Forecast volatility and Value at Risk Modeling

Romain Lafarguette, Ph.D. Amine Raboun, Ph.D.

Quants & IMF External Experts

romainlafarguette.github.io/ amineraboun.github.io/

Singapore Training Institute, 17 April 2023

1 Introduction

2 ARCH Models

3 GARCH Models

4 GARCH Extensions

5 Value At Risk

Challenges in estimating volatility

- Daily volatility is unobserved, and can not be derived from the daily returns R_t because there is only one observation in a trading day t
- Volatility of price returns is not static, It changes frequently for several reasons

Why does volatility change?

- **News Announcements**: Macroeconomic and earnings announcement. As new information arrives, uncertainty rises regarding interpreting it and reshuffling portfolios.
- **State of Uncertainty**: Brexit, Trump election, Covid-19, SVB
- **Illiquidity**: Price movement upon taking directional bets on illiquid assets is high
- **Volatility Feedback**: Market-Makers behavior, fleeing the order book when volatility increase (when it matters the most)
- **Leverage**:An price declines, companies become more leveraged (debt-to-equity ratio up) and riskier

Volatility Estimates

1. **Realized Volatility**, but on what window ?

$$\hat{\sigma} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(r_t - \hat{\mu})}$$

2. **Implied volatility** the volatility which when input in an option pricing model (such as Black-Scholes) will return the market price of the option. Example: the CBOE Volatility Index (ticker: VIX).
3. **High Frequency Data Estimators**. The realized volatility is computed as the sum of squared intraday returns (Andersen and Bollerslev, 1998).
4. The **Conditional Volatility** issued from dynamic models such as the **ARCH** and **GARCH** type models $\mathbb{E}_t[\sigma_{t+1}^2]$

Volatility Estimates

1. **Realized Volatility**, but on what window ?

$$\hat{\sigma} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(r_t - \hat{\mu})}$$

2. **Implied volatility** the volatility which when input in an option pricing model (such as Black-Scholes) will return the market price of the option. Example: the CBOE Volatility Index (ticker: VIX).
3. **High Frequency Data Estimators**. The realized volatility is computed as the sum of squared intraday returns (Andersen and Bollerslev, 1998).

4. The **Conditional Volatility** issued from dynamic models such as the **ARCH** and **GARCH** type models $\mathbb{E}_t[\sigma_{t+1}^2]$

Stylized Facts of Financial Time series

How ARCH/GARCH models cover the properties of financial time series?

1. The returns are stationary
2. Absence of autocorrelations
3. Heavy tails
4. Asymmetry
5. Volatility clustering
6. Aggregational Gaussianity
7. ARCH effect
8. Leverage effect

ARCH Models

The ARCH model has been introduced by **Engle (1982)**

ARCH = **A**uto**R**egressive **C**onditional **H**eteroskedasticity
Robert F. Engle Nobel Prize 2003

ARCH Models

The ARCH model has been introduced by **Engle (1982)**

ARCH = **A**uto**R**egressive **C**onditional **H**eteroskedasticity
Robert F. Engle Nobel Prize 2003

- The squared return follows an **autoregressive model**.
- The term **heteroscedasticity** refers to a time-varying variance.
- In an ARCH model, it is the **conditional variance** (and not the variance itself) that changes with time, in a specific way, depending on the available data.

ARCH Models

Definition: ARCH(q)

The process $X_t, t \in \mathbb{Z}$ is said to be an ARCH(q) process, if

$$X_t = Z_t \sigma_t$$

where Z_t is a sequence of independent and identically distributed (i.i.d.) random variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and σ_t is a non-negative process such that

$$\sigma_t^2 = \alpha_0 + \sum_{i-1}^{P} \alpha_i X_{t-i}^2$$

with $\alpha_0 > 0,\ \alpha_i \in \mathbb{R}, \quad \forall i < P, \alpha_p \in \mathbb{R}^*$ and $\sum_{i-1}^{P} \alpha_i < 1$

Focus on ARCH(1)

Definition: ARCH(1)

The process $X_t, t \in \mathbb{Z}$ is said to be an ARCH(1) process, if

$$X_t = Z_t \sigma_t$$

where Z_t is a sequence of independent and identically distributed (i.i.d.) random variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and σ_t is a non-negative process such that

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$$

with $\alpha_0 > 0$ and $0 \leqslant \alpha_1 < 1$

Conditional Variance

Definition

The process σ_t^2 corresponds to the **conditional variance** of X_t

$$\mathbb{V}(X_t|\mathcal{F}_{t-1}) \equiv \mathbb{V}(X_t|\underline{X}_{t-1}) = \sigma_t^2$$

where $\mathcal{F}_{t-1} \equiv \underline{X}_{t-1} = \{X_{t-1}, X_{t-2}, ...\}$ is the information set available at time $t-1$

Some authors denote the conditional variance by h_t, with

$$X_t = Z_t\sqrt{h_t}$$
$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2$$

Interpretation

Consider an ARCH(1) process

$$X_t = Z_t \sigma_t$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$$

Then we have

$$\mathbb{V}(X_t | \underline{X}_{t-1}) = \mathbb{V}(Z_t \sigma_t | \underline{X}_{t-1}) = \sigma_t^2 \mathbb{V}(Z_t | \underline{X}_{t-1}) = \sigma_t^2 \mathbb{V}(Z_t) = \sigma_t^2$$

- Given the past information $\underline{X}_{t-1}$, the conditional variance $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$ is **deterministic**, since x_{t-1} is a constant
- The process $Z_t, t \in \mathbb{Z}$ is an IID noise, so $\mathbb{V}(Z_t | \underline{X}_{t-1}) = \mathbb{V}(Z_t)$ i.e. there is no memory in Z_t
- The normalization $\mathbb{V}(Z_t) = 1$ is not a restriction: the scaling implied by any other variance would be absorbed by the parameters α_0 and α_1

