Unconditional and conditional heavy-tailed distribution for returns of cryptocurrencies

Quang Van Tran

KSI, Faculty of Nuclear Sciences and Physical Engineering, CTU in Prague

SPMS conference 2024
Dobřichovice, Czech Republic, June 23, 2024

Motivation

Figure 1: The dynamics of BTC price and returns

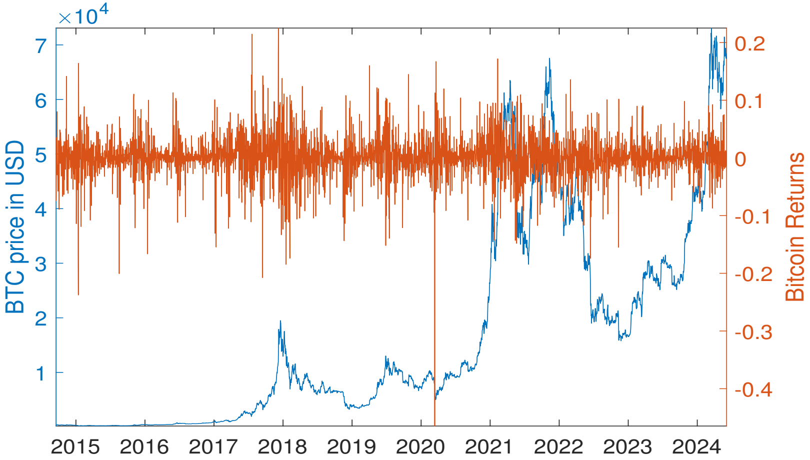
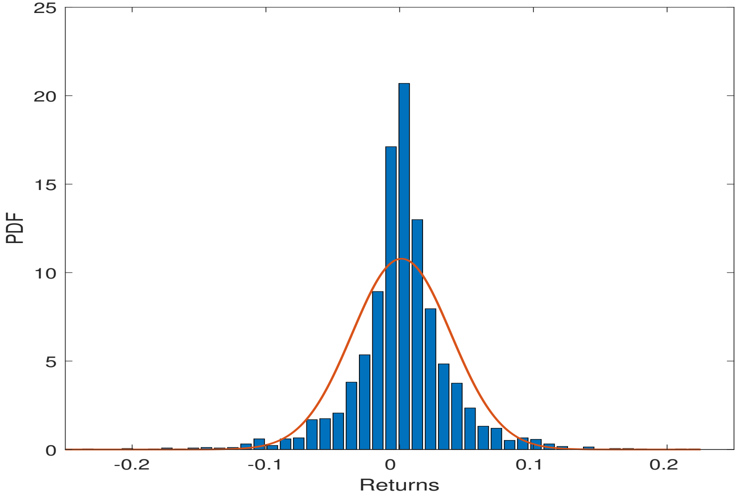


Figure 2: The density of returns of Bitcoin



Motivation and Research Objective

- Returns of financial assets exhibit two stylized properties: leptokurtic property and heteroskedasticity and normal distribution is not suitable for modeling conditional and unconditional
- Returns of cryptocurrencies are no exception and they rae extremely volatile
- Six alternative distributions have been proposed to deal with the first problems: t-distribution, generalized normal distribution, Normal Inverse Gaussian distribution (NIG), Alpha stable distribution
- Four GARCH models are chosen to deal with the heteroskedasticity problem
- We verify on five most traded cryptocurrencies using their daily price series

Heavy tail distributions

We choose six tail distributions to model the returns of cryptocurrencies

- Student t-distribution
- Generalized normal distribution (GN)
- Normal Inverse Gaussian distribution (NIG)
- Alpha stable distribution (ALPHA)
- Two-sided Regularized Generalized Gamma distribution (Tran and Kukal, Finance Research Letters) (TSRGG)
- Two-shape parameter Generalized t-distribution (Tran and Kukal, Physica A) (NEW)

GARCH models

We use four tail GARCH models to model the heteroskedasticity of the returns of cryptocurrencies

- Plain GARCH
- Threshhold GARCH
- Range GARCH
- Exponential GARCH

Conditional and unconditional returns

- Unconditional returns are modeled by one of the chosen distributions without any additional manipulation with return data
- Conditional returns are jointly modeled with ARMA model for the conditional mean process and GARCH model for the conditional variance process

