Volatility Modeling

MIT 18.642

Dr. Kempthorne

Fall 2024

Defining Volatility

Basic Definition

- Annualized standard deviation of the change in price or value of a financial security.

Estimation/Prediction Approaches

- Historical/sample volatility measures.
- Geometric Brownian Motion Model
- Poisson Jump Diffusion Model
- ARCH/GARCH Models
- Stochastic Volatility (SV) Models
- Implied volatility from options/derivatives

Historical Volatility

Computing volatility from historical series of actual prices

- Prices of an asset at $(T+1)$ time points
  $$\{P_t, t = 0, 1, 2, \ldots, T\}$$
- Returns of the asset for T time periods
  $$R_t = log(P_t/P_{t-1}), t = 1, 2, \ldots, T$$
- $\{R_t\}$ assumed covariance stationary with
  $$\sigma = \sqrt{var(R_t)} = \sqrt{E[(R_t - E[R_t])^2]}$$
  with sample estimate:
  $$\hat{\sigma} = \sqrt{\tfrac{1}{T-1}\sum_{t=1}^{T}(R_t - \bar{R})^2}, \text{ with } \bar{R} = \tfrac{1}{T}\sum_{1}^{T} R_t.$$
- Annualized values
  $$\widehat{vol} = \begin{cases} \sqrt{252}\hat{\sigma} & \text{(daily prices for 252 business days/year)} \\ \sqrt{52}\hat{\sigma} & \text{(weekly prices)} \\ \sqrt{12}\hat{\sigma} & \text{(monthly prices)} \end{cases}$$

Prediction Methods Based on Historical Volatility

Definition For time period t, define the **sample volatility**

$$\hat{\sigma}_t = \text{sample standard deviation of period } t \text{ returns}$$

- If t indexes months with daily data, then $\hat{\sigma}_t$ is the sample standard deviation of daily returns in month t.
- If t indexes days with daily data, then $\hat{\sigma}_t^2 = R_t^2$.
- With high-frequency data, daily σ_t is derived from cumulating squared intra-day returns.

Historical Average: $\tilde{\sigma}_{t+1}^2 = \frac{1}{t}\sum_1^t \hat{\sigma}_j^2$

(uses all available data)

Simple Moving Average: $\tilde{\sigma}_{t+1}^2 = \frac{1}{m}\sum_0^{m-1} \hat{\sigma}_{t-j}^2$

(uses last m single-period sample estimates)

Exponential Moving Average: $\tilde{\sigma}_{t+1}^2 = (1-\beta)\hat{\sigma}_t^2 + \beta\tilde{\sigma}_t^2$ $0 \leq \beta \leq 1$

(uses all available data)

Exponential Weighted Moving Average:

$\tilde{\sigma}_{t+1}^2 = \sum_{j=0}^{m-1}(\beta^j \hat{\sigma}_{t-j}^2)/[\sum_{j=0}^{m-1}\beta^j]$ (uses last m)

single-period sample estimates).

Predictions Based on Historical Volatility

Simple Regression:

$$\tilde{\sigma}^2_{t+1} = \gamma_{1,t}\hat{\sigma}^2_t + \gamma_{1,t}\hat{\sigma}^2_{t-1} + \cdots + \gamma_{p,t}\hat{\sigma}^2_{t-p+1} + u_t$$

Regression can be fit using all data or last m (rolling-windows).

Note: similar but different from auto-regression model of $\hat{\sigma}^2_t$

Trade-Offs

- Use more data to increase precision of estimators
- Use data closer to time t for estimation of σ_t.

Evaluate out-of-sample performance

- Distinguish assets and asset-classes
- Consider different sampling frequencies and forecast horizons
- Apply performance measures (MSE, MAE, MAPE, etc.)

