Skoltech

Stochastic Modelling course

Multivariate stable distributions and their applications for modelling cryptocurrency-returns

Student: Bogdan Alexandrov

Moscow 2023

Contents

1. Preliminary Theory

Many banks and funds are engaged in the task of hedging, that is, distributing capital across various assets in order to reduce the risk of its depreciation. The data they use is often daily ohlcv data collected over several years. And even under such conditions, not much data is collected, only 3-5 thousand observations. Therefore, simulation of plausible data is an attractive area for funds. There are many different approaches to this problem, but the authors of the article suggested generating cryptocurrency logs using stable distributions.

Definition 1 *Distribution $p(*)$ is called stable, if $\forall X_1, X_2, X \sim p(*), \forall a, b, c, d \in \boldsymbol{R}^+$ following condition holds: $aX_1 + bX_2 \overset{d}{=} cX + d$*

Stable distributions described by characteristic function:

$$\phi(t) = \begin{cases} \exp\{-\gamma^\alpha|t|^\alpha * (1 + i\beta \tan\frac{\pi\alpha}{2} * signt)(|\gamma t|^{1-\alpha} - 1) + i\delta t\}, \alpha \neq 1 \\ \exp\{-\gamma|t| * (1 + i\beta\frac{2}{\pi} * sign(t) * \log\gamma|t|) + i\delta t\}, \alpha = 1 \end{cases}$$

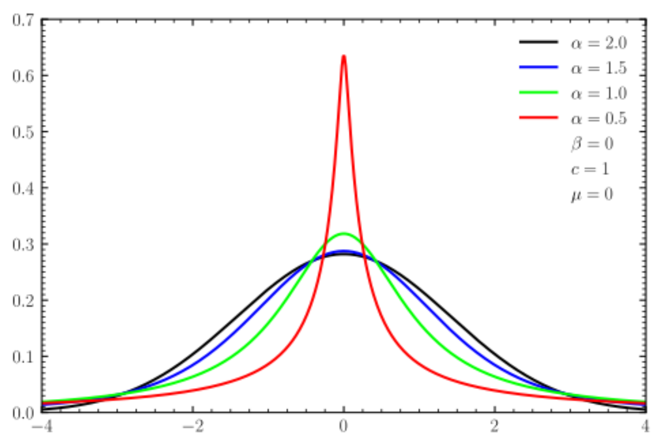


Figure 1: Stable distributions

Authors show that in multivariate case stable distribution is non parametric.

Theorem 1 *Let Λ be a finite measure on S_d, where $S_d = \{s \in R^d : \|s\|_2 = 1\}$, the surface of the unit ball. This measure is called the spectral measure. The d-dimensional variable X is stable, denoted by $X \sim S(\alpha, \Lambda, \delta)$, where $0 < \alpha \leq 2$ and $\delta \in R^d$, if and only if its characteristic function is*

$$\phi_X(t) = e^{-I_X(t) + it^T\delta}$$

, *where*

$$I_X(t) = \int_{S_d} \psi(t^T s; \alpha) \Lambda(ds)$$

and

$$\psi(u;\alpha) = \begin{cases} |u|^{\alpha}(1 - i \tan \frac{\pi\alpha}{2} * sign(u)), & \alpha \neq 1 \\ |u|(1 + i\frac{2}{\pi} sign(u) * \log|u|), & \alpha = 1 \end{cases}$$

1.3. Theorem. *Let be* $\mathbf{u}^\mathsf{T}\mathbf{X} \sim S(\alpha, \beta(\mathbf{u}), \gamma(\mathbf{u}), \delta(\mathbf{u}))$. *Then the parameter functions determining* $\mathbf{X}$ *can be written in the following form:*

$$\gamma(\mathbf{u}) = \left(\int_{S_d} |\mathbf{u}^\mathsf{T}\mathbf{s}|^{\alpha} \Lambda(d\mathbf{s}) \right)^{1/\alpha} \tag{2}$$

$$\beta(\mathbf{u}) = \gamma(\mathbf{u})^{-\alpha} \int_{S_d} |\mathbf{u}^\mathsf{T}\mathbf{s}|^{\alpha} \operatorname{sign}(\mathbf{u}^\mathsf{T}\mathbf{s}) \Lambda(d\mathbf{s}) \tag{3}$$

