

SCTMG 2020

International Scientific Conference «Social and Cultural Transformations in the Context of Modern Globalism»

STUDENT'S DISTRIBUTIONS AS AN ALTERNATIVE TO SUSTAINABLE DISTRIBUTIONS IN THE DIGITAL ECONOMY

Umar Aptievich Bachaev (a)*, Tatiana Anatolievna Karpova (b)

*Corresponding author

(a) Financial University under the Government of the Russian Federation, 49, Leningradsky prospect, Moscow, 125468, Russia, 171086@edu.fa.ru,

(b) Financial University under the Government of the Russian Federation, 49, Leningradsky prospect, Moscow, 125468, Russia, TAKarpova@fa.ru

Abstract

As the digital economy evolves, the data used for statistical analysis becomes more complex, dynamic, and the normal distribution law used to describe more trivial data becomes less suitable for describing complex processes. In this regard, the article is devoted to the study of the Student's distribution as an alternative to stable distribution laws, including the normal one. The article considers the main properties of stable distributions of statistics, their representation through the characteristic exponent, as well as the Student's distribution with n degrees of freedom. The interrelationships of these distribution classes are shown, as well as their differences. The paper also describes the meaning of Kolmogorov-Smirnov's consent criterion, which has been repeatedly applied in the practical part of the study. Using Python 3 programming language on the Jupyter Notebook platform, a statistical analysis of the XRP (Ripple) cryptocurrency course for 365 calendar days in 2018 was carried out. The analysis consists in testing three main hypotheses: H_1 , H_2 , H_3 about subordination of the XRP course to the Student's distribution, the normal law of distribution and the Cauchy distribution, respectively. The analysis has shown that the studied data are indeed subject to the Student's distribution at the level of significance α less than 0.09043713341694139. Testing of H_2 , H_3 hypotheses has shown that the normal law of distribution and distribution of Cauchy is not suitable for describing the rate of cryptocurrency for 2018 and the hypotheses were rejected.

2357-1330 © 2020 Published by European Publisher.

Keywords: Student's distribution, stable distribution, normal distribution law, Kolmogorov-Smirnov testa.



This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 Unported License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

The concept of "sustainable distribution" was first introduced by the French mathematician P. Levi in 1925. The concept was formulated as a result of studying the properties of the sum of independent equally distributed random variables. However, the practice of applying stable laws already existed in 1919 in astronomy. Later this type of distribution began to be introduced in other scientific spheres, such as economics, physics, chemistry, geology.

Sustainable distribution laws have also become widely applied in the financial sphere. Stochastic processes with stably distributed increments replaced the variants of Brownian motion based on normal distribution, with which the analysis of dynamics of financial indicators (prices and return on assets) in a continuous period of time was connected. Mandelbrot's first arguments against "normality" were made in 1963, showing empirically that asset price dynamics included a discrete component. He described the dynamics of the price increment process as a diffusion process like Levi with infinite dispersion (Morozova & Pirlik, 2009).

2. Problem Statement

The problem with this issue is that from year to year, all statistical systems become more dynamic and it is impossible to use the normal distribution law for statistical analysis. With the advent of heavy tails, the Student's distribution begins to be used as an alternative to the Student's distribution, but the Student's distribution is rarely used for data analysis, especially for cryptocurrency. The distribution of the Student as a severe tailed one is not yet fully investigated, and this study will help to get an overview of this type of distribution.

3. Research Questions

Stable distributions and their properties. The law of distribution of a random variable X is stable in the case when for any independent and equally distributed random variables X_1 и X_2 there are such constants α, β, γ and δ , that:

$$(\alpha X_1 + \beta X_2) \sim (\gamma X + \delta),$$

the sign « \sim » shows the same distribution of random variables.

The law of distribution is called strictly sustainable if the above condition is met for any α and β , for some $\gamma > 0$ and for $\delta = 0$.

The distribution is called symmetrically stable when it is symmetrically around 0 and stable, for example $X \sim -X$, where X – it's a steady random value (Rimmer & Nolan, 2005).