Benchmark Methodology: RiskMetrics, see Technical Document

https://www.msci.com/documents/10199/
5915b101-4206-4ba0-aee2-3449d5c7e95a

Geometric Brownian Motion (GBM)

For $\{S(t)\}$ the price of a security/portfolio at time t:

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t),$$

where

- σ is the volatility of the security's price
- μ is mean return (per unit time).
- $dS(t)$ infinitesimal increment in price
- $dW(t)$ infinitesimal increment of a standard Brownian Motion/Wiener Process
    - Increments $[W(t') - W(t)]$ are Gaussian with mean zero and variance $(t' - t)$.
    - Increments on disjoint time intervals are independent.
      For $t_1 < t_2 < t_3 < t_4$,
      $[W(t_2) - W(t_1)]$ and $[W(t_4) - W(t_3)]$ are independent

Geometric Brownian Motion (GBM)

Sample Data from Process:

- Prices: $\{S(t), t = t_0, t_1, \ldots, t_n\}$
- Returns: $\{R_j = log[S(t_j)/S(t_{j-1})], j = 1, 2, \ldots, n\}$
  indep. r.v.'s: $R_j \sim N(\mu_* \Delta_j, \sigma^2 \Delta_j)$, where
  $\Delta_j = (t_j - t_{j-1})$ and $\mu_* = [\mu - \sigma^2/2]$

($\{log[S(t)]\}$ is Brownian Motion with drift μ^* and volatility σ^2.)

Maximum-Likelihood Parameter Estimation

- If $\Delta_j \equiv 1$, then

$$\begin{aligned} \hat{\mu}_* &= \bar{R} = \frac{1}{n}\sum_1^n R_t \\ \hat{\sigma}^2 &= \frac{1}{n}\sum_1^n (R_t - \bar{R})^2 \end{aligned}$$

- If Δ_j varies ... Exercise.

Garman-Klass Estimator:

- Sample information more than period-close prices, also have period-high, period-low, and period-open prices.
- Assume $\mu = 0$, $\Delta_j \equiv 1$ (e.g., daily) and let $f \in (0,1)$ denote the fraction of the day prior to the market open.

$$
\begin{aligned}
C_j &= log[S(t_j)] \\
O_j &= log[S(t_{j-1} + f)] \\
\\
H_j &= \max_{t_{j-1}+f \leq t \leq t_j} log[S(t)] \\
\\
L_j &= \min_{t_{j-1}+f \leq t \leq t_j} log[S(t)]
\end{aligned}
$$

Garman-Klass Estimator

Using data from the first period:

- $\hat{\sigma}_0^2 = (C_1 - C_0)^2$: Close-to-Close squared return
  $E[\hat{\sigma}_0^2] = \sigma^2$, and $var[\hat{\sigma}_0^2] = 2(\sigma^2)^2 = 2\sigma^4$.

Garman-Klass Estimator

Using data from the first period:

- $\hat{\sigma}_0^2 = (C_1 - C_0)^2$: Close-to-Close squared return
  $E[\hat{\sigma}_0^2] = \sigma^2$, and $var[\hat{\sigma}_0^2] = 2(\sigma^2)^2 = 2\sigma^4$.
- $\hat{\sigma}_1^2 = \frac{(O_1 - C_0)^2}{f}$: Close-to-Open squared return
  $E[\hat{\sigma}_1^2] = \sigma^2$, and $var[\hat{\sigma}_1^2] = 2(\sigma^2)^2 = 2\sigma^4$.

Garman-Klass Estimator

Using data from the first period:

- $\hat{\sigma}_0^2 = (C_1 - C_0)^2$: Close-to-Close squared return
  $E[\hat{\sigma}_0^2] = \sigma^2$, and $var[\hat{\sigma}_0^2] = 2(\sigma^2)^2 = 2\sigma^4$.
- $\hat{\sigma}_1^2 = \frac{(O_1 - C_0)^2}{f}$: Close-to-Open squared return
  $E[\hat{\sigma}_1^2] = \sigma^2$, and $var[\hat{\sigma}_1^2] = 2(\sigma^2)^2 = 2\sigma^4$.
- $\hat{\sigma}_2^2 = \frac{(C_1 - O_1)^2}{1-f}$: Open-to-Close squared return
  $E[\hat{\sigma}_2^2] = \sigma^2$, and $var[\hat{\sigma}_2^2] = 2(\sigma^2)^2 = 2\sigma^4$.