ARMA and GARCH model

An ARMA-GARCH model of a random variable representing daily returns of a cryptocurrency r_t is a two-part econometric model where the conditional mean process is

$$r_t = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 \varepsilon_{t-1} + \varepsilon_t, \tag{1}$$

where $\alpha_0, \alpha_1, \alpha_2$ are parameters. The error term (or innovation term) ε_t is often linked to the conditional variance via this assumption

$$\varepsilon_t = z_t \sigma_t, \tag{2}$$

where $z_t \sim Z(0,1)$ and the conditional variance term σ_t^2 follows a GARCH(1,1) process defined as

$$\sigma_t^2 = \gamma_0 + \gamma_1 \sigma_{t-1}^2 + \gamma_2 \varepsilon_{t-1}^2, \tag{3}$$

where $\gamma_0, \gamma_1, \gamma_2$ are the parameters of the GARCH(1,1) process which have to be determined

Data and Descriptive statistics

We select five cryptocurrencies: Bitcoin, Binance coin, Ethereum, Solana, Ripple to verify our approach. Data are converted to logarithmic returns and their descriptive statistics are displayed below

Characteristic	Bitcoin	B2B	Ethereum	Solana	Ripple
Mean	1.10E-03	1.56E-03	2.44E-03	3.33E-03	1.38E-03
Median	1.21E-03	7.25E-04	5.93E-04	-5.84E-05	-8.54E-04
Minimum	-0.434	-0.572	-0.563	-0.548	-0.550
Maximum	0.287	0.756	0.373	0.384	0.881
1st quartile	-0.013	-0.018	-0.021	-0.033	-0.020
3rd quartile	0.017	0.022	0.026	0.038	0.019
Std. deviation	0.037	0.057	0.055	0.072	0.062
Skewness	-0.609	0.868	-0.270	-0.264	1.962
Kurtosis	13.082	36.720	11.023	9.220	31.349
Num of Obs	3694	2233	3101	1397	3537

Estimation of parameters of the new distribution

Parameters of the new distribution are estimated jointly using maximum likelihood estimation method (MLE). The MLE procedure is performed as follows:

$$\hat{\theta} = -\arg\min_{\theta \in \Theta} \sum_{i=1}^{n} \ln f(X_i;\theta) = -\arg\min_{\theta \in \Theta} \ln L(\theta), \tag{4}$$

where Θ is the set of all admissible parameters and $f(X_i;\theta)$ is the density function of the corresponding distribution. The sign "-" is added so that the minimalization procedure can be applied. The MLE estimator has the following property

$$\sqrt{n}(\hat{\theta} - \theta_{true}) \sim N(0, \mathscr{I}^{-1}),$$

where $\mathscr{I} = -\mathbb{E}\left[\frac{\partial^2 \ln L}{\partial\theta\partial\theta'}\right]$ is the so called Fisher information matrix.

Model quality evaluation

The quality of a model is measured by the AIC criterion. The AIC value of a distribution is

