GROUP WORK PROJECT # 2 MScFE 622: Stochastic Modeling

GROUP NUMBER: 4839

FULL LEGAL NAME	LOCATION (COUNTRY)	EMAIL ADDRESS	MARK X FOR ANY NON-CONTRIBUTING MEMBER
Marin Stoyanov	Bulgaria	azonealerts@gmx.com	
Panagiotis Lois	Grenoble (France)	loispanagiotis@outlook.com	

Statement of integrity: By typing the names of all group members in the text boxes below, you confirm that the assignment submitted is original work produced by the group (excluding any non-contributing members identified with an "X" above).

Team member 1	Marin Stoyanov
Team member 2	Panagiotis Lois
Team member 3	

Use the box below to explain any attempts to reach out to a non-contributing member. Type (N/A) if all members contributed.
Note: You may be required to provide proof of your outreach to non-contributing members upon request.

N/A

Group Number: 4839

Step 1

We have collected financial time series of the following asset classes: i) Currencies (EUR/USD, GBP/JPY, CAD/USD,), ii) Crypto-currencies (BTC/USD, ETH/USD), and indices (SP500, NASDAQ). The time span is the same for all of them, from the start of 2019 till the end of 2022. Since we are more interested in the deviation of the asset over time, we converted the financial time series into cross-sectional dispersion in daily returns which is more convenient for the risk assessment of the model.

Step 2

Question a)

In order to get a better fit for the model we chose to use the indices asset class, and in particular the instrument SP500 (ticker symbol: `^GSPC`). We chose SP500 because it is a broader index of the 500 most important companies in the US economy and is used as a benchmark to the US economy as a whole. If we compare it with NASDAQ where the biggest tech companies are situated it will bring general information for the IT sector, while SP500 will bring information for all the 12 sectors that it consists of, and the information will be more valuable towards the entire economy rather than just the IT sector.

The returns of this instrument for the given period looks like this:

