(x_{t-1}^2) = \text{Var}(x_t)$$

$$\text{Var}(x_t) = \frac{c}{1 - \alpha}.$$

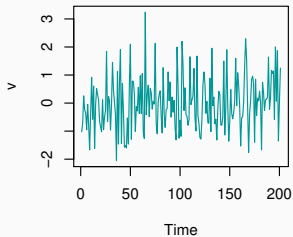
Covariâncias incondicionais

$$\begin{aligned} \text{Cov}(x_t, x_{t+k}) &= E(x_t x_{t+k}) = E[E(x_t x_{t+k}) | x_{t+k-1}, \dots, x_{t-1}] \\ &= E[x_t E(\epsilon_{t+k} \sqrt{c + \alpha x_{t+k-1}^2})] = 0, \text{ para } k > 0. \end{aligned}$$

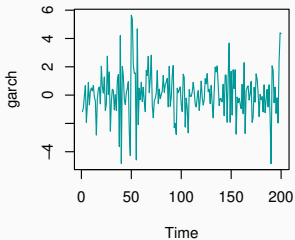
Ou seja, ao postular este modelo assumimos que os $\{x_t\}$ são não correlacionados.

Processos ARCH simulados, com $\epsilon_t \sim N(0, 1)$ e $n = 200$

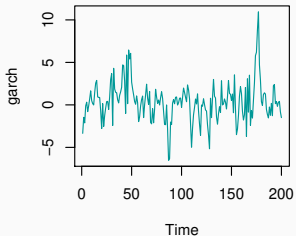
ϵ_t



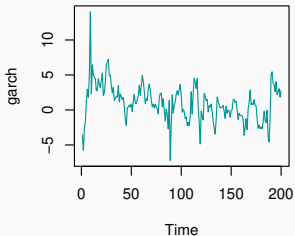
$x_t = \epsilon_t \sqrt{1 + 0.8x_{t-1}^2}$



$x_t = 0.5x_{t-1} + \epsilon_t \sqrt{1 + 0.8x_{t-1}^2}$

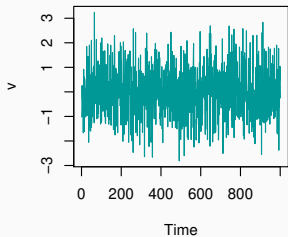


$x_t = 0.9x_{t-1} + \epsilon_t \sqrt{1 + 0.8x_{t-1}^2}$

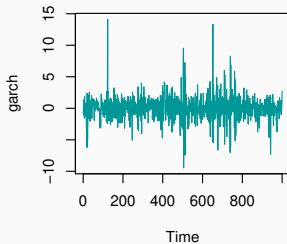


Processos ARCH simulados, $n = 1000$

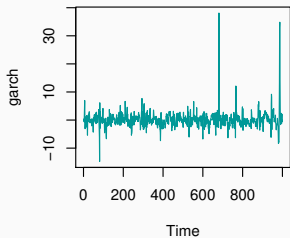
$$\varepsilon_t$$



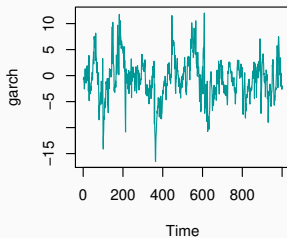
$$x_t = \varepsilon_t \sqrt{1 + 0.8x_{t-1}^2}$$



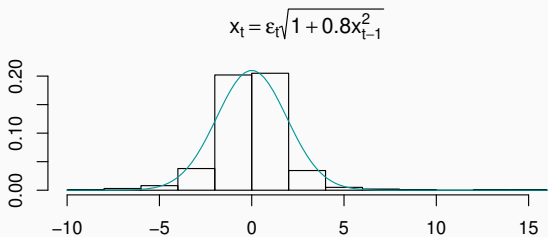
$$x_t = 0.5x_{t-1} + \varepsilon_t \sqrt{1 + 0.8x_{t-1}^2}$$



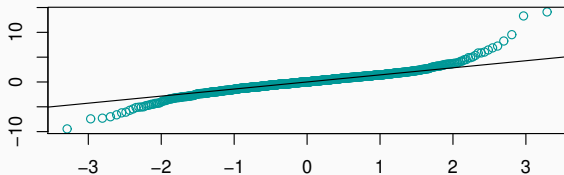
$$x_t = 0.9x_{t-1} + \varepsilon_t \sqrt{1 + 0.8x_{t-1}^2}$$



Processos ARCH simulados, $n = 1000$



Q-Q plot Normal



O modelo ARCH pode ser reescrito como um AR(1) estacionário para x_t^2 com erros não normais.

- Some x_t^2 e subtraia h_t de ambos os lados na equação de volatilidade,

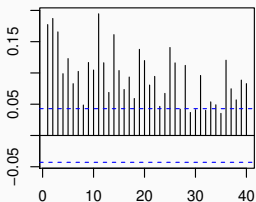
$$\begin{aligned}h_t &= c + \alpha x_{t-1}^2 \\x_t^2 + h_t - h_t &= c + \alpha x_{t-1}^2 + (x_t^2 - h_t) \\x_t^2 &= c + \alpha x_{t-1}^2 + \nu_t\end{aligned}$$

sendo $\nu_t = x_t^2 - h_t = h_t^2(\epsilon_t^2 - 1)$.

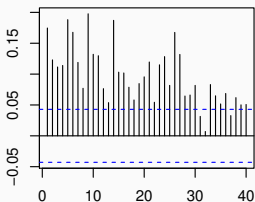
- ν_t tem média zero e variância não constante.
- Se $\epsilon_t \sim N(0, 1)$, $\epsilon_t^2 \sim \chi_1^2$.
- A função de autocorrelação de $\{x_t^2\}$ é dada por $\rho(k) = \alpha^k$ e o correlograma amostral deve apresentar um decaimento exponencial para zero.

Autocorrelações amostrais dos quadrados de retornos de taxas de câmbio em relação ao Dolar Americano, de janeiro/1991 a dezembro/1998.

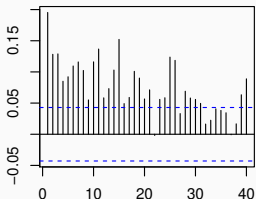
Libra



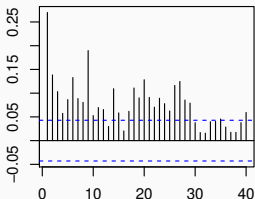
Dolar Canadense



Marco



lenc



Curtose induzida pelo modelo ARCH

O coeficiente de curtose de X_t é dado por,

$$\kappa = \frac{E(X_t^4)}{[Var(X_t)]^2}.$$

- Quarto momento condicional,

$$E(X_t^4 | \mathcal{F}_{t-1}) = E(\epsilon_t^4 h_t^2 | \mathcal{F}_{t-1}) = \lambda E(h_t^2 | \mathcal{F}_{t-1}) = \lambda (c + \alpha X_{t-1}^2)^2.$$

sendo $\lambda = E(\epsilon_t^4)$.

- Quarto momento incondicional,

$$E(X_t^4) = E[E(X_t^4 | \mathcal{F}_{t-1})] = \lambda E(c^2 + \alpha^2 X_{t-1}^4 + 2c\alpha X_{t-1}^2).$$

Se o processo é estacionário de quarta ordem podemos escrever,

$$E(X_t^4) = E(X_{t-1}^4) = \mu_4.$$

Portanto,

$$\begin{aligned}\mu_4 &= \lambda \left(c^2 + \alpha^2 \mu_4 + 2c\alpha \frac{c}{1-\alpha} \right) = \lambda c^2 \left(\frac{1+\alpha}{1-\alpha} \right) + \lambda \alpha^2 \mu_4 \\ \mu_4 &= \frac{\lambda c^2 (1+\alpha)}{(1-\alpha)(1-\lambda \alpha^2)}\end{aligned}$$

O coeficiente de curtose então fica,

$$\kappa = \frac{\lambda c^2 (1+\alpha)}{(1-\alpha)(1-\lambda \alpha^2)} \frac{(1-\alpha)^2}{c^2} = \lambda \frac{1-\alpha^2}{1-\lambda \alpha^2},$$

para $\lambda \alpha^2 < 1$.

Note que $\kappa > \lambda$, qualquer que seja a distribuição de ϵ_t .

- Se $\epsilon_t \sim N(0, 1)$ então $\lambda = 3$.
- Se ϵ_t tem distribuição t -Student com média 0, variância 1 e ν graus de liberdade então,

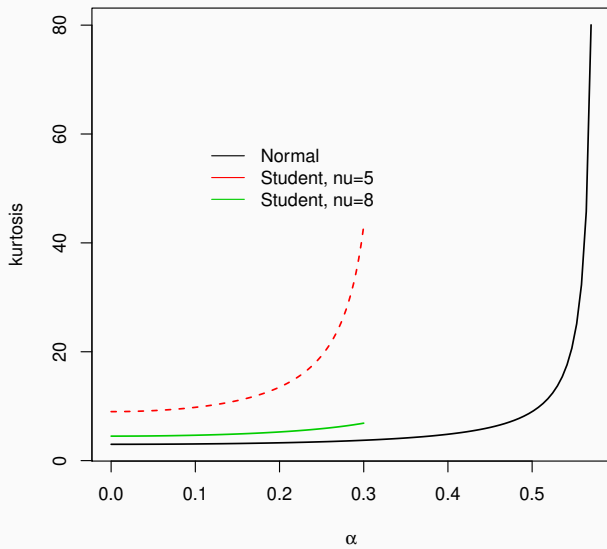
$$\lambda = 3 \frac{\nu - 2}{\nu - 4}.$$

- O coeficiente de curtose cresce mais rápido com α se usarmos a distribuição t .

Comparando curtoses

	α	normal	Student $\nu = 8$	Student $\nu = 5$
1	0.00	3.00	4.50	9.00
2	0.04	3.01	4.52	9.10
3	0.07	3.03	4.59	9.41
4	0.11	3.08	4.70	9.98
5	0.15	3.14	4.88	10.92
6	0.18	3.22	5.12	12.47
7	0.22	3.34	5.47	15.17
8	0.26	3.49	5.97	20.65
9	0.29	3.70	6.71	36.46
10	0.33	3.97	7.86	403.01

kurtosis versus alpha



- A forma da distribuição incondicional de x_t é desconhecida, mesmo que a distribuição de ϵ_t seja conhecida.
- Se a distribuição condicional de ϵ_t for simétrica então a distribuição incondicional de x_t também é simétrica, porém com caudas mais pesadas.
- Se a distribuição condicional de ϵ_t for assimétrica então a distribuição incondicional de x_t também é assimétrica.

