SOLUTION to January 2021 exam
Financial Econometrics A

Question A:

Consider the ARCH model given by

$$x_t = \sigma_t z_t$$
$$\sigma_t^2(\alpha) = \omega + \alpha x_{t-1}^2$$

with z_t $t_v(0,1)$ distributed, x_0 fixed and $t = 1, 2, ..., T$. With $v > 2$ and $\omega > 0$ fixed, the log-likelihood function in terms of $\alpha \geq 0$ is given by

$$\ell_T^{t_v}(\alpha) = -\tfrac{1}{2}\sum_{t=1}^{T}\left(\log\sigma_t^2(\alpha) + (v+1)\log\left(1 + \frac{x_t^2}{\sigma_t^2(\alpha)(v-2)}\right)\right).$$

As usual with α set to the true value $\alpha = \alpha_0$, we set $\sigma_t^2(\alpha_0) = \sigma_t^2$.

Question A.1: We wish to find a value for α, α_a say such that x_t is stationary, weakly mixing and $E|x_t| < \infty$ for $\alpha \in [0, \alpha_a)$ and $v = 4$. To do so apply the drift function

$$\delta(x) = 1 + |x|,$$

and use $E|z_t| = \sqrt{2}/2 \simeq 0.7$ to find α_a. It follows that $\alpha_a > 2$. Discuss this by comparing with the ARCH(1) model where z_t are iidN$(0,1)$.

Hint: Recall the inequality that $|a+b|^\delta < |a|^\delta + |b|^\delta$ for $a, b \in \mathbb{R}$ and $\delta \in (0,1)$.

Solution Question A.1: By definition

$$\begin{aligned} E(\delta(x_t)|x_{t-1} = x) &= 1 + |\sigma_t|E|z_t| \\ &= 1 + \left(\omega + \alpha x^2\right)^{1/2} E|z_t| \\ &\leq 1 + \omega^{1/2} + 0.7\alpha^{1/2}|x| \end{aligned}$$

and hence $\alpha_a^{1/2} 0.7 < 1$ or, $\alpha_a < (10/7)^2 = 2.0408$. Note that if $\sqrt{2/2}$ was used all the way instead of the proposed value 0.7, $\alpha_a = 2$.

For the Gaussian case $\alpha_a = \pi/2 > 1$.

Question A.2: It follows that

$$S_T = \partial \ell_T^{t_v}(\alpha)/\partial\alpha\Big|_{\alpha=\alpha_0, v=4} = -\tfrac{1}{2}\sum_{t=1}^{T}\frac{x_{t-1}^2}{\sigma_t^2}\eta_t,$$

with η_t iid with $E\eta_t = 0$, and

$$\eta_t = \left(1 - 5\frac{z_t^2/2}{1+z_t^2/2}\right).$$

Argue that $E\left(\frac{z_t^2/2}{1+z_t^2/2}\right)^2 \leq 1$, and hence that $\sigma_\eta^2 = E\eta_t^2 < \infty$.

Argue that $\gamma = E\left(x_{t-1}^2/\sigma_t^2\right)^2 \leq 1/\alpha_0^2$ for any $\alpha_0 > 0$.

Solution Question A.2:

It follows that as $\frac{y^2}{1+y^2} \leq 1$, that

$$\begin{aligned} E\eta_t^2 &= 1 + 25E\left(\frac{z_t^2/2}{1+z_t^2/2}\right)^2 - 10E\left(\frac{z_t^2/2}{1+z_t^2/2}\right) \\ &\leq 1 + 25 + 10. \end{aligned}$$

Likewise

$$E\left(x_{t-1}^4/\sigma_t^4\right) = E\left(\frac{x_{t-1}^2}{\omega + \alpha_0 x_{t-1}^2}\right)^2 \leq 1/\alpha_0^2.$$

Question A.3: Use Question A.2 (and A.1) to show that

$$T^{-1}\sum_{t=1}^{T}\left(\frac{x_{t-1}^2}{\sigma_t^2}\right)^2 \xrightarrow{P} \gamma.$$

Next, show that

$$T^{-1/2}S_T = -T^{-1/2}\tfrac{1}{2}\sum_{t=1}^{T}\frac{x_{t-1}^2}{\sigma_t^2}\eta_t \xrightarrow{D} N\left(0, \gamma\sigma_\eta^2/4\right)$$

for any $\alpha_0 \in (0, \alpha_a)$. Similarly one can show that if one relaxes v is known,

$$S_T^v = \partial \ell_T^{t_v}(\alpha, v)/\partial v\Big|_{\alpha=\alpha_0, v=v_0}$$

is asymptotically Gaussian distributed.

Discuss implications of these results and in particular why $\alpha_0 > 0$ is an important assumption.

