# Multi-period Value-at-Risk Scaling Rules: Calculations and Approximations

**ZHOU, Pengpeng**

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Doctor of Philosophy
in
Statistics

The Chinese University of Hong Kong
August, 2011

UMI Number: 3500848

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent on the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI
Dissertation Publishing

UMI 3500848

Copyright 2012 by ProQuest LLC.

All rights reserved. This edition of the work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

# Thesis/Assessment Committee

Professor GU Ming Gao (Chair)
Professor CHAN Ngai Hang (Thesis Supervisor)
Professor WONG Hoi Ying (Committee Member)
Professor TIAN Maozai (External Examiner)

Abstract of thesis entitled:
Multi-period Value-at-Risk Scaling Rules: Calculations and Approximations
Submitted by ZHOU, Pengpeng
for the degree of Doctor of Philosophy
at The Chinese University of Hong Kong in August, 2011

The thesis firstly introduces a commonly used risk measure in the financial market - Value-at-Risk (VaR) and then the research about multi-period risk management is proposed. A general tool for multi-period Value-at-Risk (VaR), proposed by the Basel Committee on Banking Supervision (1996), is the *square-root-of-time rule* (SQRT rule), which is derived based on the Gaussian distributional assumption. Owing to the theoretical limitations of Gaussian and the lessons from the financial catastrophe, this thesis develops new scaling rules based on the distributions that belong to the so-called convolution equivalent class and the semi-heavy tailed distribution class in which the tails of distribution seem adequate and comply with the empirical tail property of real financial data. In this thesis, under some regularity conditions, a result about multi-period VaR scaling approach based on convolution equivalence assumption (CE rule) is derived and proved, which may provide a conservative risk value to regulators. Furthermore, this thesis provides a precise numerical multi-period VaR scaling approach based on the semi-heavy tail assumption (SH rule), which is a numerical method that can be considered as an alternative to the SQRT rule and an internal scaling model for risk managers. Based on the assumption of a specific the semi-heavy form in the tail, we devise a semi-parametric estimation for single-period VaR calculation. The steps for using the two rules (Denoted by the SP-CE rule and the SP-SH rule) are summarized. For the whole parametric distributional assumption such as the example of the Normal Inverse Gaussian (NIG) distribution, we give specific scaling rules: the NIG-CE rule and the NIG-SH rule. The thesis also derives the saddlepoint approximation to the NIG model's multi-period VaR for internal risk management. Simulations and real data analysis evaluated and verified the feasibility of the CE rule , the SH rule and the approximated VaR method. It is found that, unlike the SQRT rule, the newly derived scaling rule has the advantage that captures the long horizon risk in a feasible way that can help regulators and risk managers. Specifically, the CE rule proposed under convolution equivalence assumption is highly recommended to regulators. About the internal supervision, the semiparametric estimation of the single-period VaR combined with the semi-heavy rule (denoted by SP-SH) would be the preferred choice. In the parametric modeling, the NIG fitting combined with the semi-heavy rule (denoted by NIG-SH) is reasonable. The saddlepoint approximation provides a fast and accurate VaR when the assumption is close to the true one.

**Keywords:** Value-at-risk (VaR), convolution equivalent, semi-heavy tail, saddlepoint approximation, backtesting.

# 摘要

本文首先回顧了在金融市場應用甚廣的風險測量工具-風險值。初步了解之後，有關長期風險管理的問題得以提出來研究。一般通用的有關長期風險的法則，是由巴塞爾委員會於1996年提出的時間平方根法則（稱為SQRT法則），它基於金融資產的收益服從正態分佈的假設，已經被金融監管機構的監管員和金融企業內部風險管理經理廣為接受。考慮到正態分佈的理論局限性，還有金融危機給予的教訓，本文試圖在符合尾部厚於正態以及輕於冪函數的分佈族-卷積等價分佈族和半厚尾分佈族中建立新型長期風險管理法則。在一定正則條件下，本文基於卷積等價的分佈假設推導和證明了一個遞推定理（稱為CE法則），並將之推薦給外部金融機構監管者來使用。更進一步，在半厚尾的分佈假設前提下，本文提供一個準確的可以通過簡單計算而得到長期風險值的計算法則（稱為SH法則）。這個計算法則可以為風險管理經理提供在平方根法則外的內部監管備選模型。這兩個法則（CE法則和SH法則）的條件以及其背景中卷積等價分佈族和半厚尾分佈族的關係在過程中會涉及和討論。假設資本收益分佈尾部服從半厚尾假設，我們發展出一個運用順序統計量來構造的短期風險值的半參數計算方法。本文總結了這個方法與前面兩種法則相結合來運用的步驟（分別表示為SP-CE規則和SP-SH規則）。當假設收益分佈服從一個全參數分佈結構時，以正態逆高斯（NIG）分佈為例，我們給出了其相應的長期風險遞推規則-分別為NIG-CE規則和NIG-SH規則。我們同時還推導了在這個分佈假設下長期風險值的鞍點逼近，作為一個內部風險管理模型來推出。模擬和實例檢測和驗證了CE法則、SH法則和鞍點逼近方法的可行性。研究發現，不同於SQRT法則，新的遞推法則以可行的方式更好地抓住了長期風險的特徵，能夠幫助金融機構監管者和風險管理經理更好地處理風險。特別地，在卷積等價假設下得到的CE法則可推薦給金融監管者。在企業內部風險管理方面，短期風險值的半參數計算方法結合SH法則（記作SP-SH規則）為表現最好的模型。全參數模型例如擬合NIG分佈結合SH法則（記作NIG-SH規則）在高分位表現上確也不錯。在分佈假設接近真實的時候，鞍點逼近提供了一個快速而準確的長期風險值近似方法。文章最後，總結之餘同時提出了進一步的討論。

**關鍵字：** 風險值（VaR），卷積等價，半厚尾，鞍點逼近，平方根法則，後驗檢驗。

# Acknowledgement

I would like to express my deepest gratitude to my supervisor, Professor CHAN Ngai Hang, for his generosity of supervision and cordial support during this research program. Throughout the whole study period, Professor CHAN helps me conquer a lot of difficulties from time to time. Without his guidance, this dissertation would not appear in this form.