Distribuição preditiva de X_{n+1} ,

$$p(x_{n+1}|\mathbf{x}) = \int p(x_{n+1}|\mathbf{x}, \boldsymbol{\theta})p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}.$$

Se $\epsilon_t \sim N(0, 1)$ temos que,

$$x_{n+1}|\mathbf{x}, \boldsymbol{\theta} \sim N(0, c + \alpha x_n^2).$$

Os momentos da distribuição preditiva são,

$$E(x_{n+1}|\mathbf{x}) = E[E(x_{n+1}|\mathbf{x}, \boldsymbol{\theta})] = 0$$

$$\begin{aligned} \text{Var}(x_{n+1}|\mathbf{x}) &= E[\text{Var}(x_{n+1}|\mathbf{x}, \boldsymbol{\theta})] + \text{Var}[E(x_{n+1}|\mathbf{x}, \boldsymbol{\theta})] \\ &= E[c + \alpha x_n^2] = E(c|\mathbf{x}) + E(\alpha|\mathbf{x})x_n^2. \end{aligned}$$

Distribuição preditiva de X_{n+2} ,

$$\begin{aligned} p(x_{n+2}|\mathbf{x}) &= \int p(x_{n+2}|\mathbf{x}, \boldsymbol{\theta})p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta} \\ &= \int \int p(x_{n+2}|x_{n+1}, \mathbf{x}, \boldsymbol{\theta})p(x_{n+1}|\mathbf{x}, \boldsymbol{\theta})dx_{n+1} p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta} \end{aligned}$$

Se $\epsilon_t \sim N(0, 1)$ temos que,

$$\begin{aligned} x_{n+1}|\mathbf{x}, \boldsymbol{\theta} &\sim N(0, c + \alpha x_n^2) \\ x_{n+2}|x_{n+1}, \mathbf{x}, \boldsymbol{\theta} &\sim N(0, c + \alpha x_{n+1}^2) \end{aligned}$$

mas a integral interna não tem solução analítica.

Os momentos da distribuição preditiva são,

$$E(x_{n+2}|\mathbf{x}) = E[E(x_{n+2}|\mathbf{x}, \boldsymbol{\theta})] = 0$$

$$\begin{aligned} \text{Var}(x_{n+2}|\mathbf{x}) &= E[\text{Var}(x_{n+2}|\mathbf{x}, \boldsymbol{\theta})] + \text{Var}[E(x_{n+2}|\mathbf{x}, \boldsymbol{\theta})] \\ &= E[\text{Var}(x_{n+2}|\mathbf{x}, \boldsymbol{\theta})] \end{aligned}$$

Mas note que,

$$\begin{aligned} \text{Var}(x_{n+2}|\mathbf{x}, \boldsymbol{\theta}) &= E[\text{Var}(x_{n+2}|x_{n+1}, \mathbf{x}, \boldsymbol{\theta})] + \text{Var}[E(x_{n+2}|x_{n+1}, \mathbf{x}, \boldsymbol{\theta})] \\ &= E[\text{Var}(x_{n+2}|x_{n+1}, \mathbf{x}, \boldsymbol{\theta})] \\ &= E(c + \alpha x_{n+1}^2) \\ &= c + \alpha E(x_{n+1}^2|\mathbf{x}, \boldsymbol{\theta}) = c + \alpha(c + \alpha x_n^2) \end{aligned}$$

O Modelo ARCH(p)

Estas idéias podem ser generalizadas para processos ARCH(p).

Definição

O modelo ARCH de ordem p , ou ARCH(p) é dado por,

$$\begin{aligned}x_t &= \epsilon_t \sqrt{h_t} \\ h_t &= c + \alpha_1 x_{t-1}^2 + \cdots + \alpha_p x_{t-p}^2\end{aligned}$$

sendo ϵ_t independentes com média 0 e variância 1.

Portanto, a variância condicional é modelada como,

$$h_t = E(x_t^2 | \mathcal{F}_{t-1}) = c + \alpha_1 x_{t-1}^2 + \cdots + \alpha_p x_{t-p}^2.$$

Restrições no modelo,

- Restrições de positividade, $c > 0$ e $\alpha_1 \geq 0, \dots, \alpha_p \geq 0$.
- Restrições de estacionariedade: raízes de $1 - \alpha_1 B - \dots - \alpha_p B^p = 0$ fora do círculo unitário.
- Juntando estas restrições equivale a impor a restrição,

$$c > 0 \quad \text{e} \quad \sum_{i=1}^p \alpha_i < 1.$$

O modelo ARCH(p) pode ser reescrito como um processo AR(p) para x_t^2 somando x_t^2 e subtraindo h_t de ambos os lados na equação de volatilidade,

$$x_t^2 = c + \alpha_1 x_{t-1}^2 + \cdots + \alpha_p x_{t-p}^2 + \nu_t.$$

com $\nu_t = h_t(\epsilon_t^2 - 1)$.

A função de autocorrelação parcial de x_t^2 indica a ordem p do processo.

Modelos GARCH

Definição

O modelo ARCH Generalizado de ordem p e q , GARCH(p, q) é dado por,

$$\begin{aligned}x_t &= \epsilon_t \sqrt{h_t} \\h_t &= c + \sum_{i=1}^p \alpha_i x_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}.\end{aligned}$$

sendo ϵ_t independentes com média 0 e variância 1.

A variância condicional depende também de seus valores passados.

Restrições do modelo,

- Restrições de positividade, $c > 0$, $\alpha_i \geq 0$, $i = 1, \dots, p$,
 $\beta_j \geq 0$, $j = 1, \dots, q$.
- Restrições de estacionariedade,

$$\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1.$$

Pode-se mostrar que a variância incondicional é dada por,

$$\text{Var}(X_t) = \frac{c}{1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j}.$$

- Em geral mais parcimoniosos do que modelos ARCH.
- Na maioria dos casos um modelo GARCH(1,1) é suficiente.
- Variância incondicional no GARCH(1,1),

$$\text{Var}(X_t) = \frac{c}{1 - \alpha - \beta}.$$

- O parâmetro $\alpha + \beta$ é chamado de *persistência* e em geral é estimado próximo de 1.

Estimação

Para uma série temporal x_1, \dots, x_n observada e um modelo GARCH(p, q), denota-se o vetor de parâmetros por

$$\theta = (c, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q).$$

A função de verossimilhança completa fica,

$$p(x_1, \dots, x_n | \theta) = p(x_1, \dots, x_p | \theta) \prod_{t=p+1}^n p(x_t | x_{t-1}, \dots, x_{t-p}, \theta).$$

Se $\epsilon_t \sim N(0, 1)$ segue que,

$$X_t | x_{t-1}, \dots, x_{t-p} \sim N(0, h_t)$$

e portanto

$$p(x_1, \dots, x_n | \theta) = p(x_1, \dots, x_p | \theta) \prod_{t=p+1}^n (2\pi h_t)^{-1/2} \exp\left(-\frac{x_t^2}{2h_t}\right).$$

A distribuição incondicional de X_t não é normal e tem caudas pesadas.

Em geral o número de observações será grande o suficiente para que o termo $p(x_1, \dots, x_p | \theta)$ possa ser desprezado.

Por exemplo, para um modelo ARCH(1) a função log-verossimilhança fica,

$$-\frac{1}{2} \sum_{t=2}^n \left[\log(2\pi) + \log(c + \alpha x_{t-1}^2) + \frac{x_t^2}{c + \alpha x_{t-1}^2} \right].$$

Para um modelo GARCH(1,1) a função log-verossimilhança fica,

$$-\frac{1}{2} \sum_{t=2}^n \left[\log(2\pi) + \log(c + \alpha x_{t-1}^2 + \beta h_{t-1}) + \frac{x_t^2}{c + \alpha x_{t-1}^2 + \beta h_{t-1}} \right].$$

- Algum algoritmo de otimização não linear deverá ser utilizado e nada garante sua convergência para um ótimo global.
- No R pode-se usar a função `garch` do pacote `tseries` para fazer a estimação por máxima verossimilhança.

- Existe evidência de que as séries de retornos apresentam curtose observada ainda maior do que as do modelo GARCH com erros normais.
- Costuma-se usar distribuições com caudas mais pesadas para os erros, especialmente a t -Student e a exponencial potência.

Forma geral da função de verossimilhança (condicional),

$$p(x_1, \dots, x_n | \theta) = \prod_{t=s+1}^n \frac{1}{\sqrt{h_t}} p_\epsilon \left(\frac{x_t}{\sqrt{h_t}} \right)$$

sendo $s = \max(p, q)$ e p_ϵ a densidade de ϵ_t .

Distribuição Exponential Potência com média 0 e variância 1

$$f(\epsilon_t) = \frac{\nu}{\lambda 2^{1+1/\nu} \Gamma(1/\nu)} \exp \left\{ -\frac{1}{2} \left| \frac{\epsilon_t}{\lambda} \right|^\nu \right\}$$

$$\lambda^2 = \frac{1}{2^{2/\nu}} \frac{\Gamma(1/\nu)}{\Gamma(3/\nu)}$$

sendo $\nu > 0$ o parâmetro de forma e $\Gamma(\cdot)$ é a função Gama.

Casos particulares,

- distribuição de Laplace (ou exponencial dupla) quando $\nu = 1$,
- distribuição normal padrão quando $\nu = 2$.

A curtose de ϵ_t é dada por,

$$\Gamma(1/\nu)\Gamma(5/\nu)/\Gamma(3/\nu)^2 - 3$$

portanto quando $\nu < 2$ esta distribuição reproduz caudas pesadas.

Distribuição t -Student com média 0 e variância 1

Se Y tem distribuição t com parâmetros de locação 0, escala 1 e graus de liberdade ν , $Y \sim t(0, 1, \nu)$, então

$$E(Y) = 0$$
$$\text{Var}(Y) = \left(\frac{\nu}{\nu - 2} \right), \nu > 2.$$

Portanto, se $X = Y\sqrt{(\nu - 2)/\nu}$, segue que $E(X) = 0$ e $\text{Var}(X) = 1$ e,

$$f(x) = f_y \left(x \left(\frac{\nu}{\nu - 2} \right)^{1/2} \right) \left(\frac{\nu}{\nu - 2} \right)^{1/2}, \nu > 2.$$

Se ϵ_t tem distribuição t -Student com média 0 e variância 1,

$$f(\epsilon_t) = \frac{1}{\sqrt{\pi(\nu - 2)}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \left\{ 1 + \frac{\epsilon_t^2}{\nu - 2} \right\}^{-(\nu+1)/2}, \quad \nu > 2. \quad (1)$$

Quando $\nu \rightarrow \infty$ esta distribuição se aproxima da distribuição normal padrão.