Solution Question A.3:

First by standard application of LLN for weakly mixing processes (be precise in reference) as by Question A.1 x_t is weakly mixing for $\alpha_0 < \alpha_a$. As $E\left(\frac{x_{t-1}^2}{\sigma_t^2}\eta_t\right) = 0$, again by standard application of the CLT for weakly mixing processes (be precise in reference) the score converges in distribution.

Implications: Together with regularity conditions on the second and third order derivatives (be precise) of the log-likelihood function implies $\sqrt{T}(\hat{\alpha} - \alpha_0)$ and $\sqrt{T}(\hat{v} - v_0)$ are asymptotically Gaussian. Note that if $\alpha_0 = 0$ this does not (necessarily) hold (may expand here, see also next question).

Question A.4: With a sample of $T = 1000$ observations, it follows that the MLE of v and α, and corresponding LR test statistics for the hypotheses $H_v : v = 4$ and $H_\alpha : \alpha = 0$ are given by:

MLE	Value	Hypothesis	LR statistic
$\hat{v}$	4.7	$H_v : v = 4$	$LR(v = 4) = 2.2$
$\hat{\alpha}$	0.06	$H_\alpha : \alpha = 0$	$LR(\alpha = 0) = 3.1$

Discuss if you would reject H_v and/or H_α. Be precise about which asymptotic distribution(s) and quantiles you are applying.

Solution Question A.4:

Given the results in A.3, would expect LR test for interior points to be asymptotically χ_1^2 distributed (H_v), while for H_α would expect $\frac{1}{2}\chi_1^2$ (be precise in reference to notes). Specifically:

H_v accepted based on 5% quantile of the χ_1^2 distribution (3.84) as under the hypothesis H_v the v parameter is an interior point.

H_α rejected based on the 10% quantile of the χ_1^2 (5% quantile of the $\frac{1}{2}\chi_1^2$) distribution (2.7) as under H_α, α is a boundary point.

Question B:

In order to introduce a stochastic jump in log-returns x_t, consider the "Jump-ARCH" model for x_t as given by

$$x_t = \varepsilon_t + J_t, \quad t = 1, 2, ..., T.$$

With the initial values x_0 and x_{-1} fixed, ε_t is an "ARCH" component as given by

$$\varepsilon_t = \sigma_t z_t,$$

$$\sigma_t^2 = \omega + \alpha x_{t-1}^2, \quad \omega > 0, \quad \alpha \geq 0,$$

with z_t iid, $\mathrm{N}(0,1)$. Note that it is x_t lagged that enters the σ_t^2 and not as for a standard ARCH model, ε_t lagged.

Next the J_t is a "jump" component which is given by a sum of a random number s_t of random $\eta_{t,i}$ variables which are $\mathrm{iidN}(0,\gamma)$, with $\gamma > 0$. That is,

$$J_t = \eta_{t,1} + ... + \eta_{t,s_t} = \sum_{i=1}^{s_t} \eta_{t,i},$$

with s_t stochastic and taking values in $1, 2$ or 3. More specifically, s_t we

consider here the case where s_t is given by a 3-state Markov chain with constant transition probabilities $p_{ij} = P(s_t = j | s_{t-1} = i) \in [0,1]$ for $i, j = 1, 2, 3$, such that $\sum_{j=1}^{3} p_{ij} = 1$ for $i = 1, 2, 3$.

Throughout, we assume that the processes $(z_t)_{t=1,2,...}$ and $(\eta_{t,i})_{t=1,2,....}$ are independent for every $i = 1, 2, 3$, and that the processes $(\eta_{t,i})$ and $(\eta_{t,j})$ are independent for $i \neq j$. Lastly, we also assume that the Markov chain (s_t) is independent of (z_t) and $(\eta_{t,i})$ for every $i = 1, 2, 3$.

Question B.1: State conditions on the transition probabilities $(p_{ij})_{i,j=1,2,3}$ which implies that s_t is weakly mixing.
Note: You do not have to provide any derivations.

Solution Question B.1: Consider the transition matrix defined by

$$P = \begin{pmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \\ p_{13} & p_{23} & p_{33} \end{pmatrix}.$$

The Markov chain (s_t) is weakly mixing if for any $i, j = 1, 2, 3$ there exists an $m_{ij} \in \mathbb{N}$ such that $P(s_{t+m_{ij}} = i | s_t = j) > 0$ *and* the eigenvalues, $\lambda_1 \geq$

$\lambda_2 \geq \lambda_3$, of P satisfy that $\lambda_1 = 1$ and $|\lambda_2| < 1$ and $|\lambda_3| < 1$. In particular, a sufficient condition for weakly mixing is that $p_{ij} > 0$ for all $i, j = 1, 2, 3$.

Note in particular that the chain is not necessarily weakly mixing if $p_{11}, p_{22}, p_{33} < 1$ and $p_{11} + p_{22} + p_{33} > 0$: A counter example is the case where $p_{12} = p_{21} = 1$ and $p_{33} \in (0, 1)$.