Key properties of ARCH models

If $X_t,\ t \in \mathbb{Z}$ has an ARCH(1) representation with Gaussian innovations, then

1. X_t^2 has an AR(1) representation
2. X_t is a martingale difference
3. X_t is a stationary process under some conditions on the parameters
4. X_t is (unconditionally) Homoscedastic
5. X_t is conditionally Heteroscedastic
6. The (marginal) distribution of X_t is leptokurtic
7. The conditional distribution of X_t is normal

Property 1:

Definition

If $\{X_t, t \in \mathbb{Z}\}$ has an ARCH(1) representation, with

$$X_t = Z_t \sigma_t$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$$

If $\{X_t^2, t \in \mathbb{Z}\}$ has an AR(1) representation, with

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \nu_t$$

where ν_t is an innovation process

$$\mathbb{E}[\nu_t | \underline{X}_{t-1}] = 0$$

Consequences X_t^2 and X_{t-k}^2 are correlated $\Rightarrow$ **ARCH effect**

$$\rho_k = \mathbb{C}orr(X_t^2, X_{t-k}^2) \neq 0$$

especially for small values of k

Property 2:

Definition

if $\{X_t, t \in \mathbb{Z}\}$ is an ARCH(1) process, then it is a martingale difference

$$\mathbb{E}[X_t|\mathcal{F}_{t-1}] \equiv \mathbb{E}[X_t|\underline{X}_{t-1}] = 0$$

Consequences

- The very best (linear or nonlinear) predictor of X_t based on the available information at time $t-1$ 1 is simply the trivial predictor, namely the series mean, 0.
- In terms of point forecasting of the series itself, then, the ARCH models offer no advantages over the linear ARMA models.
- This property implies that $\mathbb{C}ov(X_t, X_{t-k}) = 0 \quad \forall k \neq 0$, i.e that the process X_t has **no memory**

Innovation	ARCH model	Output
Z_t i.i.d noise **No correlation** between Z_t and Z_{t-k}	$X_t = Z_t \sigma_t$ $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$	X_t is martingale difference **No correlation** between X_t and X_{t-k} but $\mathbb{C}ov(X_t^2, X_{t-k}^2) \neq 0$

ARCH effect

The daily squared returns often exhibit significant correlations. These autocorrelations are often referred to as an ARCH effect.

Absence of autocorrelation

The **autocorrelation** of asset returns R_t are often insignificant, except for very small intraday time scales ($\approx$ 20 minutes) for which microstructure effects come into play.

Property 3:

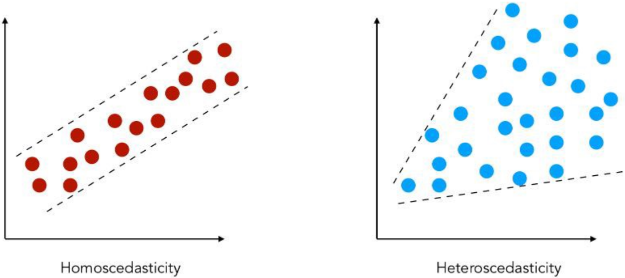
Definition

if $\{X_t, t \in \mathbb{Z}\}$ is an ARCH(1) process, then its two first unconditional moments are finite and constant

$$\mathbb{E}[X_t] = 0, \qquad \mathbb{V}(X_t) = \frac{\alpha_0}{1-\alpha_1}, \qquad \mathbb{C}ov(X_t, X_{t-k}) = 0 \ \forall k \neq 0$$

with $\alpha_0 > 0$ and $0 \leqslant \alpha_1 < 1$

Consequences



- An ARCH(1) process is unconditionally **homoscedastic**
  Unconditional Variance $\mathbb{V}(X_t) = \frac{\alpha_0}{1-\alpha_1} = cst \quad \forall t$
- An ARCH(1) process is conditionally **heteroscedastic**
  Conditional Variance $\mathbb{V}(X_t|\mathcal{F}_{t-1}) = \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$ varies with $\mathcal{F}_{t-1}$
- An ARCH(1) process is (weakly) **stationary**

Property 4:

Definition

if $\{X_t, t \in \mathbb{Z}\}$ is an ARCH(1) process with Gaussian innovations $Z_t \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$, then, its unconditional and conditional Kurtosis coefficients are equal to

$$\mathbb{K}(X_t|\underline{X}_{t-1}) = \frac{\mathbb{E}(X_t^4|\underline{X}_{t-1})}{\mathbb{V}(X_t|\underline{X}_{t-1})^2} = 3$$

$$\mathbb{K}(X_t) = \frac{\mathbb{E}(X_t^4}{\mathbb{V}(X_t)^2} = 3\left(\frac{1-\alpha_1^2}{1-3\alpha_1^2}\right) > 3 \quad if \ \alpha_1 < \sqrt{\frac{1}{3}}$$

Consequences

- The return distribution often exhibits **heavier tails** than those of a normal distribution.
- Even if the innovation Z_t has a normal distribution, the **marginal distribution** of X_t is not Gaussian
- If the innovation Z_t has a normal distribution, the **conditional distribution** of X_t is Gaussian. $X_t|\underline{X}_{t-1} \sim \mathcal{N}(0, \sigma_t^2)$

ARCH models Properties Summary

if $X_t,\ t \in \mathbb{Z}$ is an **ARCH(1)** process with **Gaussian innovation**, then

	Propertry	Consequences / Interpretation
P1	X_t^2 is an AR(1) process	ARCH effect: $\mathbb{C}ov(X_t^2, X_{t-k}^2) \neq 0$ for "small" k
P2	X_t is a martingale difference	$\mathbb{E}[X - t \vert X_{t-1}] = 0$ and $\mathbb{C}ov(X_t, X_{t-k}) = 0 \quad \forall k \neq 0$
P3	$\mathbb{E}[X_t] = 0$, $\mathbb{V}[X_t] = \frac{\alpha_0}{1-\alpha_1}$, $\mathbb{V}[X_t \vert X_{t-1}] = \sigma_t^2$	$\{X_t\}$ is stationary, unconditionally homoscedastic, and conditionally heteroscedastic
P4	$\mathbb{K}[X_t] > 3$, $\mathbb{K}[X_t \vert X_{t-1}] = 3$	The ARCH model generates leptokurtosis, The **marginal distribution** of X_t is not Gaussian, The **conditional distribution** of X_t is Gaussian