Distribuição t -Student como mistura de normais

Sejam U e W tais que,

$$U|W = w \sim N\left(0, \frac{\nu - 2}{\nu} w\right), \nu > 2$$
$$W \sim \text{Gama-Inversa}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$$

Então U tem distribuição t -Student com média 0, variância 1 e $\nu > 2$ graus de liberdade.

$$f(u|w) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\frac{\nu-2}{\nu}w}} \exp \left\{ -\frac{u^2}{2 \left(\frac{\nu-2}{\nu} \right) w} \right\}$$

$$\propto w^{-1/2} \exp \left\{ -\frac{u^2}{2 \left(\frac{\nu-2}{\nu} \right) w} \right\}$$

$$f(w) = \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2 + 1)} w^{-(\nu/2+1)} \exp \{-\nu/2w\}$$

A densidade marginal de u é,

$$\begin{aligned} f(u) &\propto \int_0^\infty w^{-(\nu+1)/2-1} \exp \left\{ -\frac{1}{2w} \left[\nu + \frac{u^2}{\nu} \right] \right\} dw \\ &\propto \left[\nu + \frac{u^2}{\nu} \right]^{-(\nu+1)/2} \\ &\propto \left[1 + \frac{u^2}{\nu-2} \right]^{-(\nu+1)/2} . \end{aligned}$$

Portanto, u tem distribuição t -Student com média 0, variância 1 e $\nu > 2$ graus de liberdade. Compare com a expressão (1).

$$x_t = u_t \sqrt{h_t}, \quad t = 1, \dots, n$$

$$h_t = c + \sum_{i=1}^p \alpha_i x_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}$$

$$u_t = \epsilon_t \sqrt{\frac{\nu - 2}{\nu}} w_t$$

$$\epsilon_t \sim N(0, 1)$$

$$u_t | w_t \sim N\left(0, \frac{\nu}{\nu - 2} w_t\right)$$

$$w_t \sim IG\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$$

Pacote bayesGARCH para o modelo GARCH(1,1)

Estimação Bayesiana do modelo GARCH(1,1) com erros *t*-Student.
Disponível em
<http://cran.r-project.org/web/packages/bayesGARCH>.

Referência: Ardia and Hoogerheide (2010) Bayesian Estimation of the GARCH(1,1) Model with Student-t Innovations. The R Journal, 2, 41-47.

Disponível em <https://journal.r-project.org/archive/>.

O pacote assume o seguinte modelo e distribuições,

$$x_t = \epsilon_t \sqrt{\left(\frac{\nu - 2}{\nu}\right) w_t} \sqrt{h_t}, \quad t = 1, \dots, n$$

$$h_t = \alpha_0 + \alpha_1 x_{t-i}^2 + \beta h_{t-j}$$

$$\epsilon_t \sim N(0, 1)$$

$$w_t \sim IG\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$$

Equivalentemente,

$$u_t = \epsilon_t \sqrt{\left(\frac{\nu - 2}{\nu}\right) w_t}$$

tem distribuição t -Student com média 0, variância 1 e ν graus de liberdade.

Distribuição condicional de X_t ,

$$X_t | w_t, \mathcal{F}_{t-1} \sim N\left(0, \frac{\nu - 2}{\nu} w_t h_t\right).$$

Distribuições a priori

$$\alpha \sim N_2(\mu_\alpha, \Sigma_\alpha), \text{ truncada em } \alpha \in \mathbb{R}_+^2$$

$$\beta \sim N(\mu_\beta, \Sigma_\beta), \text{ truncada em } \beta \in \mathbb{R}_+$$

$$p(\mathbf{w}|\nu) = \prod_{t=1}^n p(w_t|\nu)$$

$$p(\nu) = \lambda \exp[-\lambda(\nu - \delta)], \nu > \delta, \delta \geq 2.$$

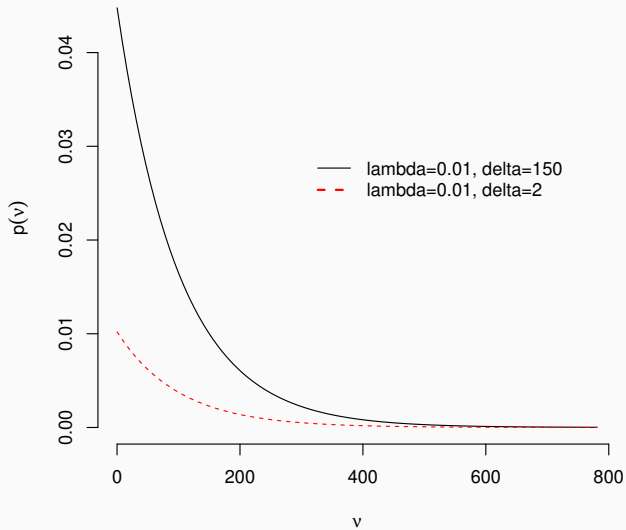
- Assume-se que, condicionalmente a ν , as variáveis latentes w_t são independentes e identicamente distribuídas a priori.
- A distribuição a priori de ν é uma exponencial deslocada, truncada em $\nu > \delta \geq 2$.
- Densidade a priori conjunta,

$$p(\boldsymbol{\theta}) = p(\boldsymbol{\alpha}) p(\beta) p(\mathbf{w}|\nu) p(\nu).$$

- Com restrição de estacionariedade,

$$p(\boldsymbol{\theta}) = p(\boldsymbol{\alpha}) p(\beta) I(\alpha_1 + \beta < 1) p(\mathbf{w}|\nu) p(\nu).$$

- O modelo com erros normais é obtido fazendo-se $\lambda = 100$ e $\delta = 500$.



Exemplo. Modelo GARCH(1,1) com erros t para os retornos diários de câmbio Libra Esterlina versus Dolar Americano. Período amostral 1/10/81 a 28/6/85.

Hiperparâmetros especificados,

$$\mu_{\alpha} = (0, 0)$$

$$\Sigma_{\alpha} = 1000\mathbf{I}_2$$

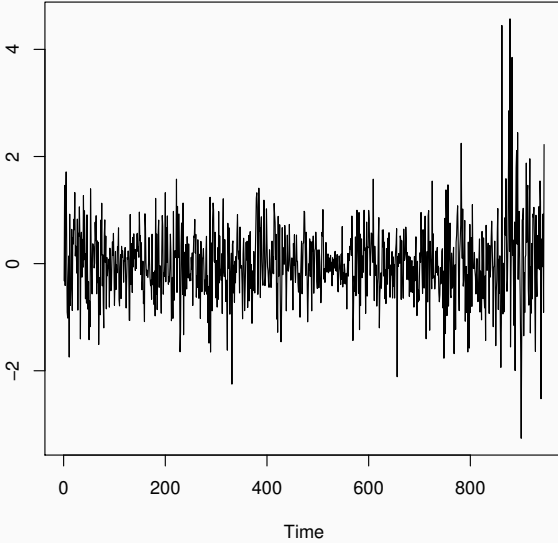
$$\mu_{\beta} = 0$$

$$\Sigma_{\beta} = 1000$$

$$\lambda = 0.01$$

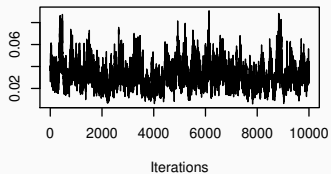
$$\delta = 2$$

Demeaned log returns

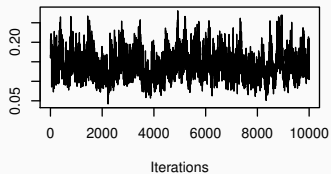



```
> library(bayesGARCH)
> niter= 5000
> nburn= 2000
> start= c(0.01,0.1,0.7,20) # valores iniciais
> mcmc = list(n.chain=1,l.chain=niter,start.val=start,digits=4)
> m = bayesGARCH(y,control=mcmc)
> theta= formSmpl(m,l.bi=nburn,batch.size=1)
> library(coda)
> theta= as.mcmc(transform(theta,pers=alpha1+beta))
```

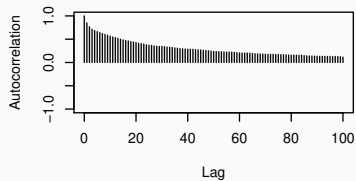
Trace of alpha0



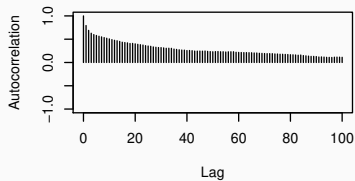
Trace of alpha1



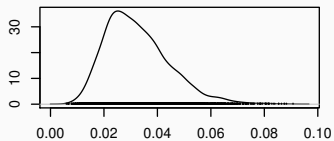
alpha0



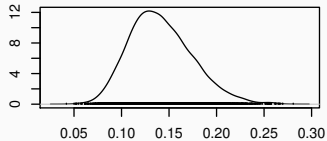
alpha1



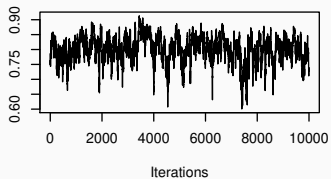
Density of alpha0



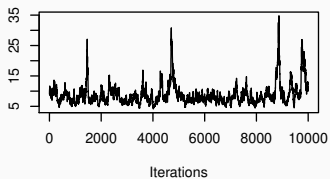
Density of alpha1



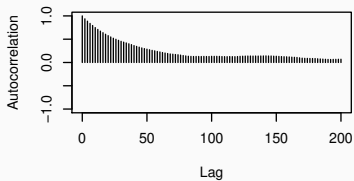
Trace of beta



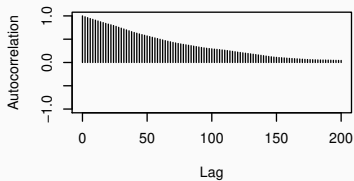
Trace of nu



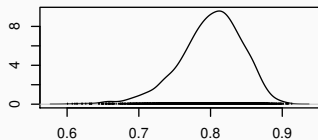
beta



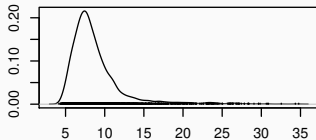
nu

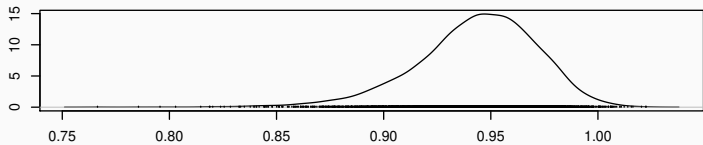
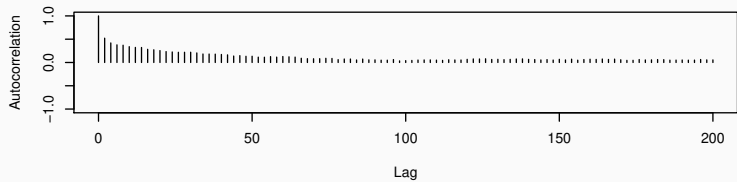
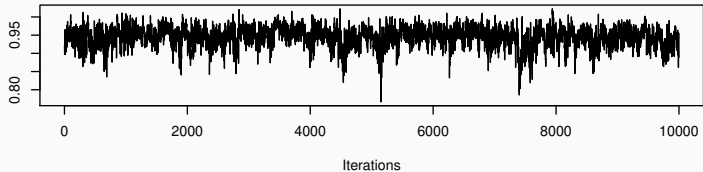


