Estimation of Stable distribution and Its Application to Credit Risk

A Dissertation presented

by

Hua Mo

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the degree of

Doctor of Philosophy

in

Applied Mathematics and Statistics

Stony Brook University

Jan 2016

Stony Brook University

The Graduate School

Hua Mo

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation

Svetlozar(Zari) Rachev – Dissertation Advisor
Professor
Department of Applied Mathematics and Statistics

James Glimm – Chairperson of Defense
Distinguished Professor
Department of Applied Mathematics and Statistics

Aaron Kim – Dissertation Co-advisor
Assistant Professor
College of Business

Keli Xiao – Outside Committee Member
Assistant Professor
College of Business

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Estimation of Stable distribution and Its Application to Credit Risk

by

Hua Mo

Doctor of Philosophy

in

Applied Mathematics and Statistics

Stony Brook University

2016

To capture the heavy tails and the volatility clustering of asset returns is always an important topic in financial market. We studies two projects related to the Alpha Stable distribution and Classical Tempered Stable(CTS) distribution respectively which both have desired properties to accommodate heavy-tails and capture skewness in financial series. (1) In the major part of the first project, we introduce the algorithm of indirect inference method. By using the skewed-t distribution as an auxiliary model which is easier to handle, we can estimate the parameters of the Alpha Stable distribution since these two models have the same numbers of parameters and each of them plays a similar role. We also estimate of the parameters of the alpha stable distribution with McColloch method, Characteristic Function Based method and MLE method respectively. Finally, we provide an empirical application on S&P 500 returns and make comparisons between these four methods. (2) In the second project, we discuss the Gaussian firm value model and the Classical Tempered Stable firm value model. By pointing out the drawbacks of application of Merton's model on firm value, we introduce the classical tempered stable distribution and make the market firm value process follows a CTS distribution instead of Gaussian distribution. We estimate the parameters of the CTS, and calculate the firm value and default probability. By comparing these two models, the results suggest that CTS firm value model has a better potential to predict the default probability of a firm since it can better capture the heavy tails of the asset returns.

Dedicated to my beloved family

Table of Contents

List of Figures

List of Tables

Acknowledgements

Foremost, I would like to give my greatest thanks to Prof. Aaron Kim and Prof. Svetlozar (Zari) Rachev. I have been grateful that Prof. Rachev offered me the precious opportunity to work in his group since the beginning of my Phd study, and later introduced me to Prof. Aaron Kim, who becomes my main dissertation advisor. It's my great honor to have both of them as my advisors. The projects that I have worked on are all very exciting and challenging. My deepest thanks goes to my advisors, for their support and advice, for their patience, motivation, enthusiasm, immense knowledge, and for their dedication of a great amount of time of guidance.

Besides my advisors, I would like to give my thanks to my committee chair Prof. James Glimm, and my external committee member Prof. Keli Xiao for being supportive.

Last but not least, I want to thank my family for always being considerate and supportive for my PhD study, and our group members and my friends at Stony Brook University, who encouraged and supported me.

It has been a great journey for me to fulfill all the requirements for a Ph.D Degree!

Part I
Estimation of Alpha Stable Distribution

1 Introduction

The central limit theorem is one of the most important applications in statistics. It expresses the fact that a sum of many independent and identically distributed (i.i.d.) random variables, or alternatively, random variables with specific types of dependence, will tend to be a normal distribution regardless of the individual shape. The consequence of this result is that the normal distribution is widespread in statistical inference.

However, in some specific cases, especially in the field of finance, it's empirically observed that assets returns are heavy-tailed and leptokurtic. Mandelbrot and Fama suggested that the financial assets returns have stable non-Gaussian distributions in the 60's. This means the central limit theory fails and one should not expect the normal Gaussian distribution as a limit law in such case. The α-stable distribution accommodates heavy-tailed financial series and therefore produces more reliable measures of tail risk such as VaR. Besides, the α-stable distribution can capture the skewness of one distribution, which is another property of financial assets.

There are several equivalent ways to define the class of the α-stable distribution.

Definition 1. *A random variable X is said to have stable distribution if there is a positive number C_n and a real number D_n such that the sum of n independent copies of X, $X_1 + X_2 + \cdots + X_n$,has the same distribution as $C_n X + D_n$,that is*

$$X_1 + X_2 + \cdots + X_n \overset{d}{=} C_n X + D_n$$

Definition 2. *where $\overset{d}{\to}$ denotes convergence in distribution.*

Definition 3. *A random variable X is said to have a stable distribution if it has a domain of attraction, i.e. if there is a sequence of iid random variables $Y_1, Y_2, \cdots$ and sequence of positive numbers $\{d_n\}$ and real numbers $\{a_n\}$ such that*

$$\frac{Y_1 + Y_2 + ... + Y_n}{d_n} + a_n \xrightarrow{d} X$$

where $\xrightarrow{d}$ denotes convergence in distribution.

Definition 4. *A random variable X is said to have a stable distribution if there are parameters $0 < \alpha \leq 2, \sigma > 0, -1 \leq \beta \leq 1, \mu \in R$ such that its characteristic function (chf) has the following form*

$$\varphi(t) = Ee^{itX} = \{^{\exp\{-\sigma^\alpha|t|^\alpha(1-i\beta\frac{t}{|t|}\tan(\frac{\pi\alpha}{2}))+iut\},\alpha\neq 1}_{\exp\{-\sigma|t|(1+i\beta\frac{2}{\pi}\frac{t}{|t|}\ln(|t|))+iut\},\alpha=1}$$

where $\frac{t}{|t|} = 0$ if t=0. Here $\alpha \in (0, 2]$,measures the tail thickness, $\beta \in [-1, 1]$ measures the degrees of asymmetry, $\sigma > 0$ determines the scale and $\mu \in R$ determines the location. Since the parameters in the above equation is not continuous in the parameter space, it's usually recommended to use the following representation of the chf

$$\varphi(t) = Ee^{itX} = \{^{\exp\{-|\sigma^\alpha t^\alpha|+i\sigma t\beta(|\sigma t|^{\alpha-1}-1)\tan(\frac{\pi\alpha}{2}))+iu_1 t\},\alpha\neq 1}_{\exp\{-|\sigma t|+i\sigma t\beta\frac{2}{\pi}\ln|\sigma t|+iu_1 t\},\alpha=1}$$

where $0 < \alpha \leq 2, \sigma > 0, -1 \leq \beta \leq 1, \mu_1 \in R, \mu_1 = \{^{u+\beta\sigma\tan\frac{\pi\alpha}{2},\alpha\neq 1}_{u,\alpha=1}$

It's a nontrivial task to estimate the parameters of the α-stable distribution since there are no closed form expressions for the density functions except for some specific cases. In this paper we will talk about four parameter estimation techniques: quantile method, characteristic function based method, maximum likelihood and indirect inference. The quantile method is produced by $FamaandRoll$ and modified by $McCulloch$. The Chf based methods include several methods, in this paper we just talk about regression type estimator of Kogon-Williams. The maximum likelihood estimation (MLE) theory was first proposed by DuMouchel. We emphasized the method of the indirect inference. It is a method particularly suited to situations where the model of interest is difficult to estimate but relatively easy to simulate. Indirect inference involves the use of an auxiliary model. Auxiliary parameters are estimated through maximization likelihood method and also have a one-to-one correspondence with the parameters of the stable distribution.

This report organized as follows, Section 2 introduce the four parameters estimation techniques: McCulloch, characteristic function based method, maximum likelihood estimation and indirect inference. To emphasize the application of the indirect inference, we set an example and detail its application. Section 3 shows the empirical illustration. We resample the data and compare the spread of the parameters with different methods. We also fit the sample data pdf and the estimated data pdf and test the fitness. Finally we compare the tail event between these four methods respectively. Section 4 make a conclusion, and section 5 envision future research on this topic. Section 6 is the acknowledgement and section 7 is the reference.

2. Parameters Estimation Techniques

2.1 Quantile Method of McCulloch

In fact, this estimation method can be seen as a specific case of indirect inference. According to McCulloch, the four parameters of a stable distribution can be estimated from five predetermined sample quantiles, for $0.6 < \alpha \leq 2$ and $-1 \leq \beta \leq 1$.Suppose we have n independent variables X_i from the the stable distribution whose parameters are to be estimated. Denote x_p is the p-th quantile if F(x_p)=p, where F(x) is the cdf of a random variable. Let $\hat{x}_p$be the corresponding sample quantile. Define

$$v_\alpha = \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}}$$

This index v_α is independent of σ and μ. let $\hat{v}_\alpha$be the corresponding sample value

$$\hat{v}_\alpha = \frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.75} - \hat{x}_{0.25}}$$

The statistic $\hat{v}_\alpha$is the continuous estimator of v_α.Define

$$v_\alpha = \phi_1(\alpha, \beta)$$

$$v_\beta = \frac{x_{0.95} + x_{0.05} - 2x_{0.50}}{x_{0.95} - x_{0.05}}$$

and let $\hat{v}_\beta$ be the corresponding sample and

$$\hat{v}_\beta = \frac{\hat{x}_{0.95} + \hat{x}_{0.05} - 2\hat{x}_{0.50}}{\hat{x}_{0.95} - \hat{x}_{0.05}}$$

The statistic $\hat{v}_\beta$ is the continuous estimator of v_β. The value of v_α and v_β as functions of $\phi_1(\alpha, \beta)$ and $\phi_2(\alpha, \beta)$ are tabulated in Table 1 and 2. Table 1 and 2 are derived from DuMouchel's tabulation [1971] of the stable distribution.

Table 1: $v_\alpha = \phi_1(\alpha, \beta)$

α \ β	0.00	0.25	0.50	0.75	1.00
2.00	2.439	2.439	2.439	2.439	2.439
1.90	2.512	2.512	2.513	2.513	2.515
1.80	2.608	2.609	2.610	2.613	2.617
1.70	2.737	2.738	2.739	2.742	2.746
1.60	2.912	2.909	2.904	2.900	2.902
1.50	3.148	3.136	3.112	3.092	3.089
1.40	3.464	3.436	3.378	3.331	3.316
1.30	3.882	3.834	3.720	3.626	3.600
1.20	4.447	4.365	4.171	4.005	3.963
1.10	5.217	5.084	4.778	4.512	4.451
1.00	6.314	6.098	5.624	5.220	5.126
0.90	7.910	7.590	6.861	6.260	6.124
0.80	10.448	9.934	8.779	7.900	7.687
0.70	14.838	13.954	12.042	10.722	10.370
0.60	23.483	21.768	18.332	16.216	15.584
0.50	44.281	40.137	33.002	29.140	27.782

Note that $\phi_1(\alpha, \beta) = \phi_1(\alpha, -\beta)$

Table 2: $v_\beta = \phi_2(\alpha, \beta)$

		β				
		0.00	0.25	0.50	0.75	1.00
	2.00	0.00	0.00	0.00	0.00	0.00
	1.90	0.00	0.018	0.036	0.053	0.071
	1.80	0.00	0.039	0.077	0.113	0.148
	1.70	0.00	0.063	0.123	0.178	0.228
	1.60	0.00	0.089	0.174	0.248	0.309
	1.50	0.00	0.118	0.228	0.320	0.390
	1.40	0.00	0.148	0.285	0.394	0.469
α	1.30	0.00	0.177	0.342	0.470	0.546
	1.20	0.00	0.206	0.399	0.547	0.621
	1.10	0.00	0.236	0.456	0.624	0.693
	1.00	0.00	0.268	0.513	0.699	0.762
	0.90	0.00	0.303	0.573	0.770	0.825
	0.80	0.00	0.341	0.634	0.834	0.881
	0.70	0.00	0.387	0.699	0.890	0.927
	0.60	0.00	0.441	0.768	0.936	0.962
	0.50	0.00	0.510	0.838	0.970	0.985

Note that$\phi_2(\alpha, \beta) = -\phi_2(\alpha, \beta)$

The relationship

$$\begin{aligned} v_\alpha &= \phi_1(\alpha, \beta) \\ v_\beta &= \phi_2(\alpha, \beta) \end{aligned}$$

may be inverted to produce the relationship

$$\begin{aligned} \alpha &= \psi_1(v_\alpha, v_\beta) \\ \beta &= \psi_2(v_\alpha, v_\beta) \end{aligned}$$

The parameter α and β may be estimated by

$$\begin{aligned} \hat{\alpha} &= \psi_1(\hat{v}_\alpha, \hat{v}_\beta) \\ \hat{\beta} &= \psi_2(\hat{v}_\alpha, \hat{v}_\beta) \end{aligned}$$

The function $\psi_1(v_\alpha, v_\beta)$ and $\psi_2(v_\alpha, v_\beta)$ are tabulated in Tables 3 and 4.

Table 3: $\alpha = \psi_1(v_\alpha, v_\beta)$

ν_β

v_α		0	0.1	0.2	0.3	0.5	0.7	1
	2.5	1.916	1.924	1.956	1.998	2	2	2
	3.5	1.391	1.386	1.364	1.337	1.319	1.300	1.291
	4.5	1.193	1.187	1.166	1.129	1.093	1.066	1.053
	5.5	1.074	1.068	1.050	1.015	0.976	0.944	0.927
	6.5	0.988	0.983	0.969	0.937	0.901	0.866	0.849
	7	0.957	0.951	0.938	0.905	0.872	0.836	0.815
	9	0.857	0.852	0.841	0.814	0.784	0.752	0.730
	10	0.818	0.812	0.801	0.780	0.758	0.720	0.691
	12	0.765	0.760	0.751	0.729	0.701	0.676	0.655
	16	0.687	0.683	0.677	0.663	0.642	0.610	0.586
	20	0.640	0.636	0.628	0.615	0.596	0.576	0.555
	25	0.593	0.590	0.586	0.579	0.563	0.541	0.513

Note that $\psi_1(v_\alpha, v_\beta) = \psi_1(v_\alpha, -v_\beta)$

Table 4: $\beta = \psi_2(v_\alpha, v_\beta)$

v_β

v_α		0	0.1	0.2	0.3	0.5	0.7	1
	2.5	0	1	1	1	0	0	0
	3.5	0	0.165	0.499	0.943	1	1	1
	4.5	0	0.119	0.355	0.596	1	1	1
	5.5	0	0.102	0.302	0.498	0.731	1	1
	6.5	0	0.091	0.272	0.453	0.662	1	1
	7	0	0.088	0.259	0.435	0.639	1	1
	9	0	0.078	0.231	0.391	0.597	1.968	1
	10	0	0.074	0.220	0.377	0.546	0.874	1
	12	0	0.070	0.207	0.354	0.501	0.788	1
	16	0	0.063	0.189	0.324	0.470	0.708	1
	20	0	0.060	0.177	0.303	0.446	0.677	1
	25	0	0.056	0.167	0.285	0.428	0.650	1

Note that $\psi_2(v_\alpha, v_\beta) = -\psi_2(v_\alpha, -v_\beta)$

Since there are no closed form expression for the functions $\psi_1(v_\alpha, v_\beta)$ and $\psi_2(v_\alpha, v_\beta)$,we estimate α and β from $\hat{v}_\alpha$and $\hat{v}_\beta$according to the Table 3 and 4. Observing the tables we can find that when $\hat{v}_\alpha$is below 2.439 $\hat{\alpha}$ will be set equal to 2 and $\hat{\beta}$ equal to 0.