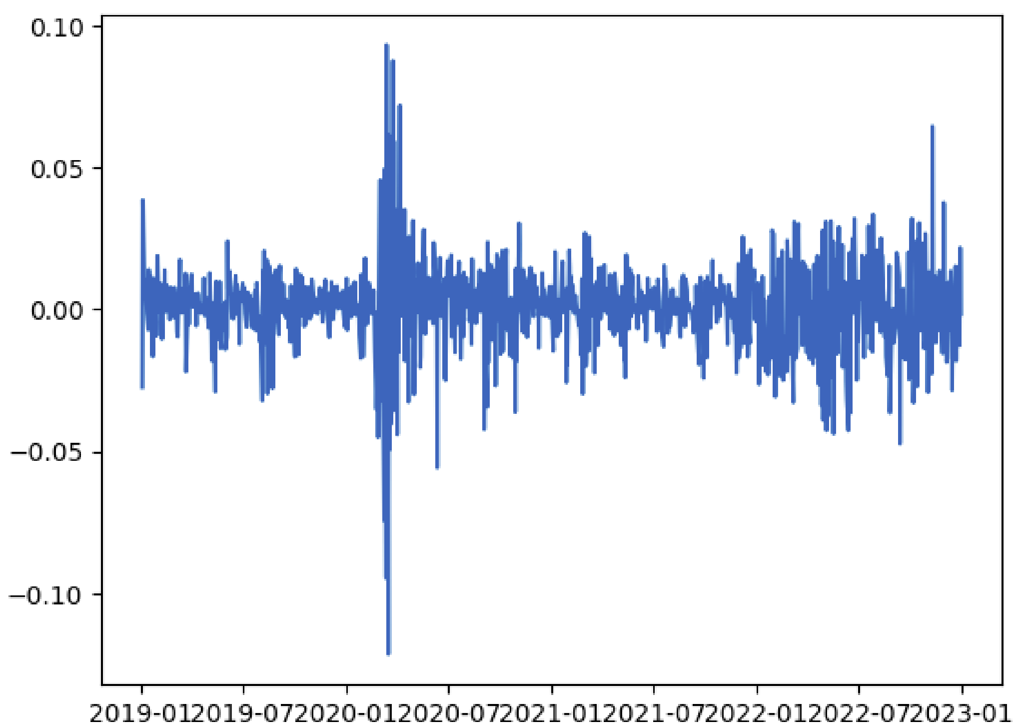

Fig.1. SP500 returns from the start of 2019 till the end of 2022

It is clearly visible that there are 2 enormous spikes in the beginning of 2020 around March which is exactly when the COVID-19 crisis began.

In figure 2 we are plotting the returns of SP500 and NASDAQ overlaid onto each other so that we could see the differences.

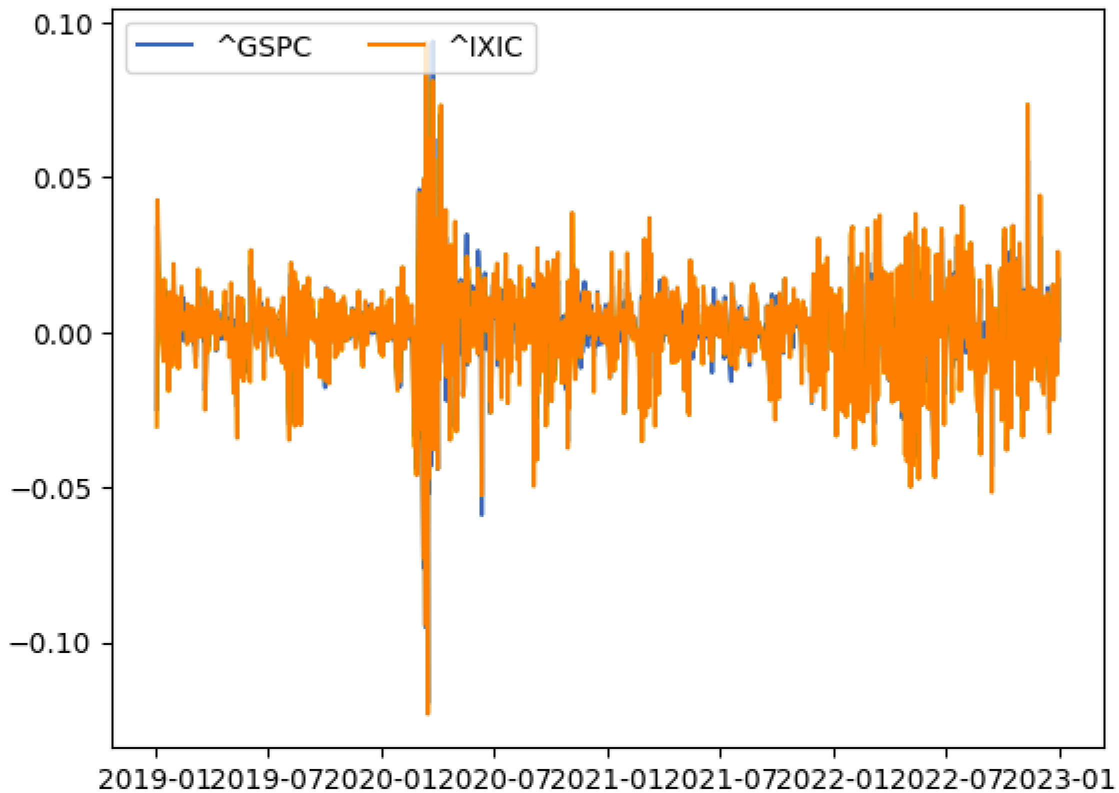

Fig.2. SP500 and NASDAQ returns comparison

As we can see the volatility of the returns of NASDAQ is bigger that the one of SP500 and this could be explained with the fact that the IT sector is the first to hit the lows when a crisis begins and so the selloff was way bigger there compared to SP500. This is quite normal and expected since nasdaq is presenting the IT sector, while SP500 has inside different sectors like: Health Care, Financials, Consumer Discretionary, Communication Services, Industrials, Consumer Staples, Energy and more, which makes it much more diversified and crisis resilient, because some sectors will fall other will grom (or at least stay flat) and this makes the volatility of the returns much less severe.