Density of beta



Density of nu





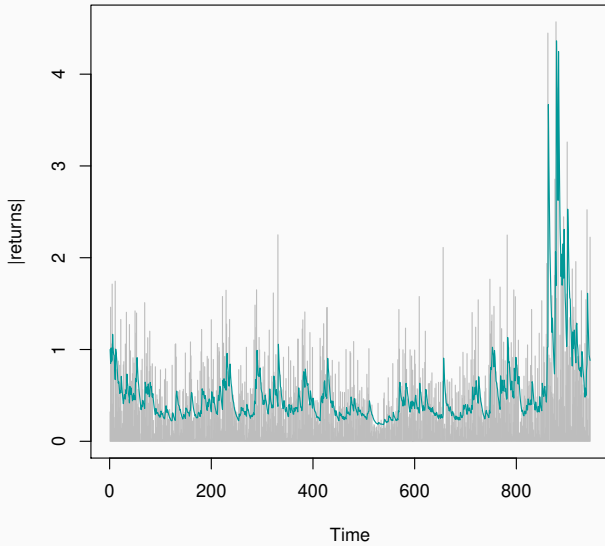
1. Empirical mean and standard deviation for each variable, plus standard error of the mean:

	Mean	SD	Naive SE	Time-series SE
alpha0	0.0325	0.01159	0.0001159	0.0007604
alpha1	0.1422	0.03194	0.0003194	0.0019490
beta	0.8023	0.04248	0.0004248	0.0033913
nu	8.1107	2.19581	0.0219592	0.2391872
pers	0.9445	0.02788	0.0002788	0.0013786

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
alpha0	0.01491	0.02424	0.03074	0.03884	0.05995
alpha1	0.08806	0.11983	0.13936	0.16114	0.21285
beta	0.70469	0.77781	0.80670	0.83206	0.87353
nu	5.22076	6.58241	7.65092	9.10996	13.60883
pers	0.88174	0.92822	0.94725	0.96349	0.99117

Estimated Volatility



Análise de Resíduos

- No modelo normal,

$$\epsilon_t = x_t h_t^{-1/2}.$$

- No modelo com erros t como mistura de normais,

$$\epsilon_t = x_t \left[\frac{\nu - 2}{\nu} w_t h_t \right]^{-1/2}.$$

Distribuição preditiva

- Obter uma amostra de N valores simulados $(\theta^{(1)}, \dots, \theta^{(N)})$,
- Obter uma amostra de tamanho N de h_{n+1} ,

$$h_{n+1}^{(j)} = c^{(j)} + \alpha^{(j)} x_n^2 + \beta^{(j)} h_n^{(j)}, \quad j = 1, \dots, N$$

- Obter uma amostra de tamanho N de x_{n+1} ,

$$x_{n+1}^{(j)} = \epsilon_{n+1}^{(j)} \sqrt{h_{n+1}^{(j)}}, \quad j = 1, \dots, N$$

- Obter uma amostra de tamanho N de h_{n+2} ,

$$h_{n+2}^{(j)} = c^{(j)} + \alpha^{(j)} x_{n+1}^2 + \beta^{(j)} h_{n+1}^{(j)}, \quad j = 1, \dots, N$$

Pacote fGarch para modelos GARCH

Estimação clássica de modelos GARCH com erros *t*-Student.

Disponível em

<https://cran.r-project.org/web/packages/fGarch>.

```
> library(fGarch)
```

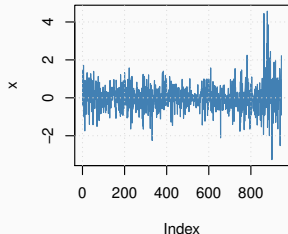
```
> args(garchFit)
```

```
function (formula = ~garch(1, 1), data = dem2gbp, init.rec = c("mci",  
  "uev"), delta = 2, skew = 1, shape = 4, cond.dist = c("norm",  
  "snorm", "ged", "sged", "std", "sstd", "snig", "QMLE"), include.mean,  
  include.delta = NULL, include.skew = NULL, include.shape = NULL,  
  leverage = NULL, trace = TRUE, algorithm = c("nlminb", "lbfgsb",  
  "nlminb+nm", "lbfgsb+nm"), hessian = c("ropt", "rcd"),  
  control = list(), title = NULL, description = NULL, ...)
```

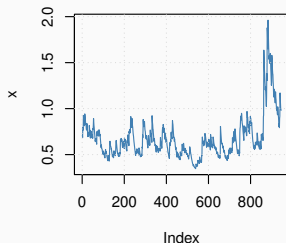
```
NULL
```

Exemplo 1 (cont.): Modelo GARCH(1,1) com erros normais para os retornos diários de câmbio Libra Esterlina versus Dolar Americano. Período amostral 1/10/81 a 28/6/85.

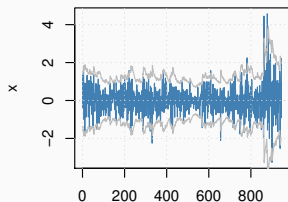
Time Series



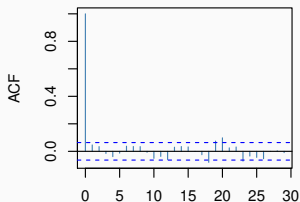
Conditional SD



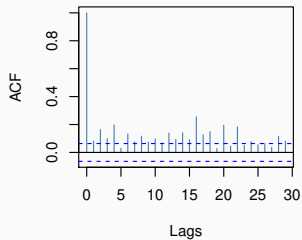
Series with 2 Conditional SD Superimp



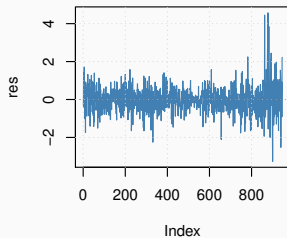
ACF of Observations



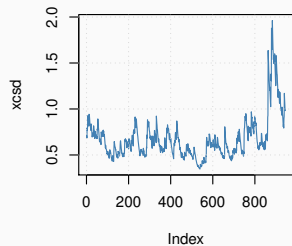
ACF of Squared Observations



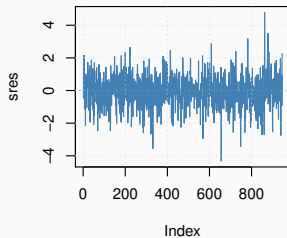
Residuals



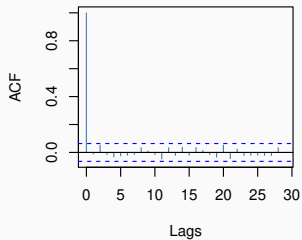
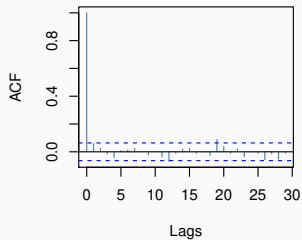
Conditional SD's



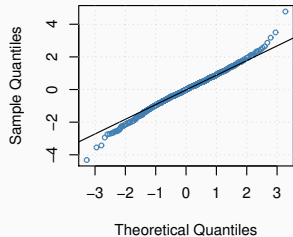
Standardized Residuals



ACF of Standardized Residuals ACF of Squared Standardized Residu



qnorm - QQ Plot



Exemplo 1 (cont.): Modelo GARCH(1,1) para os retornos diários de câmbio Libra Esterlina versus Dolar Americano. Período amostral 1/10/81 a 28/6/85. Erros com distribuição t assimétrica.

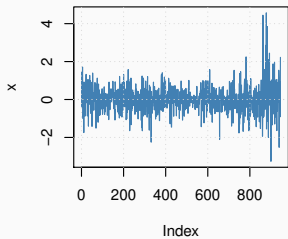
Uma forma de introduzir assimetria na densidade dada em (1) é,

$$f(\epsilon_t) = \frac{2}{\gamma + 1/\gamma} \frac{1}{\sqrt{\pi(\nu - 2)}} \times$$
$$\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \left[1 + \frac{\epsilon_t^2}{\nu - 2} \left\{ \frac{1}{\gamma^2} I_{[0,\infty)}(\epsilon_t) + \gamma^2 I_{[-\infty,0)}(\epsilon_t) \right\} \right]^{-\frac{\nu+1}{2}}$$

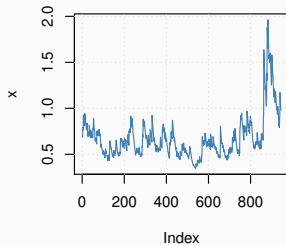
- $\gamma > 0$ é o parâmetro de assimetria.
- Se $\gamma = 1$ a densidade (1) é recuperada.