I should also thank Professor GU Ming-gao , Professor WONG Hoi Ying for their kind assistances on an earlier version of this thesis, and thank Professor TIAN Mao-zai for his everlasting support and encouragement during my study period.

It is also a pleasure to give my thanks to all my friends and family members. They give me warm support when I am overwhelmed by pressures and difficulties.

This thesis is dedicated to the memory my grandmother Mrs. Tingyu J. Zhou (1921 - 2009) for her love, endless support.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

People have to admit the fact that *risk* may happen almost everywhere, which have a significant impact on human activities. Risk may bring undesirable event and irresistible disaster that may destroy the existing social framework, hurt people's confidence and cause irreparable pain. Therefore, it is important that, before the occurance of possible risks, it is necessary to assess the risk status, detect the system vulnerability and enhance the prevention capability. We generally call this process *risk management*. In financial economics, the turmoil in stock market, the movements of exchange rates, the wrong commercial decision and so on, would drive to wreck the financial stability of many companies and put them at risk. Among all kinds of risks in finance, the most important concern would be the so-called *market risk*. The market risk calculates the uncertain loss that may occur and arise from the unforeseen movements in market prices of risk factors. Typical risk factors are equity prices, stock market indexes, interest rates, foreign exchange rates and commodity prices.

With the development of globalization of financial markets, external regulators from different countries quickly move toward to risk-based consensus on improving supervision, while risk managers from various firms begin to play more and more important roles on conducting internal risk management. The potential loss due to market risk must be measured in a number of ways. Traditionally, the most popularly used risk measure is the *Value-at-Risk (VaR)*.

## 1.1 Value-at-Risk

Value-at-Risk refers to the worst outcome that is likely to occur at a given confidence level. The practical origins of VaR is said from 1922 in the New York Stock Exchange, while the mathematical roots of VaR were developed already in the context of portfolio theory by Markowitz *et al.* in the 1950s. Since 1994, J.P. Morgan proposed a risk measure, now called the Value-at-Risk (VaR) to quantify the firms' exposure to market risk and made available its RiskMetrics system on the Internet. From then on VaR became widespread and prominent. With the development of the global finance, the Basel Committee on Banking Supervision adopted VaR as a standard tool for regulators. Nowadays VaR is not only used to manage market risk exposure but also other forms of risk, such as the credit risk, the operational risk, the liquidity risk and so on.

In financial theory, the relationship between risk and return are very important. To give the quantitative definition of VaR, consider the log-return first.

Let $P_t$, $t = 1, \ldots, n$ be the observed daily price of an asset, so the simple return of the asset between time $t$ and time $t-1$ is denoted by $r_t = (P_t - P_{t-1})/P_{t-1}$, $t = 1, \ldots, n$. The simple return can be seen an intuitive description of relative price change, but for a variety of reasons it is much easier to use *log-return* (also called the continuously compounded return).

$$R_t = \log(\frac{P_t}{P_{t-1}}) = \log(P_t) - \log(P_{t-1})$$

The distribution of the log-return $R_t$ is known as the *return* *distribution*. Herein, when return $X_i$, $i = 1, \ldots, n$ is referred, it means log-return of an asset.

**Definition 1.1.** Let $X$ be a r.v. whose cumulative distribution function $G(x)$ describes the profit and return distribution (or Propfit & Loss, P & L) of the risky financial asset $X$. For a given confidence level $p \in (0, 1)$, the *Value-at-risk (VaR)* of an asset $X$ at level $p$ is given by

$$VaR_p \equiv \inf_x \{x : \overline{G(x)} \leq 1 - p\}, \tag{1.1}$$

where $\overline{G}(y) = P(Y > y)$.

From the mathematical concept from (1.1), $VaR$ can be seen as the quantile function of an return distribution. Fig. 1.1 from Khindanova & Rachev (2000) displays the examples of VaR under the continuous return distribution and discontinuous distribution, respectively. In practice, the confidence level $p$ is usually selected with 0.95 and 0.99. (Individual trading is typically set at the level 0.95. 0.99 is required by the Basel Committee. For example, banks include Goldman Sachs and Merrill Lynch employ 95% internally while report 99% to regulators.) We now briefly review different aspects of VaR.

◇ **VaR - a risk measure internally and externally**

Choosing a proper risk measure is of great supervisory importance.

The definition of *coherent risk statistic* proposed by Artzner *et al.* (1999) claim that it is better for a risk measure $\rho$ to satisfy: (1) Translation invariance: $\rho(X + a) = \rho(X) + a$; (2) Positive homogeneity: $\rho(\lambda X) = \lambda\rho(X)$; (3) Monotonicity: $\rho(X) \leq \rho(Y),\ if\ X \leq Y$; (4) Subadditivity: $\rho(X + Y) \leq \rho(X) + \rho(Y)$. These axioms are widely accepted. It should be noted that VaR fails to satisfy the last one so that VaR has been criticized as a risk measure. But VaR is still the most prominent measure. Danielsson *et al.* (2005) showed that VaR is subadditive in the

