Financial Econometrics Exam Winter 2020

Flow ID: 61

December 22, 2020

1 Theoretical part

During the course, we have widely employed the Gaussian and the Student's t distributions to model the conditional distribution of financial r eturns. However, many other distributions are available, and it is not always easy to choose the best one. For example, Nelson (1991) proposed to model the conditional distribution of financial returns using the Generalized Error Distribution (GED). Specifically, for a generic random variable $Y \in \Re$ with mean $\mu \in \Re$ and scale $\varphi > 0$, the probability density function (pdf) evaluated in y is:

$$p(y;\mu,\varphi,v) = \left[2^{1+1/v}\varphi\Gamma(1+1/v)\right]^{-1} \exp\left(-\frac{|(y-\mu)/\varphi|^v}{2}\right)$$

where $\Gamma(\cdot)$ is the Gamma function, and $v > 0$ is a tail-thickness parameter. The standard deviation σ is related to the scale by the formula:

$$\sigma = 2^{1/v}\sqrt{\frac{\Gamma(3/v)}{\Gamma(1/v)}}\varphi$$

Consider a Generalised Autoregressive Score (GAS) model, where the conditional distribution of financial returns $Y_t \mid \mathcal{F}_{t-1}$ is GED with location 0 ($\mu = 0$) and time varying scale φ_t:

$$Y_t \mid \mathcal{F}_{t-1} \sim GED\left(0, \varphi_t, v\right)$$

where $\mathcal{F}_{t-1}$ represents the filtration generated by the process $\{Y_s, s \leq t\}$.

1.1 Question 1.1 - Derive the GAS

Derive the GAS updating equation for φ_t using an exponential link function, $\varphi_t = \exp\left(\tilde{\varphi}_t\right)$, and imposing unit scaling, i.e. do not scale the score ($d = 0$). Use s_t to label the score of $Y_t \mid \mathcal{F}_{t-1}$ with respect to $\tilde{\varphi}_t$. Evidently, s_t is going to be a function of the realization y_t, the tail-thickness parameter , and the reparametrized scale $\tilde{\varphi}_t$, i.e. $s_t = s\left(y_t, v, \tilde{\varphi}_t\right)$. Compute $E\left[s_t \mid \mathcal{F}_{t-1}\right]$ and $E\left[s_t\right]$.

First of all we want to derive the updating equation for φ_t using the exponential link function, $\varphi_t = \exp\left(\tilde{\varphi}_t\right)$ while imposing unit scaling. To do so we start out by calculating the score, which we know is given as

$$\nabla_t = \frac{\partial \log p\left(y_t \mid \mathbf{y}_{1:t-1}; \varphi\right)}{\partial \varphi_t} = \frac{\partial \ell_t}{\partial \varphi_t}$$

We start out by taking the natural logarithm

$$\log\left(p(y;\mu,\varphi,v)\right) = -\left(1+\frac{1}{v}\right)\ln\left(2\right) - \ln\left(\varphi_t\right) - \ln\left(\Gamma\left(1+\frac{1}{v}\right)\right) - \frac{1}{2}\left|\frac{y_t-\mu}{\varphi_t}\right|^v = \ell_t$$

We have given from the assignment that the location is zero, i.e. $\mu = 0$. We then proceed to differentiate wrt. to φ_t. Notice we use the rule for derivative of absolute values, i.e. $|x|' = \frac{x}{|x|}$

$$\begin{aligned}
\nabla_t = \frac{\partial \ell_t}{\partial \varphi_t} &= -\frac{1}{\varphi_t} + \frac{1}{2}\frac{\varphi_t v \left|y_t\right|^v}{\left|\varphi_t\right|^{v+2}} \\
&= \frac{1}{\varphi_t}\left(\frac{1}{2}\frac{\varphi_t^2 v \left|y_t\right|^v}{\left|\varphi_t\right|^{v+2}} - 1\right) \\
&= \frac{1}{\exp\left(\tilde{\varphi}_t\right)}\left(\frac{1}{2}\frac{v \left|y_t\right|^v}{\exp\left(\tilde{\varphi}_t\right)^v} - 1\right)
\end{aligned}$$

We know that when we have unit scaling scaling, i.e. $d = 0$, and therefore do not scale the score, $u_t = s_t$.

$$\begin{aligned}
\varphi_t &= \exp\left(\tilde{\varphi}_t\right) \\
\tilde{\varphi}_t &= \omega + \alpha \tilde{u}_{t-1} + \beta \tilde{\varphi_{t-1}}
\end{aligned}$$

where

$$\begin{aligned}
s_t = \frac{\partial \ell_t}{\partial \tilde{\varphi}_t} &= \frac{\partial \varphi_t}{\partial \tilde{\varphi}_t}\frac{\partial \ell_t}{\partial \varphi_t} \\
&= \varphi_t * \frac{1}{\varphi_t}\left(\frac{1}{2}\frac{\varphi_t^2 v \left|y_t\right|^v}{\left|\varphi_t\right|^{v+2}} - 1\right) \\
&= \frac{1}{2}\frac{\varphi_t^2 v \left|y_t\right|^v}{\left|\varphi_t\right|^{v+2}} - 1 \\
&= \frac{1}{2}\frac{\exp\left(\tilde{\varphi}_t\right)^2 v \left|y_t\right|^v}{\left|\exp\left(\tilde{\varphi}_t\right)\right|^{v+2}} - 1 \\
&= \frac{1}{2}\frac{v \left|y_t\right|^v}{\exp\left(\tilde{\varphi}_t\right)^v} - 1
\end{aligned}$$

We can therefore write the GAS model as

$$\begin{aligned}
\varphi_t &= \exp\left(\tilde{\varphi}_t\right) \\
\tilde{\varphi}_t &= \omega + \alpha\left[\frac{1}{2}\frac{v \left|y_{t-1}\right|^v}{\exp\left(\tilde{\varphi_{t-1}}\right)^v} - 1\right] + \beta \tilde{\varphi_{t-1}}
\end{aligned}$$

We can unfold to find the uncondtional mean

$$\begin{aligned}
\tilde{\varphi}_t &= \omega + \alpha \tilde{u_{t-1}} + \beta\left(\omega + \alpha \tilde{u}_{t-2} + \beta\left(\omega + \alpha \tilde{u}_{t-3} + \beta \tilde{\varphi_{t-3}}\right)\right) \\
&= ... \\
E\left[\tilde{\varphi}_t\right] &= \frac{\omega}{1-\beta}
\end{aligned}$$

We now proceed to compute the conditional expectation of the score, i.e. $E\left[s_t \mid \mathcal{F}_{t-1}\right]$

$$E\left[s_t \mid \mathcal{F}_{t-1}\right]=E\left[\frac{1}{2}\frac{v\left|y_t\right|^v}{\exp\left(\tilde{\varphi}_t\right)^v}-1 \mid \mathcal{F}_{t-1}\right]$$

where we remember that $s_t=\frac{\partial\varphi_t}{\partial\tilde{\varphi}_t}\frac{\partial\ell_t}{\partial\varphi_t}=\varphi_t*\nabla_t$. Hence, we can rewrite

$$E\left[s_t \mid \mathcal{F}_{t-1}\right]=E\left[\varphi_t\nabla_t \mid \mathcal{F}_{t-1}\right]$$

We say φ_t is adapted to the filtration and therefore it is known at time $t-1$, and therefore we can move it out of the expectation. Furthermore, we know that one of the properties of the score is, $\mathbb{E}_{t-1}\left[\nabla_t\right]=0\Rightarrow\mathbb{E}_t\left[\nabla_t \mid \mathcal{F}_{t-1}\right]=0$

$$E\left[s_t \mid \mathcal{F}_{t-1}\right]=\varphi_t E\left[\nabla_t \mid \mathcal{F}_{t-1}\right]=0$$

We now proceed to compute the unconditional expectation of the score, i.e. $E\left[s_t\right]$. Using law of iterated expectations (i.e. $E[X]=E[E[X \mid Y]]$) will give us

$$E\left[s_t\right]=E\left[\underbrace{E\left[s_t \mid \mathcal{F}_{t-1}\right]}_{=0}\right]=0$$

1.2 Question 1.2 - Log-likelihood

Write the log likelihood of the model and state the model parameter constraints. Which constraint do we need to impose to ensure that the process of $\{\tilde{\varphi}_t, t>0\}$ is covariance stationary (assume that $E\left[s_t^2\right]<\infty$ for all values of $\tilde{\varphi}_t$ and v)?