	Estimate	Std. Error	t value	Pr(> t)
omega	0.007436226	0.004437778	1.675664	0.0938039832
alpha1	0.078232859	0.023573548	3.318671	0.0009044679
beta1	0.908461540	0.026810207	33.884913	0.0000000000
skew	0.944896272	0.043377308	21.783193	0.0000000000
shape	9.224615222	2.473775823	3.728962	0.0001922704

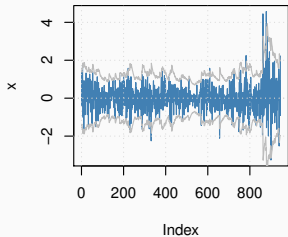
Time Series



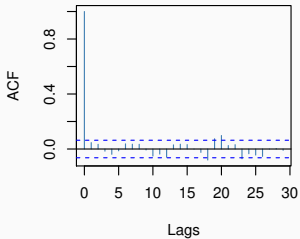
Conditional SD



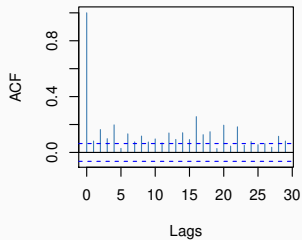
Series with 2 Conditional SD Superimp



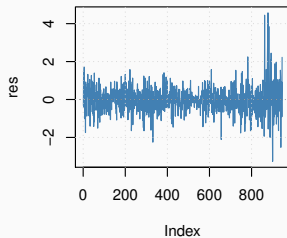
ACF of Observations



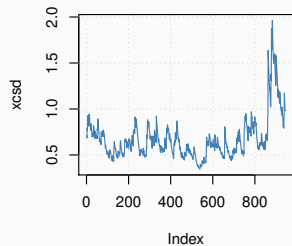
ACF of Squared Observations



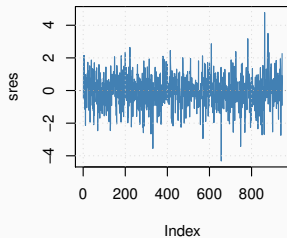
Residuals



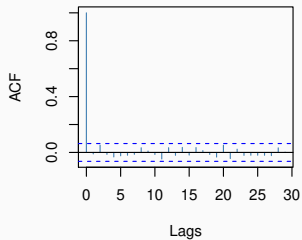
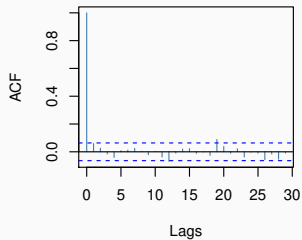
Conditional SD's



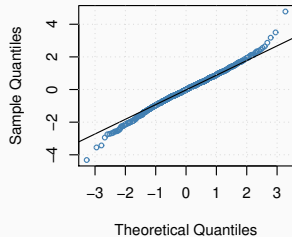
Standardized Residuals



ACF of Standardized Residuals ACF of Squared Standardized Residu



qnorm - QQ Plot



Modelo EGARCH

- Em geral, o mercado tem baixa volatilidade quando está em alta e alta volatilidade quando está em baixa.
- Tal assimetria não é levada em conta pelos modelos GARCH.
- Modelos GARCH exponencial (EGARCH),

$$x_t = \epsilon_t \sqrt{h_t}$$
$$\log(h_t) = c + \alpha \left| \frac{x_{t-1}}{\sqrt{h_{t-1}}} \right| + \gamma \frac{x_{t-1}}{\sqrt{h_{t-1}}} + \beta \log(h_{t-1}).$$

- Os parâmetros c , α e β podem ser estimados de forma irrestrita.
- O parâmetro γ é restrito a $\gamma < 0$ indicando que a volatilidade aumenta quando $\epsilon_{t-1} < 0$.

Modelos GJR-GARCH

Referência: Glosten, L.R., Jagannathan, R. and Runkle, D.E. (1993) "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks". The Journal of Finance, vol. 48, 1791-1801.

$$\begin{aligned}x_t &= \epsilon_t \sqrt{h_t} \\h_t &= c + \alpha x_{t-1}^2 + \lambda x_{t-1}^2 I(x_{t-1} < 0) + \beta h_{t-1}\end{aligned}$$

sendo $I(\cdot)$ a função indicadora, $I(x_{t-1} < 0) = 1$ se $x_{t-1} < 0$ e $I(x_{t-1} < 0) = 0$ caso contrário.

Modelos APARCH (Asymmetric Power ARCH)

Referência: Ding, Z., Granger C.W.J., Engle R.F. (1993) "A Long Memory Property of Stock Market Returns and a New Model", Journal of Empirical Finance 1, 83-106.

$$\begin{aligned}x_t &= \epsilon_t \sqrt{h_t} \\ \sigma_t^\delta &= c + \alpha(|x_{t-1}| - \gamma x_{t-1})^\delta + \beta \sigma_{t-1}^\delta\end{aligned}$$

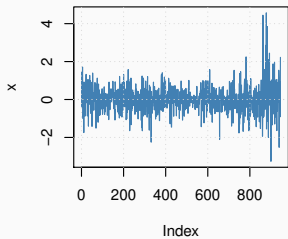
sendo $\sigma_t = \sqrt{h_t}$

- O parametro $\gamma \in (-1, 1)$ reflete o efeito alavancagem.
- Se $\gamma > 0$: informação negativa tem impacto maior sobre a volatilidade de preços do que informação positiva
- Restrição de positividade: $c > 0, \alpha \geq 0, \beta \geq 0$.
- Restrição de estacionariedade: $\alpha + \beta < 1$.
- Se $\delta = 2$ e $\gamma = 0$ temos o modelo GARCH(1,1).

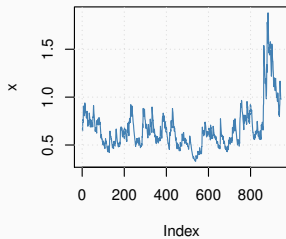
Exemplo 1 (cont.): Modelo APARCH(1,1) com erros normais para os retornos diários de câmbio Libra Esterlina versus Dolar Americano. Período amostral 1/10/81 a 28/6/85. Pacote fGarch.

	Estimate	Std. Error	t value	Pr(> t)
omega	0.010922550	0.005917884	1.84568527	6.493792e-02
alpha1	0.104041921	0.026543550	3.91966866	8.867079e-05
gamma1	0.007982925	0.079833791	0.09999431	9.203488e-01
beta1	0.890229719	0.027688433	32.15168258	0.000000e+00
delta	1.674897634	0.474893574	3.52689050	4.204706e-04

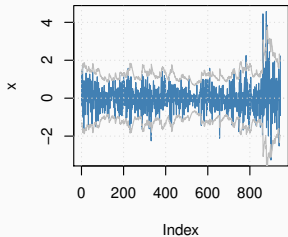
Time Series



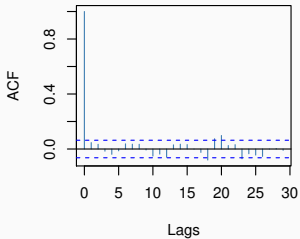
Conditional SD



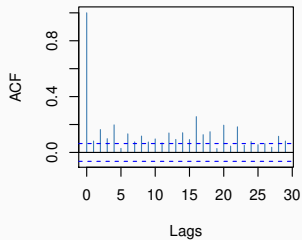
Series with 2 Conditional SD Superimp



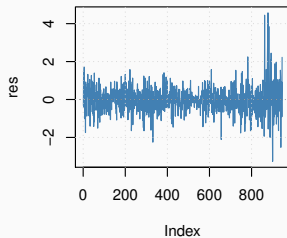
ACF of Observations



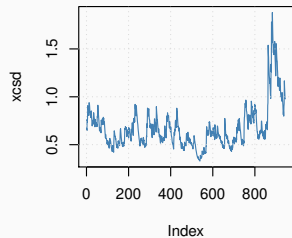
ACF of Squared Observations



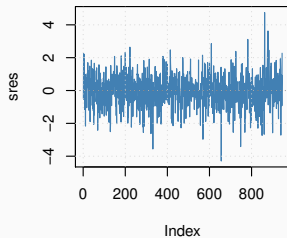
Residuals



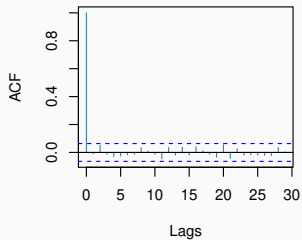
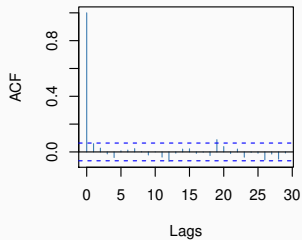
Conditional SD's



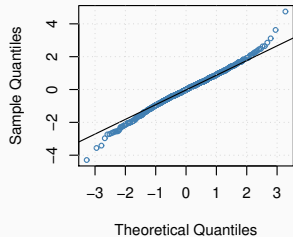
Standardized Residuals



ACF of Standardized Residuals ACF of Squared Standardized Residu



qnorm - QQ Plot



Volatilidade Estocástica

Nos modelos GARCH a evolução de h_t é determinística.

Pode ser mais razoável assumir que a variância condicional varia estocasticamente.

Como alternativa assume-se que h_t segue um processo estocástico.

$$\begin{aligned}X_t &= \epsilon_t \exp(h_t/2), \quad \epsilon_t \sim N(0, 1) \\h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2)\end{aligned}$$

com $h_t = \log(\sigma_t^2)$ sendo,

$$\text{Var}(X_t | \mathcal{F}_{t-1}) = \exp(h_t) = \sigma_t^2.$$

- Restrição de estacionariedade, $|\phi| < 1$ e ϕ pode ser interpretado como a persistencia na volatilidade.
- Não há necessidade de restrições de positividade.
- Os erros ϵ_t e η_t são assumidos independentes.
- O fator de escala $\beta = \exp(\mu/2)$ é a volatilidade modal.
- σ_η^2 é a volatilidade das log-volatilidades.

- O modelo pode ser estendido para uma estrutura AR(p) em h_t ,

$$X_t = \epsilon_t \exp(h_t/2), \quad \epsilon_t \sim N(0, 1)$$

$$h_t = \mu + \sum_{i=1}^p \phi_i (h_{t-i} - \mu) + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2)$$

- Estimar modelos de volatilidade estocástica é mais difícil porque a função de verossimilhança é difícil de ser avaliada.

- $E(X_t) = E(\epsilon_t e^{h_t/2}) = E(e^{h_t/2})E(\epsilon_t) = 0$, já que h_t e ϵ_t são independentes.
- $Var(X_t) = E(X_t^2) = E(e^{h_t} \epsilon_t^2) = E(e^{h_t})E(\epsilon_t^2) = E(e^{h_t})$.
- Como estamos assumindo que h_t é estacionária,

$$E(h_t) = \mu = c/(1 - \alpha) \quad \text{e} \quad Var(h_t) = \sigma^2 = \sigma_\eta^2/(1 - \alpha^2)$$

e a distribuição incondicional é $h_t \sim N(\mu, \sigma^2)$.