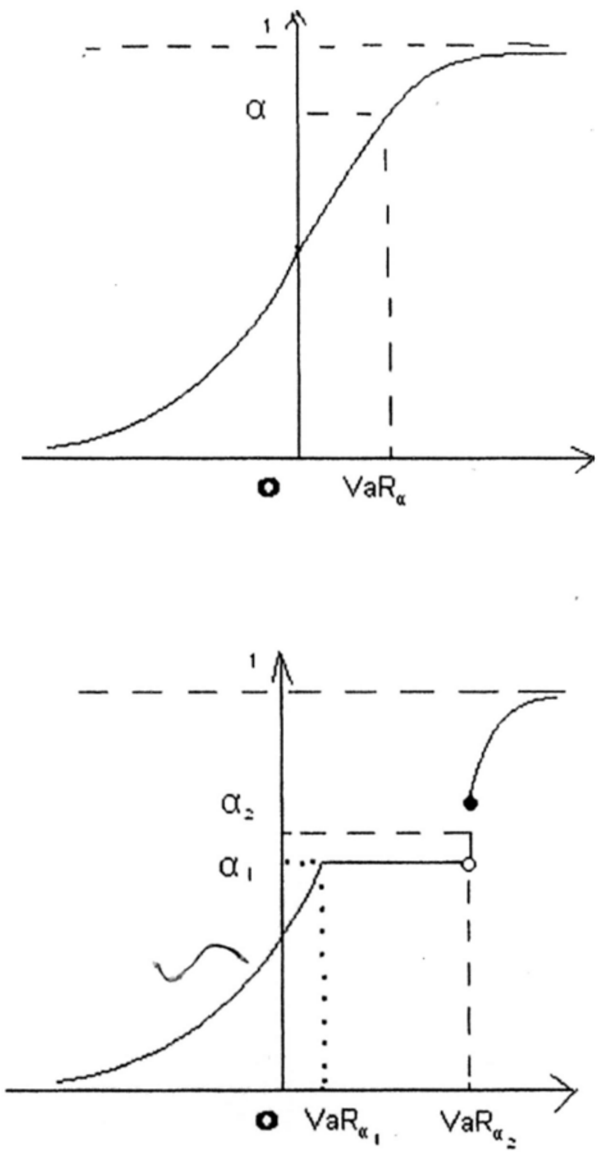


Fig. 1.1: VaR for continuous and discontinuous return distribution.

tail regions if the tails in the joint distribution are not extremely fat. Since asset returns with extremely fat tails are no frequently found, once the extremal events appears, it can be treated as special cases. Therefore, VaR is feasible for internal management.

Heyde, Kou and Peng (2007) clarified how to choose the risk measures, that depends on whether it is for external or internal purpose. The risk measure $\breve{\rho}$ belonging to the proposed class called the *natural risk statistics*, which relaxed subadditivity into comonotonic data, should satisfy the following axioms: (1) Positive homogeneity and translation invariance: $\breve{\rho}(aX + b) = a\breve{\rho}(X)+b, \ \forall\ a \geq 0$; (2) Monotonicity: $\breve{\rho}(X) \leq \breve{\rho}(Y),\ if\ X \leq Y$; (3) Comonotonic subadditivity: $\breve{\rho}(X+Y) \leq \breve{\rho}(X)+\breve{\rho}(Y)$, if $X, Y$ are comonotonic under each scenario; (4) Empirical law invariance: $\breve{\rho}(X) = \breve{\rho}(Y)$, if $X$ and $Y$ have the same order statistics under each scenario. VaR is a special case of the natural risk statistics. Therefore VaR is a rational choice for external regulation. References about the external risk measure can be also found in Ahmed *et al.* (2008), Eling and Tibiletti (2010).

Above all, VaR can be useful both for regulatory purpose and internal purpose. Thus, we can separate the external model with regulatory purpose (such as helping regulators to formulate the guidance on regulatory capital) from the internal model with the business purpose (such as helping the calculation of economic capital). Therefore, VaR calculated from external model should be prudential, while VaR from internal model should pay more attention to accuracy.

**◇ VaR - calculation and evaluation**

VaR is a versatile model. The calculation approaches can be divided into three categories. (A) *Parametric* models. The major representatives of the Gaussian parametric family, are the Delta-normal method (J.P. Morgan - RiskMetrics, 1996). The related improvement is Delata-gamma method (Fong and Vasicek, 1997). There are also non-Gaussian parametric models, for instance, stable distributions (Cheng and Rachev, 1994; Mittnik *et al.*, 2002; Frain, 2008). (B) *Nonparametric* models. For example, bootstrapped historical simulation (Duffie and Pan, 1997) and kernel methods (Chen and Tang, 2005; Epperlein and Smillie, 2006; Cai and Wang, 2008; Huang, 2009). (C) *Semiparametric* models. For example, Extreme value approaches (Embrechts *et al.*, 1997; McNeil and Frey, 2000; Neftci, 2000; Drees, 2003), CAViaR (Engle and Manganell, 2004; Taylor, 2008; Yu *et al.*, 2010) and others ( Fan ,2003; Chen *et al.*, 2006; Chen *et al.*, 2008). Gourieroux *et al.* (2000), Khindanova and Rachev (2000) summarized several methods on VaR calculation. Literatures on the modeling of volatility have Exponential weighted moving average (EWMA) (J.P. Morgan, 1996), ARCH (Engle, 2002), GARCH (Bollerslev, 1986; Bollerslve *et al.* , 1992; Angelidis *et al.*, 2004), threshold-ARCH (TARCH) (Glosten *et al.*, 1993; Zakoian, 1994), exponential GARCH (EGARCH) (Nelson, 1991), Heston model (1993). Poon and Granger (2003) gave a review on volatility modeling.