$$\delta(\mathbf{u}) = \begin{cases} \mathbf{u}^\mathsf{T}\boldsymbol{\delta} & \alpha \neq 1 \\ \mathbf{u}^\mathsf{T}\boldsymbol{\delta} - \frac{2}{\pi} \int_{S_d} \mathbf{u}^\mathsf{T}\mathbf{s} \cdot \log\big(|\mathbf{u}^\mathsf{T}\mathbf{s}|\big) \Lambda(d\mathbf{s}) & \alpha = 1. \end{cases} \tag{4}$$

Using these, $I_{\mathbf{X}}(\mathbf{t})$ *can be written as*

$$I_{\mathbf{X}}(\mathbf{t}) = \begin{cases} \gamma^{\alpha}(\mathbf{t})(1 - i\beta(\mathbf{t}) \tan \frac{\pi\alpha}{2}) & \alpha \neq 1 \\ \gamma(\mathbf{t})(1 - i\delta(\mathbf{t})) & \alpha = 1. \end{cases} \tag{5}$$

Theorem above is important for parameters estimation. Also, since Λ measure is continuous and we want to estimate it, authors propose to make it discrete:

$$\Lambda(*) = \sum_{i=1}^{n} \lambda_i \delta_{s_i}(*)$$

So the weights will exist only on points s_i.

2. Parameter estimation procedure

Suppose we work in 2-dimensional case. Firstly, let's assume, that $X_1, X_2, \dots X_m$ - bivariate samples from stable distribution. Λ is discrete and concentrated on n points.

1. Step 1

   Firstly, let's eliminate δ vector, since it's just shifting circle on coordinate plane. Marginal elements of vector can be estimated from marginal distributions using quantile methods [1] or MLE method.

2. Step 2

   Take n points equal to $s_j = (\cos \frac{2*\pi(j-1)}{n}, \sin \frac{2*\pi(j-1)}{n})$, $j = 1, \dots, n$. These are supports of density. Then we need to determine values from unit sphere for characteristic functions $t_1, ..., t_n \in S_2$. Authors proposed to make them equal to s: $t_j = s_j$. After that it is needed to calculate projections $< t_j, X_1 >, < t_j, X_2 >, ..., < t_j, X_m >$ - all set of observation with every projection point.

3. Step 3

   For every projected group it is needed to estimate parameters γ and β. Also, we estimate $\alpha(i)$ parameter for every group, and since we need one parameter, we just take mean of them: $\alpha^* = \frac{\sum_{i=1}^{n} \alpha(i)}{n}$. With estimated α, β and γ we calculate $I_X(t)$.

4. Step 4

   Since $I_X(t_i) = \sum_{j=1}^{n} \psi(t^T s_j; \alpha^*)\lambda_j$, we can write

   $$\Psi\lambda = I_X$$

   , where $I_{Xi} = I_X(t_i)$ and $\Psi_{ij} = \psi(t_i^T s_j; \alpha^*)$. To solve this equation it is needed to calculate inverse matrix of Ψ. But in such case λ vector can be imaginary. To solve this problem authors proposed to solve following problem [2]:

   $$\min_{\lambda} \|c - A\lambda\|^2, \ \lambda \geq 0$$

   , where $c = [ReI_1, ReI_2, \dots ReI_k, ImI_1, \dots, ImI_k]$, $n = 2k$ and,

   $$a_{ij} = \begin{cases} Re\psi_{ij}, 1 \leq i \leq k \\ Im\psi_{ij}, k+1 \leq i \leq n \end{cases}$$

3. Data

Data was taken from kaggle.com for Bitcoin, Litecoin and XRP(Ripple) dayly historical prices from April of 2013 to February of 2018. Price charts:

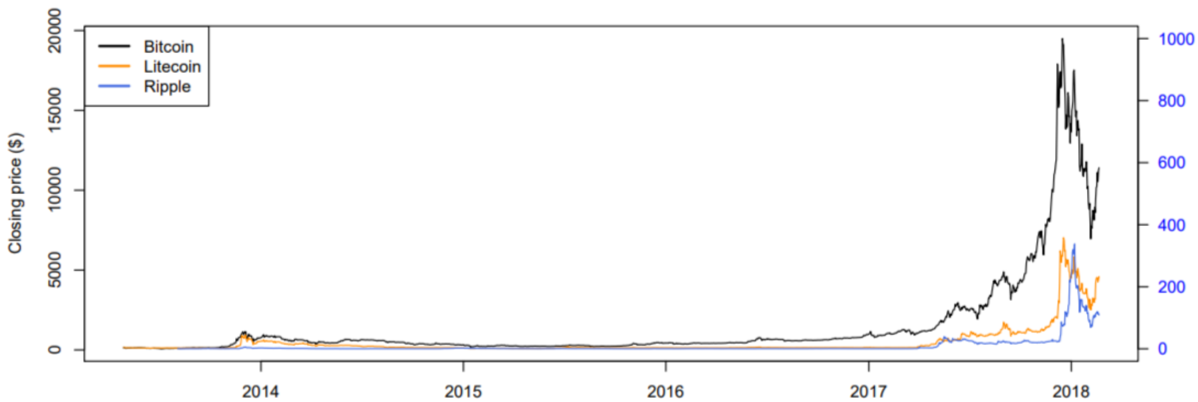


Figure 2: Price chart from paper

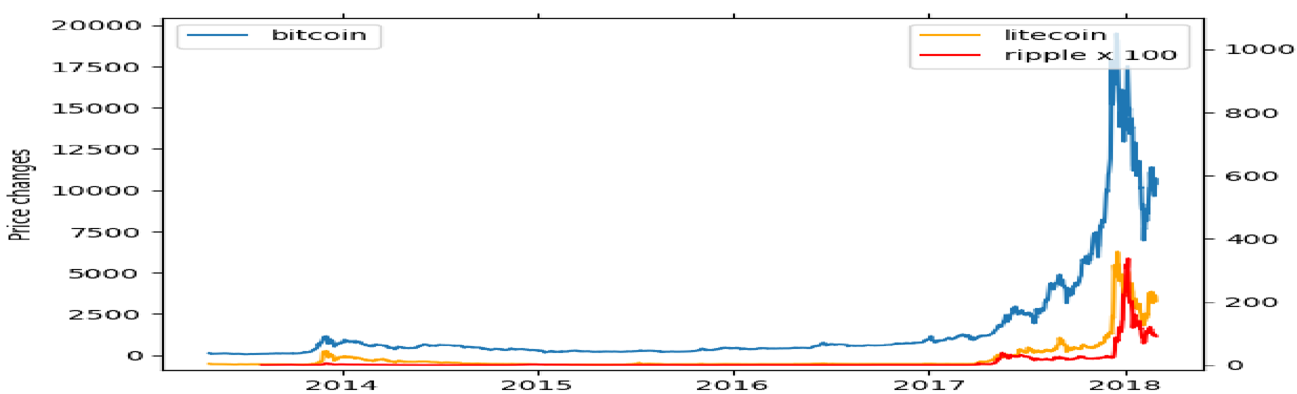


Figure 3: My price chart

As we can see, close prices charts are pretty similar. Next, let's look at log-returns graphs.

Log returns are seem to be similar too. However, to get such result one needs to multiply log returns by hundred. Another graph that plotted authors is gain over losses, and precisely - ratio between high and low quantiles.

