Unit root test and GARCH (5.2 - 5.3)

Consider a AR(p) model $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + W_t$, we want to test if $\Phi(1) = 0$, where $\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p$, i.e. we want to test if $\sum_{j=1}^{p} \phi_j = 1$

Other approaches:

Fit AR(p) model to get $\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p$, then

(1) To see if the roots of $\hat{\Phi}(z) = 1 - \sum_{j=1}^{p} \hat{\phi}_j z^j$ are close to 1

(2) To see if $\hat{\rho}(h)$ decays to 0 fast as h increases

Both methods are not a proper hypothesis test.

Rewrite X_t as $X_t = \beta_1 X_{t-1} - \sum_{j=1}^{p-1} \beta_{j+1} \Delta X_{t-j} + W_t$

or $\Delta X_t = \gamma X_{t-1} - \sum_{j=1}^{p-1} \beta_{j+1} \Delta X_{t-j} + W_t$

and we test $H_0: \gamma = 0$

Under the assumption that X_t follows AR(p) and ΔX_t is stationary we can estimate the distribution of $\hat{\gamma}$, which is the estimate of γ by regressing ΔX_t on $X_{t-1}, \Delta X_{t-1}, \Delta X_{t-2}, \dots, \Delta X_{t-p+1}$

If W_t's are uncorrelated (i.e. usual AR(p)), we can use ADF test. The R function is adf.test

If W_t's are correlated (e.g. the error term in ARMA(p,q)), we can use PP test. The R function is pp.test

Note that the "alternative hypothesis: stationary" is only valid when the assumption is true.

Typically, for financial series, the return r_t, does not have a constant conditional variance, and highly volatile periods tend to be clustered together (i.e. r_t^2 depends on its past values)

$\Rightarrow$ W_t's iid assumption does not hold for r_t

Suppose r_t follows ARMA(p,q) invertible, then $r_t = \sum_{j=1}^{\infty} \pi_j r_{t-j} + W_t$

The conditional variance is $Var(r_t | r_{t-1}, r_{t-2}, \dots) = Var(W_t | r_{t-1}, \dots) = \sigma_w^2$

if $W_t \sim iid(0, \sigma_w^2)$

While we need a new model for non-constant conditional variance, we would like to keep the following properties for W_t

1. W_t's are white noises, i.e. $EW_t = 0$, $Var(W_t) = E(W_t^2) = \text{constant}$

   $Cov(W_t, W_s) = 0$ for $t \neq s$

   so that the considered model is still ARMA(p,q)

2. $E(W_t \mid r_n, r_{n-1}, \dots) = \begin{cases} 0 & \text{for } t > n \\ W_t & \text{for } t \leq n \end{cases}$

   so that the formula for $\tilde{r}_{n+m}$ is still valid

3. W_{n+t} and W_{n+s} are uncorrelated given $r_n, r_{n-1}, \dots$ so that,

   from $r_{n+m} - \tilde{r}_{n+m} = \sum_{j=0}^{m-1} \psi_j W_{n+m-j} = \psi_0 W_{n+m} + \psi_1 W_{n+m-1} + \dots + \psi_{m-1} W_{n+1}$,

   $Var(r_{n+m} - \tilde{r}_{n+m} \mid r_n, r_{n-1}, \dots) = E((r_{n+m} - \tilde{r}_{n+m})^2 \mid r_n, r_{n-1}, \dots) = \sum_{j=0}^{m-1} \psi_j^2 E(W_{n+m-j}^2 \mid r_n, r_{n-1}, \dots)$

Therefore, it is natural for us to consider $W_t = \sigma_t \varepsilon_t$, where $\varepsilon_t \sim iid(0,1)$, $E\sigma_t^2 = \sigma^2$ independent of t and σ_t and ε_t are independent. Then, we have

1. $E(W_t) = E(\sigma_t) E(\varepsilon_t) = 0$, $E(W_t^2) = E(\sigma_t^2) E(\varepsilon_t^2) = \sigma^2$

   $Cov(W_t, W_s) = E(\sigma_t \sigma_s \varepsilon_s) E(\varepsilon_t) = 0$ for $t > s$