The evaluation of VaR is called *backtesting* as laid out by the Basel Committee, which compares the actual data with the calculated values. Since future does not always depend on the past, the evaluation of regulatory VaR is challenging. Actually, backtesting may be useful for the accuracy of interval VaR. Literatures dealing with the evaluation of internal VaR have Kupiec (1995), Christoffersen (1998), Engle and Manganelli's DQ test (2004), Christoffersen and Pelletier (2004), Berkowitz et al. (2009). However, practitioners and academics have not reached a common conclusion for the best performing model.

◇ **VaR - from Basel II to Basel III**

The principles devised by the Basel Committee are of central importance in banking sector. From the Basel I Accord (1988) to the current Basel II Accord (1996), the supervision on its member financial institutions are gradually strengthened. VaR has been adopted to determine the capital requirements. According to the Basel II, the *Market Required Capital (MRC)* is calculated by $MRC_t = \max(k\frac{1}{60}\sum_{i=1}^{60} VaR_{t-i},\ VaR_{t-1})$, where $VaR_t$ is the VaR with confidence level 99% by implementing internal models on day $t$ over the preceding 60 days, and $k$ is the *multiplier* ranges from 3 to 4 depending on the backtesting results.

As evidenced by the financial crisis, it is found that the measure for the capital requirement is not conservative enough and there comes the Basel III Accord (2009). The new principle will charge more capital via

$$\max(k\frac{1}{60}\sum_{i=1}^{60} VaR_{t-i},\ VaR_{t-1}) + \max(l\frac{1}{60}\sum_{i=1}^{60} sVaR_{t-i},\ sVaR_{t-1}).$$

It is better the regulatory capital depends on more conservative VaR. More details on VaR can be referred to Jorison (2007).

## 1.2 Multi-period Scaling Problem

The VaR models have been well established and accepted in the single-period risk management. Banks use their daily VaR for controlling internal risks every day. However, it is not enough to only consider the single-period return. We should also consider the long-term development. Undoubtedly, the multi-period model will become increasingly important.

For internal multi-period risk supervision, the *holding period* can take any time length, and depends on the liquidity of assets and the frequency of trading transactions. Beder (1995) analyzed the significance of holding period and concluded the necessity of larger VaR for long horizons. Actually, in practical external model, with regulatory purpose, the holding period required by regulators is set to 10-day. And the multi-period VaR model advocated by The Basel Committee (1996), is called the *square-root-of-time rule* (SQRT rule), which is derived based on the assumption that the asset returns follow the Gaussian distribution. The banking supervision institution would ask its member to report their 10-day VaR for potential long-horizon risk regulation. By using SQRT rule, the $n$-day VaR can be represented through the 1-day VaR.

$$VaR^{(n)} = \sqrt{n}VaR^{(1)}.$$

The Gaussian-based scaling approach is simple and fast but it has a serious drawback that it fails to capture the feature of the heavy tails. Kaufmann (2005) carefully studied the long-term risk management and found the SQRT rule can perform a good approximation under many all single-period financial models.

However, different models may have different scaling rules. For example, the multi-period model for complex model such as AR(1)-GARCH(1,1) cannot be calculated analytically. The SQRT rule as a routine is uncritically used and accepted by many banks and financial institutions. Operationally, some financial institutions are even interested to extrapolate 1-day to 252-day VaR. Many banks use it as an internal risk management model. For regulators the misuse is also widespread. Kupiec and O'Brien (1995) pointed out that the simple SQRT rule does not hold when the market risk factors are non-Gaussian. Dielbold *et al.* (1998) firstly criticized that an inappropriate use of SQRT rule would overestimate the variability of long-horizon volatility. However, the SQRT rule still occupied the long-run risk model. The Basel II Accord has proposed a *multiplier* (or multiplication factor) between 3 and 4 on banks' internal 99% VaR with holding period 10. The multiplier requires banks to hold more buffering capital. Stahl (1997) tried to advance the theoretical justification for the multiplier based on Chebyshev Inequality. Stahl's view is seriously refuted by Danielson *et al.*

(1998) and they criticized a lot on the conservatism of existing rule and studied new scaling approach based on the assumption of extreme returns.

Comparing with the Gaussian distribution, financial data always exhibits some asymmetric, fatter tail and excess kurtosis (leptokurtic or platokurtic) properties. Many papers (Danielson and De Vries, 1998; 2000; Dacorogna *et al.*, 2001) criticized the Gaussian assumption for its underestimation of short-term potential risk and pointed out the overestimation of long-term risk by using the SQRT rule. According to their suggestions, they proposed *α-root rule*, which is derived from the heavy-tailed (or power-law tailed) distributions.

$$VaR_p^{(n)} \sim n^{1/\alpha} VaR_p^{(1)}, \text{ as } p \to 1,$$

where $\alpha$ represents the tail index in the heavy-tailed return distribution $P(X_t < -x) = x^{-\alpha}L(x),\ as\ x \to \infty$, and $L$ is a slowly varying function satisfying $\lim_{x\to\infty} L(sx)/L(x) = 1$ for all $s > 0$. The natural estimator for the tail index is the so-called Hill estimator (Hill, 1975). Danielsson and Vries (2000) developed a semiparametric estimator of VaR by using order statistics based on the assumption of power-law tailed distribution. Many banks with risk-seeking spirit began to employ $\alpha-$root rule internally.

Another scaling rule similar to the $\alpha-$root rule depends on the the hypothesis of self-similarity. The self-similarity is one of the empirical properties of real asset returns (Cont, 2001). Theoretically, self-similarity means the distribution of asset returns at day $n$ equals in law to $n^\gamma$ multiplies with the distribution at day 1. The coefficient $\gamma$ represents the Hurst index (or scaling exponent) and the scaling rule for VaR has $VaR^{(n)} = n^\gamma VaR^{(1)}$, which can be found in Mentegna and Stanley (1995) for the Lévy stable case. Then Provizionatou *et al.* (2005) derived two empirical calculating or estimating approaches for the scaling exponent which has locally determined time-varying properties.