- e^{h_t} segue uma distribuição log-normal com parâmetros μ e σ^2 ,

$$\begin{aligned} E(e^{h_t}) &= e^{\mu + \sigma^2/2} = Var(X_t) \\ Var(e^{h_t}) &= (e^{\sigma^2} - 1)e^{2\mu + \sigma^2} \end{aligned}$$

- $E(X_t^4) = E(\epsilon_t^4 e^{2h_t}) = E(\epsilon_t^4)E(e^{2h_t})$.
- Se $\epsilon_t \sim N(0, 1)$ então $E(\epsilon_t^4) = 3$ e $E(X_t^4) = 3E(e^{2h_t})$ sendo,

$$\begin{aligned} E(e^{2h_t}) &= \text{Var}(e^{h_t}) + E^2(e^{h_t}) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2} + (e^{\mu + \sigma^2/2})^2 \\ &= e^{2\mu + \sigma^2}(1 + e^{\sigma^2} - 1) = e^{2\mu + 2\sigma^2}. \end{aligned}$$

- Portanto, a curtose é dada por,

$$\kappa = \frac{3 e^{2\mu + 2\sigma^2}}{e^{2\mu + \sigma^2}} = 3e^{\sigma^2}.$$

- κ é sempre maior do que 3 pois $e^{\sigma^2} > 1$. Um resultado mais geral é que $\kappa = E(\epsilon_t^4)e^{\sigma^2}$ ou seja a curtose induzida por este modelo é sempre maior do que a curtose de ϵ_t .

$$\begin{aligned} f(\mathbf{x}, \mathbf{h} | \mu, \phi, \sigma_\eta^2) &= f(\mathbf{x} | \mathbf{h}, \mu, \phi, \sigma_\eta^2) f(\mathbf{h} | \mu, \phi, \sigma_\eta^2) \\ &= \prod_{t=1}^n f(x_t | h_t) \prod_{t=2}^n f(h_t | h_{t-1}, \mu, \phi, \sigma_\eta^2) f(h_1 | \phi, \sigma_\eta^2) \end{aligned}$$

$$x_t | h_t \sim N(0, e^{h_t})$$

$$h_1 | \phi, \sigma_\eta \sim N\left(\mu, \frac{\sigma_\eta^2}{1 - \phi^2}\right)$$

$$h_t | h_{t-1}, \mu, \phi, \sigma_\eta \sim N(\mu + \phi(h_{t-1} - \mu), \sigma_\eta^2), t = 2, \dots, n$$

h_1 tem a distribuição estacionária de um processo AR(1).

$$f(x_t|h_t) \propto e^{-h_t/2} \exp\left\{-\frac{1}{2e^{h_t}}x_t^2\right\}$$

$$f(h_t|h_{t-1}, \mu, \phi, \sigma_\eta) \propto \frac{1}{\sigma_\eta} \exp\left\{-\frac{1}{2\sigma_\eta^2}(h_t - \mu - \phi(h_{t-1} - \mu))^2\right\},$$

$t = 2, \dots, n$

$$f(h_1|\phi, \sigma_\eta) \propto \frac{\sqrt{1-\phi^2}}{\sigma_\eta} \exp\left\{-\frac{1-\phi^2}{2\sigma_\eta^2}(h_1 - \mu)^2\right\}$$

$$\mu \sim N(b_\mu, B_\mu)$$

$$\sigma_\eta^2 \sim \text{Gamma}\left(\frac{1}{2}, \frac{1}{2B}\right), E(\sigma_\eta^2) = B, \text{Var}(\sigma_\eta^2) = 2B^2$$

$$\frac{\phi + 1}{2} \sim \text{Beta}(a_0, b_0)$$

$$p(\phi) = \frac{1}{2B(a_0, b_0)} \left(\frac{1 + \phi}{2}\right)^{a_0 - 1} \left(\frac{1 - \phi}{2}\right)^{b_0 - 1}$$

$$E(\phi) = \frac{2a_0}{a_0 + b_0} - 1$$

$$\text{Var}(\phi) = \frac{4a_0b_0}{(a_0 + b_0)^2(a_0 + b_0 + 1)}$$

- Em geral é suficiente especificar uma priori pouco informativa para μ com $b_\mu = 0$ e $B_\mu = 0 \geq 100$ para dados diários.
- O valor de B na distribuição a priori de σ_η^2 não costuma ter muita influencia nos resultados mas não deve ser muito pequeno.
- Os valores de a_0 e b_0 na distribuição a priori de ϕ podem ter bastante influencia na forma da distribuição a posteriori especialmente para séries não muito longas ($n < 1000$).
- Os valores $a_0 = 20$ e $b_0 = 1.5$ implicam em $E(\phi) = 0.86$, $Var(\phi) = 0.11^2$, colocam pouca probabilidade em $\phi < 0$ e costumam funcionar bem na prática.

Pacote `stochvol` para o modelo SV

Estimação Bayesiana do modelo SV com erros normais. Disponível em

<https://cran.r-project.org/web/packages/stochvol/index.html>

Retorna valores simulados via MCMC da distribuição conjunta de $\mathbf{h}, \mu, \phi, \sigma_{\eta} | \mathbf{x}$.

```
> library(stochvol)
> args(svsample)
```

```
function (y, draws = 10000, burnin = 1000, designmatrix = NA,
  priormu = c(0, 100), priorphi = c(5, 1.5), priorsigma = 1,
  priornu = NA, priorbeta = c(0, 10000), priorlatent0 = "stationary",
  thinpara = 1, thinlatent = 1, thintime = 1, keeptau = FALSE,
  quiet = FALSE, startpara, startlatent, expert, ...)
NULL
```

Exemplo. Modelo SV com erros normais para os retornos diários de câmbio Libra Esterlina versus Dolar Americano. Período amostral 1/10/81 a 28/6/85.

Hiperparâmetros especificados,

$$b_{\mu} = 0$$

$$B_{\mu} = 10$$

$$a_0 = 20$$

$$b_0 = 1.5$$

$$B = 0.2$$

```
> library(stochvol)
```

```
> draws = 2000
```

```
> burnin= 2000
```

```
> priormu = c(0,10)
```

```
> priorphi= c(20,1.5)
```

```
> priorsigma = 0.2
```

```
> v = svsample(y,draws,burnin,priormu,designmatrix=NA,prior
```

Iterations = 2001:4000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 2000

1. Empirical mean and standard deviation for each variable,
plus standard error of the mean:

	Mean	SD	Naive SE	Time-series SE
mu	-0.9276	0.27451	0.0061381	0.009532
phi	0.9696	0.01214	0.0002714	0.001142
sigma	0.1910	0.03123	0.0006984	0.004559

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
mu	-1.3907	-1.0771	-0.9413	-0.7958	-0.4279
phi	0.9438	0.9616	0.9705	0.9783	0.9906
sigma	0.1349	0.1680	0.1889	0.2128	0.2560

Summary of 2000 MCMC draws after a burn-in of 2000.

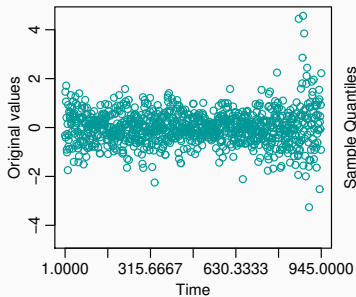
Prior distributions:

```
mu          ~ Normal(mean = 0, sd = 10)
(phi+1)/2   ~ Beta(a0 = 20, b0 = 1.5)
sigma^2     ~ 0.2 * Chisq(df = 1)
```

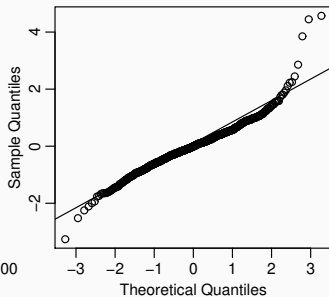
Posterior draws of parameters (thinning = 1):

	mean	sd	5%	50%	95%	ESS
mu	-0.928	0.275	-1.30	-0.941	-0.534	829
phi	0.970	0.012	0.95	0.971	0.988	113
sigma	0.191	0.031	0.14	0.189	0.244	47
exp(mu/2)	0.636	0.163	0.52	0.625	0.766	829
sigma^2	0.037	0.012	0.02	0.036	0.059	47

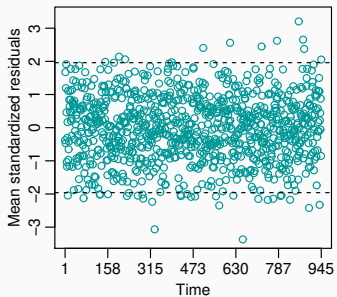
Original data



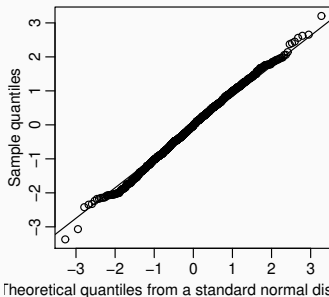
Q-Q plot for original data



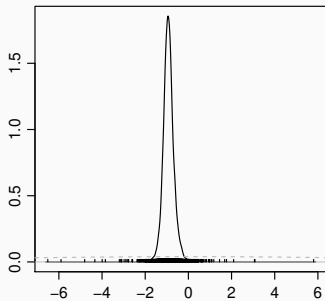
Residual plot



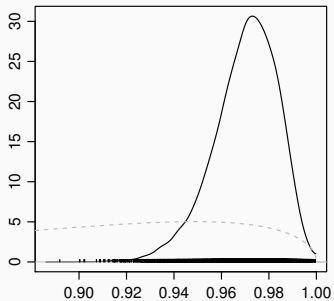
Q-Q plot for mean standardized resid



Density of mu

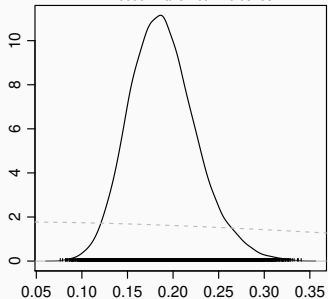


Density of phi



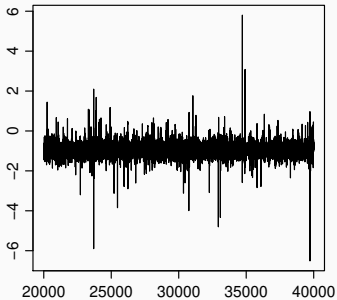
Density of sigma

N = 20000 Bandwidth = 0.03153

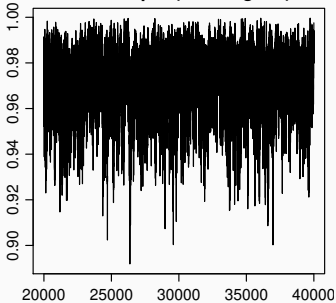


N = 20000 Bandwidth = 0.001947

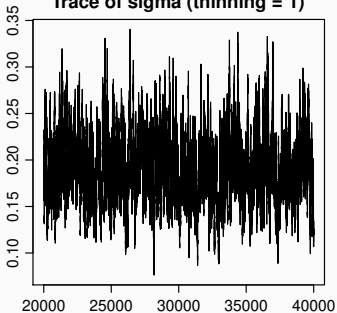
Trace of mu (thinning = 1)