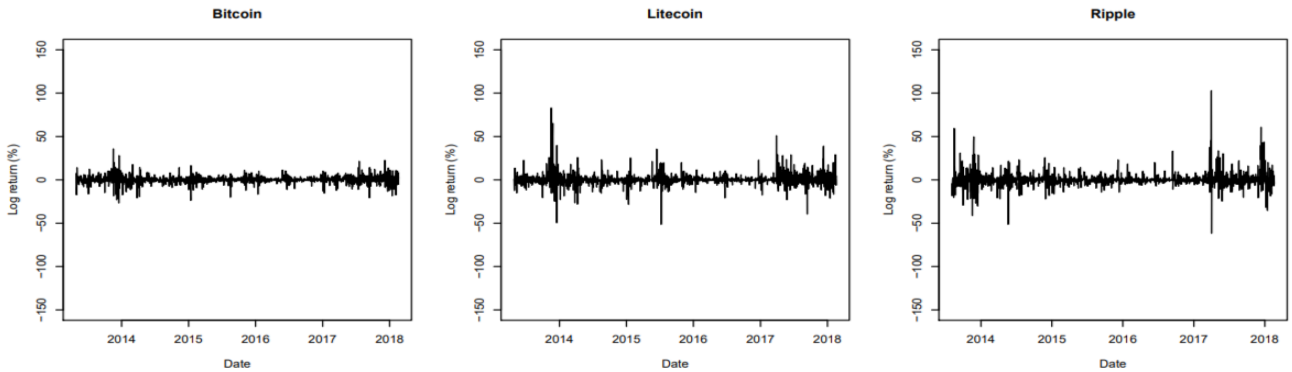


Figure 4: Log returns from paper

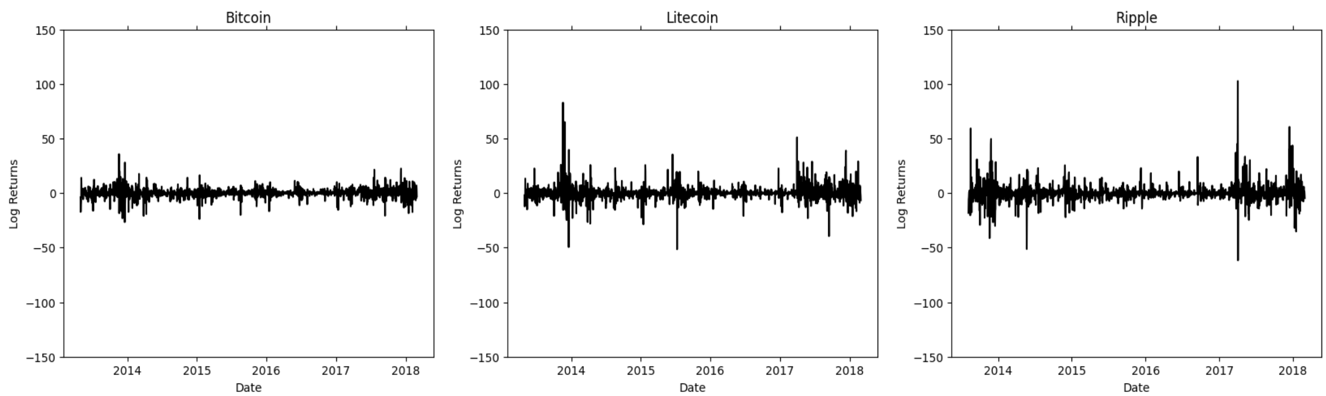
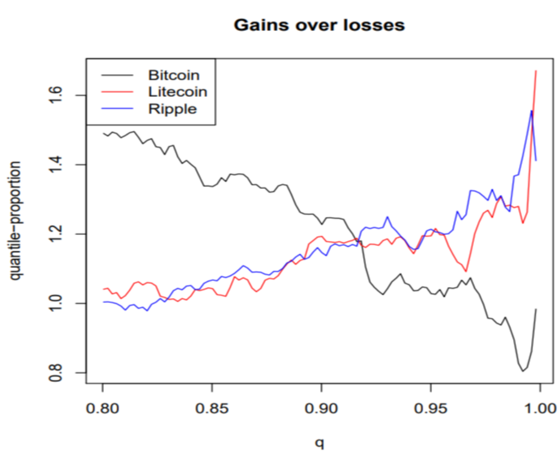
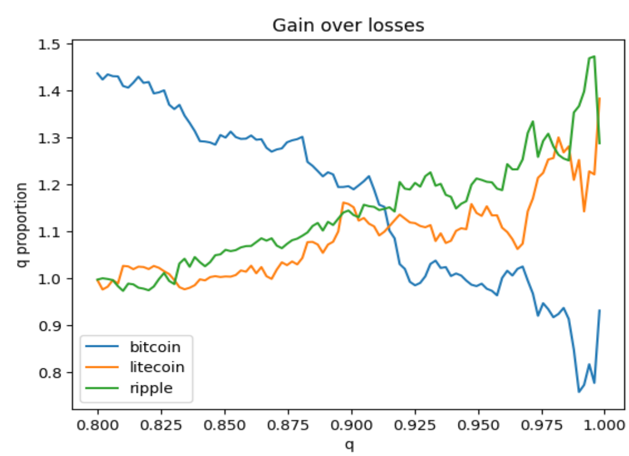