Question B.2: State the conditional density of x_t given J_t and $(x_{t-1}, \ldots, x_0)$. That is, give an expression for

$$f(x_t | J_t, x_{t-1}, \ldots, x_0),$$

for $t \geq 1$. Use this to give an interpretation of the model for the log-returns x_t conditional on J_t and past x_t's.

Solution Question B.2: Using that $(J_t, x_{t-1}, \ldots, x_0)$ and z_t are independent, we have that $\varepsilon_t | (J_t, x_{t-1}, \ldots, x_0) \overset{d}{=} \varepsilon_t | (x_{t-1}) \overset{d}{=} N(0, \sigma_t^2)$. Hence, $x_t | (J_t, x_{t-1}, \ldots, x_0) \overset{d}{=} \varepsilon_t + J_t | (J_t, x_{t-1}, \ldots, x_0) \overset{d}{=} N(J_t, \sigma_t^2)$. So we have that

$$f(x_t | J_t, x_{t-1}, \ldots, x_0) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(x_t - J_t)^2}{2\sigma_t^2} \right).$$

Hence the "Jumps" J_t, when conditioned upon, may be considered as a change in the level of log-returns.

Question B.3: Here we consider the model without conditioning on the "Jumps" J_t but conditional on the value of s_t (and lagged x_t's).

Argue that J_t conditional on s_t is $N\left(0, \sum_{j=1}^{s_t} \gamma\right)$ distributed.

Next, use this to show that for $t \geq 1$ the conditional density of x_t given $s_t = i$ and $(x_{t-1}, \ldots, x_0)$ is given by

$$f(x_t | s_t = i, x_{t-1}, \ldots, x_0) = \frac{1}{\sqrt{2\pi(\sigma_t^2 + i\gamma)}} \exp\left(-\frac{x_t^2}{2(\sigma_t^2 + i\gamma)} \right), \quad i = 1, 2, 3.$$

Give an interpretation of the model in this case.

Solution Question B.3: As J_t is independent of $(x_{t-1}, \ldots, x_0)$, $J_t | \{s_t = i\}, (x_{t-1}, \ldots, x_0)$ is given by $\sum_{j=1}^{i} \eta_{t,j} \overset{d}{=} N(0, \gamma i)$. Moreover, $\varepsilon_t | \{s_t = i\}, (x_{t-1}, \ldots, x_0) \overset{d}{=} (\varepsilon_t | x_{t-1}) \overset{d}{=} N(0, \sigma_t^2)$. Hence by independence of J_t and z_t, we have that

$$\begin{pmatrix} J_t \\ \varepsilon_t \end{pmatrix} | \{s_t = i\}, (x_{t-1}, \ldots, x_0) \overset{d}{=} N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \gamma i & 0 \\ 0 & \sigma_t^2 \end{pmatrix} \right).$$

Hence, using that $x_t = J_t + \varepsilon_t$, we have that $x_t | \{s_t = i\}, (x_{t-1}, \ldots, x_0) \overset{d}{=} N(0, \gamma i + \sigma_t^2)$, such that

$$f(x_t | s_t = i, x_{t-1}, \ldots, x_0) = f(x_t | s_t = i, x_{t-1}) = \frac{1}{\sqrt{2\pi(\sigma_t^2 + i\gamma)}} \exp\left(-\frac{x_t^2}{2(\sigma_t^2 + i\gamma)}\right).$$

Hence, the "Jumps", when only conditioned on the state variable s_t, may be considered as a change in the level of the conditional variance, $V(x_t | s_t, x_{t-1}) = \sigma_t^2 + s_t \gamma$.

Question B.4: Let $\theta = (\omega, \alpha, \gamma, p_{11}, p_{12}, p_{21}, p_{22}, p_{31}, p_{32})'$ denote the model parameters. The log-likelihood function is given by

$$L_T(\theta) = \sum_{t=1}^{T} \log f_\theta(x_t | x_{t-1}, \ldots, x_0).$$

Explain how you would estimate θ. In particular, explain how the log-likelihood function can be evaluated.

Solution Question B.4: As the log-likelihood function is given by

$$L_T(\theta) = \sum_{t=1}^{T} \log f_\theta(x_t | x_{t-1}, \ldots, x_0),$$

as usual, evaluating the log-likelihood function boils down to computing $f_\theta(x_t | x_{t-1}, \ldots, x_0)$ for $t = 1, \ldots, T$. This is done using a filtering algorithm. Specifically, note that

$$f_\theta(x_t | x_{t-1}, \ldots, x_0) = \sum_{i=1}^{3} f_\theta(x_t | s_t = i, x_{t-1}) P_\theta(s_t = i | x_{t-1}, \ldots, x_0),$$

where $f_\theta(x_t | s_t = i, x_{t-1})$ was derived in Question B.3. The predicted probabilities $P_\theta(s_t = i | x_{t-1}, \ldots, x_0)$ can be computed using a filtering algorithm. Details should be included.