Now we want to write the log likelihood of the model

$$Y_{1:T}=\left(y_1,\ldots,y_T\right)'$$

$$\begin{aligned}
L\left(\theta \mid Y_{1:T}\right)&=\sum_{t=1}^{T}\ln\left(p(y;\mu,\varphi,v)\right)\\
&=T\left(-\ln\left(\Gamma\left(1+\frac{1}{v}\right)-\ln(2)\left(1+\frac{1}{v}\right)\right)\right)-\sum_{t=1}^{T}\left(\ln\left(\varphi_t\right)+\frac{1}{2}\left|\frac{y_t}{\varphi_t}\right|^v\right)
\end{aligned}$$

For covariance stationarity we require

$$\begin{array}{ll}
(i): & E\left(y_t\right)=\mu_y=\text{ constant}\\
(ii): & \operatorname{var}\left(y_t\right)=\sigma_y^2=\text{ constant}\\
(iii): & \operatorname{cov}\left(y_t, y_{t-s}\right)=\gamma_s=f(s)
\end{array}$$

Our GAS model is given as

$$\begin{aligned}
\tilde{\varphi_t}&=\omega+\alpha u_{t-1}^{\sim}+\beta\varphi_{t-1}^{\sim}\\
&=\frac{\omega}{1-\beta}+\alpha\sum_{i=0}^{\infty}\beta^i u_{t-i-1}^{\sim}
\end{aligned}$$

Where we made use of iteratively recursion and used the sum properties of a geometric series. To ensure that the proces of $\tilde{\varphi}_t$ is weakly stationary (covariance stationary) we look at the mean, variance and covariance

We calculate the mean as

$$\begin{aligned} E\left[\tilde{\varphi}_t\right] &= E\left[\frac{\omega}{1-\beta} + \alpha \sum_{i=0}^{\infty} \beta^i u_{t\tilde{-}i-1}\right] \\ &= \frac{\omega}{1-\beta} + \alpha \sum_{i=0}^{\infty} \beta^i \underbrace{E\left[s_{t-i-1}\right]}_{=0} \\ &= \frac{\omega}{1-\beta} = \text{constant} \end{aligned}$$

Notice that $E\left(s_t\right) = E\left(\tilde{u}_t\right) = 0$ for all t, and therefore the mean is constant. Where necessary conditions for covariance stationarity are $\mid \beta \mid < 1$.

We can calculate the variance as

$$\begin{aligned} Var\left[\tilde{\varphi}_t\right] &= \mathbb{E}\left[\left(\tilde{\varphi}_t - E\left[\tilde{\varphi}_t\right]\right)^2\right] \\ &= \mathbb{E}\left[\left(\frac{\omega}{1-\beta} + \alpha \sum_{i=0}^{\infty} \beta^i u_{t\tilde{-}i-1} - \frac{\omega}{1-\beta}\right)^2\right] \\ &= \mathbb{E}\left[\alpha^2 \sum_{i=0}^{\infty} \beta^{2i} u^{2\,\tilde{}}_{t-i-1}\right] \\ &= \mathbb{E}\left[\alpha^2 \sum_{i=0}^{\infty} \beta^{2i} u^{2\,\tilde{}}_{t-i-1}\right] \\ &= \alpha^2 \sum_{i=0}^{\infty} \beta^{2i} \underbrace{E\left[s^2_{t-i-1}\right]}_{<\infty} = \text{constant} \end{aligned}$$

And notice that the variance is constant, as we assume $E\left[s_t^2\right] < \infty$. Also we need to have $\mid \beta \mid < 1$ and $\alpha < \infty$

Now we turn to the covariance where we want to show that it is a function of the distance between scale parameter

$$
\begin{aligned}
cov\left(\tilde{\varphi}_t, \tilde{\varphi_{t-h}}\right) &= \mathbb{E}\left[\left(\tilde{\varphi}_t - \mathbb{E}\left[\tilde{\varphi}_t\right]\right) * \left(\tilde{\varphi_{t-h}} - \mathbb{E}\left[\tilde{\varphi_{t-h}}\right]\right)\right] \\
&= \mathbb{E}\left[\left(\frac{\omega}{1-\beta} + \alpha\sum_{i=0}^{\infty}\beta^i \tilde{u_{t-i-1}} - \frac{\omega}{1-\beta}\right) * \left(\frac{\omega}{1-\beta} + \alpha\sum_{s=0}^{\infty}\beta^i \tilde{u_{t-i-1-h}} - \frac{\omega}{1-\beta}\right)\right] \\
&= \mathbb{E}\left[\left(\alpha\sum_{i=0}^{\infty}\beta^i \tilde{u_{t-i-1}}\right) * \left(\alpha\sum_{i=0}^{\infty}\beta^i \tilde{u_{t-i-1-h}}\right)\right] \\
&= \alpha^2\mathbb{E}\left[\left(\sum_{i=0}^{\infty}\beta^i \tilde{u_{t-i-1}}\right) * \left(\sum_{i=0}^{\infty}\beta^i \tilde{u_{t-i-1-h}}\right)\right]
\end{aligned}
$$

Notice we make use of the fact that

$$
\mathbb{E}\left[\tilde{u_t}\tilde{u_{t-j}}\right] = \mathbb{E}\left[\mathbb{E}\left[\tilde{u_t}\tilde{u_{t-j}} \mid \mathcal{F}_{t-1}\right]\right] = \mathbb{E}\left[\tilde{u_{t-j}}\underbrace{\mathbb{E}\left[\tilde{u_t} \mid \mathcal{F}_{t-1}\right]}_{=0}\right] = 0, \forall j
$$

Where we know the last equation holds as we showed this in exercise 1.1. This means that all the cross terms will be equal to zero when multiyplying the two sums together, while only the similar terms will sustain

$$
\begin{aligned}
\alpha^2\mathbb{E}\left[\left(\sum_{i=0}^{\infty}\beta^i \tilde{u_{t-i-1}}\right) * \left(\sum_{i=0}^{\infty}\beta^i \tilde{u_{t-i-1-h}}\right)\right] &= \alpha^2\beta^h\mathbb{E}\left[\alpha^2\sum_{i=0}^{\infty}\beta^{2i}\tilde{u^2_{t-i-1}}\right] \\
&= \beta^h Var\left[\tilde{\varphi}_t\right] \\
&= \gamma_h = f(h)
\end{aligned}
$$

We therefore find that the covariance becomes a function of the distance h between the scale parameters $\tilde{\varphi}_t$ and $\tilde{\varphi_{t-h}}$, when $E\left[s_t^2\right] < \infty$.

To summerize the contraint that ensures that the process of $\{\tilde{\varphi}_t, t > 0\}$ is covariance stationary for all values of $\tilde{\varphi}_t$ and v is $\mid \beta \mid < 1$

1.3 Question 1.3 - Comparison to Gaussian dist

When $v = 2$, the GED distribution collapses to the Gaussian distribution. Compare the functions $f_1(y_t) = s(y_t, 1, 0)$ and $f_2(y_t) = s(y_t, 2, 0)$. Which conclusion you draw about the response of the scale parameter to different values of y_t in the case $v = 1$ and $v = 2$?

We start out by inserting the values in a our score functions and get.

$$
f_1(y_t) = s(y_t, 1, 0) = \frac{\mid y_t \mid}{2} - 1
$$

$$
f_2(y_t) = s(y_t, 2, 0) = |y_t|^2 - 1
$$

If we insert these two different functions of the scores back into the scale parameter. We can then see the response to diferent values of y_t

$$\tilde{\varphi_t} = \omega + \alpha\left(\frac{\mid y_{t-1}\mid}{2} - 1\right) + \beta\tilde{\varphi_{t-1}}$$

$$\tilde{\varphi_t} = \omega + \alpha\left(|y_{t-1}|^2 - 1\right) + \beta\tilde{\varphi_{t-1}}$$

From these equations we immediately notice that the scale parameter reacts more strongly to changes in y_{t-1} in the normal distribution (when $v = 2$) than when v=1. We could also differentiate to see that

$$\frac{\partial s\left(y_t, 1, 0\right)}{\partial \mid y_t \mid} = \frac{\partial}{\partial \mid y_t \mid}\left(\frac{\mid y_t \mid}{2} - 1\right) = \frac{1}{2}$$

$$\frac{\partial s\left(y_t, 1, 0\right)}{\partial \mid y_t \mid} = \frac{\partial}{\partial \mid y_t \mid}\left(|y_t|^2 - 1\right) = 2 \mid y_t \mid$$

Hence, we can conclude that the response of the scale parameter reacts more strongly when $v = 2$ compared to $v = 1$. When $v = 2$ the distibution collapses to the normal distribution, and when $v = 1$ it collapses to a double exponential distribution. When $v < 2$, the distribution has thicker tails than in the normal distribution and when $v > 2$ then tails get thinner than the normal distribution. The reaction in the scale parameter, φ_t for different values of $y_t$in the two cases is therefore quite different. If our score changes a lot, the scale while be affected a lot for large observations of y_t. As taught in class the normal distribution overreacted at extreme observations, while the t distribution might be better to use, as it has thicker tails. In general we can conclude that a larger v indicates thinner tails, which in turn means a larger reaction on extreme oberservations. If v is large, it might therefore overreact in case of large observations.

