Model Selection for Stable Distributions over the Chinese ETF Market

Introduction

Financial markets frequently display price movements that deviate from normality assumptions, such that VaR calculations based on normal distributions could underestimate the risk of severe losses over extreme market events. Such deviations in US stock indices have been commented upon by numerous studies, and some researchers aim to use a new family of distributions, the stable distributions, as alternative statistical models. This small write-up focuses on the feasibility of such stable distributions on the Chinese ETF market. We commence with a recapitulation on the basic tenets of stable distribution. Then, we shall discuss how model selection can be made with the aid of likelihood-ratio tests as well as AIC and BIC. Eventually, we shall apply the concepts to two ETFs traded on the Chinese ETF market to demonstrate the principles and implications.

Section 1. Statistical models on stochastic processes from stable distributions

1.1. Generalization from normal to stable distribution

Let $(X_t)_{t\geq 0}$ be a discrete-time stochastic process with initial value $X_0 = x_0$ with probability one. The normality approach dictates that the associated process $\ln(X_t)_{t\geq 0}$ has independent increments that follow a normal distribution with

$$\ln(X_{t+1}) - \ln(X_t) = \ln(X_{t+1}/X_t) \sim \mathrm{N}(\mu, \sigma^2),$$

such that the return $R_t = X_t/x_0$ at time t follows

$$\ln(R_t) = \ln(X_t/x_0) \sim \mathrm{N}(\mu t, \sigma^2 t).$$

Nonetheless, it is well-established that for stochastic movements in share prices, this normality assumption does not hold, such that indiscriminate use of normal distributions tends to underestimate the tail probability for extreme market events. A natural remedy which assimilates the above-mentioned protocol is to use independent stable distributions instead to model the increments, such that

$$\ln(X_{t+1}) - \ln(X_t) = \ln(X_{t+1}/X_t) \sim \mathrm{S}_\alpha(\beta, c, \mu)$$

with α, β, c and μ as the stability, skewness, scale and location parameters respectively. The distribution for R_t then becomes

$$\ln(R_t) = \ln(X_t/x_0) \sim \mathrm{S}_\alpha\left(\beta, c\sqrt{t}, \mu t\right)$$

based on closure of stable distribution upon independent sums. Comparatively, the stable model is a superset of the normal model, and the most crucial concern is whether α deviates substantially from 2, and whether β deviates substantially from zero. For share price models, the former concerns whether $\ln(R_t)$ has heavier tails that preclude application of SLLN and CLT for mean returns over protracted durations, while the latter concerns whether complications of bullish and bearish market events are symmetrically distributed. Since the normal model can be viewed as a reduced model to the stable model, we naturally consider likelihood-ratio tests to compare different models, which will be discussed in the next section.

1.2. Likelihood-ratio tests for nested stable models

Since our main concerns are whether or not to constrain the stability and skewness parameters, we naturally have four permutations. If there is also concern whether the location parameter deviates substantially from zero, we would also like to assess if $\mu = 0$, which results in 8 possible permutations. But when $\alpha = 0$, whether or not β is constrained becomes irrelevant, such that the total number of candidate models amounts to 6, which are presented in Table 1.2.1.

Table 1.2.1: List of models and parameter constraints

Model	α	β	μ	c	No. of estimates
$\ln(R_t) \sim \mathrm{N}(0, \sigma^2 t)$	−	−	−	+	1
$\ln(R_t) \sim \mathrm{S}_\alpha\left(0, c\sqrt{t}, 0\right)$	+	−	−	+	2
$\ln(R_t) \sim \mathrm{S}_\alpha\left(\beta, c\sqrt{t}, 0\right)$	+	+	−	+	3
$\ln(R_t) \sim \mathrm{N}(\mu t, \sigma^2 t)$	−	−	+	+	2
$\ln(R_t) \sim \mathrm{S}_\alpha\left(0, c\sqrt{t}, \mu t\right)$	+	−	+	+	3
$\ln(R_t) \sim \mathrm{S}_\alpha\left(\beta, c\sqrt{t}, \mu t\right)$	+	+	+	+	4
+: Parameter that requires an estimate.					
−: Parameter with constraint.					