2. $E(W_t \mid r_n, r_{n-1}, \dots) = E(\varepsilon_t) E(\sigma_t \mid r_n, r_{n-1}, \dots) = 0$ for $t > n$

   $E(W_t \mid r_n, r_{n-1}, \dots) = W_t$ comes from the assumption that r_t is invertible

3. $Cov(W_{n+t}, W_{n+s} \mid r_n, r_{n-1}, \dots) = E(\sigma_{n+t}\varepsilon_{n+t}\,\sigma_{n+s}\varepsilon_{n+s} \mid r_n, r_{n-1}, \dots)$

(assume $t>s$) $= E(\varepsilon_{n+t})\,E(\varepsilon_{n+s}\sigma_{n+s}\sigma_{n+t} \mid r_n, r_{n-1}, \dots) = 0$

Note that $E(W_t^2) = E(\sigma_t^2)$. Therefore, if W_t^2 is stationary, we have $E\sigma_t^2 = \sigma^2$. For simplicity, we start from AR(0) model, i.e. $r_t = W_t = \sigma_t\varepsilon_t$. Then checking if $W_t^2 (= r_t^2)$ is stationary and fitting models for W_t^2 are strict forward as we have observations $r_1, \dots, r_n$ (but we don't have the corresponding $\sigma_1^2, \dots, \sigma_n^2$)

Autoregressive Conditionally Heteroscedastic (ARCH)

ARCH(1) $\quad r_t = \sigma_t\varepsilon_t \qquad \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 \qquad (5.37)$

If $|\alpha_1|<1$ and $Var(r_t) = E\sigma_t^2 < \infty \ \forall t$, then

$$\begin{aligned} E\sigma_t^2 &= \alpha_0 + \alpha_1 E\sigma_{t-1}^2 = \alpha_0 + \alpha_1(\alpha_0 + \alpha_1 E\sigma_{t-2}^2) \\ &= \alpha_0(1+\alpha_1+\alpha_1^2+\dots+\alpha_1^{k-1}) + \alpha_1^k E\sigma_{t-k}^2 \\ &= \alpha_0\left(\frac{1}{1-\alpha_1}\right) \qquad (\text{take } k \to \infty) \end{aligned}$$

Or, we can consider $E\sigma_t^2 = \alpha_0 + \alpha_1 E\sigma_{t-1}^2 + \tilde{w}_t$, with $\tilde{w}_t = 0 \ \forall t$, as a AR(1) model for $E\sigma_t^2$ so that $E\sigma_t^2$ is stationary if $|\alpha_1|<1$ and hence $E\sigma_t^2 = \sigma^2 \ \forall t$ and $\sigma^2 = \alpha_0 + \alpha_1\sigma^2 \Rightarrow \sigma^2 = \frac{\alpha_0}{1-\alpha_1}$

We can rewrite (5.37) as $\quad r_t = \sigma_t\varepsilon_t$

$$r_t^2 - (\alpha_0 + \alpha_1 r_{t-1}^2) = \sigma_t^2(\varepsilon_t^2 - 1) \overset{let}{=} V_t, \text{ where}$$

$EV_t = 0$, $Cov(V_t, V_s) = E(V_tV_s) = E(\varepsilon_t^2 - 1)E(\sigma_t^2 V_s) = 0$ for $t>s$

Note that $\sigma_t^4 = (\alpha_0 + \alpha_1 r_{t-1}^2)^2 = \alpha_0^2 + 2\alpha_0\alpha_1 r_{t-1}^2 + \alpha_1^2 r_{t-1}^4$

$$\Rightarrow E\sigma_t^4 = (\alpha_0^2 + 2\alpha_0\alpha_1\sigma^2) + \alpha_1^2(E\varepsilon_{t-1}^4)E\sigma_{t-1}^4$$

$\therefore$ If $\alpha_1^2(E\varepsilon_{t-1}^4) < 1$ (e.g. $\varepsilon_t \sim N(0,1) \Rightarrow E\varepsilon_t^4 = 3$), then $E\sigma_t^4$ is also "causal stationary" and hence

$Var(V_t) = EV_t^2 = E\sigma_t^4\,E(\varepsilon_t^2 - 1)$ does not depend on t

$\therefore V_t$ is a white noise and r_t^2 follows causal stationary AR(1) model.