ARCH(1) models fits most of the stylized facts of financial series

The properties of the ARCH(1) allow to capture most of the stylized facts of financial data (cf. Chapter 1)

1. **The returns are stationary** $\Rightarrow X_t$ is stationary
2. **Absence of autocorrelations** $\Rightarrow X_t$ is a martingale difference
3. **Heavy tails** $\Rightarrow \mathbb{K}(X_t)$ may be larger than 3 given the value of α_1
4. **Asymmetry**
5. **Volatility clustering** $\Rightarrow \mathbb{C}ov(X_t^2, X_{t-k}^2) \neq 0$
6. **Aggregational Gaussianity** $\Rightarrow$ The marginal distribution of X_t is not normal
7. **ARCH effect** $\Rightarrow {X_t}^2$ has an AR(1) representation and $\mathbb{C}ov(X_t^2, X_{t-k}^2) \neq 0$
8. **Leverage effect**

Weaknesses of ARCH Models

Tsay (2002) identifies three main limits of the ARCH models.

1. The model assumes that positive and negative shocks have the same effects on volatility because it depends on the square of the previous shocks. In practice, the return of a financial asset responds differently to positive and negative shocks.
2. The ARCH model is rather restrictive. For instance, the fourth moment $\mathbb{E}(X_t^4)$ exists only if $\alpha_1^2 < \frac{1}{3}$
3. The ARCH model does not provide any insight for understanding the source of volatility. It only provides a *mechanical way* to describe the behavior of the conditional variance. It gives no indication of what causes such behavior to occur.

Building an ARCH model

Denote by R_t the **daily return** of an asset or a portfolio at time t.

$$R_t = \underbrace{\mu_t}_{\substack{\text{conditional}\\ \text{mean model}}} + \underbrace{\epsilon_t}_{\substack{\text{innovation}\\ \text{(martingale diff)}}}$$

$$\epsilon_t = Z_t \sigma_t$$

$$\sigma_t^2 = \underbrace{\alpha_0 + \sum_{i-1}^{P} \alpha_i X_{t-i}^2}_{\substack{\text{conditional}\\ \text{variance model}}}$$

follows the structure of a **(conditional) volatility model**

$$\mu_t \equiv \mathbb{E}(R_t | \mathcal{F}_{t-1}) = \mu_t(\underline{R}_{t-1}; \theta)$$

$$\sigma_t^2 \equiv \mathbb{V}(R_t | \mathcal{F}_{t-1}) = \sigma_t^2(\underline{R}_{t-1}; \theta)$$

where θ denotes the **set of parameters** for the conditional mean and variance and μ_t is typically an **ARMA**-type model .

Example

Example ARMA(1,1)-ARCH(1)

The process $\{R_t,\ t \in \mathbb{Z}\}$ has an **ARMA(1,1)-ARCH(1)** representation if

$$R_t = \phi_0 + \phi_1 R_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t$$

$$\epsilon_t = Z_t \sigma_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$$

where Z_t is a sequence of i.i.d. variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$.
We have

$$\mu_t = \mathbb{E}(R_t | \mathcal{F}_{t-1}) = \phi_0 + \phi_1 R_{t-1}$$

$$\sigma_t^2 = \mathbb{V}(R_t | \mathcal{F}_{t-1}) = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$$

and $\theta = (\phi_0, \phi_1, \theta_1, \alpha_0, \alpha_1)'$ is the vector of parameters to estimate

Model Checking

- For an ARCH model, the **standardized innovations** $Z_t = \frac{\epsilon_t}{\sigma_t}$ are i.i.d. random variates (following either a standard normal or Student-t distribution).
- Therefore, one can check the adequacy of a fitted ARCH model by examining the series of **standardized residuals**

$$\hat{Z}_t = \frac{\hat{\epsilon}_t}{\hat{\sigma}_t}$$

Tsay (2002) recommends three types of tests on the series $\{\hat{z}_t\}_{t=1}^T$

1. The Ljung-Box Q-statistics of $\hat{z}_t$ can be used to check the adequacy of the mean equation.
2. The Ljung-Box Q-statistics of $\hat{z}_t{}^2$ can be used to check the adequacy of the volatility equation.
3. The skewness, kurtosis, and QQ-plot of $\hat{z}_t$ can be used to check the validity of the distribution assumption on Z_t.

Forecasting

We have to distinguish:

- The forecasts on the series R_t itself (typically the returns).
- The forecasts on the volatility (or the variance) of R_t.

Forecasting of the series R_t

The best linear forecast of R_t given the information set $\mathcal{F}_{t-1}$ will be no different **with or without** an ARCH error

Example AR(1)-ARCH(1)

$$R_t = \phi_0 + \phi_1 R_{t-1} + \epsilon_t$$

$$\epsilon_t = Z_t \sigma_t, \quad Z_t \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$$

then, the conditional variance forecast at horizon h = 1 is given by

$$\hat{\sigma}_{t+1|t}^2 = \mathbb{V}(R_{t+1}|\mathcal{F}_t) = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 = \alpha_0 + \alpha_1 \left(R_t - \phi_0 - \phi_1 R_{t-1}\right)^2$$

GARCH Models

Due to the large persistence in volatility, ARCH models often require a large p to fit the data. A more parsimonious specification is provided by **GARCH** models.