The above definitions of sustainable distribution laws are considered "simple". These terms include only the identity of the laws and their aggregation, but they do not give a complete picture of sustainable laws as a parametric class. The definition of sustainability, which allows the properties of such distributions to be investigated, leads to the following statement. Random value distribution law X is stable, when $X \sim (a\varepsilon + b)$, where $a > 0, b \in R, \varepsilon$ – a random value, the distribution of which is given by a characteristic exponent or a negative natural logarithm of a characteristic function (Daniel, 2014):

$$-\ln E[e^{-it\varepsilon}] = \begin{cases} |t|^\alpha \left[1 - i\beta \operatorname{tg} \left(\frac{\pi\alpha}{2} \right) \operatorname{sign} t \right], & \alpha \neq 1 \\ |t| \left[1 - i\beta \left(\frac{2}{\pi} \right) \ln|t| \operatorname{sign} t \right], & \alpha = 1 \end{cases}$$

In the above formula $\operatorname{sign} t$ is defined as:

$$\operatorname{sign} t = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0 \end{cases}$$

At $\beta = 0, b = 0$ such distributions are symmetrical around zero and seem to be a characteristic exponent (Daniel, 2014)

$$-\ln E[e^{-it\varepsilon}] = a^\alpha |t|^\alpha.$$

Indication:

$$X \sim S_\alpha(\beta, \gamma, \delta),$$

where β, γ, δ – are parameters showing the slope, scale and shift of a random value, and α – the stability index.

Then, the law of random magnitude $X \sim S_\alpha(\beta, \gamma, \delta)$ is true only when

$$X \sim \begin{cases} \gamma \left(\varepsilon - \beta \operatorname{tg} \left(\frac{\pi\alpha}{2} \right) \right) + \delta, & \alpha \neq 1 \\ \gamma \varepsilon + \delta, & \alpha = 1 \end{cases},$$

Where the distribution of ε is defined by the characteristic function defined earlier.

Then you can plot the distribution through the characteristic exponent for $X \sim S_\alpha(\beta, \gamma, \delta)$:

$$-\ln E[e^{-it\varepsilon}] = \begin{cases} \gamma^\alpha |t|^\alpha \left[1 - i\beta \operatorname{tg} \left(\frac{\pi\alpha}{2} \right) \operatorname{sign} t (|\gamma t|^{1-\alpha} - 1) \right] + \delta, & \alpha \neq 1 \\ \gamma |t| \left[1 - i\beta \left(\frac{2}{\pi} \right) \ln|t| \operatorname{sign} t \ln|\gamma t| \right] + i\delta t, & \alpha = 1 \end{cases}.$$

The characteristic exponent is the main parameterization of the sustainable distribution of the general view.

Stable distributions arise as limits of sums of equally distributed random variables. However, not every sum limit has a stable distribution. There is a wider class of distributions, for which stable laws are a special case.

The distribution of the random variable X is infinitely divisible if $\forall n \geq 1$ you can find independent and equally distributed random variables X_{n1}, \dots, X_{nn} , so that $\sum_{k=1}^n X_{nk}$.

It can be seen that sustainable distributions are infinitely divisible, but not every infinitely divisible distribution is sustainable.

One of the main features of sustainable laws is that except in a few specific cases, sustainable laws do not have density and distribution functions. Specific cases: $N(\mu, \sigma^2) = S_2\left(0, \frac{\sigma}{\sqrt{2}}, \mu\right)$, $Co(\gamma, \delta) = S_1(0, \gamma, \delta)$, $Levy(\gamma, \delta) = S_{\frac{1}{2}}(1, \gamma, \delta)$.