Since our aim is to model a Hidden Markov model, we will use the dispersion of daily returns as it can make an investment decision more informative, taking into consideration the asset's historical volatility.

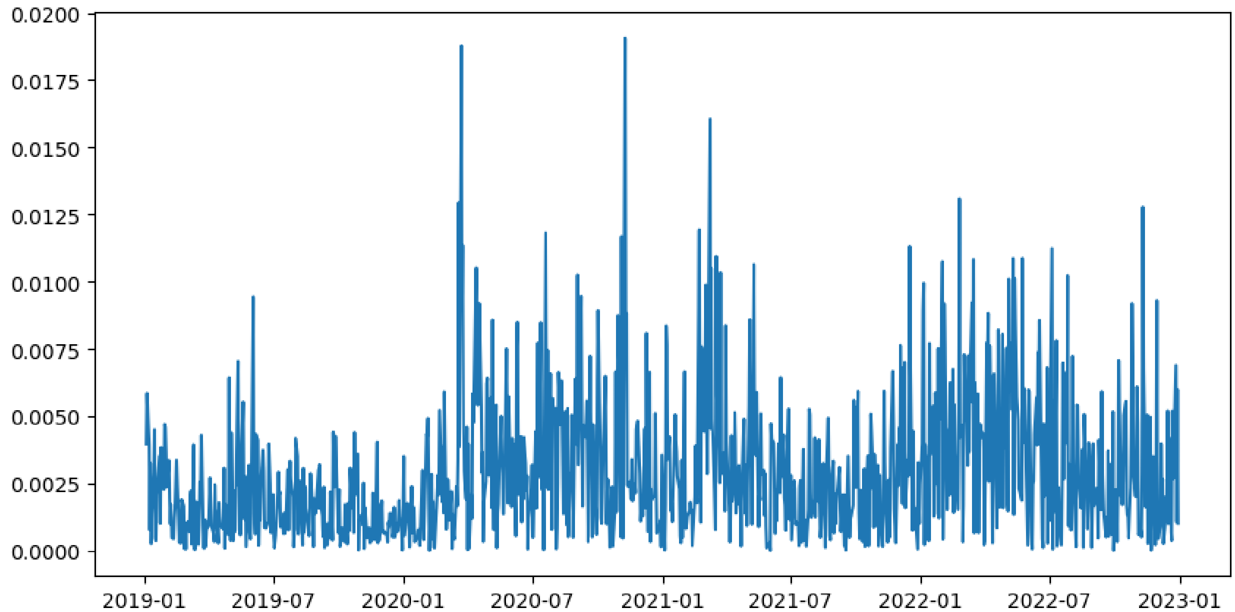

Fig. 3. Dispersion of daily cryptocurrencies returns

As we notice, there are some periods where the dispersion of returns is higher than average. High dispersion gives a sign that the time series is unpredictable and difficult to consult. The highest spikes are located in the second, and last quarter of 2020.

Question b)

i) Different number of states N

First, we compute the model for a different number of states. This is simple to implement as we only need to set the number of states to the desired amount. The initial conditions of the model, (mu, sigma, and the transition matrix) have to be set accordingly. For example, for 4 states, we would create a vector of 4x4 that addresses the 4 states.

ii) Different expectation across states (different "mus") but constant variance (same "sigma")

For constant volatility between time periods we define the parameter sigma as common between all steps such as:

$$\frac{\partial \sum_{i=1}^{N}\sum_{t=1}^{T}\xi_{t|T}^{(k-1)}(i)\, log\phi_{it}}{\partial\sigma} = \sum_{i=1}^{N}\sum_{t=1}^{T}\xi_{t|T}^{(k-1)}(i)\left(\frac{(y_t-\mu_i)^2}{\sigma^2}-1\right) = 0$$