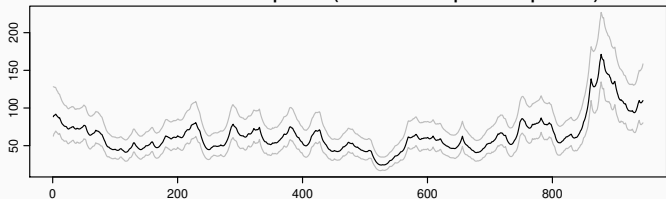
Trace of phi (thinning = 1)



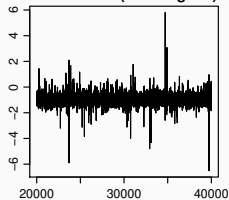
Trace of sigma (thinning = 1)



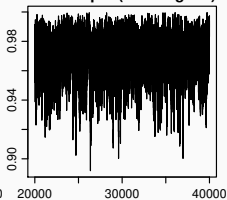
Estimated volatilities in percent (5% / 50% / 95% posterior quantiles)



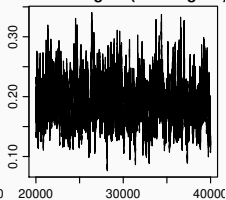
Trace of μ (thinning = 1)



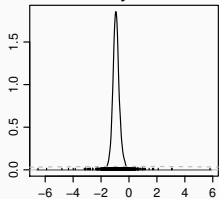
Trace of ϕ (thinning = 1)



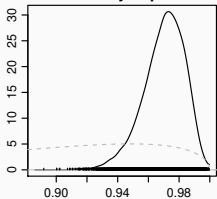
Trace of σ (thinning = 1)



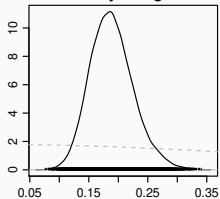
Density of μ



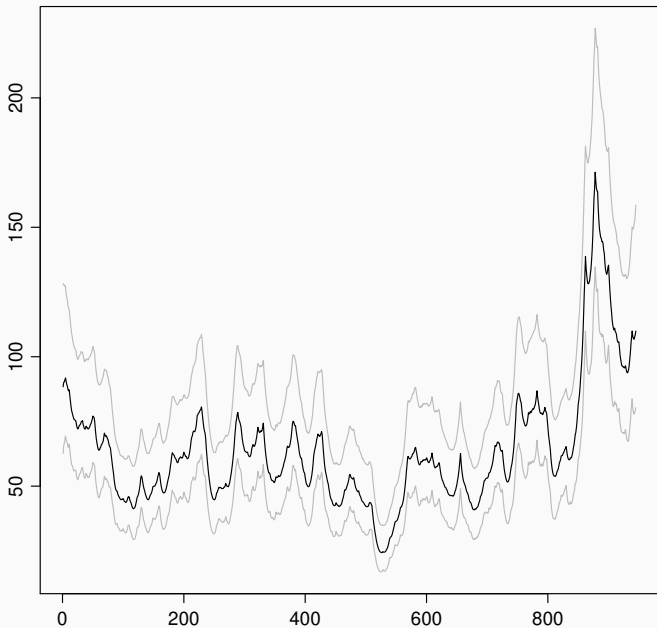
Density of ϕ



Density of σ



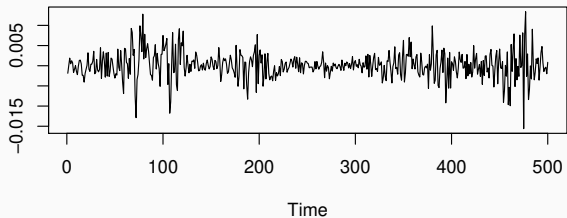
Estimated volatilities in percent (5% / 50% / 95% posterior quantil



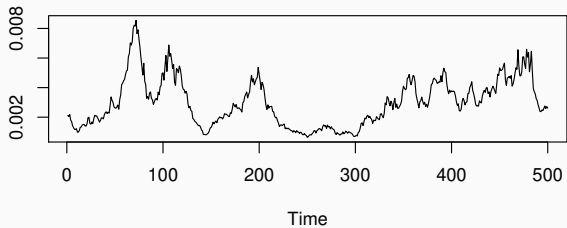
Exemplo. Estimando um modelo SV altamente persistente ($\phi = 0.999$).

```
> phi = 0.999
> sim = svsim(500, mu = -10, phi = phi, sigma = 0.2)
> y = sim$y
> draws = 5000
> burnin= 5000
> priormu = c(-10,1)
> priorphi= c(20,1.5)
> priorsigma = 0.2
> draws = svsample(y,draws,burnin,priormu,designmatrix=NA,
+                priorphi,priorsigma,quiet=T)
```

Log-returns



Volatilities



Iterations = 5001:10000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 5000

1. Empirical mean and standard deviation for each variable, plus standard error of the mean:

	Mean	SD	Naive SE	Time-series SE
mu	-11.6053	0.57801	0.0081742	0.0159942
phi	0.9728	0.01319	0.0001865	0.0005096
sigma	0.2726	0.04102	0.0005802	0.0028167

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
mu	-12.4700	-11.9675	-11.6974	-11.3625	-10.1245
phi	0.9430	0.9651	0.9740	0.9823	0.9944
sigma	0.2003	0.2443	0.2691	0.2985	0.3604

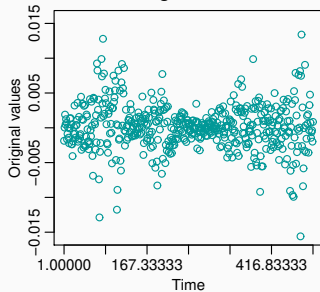
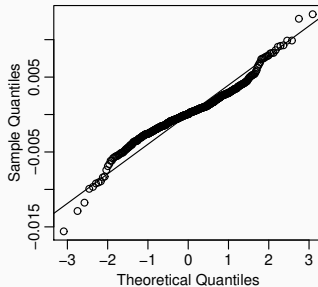
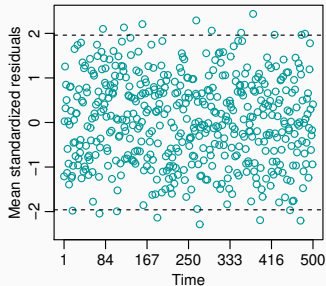
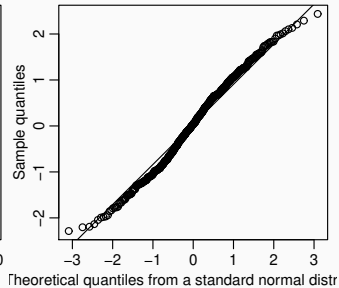
Summary of 5000 MCMC draws after a burn-in of 5000.

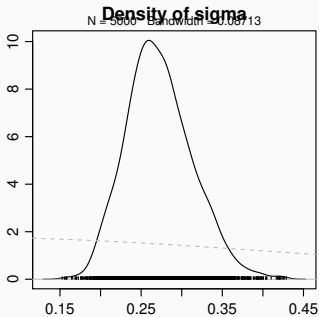
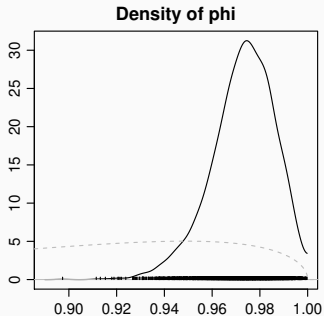
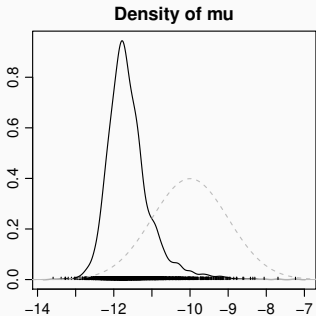
Prior distributions:

mu ~ Normal(mean = -10, sd = 1)
(phi+1)/2 ~ Beta(a0 = 20, b0 = 1.5)
sigma^2 ~ 0.2 * Chisq(df = 1)

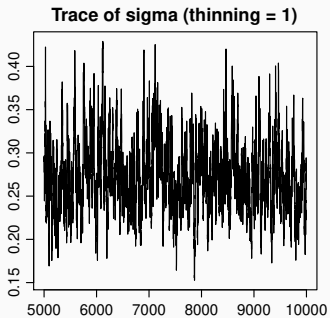
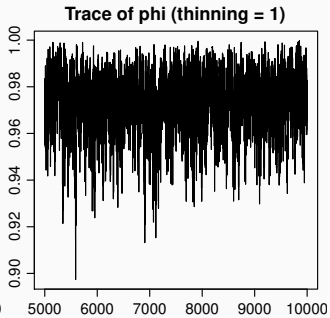
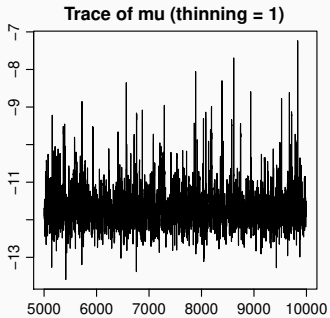
Posterior draws of parameters (thinning = 1):

	mean	sd	5%	50%	95%	ESS
mu	-11.6053	0.5780	-12.3300	-11.6974	-10.5413	1306
phi	0.9728	0.0132	0.9487	0.9740	0.9921	670
sigma	0.2726	0.0410	0.2095	0.2691	0.3439	212
exp(mu/2)	0.0032	0.0013	0.0021	0.0029	0.0051	1306
sigma^2	0.0760	0.0232	0.0439	0.0724	0.1183	212