Garman-Klass Estimator

Using data from the first period:

- $\hat{\sigma}_0^2 = (C_1 - C_0)^2$: Close-to-Close squared return
  $$E[\hat{\sigma}_0^2] = \sigma^2, \text{ and } var[\hat{\sigma}_0^2] = 2(\sigma^2)^2 = 2\sigma^4.$$
- $\hat{\sigma}_1^2 = \frac{(O_1 - C_0)^2}{f}$: Close-to-Open squared return
  $$E[\hat{\sigma}_1^2] = \sigma^2, \text{ and } var[\hat{\sigma}_1^2] = 2(\sigma^2)^2 = 2\sigma^4.$$
- $\hat{\sigma}_2^2 = \frac{(C_1 - O_1)^2}{1-f}$: Open-to-Close squared return
  $$E[\hat{\sigma}_2^2] = \sigma^2, \text{ and } var[\hat{\sigma}_2^2] = 2(\sigma^2)^2 = 2\sigma^4.$$

  Note: $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ are independent!

- $\hat{\sigma}_*^2 = \frac{1}{2}\hat{\sigma}_1^2 + \frac{1}{2}\hat{\sigma}_2^2$
  $$E[\hat{\sigma}_*^2] = \sigma^2, \text{ and } var[\hat{\sigma}_*^2] = \sigma^4.$$

Garman-Klass Estimator

Using data from the first period:

- $\hat{\sigma}_0^2 = (C_1 - C_0)^2$: Close-to-Close squared return
  $$E[\hat{\sigma}_0^2] = \sigma^2, \text{ and } var[\hat{\sigma}_0^2] = 2(\sigma^2)^2 = 2\sigma^4.$$
- $\hat{\sigma}_1^2 = \frac{(O_1 - C_0)^2}{f}$: Close-to-Open squared return
  $$E[\hat{\sigma}_1^2] = \sigma^2, \text{ and } var[\hat{\sigma}_1^2] = 2(\sigma^2)^2 = 2\sigma^4.$$
- $\hat{\sigma}_2^2 = \frac{(C_1 - O_1)^2}{1-f}$: Open-to-Close squared return
  $$E[\hat{\sigma}_2^2] = \sigma^2, \text{ and } var[\hat{\sigma}_2^2] = 2(\sigma^2)^2 = 2\sigma^4.$$

  Note: $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ are independent!

- $\hat{\sigma}_*^2 = \frac{1}{2}\hat{\sigma}_1^2 + \frac{1}{2}\hat{\sigma}_2^2$
  $$E[\hat{\sigma}_*^2] = \sigma^2, \text{ and } var[\hat{\sigma}_*^2] = \sigma^4.$$

$\Longrightarrow eff(\hat{\sigma}_*^2) = \frac{var(\hat{\sigma}_0^2)}{var(\hat{\sigma}_*^2)} = 2.$

Parkinson (1976): With $f = 0$, defines

$$\hat{\sigma}_3^2 = \frac{(H_1 - L_1)^2}{4(log 2)}$$ and shows $eff(\hat{\sigma}_3^2) \approx 5.2$.

Garman and Klass (1980) show that for any $0 < f < 1$:

- $\hat{\sigma}_4^2 = a \times \hat{\sigma}_1^2 + (1 - a)\hat{\sigma}_3^2$
  has minimum variance when $a \approx 0.17$, independent of f and $Eff(\hat{\sigma}_4^2) \approx 6.2$.
- "Best Analytic Scale-Invariant Estimator"'
  $\hat{\sigma}_{**}^2 = 0.511(u_1 - d_1)^2 - 0.019\{c_1(u_1 + d_1) - 2u_1 d_1\} - 0.383c_1^2$,
  where the normalized high/low/close are:
  $$\begin{aligned} u_j &= H_j - O_j \\ d_j &= L_j - O_j \\ c_j &= C_j - O_j \end{aligned}$$
  and $Eff(\hat{\sigma}_{**}^2) \approx 7.4$