So our new guess for the parameter sigma is given by:

$$\sigma^{(k)} = \sqrt{\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\xi_{t|T}^{(k-1)}(i)(y_t-\mu_i^k)^2}{T}}$$

For expectation across states that differs, we calculate the weighted average estimated probability that the process is in state i at each date t:

$$\mu_i^k = \frac{\sum_{t=1}^{T}\xi_{t|T}^{(k-1)}(i)yt}{\sum_{i=1}^{M}\xi_{t|T}^{(k-1)}(i))}$$

Below we see the estimated probabilities of the two states having different "mus" but same sigma.

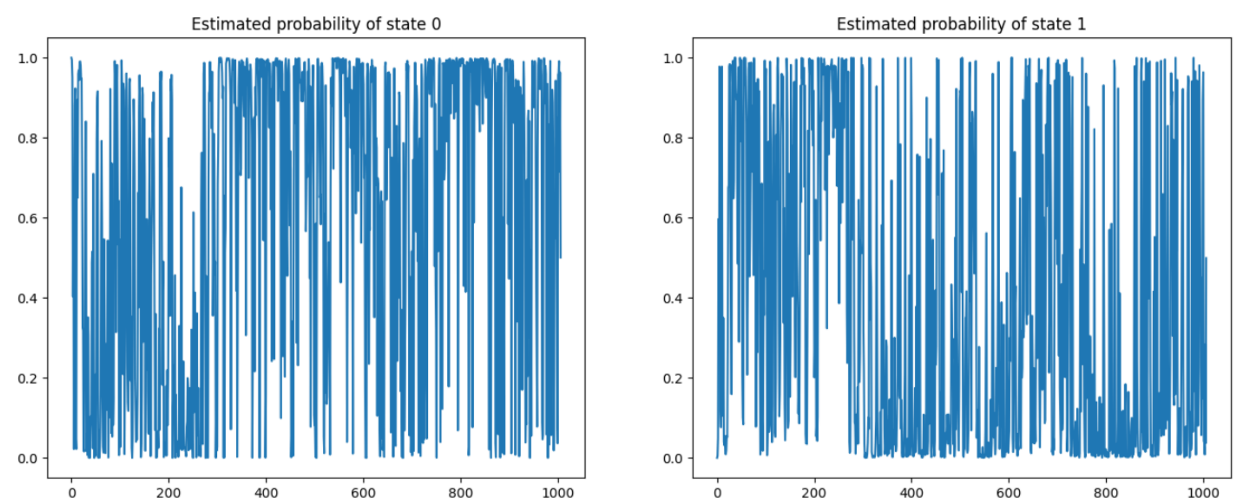

Fig.4 Model with same sigmas and different mus

iii) Different variance across states (different "sigma") but constant expectation (same "mus")

For constant expectation across states we compute "mus" as the same average across states. So the new "mu" can be given by:

$$\mu^{k} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} \xi_{t|T}^{(k-1)}(i) yt / \left(\sigma_{i}^{k}\right)^{2}}{\sum_{i=1}^{N}\sum_{t=1}^{T} \xi_{t|T}^{(k-1)}(i)}$$

And the new estimates for each σ_{i}^{k} should be:

$$\sigma_{i}^{k} = \sqrt{\frac{\sum_{t=1}^{T} \xi_{t|T}^{(k-1)}(i)\left(y_{t} - \mu_{i}^{k}\right)^{2}}{\sum_{t=1}^{T} \xi_{t|T}^{(k-1)}(i)}}$$

Below we see the estimated probabilities of the two states having different sigma but same “mus”.