McCulloch also provide the estimation formula of v_σ:

$$v_\sigma = \frac{x_{0.75} - x_{0.25}}{\sigma} = \phi_3(\alpha, \beta)$$

The function of $\phi_3(\alpha, \beta)$ is given in Table 5:

Table 5: $v_\sigma = \phi_3(\alpha, \beta)$

	β				
	0	0.25	0.50	0.75	1.00
2.00	1.908	1.908	1.908	1.908	1.908
1.90	1.914	1.915	1.916	1.918	1.921
1.80	1.921	1.922	1.927	1.936	1.947
1.70	1.927	1.93	1.943	1.961	1.987
1.60	1.933	1.94	1.962	1.997	2.043
1.50	1.939	1.952	1.988	2.045	2.116
1.40	1.946	1.967	2.022	2.106	2.211
α 1.30	1.955	1.984	2.067	2.188	2.333
1.20	1.965	2.007	2.125	2.294	2.491
1.10	1.98	2.04	2.205	2.435	2.696
1.00	2	2.085	2.311	2.624	2.973
0.90	2.04	2.149	2.461	2.886	3.356
0.80	2.098	2.244	2.676	3.265	3.912
0.70	2.189	2.392	3.004	3.844	4.775
0.60	2.337	2.635	3.542	4.808	6.247
0.50	2.588	3.073	4.534	6.636	9.144

Note that $\phi_3(\alpha, \beta) = \phi_3(\alpha, -\beta)$

Replacing α and β with the estimator found according to the above method, we can get the estimator $\hat{\sigma}$ by table 5 and the following formula:

$$\hat{\sigma} = \frac{\hat{x}_{0.75} - \hat{x}_{0.25}}{\phi_3(\hat{\alpha}, \hat{\beta})}$$

McCulloch also provides estimator for the location parameter μ :

$$v_u = \frac{u - x_{0.5}}{\sigma}$$

However, when $\alpha \to 1$ and $\beta \neq 0$,there will be an discontinuity and makes the linear interpolation highly incorrect. To solve this problem, McCulloch proposes the following native location parameter:

$$\begin{aligned} \zeta &= \{^{u+\beta\sigma\tan\frac{\pi\alpha}{2},\alpha\neq 1}_{u,\alpha=1} \Rightarrow \\ u &= \{^{\zeta-\beta\sigma\tan\frac{\pi\alpha}{2},\alpha\neq 1}_{\zeta,\alpha=1} \end{aligned}$$

Table 6 shows the behavior of $v_\zeta = \phi_4(\alpha, \beta)$:

Table 6: $v_\zeta = \phi_4(\alpha, \beta)$

		β				
		0.00	0.25	0.50	0.75	1.00
	2.00	0.00	0.00	0.00	0.00	0.00
	1.90	0.00	-0.017	-0.032	-0.049	-0.064
	1.80	0.00	-0.03	-0.061	-0.092	-0.123
	1.70	0.00	-0.043	-0.088	-0.132	-0.179
	1.60	0.00	-0.056	-0.111	-0.17	-0.232
	1.50	0.00	-0.066	-0.134	-0.206	-0.283
	1.40	0.00	-0.075	-0.154	-0.241	-0.335
α	1.30	0.00	-0.084	-0.173	-0.276	-0.39
	1.20	0.00	-0.09	-0.192	-0.31	-0.447
	1.10	0.00	-0.095	-0.208	-0.346	-0.508
	1.00	0.00	-0.098	-0.223	-0.383	-0.576
	0.90	0.00	-0.099	-0.237	-0.424	-0.652
	0.80	0.00	-0.096	-0.25	-0.469	-0.742
	0.70	0.00	-0.089	-0.262	-0.52	-0.853
	0.60	0.00	-0.078	-0.272	-0.581	-0.997
	0.50	0.00	-0.061	-0.279	-0.659	-1.198

Note that $\phi_4(\alpha, \beta)$=$\phi_4(\alpha, -\beta)$

The parameter now can be estimated by

$$\hat{\zeta} = \hat{x}_{0.5} + \hat{\sigma}\phi_4(\hat{\alpha}, \hat{\beta})$$

So we can get the estimator of the location parameter by

$$\hat{u} = \{^{\hat{\zeta} - \hat{\beta}\hat{\sigma}\tan\frac{\pi\hat{\alpha}}{2}, \alpha \neq 1}_{\hat{\zeta}, \alpha = 1}$$

2.2 Chf Based Method

Given a sequence of independent and identically distributed (i.i.d.) random variables $x_{1,}x_{2,}...x_n$, define the sample characteristic function by

$$\hat{\varphi}(t) = \frac{1}{n}\sum_{j=1}^{n} e^{itx_j}, t \in R$$

Since $\left|\hat{\varphi}(t)\right|$ is bounded by unity all moments of $\hat{\varphi}(t)$ which are finite and, for any fixed , it is the sample average of i.i.d. random variables . Hence, by the law of large numbers, it is a consistent estimator of the chf.

Here we will talk about the regression-type estimator. Regression-type estimator method is first presented by Koutrouveils. This method starts with an initial estimate of the parameters and proceeds iteratively until some pre-specified convergence criterion is satisfied. Each iteration consists of two weighted regression runs. The number of points to be used in these regressions depends on the sample size and starting values of α.However, this technique involves a lot of computations because of the required iterations. Then Kogon and William suggest a new method to eliminated this iteration procedure and reduce the amount of computations.

The linear equations directly follows from the form of the logarithm of the chf:

$$\ln[-\Re(\ln\varphi(t))] = \alpha\ln\sigma + \alpha\ln|t|$$

$$\Im(\ln\varphi(t)) = u_1 t + \beta\sigma t(|\sigma t|^{\alpha-1} - 1)\tan\frac{\pi\alpha}{2}$$

The parameter estimators can be constructed using the method of least squares after replacing the chf for the sample chf. Kogon and Williams give that the best choice is $\mathrm{t}_k = \{0.1 + 0.1k, k = 0, 1, ...9\}-$ they are 10 equally spaced points in the interval [0.1,1]. The algorithm is as follows:

1. We first use the technique of McCulloch for the preliminary estimator of σ_0 and u_{01}. Once the initial σ_0 and u_{01} have been found, the data sample $\{x_1,x_2,...,x_n\}$ are normalized:

$$x_j^{'} = \frac{x_j - \hat{u}_{01}}{\hat{\sigma}_0}, j = 1, 2, ..., n$$

2. let y_k=ln$[-\Re(\ln \varphi(t))], w_k = \ln |t_k|$ where $t_k = \{0.1 + 0.1k, k = 0, 1, ...9\}$ and ε_k denotes the error term, then we get the regression equation:

$$y_k = b + \alpha w_k + \varepsilon_k, k = 0, 1, ...9$$

Using the normalized sample $\{x_n^{'}\}$ we can find the $\hat{\alpha}$ and $\hat{b}$ according to the method of OLS.

3. Let $z_k = \Im(\ln \varphi(t)), v_k = \hat{\sigma}_1 t_k(\left|\hat{\sigma}_1 t_k\right|^{\hat{\alpha}-1} - 1)\ \tan \frac{\pi\hat{\alpha}}{2}\ , t_k = \{0.1 + 0.1k, k = 0, 1, ...9\}$,then

$$z_k = u_{11} t_k + \beta v_k + \eta_k, k = 0, 1, ...9$$

The skewness parameter $\hat{\beta}$ and the modified location parameter $\hat{u}_{11}$can be estimated from the above formula

4. The final estimate of the spread and modified location parameters are found to be

$$\begin{aligned} \hat{\sigma} &= \hat{\sigma}_0 \hat{\sigma}_1, \\ \hat{u}_1 &= \hat{u}_{01} + \hat{\sigma}_0 \hat{u}_{11} \end{aligned}$$

and the final estimate of the location parameter is from the estimator of $\hat{u}_1$ by

$$\hat{u} = \hat{u}_1 - \hat{\beta}\hat{\sigma} \tan \frac{\hat{\alpha}\pi}{2}$$

2.3 Maximum Likelihood

The method of maximum likelihood is very popular in many applications due to the good asymptotic properties of the estimates. The likelihood function is defined as

$$L(x_1,x_2,...x_n \mid \theta) = \prod_{k=1}^{n} f(x_k \mid \theta)$$

where $\{x_{1,}x_{2,}...x_n\}$ is a sample of i.i.d observations of a r.v. X. $f(x \mid \theta)$ is the pdf of X, and θ is a vector of parameters. MLE are found by searching for the parameter values which can maximize the likelihood function, that is, to maximize the log-likelihood function:

$$\overset{\wedge}{\theta}_n = \arg\max_{\theta} \log(L(x_{1,}x_{2,}...x_n \mid \theta))$$

If we want to derive the expression for MLE analytically we need know the close-form pdf of stable laws. Here we use the FFT-method to do the transform and estimate the density function.

2.4 Indirect Inference

First introduces by Smith and Gourieroux, Monfort and Renault (1993), indirect inference is an inferential approach which is quite suited for the situation where the models of interest is difficult to estimate but relatively easy to simulate.

The underlying principle is as followed: suppose we have a sample of T observations y and a model whose likelihood function $L^*(y;\theta)$ is difficult to handle and maximize. The maximum likelihood estimate of $\theta \in \Theta$,given by

$$\overset{\wedge}{\theta} = \max_{\theta\in\Theta} \sum_{t=1}^{T} \ln L^*(\theta; y_t),$$

is unavailable. Let us take an alternative model, depending on a parameter vector $\zeta \in Z$,which is indicated as auxiliary model and easier to estimate, suppose we use this model in place of the original one. Since the model is unspecified, the ML-estimator

$$\overset{\wedge}{\zeta} = \max_{\overset{\wedge}{\zeta}\in Z} \sum_{t=1}^{T} \ln \tilde{L}(\zeta; y_t),$$

is unnecessarily consistent: the idea is to exploit simulations performed under the original model to correct the inconsistency.

The first step is to compute the quasi MLE of ζ,which is demoted as $\overset{\wedge}{\zeta}$. Next, one simulates a set of S vectors of size T from the original model on the basis of an arbitrary parameters vectors $\overset{\wedge}{\theta}^{(0)}$. Let us denote each one of those vectors as y$^{(s)}(\hat{\theta}^{(0)})$.The simulated values are then estimated using the auxiliary model, yielding

$$\overset{\wedge}{\zeta}(\hat{\theta}^{(0)}) = \max_{\zeta \in Z} \sum_{s=1}^{S} \sum_{t=1}^{T} \ln \tilde{L}[\zeta; y^{(s)}(\hat{\theta}^{(0)})].$$

The idea is to numerically update the initial guess $\hat{\theta}^{(0)}$ in order to minimize the distance

$$[\overset{\wedge}{\zeta} - \overset{\wedge}{\zeta}(\theta)]'[\overset{\wedge}{\zeta} - \overset{\wedge}{\zeta}(\theta)],$$

The indirect inference estimators are consistent and asymptotically normal under certain conditions. The most difficult part is to establish the binding function which maps the parameters of the auxiliary model onto the true model–alpha-stable distribution. The auxiliary model we decide to use is the skewed-t distribution introduced by Azzalini&Capitanio in 2003, since it has similarities to the alpha-stable distribution. Table 7 is the relationship between these parameters.

Table 7: Relation structural and auxiliary parameters

characteristic	Structural	Auxiliary
Tail thickness	α	ν
Skewness	β	γ
Scale	σ	λ
Location	μ	ω

The skewed t distribution has the pdf as follows:

$$f(x;\nu,\gamma,\lambda,\omega) = \frac{\Gamma(1/2+\nu)}{\rho\lambda(\pi\nu)^{1/2}\Gamma(\nu)(\frac{|x-\omega+m|^2}{\nu(\rho\lambda)^2(\gamma sign(x-\omega+m)+1)^2}+1)^{1/2+\nu}}$$

where

$$m = \frac{2\rho\lambda\gamma\nu^{1/2}\Gamma(\nu-1/2)}{\pi^{1/2}\Gamma(\nu+1/2)}$$

and

$$\rho = \frac{1}{\nu^{1/2}\sqrt{(3\gamma^2+1)(\frac{1}{2\nu-2})-\frac{4\gamma^2}{\pi}(\frac{\Gamma(\nu-1/2)}{\Gamma(\nu)})^2}}$$

In this paper, we do a little change on the algorithm, the steps are as follows:

1. Use an alternative model, depending on a parameter vector ς,which is indicated as an auxiliary model who has closed form density function and is easier to estimate. The auxiliary we use here is skewed t distribution.

2. The sample data is then estimated using the auxiliary model with MLE method, yielding

$$\hat{\varsigma} = (\nu, \gamma, \lambda, \omega)$$

then we can get the probability density function $f(x)$. According to the FFT, we can also get the pdf of alpha stable distribution $g(x)$.

3. Find the parameter $\theta = (\alpha, \beta, \sigma, \mu)$ in order to minimize the distance

$$\sqrt{\sum_{X \in P} |f(x) - g(x)^2|}$$

where P=$\{x_1, x_{2,} ... x_N\}$.

In our application , we will estimate the parameters of the return distribution of S&P 500 index with daily price from 1/1/2009 to 7/27/2012. There are totally 901 daily returns.

First step, we use MLE to estimate the parameters of skew-t distributions. The results are as following:

$$\{\upsilon = 2.7988, \gamma = 0.2682, \lambda = 0.0088, \omega = -0.0030\}$$

Secondly, we find the α-stable distribution parameters which minimize the distance between probability density of the skewed-t distribution and alpha-stable distribution, which is $\sqrt{\sum_{X \in P} |f(x) - g(x)^2|}$.

The numerical estimation result as following:

$$\{\alpha = 1.0001, \beta = 0.2682, \sigma = 5.0050, \mu = -0.0032\}$$

3. Comparison between Methods

We want to compare these four methods on simulated sample with different sample size. In our simulation,we generate different size alpha-stable samples with fixed parameters, and apply these methods to estimate the parameters of alpha-stable distribution. I set $\alpha = 1.2, \beta = 0.5, \sigma = 1, \mu = 0$.

3.1 Compare the Spread

To test the stabilities of the outcomes, I resample 100 groups of the estimated stable parameters. The following are the boxplot diagrams:

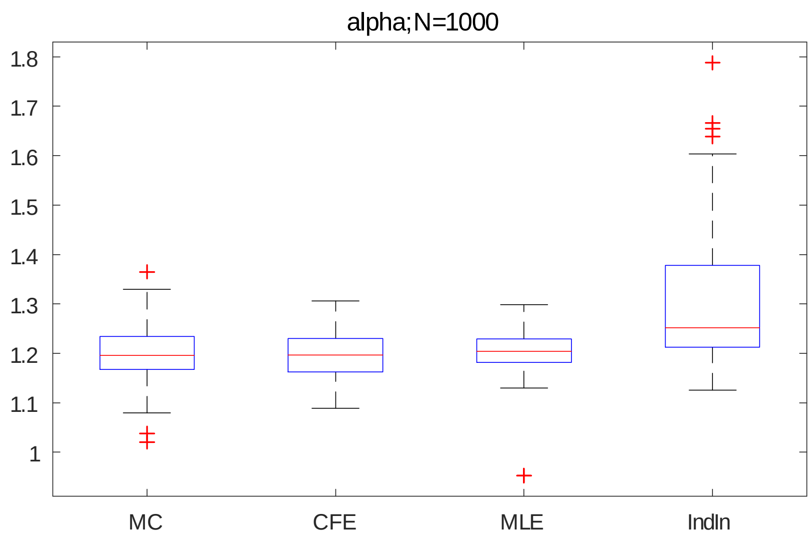


Figure 1: Boxplot Diagrams of α,N=1000

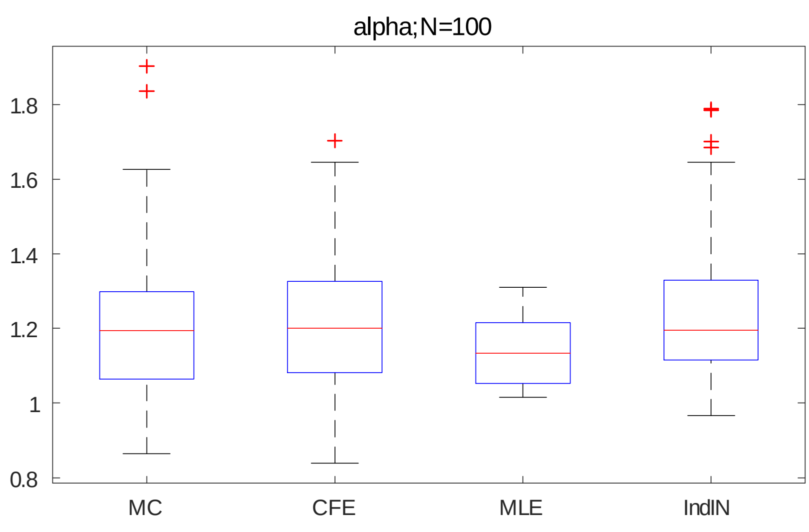


Figure 2: Boxplot Diagrams of α,N=100

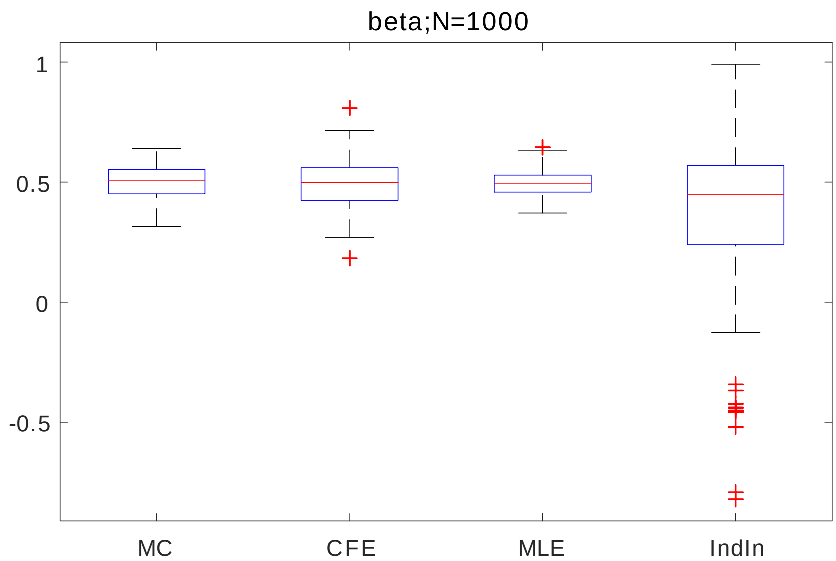


Figure 3: Boxplot Diagrams of β,N=1000

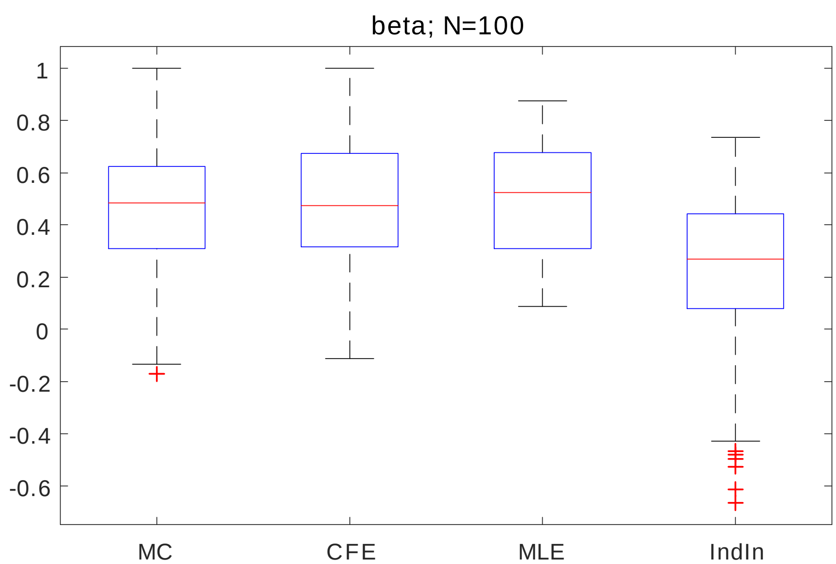


Figure 4: Boxplot Diagrams of β,N=100

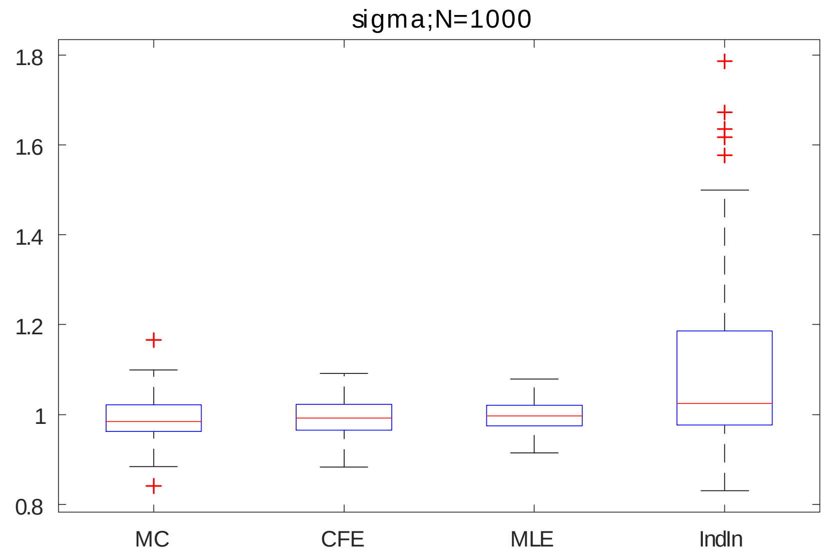


FIGURE 1. Figure 5: Boxplot Diagrams of σ,N=1000

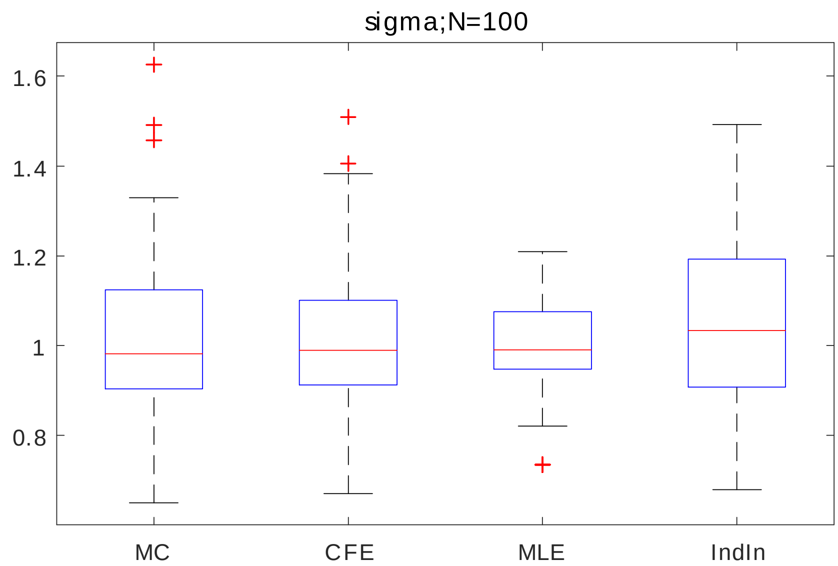


Figure 6: Boxplot Diagrams of σ,N=100

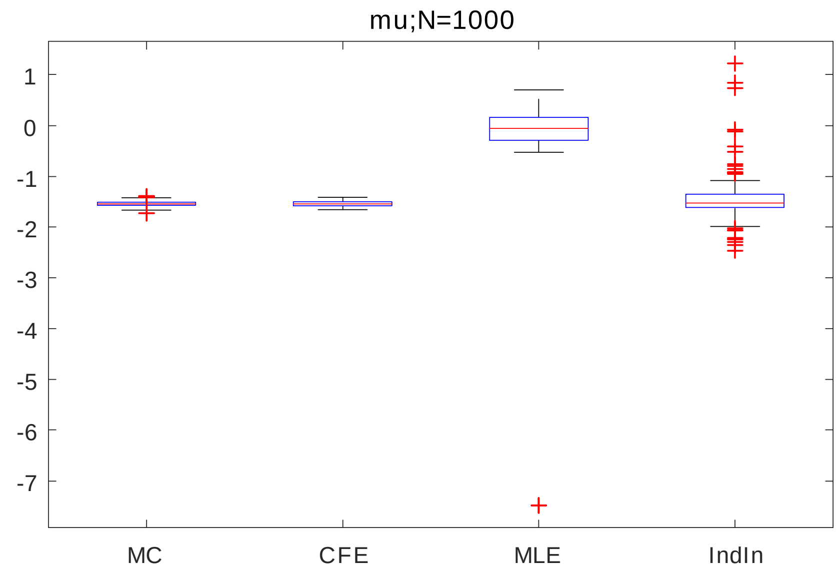


Figure 7: Boxplot Diagrams of μ,N=1000

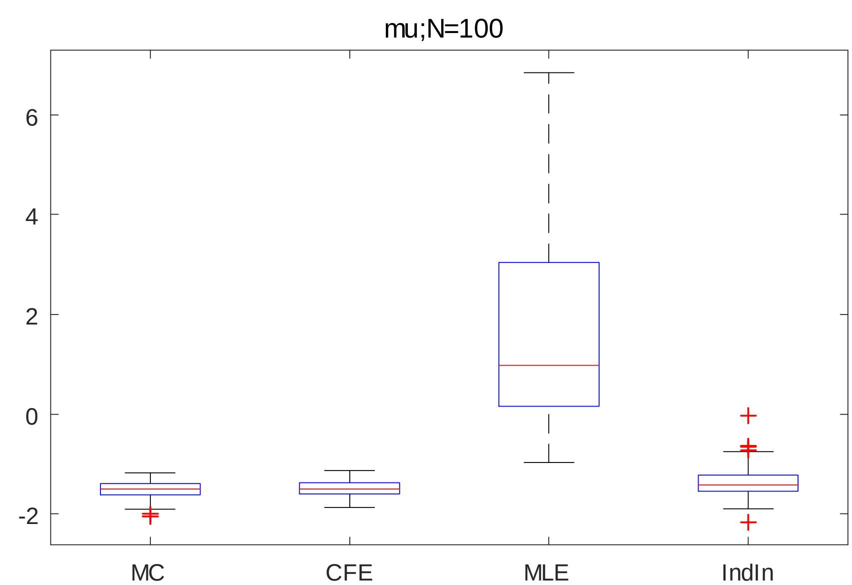


Figure 8: Boxplot Diagrams of μ,N=100

Table 8 : Mean and Standard Derivation of Estimate Parameters
According to the Four Methods; N=1000

	MC	CFE	MLE	IndIn
mean of α	1.1961	1.1956	1.2015	1.3025
std of α	0.0590	0.0464	0.0492	0.1384
mean of β	0.4999	0.4925	0.4939	0.3434
std of β	0.0740	0.1057	0.0597	0.3861
mean of σ	0.9875	0.9932	0.9971	1.1054
std of σ	0.0495	0.0444	0.0346	0.1953
mean of μ	-1.5424	-1.5421	-0.1829	-1.4050
std of μ	0.0554	0.0534	0.0512	0.5693

Table 9 : Mean and Standard Derivation of Estimate Parameters
According to the Four Methods; N=100

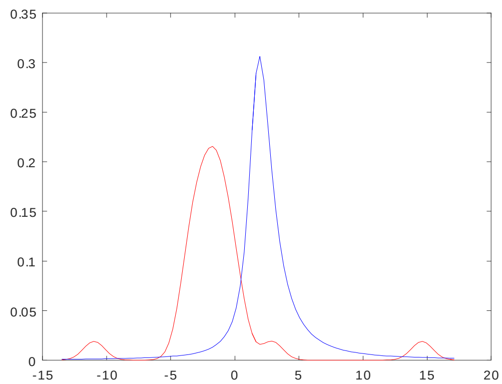
	MC	CFE	MLE	IndIn
mean of α	1.1974	1.2094	1.2056	1.2449
std of α	0.1925	0.1771	0.1665	0.1861
mean of β	0.4742	0.4901	0.4743	0.2005
std of β	0.2496	0.2767	0.2225	0.3448
mean of σ	1.0091	1.0107	1.0006	1.0510
std of σ	0.1753	0.1616	0.1212	0.1966
mean of μ	-1.5230	-1.5058	1.7595	-1.3621
std of μ	0.1854	0.1648	0.1512	0.3188

According to the above figures and tables, we can see that all the methods have better performance when sample size is getting larger. Compare figure 1 and figure 2, it's very obvious that when the sample size is large, MLE appears to be the most superior approach, the mean is very close to 2 and the standard deviation is small. MC and CFE also work well. The estimator of IndIn has the worst performance when the size is large, in terms of the mean and the standard deviation. MLE work less well when the sample size is small, also do MC and CFE. But IndIn works much better when the sample size is small compared with large size. Observing figure 3 and 4, figure 5 and 6, figure 7 and 8, there are same situations for β, σ and μ.

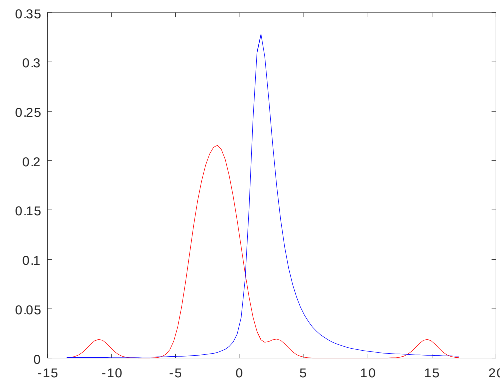
3.2 Compare the Fitness

To compare these methods more directly and exam how they perform in the tail exam, I am going to set a specific example in this paper. Here we also set $\alpha = 1.2, \beta = 0.5, \sigma = 1, \mu = 0$.In our simulation, we generate different sample size of the fixed parameters, and apply MC, CFE, MLE and indirect inference to estimate parameters respectively.