- If $0 < f < 1$ then the opening price O_1 may differ from C_0 and the composite estimator is

$$\hat{\sigma}^2_{GK} = a\frac{(O_1 - C_0)^2}{f} + (1-a)\frac{\sigma^2_{**}}{(1-f)}$$

which has minimum variance when $a = 0.12$ and

$$Eff(\hat{\sigma}^2_{GK}) \approx 8.4.$$

References/ Extensions

- **Garman Klass (1980):** `https://www.cmegroup.com/trading/fx/files/a_estimation_of_security_price.pdf`
- **Garman, M. B. and Klass, M. J. (1980)** On the estimation of security price volatilities from historical data. Journal of business, pages 67-78
- **Parkinson, M. (1980):** The extreme value method for estimating the variance of the rate of return. Journal of business, pages 61-65.
- **Rogers, L. C. G. and Satchell, S. E. (1991):** Estimating variance from high, low and closing prices. The Annals of Applied Probability, pages 504-512.
- **Rogers, L. C., Satchell, S. E., and Yoon, Y. (1994)**. Estimating the volatility of stock prices: a comparison of methods that use high and low prices. Applied Financial Economics, 4(3):241-247.
- **Yang, D. and Zhang, Q. (2000):** Drift-independent volatility estimation based on high, low, open, and close prices. The Journal of Business, 73(3):477-492.

Estimating Historical Volatility of the S&P 500 Index

(Case Study using R)

Poisson Jump Diffusions

For $\{S(t)\}$ the stochastic process for the price of the security/portfolio at time t,

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t) + \gamma\sigma Z(t) d\Pi(t),$$

where

- $dS(t) =$ infinitesimal increment in price.
- $\mu =$ mean return (per unit time)
- $\sigma =$ diffusion volatility of the security's price process
- $dW(t)=$ increment of standard Wiener Process
- $d\Pi(t) =$ increment of a Poisson Process with rate λ, modeling the jump process.
- $(\gamma\sigma) \times Z(t)$, the magnitude a return jump/shock
  $Z(t)$ i.i.d $N(0,1)$ r.v.'s and
  $\gamma=$ scale(σ units) of jump magnitudes.

Poisson Jump Diffusions

Maximum-Likelihood Estimation of the PJD Model

- Model is a Poisson mixture of Gaussian Distributions.
- Moment-generating function derived as that of random sum of independent random variables.
- Likelihood function product of infinite sums
- EM Algorithm* expressible in closed form
    - Jumps treated as latent variables which simplify computations
    - Algorithm provides a posteriori estimates of number of jumps per time period.

* See Pickard, Kempthorne, Zakaria (1987).

Laplace Distribution: Brownian Motion With Exponential Time Increments

Laplace Distribution: $X \sim Laplace(\mu, \sigma)$

- Probability density function:

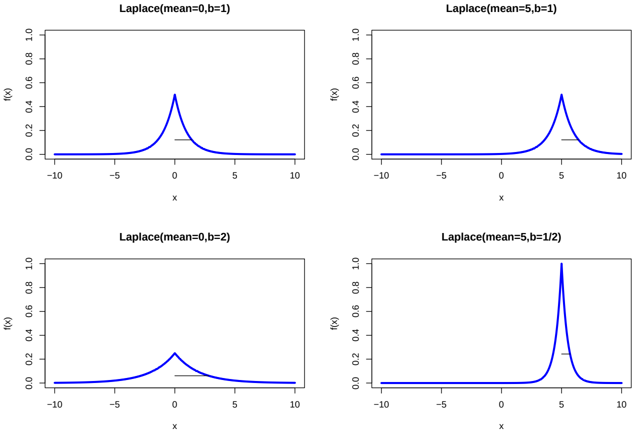
$$f(x \mid \mu, b) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}, -\infty < x < \infty.$$

- Two parameters:

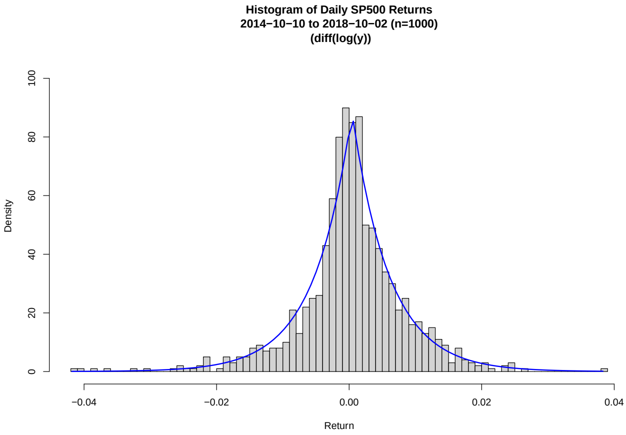
$$\begin{array}{lllllll} mean & = & E[X] & = & \mu & = & \int_{-\infty}^{+\infty} x f(x \mid \mu, b) dx \\ variance & = & Var[X] & = & 2 \times b^2 & = & \int_{-\infty}^{+\infty} (x-\mu)^2 f(x \mid \mu, b) dx \end{array}$$

- History/motivation
  - Laplace (1774) "First Law of Errors"
  - Brownian motion observed at exponential times
  - Geometric sum of normal random variables

Four Laplace Models

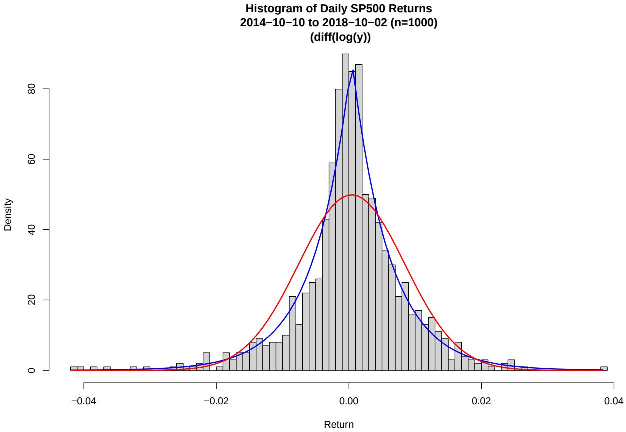


Daily Returns of SP500 Index



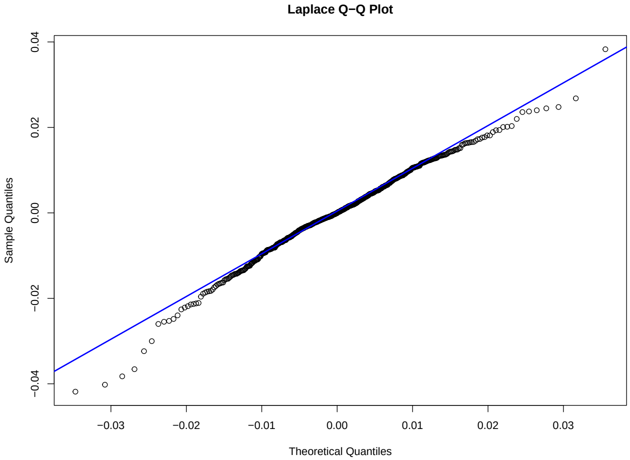
Laplace Model: $\hat{\mu} = Mean = 0.0004273$ and $\hat{b} = \sqrt{Variance/2} = 0.005652$
(Method-of-Moments estimates)
Annualized Volatility: $\hat{\sigma} = 0.1268901$

Daily Returns of SP500 Index



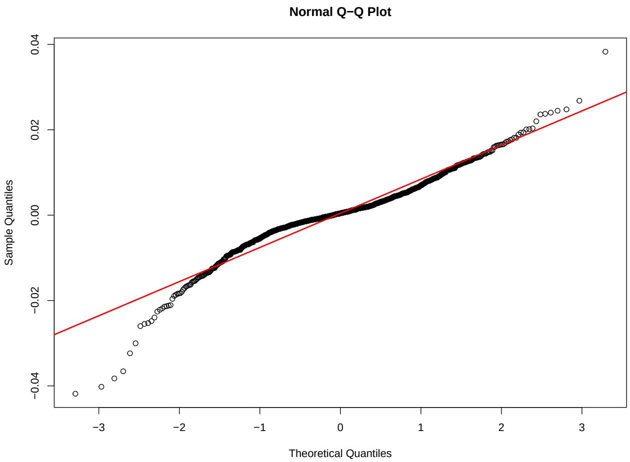
Laplace Model (Blue) versus Normal Model (Red)