2 Computational part

Made in the R script: ExamScript20

3 Empirical part

3.1 (a)

We consider the general dynamic volatility model

$$y_t = \sigma_t \varepsilon_t$$

Write down the GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) models for σ_t^2 .

GARCH$(1, 1)$:
The GARCH(1,1) model is given as

$$\sigma_t^2 = \omega + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

We can rewrite this by recursion into

$$\begin{aligned}
\sigma_t^2 &= \omega + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\
&= \omega + \alpha_1 \sigma_{t-1}^2 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\
&= \omega + \left(\beta_1 + \alpha_1 \varepsilon_{t-1}^2\right) \sigma_{t-1}^2 \\
&= \omega + \left(\beta_1 + \alpha_1 \varepsilon_{t-1}^2\right) \left[\omega + \left(\beta_1 + \alpha_1 \varepsilon_{t-2}^2\right) \sigma_{t-2}^2\right] \\
&= \omega \left[1 + \left(\beta_1 + \alpha_1 \varepsilon_{t-1}^2\right)\right] + \left(\beta_1 + \alpha_1 \varepsilon_{t-2}^2\right) \left(\beta_1 + \alpha_1 \varepsilon_{t-1}^2\right) \sigma_{t-2}^2 \\
&= \ldots \\
&= \omega \left[1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} \left(\beta_1 + \alpha_1 \varepsilon_{t-i}^2\right)\right]
\end{aligned}$$

We know that for weak stationarity it must hold that

$$\begin{aligned}
E\left(y_t\right) &= \mu_y < \pm\infty \\
var\left(y_t\right) &= \sigma_y^2 < \infty \\
cov\left(y_t, y_{t-s}\right) &= \gamma_s = f(s)
\end{aligned}$$

whereas for strong stationarity it must hold that

$$F_X\left(x_{t_1+h}, \ldots, x_{t_n+h}\right) = F_X\left(x_{t_1}, \ldots x_{t_n}\right)$$

We start out by calculating the unconditional mean for $y_t$as

$$\begin{aligned}
\mathbb{E}\left[y_t\right] &= \mathbb{E}\left[\sigma_t \varepsilon_t\right] \\
&= \mathbb{E}\left[\sigma_t\right] \underbrace{\mathbb{E}\left[\varepsilon_t\right]}_{=0} \\
&= 0
\end{aligned}$$

And the variance is calculated as

$$\begin{aligned}
Var\left(y_t\right) = \mathbb{E}\left[y_t^2\right] = \mathbb{E}\left[\sigma_t^2\right] &= \mathbb{E}\left[\omega \left[1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} \left(\beta_1 + \alpha_1 \varepsilon_{t-i}^2\right)\right]\right] \\
&= \omega \left[1 + \sum_{k=1}^{\infty} \mathbb{E}\left[\prod_{i=1}^{k} \left(\beta_1 + \alpha_1 \varepsilon_{t-i}^2\right)\right]\right]
\end{aligned}$$

Where we use that $\mathbb{E}\left[\prod_{i=1}^{k} \underbrace{\left(\beta_1 + \alpha_1 \varepsilon_{t-i}^2\right)}_{b_{t-i}}\right] = \mathbb{E}\left[b_t\right]^k \rightarrow \mathbb{E}\left[b_t\right] = \mathbb{E}\left[\beta_1 + \alpha_1 \varepsilon_t^2\right] = \alpha_1 + \beta_1$

$$\begin{aligned}
\mathbb{E}\left[y_t^2\right] = \mathbb{E}\left[\sigma_t^2\right] &= \omega\left[1+\sum_{k=1}^{\infty}\left(\alpha_1+\beta_1\right)^k\right] \\
&= \omega\left[\sum_{k=0}^{\infty}\left(\alpha_1+\beta_1\right)^k\right] \\
&= \frac{\omega}{1-\left(\alpha_1+\beta_1\right)} \\
&= \frac{\omega}{1-\alpha_1-\beta_1}
\end{aligned}$$

Where we have used rules for a geometric series. We notice that that the process is covariance stationary if and only if $\alpha_1 + \beta_1 < 1$. When these parameter constraints are satisfied we say that the process are stable. If these are not satisfied the process would be explosive. Futhermore we need $\omega > 0, \beta > 0, \alpha > 0$ for positivity. We then proceed to look at the covariance.

$$\begin{aligned}
cov\left(y_t, y_{t-s}\right) &= \mathbb{E}\left[y_t y_{t-s}\right] - \underbrace{\mathbb{E}\left[y_t\right]\mathbb{E}\left[y_{t-s}\right]}_{=0} \\
&= \mathbb{E}\left[\mathbb{E}\left[y_t y_{t-s} \mid \mathcal{F}_{t-1}\right]\right] \\
&= \mathbb{E}\left[y_{t-s}\underbrace{\mathbb{E}\left[y_t \mid \mathcal{F}_{t-1}\right]}_{=0}\right] = 0
\end{aligned}$$

Where we have used the law of iterated expectations. The last equation holds because y_t represents a Martingale difference sequence, and therefore, $\mathbb{E}\left[y_t \mid \mathcal{F}_{t-1}\right] = 0$ for all t, where $\mathcal{F}_{t-1}$ represents the filtration genereated up to time t-1. This means that the process has weak stationarity.

This is, however, only a sufficient and not necessary condition for process to be strictly stationary. Nelson showed in 1990, that if $\omega > 0$ we have $\sigma_t^2 < \infty$ almost surely and that the process of y_t will be strictly stationary if and only if

$$\mathbb{E}\left[\ln\left(\beta + \alpha\varepsilon_t^2\right)\right] < 0$$

by Jensen inequality we have

$$\mathbb{E}\left[\ln\left(\beta + \alpha\varepsilon_t^2\right)\right] < \ln\left(\mathbb{E}\left[\left(\beta + \alpha\varepsilon_t^2\right)\right]\right) = \ln(\alpha+\beta)$$

We notice that when $\alpha + \beta = 1$, the model is strictly stationary, even though we just found out it is not weakly stationary under those condition, as the variance doesn't exist.

Proof:

We define two processes

$$
\begin{aligned}
\sigma_t^2 &= \sigma_0^2 \prod_{i=1}^{t} \left(\beta + \alpha\varepsilon_{t-i}^2\right) + \omega \left[1 + \sum_{k=1}^{t-1} \prod_{i=1}^{k} \left(\beta + \alpha\varepsilon_{t-i}^2\right)\right] \\
{}_u\sigma_t^2 &= \omega \left[1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} \left(\beta + \alpha\varepsilon_{t-i}^2\right)\right]
\end{aligned}
$$

where σ_t^2 is the process initialized at time $t = 0$ at σ_0^2 and ${}_u\sigma_t^2$ is the process that goes back to the infinite past.We start by looking at Theorem 1 of Nelson (1990) where $\omega = 0$

(a) ${}_u\sigma_t^2 = 0$ almost surely

(b) We take notice the second terms becomes zero and then takes the natural logarithm.

$$
\sigma_t^2 = \sigma_0^2 \prod_{i=1}^{t} \left(\beta + \alpha\varepsilon_{t-i}^2\right)
$$

$$
\begin{aligned}
\log\left(\sigma_t^2\right) &= \log\left(\sigma_0^2 \prod_{i=1}^{t} \left(\beta + \alpha\varepsilon_{t-i}^2\right)\right) \\
&= \log(\sigma_0^2) + \sum_{i=1}^{t} \log\left(\beta + \alpha\varepsilon_{t-i}^2\right)
\end{aligned}
$$

If we define a function $x_t = \log\left(\beta + \alpha\varepsilon_t^2\right)$ we have

$$
\log\left(\sigma_t^2\right) = \log(\sigma_0^2) + \sum_{i=1}^{t} x_t
$$

We further define $\mathbb{E}\left[x_t\right] = \mathbb{E}\left[\log\left(\beta + \alpha\varepsilon_t^2\right)\right] = \bar{x}$, and then $\tilde{x}_t = x - \bar{x}$ where there expectation of $\tilde{x}$is zero per definition.

$$
\begin{aligned}
\log\left(\sigma_t^2\right) &= \log(\sigma_0^2) + \sum_{i=1}^{t} \left(\tilde{x}_t + \bar{x}\right) \\
&= \log(\sigma_0^2) + t\bar{x} + \sum_{i=1}^{t} \tilde{x}_t
\end{aligned}
$$

We immediately recognize that this is a random walk with a drift. This process is defined as

$$
\log\left(\sigma_t^2\right) = \bar{x} + \log\left(\sigma_{t-1}^2\right) + \tilde{x}_t
$$

where we see we get the above equation by iteratively substition.