GARCH = **G**eneralized **A**uto**R**egressive **C**onditional **H**eteroskedasticity

Bollerslev, T. (1986)

Definition: GARCH Model

The stochastic process $\{\epsilon_t,\ t \in \mathbb{Z}\}$ is said to be a **GARCH**(p,q) process if:

$$\epsilon_t = Z_t \sigma_t$$

where Z_t is a sequence of i.i.d variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and σ_t is a non-negative process such that:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$$

with $\omega > 0$, $\forall\ i, (\alpha_i, \beta_i)\ \in\ \mathbb{R}^{+,2}$ and $\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{p} \beta_i < 1$

GARCH: Intuition

- The conditional variance of a GARCH(p, q) depends on:
  - The first p lag of the ϵ_t^2 (e.g. the squared error terms)
  - The first q lag of the conditional variance σ^2

$$\sigma_t^2 = \omega + \underbrace{\sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2}_{\text{ARCH Components}} + \underbrace{\sum_{i=1}^{p} \beta_i \sigma_{t-i}^2}_{\text{GARCH components}}$$

- The parameters α_i are often called the **ARCH parameters**

- The parameters β_i are often called the **GARCH parameters**

GARCH(1, 1)

Tip

GARCH(1,1) specifications are generally sufficient to capture the dynamics of the conditional variance

Special Case: GARCH(1, 1)

The process $\{\epsilon_t,\ t \in \mathbb{Z}\}$ is said to be a **GARCH**(1,1) if:

$$\epsilon_t = Z_t \sigma_t$$

where Z_t is a sequence of i.i.d variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and σ_t is a non-negative process such that:

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

with $\omega > 0$, $\alpha \geqslant 0, \beta \geqslant 0$ and $\alpha + \beta < 1$

Conditional Variance Persistences

- The conditional variance $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$ depends on two effects:
  - An **intrinsic persistence** effect through the first lag of the conditional variance
  - An **extrinsic persistence** effect

- Following a positive (or negative) shock at time *t-1*, the conditional variance at time t increases (impact effect) and thus it has an impact on $\epsilon_t = Z_t\sigma_t$

$$\text{shock } z_{t-1} \; > \; 0 \Rightarrow \; \epsilon_{t-1} \; \uparrow \Rightarrow \; \sigma_t \; \uparrow \; \ldots$$

- Starting from the next period (at time t), the effect of the shock at $t-1$ on the conditional variance at $t+1$ passes through the conditional variance at time t (intrinsic persistence)

$$... \; \Rightarrow \; \sigma_t \; \uparrow \Rightarrow \; \sigma_{t+1}^2 \; \uparrow$$

- The overall effect of a shock can be decomposed into a **contemporaneous effect**, which depends on α and a **persistence effect** that depends on β

Remarks

It is often the case that:

1. The sum of the estimates of α and β are generally close (but below 1)
2. The estimate of β is generally greater than the one of α
3. The estimate of β is generally larger than 0.90 for daily returns and the estimate of α is below 0.1

Be careful: it is not a general rule, just an observation.

GARCH Properties

GARCH process properties are similar to those of an ARCH process.

1. ARCH properties
2. ϵ_t^2 has an ARMA representation

ARMA representation

If $\{\epsilon_t,\ t \in \mathbb{Z}\}$ has a **GARCH**(p,q) representation with:

$$\epsilon_t = Z_t \sigma_t$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$$

the $\{\epsilon_t^2,\ t \in \mathbb{Z}\}$ has a **ARMA**(max(p,q), q) representation with:

$$\epsilon_t^2 = \omega + \sum_{i=1}^{p} (\alpha_i + \beta_i)\epsilon_{t-i}^2 + \nu_t - \sum_{i=1}^{p} \beta_i \nu_{t-i}$$

with $\nu = \epsilon_t^2 - \sigma_t^2$ is an innovation process, i.e $\mathbb{E}(\nu_t | \mathcal{F}_{t-1})$

Estimation

1. The set of parameters θ of an ARMA-GARCH model is estimated by Maximum Likelihood (ML) or Quasi Maximum Likelihood (QML).
2. When the model is estimated by ML, the most often used distributions for Z_t are:
   1. The **normal distribution**, $Z_t \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$. IMPORTANT: the normality assumption on Z_t does not imply that the return R_t has a normal (marginal) distribution.
   2. The **Student t-distribution**, $Z_t \overset{\text{i.i.d}}{\sim} t(\nu)$, which is symmetric and leptokurtic (if ν is "small").
   3. The **skewed Student t-distribution**, $Z_t \overset{\text{i.i.d}}{\sim} Skewed\ t(\delta, \nu)$, which is asymmetric (if $\delta \neq 1$) and leptokurtic (if ν is "small").
   4. The **Generalized Error Distribution (GED)**, $Z_t \overset{\text{i.i.d}}{\sim} GED(\nu)$, which is symmetric and leptokurtic (if $\nu < 2$).

Remark

Why consider non-Gaussian distributions for the innovation Z_t?

1. The use of a **leptokurtic distribution** for Z_t allows to increase the kurtosis of R_t.

$$\text{Kurtosis of a GARCH process} = \text{kurtosis generated by the model} \\ + \text{ kurtosis of the innovation } Z_t$$

   The kurtosis generated by the model dynamic is generally not sufficient to reproduce the level of kurtosis observed in the financial returns.

2. The use of a **skewed distribution** for Z_t allows to reproduce the skewness observed in the distribution of the financial returns.

$$\text{Skewed distribution for } Z_t \Rightarrow \text{Skewed distribution for } R_t$$

Asymmetric GARCH models

- The GARCH model assumes that positive and negative shocks have the same effects on volatility because it depends on the square of the previous shocks.
- In practice, the return of a financial asset responds differently to positive and negative shocks.
- The GARCH model does not allow to capture the **leverage effect**.