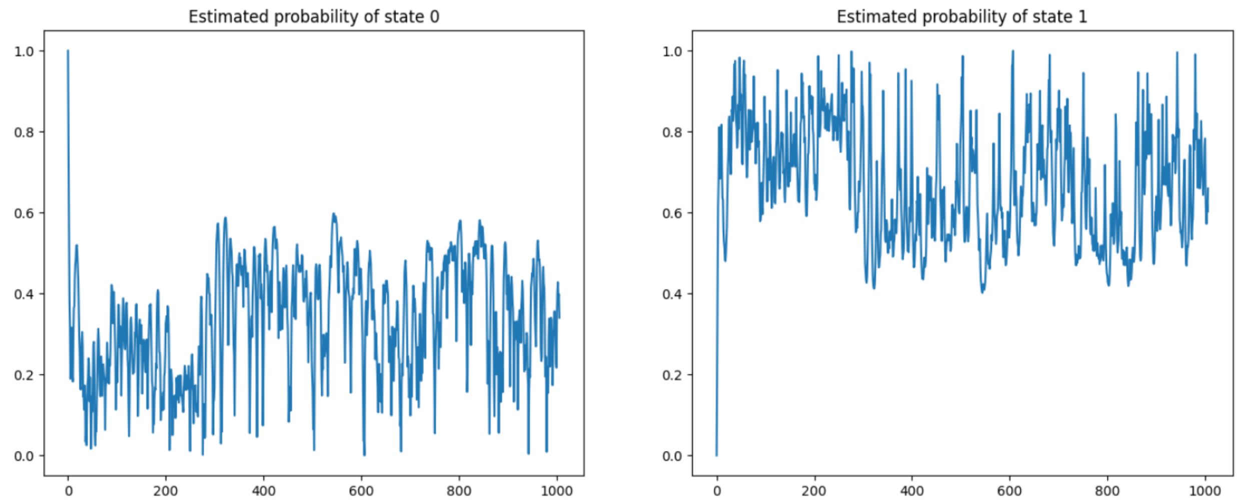

Fig.5 Model with same mus and different sigmas

iv) Different variance and different expectation across states (different “sigma”, “mus”)

Lastly, it is required to estimate a model that allows for different expectations and different variances across states. In this case, the parameters mu, and the variance sigma should vary in each state. Given that, we can combine the functions above for each parameter of the model:

For μ_{i}^{k}:

$$\mu_{i}^{k} = \frac{\sum_{t=1}^{T} \xi_{t|T}^{(k-1)}(i) yt}{\sum_{i=1}^{M} \xi_{t|T}^{(k-1)}(i))}$$

For variance σ_i^k: $$\sigma_i^k = \sqrt{\frac{\sum_{t=1}^{T} \xi_{t|T}^{(k-1)}(i)(y_t - \mu_i^k)^2}{\sum_{t=1}^{T} \xi_{t|T}^{(k-1)}(i)}}$$

Step 3

The Akaike information criterion (AIC) is used as an estimator of prediction error for statistical models like time series.[1][2][3] To compare the performance of different models we can look up to each model's value of AIC. The lower the AIC value is, the better, since every statistical model is used for estimation of a real world event and it will never be exact. There will always be some data lost and AIC is used for estimating the amount of information lost. The less information is lost, the more precise is the model.

The formula for AIC is the following [4]:

$$AIC = 2k - 2ln(\hat{L})$$

Where:

$k \text{ is the number of estimated parameters in the model}$

$\hat{L} \text{ is the maximized value of the likelihood function for the model}$

Very often AIC is combined with the Schwarz Information Criterion (SIC) which is also called the Bayesian Information Criterion (BIC). Again lower values of BIC are preferred for model selection and it is closely related to AIC. The logic behind BIC is that when creating a model it is common sense to add more variables in order the model to be as much detailed and precise as possible, but this could lead to the so-called model overfitting. Both AIC and BIC are trying to prevent this from happening by introducing a penalty term for the number of parameters and the BIC's one is larger than the AIC's. [5]

The BIC's formula looks like this [6]:

$$BIC = kln(n) - 2ln(\hat{L})$$

Where:

$\hat{L} \text{ is the maximized value of the likelihood function of the model}$

$n \text{ is the number of data points (observations)}$

$k \text{ is the number of parameters estimated by the model}$

Group Number: 4839

In order to compare these models we will present the values of parameters to the corresponding model.