Figure 5: My log returns



(a) Paper gains over losses



(b) My gains over losses

Figure 6: Gain over losses

4. Univariate case

Firstly, authors estimated univariate log-returns. You can see paper results and mine below:

Bitcoin parameters			
α	β	γ	δ
1.319	-0.014	1.768	0.064
1.292	0.043	1.636	0.033
1.254	-0.014	1.418	0.072
1.291	-0.067	1.243	0.069
1.3	0.014	1.131	0.08
1.295	-0.059	1.144	0.163
1.27	-0.005	1.136	0.217
1.273	0.001	1.167	0.244
1.225	-0.054	1.223	0.283
1.184	-0.07	1.381	0.361

Litecoin parameters			
α	β	γ	δ
1.211	0.015	2.058	-0.246
1.15	0.064	1.824	-0.24
1.117	0.016	1.607	-0.121
1.15	-0.041	1.426	0.021
1.144	-0.035	1.21	0.016
1.107	-0.027	1.169	0.012
1.063	0.046	1.175	-0.02
1.058	0.115	1.262	-0.054
1.043	0.149	1.395	-0.078
1.051	0.096	1.5	-0.054

Figure 7: Paper estimations

	alpha	beta	gamma	delta
0	1.3080	0.0753	1.74289	0.360711
1	1.2787	0.1092	1.56749	0.440316
2	1.2460	0.0936	1.52536	0.426905
3	1.2859	0.0203	1.29609	0.082699
4	1.2883	0.0471	1.19327	0.201311
5	1.2821	0.0275	1.20491	0.215563
6	1.2963	0.0607	1.29299	0.406834
7	1.2785	0.0871	1.32907	0.521391
8	1.2760	0.1409	1.33106	0.694204
9	1.1991	0.0975	1.45727	0.788911

(a) bitcoin parameters

	alpha	beta	gamma	delta
0	1.2066	0.0057	2.09311	-0.173394
1	1.1445	0.0007	1.79909	-0.164173
2	1.1184	0.0180	1.76941	0.050533
3	1.1516	-0.0333	1.47992	-0.324583
4	1.1331	-0.0042	1.28344	-0.117259
5	1.1259	0.0030	1.22666	-0.042291
6	1.0701	0.0706	1.26302	0.769714
7	1.0522	0.0802	1.38880	1.327060
8	1.0615	0.1004	1.44285	1.476980
9	1.0298	0.0996	1.63172	3.436150

(b) litecoin parameters

Figure 8: My estimations

First row in each table corresponds to the most older 1000 observations, and the last one - to the earliest. As authors mentioned in their work, α decreases from top to bottom, so results similar at this point. But estimation itself differ. Similarity holds in alpha estimations since they differ not a lot, and maybe in gamma parameters - their behavior estimation is the same.