Within the table, one model is nested within another model, if the set of parameter constraints in the former is a superset to that of the latter. For hypothesis tests, we treat the nested model as the null model, and the more complicated model as the alternative model. Let ℓ_0 and ℓ_A denote the loglikelihoods for the null and alternative models. The likelihood-ratio statistic then becomes

$$\lambda = 2(\ell_A - \ell_0) \sim \chi^2_{d_A - d_0},$$

where d_0 and d_A denote the number of estimated parameters for the null and alternative models respectively. These results then allow us to set up a forward model selection protocol as below.

1. Treat the centered normal model $\ln(R_t) \sim \mathrm{N}(0, \sigma^2 t)$ or the normal model $\ln(R_t) \sim \mathrm{N}(\mu t, \sigma^2 t)$ as the baseline model, in accordance with whether the constraint $\mu = 0$ should be considered.
2. Locate all models with one additional parameter estimate, and incorporate the most important parameter.
3. Repeat Step 2 until all other parameters are unimportant.

Similarly, the backward elimination protocol can be derived as below.

1. Treat the stable model $\ln(R_t) \sim \mathrm{S}_\alpha\left(\beta, c\sqrt{t}, \mu t\right)$ as the full model.
4. Locate all models with one less parameter estimate, and discard the least important parameter, in accordance with whether the constraint $\mu = 0$ should be considered.
2. Repeat Step 2 until no more parameters can be discarded.

1.3. One-period models and multiperiod models

The above-mentioned discussions have provided a systematic approach to select a parsimonious model that adequately captures the deviation of share price movements from normality. Nonetheless, for models that cover protracted durations, we inevitably encounter situations where the core trend is interspersed by some unforeseen market event that causes additional turbulence to price movements. Instead of a one-period model that puts all data under one distribution, these scenarios prompt the need for multiperiod models, where we compute different stable distributions for the core trend and for the period of market turbulence, to which our previous protocols shall adapt. In effect, suppose that we partition the entire dataset into n periods. We would then select one best distribution for each time period, so that the entire dataset is described by n separate distributions. Let d_i and ℓ_i denote the number of estimated parameters and the loglikelihood for the ith distribution. Let d_0 and ℓ_0 denote the results for the one-period model. The likelihood-ratio statistic is then computed as

$$\lambda = 2\left[\left(\sum_{i=1}^{n} \ell_i\right) - \ell_0\right] \sim \chi^2_{\left(\sum_{i=1}^{n} d_i\right) - d_0},$$

provided that the one-period model is nested in the multiperiod model.

1.4. Use of model selection metrics

When considerations in Section 1.3 are made, we may face an increased number of candidate models, such that the model metrics AIC and BIC can be computed to identify the best model. In practice, BIC tends to select more parsimonious models than AIC, whereas likelihood-ratio tests depend on the error level used. Therefore, it is always desirable to compare the results from all three approaches, and employ relevant market fundamentals to select the most appropriate model.