The unknown parameters α_0 and α_1 are estimated by conditional MLE with r_1 is fixed. ④

$$f(r_n, r_{n-1}, \ldots, r_2 | r_1) = f(r_n | r_{n-1}, \ldots, r_1) \cdots f(r_2 | r_1)$$
$$= \prod_{t=2}^{n} f(r_t | r_{t-1})$$

Assume $\varepsilon_t \sim N(0,1)$, then $r_t | r_{t-1} \sim N(0, \alpha_0 + \alpha_1 r_{t-1}^2)$

$$\therefore f(r_n, r_{n-1}, \ldots, r_2 | r_1) = \prod_{t=2}^{n} \frac{1}{\sqrt{2\pi(\alpha_0 + \alpha_1 r_{t-1}^2)}} e^{-\frac{r_t^2}{2(\alpha_0 + \alpha_1 r_{t-1}^2)}}$$

$$\Rightarrow \ell(\alpha_0, \alpha_1) = -\log f(r_n, \ldots, r_2 | r_1) = \frac{1}{2} \sum_{t=2}^{n} \log(2\pi) + \frac{1}{2} \sum_{t=2}^{n} \log(\alpha_0 + \alpha_1 r_{t-1}^2)$$
$$+ \frac{1}{2} \sum_{t=2}^{n} \left(\frac{r_t^2}{\alpha_0 + \alpha_1 r_{t-1}^2} \right)$$

And $\hat{\alpha}_0$ and $\hat{\alpha}_1$ is the minimizers of $\ell(\alpha_0, \alpha_1)$

For more general ARCH(p) model,

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \ldots + \alpha_p r_{t-p}^2$$

The conditional likelihood given $r_1, \ldots, r_p$ is

$$f(r_n, r_{n-1}, \ldots, r_{p+1} | r_p, \ldots, r_1) = \prod_{t=p+1}^{n} f(r_t | r_{t-1}, \ldots, r_{t-p})$$
$$= \prod_{t=p+1}^{n} \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-r_t^2/2\sigma_t^2}$$

$(\because r_t | r_{t-1}, \ldots, r_{t-p} \sim N(0, \alpha_0 + \alpha_1 r_{t-1}^2 + \ldots + \alpha_p r_{t-p}^2) = N(0, \sigma_t^2))$

Also note that

$$r_t^2 - (\alpha_0 + \alpha_1 r_{t-1}^2 + \ldots + \alpha_p r_{t-p}^2) = \sigma_t^2(\varepsilon_t^2 - 1) = V_t$$ which is an AR(p) model for suitable choices of $(\alpha_1, \ldots, \alpha_p)$ so that $E\sigma_t^2$ and $E\sigma_t^4$ are time t invariant.

Another extension of ARCH is the generalized ARCH, GARCH model

For GARCH(1,1), $r_t = \sigma_t \varepsilon_t$, $\varepsilon_t \sim iid(0,1)$

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

Consider $r_t^2 - (\alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2) = \sigma_t^2(\varepsilon_t^2 - 1)$

$$\Rightarrow r_t^2 - \alpha_0 - (\alpha_1 + \beta_1) r_{t-1}^2 = -\beta_1 (r_{t-1}^2 - \sigma_{t-1}^2) + \sigma_t^2(\varepsilon_t^2 - 1) = V_t - \beta_1 V_{t-1}$$

which is a causal ARMA(1,1) if $\alpha_1 + \beta_1 < 1$ (note that $\alpha_1, \beta_1 \geq 0$) and V_t is a white noise (constraints on α_1, β_1 so that $E\sigma_t^4$ and $E\sigma_t^2$ are constant)

(5)

For GARCH(p,q), $\sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j r_{t-j}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$

The conditional likelihood function given $r_1, \dots, r_{\max(p,q)}$ and $\sigma_1^2 = \dots = \sigma_q^2$