For the model with same sigmas and different mus across states (model ii):

```
Iteration:  3
Log-Likelihood:  -1638.8315 Change:  -0.0014
Final Estimates
Log-Likelihood:  -1638.8315 Akaike:  3295.663 Schwarz:  3339.8956
Mu:  [-1.0792 -2.6267]
Sigma:  0.7696
Transition matrix:
[0.7683 0.2317]
[0.3268 0.6732]
Initial probabilities: [9.999e-01 1.000e-04]
```

Fig. 6. Parameters' values for the model with same sigmas different mus

For the model with same mus and different sigmas (model iii):

```
Iteration:  9
Log-Likelihood:  -2171.5071 Change:  0.0001
Final Estimates
Log-Likelihood:  -2171.5071 Akaike:  4361.0141 Schwarz:  4405.2467
Mu:  0.0
Sigma:  [1.6007 2.2998]
Transition matrix:
[0.6903 0.3097]
[0.1527 0.8473]
Initial probabilities: [9.998e-01 2.000e-04]
```

Fig.7. Parameters' values for the model with same mus different sigmas

The first thing that we see is that for model ii), it took a lot less iterations (3 of them) for the convergence to appear while for model iii), there were 9 iterations, so we can say that model ii) is taking less computational time.

Comparing the AICs and BICs of these 2 models, we see that the values obtained from the model ii) with the same sigma but different "mus" are lower. AIC 3295 versus 4361, BIC 3339 versus 4405. Therefore, this model should be preferred, considering the information criterion of AIC and BIC.

Model ii) has a better balance between fit and complexity (AIC) while at the same time it avoids overfitting (BIC). Moreover, we proved that it is cost efficient (computing time). So this makes the first model much more simple, robust and efficient than the later model.

Step 4

The state-dependent autoregressive model (AR) is part of the so-called regime switching models, which allow different structures and parameters for different regimes. Regime switching models can capture various features of time series data, such as nonlinearities, asymmetries, heteroskedasticity, and persistence [7], which is exactly the case with our dataset, where due to a black swan event (covid-19), the regime switched from up trend to free falling on every market because of mass investor's fear, caused by uncertainty and panic over all sectors and investment instruments.

A state-dependent autoregressive model can be defined like this:

$$y_t = \phi_{s_t} y_{t-1} + \epsilon_t$$

Where:

y_t *is the observed time series*

ϕ_{s_t} *is the autoregressive coefficient in state* s_t

ϵ_t *is the error term with variance* μ_{s_t} $\sigma^2_{s_t}$

s_t *is the state variable and it is assumed to follow a Markov chain with transition probabilities*:

$$p_{ij} = P(s_t = j \mid s_{t-1} = i) \quad for\ i, j = 1, \ldots, K$$

For K : *number of states*

For an Auto-Regressive model of first order, we consider that the dispersion of the indices return will depend ϕ_{s_t} times the previous realization of the dispersion of returns plus a randomness.

To estimate the model we would have to calculate again the functions for likelihood, and Hamilton filtering as we need to account for t-1 periods. Estimating the mus and sigma would be similar to step 2 models. In this case we would choose a model with variable mu and sigma across states.