If $t \to \infty$ then $\log\left(\sigma_t^2\right) = \begin{cases} -\infty & \text{if } \bar{x} < 0 \\ +\infty & \text{if } \bar{x} > 0 \end{cases}$

Where our moment conditon is $\mathbb{E}\left[\log\left(\beta + \alpha\varepsilon_t^2\right)\right] = \bar{x}$. We see that if $\omega = 0$ the variance process diverges.

We now turn to another theorem from Nelson (1990) where $\omega > 0$. Where we have two different cases of $\mathbb{E}\left[\log\left(\beta + \alpha\varepsilon_t^2\right)\right]$.

First case: If we have $\omega > 0$ and $\mathbb{E}\left[\log\left(\beta + \alpha\varepsilon_t^2\right)\right] \geq 0$ then (a) $\sigma_t^2 \to \infty$ *a.s.* and (b) ${}_u\sigma_t^2 \to \infty$ *a.s.* for all t. This mean that both processes diverges. We have

$$\sigma_t^2 = \sigma_0^2 \prod_{i=1}^{t}\left(\beta + \alpha\varepsilon_{t-1}^2\right) + \omega\left[1 + \sum_{k=1}^{t-1}\prod_{i=1}^{k}\left(\beta + \alpha\varepsilon_{t-i}^2\right)\right]$$

We know the random variable σ_t^2 is the sum of the two terms and that these are both strictly positive. Therefore it most hold that σ_t^2 is greater or equal than the second term. Which is the sum of a number variables, instead we can bound it. We say that σ_t^2 is going to be greater than the largest of the sum, i.e. the supremum of $\prod_{i=1}^{k}\left(\beta + \alpha\varepsilon_{t-i}^2\right)$ up to time t-1.

$$\begin{aligned}\sigma_t^2 &\geq \omega\left[1 + \sum_{k=1}^{t-1}\prod_{i=1}^{k}\left(\beta + \alpha\varepsilon_{t-i}^2\right)\right] \\ &\geq \omega \sup_{1\leq k\leq t-1}\prod_{i=1}^{k}\left(\beta + \alpha\varepsilon_{t-i}^2\right)\end{aligned}$$

We now take the logarithm of both sides and then we proceed

$$\log\left(\sigma_t^2\right) \geq \log\left(\omega\right) + \sup_{1\leq k\leq t-1}\sum_{i=1}^{k}\log\left(\left(\beta + \alpha\varepsilon_{t-i}^2\right)\right)$$

Which is term similar to what we had in theorem 1 of Nelson (1990). Using the same arguments as before, we can show that

$$\begin{aligned}\sum_{i=1}^{k}\log\left(\beta + \alpha\varepsilon_{t-i}^2\right) \to \infty \quad iff \quad \mathbb{E}\left[\log\left(\beta + \alpha\varepsilon_t^2\right)\right] > 0 \\ \lim_{t\to\infty}\sup_{1\leq k\leq t-1}\sum_{i=1}^{k}\log\left(\beta + \alpha\varepsilon_{t-i}^2\right) = \infty \quad iff \quad \mathbb{E}\left[\log\left(\beta + \alpha\varepsilon_t^2\right)\right] = 0\end{aligned}$$

The process therefore diverges to $+\infty$ when $\mathbb{E}\left[\log\left(\beta + \alpha\varepsilon_t^2\right)\right] \geq 0$. Point b can be proved using similar arguments as it the same as the second term of the process initilized at time t.

Second case: If we have $\omega > 0$ and $\mathbb{E}\left[\log\left(\beta + \alpha\varepsilon_t^2\right)\right] < 0$ then (c) $\frac{\omega}{1-\beta}\ {}_u\sigma_t^2 < \infty$for all t *a.s.* and (d) ${}_u\sigma_t^2 - \sigma_t^2 \to 0$ *a.s.* We take the σ_t from the unconditional process and study its minimum (lower bound)

$${}_u\sigma_t^2 = \omega\left[1 + \sum_{k=1}^{\infty}\prod_{i=1}^{k}\left(\beta + \alpha\varepsilon_{t-i}^2\right)\right]$$

To do so we note that it's infimum is achieved for $\varepsilon_{t-1}^2 = 0$ for all t, so we have that

$$
\begin{aligned}
\inf \ {}_u\sigma_t^2 &= \omega \left[1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} \beta \right] \\
&= \omega \left[1 + \sum_{k=1}^{\infty} \beta^k \right] \\
&= \omega \sum_{k=1}^{\infty} \beta^k \\
&= \frac{\omega}{1-\beta}
\end{aligned}
$$

We make use of $0 < \beta < 1$ to make a sum rule of a geometric series. However this $\beta < 1$ is different from $\alpha + \beta < 1$. The upper bound relies on the fact that each element $\prod_{i=1}^{k} \left(\beta + \alpha\varepsilon_{t-i}^2\right)$ is of order $\exp(-\lambda k)$ for some constant $\lambda > 0$. We can see that the process does not diverge when k goes to infinity, because $O(\exp(-\lambda k))$ goes to 0 very fast when k grows.

To prove point (d) we subtract ${}_u\sigma_t^2$ from σ_t^2, i.e.

$$
\sigma_t^2 - {}_u\sigma_t^2 = \sigma_0^2 \prod_{i=1}^{t} \left(\beta + \alpha\varepsilon_{t-i}^2\right) - \omega \left[1 + \sum_{k=t}^{\infty} \prod_{i=1}^{k} \left(\beta + \alpha\varepsilon_{t-i}^2\right) \right] = 0
$$

We know the first term goes to zero when t goes to infinit provided $\mathbb{E}\left[\log\left(\beta + \alpha\varepsilon_t^2\right)\right] < 0$ from theorem 1. Furtermore it can be shown that the second term also goes to zero because, as in the case of the upper bound, each element $\prod_{i=1}^{k} \left(\beta + \alpha\varepsilon_{t-i}^2\right)$ is of order $exp(-\lambda k)$, for some $\lambda > 0$, and therefore it does not diverge as k goes to infinity.

EGARCH$(1,1)$

The Exponential Garch was introduced by Nelson (1991) to bypass some of the limitations of the original GARCH model. First of all the parameters α and β has to be constrained in order to ensure positively of the variance process. Furthermore empiricial evidence indicates that there is a asymmetrical response of volatility in case of shocks.

$$
\log\left(\sigma_t^2\right) = \omega + \beta_1 \log\left(\sigma_{t-1}^2\right) + \phi\varepsilon_{t-1} + \psi\left(|\varepsilon_{t-1}| - \mathbb{E}\,|\varepsilon_{t-1}|\right)
$$

Where we have the density of ε_t is given as1

$$
E\,|\varepsilon_t| = \frac{\lambda 2^{1/v}\Gamma(2/v)}{\Gamma(1/v)}
$$

where $\lambda \equiv \left[2^{(-2/v)}\Gamma(1/v)/\Gamma(3/v)\right]^{1/2}$ and $\Gamma(\cdot)$ is the gamma function and ε_t has a GED distribution, i.e. $\varepsilon_t \sim GED(0,1,v)$.

We calculate the unconditional variance by

$$\begin{aligned}\mathbb{E}\left[\log\left(\sigma_t^2\right)\right] &= \mathbb{E}\left[\omega+\beta_1\log\left(\sigma_{t-1}^2\right)+\phi\varepsilon_{t-1}+\psi\left(|\varepsilon_{t-1}|-\mathbb{E}|\varepsilon_{t-1}|\right)\right]\\ &= \omega+\beta_1\mathbb{E}\left[\log\left(\sigma_{t-1}^2\right)\right]+\phi\underbrace{\mathbb{E}\left[\varepsilon_{t-1}\right]}_{=0}+\psi\left(\underbrace{\mathbb{E}\left[|\varepsilon_{t-1}|\right]-\mathbb{E}\left[\mathbb{E}|\varepsilon_{t-1}|\right]}_{=0}\right)\end{aligned}$$

where $\mathbb{E}\left[\mathbb{E}|\varepsilon_{t-1}|\right]=\mathbb{E}\left[|\varepsilon_{t-1}|\right]$ due to law of iterated expectations

$$\begin{aligned}\mathbb{E}\left[\log\left(\sigma_t^2\right)\right] &= \omega+\beta_1\mathbb{E}\left[\log\left(\sigma_{t-1}^2\right)\right]\\ &\Longleftrightarrow\\ \mathbb{E}\left[\log\left(\sigma_t^2\right)\right] &= \frac{\omega}{1-\beta_1}\\ \mathbb{E}\left[\sigma_t^2\right] &= e^{\frac{\omega}{1-\beta_1}}\end{aligned}$$

We use this as our unconditonal variance in the R code, where we notice $\mid\beta\mid<1$