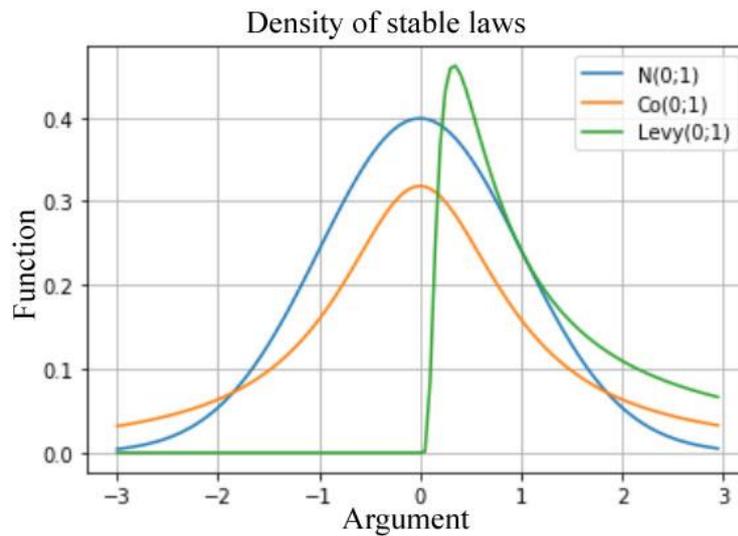


Figure 01. Densities of sustainable statistical laws

“Heavy tails” are another important feature of sustainable laws (Figure 01). It is said that distribution has heavy tails if its residual probability $P(Y > y)$ at the growth of y decreases slower than the indicative distribution E_1 . This property has been proven for $\alpha < 2$. It can be concluded from this that the property of heavy tails is true for all stable laws, except for the normal (Nolan, 2017).

Let us show the heavy tails for clarity on the graph implemented on Python 3 for the standard normal law $N(0,1)$ and Koshi's distribution $Co(0,1)$. Both laws are stable, but the latter has the properties of heavy tails as opposed to the first.

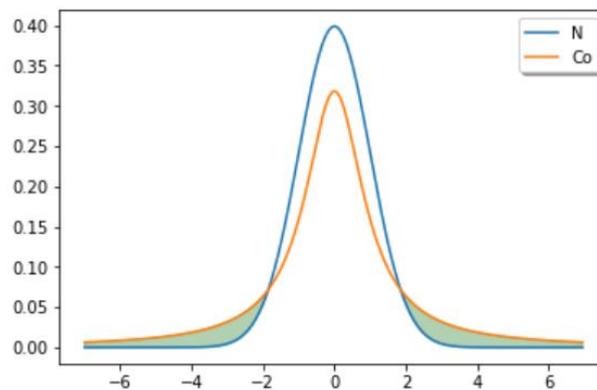


Figure 02. Normal distribution and distribution densities of Koshi

Student's distribution. Student's Distribution Properties. If $Z_0, Z_1, Z_2, \dots, Z_n$ – are independent standard normal random variables with parameters $\mu = 0, \sigma^2 = 1$. Then the distribution of a random variable t :

$$t = \frac{Z_0}{\sqrt{\frac{Z_1^2 + Z_2^2 + \dots + Z_n^2}{n}}}$$

is called the Student's distribution of n degrees of freedom (Shaw & Lee, 2008).

The Student's distribution (or t -distribution) has a probability density

$$f_n(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n} \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \forall x \in R,$$

where $\Gamma(\cdot)$ is an integral of the second kind or gamma function (Surbhi et al., 2015).

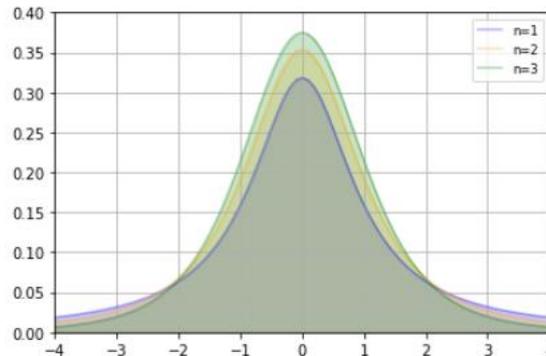


Figure 03. Student's distribution graph for different degrees of freedom

Previously, in descriptive statistics, the Student's distribution was practically not used as an analytical model. And only relatively recently appeared works in which t-distribution is used to describe the dynamics of various financial indices. The Student's distribution was started relatively recently, as it is "more complicated" than the normal law, because it is not considered to be an asymptotic approximation.