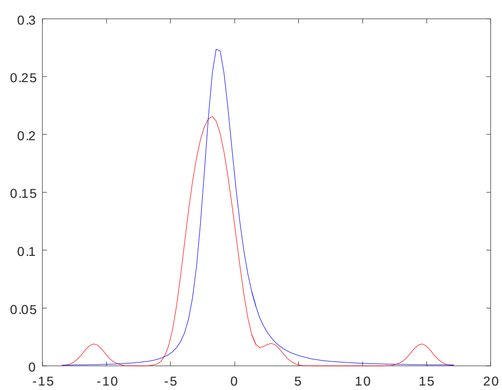
3.2.1 Fit the pdf in Figures



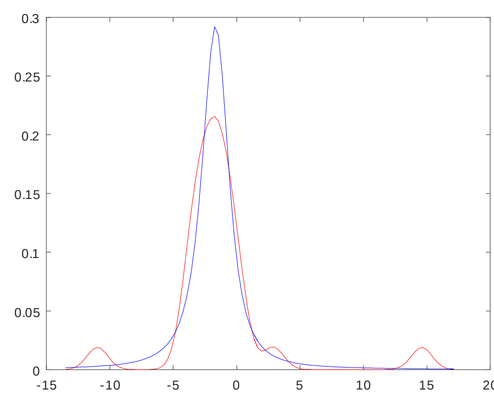
MC, N=25



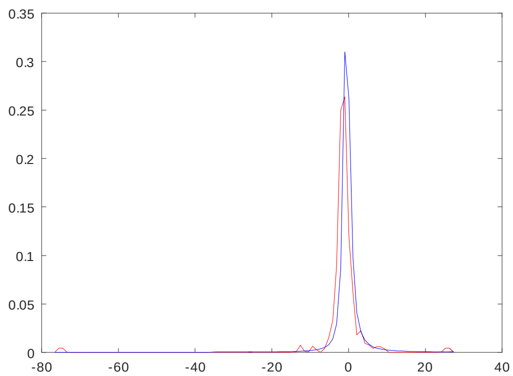
CFE, N=25



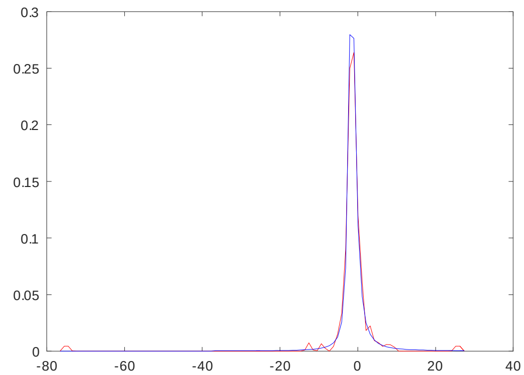
MLE, N=25



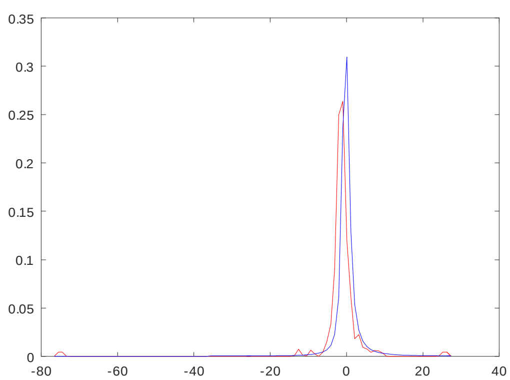
Indirect Inference, N=25



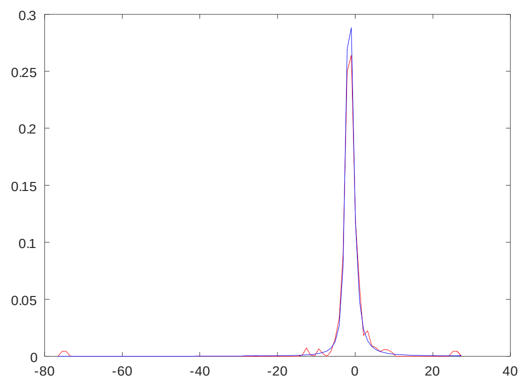
MC, N=100



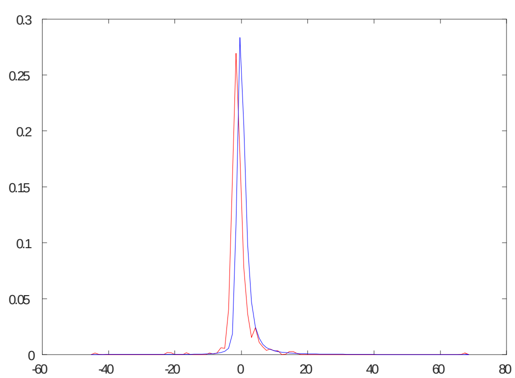
CFE, N=100



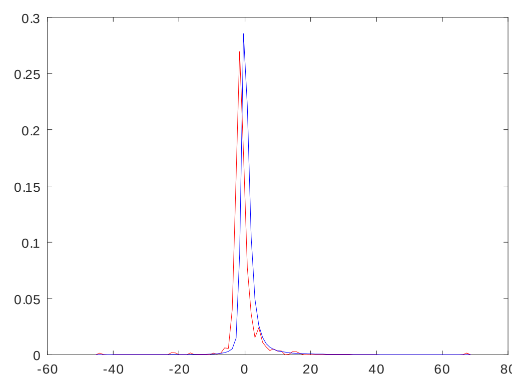
MLE, N=100



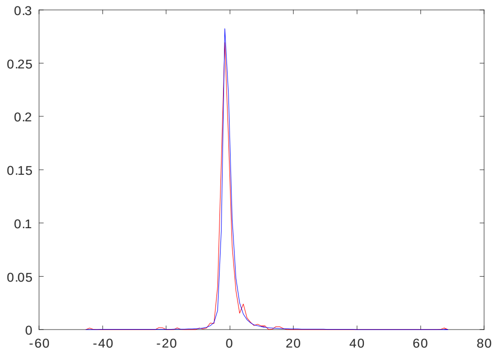
Indirect Inference, N=100



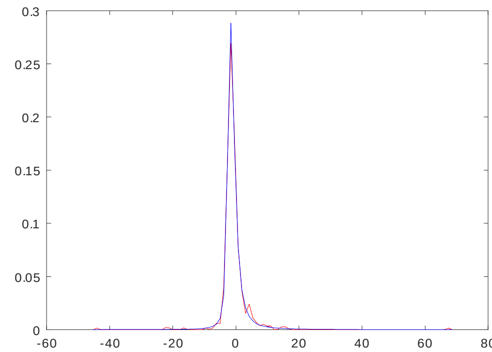
MC, N=500



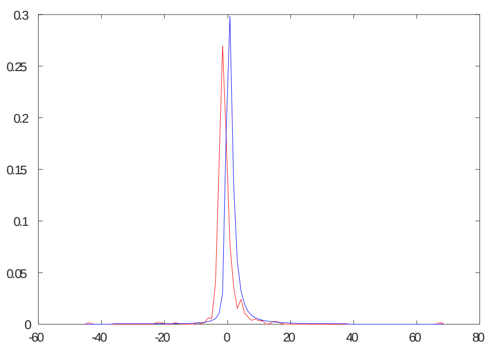
CFE, N=500



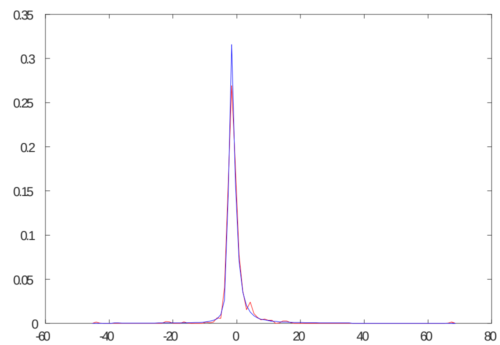
MLE, N=500



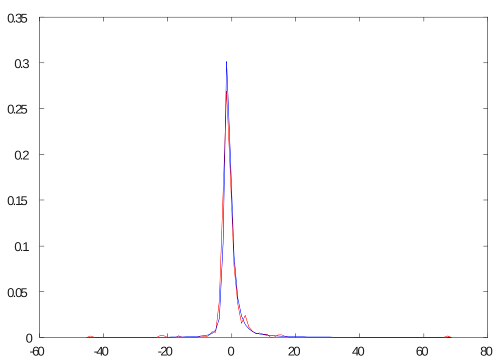
Indirect Inference, N=500



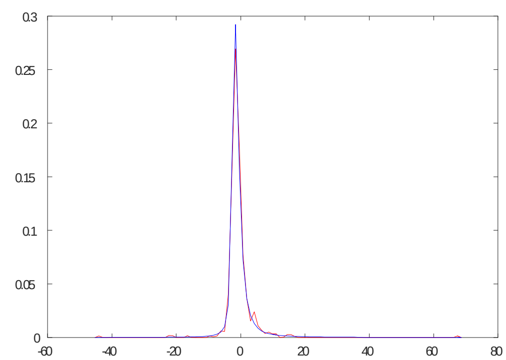
MC, N=1000



CFE, N=1000



MLE, N=1000



Indirect Inference, N=1000

Figure 9: Simulated Sample and Estimated Sample by Using MC, CFE, MLE, IndIn Fitting pdf with Different Sample Size N

3.2.2 Goodness of Fit

To compare the goodness of fit for the estimated CTS distribution, we employ two statistics: Kolmogorov-Smirnov distance statistic(KS-statistic) and Anderson-Darling statistic (AD - statistic).

Kolmogorov-Smirnov distance statistic is defined as follows:

$$KS = \sup_x \left| \hat{F}(x) - F_x(x) \right|$$

where $\hat{F}(x)$ is the empirical sample distribution and $F_x(x)$ is the cumulative distribution function of the estimated theoretical distribution. The KS-statistic emphasize deviation around the median of the fitted distribution. It is a robust measure in the sense that it focuses only on the maximum deviation between empirical and estimated distributions.

Anderson-Darling statistic can test the ability to model extreme events. In simplest version, it is a variance-weighted KS statistic:

$$AD = \sup_x \frac{\left| \hat{F}(x) - F_x(x) \right|}{\sqrt{F_x(x)(1 - F_x(x))}}$$

The AD-statistic measures the distance between the empirical and theoretical distribution functions but is rescaled by dividing the "standard deviation" of some distance. By this definition, the AD-statistic accentuates more discrepancies in the tail. The following are the outcomes of the fitness with McCulloch, Cfe, MLE and indirect inference method.

Table 10: Comparison in Terms of Goodness of Fit; N=25

	MC	CFE	MLE	IndIn
KS	0.8104	0.8617	0.2459	0.0787
AD	2.8454	4.5611	0.6137	0.3251

Table 11: Comparison in Terms of Goodness of Fit; N=100

	MC	CFE	MLE	IndIn
KS	0.3186	0.0536	0.3921	0.0568
AD	0.7173	0.2237	0.9719	0.2767

Table 12: Comparison in Terms of Goodness of Fit; N=500

	MC	CFE	MLE	IndIn
KS	0.3738	0.4072	0.1164	0.0403
AD	1.1289	1.3189	0.2455	0.0933

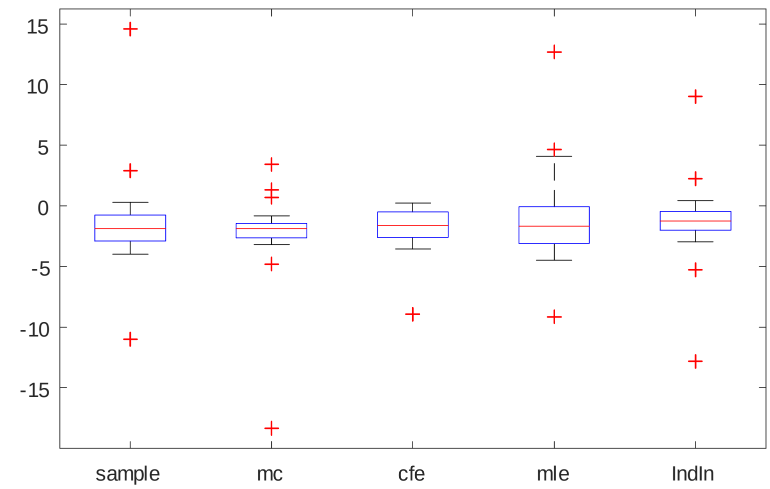
Table 13: Comparison in Terms of Goodness of Fit; N=1000

	MC	CFE	MLE	IndIn
KS	0.5526	0.0704	0.0382	0.0727
AD	0.2390	0.4565	0.1592	0.8613

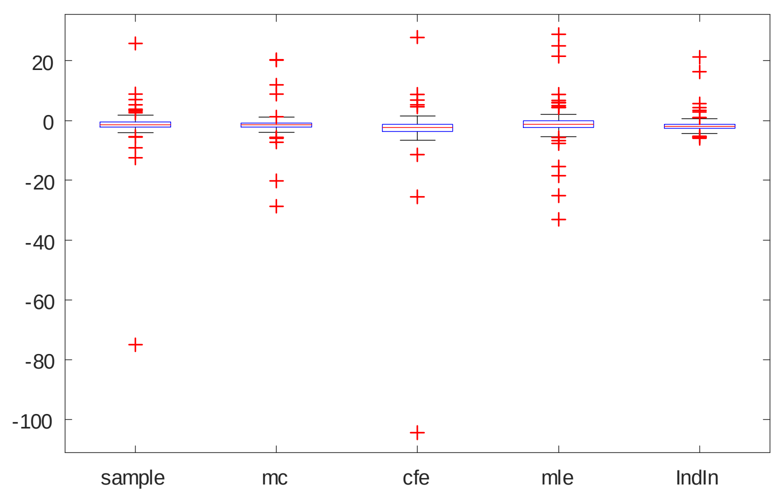
In figure 9, the blue line indicates the simulated sample distribution and the red line indicated the estimated sample distribution. We can see that all the methods have better performance when sample size is getting larger. The estimator of McCulloch has the worst overall performance no matter the sample size in terms of the Kolmogorov distance and Anderson-Darling statistics. CFE and MLE don't perform very well when the size is small, we can see that both from the figure and the fitness outcome. But when the sample becomes large, MLE has the most superior performance. The pdf fit very well, the KS and AD are both very small. Indirect inference perform the best when the sample is small, it also works well when the sample is large, but still not as good as MLE.

3.3 Compare the Heavy Tails

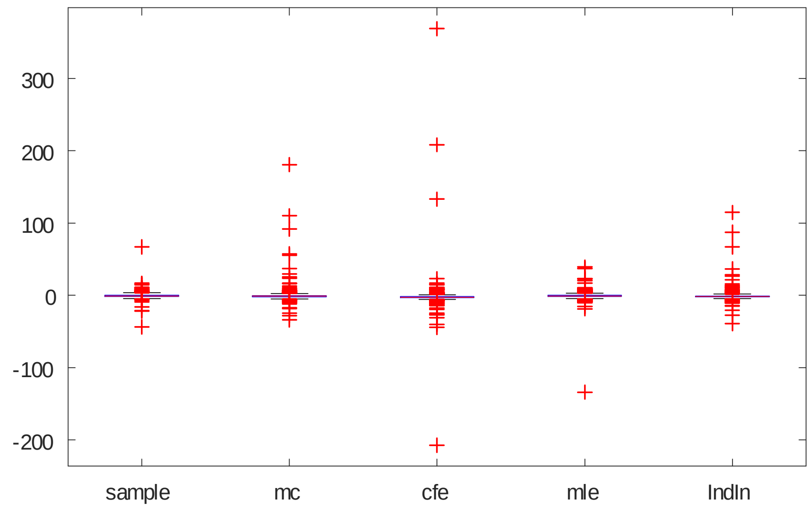
However, the above figures can't tell the performance in the tail event, that's why we will compare the boxplot of the samples.



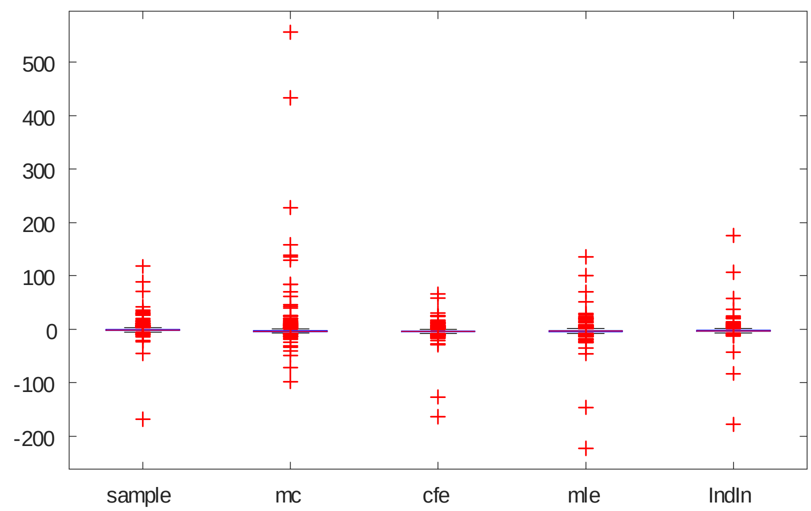
(a) N=25



(b) N=100



(C) N=500



(d) N=1000

Figure 10: Box-plots of Simulated Sample, Estimated Sample by MC,CFE,MLE and Indirect Inference on Different Sample Size N

In figure 10, (a) N=25, we can see that indirect inference (the fifth column) has most closest lower and upper whisker to simulated sample(the first column), compared to the other three methods. When sample size is getting larger, MC, CFE, MLE and IndIn they all have better performance. But obviously, when the sample size is large, eg, in (d) N=1000, we can see that MLE has closest lower and upper whisker to simulated sample ,compared to MC, CFE and indirect inference. MC doesn't work well no matter when the size is small and large. CFE performs better than MC, but worse than MLE when the size is large and worse than Indirect inference when the size is small. Therefore, indirect inference has advantage when the sample size is small and MLE has advantage when the sample size is large.

4.Conclusion

The stable distribution is very useful to model financial asset returns which usually have heavy-tails and skewed characteristic. However, it is very hard to estimate their parameters because it does not have closed form of density function. In this paper, we propose four methods to estimate the parameters. The McCulloch method uses the quantile calculation. There are also some characteristic function based methods rely on the sample chf for parameter estimation, in this paper we use the called Kogon-Williams method, whose main idea is to generate regression-type estimators. MLE is the most popular estimation method people are using. Since the density function of a stable distribution does not have a closed form but a stable series is relatively easy to simulate, we propose an indirect inference estimator method which is suited to such characteristics. In this paper, we use the skewed-t distribution as the auxiliary model. We apply this method to S&P500 data and our numerical result shows that it has very good performance.

This paper also provide comparison between the McCulloch method, the Kogon-Williams method, MLE and indirect inference method with different sample size. First I resample the data 100 times, and compare the boxplot of the four parameters. The results show that MC and CFE have relatively less diversity no matter what the size is, the MLE method has very small diversity for the large sample size and the indirect inference has less diversity when the size is small. Then we compare the fitness between simulated sample and estimated sample, the box-plots of simulated sample and estimated sample by the four methods on different sample size. MC has the worst performance on all the tests, then does CFE. For small sample size, indirect inference has better performance both on the fitness and heavy tail; For larger sample size, MLE has better performance.

5 Future Work

One of the most promising fields of applications of stable distributions is that of time series models. As one can in fact note, several empirical phenomenons that are observed over time exhibit asymmetry and leptokurtosis.

Definition 5. *Let* $(S_t)_{t\geq 0,}$ *be the asset price process and* $(y_t)_{t\geq 0}$ *be the return process of* $(S_t)_{t\geq 0}$ *defined by* $y_t = \log \frac{S_t}{S_{t-1}}$. *We propose the ARMA(1,1)-GARCH(1,1) model:*

$$\begin{cases} y_t = a y_{t-1} + b\sigma_{t-1}\varepsilon_{t-1} + \sigma_t \varepsilon_t + c \\ \sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{cases}$$

where $\varepsilon_0 = 0$,and a sequence $(\varepsilon_t)_{t\in N}$ is a sequence of independent and identically distributed real random variables. The innovation ε_t is assumed to follow the standard normal distribution. In this case, the ARMA(1,1)-GARCH(1,1) model is referred to as the normal-ARMA-GARCH model.

When the innovation ε_t is assumed to follow the alpha stable distribution, the ARMA(1,1)-GARCH(1,1) model is referred to as the α-stable-ARMA-GARCH model. α-stable-ARMA-GARCH model : ε_t ~α-stable$(\alpha, \beta, \sigma, \mu)$

We also want to use MLE method and Indirect inference method to estimate the parameters. With the MLE method, the parameters are estimated as follows:

1: Estimate parameters with student-t distributed innovation by the MLE.

2: Extract residuals using the estimated parameters.

3: Fit the parameters of the innovation distribution (alpha stable distribution) to the extracted residual using the MLE.

With the Indirect Inference method, the idea one could pursue is to use as auxiliary model the skewed-t distribution analog of the "true" model of interest, e.g. for an α-stable-ARMA(1,1)-GARCH(1,1) an auxiliary skewed-t-ARMA(1,1)-GARCH(1,1) model is used. That is to assume ε_t is a standard skewed-t distribution, estimate the parameters of skewed-t-ARMA-GARCH model, and then estimate the alpha stable distribution parameters by indirect inference method which we proposed in chapter 2.

6 References

[1] Stoyan Stoyanov, Borjana Rachev Iotova(2003). "Univariate stable laws in the field of finance-parameter estimation". FinAnalytica.

[2] Stoyan Stoyanov, Borjana Rachev Iotova(2003). "Univariate stable laws in the field of finance-approximations of density and distribution functions". FinAnalytica.

[3] Rene Garcia, Eric Renault, David Veredas(2011). "Estimation of Stable Distributions by Indirect Inference". Journal of Banking&Finance.

[4] Young Shin Kim, Svetlozar T. Rachev, Michele Leonardo(2011). "Time series analysis for financial market meltdowns". Journal of Banking&Finance.

[5] Marco J. Lombardi, Giorgio Calzolari(2007). "Indirect estimation of alpha-stable distribution and process". Journal of Statistics and Econometrics.

[6] W. H. DuMouchel(1971). "Stable Distributions in Statistical Inference", Ph.D. dissertation, Department of Statistics, Yale University.

[7] J.Huston McCulloch(1986). "Simple Consistent Estimators Of Stable Distributions Parameters". Communications in Statistics-Simulation and Computation.

[8] Gourieroux, C.Monfort, A&Renault, E.(1993), "Indirect inference". Journal of Applied Econometrics.

[9] R.j.Adler, R.Feldman and M.Taqqu(1998). "A practical Guide to Heavy Tails". Statistical Techniques and Applications.