GJR-GARCH$(1,1)$
We have the returns follow the process

$$y_t=\sigma_t\varepsilon_t,\varepsilon_t\sim GED\left(0,1,v\right)$$

Where σ_t^2 follows the GJR-GARCH$(1,1)$ process

$$\sigma_t^2=\omega+y_{t-1}^2\left(\alpha_1+\gamma_1S_{t-1}^-\right)+\beta_1\sigma_{t-1}^2$$

where

$$S_t^-=\begin{cases}1 & \text{if } r_t<0\\ 0 & \text{if } r_t\geq 0\end{cases}$$

We go through the usual steps for weak stationarity (I stated them already earlier). We have that y_t is a martingale sequence, which mean that

$$\begin{aligned}E\left[y_t\mid\mathcal{F}_{t-1}\right]&=0\ \forall\ t\\ E\left[y_t\right]&=0\ \forall\ t\end{aligned}$$

Hence, the unconditional mean for $y_t$is 0 as in the case for $GARCH(1,1)$

We can unfold the process of σ_t^2 by inserting $y_y=\sigma_t\varepsilon_t$ to investigate the parameters that ensure convergence

$$
\begin{aligned}
\sigma_t^2 &= \omega + y_{t-1}^2\left(\alpha_1 + \gamma_1 S_{t-1}^-\right) + \beta_1\sigma_{t-1}^2 \\
&= \omega + \sigma_{t-1}^2\varepsilon_{t-1}^2\left(\alpha_1 + \gamma_1 S_{t-1}^-\right) + \beta_1\sigma_{t-1}^2 \\
&= \omega + \sigma_{t-1}^2\left(\varepsilon_{t-1}^2\left(\alpha_1 + \gamma_1 S_{t-1}^-\right) + \beta_1\right) \\
&= \omega + \left(\varepsilon_{t-1}^2\left(\alpha_1 + \gamma_1 S_{t-1}^-\right) + \beta_1\right)\left(\omega + \sigma_{t-2}^2\left(\varepsilon_{t-2}^2\left(\alpha_1 + \gamma_1 S_{t-1}^-\right) + \beta_1\right)\right) \\
&= \ldots \\
&= \omega\left[1 + \sum_{k=1}^{\infty}\prod_{i=1}^{k}\left(\beta_1 + \alpha_1\varepsilon_{t-i}^2 + \gamma_1 S_{t-1}^-\varepsilon_{t-i}^2\right)\right]
\end{aligned}
$$

If we define $b_t = \beta_1 + \alpha_1\varepsilon_t^2\gamma_1 + S_t^-\varepsilon_t^2$ to simplify, we can take the expectation

$$
\begin{aligned}
\mathbb{E}\left[b_t\right] &= \mathbb{E}\left[\beta_1 + \alpha_1\varepsilon_t^2 + \gamma_1 S_t^-\varepsilon_t^2\right] \\
&= \beta_1 + \alpha_1\underbrace{\mathbb{E}\left[\varepsilon_t^2\right]}_{=1} + \gamma_1\mathbb{E}\left[S_t^-\varepsilon_t^2\right]
\end{aligned}
$$

Notice $\mathbb{E}\left[S_t^-\varepsilon_t^2\right] = 0.5$, as the GED distribution is symmetrical.

$$
\begin{gathered}
\mathbb{E}\left[Z_t^-\varepsilon_t^2\right] = \mathbb{E}\left[\varepsilon_t^2 \mid \varepsilon_t < 0\right] * \mathbb{P}\left(\varepsilon_t < 0\right) = \frac{1}{2} \\
\Downarrow \\
\mathbb{E}\left[b_t\right] = \beta_1 + \alpha_1 + \frac{\gamma_1}{2}
\end{gathered}
$$

Now looking towards the unconditional variance, and then inserting our expression yields

$$
\begin{aligned}
Var(y_y) = \mathbb{E}\left[y_t^2\right] = \mathbb{E}\left[\sigma_t^2\right] &= \mathbb{E}\left[\omega\left[1 + \sum_{k=1}^{\infty}\prod_{i=1}^{k}\left(\beta_1 + \alpha_1\varepsilon_{t-i}^2 + \gamma_1 S_{t-1}^-\varepsilon_{t-i}^2\right)\right]\right] \\
&= \omega\left[1 + \sum_{k=1}^{\infty}\prod_{i=1}^{k}\mathbb{E}\left[b_{t-i}\right]\right] \\
&= \omega\left[1 + \sum_{k=1}^{\infty}\prod_{i=1}^{k}\beta_1 + \alpha_1 + \frac{\gamma_1}{2}\right] \\
&= \omega\left[1 + \sum_{k=1}^{\infty}\left(\beta_1 + \alpha_1 + \frac{\gamma_1}{2}\right)^k\right] \\
&= \omega\sum_{k=0}^{\infty}\left(\beta_1 + \alpha_1 + \frac{\gamma_1}{2}\right)^k \\
&= \frac{\omega}{1 - \beta_1 - \alpha_1 - \frac{\gamma_1}{2}}
\end{aligned}
$$

To have weak stationarity the following must hold: $E\left[\sigma_t^2\right] = \sigma^2 < \infty$ for all t. From this equation we immediately recognize that $0 < \beta_1 + \alpha_1 + \frac{\gamma_1}{2} < 1$ and $0 < \omega$ for the process to be covariance stationary and remain positive. The last requirement for weak stationarity is that the covariance has to be a function of the the distance between the returns, i.e.

$$\begin{aligned} cov\left(y_t, y_{t-s}\right) &= \mathbb{E}\left[y_t y_{t-s}\right] - \underbrace{\mathbb{E}\left[y_t\right]\mathbb{E}\left[y_{t-s}\right]}_{=0} \\ &= \mathbb{E}\left[\mathbb{E}\left[y_t y_{t-s} \mid \mathcal{F}_{t-1}\right]\right] \\ &= \mathbb{E}\left[y_{t-s}\underbrace{\mathbb{E}\left[y_t \mid \mathcal{F}_{t-1}\right]}_{=0}\right] = 0 \end{aligned}$$

Thereby we have shown that the process exhibits weak stationarity.
To have strong/strict stationarity for the process we would instead need (the arguments I use very similar to those I used for the GARCH(1,1) process)

$$\mathbb{E}\left[\ln\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right)\varepsilon_t^2\right)\right] < 0$$

By Jensen inequality we have

$$\mathbb{E}\left[\ln\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right)\varepsilon_t^2\right)\right] < \ln\mathbb{E}\left[\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right)\varepsilon_t^2\right] = \ln\left(\alpha_1 + \gamma S_{t-1}^- + \beta_1\right)$$

We define σ_t^2 as the process initlized at time $t = 0$ at σ_0^2 and ${}_u\sigma_t^2$ as the process that goes back to the infinite past, i.e. the unconditional variance. Then we have

$$\begin{aligned} \sigma_t^2 &= \sigma_0^2 \prod_{i=1}^{t}\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right)\varepsilon_t^2\right) + \omega\left[1 + \sum_{k=1}^{t-1}\prod_{i=1}^{k}\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right)\varepsilon_t^2\right)\right] \\ u^{\sigma_t} &= \omega\left[1 + \sum_{k=1}^{\infty}\prod_{i=1}^{k}\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right)\varepsilon_t^2\right)\right] \end{aligned}$$

To have stationarity at $\omega = 0$ and when $E\left[\ln\left(\beta + \left(\alpha + \gamma S_{t-1}^-\right)\varepsilon_t^2\right)\right] < 0$:
Proofs:
If $\omega = 0$ we have

$$\begin{aligned} {}_u\sigma_t^2 &= \omega\left[1 + \sum_{k=1}^{t-1}\prod_{i=1}^{k}\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right)\varepsilon_{t-i}^2\right)\right] \\ &= 0 \cdot \left[1 + \sum_{k=1}^{t-1}\prod_{i=1}^{k}\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right)\varepsilon_{t-i}^2\right)\right] \\ &= 0 \end{aligned}$$

and

$$
\begin{aligned}
\sigma_t^2 &= \sigma_0^2 \prod_{i=1}^{t} \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right) + \omega \left[1 + \sum_{k=1}^{t-1} \prod_{i=1}^{k} \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)\right] \\
&= \sigma_0^2 \prod_{i=1}^{t} \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right) \\
\Rightarrow \log \sigma_t^2 &= \log \sigma_0^2 + \sum_{i=1}^{t} \log \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)
\end{aligned}
$$

If we define

$$
\begin{aligned}
x_t &= \log \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right) \\
E\left[x_t\right] &= E\left[\log \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)\right] \\
&= \overline{x} \\
\widetilde{x}_t &= x_t - \overline{x} \\
\Rightarrow x_t &= \widetilde{x}_t + \overline{x} \\
E\left[\widetilde{x}_t\right] &= 0
\end{aligned}
$$

Using this fact and inserting again

$$
\begin{aligned}
\log \sigma_t^2 &= \log \sigma_0^2 + \sum_{i=1}^{t} \log \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right) \\
&= \log \sigma_0^2 + \sum_{i=1}^{t} \left(\widetilde{x}_t + \overline{x}\right) \\
&= \log \sigma_0^2 + t\overline{x} + \sum_{i=1}^{t} \widetilde{x}_t
\end{aligned}
$$

We see that we will end up with a random walk with a drift given by t.