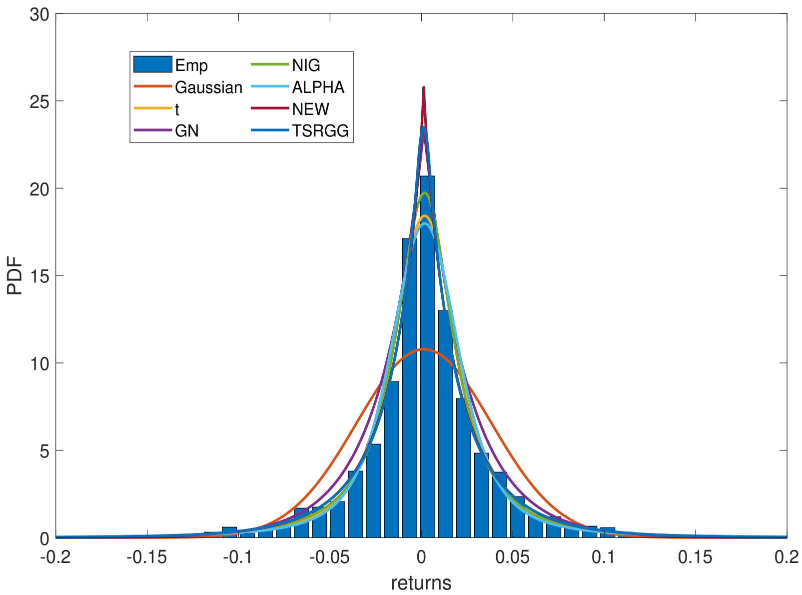
$$\mathrm{AIC} = 2k - 2\ln L, \tag{5}$$

where k is the number of parameters of a model, $\ln L$ is the optimal value of log-likelihood function

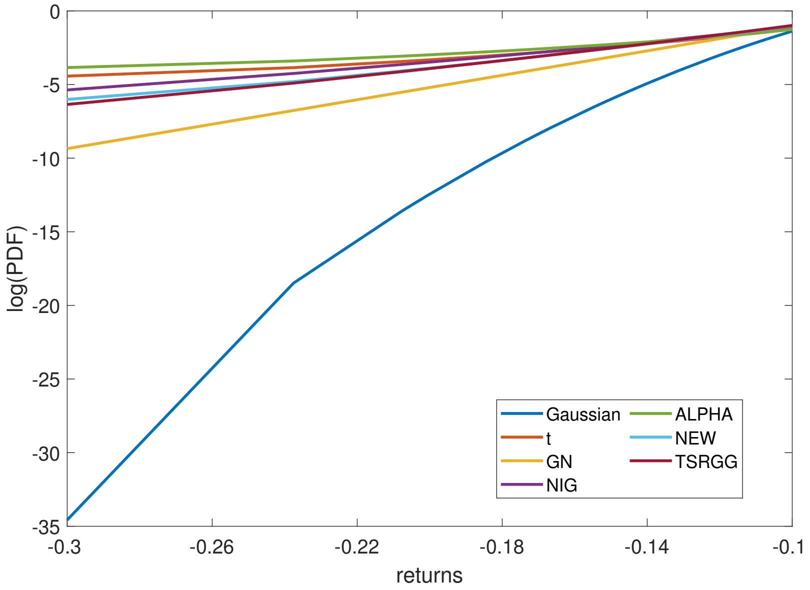
Estimation results - Unconditional Returns

Dist	Bitcoin	B2B	Ethereum	Solana	Ripple
Normal	-13788.97	-4054.65	-9167.81	-3392.22	-9533.80
t	-14964.86	-7578.55	-10003.27	-3644.00	-11899.12
GN	-15874.19	-7699.85	-10775.55	-3970.04	-12453.58
NIG	-15031.16	-7537.20	-10048.03	-3648.90	-11930.03
ALPHA	-14875.98	-7557.25	-9938.28	-3629.33	-11851.29
NEW	-15063.05	-7585.44	-10061.48	-3644.49	-11922.51
TSRGG	-15070.46	-7524.28	-10062.69	-3643.18	-11920.21

Distributions of returns of Bitcoin - Comparison



Distributions of returns of Bitcoin - Left tails



Estimation results - Plain GARCH

Dist	Bitcoin	B2B	Ethereum	Solana	Ripple
Normal	-14575.17	-6547.13	-9985.73	-3661.06	-11651.48
t	-14999.13	-7925.67	-10221.08	-3716.07	-12203.75
GN	-15518.08	-8146.16	-10577.59	-3817.44	-12772.95
NIG	-15329.79	-7891.69	-10384.93	-3736.67	-12341.17
ALPHA	-14451.33	-7662.66	-9785.58	-3527.07	-11857.88
NEW	-15391.92	-8036.40	-10414.37	-3764.71	-12435.36
TSRGG	-15416.64	-7889.07	-10415.73	-3723.03	-12446.66