[10] V.Akgiray, C.G. Lamoureux(1989). "Estimation of Stable-law Parameters: A Comparative study". Journal of Business and Economic Statistics.

Part II

Classical Tempered Stable Firm Value Model

1. Introduction

1.1 What's Credit Risk?

Credit risk is one of the oldest forms of risk and has been become one of the most popular studied topics in quantitative finance. Credit risk refers to the risk that a borrower will default on a debt by failing to make required payments, it's associated with any kind of credit-linked events, such as: changes in the credit quality, variations of credit spread and the default event, which occurs if the debtor is unable to meet its legal obligation according to the debt contract.

In the last decade from July 1990 through March 1991, United states experienced a recession which recorded the lowest economic growth rate since the Great Depression, In the mid-1980s United States experienced record defaults on bank loans and corporate bonds. The junk bond defaults jumped to over 10 percent between the years 1990 and 1991. Economists still hold different opinions on the reasons of the recession but credit risk is undoubtedly considered as a major factor. Banks then realized from the real estate default experience that factors like diversification, liquidity and regulatory changes such as risk-based capital requirements were becoming increasingly important. The factors are viewed as very important in credit risk management. In recent years, more and more banks and corporations realized that the Basel regulation do not manage credit risk adequately, therefore they have built complex mathematical-statistical models as a substitution for quantifying credit risk. These models are used to determine internal economic capital to protect them from credit risks as well as to play important roles in risk management and performance measurement processes, including performance-based compensation, customer profitability analysis, risk-based pricing and, to a lesser but growing degree, active portfolio management and capital structure decisions.

1.2 Credit Risk Management: Credit Derivatives

In recent times there have been developed many instruments to transfer credit risk. These instruments are called credit derivatives. A credit derivative is a financial instrument that transfers credit risk related to an underlying entity or a portfolio of underlying entities from one party to another without transferring the underlying. The underlyings may or may not be owned by either party in the transaction. The market for credit derivatives was created in the early 1990s in London and New York and it is the fastest growing derivative market at the moment. Here we introduce some common types of credit derivatives in the current financial market.

The first one is credit default swap (CDS). It is a financial swap agreement that the seller of the CDS will compensate the buyer (usually the creditor of the reference loan) in the event of a loan default (by the debtor) or other credit event. This is to say that the seller of the CDS insures the buyer against some reference loan defaulting. The buyer of the CDS makes a series of payments (the CDS "fee" or "spread") to the seller and, in exchange, receives a payoff if the loan defaults. It was invented by Blythe Masters from JP Morgan in 1994. The value of a default swap depends not only on the credit quality of the underlying reference entity but also on the credit quality of the writer, also referred to as the counterparty. If the counterparty defaults, the buyer of a default swap will not receive any payment if a credit event occurs. We also note that if a counterparty defaults, the premium payments end. Hence, the value of a default swap depends on the probability of counterparty default, probability of entity default and the correlation between them.

The other one is total return swap. It's a swap agreement in which one party makes payments based on a set rate, either fixed or variable, while the other party makes payments based on the return of an underlying asset, which includes both the income it generates and any capital gains. In total return swaps, the underlying asset, referred to as the reference asset, is usually an equity index, loans, or bonds. This is owned by the party receiving the set rate payment. The most important difference between a TRS and a CDS is the matter of isolating credit risk. While the default risk of a CDS is completely isolated, a TRS transfers both credit and market risk.

1.3 Credit Risk Modelling Approaches

There are three main approaches to credit risk modelling in financial market. They range from practical market methods to theory methods which rely on firm value. The first approach is the firm value models. Under the structure models of the firm value, a default event is deemed to occur for a firm when its assets reach a sufficiently low level compared to its liabilities. These models requires strong assumptions on the dynamics of the firm's asset, its debt and how its capital is structured. In this section, we will mainly focus on firm value models. We will also introduce other evaluation method of credit risk, such as intensity based model and rating based model in the next chapter. Besides these method, there have been also developed some credit derivatives such as: Total Return Swaps (TRS) and Credit Default Swaps (CDS).

1.3.1 Firm Value Model

The firm value models go back to the initial proposal of Black and Scholes (1973), It's the oldest literature to the pricing of defaultable securities in modern continuous-time finance. In their model, a defaultable security is regarded as a contingent claim on the value of the issuing firm's assets and is valued according to option pricing theory. The firm's value is assumed to follow a diffusion process and default is modeled as the first time the firm's value hits a pre-specified boundary. Because of the continuity of the processes used, the time of default is a predictable stopping time. The payoff in default is usually a constant cash payment representing the proceeds from liquidating the firm (possibly after bankruptcy costs). Two classic structural models: the Merton's model(Merton, 1974) and the first-passage-time model (Black and Cox, 1976) are inspired by the Black-Scholes model and are the representatives of this approach. Here we mainly focus on the Merton's model.

Although Merton's model is a great application of option price method, it has some drawbacks:

1. Skewness is not considered in the firm value model. Considering skewness, two firms with same leverage and same volatility may have different default probabilities.

2. Driving process for the firm value is Brownian Motion. Therefore, it can't capture the fat-tail property and big jump of firm value process.

Because of these limitations which could not address the credit risk accurately, in this paper we will introduce and apply a new firm value model: Tempered stable firm value model. In this model, the driving process for the firm value is the CTS. To avoid repetition, we will discuss the details later.

1.3.2 Intensity Based Model

In the intensity models the time of default is modeled directly as the time of the first jump of a Poisson process with random intensity (a Cox process). This approach is first proposed by Jarrow and Turnbull (1995), Madan and Unal (1998). Unlike the firm value models, assumes that the defaults are endogenous and ultimately determined by the asset value, the defaults in the intensity-based model are exogenous and determined by an externally specified intensity process that may or may not be related to the asset value. Jarrow and Turnbull consider the simplest case where the default is driven by a Poisson process with constant intensity with known payoff at default. This is changed in the Madan and Unal model where the intensity of the default is driven by an underlying stochastic process, and the payoff in default is a random variable drawn at default, it is not predictable before default. Many follow-up papers can be found in Lando (2004), Bielecki and Rutkowski (2004).

In the intensity-based model, we can express the bond price as

$$\overline{B}(t,T) = \overline{B}(t, N_t, r_t, T)$$

where N_t is the number of defaults in [0, T] in the portfolio. N_t will be modeled by the Non-homogeneous Poisson Process.

Definition 6. *N_t is s a (simple, homogeneous) Poisson process with an intensity $\lambda_t = \lambda(t) > 0, t \geq 0$,if*

i N(0)=0

ii It has independent and stationary increments.

iii $P(N_{t+s}-N_t=k)=\frac{(\int_t^{t+s}\lambda(u)du)^k}{k!}e^{-\int_t^{s+t}\lambda(u)du}$: the probability of k defaults in $[t,t+s]$.

In the Cox process, λ_t is an Ito process with mean reverting property on the risk-neutral world:

$$d\lambda_t=\mu_\lambda(t)dt+\sigma_\lambda(t)dW_t^\lambda$$

The default free interest rate is also an Ito process on the risk-neutral world:

$$dr_t=\mu_r(t)dt+\sigma_r(t)dW_t^r$$

By Ito formula and Arbitrage Pricing Theory (APT), we obtain:

$$\frac{\partial\overline{B}}{\partial t}+\frac{\partial\overline{B}}{\partial r}\mu_r(t)+\frac{1}{2}\frac{\partial^2\overline{B}}{\partial r^2}\sigma_r^2(t)-\overline{B}(t,T)(\lambda_t+r_t)=0$$

When $\lambda_t=0$, we can get the the PDE for default free bond B(t,T):

$$\frac{\partial B}{\partial t}+\frac{\partial B}{\partial r}\mu_r(t)+\frac{1}{2}\frac{\partial^2 B}{\partial r^2}\sigma_r^2(t)-B(t,T)r_t=0\Rightarrow$$

$$B(t,T)=E_t(e^{-\int_t^T r_u du})$$

Therefore, the solution of the defaultable bond is

$$\overline{B}(t,T)=B(t,T)e^{-\int_t^T\lambda_u du}$$

In terms of the yield, we have

$$B(t,T)=e^{-Y_{t,T}(T-t)};\overline{B}(t,T)=e^{-\overline{Y}_{t,T}(T-t)}$$

where $Y_{t,T}$ is the yield of default free bond during $[t,T]$, and $\overline{Y}_{t,T}$ is the yield of defaultable bond during time $[t,T]$ .Then we can get the spread

$$S(t,T)=\overline{Y}_{t,T}-Y_{t,T}=\frac{1}{T-t}(\ln B(t,T)-\ln\overline{B}(t,T))$$

$$=\frac{1}{T-t}\int_t^T\lambda_u du$$

There are some other popular intensity-based models derived from the term-structure models, with reference to Vasicek (1977), Cox, Ingersoll and Roll (1985) and Duffie, Pan and Singleton (2000):

$$\begin{aligned}\text{Vasicek:}\quad & d\lambda(t)=a(b-\lambda(t))dt+\sigma_1 dW_t\\ \text{Cox-Ingersoll-Roll:}\quad & d\lambda(t)=a(b-\lambda(t))dt+\sigma_2\sqrt{\lambda_t}dW_t\end{aligned}$$

$$\text{Affinejump: } d\lambda(t) = a(b - \lambda(t))dt + \sigma_3 dW_t + dJ_t$$

The constant a, b, σ_1, σ_2 should be calibrated from the defaultable term structure of interest rate. In the last model, $\sigma_3 = \sqrt{\sigma_1^2 + \sigma_2^2 \lambda_t}$,and J_t is an independent compound Poisson process with constant jump intensity and independent exponentially distributed jumps with mean u. These models are attractive in finance because they could yield closed-form pricing formulas and provide straightforward ways to simulate the future default intensity for the purpose of predicting the conditional default probability.

1.3.3 Rating Based Model

In the last decade, rating based models in credit risk manegement have become very popular. These systems use the rating of a company as the decisive variable and not-like the formerly used structural models the value of the firm-when it comes to evaluate the default risk of a bond or loan. The popularity is due to the straightforwardness of the approach but also to the new Capital Accord (Basel II) of the Basel Committee on Banking Supervision (2001), a regulatory body under the bank of International Settlements (BIS). Basel II allows banks to base their capital requirement on internet as well as external rating system.

The Moody's KMV is a very famous internal rating platform which can accommodate a wide variety of risk rating models. Specifically, it is used to deploy risk models across a bank's local and global lending network and to manage the data requirements of the credit rating process. Banks around the world are making the platform a critical component of their credit risk process as they prepare for Basel II IRB compliance. Moody's KMV believes that all lending institutions, whether or not they are required to comply with the new regulations, should carefully consider the credit policy issues and recommendations raised by the Basel Committee in the context of managing risk and measuring customer profitability. Moody's KMV offers several products and services that help banks develop solutions that meet regulatory requirements and contribute to improved business performance:

1. Moody's KMV Risk Advisor$^{\text{TM}}$ and RiskAnalyst$^{\text{TM}}$ allow institutions to create and deploy internal rating models based on both quantitative and qualitative criteria.

2. Moody's KMV BaselCalc$^{\text{TM}}$ gives banks an optimized calculation of required regulatory capital across any reporting jurisdiction.

3. Moody's KMV RiskCalc® Moody's KMV CreditEdge® and Moody's KMV Credit Monitor® produce Expected Default Frequency$^{\text{TM}}$ (EDF$^{\text{TM}}$) credit measures for individual borrowers.

4. Moody's KMV LossCalc$^{\text{TM}}$ produces estimates of loss in the event of default.

5. Moody's KMV Professional Services provides implementation, modeling, portfolio advisory, and benchmarking services for the banking industry.

2. Term Structure Models

2.1 Introduction

The topic of term-structure modelling for derivatives pricing has been discussed in recent years, it has a big significance to have a better understanding of the term structure of interest rate since the following reasons. Firstly, a qualitatively new dimension has been added to the modelling complexity, because of the appearance of pronounced and complex smiles in the implied volatility surfaces of caplets and swaptions. Currently, the modelling of smiles is one of the most active areas of research in interest-rate derivatives pricing, and a sufficiently amount of work has accumulated to warrant a review of its achievements and of the problems still to be tackled.Secondly, we need a term structure of interest rates to embody in the shape of the forward curve, for example, fixed income instruments typically depend on a forward curve instead of a single point. Thus pricing such instruments needs a model that can describe a stochastic time evolution of the entire forward curve. There exists a large number of term structure models based on different choices of state variables parameterizing the curve, number of dynamic factors, volatility smile characteristics, etc. In this chapter, we mainly focus on the short rate models including Hull-White model and Cox-Ingersoll-Ross(CIR) model.

2.2 Short Rate Models

A short-rate model is a mathematical model that describes the future evolution of interest rates by describing the future evolution of the short rate r(t). It can be classified into equilibrium and non-arbitrage model. Equilibrium models are also referred to as endogenous term structure models because the term structure if interest rate is an output of these models. If we have a initial zero coupon bond curve, the parameters of the equilibrium model will be calculated such that a zero coupon bond curve as close as possible to the one observed in the market can be reproduced. Vasicek model(1977) is a classic representative of equilibrium models. However, since the equilibrium models have difficulties in producing the initial condition, non-arbitrage models are built to make up this shortcoming.

We apply a zero bond maturing T and paying 1 at T for all the models. According to the risk-Neutral pricing formula, its value at time t is

$$B(t,T) = E[e^{-\int_t^T r(u)du} |\mathcal{F}_t|]$$

2.2.1 The Hull-White Model

The Hull-White model is an non-arbitrage model in which the term structure of interest rate is an input. The Hull and White model is defined that

$$dr(t) = (a(t) - b(t)r(t))dt + \sigma(t)dW_t$$

where a(t), b(t) and $\sigma(t)$ are non-random functions. The time deterministic function a(t) is chosen so that the model fits the initial term structure of the interest rate. we suppose $\psi(t) = e^{\int_0^t b(s)ds}$, $\phi(t) = \psi(t)^{-1} = e^{-\int_0^t b(s)ds}$,then according to the It$\hat{o}$-Doeblin formula, we have

$$\begin{aligned} d(r(t)\psi(t)) &= r(t)d\psi(t) + \psi(t)dr(t) + d\psi(t)dr(t) \\ &= r(t)\psi(t)b(t)dt + \psi(t)a(t)dt - \psi(t)b(t)r(t)dt + \psi(t)\sigma(t)dW_t \\ &= \psi(t)a(t)dt + \psi(t)\sigma(t)dW_t \end{aligned}$$

Integrate both sides of the formula, we get

$$\begin{aligned}
r(t)\psi(t) &= r(0)\psi(0) + \int_0^t \psi(s)a(s)ds + \int_0^t \psi(s)\sigma(s)dW_s \Longrightarrow \\
r(t) &= \phi(t)(r(0) + \int_0^t (\phi(s))^{-1}a(s)ds + \int_0^t (\phi(s))^{-1}\sigma(s)dW_s)
\end{aligned}$$

The expectation of the r(t) is

$$E[r(t)] = \phi(t)(r(0) + \int_0^t (\phi(s))^{-1}a(s)ds)$$

and the variance of r(t) is

$$Var[r(t)] = (\phi(t))^2 \int_0^t (\phi(s))^{-2}(\sigma(s))^2 ds$$

Since

$$\begin{aligned}
E[\int_0^T r(s)ds] &= \int_0^T E[r(s)]ds \\
&= \int_0^T (\phi(s)(r(0) + \int_0^s (\phi(u))^{-1}a(u)du))ds \\
&= r(0)\int_0^T \phi(s)ds + \int_0^T (\phi(u))^{-1}a(u)(\int_u^T \phi(s)ds)du
\end{aligned}$$

$$\begin{aligned}
Var[\int_0^T r(s)ds] &= Cov(\int_0^T r(s)ds, \int_0^T r(s)ds) \\
&= \int_0^T \int_0^T Cov(r(u), r(v))dudv \\
&= \int_0^T \int_0^T \phi(u)\phi(v) \int_0^{u\wedge v} (\phi(s))^{-2}(\sigma(s))^2 dsdudv) \\
&= \int_0^T (\phi(s))^{-2}(\sigma(s))^2 (\int_s^T \phi(u)du)(\int_s^T \phi(v)dv)ds \\
&= \int_0^T (\phi(s))^{-2}(\sigma(s))^2 (\int_s^T \phi(u)du)^2 ds
\end{aligned}$$

We can get the bond price

$$\begin{aligned}
B(0,T) &= E[e^{-\int_0^T r(s)ds}] \\
&= e^{-E[\int_0^T r(s)ds] + \frac{1}{2}Var(\int_0^T r(s)ds)} \\
&= e^{-C(0,T)r(0) - A(0,T)}
\end{aligned}$$

where

$$C(0,T) = \int_0^T \phi(s)ds$$

$$A(0,T) = \int_0^T (\phi(s))^{-1} a(s)(\int_s^T \phi(u)du)ds - \frac{1}{2}\int_0^T (\phi(s))^{-2}(\sigma(s))^2(\int_s^T \phi(u)du)^2 ds$$

Because r(t) in the Hull and White model is normally distributed, the interest rate can be negative. The CIR model is designed to avoid this kind of shortcoming.