Evaluating Goodness of Fit



QQ-plot displays “Goodness-of-Fit”
Sorted Observations versus Theoretical Expected Values

Evaluating Goodness of Fit



QQ-plot displays "Goodness-of-Fit"
Sorted Observations versus Theoretical Expected Values for N(0,1)

ARCH Models

ARCH models are specified relative to the discrete-time process for the price of the security/portfolio: $\{S_t, t = 1, 2, \ldots\}$

Engle (1982) models the discrete returns of the process $y_t = log(S_t/S_{t-1})$ as

$$y_t = \mu_t + \epsilon_t,$$

where μ_t is the mean return, conditional on $\mathcal{F}_{t-1}$, the information available through time $(t-1)$, and

$$\epsilon_t = Z_t \times \sigma_t,$$

where Z_t i.i.d. with $E[Z_t] = 0$, and $var[Z_t] = 1$,

$$\sigma_t^2 = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \alpha_2\epsilon_{t-2}^2 + \cdots + \alpha_p\epsilon_{t-p}^2$$

Parameter Constraints: $\alpha_j \geq 0, j = 0, 1, \ldots, p$

$\sigma_t^2 = var(R_t \mid \mathcal{F}_{t-1})$, "Conditional Heteroscedasticity" of returns .

ARCH Models

The ARCH model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \cdots + \alpha_p \epsilon_{t-p}^2$$

implies an AR model in ϵ_t^2. Add $(\epsilon_t^2 - \sigma_t^2) = u_t$ to both sides:

$$\epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \cdots + \alpha_p \epsilon_{t-p}^2 + u_t$$

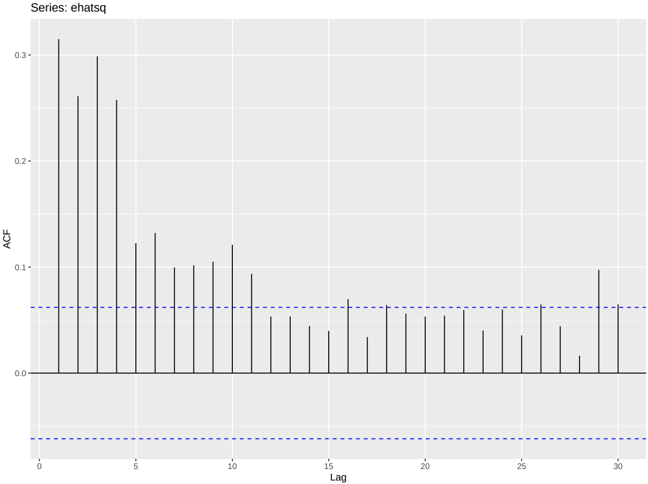
where $u_t : E[u_t \mid \mathcal{F}_t] = 0$, and $var[u_t \mid \mathcal{F}_t] = var(\epsilon_t^2) = 2\sigma_t^4$.

Lagrange Multiplier Test

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$$

- Fit linear regression on squared residuals $\hat{\epsilon}_t = y_t - \hat{\mu}_t$. (i.e., Fit an AR(p) model to $[\hat{\epsilon}_t^2]$, $t = 1, 2, \ldots, n$)
- LM test statistic $= nR^2$, where R^2 is the R-squared of the fitted AR(p) model.
  Under H_0 the r.v. nR^2 is approx. χ^2 $(df = p)$
- Note: the linear regression estimates of parameters are not MLEs under Gaussian assumptions; they correspond to quasi-maximum likelihood estimates (QMLE).