If we talk about the relationship between the Student's distribution and the stable laws of statistics, we can distinguish two properties:

- In case of large degrees of freedom, the Student's distribution converges to the standard normal distribution law, i.e. $n \rightarrow \infty t(n) \xrightarrow{d} N(0; 1)$, where \xrightarrow{d} – distribution consistency
- The Student Distribution with One Degree of Liberty is a standard Koshi distribution $t(1) = Co(0; 1)$ (Kremer, 2002)

Then, we can say that t-distribution from both sides is limited by stable laws, from below – by Koshi law, from above – by normal law. And since we consider the Student's distribution as an alternative to the stable laws, it will be reasonable to assume that t-distribution occurs just in those cases when the number of degrees of freedom is not very large and differs from 1.

4. Purpose of the Study

The purpose of this paper is to identify the relationships and differences between the Student's distribution and sustainable distributions. To show that a statistical process such as a change in the course of a cryptocurrency is subject to the Student's distribution.

5. Research Methods

To confirm the above facts we will conduct a statistical analysis in Python 3 Jupyter Notebook. To carry out the analysis, we will put forward the main hypotheses and to check them we will use the criterion of Kolmogorov-Smirnov's consent.

Hypothesis statement:

H_0^1 : Hypothesis of subordination of XRP (Ripple) course Student distribution for 2018;

H_0^2 : Hypothesis of subordination to the normal XRP (Ripple) exchange rate distribution law for 2018;

H_0^3 : Hypothesis of subordination of XRP (Ripple) course Koshi distribution for 2018.

The Kolmogorov-Smirnov Consent Criterion is used to test the hypothesis that a sample belongs to a certain distribution law.

Let us suppose that X_n – is the original sample, $F_n(x)$ – empirical distribution function, and $F(x)$ – theoretical function of distribution of the law assumed in the hypothesis H_0 .

Statistics of Kolmogorov's criterion is defined as:

$$D_n = \sup_x |F_n(x) - F(x)|.$$

Well, then, if the hypothesis is fair H_0 according to the Kolmogorov's theorem:

$$\forall t > 0: \lim_{n \rightarrow \infty} P(\sqrt{n}D_n \leq t) = K(t) = \sum_{j=-\infty}^{+\infty} (-1)^j e^{-2j^2 t^2}$$

The hypothesis is rejected if $\sqrt{n}D_n$ exceeds the quantile of distribution K_α the quantile of distribution α , but is accepted otherwise (Brailov, 2007).

6. Findings

As the data for the study, we will take the RippleXRP network cryptocurrency rate to the dollar for the period from 01.01.2018 to 31.12.2018.

Upload the data to *PythonJupyterNotebook*.

	<TICKER>	<PER>	<DATE>	<TIME>	<OPEN>	<HIGH>	<LOW>	<CLOSE>	<VOL>
0	BTSP.XRP/USD	D	2018-01-23	0	1.26264	1.43300	1.26264	1.34366	18630738
1	BTSP.XRP/USD	D	2018-01-24	0	1.34366	1.42000	1.28601	1.36000	28270598
2	BTSP.XRP/USD	D	2018-01-25	0	1.35225	1.37900	1.26953	1.30203	17036113
3	BTSP.XRP/USD	D	2018-01-26	0	1.29703	1.32000	1.09897	1.20938	40593234
4	BTSP.XRP/USD	D	2018-01-27	0	1.20759	1.24001	1.18000	1.21899	13655012
5	BTSP.XRP/USD	D	2018-01-28	0	1.21451	1.38900	1.21009	1.35000	18441605
6	BTSP.XRP/USD	D	2018-01-29	0	1.34722	1.35612	1.24000	1.26221	16742793
7	BTSP.XRP/USD	D	2018-01-30	0	1.25870	1.26349	1.06898	1.09781	40097419
8	BTSP.XRP/USD	D	2018-01-31	0	1.09781	1.15000	1.00000	1.12304	27943750
9	BTSP.XRP/USD	D	2018-02-01	0	1.12647	1.12999	0.86850	0.94185	54259341
10	BTSP.XRP/USD	D	2018-02-02	0	0.93601	0.95800	0.63252	0.89389	127492737
11	BTSP.XRP/USD	D	2018-02-03	0	0.89389	1.01915	0.77252	0.96061	45890981

Figure 04. Input data (first 12 lines)

Let us build a histogram of relative frequencies density by closing prices (by column <CLOSE>).