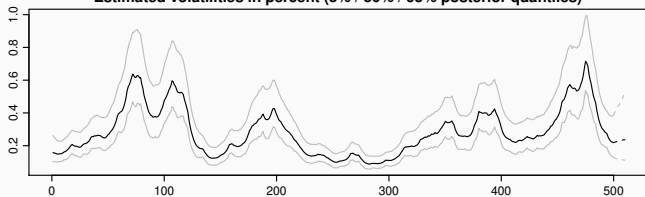
Original data**Q-Q plot for original data****Residual plot****Q-Q plot for mean standardized resid**



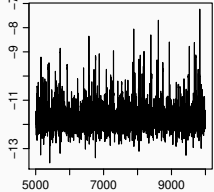
N = 5000 Bandwidth = 0.002481



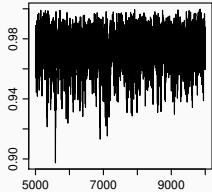
Estimated volatilities in percent (5% / 50% / 95% posterior quantiles)



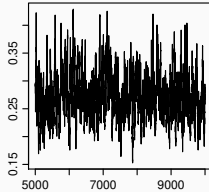
Trace of μ (thinning = 1)



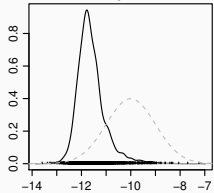
Trace of ϕ (thinning = 1)



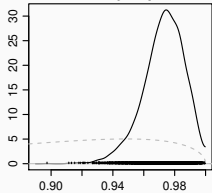
Trace of σ (thinning = 1)



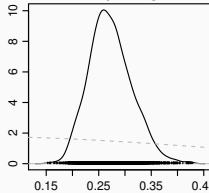
Density of μ



Density of ϕ



Density of σ



Performance preditiva e ajuste

- Densidade preditiva,

$$\begin{aligned} p(x_{n+1}|x_1, \dots, x_n) &= \int p(x_{n+1}|x_1, \dots, x_n, \boldsymbol{\theta}, \mathbf{h}) p(\boldsymbol{\theta}, \mathbf{h}|x_1, \dots, x_n) \\ &= \int p(x_{n+1}|\mathbf{x}, \boldsymbol{\theta}, \mathbf{h}) p(\boldsymbol{\theta}, \mathbf{h}|\mathbf{x}) d\boldsymbol{\theta} d\mathbf{h}. \end{aligned}$$

- No modelo SV esta integral tem dimensão $n + 3$ e não pode ser resolvida analiticamente.
- Dados $\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(N)}$ e $\mathbf{h}^{(1)}, \dots, \mathbf{h}^{(N)}$ obtém-se uma aproximação de Monte Carlo,

$$p(x_{n+1}|\mathbf{x}) \approx \frac{1}{N} \sum_{j=1}^N p(x_{n+1}|\mathbf{x}, \boldsymbol{\theta}^{(j)}, \mathbf{h}^{(j)}),$$

para um ponto arbitrário x_{n+1} .

Dados $\theta^{(1)}, \dots, \theta^{(N)}$ e $\mathbf{h}^{(1)}, \dots, \mathbf{h}^{(N)}$, para cada iteração j ,

- simule $\eta_{n+1}^{(j)} \sim N(0, \sigma_\eta^{2(j)})$,
- calcule,

$$h_{n+1}^{(j)} = \mu^{(j)} + \phi^{(j)}(h_n^{(j)} - \mu^{(j)}) + \eta_{n+1}^{(j)},$$

- simule uma amostra de tamanho L dos erros, $\epsilon_{n+1}^{(j,1)}, \dots, \epsilon_{n+1}^{(j,L)}$,
- obtenha uma amostra de retornos,

$$x_{n+1}^{(j,1)} = \epsilon_{n+1}^{(j,1)} \exp(h_{n+1}^{(j)}/2)$$

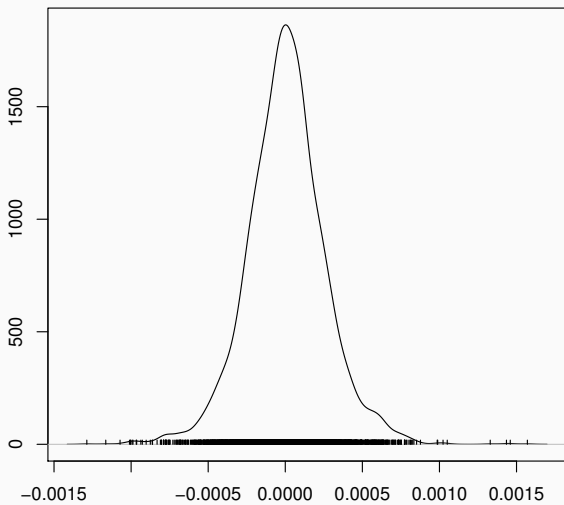
$$\vdots$$

$$x_{n+1}^{(j,L)} = \epsilon_{n+1}^{(j,L)} \exp(h_{n+1}^{(j)}/2)$$

- Calcule,

$$\tilde{x}_{n+1}^{(j)} \approx \frac{1}{L} \sum_{l=1}^L x_{n+1}^{(j,l)}$$

```
> N = nrow(draws$para)
> n = length(y)
> L = 100
> mu = draws$para[,"mu"]
> phi= draws$para[,"phi"]
> sigma = draws$para[,"sigma"]
> eps = matrix(rnorm(N*L),N,L)
> hn = draws$latent[,n]
> eta = rnorm(N,0,sigma)
> hpred = mu + phi*(hn-mu) + eta
> xpred = exp(hpred/2) * eps
> xsample = as.mcmc(apply(xpred,1,mean))
```



N = 5000 Bandwidth = 4.26e-05

Distribuição Preditiva em modelos GARCH

Dados $\theta^{(1)}, \dots, \theta^{(N)}$, para cada iteração j ,

- calcule,

$$h_{n+1}^{(j)} = c^{(j)} + \alpha^{(j)} x_n^2 + \beta^{(j)} h_n^{(j)},$$

- simule uma amostra de tamanho L dos erros, $\epsilon_{n+1}^{(j,1)}, \dots, \epsilon_{n+1}^{(j,L)}$,
- obtenha uma amostra de retornos,

$$\begin{aligned} x_{n+1}^{(j,1)} &= \epsilon_{n+1}^{(j,1)} \sqrt{h_{n+1}^{(j)}} \\ &\vdots \\ x_{n+1}^{(j,L)} &= \epsilon_{n+1}^{(j,L)} \sqrt{h_{n+1}^{(j)}} \end{aligned}$$

- VaR (*Value at Risk*): medida do risco de investimentos.
- Qual a perda máxima potencial durante um período de tempo com uma dada probabilidade?
- As perdas financeiras durante um intervalo de tempo s devem exceder o VaR durante uma proporção p do tempo.
- Definição probabilística. Num período de tempo $\nabla t = I$ houve uma perda descrita por uma variável aleatória L , tal que,

$$Pr(L \leq VaR) = \alpha.$$

- VaR é o quantil de ordem α da distribuição da variável aleatória L .
- Em geral usa-se $\alpha = 0.01$ ou $\alpha = 0.05$.
- Qual a função de distribuição de L , $F_I(L)$?

Dados os valores de $\mathbf{x} = (x_1, \dots, x_n)$, $\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(N)}$ e $\mathbf{h}^{(1)}, \dots, \mathbf{h}^{(N)}$,

- Uma estimativa pontual do valor em risco 1 passo a frente é,

$$E(\text{VaR}_{n+1}|\mathbf{x}) \approx \frac{1}{N} \sum_{j=1}^N \text{VaR}_{n+1}^{(j)} \quad (2)$$

sendo $\text{VaR}_{n+1}^{(j)}$ o valor em risco 1 passo a frente na iteração j .

- Porém $\text{VaR}_{n+1}^{(j)}$ é desconhecido e não pode ser obtido analiticamente.

- Simule,

$$h_{n+1}^{(j)} \sim N(\mu^{(j)} + \phi^{(j)}(h_n^{(j)} - \mu^{(j)}), \sigma_\eta^{2(j)})$$

- simule uma amostra de tamanho L dos erros,

$$\epsilon_{n+1}^{(j,1)}, \dots, \epsilon_{n+1}^{(j,L)} \sim N(0, 1),$$

- obtenha uma amostra de retornos em $t = n + 1$,

$$x_{n+1}^{(j,1)} = \epsilon_{n+1}^{(j,1)} \exp(h_{n+1}^{(j)}/2)$$

$$\vdots$$

$$x_{n+1}^{(j,L)} = \epsilon_{n+1}^{(j,L)} \exp(h_{n+1}^{(j)}/2)$$

- obtenha o $VaR_{n+1}^{(j)}$ aproximado como o quantil de ordem p desta amostra $x_{n+1}^{(j,1)}, \dots, x_{n+1}^{(j,L)}$.
- Repita para $j = 1, \dots, N$ e aplique a equação (2).

VaR em Modelos GARCH

No modelo GARCH(1,1) em $t + 1$,

$$\begin{aligned}x_{t+1} &= \epsilon_{t+1} \sqrt{h_{t+1}}, \\h_{t+1} &= c + \alpha x_t^2 + \beta h_t\end{aligned}$$

$$VaR_{t+1} = F_{\epsilon}^{-1}(\alpha) \sqrt{h_{t+1}}$$

sendo $F_{\epsilon}^{-1}(\alpha)$ o quantil de ordem α da distribuição de ϵ_t .

Dados os valores de $\mathbf{x} = (x_1, \dots, x_n)$ e $\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(N)}$,

- Obtenha a volatilidade no tempo $t = n + 1$,

$$h_{n+1}^{(j)} = c^{(j)} + \alpha^{(j)} x_n^2 + \beta^{(j)} h_n^{(j)}$$

- simule uma amostra de tamanho L dos erros em $t = n + 1$,

$$\epsilon_{n+1}^{(j,1)}, \dots, \epsilon_{n+1}^{(j,L)} \sim N(0, 1),$$

- obtenha uma amostra de tamanho L de retornos em $t = n + 1$,

$$x_{n+1}^{(j,1)} = \epsilon_{n+1}^{(j,1)} \sqrt{h_{n+1}^{(j)}}$$

$$\vdots$$

$$x_{n+1}^{(j,L)} = \epsilon_{n+1}^{(j,L)} \sqrt{h_{n+1}^{(j)}}$$

- obtenha o $VaR_{n+1}^{(j)}$ aproximado como o quantil de ordem p desta amostra $x_{n+1}^{(j,1)}, \dots, x_{n+1}^{(j,L)}$.
- Repita para $j = 1, \dots, N$ e aplique a equação (2).