Autocorrelations of $\hat{\epsilon}_t^2$ for SP500 Log Returns



$\hat{\epsilon}_t = diff(log(SP500)) - mean$
Autocorrelations of $\hat{\epsilon}_t^2$

Maximum Likelihood Estimation

ARCH Model:

$$
\begin{aligned}
y_t &= c + \epsilon_t \\
\epsilon_t &= z_t \sigma_t \\
\sigma_t^2 &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \cdots + \alpha_p \epsilon_{t-p}^2 \\
& \quad t = 0, 1, \ldots, T
\end{aligned}
$$

Likelihood:

$$
\begin{aligned}
L(c, \boldsymbol{\alpha}) &= p(y_1, \ldots, y_n \mid c, \alpha_0, \alpha_1, \ldots, \alpha_p) \\
&= \prod_{t=1}^{n} p(y_t \mid \mathcal{F}_{t-1}, c, \boldsymbol{\alpha}) \\
&= \prod_{t=1}^{n} [\frac{1}{\sqrt{2\pi\sigma_t^2}} exp(\frac{-1}{2} \frac{\epsilon_t^2}{\sigma_t^2})] \\
& \quad \text{where } \epsilon_t = y_t - c \text{ and} \\
& \quad \quad \sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \cdots \alpha_p \epsilon_{t-p}^2.
\end{aligned}
$$

Constraints:

- $\alpha_i \geq 0, \; i = 1, 2, \ldots, p$
- $(\alpha_1 + \cdots + \alpha_p) < 1.$

GARCH Models

Bollerslev (1986) extended ARCH models to:

GARCH(p,q) Model

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

Constraints: $\alpha_i \geq 0, \forall i$, and $\beta_j \geq 0, \forall j$

GARCH(1,1) Model

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

- Parsimonious
- Fits many financial time seriies

GARCH Models

The GARCH(1,1) model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

implies an ARMA model in ϵ_t^2. Eliminate $\sigma_{t'}^2$ using $(\epsilon_{t'}^2 - \sigma_{t'}^2) = u_{t'}$

$$\begin{aligned} \epsilon_t^2 - u_t &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1(\epsilon_{t-1}^2 - u_{t-1}) \\ \epsilon_t^2 &= \alpha_0 + (\alpha_1 + \beta_1)\epsilon_{t-1}^2 + u_t - \beta_1 u_{t-1} \end{aligned}$$

where $u_t : E[u_t \mid \mathcal{F}_t] = 0$, and $var[u_t \mid \mathcal{F}_t] = var(\epsilon_t^2) = 2\sigma_t^4$.

$\Longrightarrow GARCH(1,1)$ implies an $ARMA(1,1)$ with

$$u_t = (\epsilon_t^2 - \sigma_t^2) \sim WN(0, 2\sigma^4)$$

Stationarity of GARCH model deduced from ARMA model

$$\begin{aligned} A(L)\epsilon_t^2 &= B(L)u_t \\ \epsilon_t^2 &= [A(L)]^{-1}B(L)u_t. \end{aligned}$$

Covariance stationary:roots of $A(z)$ outside $\{|z| \leq 1\}$,i.e.,

$$|\alpha_1 + \beta_1| < 1$$

Unconditional Volatility / Long-Run Variance

GARCH(1,1): Assuming stationarity, $0 < (\alpha_1 + \beta_1) < 1$

$$\begin{aligned} \sigma_*^2 &= \alpha_0 + (\alpha_1 + \beta_1)\sigma_*^2 \\ \implies \sigma_*^2 &= \frac{\alpha_0}{(1-\alpha_1-\beta_1)} \end{aligned}$$

GARCH(p,q) implies ARMA(max(p,q), q) model

- Stationary if $0 < (\sum_1^p \alpha_i + \sum_1^q \beta_j) < 1$
- Long-Run Variance:

$$\begin{aligned} \sigma_*^2 &= \alpha_0 + (\sum_1^p \alpha_j + \sum_1^q \beta_1)\sigma_*^2 \\ \implies \sigma_*^2 &= \frac{\alpha_0}{[1-\sum_1^{max(p,q)}(\alpha_i+\beta_j)]} \end{aligned}$$