Estimation results - RGARCH

Dist	Bitcoin	B2B	Ethereum	Solana	Ripple
Normal	-15615.24	-7695.70	-10555.74	-3750.52	-11021.53
t	-16476.44	-7987.97	-11006.06	-3794.81	-12222.68
GN	-16914.87	-8225.32	-11323.59	-3882.41	-12615.53
NIG	-16804.46	-8064.44	-11174.42	-3820.18	-12512.50
ALPHA	-15903.92	-7687.25	-10565.71	-3598.52	-11757.29
NEW	-16763.90	-8095.37	-11161.09	-3842.56	-12580.99
TSRGG	-16747.96	-8085.75	-11155.85	-3841.99	-12558.82

Estimation results - TGARCH

Dist	Bitcoin	B2B	Ethereum	Solana	Ripple
Normal	14572.39	6192.28	9983.73	3659.22	11649.48
t	15159.53	7959.24	10304.02	3735.73	12326.65
GN	15898.76	8305.73	10906.18	3968.98	13055.93
NIG	15336.01	7890.03	10387.33	3742.92	12365.32
ALPHA	14893.70	7829.57	10103.50	3659.73	12145.31
NEW	15370.87	8026.45	10396.17	3762.17	12437.14
TSRGG	15414.64	7962.81	10424.61	3760.67	12452.50

Estimation results - EGARCH

Dist	Bitcoin	B2B	Ethereum	Solana	Ripple
Normal	14572.39	6192.28	9983.73	3659.22	11649.48
t	15159.53	7959.24	10304.02	3735.73	12326.65
GN	15898.76	8305.73	10906.18	3968.98	13055.93
NIG	15336.01	7890.03	10387.33	3742.92	12365.32
ALPHA	14893.70	7829.57	10103.50	3659.73	12145.31
NEW	15370.87	8026.45	10396.17	3762.17	12437.14
TSRGG	15414.64	7962.81	10424.61	3760.67	12452.50

Estimation results - GARCH with GN distribution

Model	Bitcoin	B2B	Ethereum	Solana	Ripple
Unconditional	-15874.19	-7699.85	-10775.55	-3970.04	-12453.58
GARCH	-15518.08	-8146.16	-10577.59	-3817.44	-12772.95
RGARCH	-16914.87	-8225.32	-11323.59	-3882.41	-12615.53
TGARCH	-15898.76	-8305.73	-10906.18	-3968.98	-13055.93
EGARCH	-16358.44	-8508.29	-11211.44	-4090.62	-13317.84

Extension: The Range-Exponential GARCH Model

The Range-Exponential GARCH process is defined as

$$\log(\sigma_t^2) = \gamma_0 + \gamma_1 \log(\sigma_{t-1}^2) + \gamma_2 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \gamma_3 \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma_4 \log(\sigma_{p,i-1}^2), \quad (6)$$

where

$$\sigma_{p,t}^2 = \frac{[\log(H_t/L_t)]^2}{4 \log 2}, \quad (7)$$

where H_t and L_t are the highest and lowest price of an asset in a day, respectively.

Values of AIC criterion of Range-Exponential GARCH

Distr.	Bitcoin	B2B	Ethereum	Solana	Ripple
Normal	-13950.45	-8287.67	-8185.39	-4063.95	-7444.34
t	-14973.96	-8719.32	-8662.31	-4190.29	-8569.84
GN	-15792.01	-9298.23	-9232.21	-4520.55	-9041.05
NIG	-15029.41	-8701.45	-8681.16	-4179.26	-8534.53
ALPHA	-14956.88	-8719.47	-8657.83	-4215.62	-8582.40
NEW	-14991.88	-8728.03	-8668.31	-4182.21	-8570.13
TSRGG	-14972.89	-8721.21	-8657.21	-4183.45	-8556.21

Conclusion

- For each cryptocurrency of the five ones Bitcoin, Binance coin, Ethereum, Solana, and Ripple, the most appropriate distribution is generalized normal distribution.
- This conclusion remains the same no matter whether we consider unconditional, or conditional distribution, and in case of conditional distribution, no matter which volatility model we use
- Among the four volatility models we tested, the EGARCH model is one of those which best fits for modeling returns of cryptocurrencies while in some cases the range GARCH model is also fairly good, but it is not as reliable as the previous one.
- Extension: R-GARCH and EGARCH models can be combined and the combination is a better alternative of both individual versions

Thank you for your attention!