We can see that if $t \to \infty$ we can get the following cases

$$
\begin{aligned}
\log \sigma_t^2 &\to +\infty \text{ if } \overline{x} > 0 \\
\log \sigma_t^2 &\to -\infty \text{ if } \overline{x} < 0
\end{aligned}
$$

Recall, that $\overline{x} = E\left[\log \left(\beta + \left(\alpha + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)\right]$.

We now turn to another theorem from Nelson (1990) where $\omega > 0$. Where we have two different cases of $\mathbb{E}\left[\log \left(\beta + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_t^2\right)\right]$.

First case: If we have $\omega > 0$ and $\mathbb{E}\left[\log \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_t^2\right)\right] \geq 0$ then (a) $\sigma_t^2 \to \infty$ $a.s.$ and (b) ${}_u\sigma_t^2 \to \infty$ $a.s.$ for all t. This mean that both processes diverges. We have

$$\sigma_t^2 = \sigma_0^2 \prod_{i=1}^{t} \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right) + \omega \left[1 + \sum_{k=1}^{t-1} \prod_{i=1}^{k} \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)\right]$$

We know the random variable σ_t^2 is the sum of the two terms and that these are both strictly positive. Therefore it most hold that σ_t^2 is greater or equal than the second term. Which is the sum of a number variables, instead we can bound it. We say that σ_t^2 is going to be greater than the largest of the sum, i.e. the supremum of $\prod_{i=1}^{k} \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)$ up to time t-1.

$$\begin{aligned}
\sigma_t^2 &\geq \omega \left[1 + \sum_{k=1}^{t-1} \prod_{i=1}^{k} \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)\right] \\
&\geq \omega \sup_{1 \leq k \leq t-1} \prod_{i=1}^{k} \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)
\end{aligned}$$

We now take the logarithm of both sides and then we proceed

$$\log\left(\sigma_t^2\right) \geq \log(\omega) + \sup_{1 \leq k \leq t-1} \sum_{i=1}^{k} \log\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)$$

Which is term similar to what we had in theorem 1 of Nelson (1990). Using the same arguments as before, we can show that

$$\begin{aligned}
\sum_{i=1}^{k} \log\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right) \to \infty \quad iff \quad \mathbb{E}\left[\log\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)\right] > 0 \\
\lim_{t \to \infty} \sup_{1 \leq k \leq t-1} \sum_{i=1}^{k} \log\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right) = \infty \quad iff \quad \mathbb{E}\left[\log\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)\right] = 0
\end{aligned}$$

The process therefore diverges to $+\infty$ when $\mathbb{E}\left[\log\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_t^2\right)\right] \geq 0$. Point b can be proved using similar arguments as it the same as the second term of the process initilized at time t. Second case: If we have $\omega > 0$ and $\mathbb{E}\left[\log\left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_t^2\right)\right] < 0$ then (c) $\frac{\omega}{1-\beta}\ {}_u\sigma_t^2 < \infty$for all t $a.s.$ and (d) ${}_u\sigma_t^2 - \sigma_t^2 \to 0\ a.s.$

We take the σ_t from the unconditional process and study its minimum (lower bound)

$${}_u\sigma_t^2 = \omega \left[1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} \left(\beta_1 + \left(\alpha_1 + \gamma S_{t-1}^-\right) \varepsilon_{t-i}^2\right)\right]$$

To do so we note that it's infimum is achieved for $\varepsilon_{t-1}^2 = 0$ for all t, so we have that

$$\begin{aligned}\inf \ {}_u\sigma_t^2 &= \omega\left[1+\sum_{k=1}^{\infty}\prod_{i=1}^{k}\beta\right]\\ &= \omega\left[1+\sum_{k=1}^{\infty}\beta^k\right]\\ &= \omega\sum_{k=1}^{\infty}\beta^k\\ &= \frac{\omega}{1-\beta}\end{aligned}$$

We make use of $0<\beta<1$ to make a sum rule of a geometric series. However this $\beta<1$ is different from $\alpha+\beta+0.5\gamma<1$. The upper bound relies on the fact that each element $\prod_{i=1}^{k}\left(\beta_1+\left(\alpha_1+\gamma S_{t-1}^-\right)\varepsilon_t^2\right)$ is of order $\exp(-\lambda k)$ for some constant $\lambda>0$. We can see that the process does not diverge when k goes to infinity, because $O(\exp(-\lambda k))$ goes to 0 very fast when k grows.

To prove point (d) we subtract ${}_u\sigma_t^2$ from σ_t^2, i.e.

$$\sigma_t^2 - {}_u\sigma_t^2 = \sigma_0^2\prod_{i=1}^{t}\left(\beta_1+\left(\alpha_1+\gamma S_{t-1}^-\right)\varepsilon_{t-i}^2\right)-\omega\left[1+\sum_{k=t}^{\infty}\prod_{i=1}^{k}\left(\beta_1+\left(\alpha_1+\gamma S_{t-1}^-\right)\varepsilon_{t-i}^2\right)\right]=0$$

We know the first term goes to zero when t goes to infinit provided $\mathbb{E}\left[\log\left(\beta_1+\left(\alpha_1+\gamma S_{t-1}^-\right)\varepsilon_t^2\right)\right]<0$ from theorem 1. Furtermore it can be shown that the second term also goes to zero because, as in the case of the upper bound, each element $\prod_{i=1}^{k}\left(\beta_1+\left(\alpha_1+\gamma S_{t-1}^-\right)\varepsilon_{t-i}^2\right)$ is of order $exp(-\lambda k)$, for some $\lambda>0$, and therefore it does not diverge as k goes to infinity.

Estimate these three models under the GED assumption for the innovations. You cannot use the rugarch package.

I just used the derived values for the unconditional variance in last erercise for the different models. Furthermore i used the Log likelihood function from the GED distribution given in exercise 1. Code can be found in the R script: ExamScript20.

Parameter estimates for the S&P500 dataset:

S&P500	ω	α	β	υ	ϕ	ψ	γ
GARCH(1,1)	0.01807	0.11785	0.86825	1.20855			
EGARCH(1,1)	-0.002571		0.968414	1.263433	-0.189357	0.167157	
GJR-GARCH(1,1)	0.0242447	0.0001008	0.8648455	1.2671180			0.2332405

Table 1: Parameter estimates (S&P500)

Parameter estimates for the DJI dataset:

DJI	ω	α	β	υ	ϕ	ψ	γ
GARCH(1,1)	0.01877	0.12414	0.85924	1.25344			
EGARCH(1,1)	-0.005564		0.964439	1.311472	-0.171081	0.183462	
GJR-GARCH(1,1)	0.0237581	0.0001001	0.8647457	1.3197139			0.2287202

Table 2: Parameter estimates (DJI)

Compare the filtered volatilities according to these three specifications in a figure.