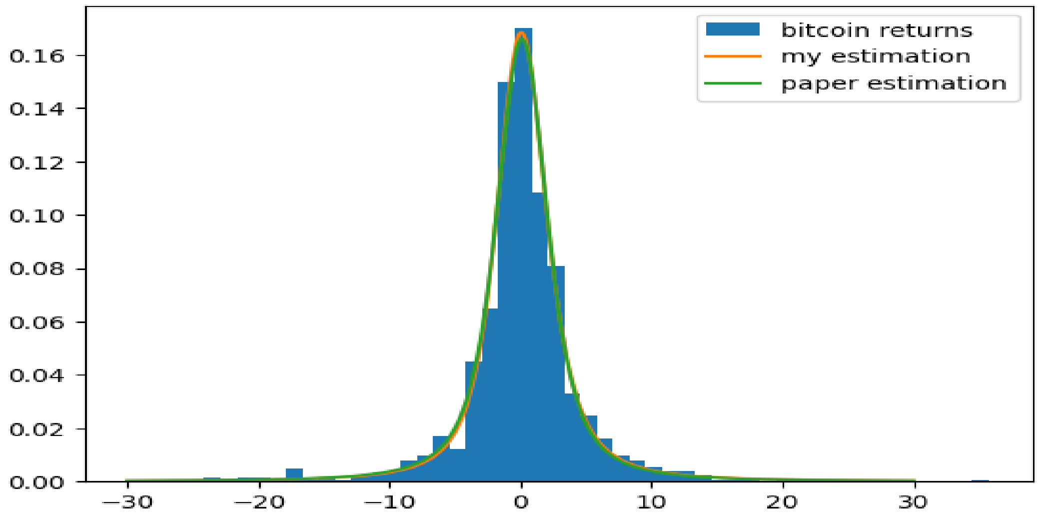


Figure 9: Comparing of estimations

Plot provided for the oldest time-series estimations. Lines are almost identical, as one can see. For estimation of parameters programm from [3] was used.

For goodness of fit test authors performed Anderson-Darling test. Paper results are following:

Bitcoin	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
Critical Value	2.428	2.387	2.399	2.664	2.343	2.426	2.55	2.659	2.57	2.491
Test statistic	1.709	1.123	0.824	1.232	1.269	1.582	1.777	2.473	3.906	4.054

Litecoin	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
Critical Value	2.625	2.591	2.334	2.577	2.358	2.366	2.594	2.548	2.172	1.967
Test statistic	0.747	0.787	1.041	1.807	1.963	1.069	1.011	1.293	1.397	1.663

Figure 10: Paper Anderson-Darling test results

	0	1	2	3	4	5	6	7	8	9
bitcoin	-3.069722	-3.190291	-3.166313	-2.969894	-2.964471	-2.765681	-2.570581	-2.423968	-2.423205	-2.169879
litecoin	-3.385945	-3.351430	-3.269589	-3.205542	-3.391588	-3.395912	-3.288543	-2.940596	-2.786535	-2.236666