Menkens (2007) analyzed the empirical relationship of VaR and self-similarity.

Within the last few years, the financial world has undergone a serious financial disaster. This questioned the current risk management practice about calculation of VaR. Someone is going to ask: is the regulation too conservative or not conservative enough? The Gaussian return may not be feasible. But as already referred, extremely fat tail seldom appears, either. In Danielsson and Zigrand (2005), by assuming the returns follow a jump diffusion processes, the authors found traditional time scaling VaR would lead to a systematic underestimation of risk and the worsen status may closely relate to the time horizon and confidence level. Courtois and Walter (2010) studied the multiplier rule and VaR with the Variance Gamma processes with drift (VGPD). They empirically verified in the multi-period problem, the risk must be highly related to the horizon while the difficulty lies in the choice of horizon. The 30-day VaR may be two times as big as the 10-day VaR. If setting longer horizon, the VaR may become much greater. The optimal horizon is about 30-day based on the 1-year forecast. Other literatures about multi-period risk management are: Dowd *et al.* (2003) talked about the inaccuracy and problems of SQRT rule and tried extrapolating volatility forecasts over longer horizons; Giannopoulos (2003) explored the VaR modeling based on general Filtered Historical Simulation. About the temporal aggregation of GARCH processes, Drost and Nijman (1993) supplied and generalized the GARCH model and derived the aggreation formula which can be applied to the field of risk management. Hafner (2008) generalized Drost and Nijman's formula into multivariate GARCH processes. Goldberg *et al.* (2008) analyzed the 5-day and 10-day VaR based on the EVT and RiskMetrics single-period model. Eberlein and Madan (2009) studied the scaling and distribution of returns at long horizons, which finally developed a mixed approach and found the advocated strategy for constructing long horizon return by running a Lévy process half the daily return while scaling the remainder at rate 1/2. There are more specific working papers on the long-term risk management from the RiskLab in ETH, Zurich (Kaufmann and Patie, 2003; Embrechts *et al.*, 2004; Brummelhuis and Kaufmann, 2004).

People had developed the scaling approach that based on heavier tailed distribution. In addition to distributions with power-law tail, there are also other heavier tailed distribution classes such as convolution equivalent tail (Watanabe and Yamamuro, 2010a) and semi-heavy tail (Barndorff-Nielsen, 1998). Those tails are heavier than tails of the Gaussian distribution and feasible for financial applications. For both external and internal risk supervisions, we can concentrate on these distributions with more accurate tail to derive new multi-period scaling rules. The thesis is just based on this kind of idea.

**◇ Contributions of the thesis**

The aim of the thesis is to develop new scaling methodologies for multi-period financial risk management.

(I) The first contribution of this thesis is, the derivation and proof of a prudential scaling rule, under a more suitable assumption than Gaussian, which is in line with the expectations of the public. The rule is named *convolution equivalent rule* (CE rule). This would contribute to the justification of increasing capital requirements. (II) The second contribution lies on the improvement of internal scaling approach. A more accurate scaling rule based on semi-heavy tailed assumption has been proved that it can provide more accurate VaR than the traditional SQRT rule. It is called *semi-heavy rule* (SH rule). (III) During the process, we show the proof of the semi-heavy tail property of the NEF-GHS distribution. (IV) We also employed a semiparametrc estimation (SP) approach of single-period VaR for the semi-heavy tailed distributional assumption. (V) The application of saddlepoint approximation to the multi-period VaR would be regarded as an additional contribution.

## 1.3 Outline of the thesis

We will show how the scaling rules (CE rule and SH rule), the semiparametric estimation (SP), the parametric framework (NIG) and the saddlepoint method derived and used in the calculation of multiperiod VaR. The structure of this thesis is organized as follows. Chapter 2 briefly introduces the background of exponential class and the related knowledge. Thus a scaling rule based on convolution equivalenc (CE rule) is derived. Chpter 3 focuses on the semi-heavy tail and constructs a numerical semiheavy tail rule (SH rule). It gives a semi-parametric estimation of the single-period VaR and explains how to apply derived scaling rule for multi-period supervision. It also takes a parametric distribution - normal inverse Gaussian (NIG) distribution as an example to show specific scaling rules. In Chapter 4, we employ the saddlepoint approximation to the multi-period VaR when the distributional assumption is the NIG distribution. Chapter 5 offers some simulations on the scaling rules, semiparametric estimation and saddlepoint approximation. It also conducts a real data analysis on stock prices of bankings. The final chapter summarizes the results and gives further discussion.

---

□ **End of chapter.**

# Chapter 2

# Exponential Class and Scaling Rule

**Summary**

This chapter firstly introduces the definitions of two distribution classes - the exponential class and the convolution equivalent class. Then we give some theoretical background about these classes, their properties, the relationships between classes and some related theorems. Consequently, a scaling theorem relating single-period VaR to multi-period VaR based on the convolution equivalence assumption will be derived and proved. Finally we will take the NEF-GHS distribution as an example to exhibit the rule.

The inadequacy of the Gaussian distribution have been frequently discussed. When managing risk, people have always reported that the financial asset return follows a heavier tailed distribution so that modeling the extreme market risk should become the goal. How heavy the tail of distribution should be? It is a problem. A lot of alternative distributions with property of heavier tail were proposed to fit the financial data, such as the stable distribution (McCulloch, 1996), the power-law tailed distribution (Praetz, 1972) and the generalized extreme value (GEV) distribution (Embrechts *et al.*, 1997). The referred distributions have all received considerable attention, but it seems that the extremely heavy tail may be too heavy as no extreme event happens. There are another class of distributions with exponential-like tails, the distributions with tail that are lighter than the power-law tailed distribution. Some distributions in these classes have heavier tails than the Gaussian tail. In this chapter, we shall introduce one such class - convolution equivalent class. Let us begin to introduce the class.