Section 2. Case studies on selected ETF in the Chinese ETF market

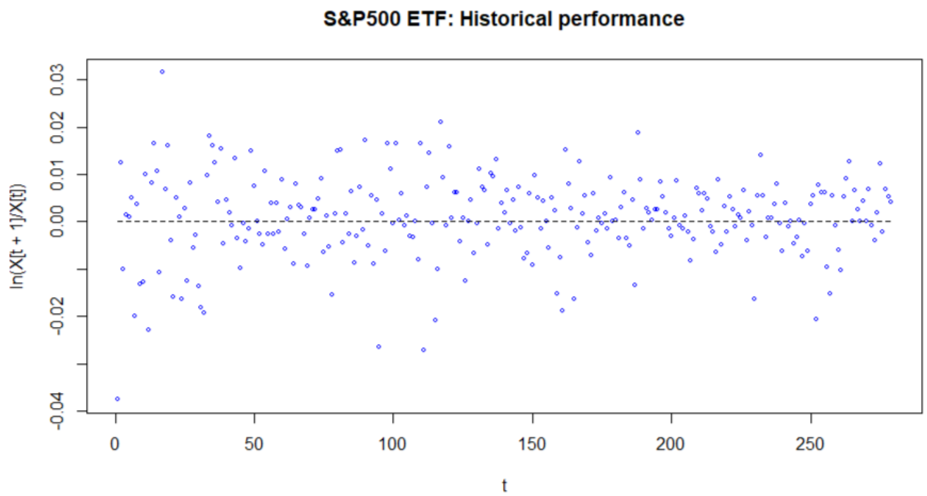
This section includes case studies on selected ETFs in the Chinese ETF market, which demonstrates exponential price movements amenable for analysis with stable distributions. In all cases, we must bear in mind that most ETFs do not show such trends, such that the application of the stable distributions, just as any other models, must be substantiated by evidence from historical performance and relevant market forecasts. Nonetheless, this write-up focuses only on the statistical procedures for model construction and selection, and would not cover contents on fundamental analysis. We will construct relevant models as delineated in Section 1, and select the appropriate models. Should there be substantial deviation from normality, we shall compute the VaRs for different probabilities and durations to quantify such deviations.

2.1. Case 1: S&P500 ETF

The Chinese A-shares market usually displays different patterns from the US stock market. Nonetheless, a Chinese trader could obtain exposure to the US market by purchase of relevant

ETFs issued in the Chinese ETF market. Since ETFs for the US market closely track the relevant US stock indices, it is of interest to assess whether prices in these ETFs show similar deviations from normality. In this case, we shall evaluate the historical data for an S&P500 ETF from 2020-09-08 to 2021-11-05, of which the daily increments are computed and displayed in Plot 2.1.1.

Plot 2.1.1. S&P500 ETF historical performance



The patterns for increments are rather consistent over different time periods, such that a one-period model is adequate. The likelihood results, AICs and BICs for the 6 candidate models are shown in Table 2.1.1.

Table 2.1.1: Model selection results for S&P500 ETF

Model	α	β	μ	c	No. of estimates	ℓ	AIC	BIC
$\ln(R_t) \sim \mathrm{N}(0, \sigma^2 t)$	−	−	−	+	1	923.70	-1845.4	-1841.7
$\ln(R_t) \sim \mathrm{S}_\alpha(0, c\sqrt{t}, 0)$	+	−	−	+	2	928.52	-1853.0	-1845.8
$\ln(R_t) \sim \mathrm{S}_\alpha(\beta, c\sqrt{t}, 0)$	+	+	−	+	3	933.18	-1860.4	-1849.5
$\ln(R_t) \sim \mathrm{N}(\mu t, \sigma^2 t)$	−	−	+	+	2	924.92	-1843.8	-1832.9
$\ln(R_t) \sim \mathrm{S}_\alpha(0, c\sqrt{t}, \mu t)$	+	−	+	+	3	931.56	-1857.1	-1846.2
$\ln(R_t) \sim \mathrm{S}_\alpha(\beta, c\sqrt{t}, \mu t)$	+	+	+	+	4	933.68	-1859.4	-1844.8
Blue: Preferred model when $\mu = 0$ is considered.								
Red: Preferred model when $\mu \neq 0$ is assumed.								

When models that constrain $\mu = 0$ are considered, forward selection and backward elimination result in the selection paths as

- $\ln(R_t) \sim \mathrm{N}(0, \sigma^2 t) \to \ln(R_t) \sim \mathrm{S}_\alpha(0, c\sqrt{t}, 0) \to \ln(R_t) \sim \mathrm{S}_\alpha(\beta, c\sqrt{t}, 0)$, and
- $\ln(R_t) \sim \mathrm{S}_\alpha(\beta, c\sqrt{t}, \mu t) \to \ln(R_t) \sim \mathrm{S}_\alpha(\beta, c\sqrt{t}, 0)$,

at 5% error level, such that the symmetric stable model is selected in both directions. When $\mu \neq 0$ is assumed, forward selection and backward selection results in the selection paths as