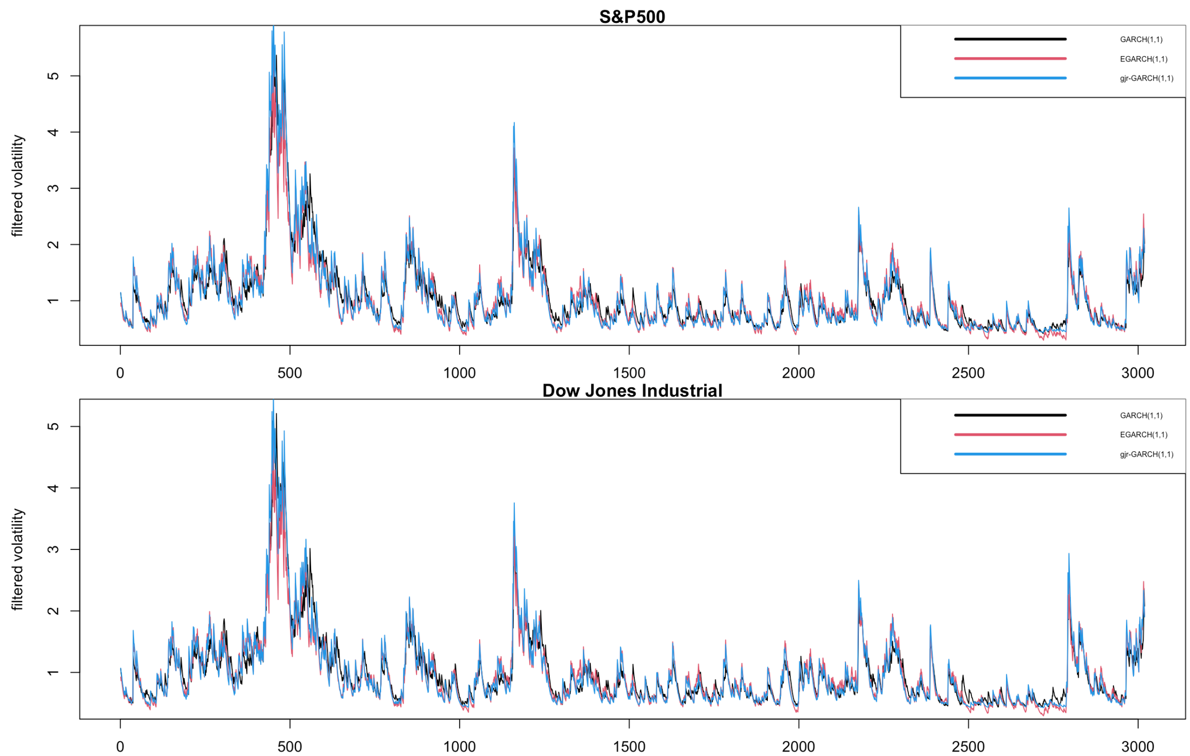

Figure 1: Figure of the filterered volatilities for the three model specifications

From figure 1 we notice that the filtered volatilites are quite similar for the three model specifications. We notice that the GJR-GARCH(1,1) model reacts more strongly to extreme observations. This is in line, with the fact that υ is higher for GJR-GARCH(1,1) in both data sets, and as we discovered in part c of the theoretical part, a higher υ implies a larger reaction to extreme observations as the distibution has thinner tails.

Select the best model using the BIC criteria.

BIC	S&P500	DJI
GARCH(1,1)	2.658917	2.557497
EGARCH(1,1)	2.617663	2.520695
GJR-GARCH(1,1)	2.623025	2.521437

Table 3: BIC for the three model specifications

From the BIC we would choose EGARCH model, as it has the lowest BIC for both data sets. Notice I have used the average BIC, but that should affect the result, since the model has the same number of observations. We also notice that the three models are all quite close in terms of BIC.

3.2 (B) Estimate the GAS-GED model on the DJI and SP500 series.

See the R code. File name: "ExamScript20"

Parameter estimates using the GAS-GED model

GAS GED	ω	α	β	υ
S&P500	-0.01299366	0.07353581	0.98267801	1.19310866
DJI	-0.01433778	0.07378180	0.98082733	1.23046056

Table 4: Parameter estimates using the GAS-GED model

Compare the filtered volatility of the GAS-GED model with that of the GARCH(1,1) model in a figure.

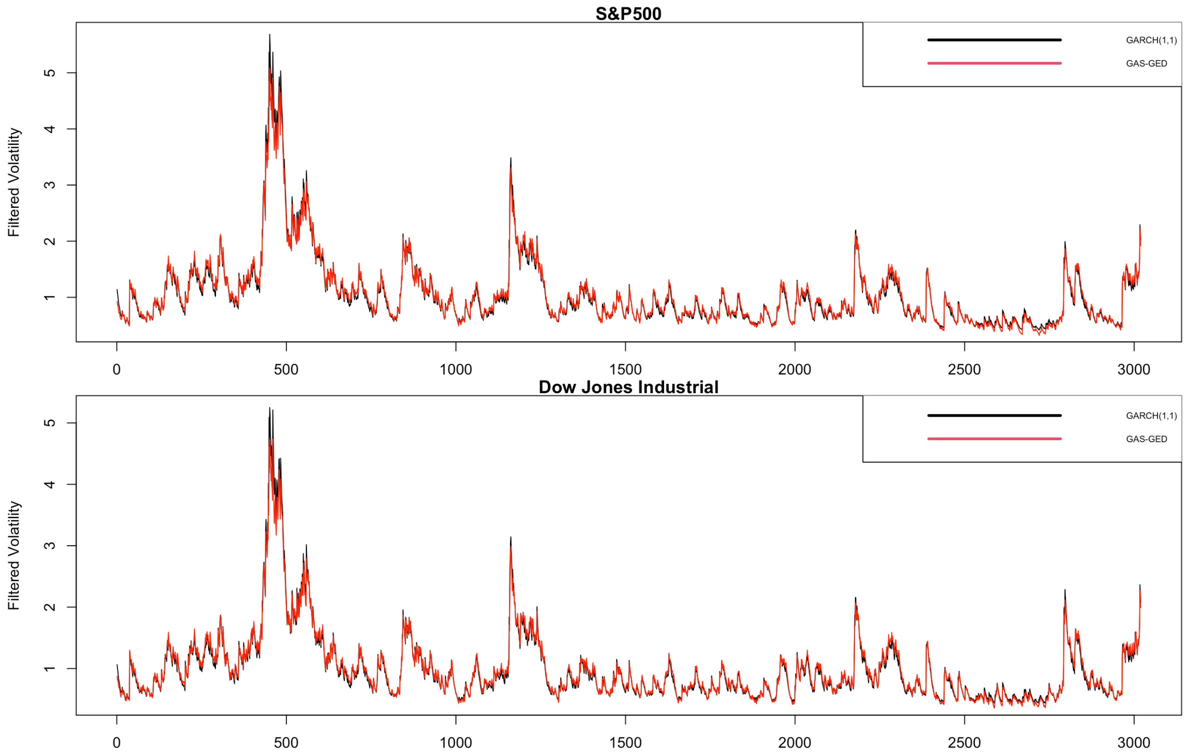

Figure 2: Figure of the filterered volatilities for GARCH(1,1) and GAS-GED on the two data sets

We notice once again, that the two graphs are quite similar. However we notice some differences, especially in periods with high volatility and extreme oberservations. If we compare the υ from the two models (using DJI data) we get $\upsilon_{GARCH} = 1.25344$ and $\upsilon_{GAS-GED} = 1.23046056$. We see that the thickness parameter is larger when using the GARCH(1,1) model specification. This is evident in this graph aswell. If we look at time $= 500$, which is right at the peak of the financial crisis, we see that the filtered volatility for the GARCH model reacts more to the extreme observations, that we saw through the financial crisis. Remember, a higher υmeant a distribution with thinner tails, and more strong reaction to these extreme observations.

Compare VaR at levels $\alpha = 1\%$ and $\alpha = 5\%$ estimated with the GAS-GED and GARCH(1,1) models in a figure.