2.2.2 Cox-Ingersoll-Ross (CIR) Model

The CIR model for the interest rate r(t) is

$$dr(t) = (a - br(t))dt + \sigma\sqrt{r(t)}dW_t$$

The advantage of CIR model over Vasicek model and the Hull-White model is that the interest rate can not be negative. However, it doesn't have a closed form solution. Consider the price at time $t \in [0,t]$ of a zero coupon bond

$$B(t,T) = E[e^{-\int_t^T r(u)du} \mid \mathcal{F}_t]$$

We have

$$\begin{aligned}
& d(e^{-\int_0^t r(u)du}B(r(t),t,T)) \\
= \quad & e^{-\int_0^t r(u)du}(-r(t))B(r(t),t,T)dt + e^{-\int_0^t r(u)du}dB(r(t),t,T) \\
= \quad & e^{-\int_0^t r(u)du}(-r(t))B(r(t),t,T)dt + e^{-\int_0^t r(u)du}(B_t^{'}(r(t),t,T)dt + B_{r(t)}^{'}(r(t),t,T)dr(t) + \\
& \frac{1}{2}B_{r(t)r(t)}^{''}(r(t),t,T)dr(t)dr(t)) \\
= \quad & e^{-\int_0^t r(u)du}(-r(t))B(r(t),t,T)dt + e^{-\int_0^t r(u)du}(B_t^{'}(r(t),t,T)dt + (a-br(t))B_{r(t)}^{'}(r(t),t,T)dt + \\
& B_{r(t)}^{'}(r(t),t,T)\sigma\sqrt{r(t)}dW_t + \frac{1}{2}B_{r(t)r(t)}^{''}(r(t),t,T)\sigma^2 r(t)dt) \\
= \quad & e^{-\int_0^t r(u)du}[-r(t)B(r(t),t,T) + (B_t^{'}(r(t),t,T) + (a-br(t))B_{r(t)}^{'}(r(t),t,T) + \\
& \frac{1}{2}B_{r(t)r(t)}^{''}(r(t),t,T)\sigma^2 r(t)]dt + e^{-\int_0^t r(u)du}B_{r(t)}^{'}(r(t),t,T)\sigma\sqrt{r(t)}dW_t
\end{aligned}$$

We also find that the discount price $e^{-\int_0^t r(u)du}B(t,T)$ is a martingale since according to the tower property:

$$\begin{aligned}
E[e^{-\int_0^t r(u)du}B(t,T) \quad | \quad \mathcal{F}_s] &= E[e^{-\int_0^t r(u)du}E[e^{-\int_t^T r(u)du} \mid \mathcal{F}_t] \mid \mathcal{F}_s] \\
&= E[e^{-\int_0^s r(u)du}e^{-\int_s^t r(u)du}E[e^{-\int_t^T r(u)du} \mid \mathcal{F}_t] \mid \mathcal{F}_s] \\
&= e^{-\int_0^s r(u)du}E[e^{-\int_s^t r(u)du}E[e^{-\int_t^T r(u)du} \mid \mathcal{F}_t] \mid \mathcal{F}_s] \\
&= e^{-\int_0^s r(u)du}E[E[e^{-\int_s^T r(u)du} \mid \mathcal{F}_t] \mid \mathcal{F}_s] \\
&= e^{-\int_0^s r(u)du}E[e^{-\int_s^T r(u)du} \mid \mathcal{F}_s] \\
&= e^{-\int_0^s r(u)du}B(s,T)
\end{aligned}$$

where 0<s<t. By the martingale property, we can easily deduce that

$$-r(t)B(r(t),t,T)+(B_t^{'}(r(t),t,T)+(a-br(t))B_{r(t)}^{'}(r(t),t,T)+\frac{1}{2}B_{r(t)r(t)}^{''}(r(t),t,T)\sigma^2 r(t)=0$$

and the terminal condition is $B(r(t),T,T)=1$.In the Hull and White model case, we have the solution that $B(r(t),t,T) = e^{-rC(t,T)-A(t,T)}$.Thus $B_t^{'} = (-rC_t^{'}-A_t^{'})B$, $B_{r(t)}^{'}=-CB, B_{r(t)r(t)}^{''}=C^2B$.Then we have

$$rB(-1-C_t^{'}-bC+\frac{1}{2}\sigma^2C^2)-B(A_t^{'}+aC)=0$$

To solve $-1-C_t^{'}-bC+\frac{1}{2}\sigma^2C^2=0$,we get

$$C(t,T)=\frac{\sinh(\gamma(T-t))}{\gamma\cosh(\gamma(T-t))+\frac{1}{2}b\sinh(\gamma(T-t))}$$

$$A(t,T)=-\frac{2a}{\sigma^2}\log[\frac{\gamma e^{\frac{b(T-t)}{2}}}{\gamma\cosh(\gamma(T-t))+\frac{1}{2}b\sinh(\gamma(T-t))}]$$

where $\gamma=\frac{1}{2}\sqrt{b^2+2\sigma^2}, \sinh u=\frac{e^u-e^{-u}}{2}, \cosh u=\frac{e^u+e^{-u}}{2}$.

3 Firm Value Models

3.1 Introduction

So far we have introduced several types of credit derivatives and the term structural models. Now we will focus on the modeling method which can be used on the pricing of credit derivatives.

To be able to price credit derivatives we have to study the default risk of the underlying asset. The aim of the firm value models and intensity-based models are to model the default risk. Compared with the intensity model, the firm value model use a more fundamental approach to value the defaultable debt and also to provide a connection between debt of the firm and the values of equity.

Firm's value models assume a fundamental process V, denoting the total value of the assets of the firm that has issued the bonds. V is described as a stochastic process, influenced by the prices of all securities issued by the firm. A very important point of this type of model is that all claims on the firm's value are modelled as derivative securities with the firm's value as underlying. Black and Scholes (1973) and Merton (1974) were the first people modeling credit risk with what we know today as a firm's value model. Modeling credit risk means modeling default probability. The market value of the firm is simply the sum of the market value of the firm's debt and the value of its equity. If both these quantities were readily observable, calculating default probabilities would be trivial. While equity values are readily available, reliable data on the market value of debt is generally unavailable.

The value of a firm is an economic measure reflecting the market value of a whole business. The firm value can be measured by the price at which the total of the firm's liabilities can be bought or sold. It is a sum of claims of all claimants: creditors (secured and unsecured) and equity holders (preferred and common).

The market value of the asset V_t = market value of equity + market value of debt = S_t+ B_t , where S_t is the stock price and B_t=$B(t,T)$ is the bond price with maturity T and principle D.

We assume that the firm issues a single class of debt, a zero-coupon bond, with a face value D payable at T. Default may happen only at date T. If default happens, creditors take over the firm without incurring any distress costs and realize an amount V_T. Otherwise, they receive D. That is when $V_T > D$,the stock price $S_T = V_T - D$,and the bond price $B(T,T) = D$;When $V_T \leq D$,$S_T = 0$, $B(T,T) = V_T$.Therefore we can get the stock price and the bond price at maturity T are:

$$S_T = max(V_T - D, 0)$$

$$B(T,T) = min(D, V_T) = D - \max(D - V_T, 0)$$

We can also get:

$$S_t = E_{Q,t}[e^{-r(T-t)} S_T | \mathcal{F}(t)]$$

$$B(t,T) = E_{Q,t}[e^{-r(T-t)} B(T,T) | \mathcal{F}(t)]$$

where r is the risk free return, Q is risk neutral measure and $E_{Q,t}$ is conditional expectation under information until time t and measure Q.

3.2 Gaussian Firm Value Model

The Merton model generates the probability of default for each firm at any given point in time. To calculate the probability, the model subtracts the face value of the firm's existing debt from an estimate of the future market value of the firm and then divides this difference by an estimate of the volatility of the firm (scaled to reflect the horizon of the forecast). The resulting score, which is referred to as the distance to default, is then substituted into a cumulative density function to calculate the probability that the value of the firm will be less than the face value of debt at the forecasting horizon.

3.2.1 Merton Firm Value Process

The merton model makes an important assumption, which is the total value of a firm follows the geometric Brownian motion:

$$dV_t = u_v V_t dt + \sigma_v V_t dW_t$$

or

$$V_t = V_0 \exp(u_v t - \frac{\sigma_v^2}{2} t + \sigma_v W_t)$$

where the variable σ_v is the volatility of firm value, the variable u_v is the expected rate of return, and W_t is a Wiener process under market measure P. Under the risk neutral measure Q, we have

$$dV_t = rV_t dt + \sigma_v V_t d\widetilde{W_t}$$

or

$$V_t = V_0 \exp(rt - \frac{\sigma_v^2}{2}t + \sigma_v \widetilde{W_t})$$

where $\widetilde{W_t}$ is a Wiener process under the risk neutral measure Q.

The basic feature of Merton's model is that the firm's value drifts upwards over time and so its leverage falls. The second critical assumption of the Merton model is that the firm has issued just one discount bond maturing at time T. The company defaults if the value of its assets is less than the promised debt repayment at time T. The equity of the company is a European call option on the assets of the company with maturity T and a strike price equal to the face value of the debt. The model can be used to estimate the risk-neutral probability that the company will default as well as the credit spread on the debt. As inputs, Merton's model requires the current value of the company's assets, the volatility of the company's assets, the outstanding debt, and the debt maturity. In order to make the model analytically tractable, one has to estimate the current value and volatility of the company's assets from the market value of the company's equity and the equity's instantaneous volatility. A debt maturity date is chosen and debt payments are mapped into a single payment on the debt maturity date in some way. The rest of the implicit assumptions are the absence of transaction costs, bankruptcy costs, taxes or problems with indivisibility of assets, continuous time trading; unrestricted borrowing and lending at a constant interest rate r, no restrictions on the short selling of the assets and few more.

3.2.2 Bond and Stock Prices

From now on in this model, the prices of both debt $B(t,T)$ and equity S_t are functions of the firm's value V and the time t. What Black and Scholes (1973) and Merton (1974) did was a breakthrough. They showed that both equity and debt of the firm can be seen as derivative securities on the value V of firm's assets. The representation of the payoff to creditors makes it clear that the creditors short a put option written on the assets of the borrowing firm with a strike price equal to D, the face value of debt.

In addition, once we recognize that the borrower (equity holders in Merton's model), (a) owns the firm, (b) borrowed the amount D at t=0 , and (c) owns a put option on the assets of the firm with a strike price equal to D, it is immediate, by a put-call parity relationship, that equity is a call option on the assets of the borrowing firm with a strike price equal to D, the face value of debt. Therefore we can express stock and bond price as follows respectively:

$$S_t = Call_{BS}(V_t, D, r, T-t, \sigma)$$

$$B(t,T) = D - Put_{BS}(V_t, D, r, T-t, \sigma)$$

Since $V_t = V_0 \exp(rt - \frac{\sigma_v^2}{2}t + \sigma_v \widetilde{W}_t)$,we may thus write

$$\begin{aligned} V_T &= V_t \exp\left\{\sigma(\widetilde{W_T} - \widetilde{W}_t) + (r - \frac{\sigma^2}{2})(T-t)\right\} \\ &= V_t \exp\left\{\sigma\sqrt{T-t}Y + (r - \frac{\sigma^2}{2})(T-t)\right\} \end{aligned}$$

where Y is the standard normal random variable

$$Y = -\frac{\widetilde{W_T} - \widetilde{W}_t}{\sqrt{T-t}} \text{\textasciitilde} N(0,1)$$

We see that $\exp\left\{\sigma\sqrt{T-t}Y + (r - \frac{\sigma^2}{2})(T-t)\right\}$ is independent of $\mathcal{F}$(t), therefore,

$$\begin{aligned} S_t &= E_{Q,t}[e^{-r(T-t)}(V_T - D)^+ | \mathcal{F}(t)] \\ &= E_{Q,t}[e^{-r(T-t)}(V_t \exp\left\{-\sigma\sqrt{T-t}Y + (r - \frac{\sigma^2}{2})(T-t)\right\} - D)^+] \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-r(T-t)}(V_t \exp\left\{-\sigma\sqrt{T-t}y + (r - \frac{\sigma^2}{2})(T-t)\right\} - D)^+ e^{-\frac{y^2}{2}} dy \end{aligned}$$

The integrand is positive only if $V_t \exp\left\{-\sigma\sqrt{T-t}y + (r - \frac{\sigma^2}{2})(T-t)\right\} - D)$ is positive, which is

$$y < d_1 = \frac{1}{\sigma\sqrt{T-t}}[\log\frac{V_t}{D} + (r - \frac{\sigma^2}{2})(T-t)]$$

Then we can get

$$
\begin{aligned}
S_t &= \int_{-\infty}^{d_1} \frac{1}{\sqrt{2\pi}} e^{-r(T-t)} (V_t \exp\left\{-\sigma\sqrt{T-t}y + (r - \frac{\sigma^2}{2})(T-t)\right\} - D) e^{-\frac{y^2}{2}} dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} V_t \exp\{-\frac{y^2}{2} - \sigma\sqrt{T-t}y - \frac{\sigma^2}{2}(T-t)\} dy - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} \exp\{-r(T-t) - \frac{y^2}{2}\} D dy \\
&= \frac{1}{\sqrt{2\pi}} V_t \int_{-\infty}^{d_1} \exp\{-\frac{1}{2}(y + \sigma\sqrt{T-t})^2\} dy - e^{-r(T-t)} D N(d_1) \\
&= \frac{1}{\sqrt{2\pi}} V_t \int_{-\infty}^{d_1+\sigma\sqrt{T-t}} e^{-\frac{1}{2}z^2} dz - e^{-r(T-t)} D N(d_1) \\
&= V_t N(d_2) - e^{-r(T-t)} D N(d_1)
\end{aligned}
$$

where

$$
d_2 = d_1 + \sigma\sqrt{T-t} = \frac{1}{\sigma\sqrt{T-t}}[\log\frac{V_t}{D} + (r + \frac{\sigma^2}{2})(T-t)]
$$

Using the same method, we can get the bond price which is

$$
B(t,T) = e^{-r(T-t)} D N(d_1) + V_t N(-d_2))
$$

3.2.3 Credit Spread and Probability of Default

Merton's model allows us to compute (under the risk-neutral probability measure), respectively, the credit spread and the probability of default as follows:

$$
\begin{aligned}
\text{Spread} \quad &: \quad s(t,T) = spotrate\ R(t,T) - r \\
&= \frac{\log D - \log B(t,T)}{T-t} - r \\
&= -r - \frac{1}{T-t}\log\frac{B(t,T)}{D} \\
&= -r - \frac{1}{T-t}\log(\frac{e^{-r(T-t)} D N(d_1) + V_t N(-d_2))}{D}) \\
&= -r - \frac{1}{T-t}\log(e^{-r(T-t)} N(d_1) + M_t N(-d_2)) \\
&= -\frac{1}{T-t}\log(N(d_1) + \frac{M_t}{e^{-r(T-t)}} N(-d_2))
\end{aligned}
$$

where $M_t = \frac{V_t}{D}$.

$$\begin{aligned}
\text{Probability of default} &= P(V_T < D) \\
&= P(V_t \exp\left\{\sigma\sqrt{T-t}Y + (r - \frac{\sigma^2}{2})(T-t)\right\} < D) \\
&= P(Y < \frac{-\log\frac{V_t}{D} - (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}) \\
&= N(-d_1)
\end{aligned}$$

3.2.4 Simulation Results

By applying Ito lemma, we get

$$\begin{aligned}
\sigma_S &= \frac{\partial S}{\partial V}\frac{V}{S}\sigma_V \\
&= N(d_2)\frac{V}{S}\sigma_V
\end{aligned}$$

$$S_t = V_t N(d_2) - e^{-r(T-t)}DN(d_1)$$

Since risk free rate r, principal D of the bond, stock price history S_t are observed in the market, and σ_S can be calculated empirically using historical data of S_t, we can calculate V_t and σ_V by solving the above two formulas numerically:

$$V_t = \frac{S_t + e^{-r(T-t)}DN(d_1)}{N(d_2)}$$

$$\sigma_V = \frac{S\sigma_S}{S_t + e^{-r(T-t)}DN(d_1)}$$

Then if we know the volatility of stock price σ_S, we can calculate the volatility of firm value σ_V.On the contrary, if we know σ_V,we can calculate σ_S.Consider a company at time $t = 0$, we assume the equity is $S_t = 3m$, then we can get

Table 14: Merton Model Prediction Results(%)

σ_S	10.00	15.00	20.00	25.00	30.00	40.00	50.00	60.00	70.00	80.00
σ_V	2.33	3.49	4.65	5.81	6.98	9.32	11.76	14.39	17.32	20.64
spread	0.00	0.00	0.00	0.00	0.00	0.01	0.06	0.22	0.57	1.22
default Prob	0.00	0.00	0.00	0.00	0.01	0.26	1.43	3.96	7.84	12.84

Table 14 is the simulation results using the merton firm value model. We set the volatility of stock price σ_S as 10%, 15%, 20%, 25%, 30%, 40%, 50%, 60%, 70%, 80%, and get the corresponding volatility of firm value, the spread and the default probability respectively. According to the results, we can see that when σ_S is 30%, the spread is 0 and the default probability is only 0.01%, and when σ_S is as high as 80%, the spread is 1.22 and the default probability is only 12.84%, which is nonsense and impossible in the real financial market. The reason that leads to these results is that the merton model fails to capture the heavy tail, and it can't predict the default probability of a firm correctly.