- $\ln(R_t) \sim \mathrm{N}(\mu t, \sigma^2 t) \to \ln(R_t) \sim \mathrm{S}_\alpha(0, c\sqrt{t}, \mu t) \to \ln(R_t) \sim \mathrm{S}_\alpha(\beta, c\sqrt{t}, \mu t)$, and
- $\ln(R_t) \sim \mathrm{S}_\alpha(\beta, c\sqrt{t}, \mu t)$,

such that the full stable model is selected in both directions. Comparatively, AIC yields the same selection as the likelihood-ratio tests whether or not $\mu = 0$ is constrained. In contrast, BIC yields a more conservative model when μ is not constrained. When such discrepancies exist, selection of the most appropriate model would depend on the market views of different investors. If an investor has no prior forecast for the market, it may be reasonable to accept this lack of acumen, and apply the constraint $\mu = 0$, such that models with positive trends are only selected when it adequately exceeds the volatilities involved. In contrast, suppose that another investor expects the continuation of an expansionary economic policy, an attenuated covid status, or favorable post-covid market recovery, he or she may accept a positive μ, and discard all models with the constraint $\mu = 0$. The views on skewness should also be supported with the investor's risk appetite. Nonetheless, all the above-mentioned discussions show unequivocally that the normality assumption is not valid for S&P500 ETF, and investment returns demonstrate heavy tails as manifested in stability parameters much less than 2. This means that we should best obtain a detailed description of the return distributions for the stable model, and assess relevant deviations of VaRs with that from the normality assumption. Suppose that the stable model $\ln(R_t) \sim S_\alpha(\beta, c\sqrt{t}, \mu t)$ is chosen, and that we hold the ETF for one year with approximately $t = 250$. The comparative deciles for the stable and the normal models for the returns R_t are consolidated in Table 2.1.2.

Table 2.1.2. Comparative deciles for stable and normal models

Probability	Stable model decile	Normal model decile
0.9	37.0%	46.7%
0.8	30.1%	38.1%
0.7	25.3%	32.1%
0.6	21.4%	27.2%
0.5	17.7%	22.8%
0.4	14.1%	18.6%
0.3	10.2%	14.2%
0.2	5.60%	9.27%
0.1	-1.17%	2.80%

In this case, the normal distribution results in more optimistic predictions on returns, possibly due to its failure to capture skewness and heavy tails. We then proceed with the comparison of their 5% and 1% VaRs for different time periods, as shown in Table 2.1.3.

Table 2.1.3. VaR analysis for models

Time period	5% VaR by model		1% VaR by model	
	Stable	Normal	Stable	Normal
$t = 30$	-5.95%	-5.30%	-13.0%	-8.36%
$t = 60$	-7.33%	-6.07%	-17.0%	-10.0%
$t = 120$	-8.27%	-5.79%	-21.5%	-11.8%
$t = 240$	-7.65%	-2.61%	-25.9%	-11.2%

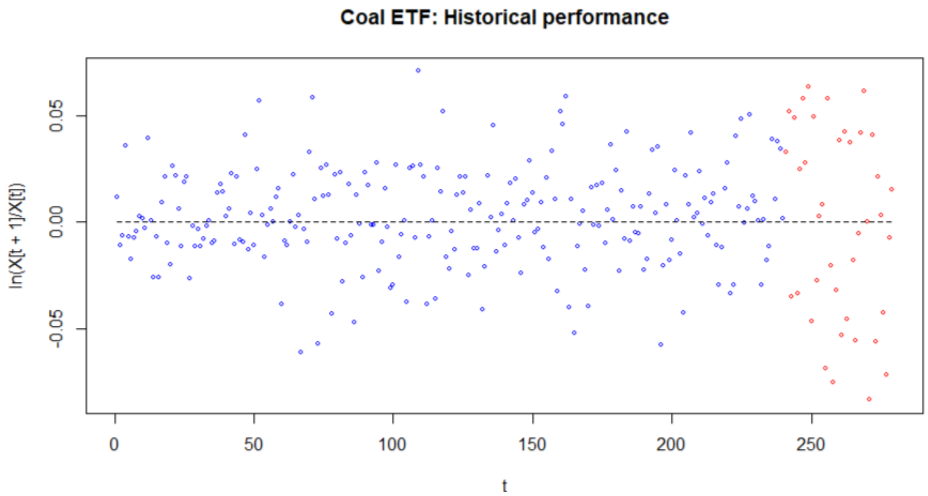
Similarly, normal distribution results in more optimistic predictions on VaRs, where the deviations are especially pronounced at 1% VaR. Therefore, from the perspective of investment risk, the stable model produces more conservative results on VaR analysis.