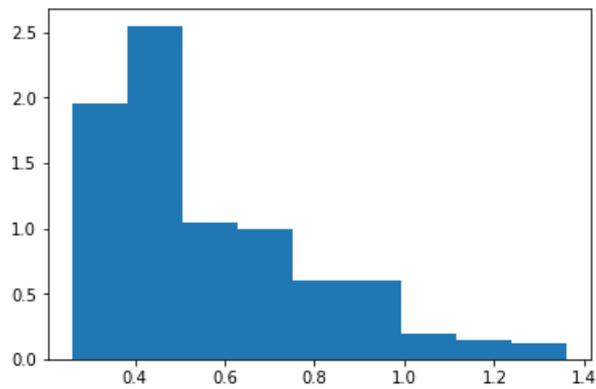


Figure 05. Relative frequency density histogram

Let us test the hypothesis H_0^1 . Fig.5 shows that the histogram is beveled on the right side, so we will select the parameters for the Student's non-central distribution using the fit function.

```
stats.nct.fit(course)
(4.67595531116085,
19.089683534241125,
-0.03062055788396588,
0.026061979893345032)
```

Figure 06. Parameter selection for the Student's non-central distribution

For parameters $df=4.67595531116085$; $nc=19.089683534241125$; $loc=-0.03062055788396588$; $scale=0.026061979893345032$ superimpose a density graph on the histogram.

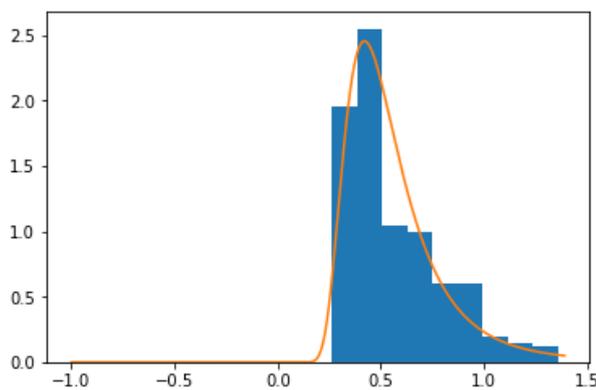


Figure 07. The Student's Non-Central Distribution Graph superimposed on the histogram

We will run a Kolmogorov-Smirnov test

```
KstestResult(statistic=0.06659022444167023, pvalue=0.09043713341694139)
```

Figure 08. Kolmogorov-Smirnov's test for the Student's non-central distribution

The test results show that $\forall \alpha \leq 0.09043713341694139$ the hypothesis is accepted. Let us check it out for $\alpha = 0.05$.

We have $D_{0,05} = \frac{1,36}{\sqrt{344}} = 0.07332628577833812$. As $D_{0,05} >$ statistic, is that way, the hypothesis is accepted.

Let us test the hypothesis H_0^2 . To do this, select the parameters for normal distribution.

```
stats.norm.fit(course)
(0.565534011627907, 0.23122284360134682)
```

Figure 09. Selection of parameters for normal distribution

Let us build a normal distribution chart with parameters $m=0.565534011627907, \sigma=0.23122284360134682$.

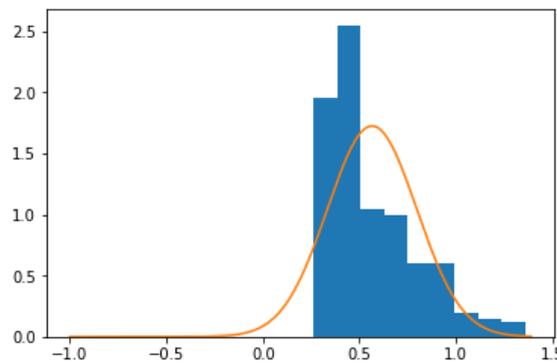


Figure 10. Normal distribution graph superimposed on the histogram. We'll run a Kolmogorov-Smirnov test

```
KstestResult(statistic=0.15099570722372313, pvalue=2.594775577854376e-07)
```

Figure 11. Kolmogorov-Smirnov test for normal distribution

Value p value is very small, and we can say that the hypothesis is rejected. Let us check it for the value level $\alpha = 0.01$. $D_{0,01} = \frac{1,63}{\sqrt{344}} \approx 0.0878837101608023$. In this case, the value $D_{0,01}$ is less than the empirical value of statistics, which means we are in a critical area $K_{0,01}$. Therefore, the hypothesis is rejected.