GARCH Model Estimation

Maximum Likelihood Estimation

GARCH Model:

$$\begin{array}{rcl} y_t & = & c + \epsilon_t \\ \epsilon_t & = & z_t \sigma_t \\ \sigma_t^2 & = & \alpha_0 + \sum_1^p \alpha_i \epsilon_{t-i}^2 + \sum_1^q \beta_j \sigma_{t-j}^2 \end{array}$$

$$t = 0, 1, \ldots, T$$

Likelihood:

$$\begin{array}{rcl} L(c, \boldsymbol{\alpha}, \boldsymbol{\beta}) & = & p(y_1, \ldots, y_T \mid c, \alpha_0, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q) \\ & = & \prod_{t=1}^T p(y_t \mid \mathcal{F}_{t-1}, c, \boldsymbol{\alpha}, \boldsymbol{\beta}) \\ & = & \prod_{t=1}^T [\frac{1}{\sqrt{2\pi\sigma_t^2}} exp(\frac{-1}{2} \frac{\epsilon_t^2}{\sigma_t^2})] \end{array}$$

where $\epsilon_t = y_t - c$ and

$$\sigma_t^2 = \alpha_0 + \sum_1^p \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2.$$

Constraints: $\alpha_i \geq 0, \forall i$, $\beta_j \geq 0, \forall j$, and $0 < (\sum_1^p \alpha_u + \sum_1^1 \beta_j) < 1$.

GARCH Model
Estimation/Evaluation/Model-Selection

- Maximum-Likelihood Estimates: $\hat{c}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}$

  $\Longrightarrow \hat{\epsilon}_t$ and $\hat{\sigma}_t^2$ $(t = T, T-1, \ldots)$
- Standardized Residuals

  $\hat{\epsilon}_t / \hat{\sigma}_t$: should be uncorrelated
- Squared Standardized Residuals

  $(\hat{\epsilon}_t / \hat{\sigma}_t)^2$: should be uncorrelated

Testing Normality of Residuals

- Normal QQ Plots
- Jarque-Bera test
- Shapiro-Wilk test
- MLE Percentiles Goodness-of-Fit Test
- Kolmogorov-Smirnov Goodness-of-Fit Test

Model Selection: Apply model-selection critera

- Akaike Information Criterion (AIC)
- Bayes Information Criterion (BIC)

Stylized Features of Returns/Volatility

Volatility Clustering

- Large ϵ_t^2 follow large ϵ_{t-1}^2
- Small ϵ_t^2 follow small ϵ_{t-1}^2

GARCH models can prescribe

- Large σ_t^2 follow large σ_{t-1}^2
- Small σ_t^2 follow small σ_{t-1}^2

Heavy Tails / Fat Tails

- Returns distribution has heavier tails (higher Kurtosis) than Gaussian
- GARCH(p,q) models are stochastic mixture of Gaussian distributions with higher kurtosis.

Engle, Bollerslev, and Nelson (1994)

Stylized Features of Returns/Volatility

Volatility Mean Reversion

GARCH(1,1) Model

- Long-run average volatility: $\sigma_*^2 = \frac{\alpha_0}{1-\alpha_1-\beta_1}$
- Mean-Reversion to Long-Run Average

$$\epsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1)\epsilon_{t-1}^2 + u_t - \beta_1 u_{t-1}$$

Substituting: $\alpha_0 = (1 - \alpha_1 - \beta_1)\sigma_*^2$

$$(\epsilon_t^2 - \sigma_*^2) = (\alpha_1 + \beta_1)(\epsilon_{t-1}^2 - \sigma_*^2) + u_t - \beta_1 u_{t-1}$$

$0 < (\alpha_1 + \beta_1) < 1 \Longrightarrow$ Mean Reversion!

Extended Ornstein-Uhlenbeck(OU) Process for ϵ_t^2 with MA(1) errors.

Extensions of GARCH Models

EGARCH Nelson (1992)

TGARCH Glosten, Jagannathan, Runkler (1993)

PGARCH Ding, Engle, Granger

GARCH-In-Mean

Non-Gaussian Distributions

- $t-$Distributions
- Generalized Error Distributions (GED)