Asymmetric GARCH models

Stylized Fact 8: Leverage Effect

Asset returns are negatively correlated with the changes in their volatilities: this negative correlation is called the leverage effect

- As asset prices decline, companies become more leveraged (debt-to-equity ratios increase) and riskier, and hence their stock prices become more volatile.
- On the other hand, when stock prices become more volatile, investors demand high returns, and hence stock prices go down.
- Many asymmetric GARCH models have been proposed: **GJR-GARCH, TGARCH, EGARCH**, APARCH, VSGARCH, QGARCH, LSTGARCH, ANSTGARCH, etc.

GJR-GARCH

One of the most often used asymmetric models is the **GJR-GARCH** model, where "GJR" stands for Glosten, Jagannathan, and Runkle (1993).

Definition

The process $\{\epsilon_t, t \in \mathbb{Z}\}$ is to be a **GJR-GARCH(1,1)** process, if

$$\epsilon_t = Z_t \sigma_t$$

where Z_t is i.i.d with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and

$$\sigma_t^2 = w + \alpha \epsilon_{t-1}^2 + \gamma \mathbb{I}_{(\epsilon_{t-1} > 0)} \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

with $w > 0, \quad \alpha > 0, \quad \beta > 0, \quad \gamma \in \mathbb{R}$ and where $\mathbb{I}_{(.)}$ is the indicator function that takes a value 1 if the condition is true and 0 otherwise.

GJR-GARCH Interpretation

- The term ϵ_t can be interpreted as a shock (surprise) on the return, since

$$\epsilon_t = R_t - \mu_t = R_t - \mathbb{E}(R_t|\mathcal{F}_{t-1})$$

- In a **GJR-GARCH** model, the influence of the past return shock ϵ_t on the current conditional variance σ_t^2 depends on its sign

$$\sigma_t^2 = w + \alpha\epsilon_{t-1}^2 + \gamma\mathbb{I}_{(\epsilon_{t-1}>0)}\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

$$\frac{\partial\sigma_t^2}{\partial\epsilon_{t-1}^2} = \begin{Bmatrix} \alpha+\gamma & if \epsilon_{t-1} < 0 \\ \alpha & otherwise \end{Bmatrix}.$$

- A **leverage effect** implies that $\gamma > 0$, i.e. the increase in volatility caused by a negative return is larger than the appreciation due a positive return of the same magnitude.

TGARCH model

- The TGARCH, where "T" stands for **Threshold**, is an asymmetric GARCH model designed to capture the leverage effect.
- The TGARCH is similar to the GJR model, different only because of the use of the **conditional volatility**, instead of the variance, in the specification.
- The TGARCH has been introduced by Zakoian (1994).

TGARCH

Definition

The process $\{\epsilon_t, t \in \mathbb{Z}\}$ is to be a **TGARCH(1,1)** process, if

$$\epsilon_t = Z_t \sigma_t$$

where Z_t is i.i.d with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and

$$\sqrt{\sigma_t^2} = w + \alpha_+ \epsilon_{t-1} \mathbb{I}_{(\epsilon_{t-1} > 0)} + \alpha_- \epsilon_{t-1} \mathbb{I}_{(\epsilon_{t-1} < 0)} + \beta \sqrt{\sigma_{t-1}^2}$$

with $(w, \alpha_+, \alpha_-, \beta) \in \mathbb{R}^4$ and $\mathbb{I}_{(.)}$ is the indicator function

TARCH Interpretation

- One advantage of the TGARCH is that it does not require any **positivity constraints** on the parameters, since we have $\forall(w, \alpha_+, \alpha_-, \beta) \in \mathbb{R}^4$

$$\sigma_t = \left(w + \alpha_+ \epsilon_{t-1} \mathbb{I}_{(\epsilon_{t-1}>0)} + \alpha_- \epsilon_{t-1} \mathbb{I}_{(\epsilon_{t-1}<0)} + \beta \sqrt{\sigma_{t-1}^2}\right) \geqslant 0$$

- The **TGARCH** allows to capture an **asymmetry** between positive and negative shocks, as

$$\frac{\partial \sigma_t}{\partial \epsilon_{t-1}} = \begin{Bmatrix} \alpha_- & if \epsilon_{t-1} < 0 \\ \alpha_+ & otherwise \end{Bmatrix}.$$

- The **leverage effect** implies that $|\alpha_- > \alpha_+|$, i.e. the increase in volatility caused by a negative return is larger than the appreciation due to a positive return of the same magnitude

EGARCH model

- "E" stands for **Exponential**. The EGARCH has been introduced by Nelson (1991).
- It was designed to capture both (1) the asymmetric effects and (2) the effects of "big" shocks.

Definition

The process $\{\epsilon_t, t \in \mathbb{Z}\}$ is to be a **EGARCH(1,1)** process, if

$$\epsilon_t = Z_t \sigma_t$$

$$ln(\sigma_t^2) = w + \alpha \; Z_{t-1} + \gamma \left(|Z_{t-1}| - \mathbb{E}(|Z_{t-1}|)\right) + \beta \; ln(\sigma_{t-1}^2)$$

with $(w, \alpha, \beta, \gamma) \in \mathbb{R}^4$ and where Z_t is i.i.d with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$,

EGARCH Interpretation

GARCH Model

$$\epsilon_t = Z_t \sigma_t$$

$$\sigma_t^2 = w + \alpha \underbrace{\epsilon_{t-1}^2}_{\text{depends on} \epsilon_{t-1}} + \beta \sigma_{t-1}^2$$

EARCH Model

$$\epsilon_t = Z_t \sigma_t$$

$$ln(\sigma_t^2) = w + \underbrace{\alpha Z_{t-1} + \gamma \left(|Z_{t-1}| - \mathbb{E}(|Z_{t-1}|) \right)}_{\text{depends on the standardized error} Z_{t-1}} + \beta ln(\sigma_{t-1}^2)$$