Figure 11: My Anderson-Darling test results

5. Bivariate case

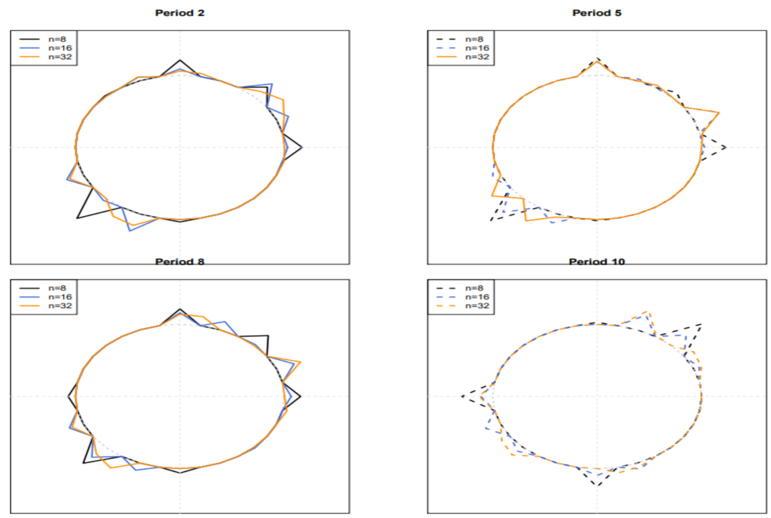


Figure 12: Paper estimated density

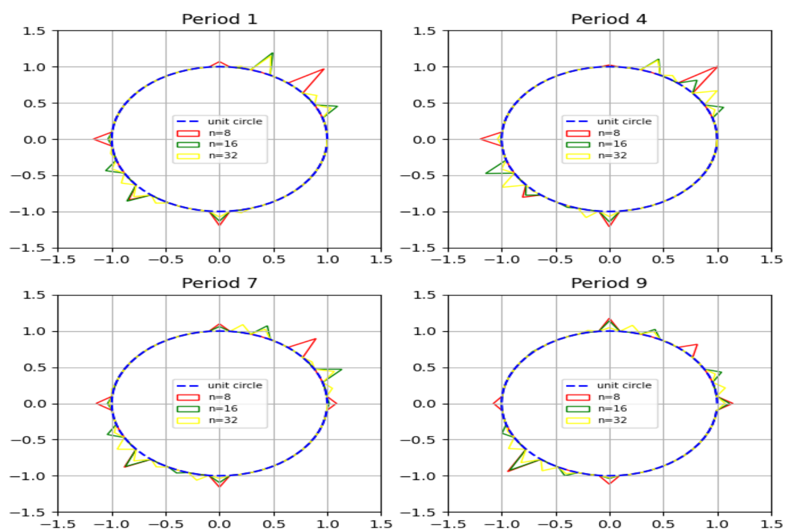


Figure 13: My estimated density

As we can compare results, they differ. It may because of different minimization problem solvers, or due to difference in estimated parameters.

Also we can compare estimated pooled alphas for bivariate case. They almost identical.

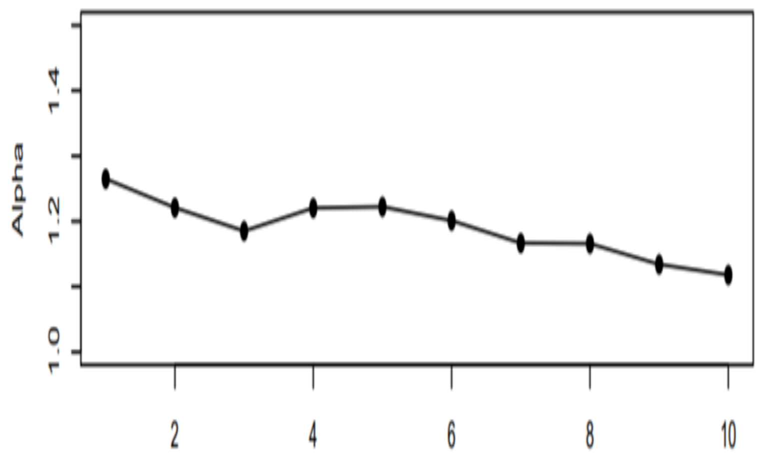


Figure 14: Paper estimated alphas

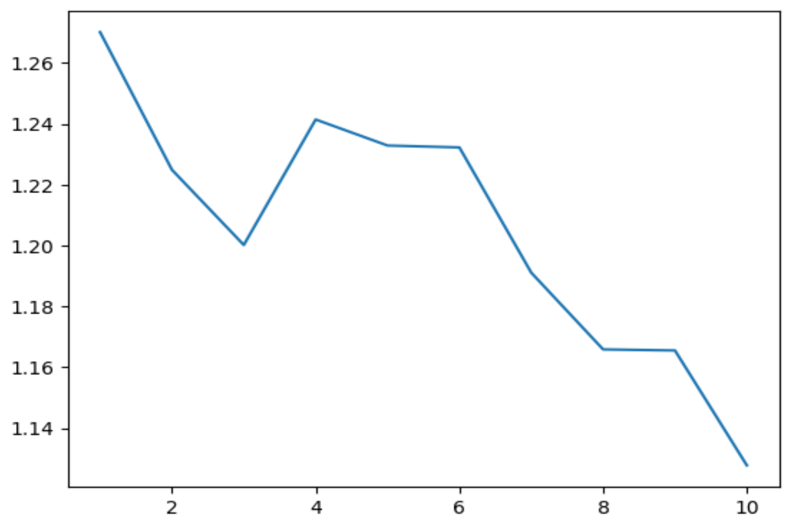


Figure 15: My estimated alphas

In conclusion we can say, that method of fitting log returns with stable distributions seems perspective, parameters estimated pretty well. But the main problem of it is overfitting on data. What about paper, authors proposed good algorithms and assumptions, but there was lack of explanation on how exactly estimate parameters, how to sample from estimated distribution. All my work can be found at [4].

6. List of References

References

[1] https://www.tandfonline.com/doi/abs/10.1080/03610918608812563

[2] https://arxiv.org/pdf/1810.09521.pdf

[3] http://robustanalysis.com/

[4] https://github.com/BogChamp/cryptocurrency_modelling