Time Series methods[1]

☐ Maximum Likelihood Estimation (MLE)

The core objective of MLE is to find the parameter value (θ) that maximize the likelihood function L(θ), which represents the probability density of the observed data ($y_1, \cdots, y_T$) given those parameters. If data points were independent, the likelihood would be a simple product of individual densities. **However, in time series models like GARCH, observations are dependent on past values,** so the likelihood is calculated using conditional densities. For practical maximization, the log-likelihood is used, transforming the product of densities into a sum.

The Formula.

$$\text{Return Equation: } R_t = a + bR_{t-1} + \varepsilon_t$$

$$\text{Error Term: } \varepsilon_t = \sigma_t z_t$$

$$\text{Variance Equation: } \sigma_t^2 = \omega + +\beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2,$$

$$\omega > 0, \alpha \ and \ \beta \geq 0$$

For Estimation of GARCH (1, 1), While simultaneous estimation of all parameters is ideal, a simplified two-estimation method is introduced. **Step 1,** Estimate the return equation ($R_t = \alpha + \beta R_{t-1} + \varepsilon_t$) using **Ordinary Least Squares (OLS)** and calculate the residuals ($\hat{\varepsilon}_t$). **Step 2,** Treat these residuals as the error term and estimate the GARCH (1, 1) parameters (ω, α, β) using **MLE.** The likelihood is expressed as $L = f(\varepsilon_1, \cdots, \varepsilon_T | \ \omega, \alpha, \beta)$.

To compute the likelihood for a GARCH model, the following assumptions and processes are used. **Initial Values:** Starting values of variance (σ_0^2) and errors (ε_0^2) are typically set to the sample variance of the residuals or the unconditional expectation $\frac{\omega}{(1-\alpha-\beta)}$. **Distribution:** The error term z_t is usually assumed to follow a standard normal distribution. **Recursive Calculation:** Given initial values ($\sigma_0^2, \varepsilon_0^2, \omega, \alpha, \beta$), the variance for each step (σ_t^2) is calculated recursively, allowing for the calculation of the probability density at each time t. The density for ε_1 is $f(\varepsilon_1 | \sigma_0^2, \varepsilon_0^2, \omega, \alpha, \beta) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp(-\frac{\varepsilon_1^2}{2\sigma_1^2})$.

For practical maximization, the log-likelihood is used $\log L = \sum_{t=1}^{T} \log f(\varepsilon_t | \ldots)$. Log density formula is $\log f(\varepsilon_t | \ldots) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_t^2) - \frac{\varepsilon_t^2}{2\sigma_t^2}$. As $T \to \infty$, the distribution of the MLE estimator $\hat{\theta}$ is $\sqrt{T}(\hat{\theta}_T^{ML} - \theta) \sim N(0, A^{-1})$, where $A = -\frac{1}{T} E\left(\frac{\partial^2 logL(\theta)}{\partial\theta\partial\theta'}\right) \approx -\frac{1}{T}\frac{\partial^2 logL(\theta)}{\partial\theta\partial\theta'}|_{\theta=\hat{\theta}_T^{ML}}$. The variance of $\hat{\theta}_T^{ML}$ is approximated by the inverse of the Hessian matrix of the log-likelihood.

[1] *The mathematical derivations presented in this section are based on the lecture notes of Professor Jouchi Nakajima (Hitotsubashi University) and Professor In-su Kim (Jeonbuk National University).*

Information Criteria is used to compare models such as AIC or SBIC. AIC is expressed as $-2\log L + 2k$, and SBIC is $-2\log L + klog(T)$, where k is the number of parameters. If P_t is the asset price at time t, return y_t are calculated as: Percentage change $y_t = \frac{P_t - P_{t-1}}{P_{t-1}} \times 100$. Log return: $y_t = (\log P_t - \log P_{t-1}) \times 100$.