If $(p \geq q)$, $f(r_n, \dots, r_{p+1} \mid r_p, \dots, r_1) = \prod_{t=p+1}^{n} f(r_t \mid r_{t-1}, \dots, r_1)$

Note that $\sigma_1^2 = \dots = \sigma_q^2 = 0$ and $r_1, \dots, r_p$ are known $\Rightarrow \sigma_{q+1}^2$ is known (with $\alpha_0, \alpha_j, \beta_j$)

$\Rightarrow r_{q+1} \sim N(0, \sigma_{q+1}^2)$

$\sigma_1^2, \dots, \sigma_q^2, \sigma_{q+1}^2$ and $r_1, \dots, r_p, r_{p+1}$ are known $\Rightarrow \sigma_{q+2}^2$ is known and

$r_{q+2} \sim N(0, \sigma_{q+2}^2)$

$\therefore$ Given $r_1, \dots, r_{t-1} \Rightarrow \sigma_1^2, \dots, \sigma_{t-1}^2$ are known $\Rightarrow \sigma_t^2$ is known and $r_t \sim N(0, \sigma_t^2)$

$\therefore f(r_n, \dots, r_{p+1} \mid r_p, \dots, r_1) = \prod_{t=p+1}^{n} \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-r_t^2/2\sigma_t^2}$

For even more general ARMA(p,q) - GARCH(h,k) model

$$r_t = \phi_1 r_{t-1} + \dots + \phi_p r_{t-p} + w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q}$$

$$w_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{j=1}^{h} \alpha_j w_{t-j}^2 + \sum_{j=1}^{k} \beta_j \sigma_{t-j}^2$$

We can apply conditional MLE to estimate the parameters

(1) $r_1, \dots, r_p$ are given and $w_p = \dots = w_{p+1-q} = 0$ (for ARMA(p,q))

(2) $w_1, \dots, w_{\max(h,k)} = 0$ and $\sigma_1^2 = \dots = \sigma_k^2 = 0$ (for GARCH(h,k))

Then $f(r_n, \dots, r_{p+1} \mid r_p, \dots, r_1) = \prod_{t=p+1}^{n} f(r_t \mid r_{t-1}, \dots, r_1)$ (assume $p \geq \max(h,k)$)

$r_{p+1} \mid r_p, \dots, r_1 = \phi_1 r_p + \dots + \phi_p r_1 + \theta_1 w_p + \dots + \theta_q w_{p+1-q} + w_{p+1} \mid r_p, \dots, r_1$

(2) $\Rightarrow \sigma_1^2, \dots, \sigma_{p+1}^2$ are known $\Rightarrow w_{p+1} \mid r_p, \dots, r_1 \sim N(0, \sigma_{p+1}^2)$ (assume r_t causal)

$\therefore r_{p+1} \mid r_p, \dots, r_1 \sim N\left(\sum_{j=1}^{p} \phi_j r_{p+1-j} + \sum_{j=1}^{q} \theta_j w_{p+1-j}, \sigma_{p+1}^2\right)$

Similarly, $r_{p+2} \mid r_{p+1}, r_p, \dots, r_1 \sim N\left(\sum_{j=1}^{p} \phi_j r_{p+2-j} + \sum_{j=1}^{q} \theta_j w_{p+2-j}, \sigma_{p+2}^2\right)$

In general, $r_t \mid r_{t-1}, \dots, r_1 \sim N\left(\sum_{j=1}^{p} \phi_j r_{t-j} + \sum_{j=1}^{q} \theta_j w_{t-j}, \sigma_t^2\right)$

And hence the conditional MLE can be computed.

With the estimated parameters, one-step-ahead forecast of $\sigma_t^2 \mid r_t, \dots, r_1$

is $\hat{\sigma}_{t+1}^2 = \hat{\alpha}_0 + \sum_{j=1}^{h} \hat{\alpha}_j r_{t+1-j}^2 + \sum_{j=1}^{k} \hat{\beta}_j \hat{\sigma}_{t+1-j}^2$