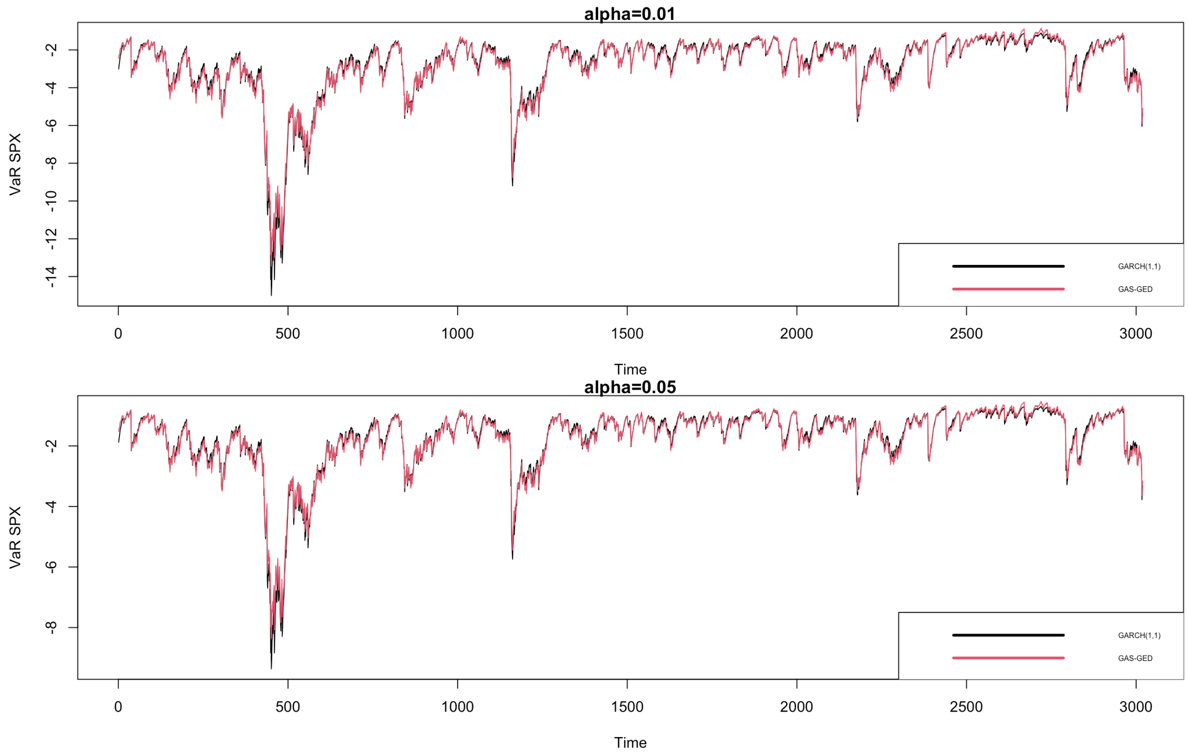

Figure 3: S&P500:VaR at levels $\alpha = 0.05$ and $\alpha = 0.01$

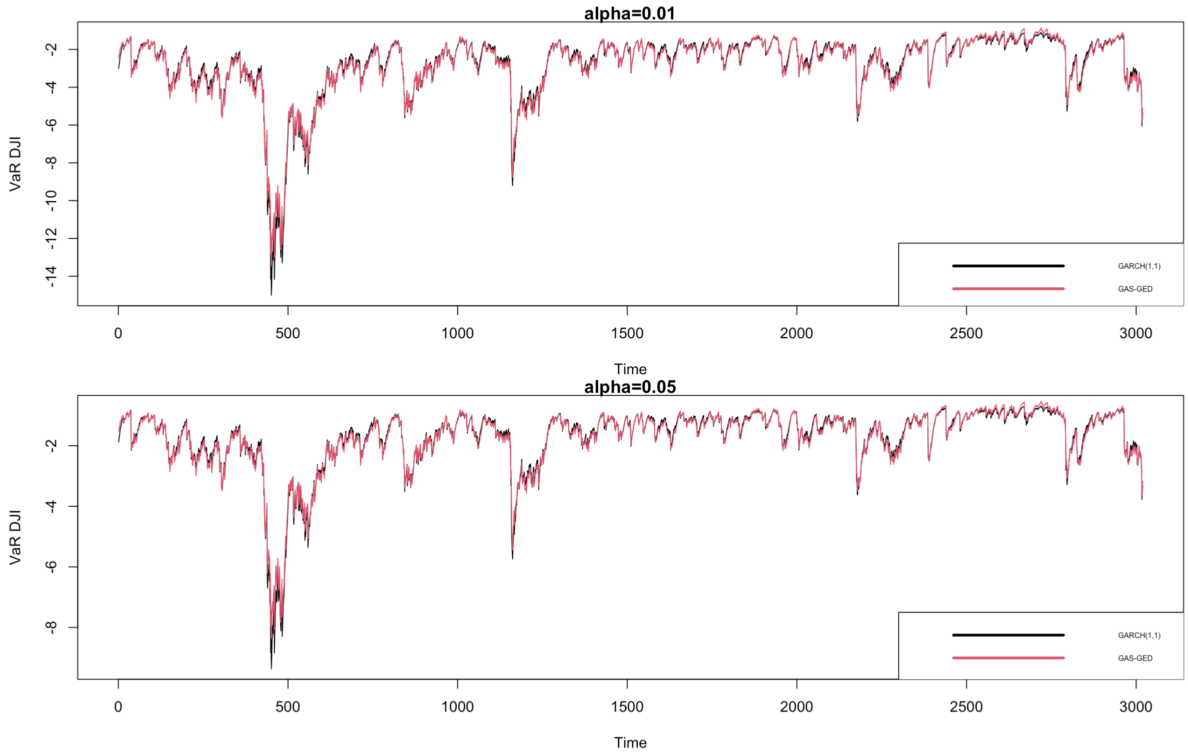

Figure 4: DJI:VaR at levels $\alpha = 0.05$ and $\alpha = 0.01$

Again, we see that the two model specifications perform quite similarly on the datasets. However, we see that minimum potential loss that the portfolio can suffer, for both confidence coefficients, is worse in periods with extreme observations, when using GARCH(1,1) compared to GAS-GED. This is perfectly in line with what we just found out, ie. that the GARCH(1,1) model has a stronger reaction to extreme observations compared to the GAS-GED model, at least in these data sets.

Select the best model between GAS-GED and GARCH(1,1) using BIC

BIC	S&P500	DJI
GARCH(1,1)	2.658917	2.557497
GAS-GED	2.661908	2.559992

Table 5: BIC for GARCH(1,1) and GAS-GED

We will choose GARCH(1,1) from a BIC perspective as it has the lowest BIC for both datasets.

To draw some general conlusions about the differences between the GAS-GED and GARCH(1,1) models, we can say that GARCH(1,1) reacts more to extreme observations as stated a bunch of times earlier. This is also evident when looking at the graps for the filtered volatility, and also for the graphs of the Value-at-Risk. The thickness parameter is estimated to be larger when using the GARCH(1,1) models and therefore its has thinner tails than the GAS-GED specification, it will therefore react more to extreme observations, which can especially be seen at the time around the financial crisis.

3.3 (C) Covariance and portfolio weights

See the R code. File name: "ExamScript20"

Compute Σ_t for each t, when DJI and SP500 both follow the GAS-GED model

Function and array can be found in the code. Would not make sense to report it here.

Compute Σ_t for each t, when DJI and SP500 both follow the GARCH(1,1) model.

Array can be found in the code.

Compute the weight ω_t associated to the Minimum Variance Portfolio (MVP) constructed using the DJI and SP500 returns at each point in time, i.e. $y_t = \omega_t y_t^{DJI} + (1-\omega_t)\, y_t^{SP500}$. |

Matrices with weights can be found in the code.

Compare the portfolio weights of the two models ω_t in a figure.

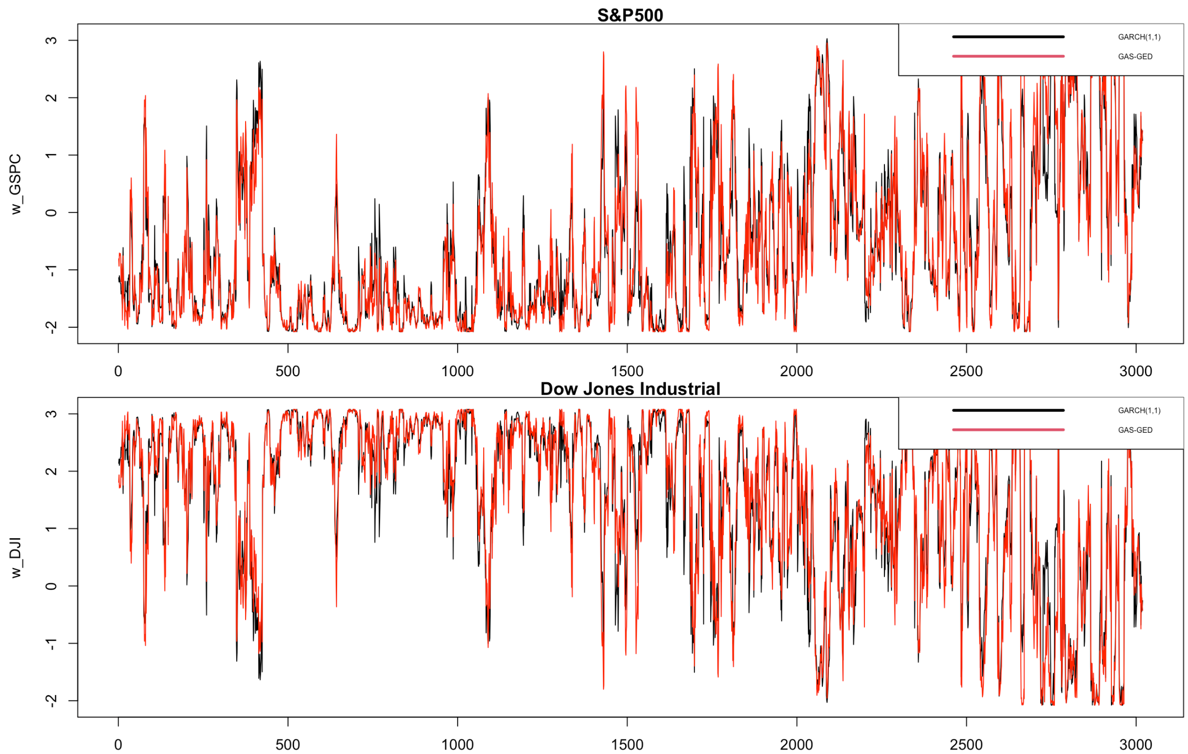

Figure 5: Portfolio weights for the two models

The portfolio weights of the two models are quite similar. We can see the weights fluctuate a lot using both model specifications. This could be due to the high correlation between the datasets. It makes sense that two indices would be correlated a lot with general market fluctuactios.