## 2.1 Basic Theoretic Background

In this section, two distribution classes with benign tail property are presented. These kinds of distribution classes are first introduced by Christyakov (1964) for $\gamma = 0$ (to be defined later), independently proposed by Chover, Ney and Wainger (1973) for $\gamma > 0$. Until now they have received a lot of attention from the academic world in probability and stochastic processes. There is little literature in studying their applications in the field of financial risk management. We now focus on this area and define the *exponential class.*

**Definition 2.1.** Let $Y$ have distribution function $G(\cdot)$. Then it has an exponential tail with rate $\gamma \geq 0$ if

$$\lim_{x\to\infty} \frac{\overline{G}(x-y)}{\overline{G}(x)} = e^{\gamma y} \qquad (-\infty < y < \infty), \tag{2.1}$$

where $\overline{G}(y) = P(Y > y)$. This class of distributions is denoted by $\mathfrak{L}_\gamma$ called the *exponential class.*

Because the tail probability of the exponential distribution $Exp(\lambda)$ satisfies $\frac{\overline{G}(x-y)}{\overline{G}(x)} = \frac{\exp[-\lambda(x-y)]}{\exp[-\lambda(x)]} = e^{\gamma y}$, it is obvious that exponential distribution with p.d.f. $g(x) = \lambda\exp(-\lambda x)$ belongs to $\mathfrak{L}_\lambda$. In the literature, a distribution $G$ in $\mathfrak{L}_\gamma$ with $\gamma > 0$ is usually said to have and an exponentially decreasing tail. For every $G \in \mathfrak{L}_\gamma$, it satisfies the property

$$\lim_{x\to\infty} \overline{G}(x)e^{\epsilon x} = \infty, \ \forall\, \epsilon > \gamma.$$

**Definition 2.2.** A cumulative distribution function (c.d.f.) $G$ with support $[0,\infty)$ is said to belong to the *convolution equivalent class* $\mathfrak{S}_\gamma$ if $G \in \mathfrak{L}_\gamma$ and

$$\lim_{x\to\infty} \frac{\overline{G*G}(x)}{\overline{G}(x)} \text{ exists and is finite,} \tag{2.2}$$

where $G*G(u) = \int_{\mathbb{R}} G(u-x)\, dG(x)$.

More generally, a distribution $G$ on $(-\infty,\infty)$ is also said to belong to $\mathfrak{S}_\gamma$ if $G_+$ does. In the literature, $G$ is said to have a convolution equivalent tail.

From the above definitions, we know that $\mathfrak{S}_\gamma \subseteq \mathfrak{L}_\gamma$. In fact, exponential distribution $Exp(\lambda)$ with p.d.f. $g(x) = \lambda\exp(-\lambda x)$ belongs to $\mathfrak{L}_\gamma - \mathfrak{S}_\gamma$ which means the exponential distribution is not convolution equivalent. Since exponential distribution just belongs to $\mathfrak{L}_\gamma$, we call $\mathfrak{L}_\gamma$ the exponential class.

In the paper of Watanabe (2008), Lemma 2.5 asserts that if $G_1, G_2 \in \mathfrak{L}_\gamma$, then $G_1 * G_2 \in \mathfrak{L}_\gamma$. Particularly, if $G \in \mathfrak{L}_\gamma$, $G^{*n} \in \mathfrak{L}_\gamma$ for all $n \geq 1$, where the representation $G^{*n}$ means the $n$-fold convolution. This is the convolution property for $\mathfrak{L}_\gamma$. From the paper of Watanabe and Yamamuro (2010b), infinite divisible distributions in $\mathfrak{S}_\gamma$ will be closed under convolution roots, or generally $\mathfrak{S}_\gamma$ does not have the convolution property as $\mathfrak{L}_\gamma$. Applying Karamata's representation, Albin and Sundén (2009) give the following representation form of distributions from the class $\mathfrak{L}_\gamma$. That is, an absolutely continuous c.d.f. $G$ belongs to $\mathfrak{L}_\gamma$ if and only if

$$G(u) = 1 - \exp\{-\int_{-\infty}^{u} (a(x) + b(x))\, \mathrm{d}x\}, \ \forall u \in \mathbf{R}$$

for some measurable functions $a$ and $b$ with $a+b \geq 0$ such that

$$\lim_{x\to\infty} a(x) = \alpha, \ \lim_{u\to\infty} \int_{-\infty}^{u} a(x)\,\mathrm{d}x = \infty \text{ and } \lim_{u\to\infty} \int_{-\infty}^{u} b(x)\,\mathrm{d}x \text{ exists.}$$

It also demonstrates distribution with form when $a(x) = \alpha \mathbf{I}_{[0,\infty)}$ and $b(x) = \beta \cos(e^x - 1)\mathbf{I}_{[0,\infty)}$, where $|\beta| \leq \alpha$ and $\alpha > 0$ are fixed. Then $G(u) = 1 - \exp\{-\int_{-\infty}^{u}(a(x) + b(x))\mathrm{d}x\}$ is in the class $\mathfrak{L}_\gamma$ because $\lim_{u\to\infty} \int_0^u \cos(e^x - 1)\,\mathrm{d}x$ exists.