2.2. Case 2: Coal ETF

The Chinese coal market experienced a steady rally in the past year, as post-covid recovery of various industrial sectors created a steady increase in energy demands. Nonetheless, this trend was recently disrupted by drastic price turbulences from the interplay of an abrupt inadequacy in coal availability and the Chinese government's interventions to control coal prices. Therefore, one topic of interest is whether the model selection approaches discussed in

Section 1 have adequate power to identify the disruption, and provide a reasonable estimate of the previous trend that would expectedly continue should coal prices stabilize. The relevant data between 2020-09-02 and 2021-11-01 are displayed in Plot 2.2.1.

Plot 2.2.1. Coal ETF historical performance



The daily returns clearly show distinct trends in two periods. The core trend lasts until 2021-08-26, which corresponds to the steady rally amid post-covid recovery. Later than 2021-08-26, the movements become substantially more volatile, due to turbulent market expectations from inadequate coal supply and government intervention.

We then compute the relevant models similar to Case 1. Since we have adequate belief on the trend for post-covid recovery, we shall not impose the constraint for $\mu = 0$, but rather treat μ as a free parameter. Subsequent computations show that likelihood-ratio tests, AICs and BICs do not notice any substantial deviation from normality, such that we choose the model $\ln(R_t) \sim \mathrm{S}_2(0, 0.0189\sqrt{t}, 0.00203t) = \mathrm{N}(0.00203t, 0.000717t)$. Then, we attempt to compute separate stable models for the two periods to obtain

$$\ln(R_t)|t \leq 240 \sim \mathrm{S}_2(0, 0.0161\sqrt{t}, 0.00256t) = \mathrm{N}(0.00256t, 0.000517t),$$

$$\ln(R_t)|t > 240 \sim \mathrm{S}_2(0, 0.0311\sqrt{t}, -0.00123t) = \mathrm{N}(-0.00123t, 0.00193t)$$

for the core trend and the period of market turbulence respectively. The two periods have $\ell_{A,t\leq240} = 567.5$ and $\ell_{A,t>240} = 66.51$ each, where for the entire two-period model we have $\ell_A = 634.0$. In contrast, the one-period model has $\ell_0 = 614.9$. The likelihood-ratio statistic is therefore $2(\ell_A - \ell_0) = 39.66$ for χ_2^2 with almost zero p-value, which concludes that we need to model the core trend and the turbulent period separately. Model selection with AIC and BIC shows a similar superiority of the two-period model, with AIC and BIC as -1260 and -1246 respectively, compared to -1222 and -1211 for the one-period model. Therefore, should an investor believe that adequate interventions have been made to restore the trend for coal prices, the model for the core trend, $\ln(R_t)|t \leq 240 \sim \mathrm{N}(0.00256t, 0.000517t)$, should be used to forecast his or her future returns.

Conclusion

This small write-up has discussed the application of stable distributions to assess potential deviations of market data from normality. We focus on various methods of model selection such as likelihood-ratio tests, AIC and BIC to compare between normal and stable models, as well as between one-period and multiperiod models. The case studies demonstrate applications to the model selection methods, and shows how such results should always be combined with forecasts on market fundamentals. Generally, these statistical tools offer systematic procedures to obtain quantitative results for price movement trends, but should always be treated as a supplement, rather than the replacement, of fundamental analysis.