3.3 Classical Tempered Stable Firm Value Model

In the Merton's firm value model, we assume that the firm value process V_t follows a geometric Brownian motion. However, the empirical evidence has shown that the Merton's Gaussian firm value model is not accurate. Firstly, in Merton's Gaussian firm value model, the leverage (M=V_t/D) is fixed, then high stock volatility means high default probability in the model which is practically not always true. Secondly, skewness is not considered in the firm value model. Considering skewness, two firms with same leverage and same volatility may have different default probabilities. Thirdly, the volatility of the firm value σ_v obtained in the Merton's model is different from the implied volatility. Finally, Driving process for the firm value is Brownian Motion. Hence, it can't capture the fat tail property and big jump of firm value process.

Because of the drawbacks of the Merton's model, the candidates for non-normal distribution that proposed for modeling extreme events are referred to as stable distributions. Stable distributions has desired property to accommodate heavy tail and capture skewness in financial series. It also has stability property that the sum of two independent stable random variables follows, up to some correction of scale and location, the same stable distribution. Despite the empirical evidence rejecting the normal distribution and in support of the stable distribution, there have been several barriers to the application of stable distribution models.

The major problem is that the variance of stable non-normal distributions is infinity. The second problem is that without a general expression for stable probability densities, one can't directly implement estimation methodologies for fitting these densities. The third problem is that the empirical tail distribution for asset returns are thinner than the stable distribution. To overcome the drawbacks , the tails of a stable random variable can be appropriately tempered or truncated in order to obtain a proper distribution. One alternative is classical tempered stable distribution (CTS).

In this chapter, we would use the classical tempered stable distribution as the driving process of firm value, and estimate the parameters of the CTS model. We then could calculate the default probability under CTS firm value model, and compare it with the default probability under Merton's firm value model.

3.3.1 Classical Tempered Stable Distribution

In financial modeling, the infinite variance characteristic of a stable distribution makes its application impossible because the infinite variance of the return lead to an infinite price for derivative instruments, clearly contradicting reality and intuition. For this reason, Menn and Rachev (2009) suggest the use of appropriately truncated stable distributions and tempered stable distributions. These techniques make the tails of the alpha stable distribution behave like fat tails but are light tails in the mathematical sense. Consequently, all moments of arbitrary order exist and are finite.

In this chapter, we will mainly discuss the CTS distribution. Let $\alpha \in (0,1) \cup (1,2), C, \lambda_+, \lambda_- > 0$,and $m \in \mathbb{R}.X$ is said to follow the CTS distribution if the characteristic function of X is given by

$$\begin{aligned} \phi_X(u) &= \phi_{CTS}(u; \alpha, C, \lambda_+, \lambda_-, m) \\ &= \exp(ium - iuC\Gamma(1-\alpha)(\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1}) + \\ &\quad C\Gamma(-\alpha)((\lambda_+ - iu)^\alpha - \lambda_+^\alpha + (\lambda_- + iu)^\alpha - \lambda_-^\alpha)) \end{aligned}$$

we denote $X \backsim CTS(\alpha, C, \lambda_+, \lambda_-, m)$.

The cumulants of X are obtained by

$$c_1(X) = m$$

$$c_n(X) = C\Gamma(n-\alpha)(\lambda_+^{\alpha-n} + (-1)^n \lambda_-^{\alpha-n}), \text{ for } n = 2, 3...$$

The role of the parameters is as follows:

(1) m determines the location of the distribution.

(2) C is the scale parameter.

(3) λ_+ and λ_- control the rate of decay on the positive and negative tails, respectively. If $\lambda_+ > \lambda_-$, the distribution is skewed to the left, if $\lambda_+ < \lambda_-$, then the distribution is skewed to the right, and if $\lambda_+ = \lambda_-$, then it is symmetric.

(4) The parameters λ_+, λ_-, and α are related to tail weights.

(5) If α approaches to 0, the CTS distribution converges to the variance gamma distribution.

3.3.2 Classical Tempered Stable Firm Value Model

Let $(X_t)_{t\geq 0}$ be the classical tempered stable process with parameters $(\alpha, C, \lambda_+, \lambda_-, m)$, where $\alpha \in (0,1) \cup (1,2), \lambda_+ > 1$. Under the risk neutral measure $\mathbb{Q}$ and $E_{\mathbb{Q}}[\exp(X_t)] = \exp(rt)$ with the risk free rate of return $r < \lambda_+$.Suppose that the firm value process $V = (V_t)_{t\geq 0}$ is given by the exponential Lėvy process model as follows:

$$V_t = V_0 \exp(X_t)$$

As we have discussed before, the stock price at time t is given by

$$\begin{aligned} S_t &= E_{Q,t}[e^{-r(T-t)}(V_T - D)^+ | \mathcal{F}(t)] \\ &= e^{-r(T-t)} E_{Q,t}[(V_t \exp(X_T - X_t) - D)^+ | \mathcal{F}(t)] \end{aligned}$$

Let $f(x)$ be the p.d.f of the random variable $X_T - X_t$.By the stationary property of the tempered stable process $X, X_T - X_t$ has the same distribution as X_{T-t}.The p.d.f. $f(x)$ is obtained by the complex inversion formula as follows:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(u+i\rho)y} \phi_{X_{T-t}}(u + i\rho) du$$

Hence we can get

$$\begin{aligned}
& E_{Q,t}[(V_t \exp((X_T - X_t) - D)^+ | \mathcal{F}(t)] \\
= & \int_{\log \frac{D}{V_t}}^{\infty} (\exp(x + \log V_t) - D) f(x) dx \\
= & \frac{1}{2\pi} \int_{\log \frac{D}{V_t}}^{\infty} (\exp(x + \log V_t) - D) \int_{-\infty}^{\infty} e^{-i(u+i\rho)x} \phi_{X_{T-t}}(u + i\rho) du dx
\end{aligned}$$

Let $z \in \mathbb{C}$.If $\mathrm{Im}(z) < -1$, then we have

$$\int_a^{\infty} e^{(1-iz)x} dx = -\frac{e^{a(1-iz)}}{1 - iz}$$

and if $\mathrm{Im}(z) < 0$

$$\int_a^{\infty} e^{izx} dx = \frac{e^{-iaz}}{iz}$$

Hence, we can get

$$\begin{aligned}
& \int_{\log \frac{D}{V_t}}^{\infty} (\exp(x + \log V_t) - D) e^{-i(u+i\rho)x} dx \\
= & V_t \int_{\log \frac{D}{V_t}}^{\infty} e^{(1-i(u+i\rho))x} dx - D \int_{\log \frac{D}{V_t}}^{\infty} e^{-i(u+i\rho)x} dx \\
= & -\frac{V_t \exp((1 - i(u + i\rho)) \log \frac{D}{V_t})}{1 - i(u + i\rho)} - \frac{D \exp(-i(u + i\rho) \log \frac{D}{V_t})}{i(u + i\rho)} \\
= & \frac{D \exp(-i(u + i\rho) \log \frac{D}{V_t})}{(iu - i - \rho)(ui - \rho)} \\
= & \frac{D^{1+\rho} \exp(-iu \log \frac{D}{V_t})}{V_t^{\rho} (iu - 1 - \rho)(ui - \rho)}
\end{aligned}$$

Therefore, by Fubini's theorem, for $\rho < -1$ we obtain the stock price formula as follows:

$$\begin{aligned}
S_t = & e^{-r(T-t)} E_{Q,t}[(V_t \exp(X_T - X_t) - D)^+ | \mathcal{F}(t)] \\
= & \frac{e^{-r(T-t)} D^{1+\rho}}{2\pi V_t^{\rho}} \int_{-\infty}^{\infty} \frac{\exp(-iu \log \frac{K}{S_t})}{(iu - 1 - \rho)(ui - \rho)} \phi_{X_{T-t}}(u + i\rho) du \\
= & \frac{e^{-r(T-t)} D^{1+\rho}}{2\pi V_t^{\rho}} \int_{-\infty}^{\infty} \frac{V_t^{iu} \phi_{X_{T-t}}(u + i\rho)}{D^{iu} (iu - 1 - \rho)(ui - \rho)} du \\
= & D e^{-r(T-t)} \frac{(D/V_t)^{\rho}}{\pi} \mathrm{Re} \int_0^{\infty} (D/V_t)^{-iu} \frac{\phi_{X_{T-t}}(u + i\rho)}{(iu - 1 - \rho)(ui - \rho)} du
\end{aligned}$$

where

$$\begin{aligned}
\phi_{X_{T-t}}(z) = & \exp(im(T - t)z + (T - t)C\Gamma(-\alpha)((\lambda_+ - iz)^{\alpha} - \lambda_+^{\alpha} + (\lambda_- + \sigma_V zi)^{\alpha} - \lambda_-^{\alpha}) \\
& -iz(T - t)C\Gamma(1 - \alpha)(\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1})
\end{aligned}$$

and

$$-\lambda_+ < \rho < -1$$

With the same method, we can get the bond price

$$\begin{aligned} B(t,T) &= D - E_{Q,t}[e^{-r(T-t)}(D-V_T)^+|\mathcal{F}(t)] \\ &= De^{-r(T-t)}(1-\frac{(D/V_t)^\rho}{\pi}\,\mathrm{Re}\int_0^\infty (D/V_t)^{-iu}\frac{\phi_{X_{T-t}}(u+i\eta)}{(iu-1-\rho)(ui-\rho)}du) \end{aligned}$$

where

$$0 < \eta < \lambda_-$$

Hence, the spread formula is

$$\begin{aligned} s(t,T) &= -r - \frac{1}{T-t}\log\frac{B(t,T)}{D} \\ &= -r - \frac{\log(e^{-r(T-t)}(1-\frac{(D/V_t)^\rho}{\pi}\,\mathrm{Re}\int_0^\infty (D/V_t)^{-iu}\frac{\phi_{Y_{T-t}}(u+i\eta)}{(iu-1-\rho)(ui-\rho)}du))}{T-t} \\ &= -\frac{\log(1-\frac{(D/V_t)^\rho}{\pi}\,\mathrm{Re}\int_0^\infty (D/V_t)^{-iu}\frac{\phi_{Y_{T-t}}(u+i\eta)}{(iu-1-\rho)(ui-\rho)}du)}{T-t} \end{aligned}$$

4. Empirical Test

4.1 Parameters Estimation

In this chapter, we will use corporate bond spread data to estimate the parameters of the CTS model and the Merton model. Table 15 is the Reuters corporate bond spread for different credit rating industrials. The evaluators obtain the spreads from brokers and traders at various firms. A generic spread for each sector is created using input from street contacts and the evaluator's expertise. A matrix is then developed based on sector, rating, and maturity.

We fit the spread data to the CTS distribution and get the parameters. Table 16 shows the estimated CTS model parameters and the firm values at time t=0 for different credit rating industrials. Table 17 shows the estimated Merton model parameters for different credit rating industrials.

Table 15: Reuters Corporate Bond Spread (%)

	Aaa/AAA	Aa2/AA	A2/A	Baa2/BBB	Ba2/BB
1 yr	0.21	0.30	0.43	0.96	1.72
2 yr	0.26	0.32	0.58	1.11	2.72
3 yr	0.38	0.42	0.71	1.32	3.31
4 yr	0.45	0.53	0.79	1.44	3.46
5 yr	0.53	0.65	0.88	1.53	3.43
6 yr	0.55	0.72	0.92	1.60	3.35
7 yr	0.61	0.83	1.02	1.78	3.36
8 yr	0.65	0.95	1.15	2.04	3.41
9 yr	0.70	1.08	1.31	2.30	3.49
10 yr	0.76	1.21	1.47	2.53	3.61
12 yr	0.84	1.37	1.65	2.77	-
15 yr	0.99	1.49	1.75	2.80	3.86
20 yr	1.22	1.53	1.75	2.69	3.64
25 yr	1.27	1.46	1.62	2.40	3.04
30 yr	1.21	1.14	1.41	2.00	-

Table 16: Estimated CTS Parameters

	α	C	λ_+	λ_-	m	V_0
Aaa/AAA	0.8049	0.5569	59.6313	3.2948	0.0153	4.0157
Aa2/AA	0.8725	0.6082	48.0487	3.9470	0.0153	3.3357
A2/A	0.8963	0.6209	52.6168	4.2247	-0.0439	2.8342
Baa2/BBB	0.7461	0.5356	54.3634	1.6673	-0.0899	4.1039
Ba2/BB	0.9614	1.2377	53.6000	6.1976	-0.0809	2.0631

Table 17: Estimated Merton Parameters

	Aaa/AAA	Aa2/AA	A2/A	Baa2/BBB	Ba2/BB
σ_V	0.3093	0.3302	0.3365	0.3997	0.4220
V_0	2.9293	2.8262	2.5692	2.5697	1.8996

By these parameters we can get the simulated spread data. Fit the spread data in figures as follows:

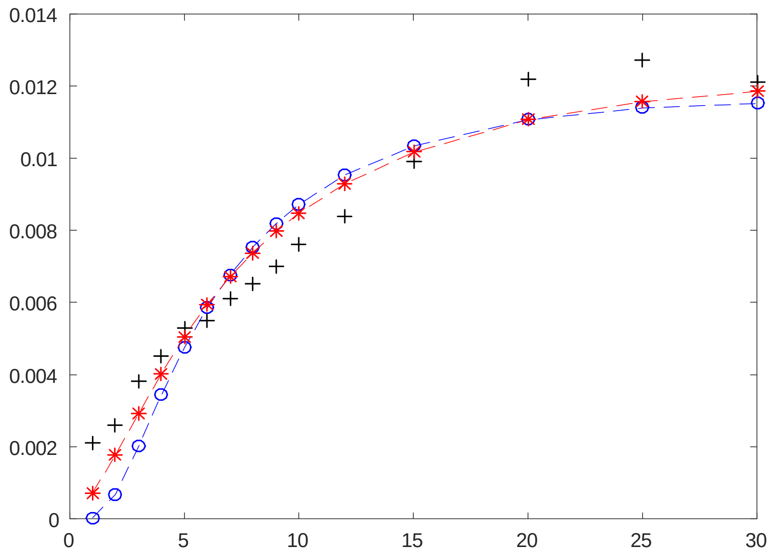


Figure 11: Spread Data of AAA Credit Rating Industrials

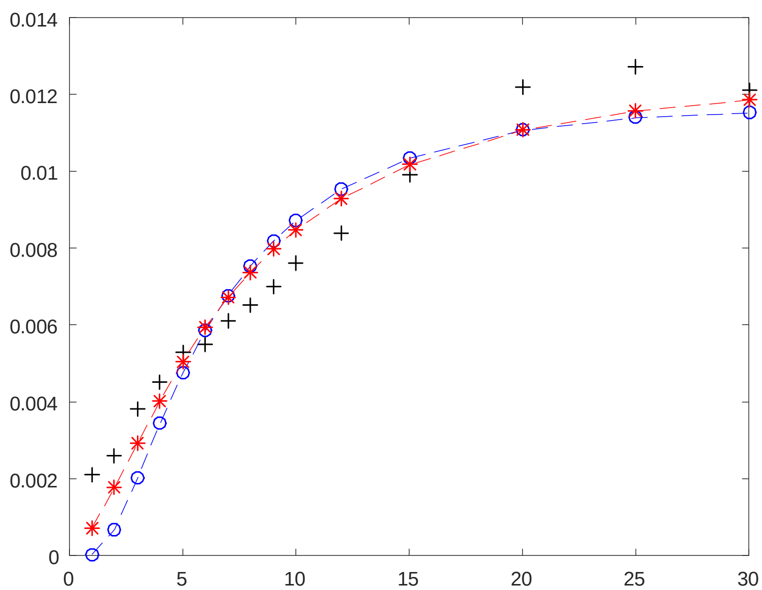


Figure 12: Spread Data of AA Credit Rating Industrials

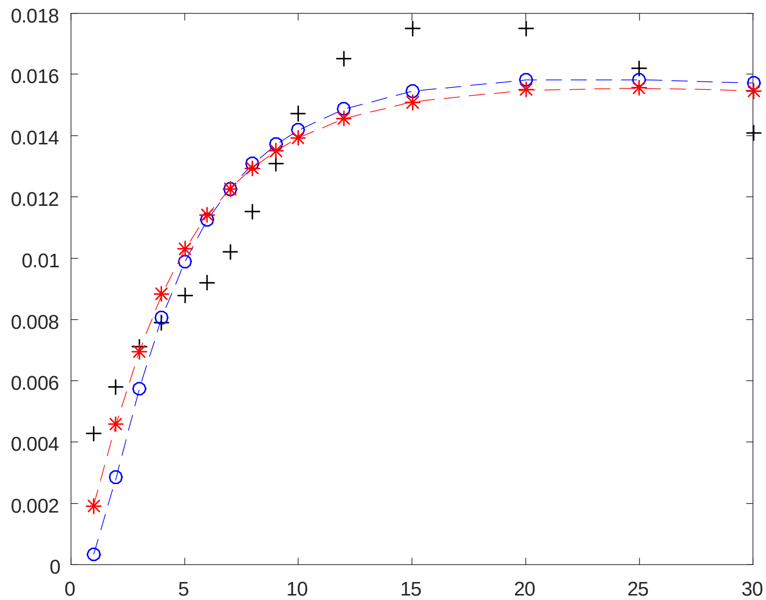


Figure 13: Spread Data of A Credit Rating Industrials

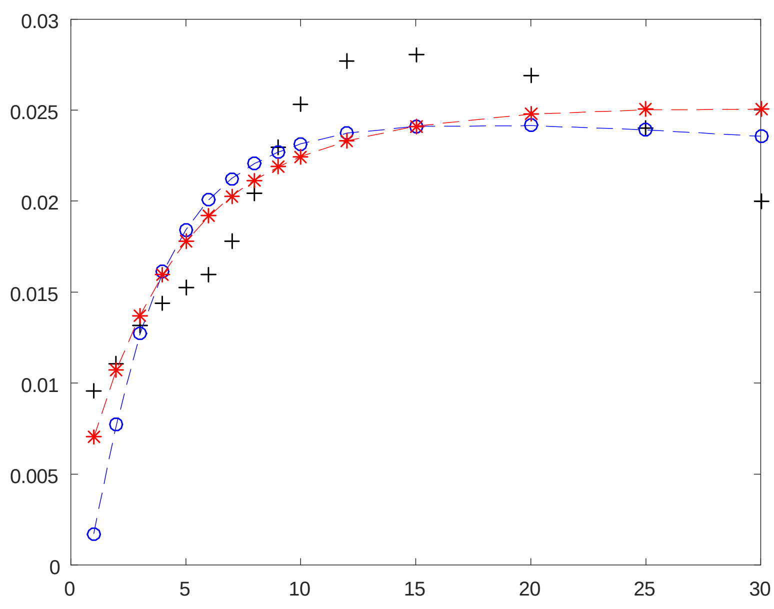


Figure 14: Spread Data of BBB Credit Rating Industrials

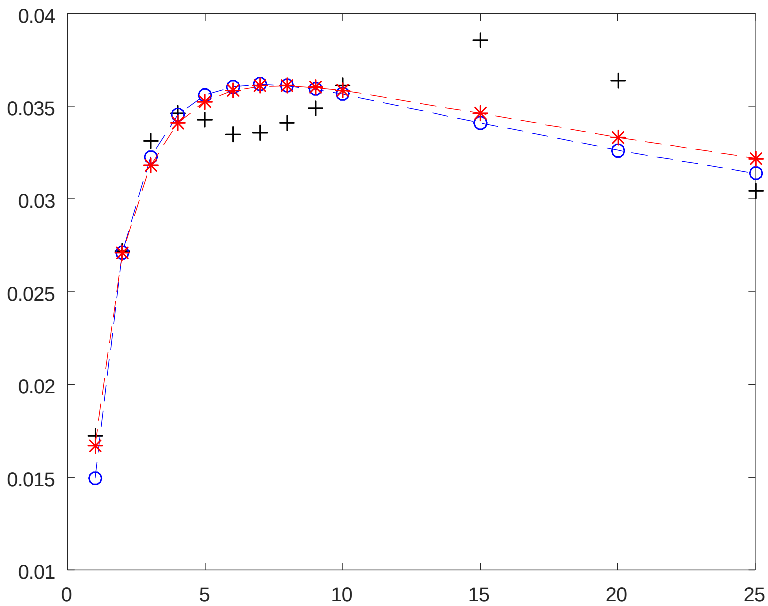


Figure 15: Spread Data of BB Credit Rating Industrials

Figure 11-Figure 15 dedicate the empirical spread data and the spread data simulated by Merton model and CTS model for different credit rating industrials. the black plus signs "+" give the empirical data, the blue lines give the Merton model simulate data, and the red lines give the CTS model simulate data. It is obvious that in each figure, the red line is closer to the black signs compared to the blue line.

To compare the goodness of fit, we use four discrepancy measures: average absolute error (AAE), average prediction error (APE), average relative prediction error (ARPE) and root mean square error (RMSE) test, the measures are as follows:

Table 18: goodness of fit - CTS model

	AAA	AA	A	BBB	BB
AAE	0	0.0011	0.0014	0.0023	0.0016
APE	0	0.1179	0.1228	0.1174	0.0481
ARPE	0.1508	0.1527	0.1484	0.1214	0.0463
RMSE	0	0.0013	0.0016	0.0027	0.0019

Table 19: goodness of fit - Merton model

	AAA	AA	A	BBB	BB
AAE	0.0011	0.0012	0.0016	0.0028	0.0017
APE	0.1533	0.1280	0.1359	0.1449	0.0527
ARPE	0.2399	0.1902	0.1908	0.1782	0.0542
RMSE	0.0012	0.0014	0.0019	0.0034	0.0022

Compare table 18 and table 19, it is easy to find that under the CTS model, all the four test results are smaller than the results under the Merton model. That means the CTS model works better than the Merton model.

4.2 Default Probabilities

The default probability under the CTS firm value model is calculated as:

$$\begin{aligned}\text{Probability of default} &= P(V_T < D)\\ &= P(V_t \exp(X_T) < D)\\ &= cdf\_CTS(\log \frac{D}{V_t})\end{aligned}$$

The default probability under the Merton firm value model is calculated as:

$$\begin{aligned}\text{Probability of default} &= P(V_T < D)\\ &= P(V_t \exp\left\{\sigma\sqrt{T-t}Y + (r - \frac{\sigma^2}{2})(T-t)\right\} < D)\\ &= P(Y < \frac{-\log \frac{V_t}{D} - (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}})\\ &= N(-d_1)\end{aligned}$$

Since we have estimated the parameters of the CTS model and the Merton model, we can then calculate the default probabilities according to the above formulas. Table 20 is the default probabilities of CTS firm Value model for different credit rating industrials at different maturity. Table 21 is the default probabilities of the Merton firm value model for different credit rating industrials at different maturity. Table 22 - Table 26 give the default probabilities and the volatility of stock price for different credit rating industrials at different maturity.

Table 20: Default Probabilities(%) of CTS Model

	Aaa/AAA	Aa2/AA	A2/A	Baa2/BBB	Ba2/BB
1 yr	0.29	0.45	1.08	2.23	8.19
2 yr	1.04	1.67	4.08	6.02	18.45
3 yr	2.07	3.23	7.80	10.30	26.09
4 yr	3.21	4.81	11.54	14.55	31.94
5 yr	4.34	6.30	15.06	18.58	36.65
6 yr	5.43	7.65	18.29	22.35	40.56
7 yr	6.45	8.86	21.25	25.84	43.89
8 yr	7.40	9.95	23.96	29.07	46.80
9 yr	8.27	10.93	26.45	32.07	49.38
10 yr	9.07	11.81	28.75	34.85	51.69
12 yr	10.49	13.31	32.84	39.85	-
15 yr	12.23	15.10	38.03	46.23	60.58
20 yr	14.39	17.21	44.91	54.69	66.83
25 yr	15.92	18.65	50.34	61.27	71.60
30 yr	17.04	19.67	54.79	66.56	-

Table 21: Default Probabilities(%) of Merton Model

	Aaa/AAA	Aa2/AA	A2/A	Baa2/BBB	Ba2/BB
1 yr	0.04	0.12	0.37	1.39	8.92
2 yr	1.05	1.99	3.52	7.48	20.39
3 yr	3.41	5.36	7.98	13.93	28.26
4 yr	6.34	9.09	12.38	19.56	34.07
5 yr	9.36	12.69	16.37	24.37	38.62
6 yr	12.29	16.03	19.94	28.50	42.37
7 yr	15.04	19.08	23.12	31.10	45.53
8 yr	17.60	21.88	25.98	35.28	48.28
9 yr	19.98	24.44	28.56	38.13	50.70
10 yr	22.18	26.79	30.92	40.69	52.87
12 yr	26.16	30.97	35.05	45.16	-
15 yr	31.22	36.24	40.20	50.65	61.20
20 yr	37.97	43.17	46.91	57.72	67.07
25 yr	43.30	48.60	52.13	63.15	71.56
30 yr	47.69	53.05	56.39	67.52	-

Table 22: Default Probabilities of Merton Model (AAA)

	$\sigma_S(\%)$	$Spread(\%)$	$Default\ probability(\%)$
1 yr	46.59	0.21	0.04
2 yr	46.07	0.26	1.05
3 yr	45.34	0.38	3.41
4 yr	44.55	0.45	6.34
5 yr	43.76	0.53	9.36
6 yr	43.01	0.55	12.29
7 yr	42.32	0.61	15.04
8 yr	41.69	0.65	17.60
9 yr	41.11	0.70	19.98
10 yr	40.57	0.76	22.18
12 yr	39.63	0.84	26.16
15 yr	38.45	0.99	31.22
20 yr	36.97	1.22	37.97
25 yr	35.88	1.27	43.30
30 yr	35.05	1.21	47.69

Table 23: Default Probabilities of Merton Model (AA)

	$\sigma_S(\%)$	$Spread(\%)$	$Default\ probability(\%)$
1 yr	50.66	0.0030	0.12
2 yr	49.90	0.0032	1.99
3 yr	48.89	0.0042	5.36
4 yr	47.84	0.0053	9.09
5 yr	46.84	0.0065	12.69
6 yr	45.94	0.0072	16.03
7 yr	45.11	0.0083	19.08
8 yr	44.37	0.0095	21.88
9 yr	43.70	0.0108	24.44
10 yr	43.08	0.0121	26.79
12 yr	42.02	0.0137	30.97
15 yr	40.72	0.0149	36.24
20 yr	39.12	0.0153	43.17
25 yr	37.96	0.0146	48.60
30 yr	37.09	0.0114	53.05

Table 24: Default Probabilities of Merton Model (A)

	$\sigma_S(\%)$	$Spread(\%)$	$Default\ probability(\%)$
1 yr	54.49	0.43	0.37
2 yr	53.27	0.58	3.52
3 yr	51.79	0.71	7.98
4 yr	50.38	0.79	12.38
5 yr	49.10	0.88	16.37
6 yr	47.97	0.92	19.94
7 yr	46.97	1.02	23.12
8 yr	46.09	1.15	25.98
9 yr	45.30	1.31	28.56
10 yr	44.59	1.47	30.92
12 yr	43.37	1.65	35.05
15 yr	41.91	1.75	40.20
20 yr	40.13	1.75	46.91
25 yr	38.86	1.62	52.13
30 yr	37.92	1.41	56.39

Table 25: Default Probabilities of Merton Model (BBB)

	$\sigma_S(\%)$	$Spread(\%)$	$Default\ probability(\%)$
1 yr	64.44	0.96	1.39
2 yr	62.18	1.11	7.48
3 yr	59.88	1.32	13.93
4 yr	57.89	1.44	19.56
5 yr	56.21	1.53	24.37
6 yr	54.78	1.60	28.50
7 yr	53.55	1.78	31.10
8 yr	52.49	2.04	35.28
9 yr	51.56	2.30	38.13
10 yr	50.74	2.53	40.69
12 yr	49.36	2.77	45.16
15 yr	47.75	2.80	50.65
20 yr	45.86	2.69	57.72
25 yr	44.55	2.40	63.15
30 yr	43.61	2.00	67.52

Table 26: Default Probabilities of Merton Model (BB)

	$\sigma_S(\%)$	$Spread(\%)$	$Default\ probability(\%)$
1 yr	82.93	1.72	8.92
2 yr	75.42	2.72	20.39
3 yr	70.21	3.31	28.26
4 yr	66.43	3.46	34.07
5 yr	63.54	3.43	38.62
6 yr	61.25	3.35	42.37
7 yr	59.38	3.36	45.53
8 yr	57.82	3.41	48.28
9 yr	56.49	3.49	50.70
10 yr	55.35	3.61	52.87
12 yr	-	-	-
15 yr	51.35	3.86	61.20
20 yr	48.95	3.64	67.07
25 yr	47.36	3.04	71.56
30 yr	-	-	-

4.3 Conclusion

According to Table 22 - Table 26, we find that when the maturity is one year after, with the Merton firm value model, the volatility of stock price σ_S is 46.59%, the corresponding default probability is 0.04% for AAA credit rating industrials. The volatility of stock price σ_S is 50.66%, the corresponding default probability is 0.12% for AA credit rating industrials. The volatility of stock price σ_S is 54.49%, the corresponding default probability is 0.37% for A credit rating industrials. The volatility of stock price σ_S is 64.44%, the corresponding default probability is 1.39% for BBB credit rating industrials. The volatility of stock price σ_S is 82.93%, the corresponding default probability is 8.92% for A credit rating industrials. It's very obvious that for the AAA rating industrials, the volatility of stock price won't be as high as around 50%. And when the volatility of stock price is as high as around 50%. the default probability can not be less than 0.5%. As we have discussed before, the merton model can't capture the heavy tails, and these results prove it again. Compared with the Merton model, the property of capturing fat tails makes the temper stable distribution a better model to predict the default probability.

5 Difficulties and Future Work

Since there is not enough data, we are not be able to do the backtest right now. We only know that the CTS firm value model is better than the Merton firm value model, but we don't know how good it is and how correct it is to predict the default probability of a firm.

The next step to do is to collect the real default probability data and do the backtest. We will also apply the CTS model on the credit derivatives such as CDS.

6 Reference

[1] Aijun Zhang (2009). "Statistical Methods in Credit Risk Modeling". Ph.D dissertation in The University of Michigan.

[2] Amit Kulkarni, Alok Kumar Mishra, Jigisha Thakker (2008). "How Good is Merton Model at Assessing Credit Risk? Evidence From India". Second Singapore International Conference on Finance 2008.

[3] A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management : Concepts, Techniques, and Tools*. Princeton University Press, 2005.

[4] Andrew Lesniewski. Short rate models. Lecture Note (2008).

[5] Basle. "Credit Risk Modelling: Current Practices and Applications". Basel Committee on Banking Supervision.

[6] Benninga, S. (1998), "Financial Modelling", Cambridge, MIT Press.

[7] Berlin, M. and Mester, L. J. (2004). Special issue of "Retail credit risk management and measurement". Journal of Banking and Finance, 28, 721-899.

[8] Brigo, D., and Mercurio, F.. *Interest Rate Models−Theory and Practice*, Springer Verlag (2006).

[9] Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985). "A theory of the term structureof interest rates". Econometrika, 53, 385-407.

[10] Collin-Dufresne, Pierre and Robert Goldstein. "Do Credit Spreads Reflect Stationary Leverage Ratios ?," Journal of Finance, 2001, v56, 1928-1957.

[11] D. C. Heath, R. A. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates. Econometrica, 60:77–105, 1992.

[12] Eom, Y.H., J. Helwege, and J.-Z. Huang (2000), "Structural Models of Corporate Bond Pricing: An Empirical Analysis", Working Paper no. 2000-16, Pennsylvania State University.

[13] Grüne. L. and W. Semmler (2005), Default Risk, Asset Pricing and Credit Control, Journal of Financial Econometrics, vol. 1: 1-28.

[14] J. James and N. Webber. *Interest Rate Modelling*. J. Wiley, Chichester, 2000.

[15] Leland, Hayne E. and Klaus Bjerre Toft. “Optimal Capital Structure, Endogenous Bankruptcy, And The Term Structure Of Credit Spreads,” Journal of Finance, 1996, v51, 987-1019.

[16] Merton, Robert C. “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates,” Journal of Finance, 1972, v29, 449-470.

[17] Merxe Tudela and Garry Young (2003). "A Merton-model approach to assessing the default risk of UK public companies". Working Paper.

[18] Nicolas Privault (2012). "*An Elementary Introduction to Stochastic Interest Rate Modeling*".

[19] R.Gibson,F.-S. Lhabitant, D. Talay. "Modeling the term structure of interest rate: A review of the literature". Lecture Note (2001).

[20] Stefan Trueck, Svetlozar T. Rachev (2009). "*Rating Based Modeling of Credit Risk*".

[21] Stockholm, Sweden (2008). "Quantitative Analysis of Credit Risk in Financial Market". Master of Science Thesis.

[22] Suresh Sundaresan. "A Review of Merton’s Model of the Firm’s Capital Structure with its Wide Applications". Annual Review of Financial Economics, 2013, Vol. 5: 21-41.

[23] Svetlozar Rachev. "Credit Risk : Intensity Based Model". Working Paper.

[24] Willi Semmler (2008), Lucas Bernard, Michael Robert. "Credit risk, credit derivatives and firm value based models". Investment Management and Financial Innovations. Vol. 5: 1-15.

[25] Wim Schoutens Jessica Cariboni. "*Levy Processes in Credit Risk* (*The Wiley Finance Series*)”.