Let us test the hypothesis H_0^3 subordination of the XRP course Koshi distribution for 2018. We will select the parameters of Koshi distribution.

```
stats.cauchy.fit(course)
(0.48066010272444215, 0.10876074731160862)
```

Figure 12. Selection of parameters for Koshi distribution

Let us superimpose a graph of Cauchy distribution with parameters $a=0.48066010272444215, b=0.10876074731160862$ on the relative frequency density histogram.

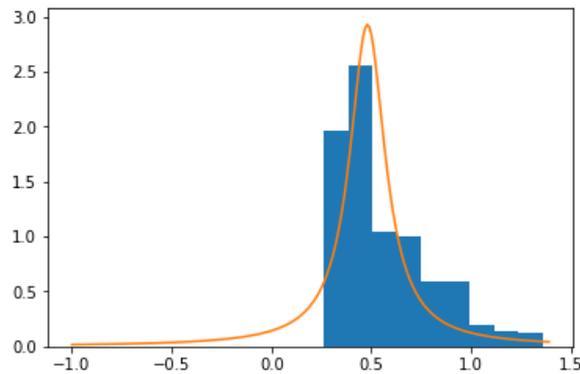


Figure 13. The Cauchy distribution graph superimposed on the histogram
 We'll run a Kolmogorov-Smirnov test

```
kstestResult(statistic=0.14835434873559805, pvalue=4.501669506711841e-07)
```

Figure 14. Kolmogorov-Smirnov test for Koshi distribution

Value pvalue it's a small one. It is reasonable to assume that the hypothesis is rejected. Let us check it for the value level $\alpha = 0.01$. $D_{0,01} = \frac{1,63}{\sqrt{344}} \approx 0.0878837101608023$. In this case, the value $D_{0,01}$ is less than the empirical value of statistics, so we are in the critical area $K_{0,01}$. As a result, this hypothesis is also rejected.

7. Conclusion

Thus, we come to the conclusion that with the development of different dynamic systems the data cannot obey the normal distribution law, so we have to resort to some alternative variant. The best such variant is the Student's distribution.

Kolmogorov-Smirnov's test has shown that the hypothesis of subordination of the Student's XRP (Ripple) course distribution for 2018 is accepted for $\alpha \leq 0.09043713341694139$, that is, for any significance level below α , the sample will be subject to the Student's non-central distribution law with parameters $df=4.67595531116085$; $nc=19.089683534241125$; $loc=-0.03062055788396588$; $scale=0.026061979893345032$.

Also by the Kolmogorov-Smirnov criterion we rejected the hypotheses about the subordination of the course to XRP-normal law and the distribution of Koshi, which also shows the need to use the Student's distribution instead of the stable laws of statistics.

References

- Brailov, A.V. (2007). *Lectures on Mathematical Statistics*. Moscow: Finacademy.
- Daniel, T.P. (2014). A SAS approach for estimating the parameters of an alpha-stable distribution. *Procedia Econ. and Finance*, 10, 68–77.
- Kremer, N. Sh. (2002). *Probability Theory and Mathematical Statistics*. Textbook for Higher Schools. Moscow: Uniti-Dana.
- Morozova, M. M., & Pirlík, V.N. (2009). Stable distribution, its modifications and financial derivatives pricing model. *Vestn. NGU*, 9(1), 97–112.

- Nolan, J.P. (2017). *Stable Distributions Models for Heavy Tailed Data*. Washington: Amer. Univer.
- Rimmer, R.H., & Nolan, J.P. (2005). Stable Distributions in Mathematica. *The Mathem. J.*, 9(4), 776–789.
- Shaw, W.T., & Lee, K.T.A. (2008). Bivariate Student distributions with variable marginal degrees of freedom and independence. *J. of Multivar. Anal.*, 99, 1276–1287.
- Surbhi, S., Harsh, V.S., & Anand, M. (2015). Secure and Robust Watermarking Using Wavelet Transform and Student t-distribution. *Procedia Comput. Sci.*, 70, 442–447.