Finally, we obtain:

The estimated probabilities of the two states in the final iteration:

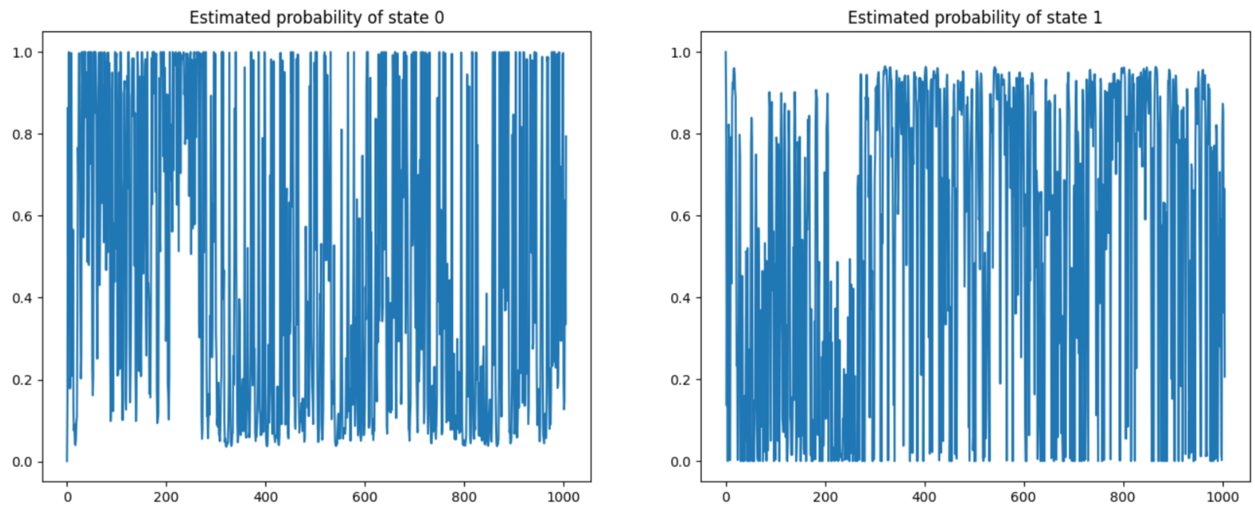

Fig.8. Estimated probabilities of the AR(1) model

And the model estimates:

```
Iteration:  6
Diff:  -0.0022263672231471856
Log-Likelihood:  -1518.414 Change:  -0.0022
Final Estimates
Log-Likelihood:  -1518.414 Akaike:  3060.8281 Schwarz:  3119.7929
Mu:  [-2.5352 -0.914 ]
Sigma:  [1.2735 0.576 ]
Transition matrix:
[0.7227 0.2773]
[0.2932 0.7068]
Initial probabilities: [0.0061 0.9939]
```

Fig.9. Estimated parameters of AR(1)

We conclude that the autoregressive model provides a better fit, given lower AIC, and BIC, obtained estimates. The mu ranges from [-2.5, -0.9] and the sigma [1.2, 0.6].

REFERENCES

1. Stoica, P.; Selen, Y. (2004), "Model-order selection: a review of information criterion rules", IEEE Signal Processing Magazine (July): 36–47, doi:10.1109/MSP.2004.1311138, S2CID 17338979
2. McElreath, Richard (2016). Statistical Rethinking: A Bayesian Course with Examples in R and Stan. CRC Press. p. 189. ISBN 978-1-4822-5344-3. AIC provides a surprisingly simple estimate of the average out-of-sample deviance.
3. Taddy, Matt (2019). Business Data Science: Combining Machine Learning and Economics to Optimize, Automate, and Accelerate Business Decisions. New York: McGraw-Hill. p. 90. ISBN 978-1-260-45277
4. Akaike, H. (1974), "A new look at the statistical model identification", IEEE Transactions on Automatic Control, 19 (6): 716–723. doi:10.1109/TAC.1974.1100705, MR 0423716.
5. Wit, Ernst; Edwin van den Heuvel; Jan-Willem Romeyn (2012). "'All models are wrong...': an introduction to model uncertainty". Statistica Neerlandica. 66 (3): 217–236. doi:10.1111/j.1467-9574.2012.00530.x. S2CID 7793470
6. Fabio Gobbi, Sabrina Mulinacci; (2020). State-Dependent Autoregressive Models: Properties, Estimation and Forecasting;