EGARCH Interpretation

- The term $(|Z_{t-1}| - \mathbb{E}(|Z_{t-1}|))$ measures the **magnitude** of the (positive or negative) shock
- If the parameter γ is positive, the **"big"** (compared to their expected value) shocks have a stronger impact on the variance than the **"small"** shocks
- The EGARCH model captures the asymmetric effects between positive and negative shocks on the returns, since

$$\frac{\partial \sigma_t}{\partial \epsilon_{t-1}} = \begin{Bmatrix} \gamma - \alpha & if z_{t-1} < 0 \\ \gamma + \alpha & otherwise \end{Bmatrix}.$$

- The **leverage effect**, i.e. the fact that negative shocks at time $t-1$ have a stronger impact on the variance at time t than positive shocks, implies that $\alpha < 0$

Daily returns with EGARCH model

Example EGARCH

Consider **AR**(1)-**EGARCH**(1,1) with Gaussian innovation for the returns $\{\mathbb{R}_t, t \in \mathbb{Z}\}$

$$R_t = \phi_0 + \phi_1 R_{t-1} + \epsilon_t$$

$$\epsilon_t = Z_t \sigma_t$$

$$ln(\sigma_t^2) = w + \alpha\ Z_{t-1} + \gamma \left(|Z_{t-1}| - \sqrt{\frac{2}{\pi}} \right) + \beta\ ln(\sigma_{t-1}^2)$$

or equivalently

$$ln(\sigma_t^2) = w + \alpha\ \left(\frac{\epsilon_{t-1}}{\sigma t - 1} \right) + \gamma \left(|\frac{\epsilon_{t-1}}{\sigma t - 1}| - \sqrt{\frac{2}{\pi}} \right) + \beta\ ln(\sigma_{t-1}^2)$$

with Z_t i.i.d $\mathcal{N}(0,1)$. The vector of parameters to be estimated is $\theta = (\phi_0, \phi_1, w, \alpha, \gamma, \beta)'$

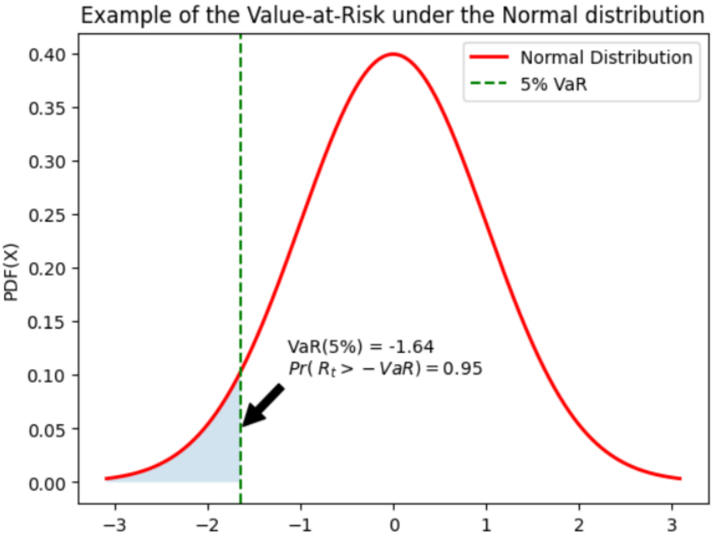
VaR: Intuitive Definition

Definition

The value at risk (VaR) defined for a hedge ratio α % corresponds to the quantile of order α of the distribution of profits and losses (P&L) associated with the holding of an asset or a portfolio of assets over a given period.

Remark: VaR is generally negative (a loss) in a P&L representation. For simplifying, we denote the VaR as a positive value by considering the opposite of the quantile

Value-At-Risk of a Normal Distribution



VaR: Formal Definition

Definition

For a hedge rate of $\alpha\%$, the Value-at-Risk, noted $VaR_t(\alpha)$, corresponds to the opposite of the fractile of order α of the distribution of profits and losses (P&L)

$$VaR_t(\alpha) = -F_{R_t}^{-1}(\alpha)$$

where F_{R_t} denotes the cumulative distribution function associated with the density function $f_{R_t}(.)$

- Sometimes the Value-at-Risk is expressed as a function of confidence level. VaR at hedge rate 1% will be Var(99%)

$$VaR(1-\alpha) = -F_{R_t}^{-1}(\alpha)$$

- the probability of observing a loss greater than the VaR over the holding period is equal by definition to the coverage rate:

$$Pr[R_t > -VaR_t(\alpha)] = \int_{-\infty}^{-VaR_t(\alpha)} f_{R_t}(r)dr = \alpha$$

VaR: Specification

The definition of Value-at-Risk is based on 3 elements

1. The distribution of profit and loss (P&L) of the portfolio or the asset
2. Level of confidence (or equivalently the hedge rate α)
3. The holding period of the asset (or the risk horizon)

Conditional Value-at-Risk

- We define the conditional distribution of P&L, based on a **set of information available at time t**, denoted as Ω_t

$$f_{R_t}(r|\Omega_t)$$

- the conditional distribution may change through time, but we usually consider the case of invariant conditional density (given the explanatory variables)

$$F_{R_t}(r|\Omega) = F_R(r|\Omega) \quad \forall t$$

Definition

For a hedge rate of $\alpha\%$, the conditional Value-at-Risk to a set of information Ω_t, noted $VaR_t(\alpha|\Omega_t)$, equals to the opposite of the fractile of order α of the P&L conditional distribution

$$VaR_t(\alpha|\Omega_t) = -F_{R_t}^{-1}(\alpha|\Omega_t)$$

where $F_{R_t}(r|\Omega_t)$ is the cumulative distribution function associated with the conditional density function $f_{R_t}(r|\Omega_t)$)

Estimation methods

Challenge

Theoretically, at each date t, the return R_t is a random variable admitting a distribution $f_{R_t}(.)$ and a fractile α. However, we have only one observation r_t of this distribution. From this single realization, **without any additional hypothesis**, it is impossible to estimate the quantiles of the distribution $f_{R_t}(.)$ at date t, i.e. the VaR