The special cases of $\mathfrak{S}_\gamma, \mathfrak{L}_\gamma$ when $\gamma = 0$ are called *subexponential distribution* $\mathfrak{S}$ and *long-tailed distribution* $\mathfrak{L}$. (that is, the distributions in $\mathfrak{S}$ or $\mathfrak{L}$ have tails that decay slower than the exponential tail.) For $G \in \mathfrak{L}$, it satisfies $\lim_{x\to\infty} \frac{\overline{G}(x-y)}{\overline{G}(x)} = 1$ and for $G \in \mathfrak{S}$, it satisfies $\lim_{x\to\infty} \frac{\overline{G*G}(x)}{\overline{G}(x)} = 2$. If the tail probability of distribution $\overline{G}$ is regularly varying (Bingham, Goldie and Teugels, 1987) at infinity with a non-positive index, which means $\lim_{u\to\infty} \frac{\overline{G}(ux)}{\overline{G}(u)} = x^\rho, \rho \leq 0$, so $G \in \mathfrak{S}$. For example, given constants $x_0, \rho > 0$, the tail probability of Pareto distribution $G(x) = 1 - (\frac{x}{x_0})^{-\rho}, x \geq x_0$ is regularly varying with index $-\rho < 0$. It is shown that $G \in \mathfrak{S} \subseteq \mathfrak{L}$. Subexponential family is useful for studying the applications of probabilistic theory in insurance and finance. Literatures about subexponential distribution can be found in Teugels (1975), Embrechts *et al.* (1979), Embrechts and Goldie (1980), Klüppelberg (1989), Goldie and Klüppelberg (1998), Rogozin (2000), Shimura and Watanabe (2005), Watanabe and Yamamuro (2010b).

Here we will pay attention to distributions that belongs to the convolution equivalent class $\mathfrak{S}_\gamma$. It is found the following theorem about the limit in its definition.

**Theorem 2.1** (Chover-Ney-Wainger, 1973a; Cline, 1986; Pakes, 2004). *For any distribution $G \in \mathfrak{S}_\gamma$ with rate $\gamma \geq 0$, the limit can* *be represented as*

$$\lim_{x\to\infty} \frac{\overline{G * G}(x)}{\overline{G}(x)} = 2M, \tag{2.3}$$

*where* $M = M_G(\gamma) = \int e^{\gamma x}\, dG(x)$.

**Proof:** The proof of this theorem can be found in Foss and Korshunov (2007), where they only use analytic and direct probabilistic methods. Earlier proofs based on the Banach algebra technique can be found in Chover, Ney and Wainger (1973), Cline (1986) and Pakes (2004).

The analytical form of the limit has been obtained from Theorem 2.1, which is about the tail of 2-fold convolution. The natural question that will be proposed immediately is what happens when it is $k$-fold convolution. The answer is given in the following theorem.

**Theorem 2.2** (Chover-Ney-Wainger, 1973b; Embrechts-Goldie, 1982). *For every* $n \geqslant 1$, *if* $G \in \mathfrak{S}_\gamma$ *with rate* $\gamma \geq 0$, *then*

$$\lim_{x\to\infty} \frac{\overline{G^{*n}}(x)}{\overline{G}(x)} = nM^{n-1}, \tag{2.4}$$

*where the representation* $G^{*n}$ *means the* $n$*-fold convolution*

$$\overbrace{G * G * \cdots * G}^{n}.$$

**Proof:** This theorem is an immediate consequence of Theorem 2.1. Details about the proof of this theorem can be found in Chover, Ney and Wainger (1973b), Embrechts and Goldie (1982).

The special case of this theorem is under the assumption of subexponential distribution class. If $G \in \mathfrak{S}$, then for every positive integer $n$, $\lim_{x\to\infty} \overline{G^{*n}}(x)/\overline{G}(x) = n$. This property has been widely used in insurance, portfolio theory and risk management. The $\alpha-$root rule (Danielsson and De Veries, 1998) has been derived based on this subexponential property (Feller, 1971).

Correspondingly, we can define the density classes just as the definitions of distribution classes. For two measurable functions $g_1$ and $g_2$ which is defined from the support $[0, \infty)$ to the image $[0, \infty)$, the convolution of two densities $g_1$ and $g_2$ is denoted by

$$g_1 \star g_2(x) = \int_0^x f(x-y)g(y)dy.$$

Denote $g^{2\star} = g \star g$ and $g^{n\star} = g \star g^{(n-1)\star}$, for every positive integer $n$ greater than 2. According to Klüppelberg (1989), a measurable function $g : [0, \infty) \to [0, \infty)$ is said to belong to the class $\mathfrak{L}\mathfrak{d}_\gamma$ for $\gamma \geq 0$ if $g(x) > 0$ for all large $x$,

$$\lim_{x\to\infty} \frac{g(x-y)}{g(x)} = e^{\gamma y}, \ y \ \in (-\infty, \infty),$$

and if $g \in \mathfrak{L}\mathfrak{d}_\gamma$, it satisfies the condition

$$\lim_{x\to\infty} \frac{g^{2\star}(x)}{g(x)} = 2c$$

exists and finite. Then we can define $g \in \mathfrak{S}\mathfrak{d}_\gamma$. From the above theorems, it is known that the constant $c$ satisfies, $c = \int_0^\infty g(x)\mathrm{d}x$. For a distribution $G$ with a density $g \in \mathfrak{L}\mathfrak{d}_\gamma$ for $\gamma > 0$, it is found that $g(x)/\overline{G} \to \gamma$. In this case, $G \in \mathfrak{S}_\gamma$ if and only if $g \in \mathfrak{S}\mathfrak{d}_\gamma$. For further detail, please refer to Klüppelberg (1989). The density classes can also be easily extended to densities in the entire real line.

More literatures about the exponential class and convolution equivalent class can be found in the following papers. Chover, Ney and Wainger (1973), Embrechts and Goldie (1980, 1982) and Cline (1987) are about the theoretical base concerning the one-sided distribution. Braverman (1995,1997) studied the probabilistic application. Pakes (2004, 2007) proved more properties and theories on the two-sided distributions. A systematic discussion about the convolution equivalent is Foss & Korshunov (2007). Albin and Sundén (2009) gave the relationship between the semi-heavy tailed distribution class to the two classes. Yu, Wang & Yang (2010) study the property of the closure of distribution from convolution equivalent class under convolution roots. Fasen (2009), Watanabe and Yamamuro(2010a), Wang & Wang (2011) give some new developments and the probabilistic applications in this area.