Estimation methods

Challenge

Theoretically, at each date t, the return R_t is a random variable admitting a distribution $f_{R_t}(.)$ and a fractile α. However, we have only one observation r_t of this distribution. From this single realization, **without any additional hypothesis**, it is impossible to estimate the quantiles of the distribution $f_{R_t}(.)$ at date t, i.e. the VaR

Several estimation methods have been proposed

- **Non-parametric Estimation**: No parametric distribution of P&L is imposed a priori. **Historical Simulation**(HS), **Weighted HS**, **Filtered HS**
- **Semi-parametric Estimation**: CAViaR method, Extreme Values Theory
- **Parametric Estimation**: ARCH, GARCH RiskMetrics

Parametric Estimation

Motivation For any elliptical distribution, the VaR forecast is a linear transformation of the variance (volatility) forecast. Predicting the variance allows to predict the VaR

Garch Model

Under the normal distribution hypothesis of conditional P&L, the VaR forecast for hedge rate α is given by

$$VaR_{t+1|t}(\alpha) = -\mu - \sqrt{h_{t+1}}\Phi^{-1}(\alpha)$$

where h_{t+1} is the conditional variance of returns derived from the GARCH model

Parametric Estimation

RiskMetrics

Developed by JP Morgan in the 90s. The forecasted conditional VaR for the hedge rate α is

$$VaR_{t+1|t}(\alpha) = -\mu - h_t \Phi^{-1}(\alpha)$$

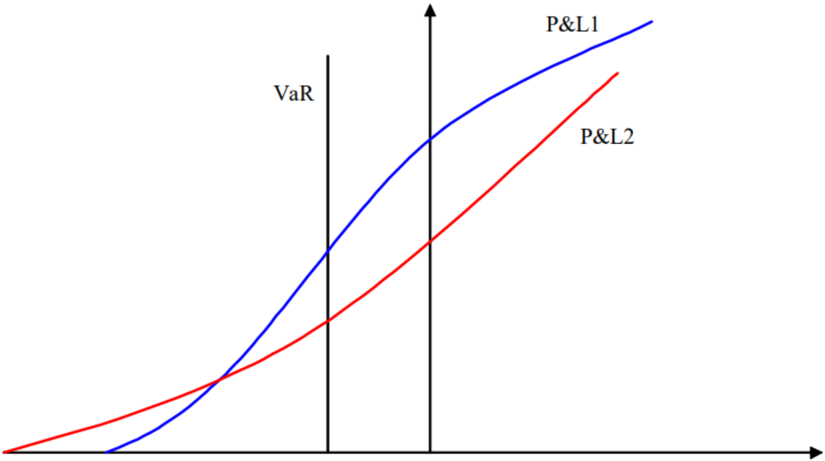
where μ is the expected return and h_t is the conditional variance

$$h_t = \lambda h_{t-1} + (1 - \lambda) r_{t-1}^2$$

where λ is a decay parameter (generally fixed to 0.97)

- The conditional variance in RiskMetrics follows a **EWMA (Exponential Weighted Moving Average)** type process: The forecasted variance is a linear function of past innovations and past variance
- RiskMetricss is a special case of GARCH

Limites of the Value-at-Risk



Limites of the Value-at-Risk

1. This risk measure does not give any information on losses beyond the VaR
2. This measure can lead some agents to voluntarily take more risk in a decentralized risk management system
3. It can lead to a decision maker to choose a project with exorbitantly large losses, as long as these losses do not affect the VaR (because they occur with low probability)
4. the VaR is not a **coherent measure of risk** because the **sub-additivity** property is not respected

$$\rho(X+Y) \leqslant \rho(X) + \rho(Y)$$

Expected Shortfall

ES: Expected shortfall

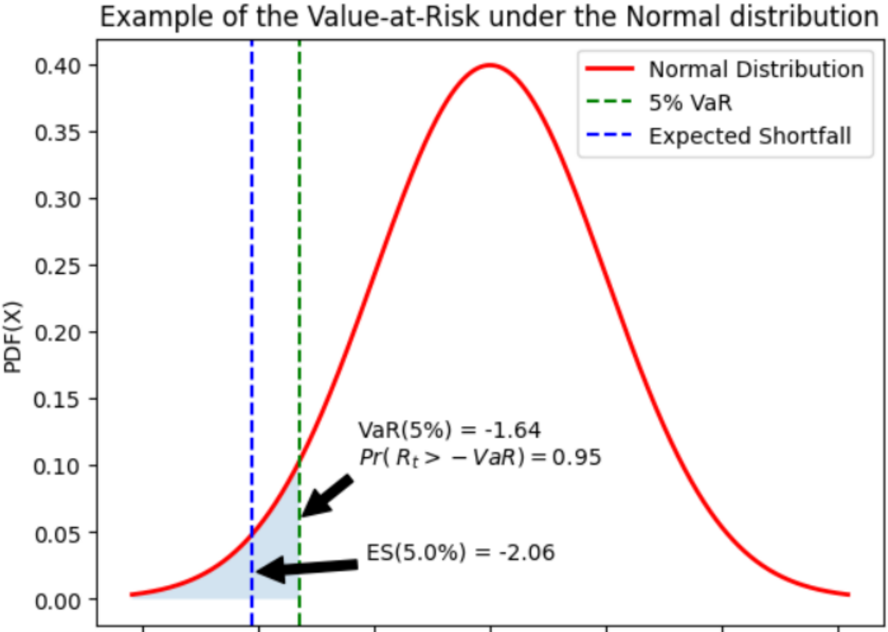
The Expected shortfall (ES) associated with a hedge rate α is the average of the $\alpha\%$ worst expected losses

$$ES_t(\alpha) = -\frac{1}{\alpha}\int_0^{\alpha} F_{R_t}^{-1}(p)dp$$

where F_{R_t} is the cumulative distribution of the density function $f_{R_t}(r)$

- **Remark 1: The Expected Shortfall is sometimes denoted the Conditional Loss or Expected Tail Loss**
- The ES gives average loss in the worst case scenario. i.e in the $\alpha\%$ situations where the losses exceed the Var

Expected Shortfall



Thank you for your attention!

Property 1: *Proof*

Consider the following ARCH process

$$X_t = Z_t \sigma_t$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$$

Add X_t^2 on both sides of the second equation and rearrange, we get

$$\begin{aligned} X_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 + (X_t^2 - \sigma_t^2) \\ &= \alpha_0 + \alpha_1 X_{t-1}^2 + (Z_t^2 \sigma_t^2 - \sigma_t^2) \\ &= \alpha_0 + \alpha_1 X_{t-1}^2 + \sigma_t^2 (Z_t^2 - 1) \end{aligned}$$

$\nu_t = X_t^2 - \sigma_t^2 = \sigma_t^2(Z_t^2 - 1)$ is an innovation, i.e $\mathbb{E}[\nu_t | \underline{X}_{t-1}] = 0$

$$\begin{aligned} \mathbb{E}[\nu_t | \underline{X}_{t-1}] &= \mathbb{E}[\sigma_t^2 (Z_t^2 - 1) | \underline{X}_{t-1}] \\ &= \sigma_t^2 \mathbb{E}[(Z_t^2 - 1) | \underline{X}_{t-1}] \\ &= \sigma_t^2 (\mathbb{E}[Z_t^2 | \underline{X}_{t-1}] - 1) \\ &= \sigma_t^2 (\mathbb{V}(Z_t) - 1) = 0 \end{aligned}$$

Property 2: *Proof*

Consider the following ARCH process

$$X_t = Z_t \sigma_t$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$$

X_t is a martingale di§erence, since

$$\begin{aligned}\mathbb{E}[X_t | \underline{X}_{t-1}] &= \mathbb{E}[Z_t \sigma_t | \underline{X}_{t-1}] \\ &= \sigma_t \mathbb{E}[Z_t | \underline{X}_{t-1}] \\ &= \sigma_t \mathbb{E}[Z_t] \\ &= 0\end{aligned}$$

Since Z_t is an i.i.d process of mean 0

Property 3: *Proof*

Consider an ARCH(1) model X_t

$$\begin{aligned}\mathbb{E}(X_t) &= \mathbb{E}(Z_t\sigma_t) \\ &= \mathbb{E}[\mathbb{E}(Z_t\sigma_t|\underline{X}_{t-1})] \\ &= \mathbb{E}[\sigma_t\mathbb{E}(Z_t|\underline{X}_{t-1})] \\ &= \mathbb{E}[\sigma_t \times 0] \\ &= 0\end{aligned}$$

Since $\mathbb{E}[X_t] = 0$, we have $\mathbb{V}(X_t) = \mathbb{E}[X_t^2]$
We know that X_t^2 has an AR(1) representation with

$$\mathbb{E}[X_t^2] = \mathbb{E}[\alpha_0 + \alpha_1 X_{t-1}^2 + \nu_t] = \alpha_0 + \alpha_1\mathbb{E}[X_{t-1}^2]$$

$$\Leftrightarrow \mathbb{E}[X_t^2] = \frac{\alpha_0}{1-\alpha_1}$$

Property 4: *Proof*

Consider an ARCH(1) model X_t, we have

$$\begin{aligned} \mathbb{E}(X_t^4|\underline{X}_{t-1}) &= \mathbb{E}(Z_t^4\sigma_t^4|\underline{X}_{t-1}) \\ &= \mathbb{E}(Z_t^4|\underline{X}_{t-1})\sigma_t^4 \\ &= \mathbb{E}(Z_t^4)(\sigma_t^2)^2 \end{aligned}$$

$Z_t \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$, so $\mathbb{E}(Z_t^4) = 3$. We get

$$\mathbb{K}(X_t^4|\underline{X}_{t-1}) = \frac{\mathbb{E}(X_t^4|\underline{X}_{t-1})}{\mathbb{V}(X_t^4|\underline{X}_{t-1})^2} = \frac{3(\sigma_t^2)^2}{(\sigma_t^2)^2} = 3$$

The conditional distribution is **mesokurtic**

Property 4: *Proof (cont'd)*

$$
\begin{aligned}
\mathbb{E}[X_t^4] &= \mathbb{E}(\mathbb{E}(X_t^4|\underline{X}_{t-1})) \\
&= 3\mathbb{E}((\alpha_0 + \alpha_1 X_{t-1}^2)^2) \\
&= 3\left(\alpha_0^2 + 2\alpha_0\alpha_1\mathbb{E}(X_{t-1}^2) + \alpha_1^2\mathbb{E}(X_{t-1}^4)\right) \\
&= 3\left(\alpha_0^2 + \frac{2\alpha_0^2\alpha_1}{1-\alpha_1} + \alpha_1^2\mathbb{E}(X_{t-1}^4)\right) \\
&= 3\alpha_0^2\left(\frac{1+\alpha_1}{1-\alpha_1}\right) + 3\alpha_1^2\mathbb{E}(X_{t-1}^4)
\end{aligned}
$$

X_t is stationary, then $\mathbb{E}[X_t^4] = \mathbb{E}[X_{t-1}^4]$, and we know $\mathbb{V}(X_t) = \frac{\alpha_0}{(1-\alpha_1)}$

$$
\mathbb{E}[X_t^4] = \frac{3\alpha_0^2(1+\alpha_1)}{(1-3\alpha_1^2)(1-\alpha_1))}
$$

$$
\mathbb{K}(X_t) = \frac{\mathbb{E}(X_t^4}{\mathbb{V}(X_t)^2} = \frac{3\alpha_0^2(1+\alpha_1)}{(1-3\alpha_1^2)(1-\alpha_1))}\frac{\alpha_0}{(1-\alpha_1)} = 3\left(\frac{1-\alpha_1^2}{1-3\alpha_1^2}\right) > 3
$$

The Kurtosis is finite and positive as soon as $\alpha_1^2 < 1/3$. Moreover, the conditional distribution is **leptokurtic**.