**Example 2.1.** *Inverse Gaussian (IG) Distribution*

The convolution equivalent class $\mathfrak{S}_\gamma$ with $\gamma > 0$ contains *inverse Gaussian (IG) distributions* (also known as the Wald distribution, see Johnson, Kotz, & Balakrishnan, 1994), which has been widely applied. The IG distribution forms a two-parameter family of continuous probability distributions with support on $(0, \infty)$. The general forms of the probability density and cumulative distribution of an inverse Gaussian distribution are given by

$$f_{IG}(x) = \frac{\alpha}{\sqrt{2\pi\beta x^3}} \exp\left\{-\frac{(\beta x - \alpha)^2}{2\beta x}\right\}, \tag{2.5}$$

$$F_{IG}(x) = \Phi\left(\frac{1}{\beta x}(\beta x - \alpha)\right) + \exp\{2\alpha\}\, \Phi\left(-\left(\frac{1}{\beta x}(\beta x + \alpha)\right)\right),$$

for $x > 0$, where $\alpha > 0, \beta > 0$ and $\Phi(\cdot)$ is standard normal distribution. The IG distribution with parameters $\alpha$ and $\beta$ is denoted by IG$(\alpha, \beta)$. According to the main theorem of Embretchs (1983), IG$(\alpha, \beta)$ belongs to the class $\mathfrak{S}_{\beta/2}$.

If $X \sim$IG$(\alpha, \beta)$, the mean, variance, skewness and kurtosis are obtained,

$$\mu = E(X) = \alpha/\beta,\ Var(X) = \alpha/\beta^2$$

$$Skewness(X) = 3\sqrt{1/\alpha},\ Kurtosis(X) = 15/\alpha.$$

The moment generating function of IG$(\mu, \nu)$ is given by

$$M(t) = \exp\{\alpha(1 - \sqrt{1 - \frac{2t}{\beta}})\}.$$

Let $\mu = \alpha/\beta, \nu = \alpha^2/\beta$. We have another parameterization of the IG distribution which is denoted by IG$(\mu, \nu)$. Some basic properites of the inverse Gaussian (IG) distributions are:

1. If a random variable $X$ is distributed by IG$(\mu, \nu)$, then for each $t > 0$, the random variable $tX$ is distributed by IG$(t\mu, t\nu)$.

2. If $X_1, \ldots, X_n$ are independent random variables with $X_i$ distributed by IG$(\mu_0\omega_i, \nu_0\omega_i^2)$, $i = 1, \ldots, n$, then $\sum_{i=1}^n X_i$ is distributed by IG$(\mu_0\omega_{(n)}, \nu_0\omega_{(n)}^2)$ where $\omega_{(n)} = \sum_{i=1}^n \omega_i$.

The inverse Gaussian (IG) distribution has several properties analogous to the Gaussian distribution. Its cumulant generating function (denoted by c.g.f., logarithm of the moment generating function (m.g.f.)) can be represented as the inverse of the c.g.f. of a Gaussian random variable. While the Gaussian describes a Brownian Motion's level at a fixed time, the inverse Gaussian describes the distribution of the time a Brownian Motion with positive drift takes to reach a fixed positive level. The inverse Gaussian (IG) distributions have been applied to a wide range of fields. Most of these applications are based on the idea of first passage times of a Brownian motion with drift. These fields include actuarial science, demography, employment management, finance, and even linguistics; see Seshadri (1999) and references therein.

The *normal inverse Gaussian (NIG)* is a convolution equivalent distribution to be discussed in the next chapter. Later we shall prove how it belongs to the convolution equivalent class. Actually, the NIG distribution is a normal mixture distribution. A random variable $X$ with the distribution NIG$(\alpha, \beta, \mu, \delta)$ is defined as follows. If

$$X|Y = y \sim N(\mu + \beta y, y)$$

$$Y \sim \mathrm{IG}(\delta\gamma, \gamma^2) \text{ with } \gamma = \sqrt{\alpha^2 - \beta},$$

where $0 \le |\beta| < \alpha$ and $\delta > 0$. We then define $X \sim \mathrm{NIG}(\alpha, \beta, \delta, \mu)$ and denote the density (p.d.f) by $f_{NIG}(x; \alpha, \beta, \mu, \delta)$. Then

$$f_{NIG}(x; \alpha, \beta, \mu, \delta) = \int_0^\infty f_G(x; \mu + \beta y, y) \cdot f_{IG}(y; \delta\gamma, \gamma^2) \,\mathrm{d}y,$$

where $f_G(x; \mu+\beta y, y)$ represents the Gaussian distribution with the mean parameter equal to $\mu+\beta y$ and the variance parameter equal to $y$ and $f_{IG}(y; \delta\gamma, \gamma^2)$ represents the inverse Gaussian distribution with density (2.5) with parameters $\alpha = \delta\gamma, \beta = \gamma^2$. The NIG distribution belongs to $\mathfrak{S}_{\alpha-\beta}$ and it has a closed form p.d.f., that we will see in the next chapter.

The IG distribution have many parameterizations. For example, when $a = \sqrt{\beta}, b = \alpha/\sqrt{\beta}$, it becomes a parameterization from Barndorff-Nielsen (1998) of the IG distribution which is denoted by $\mathrm{IG}(a, b)$. It can be generalized to a distribution class the generalized inverse Gaussian (GIG) distribution by adding one parameter $p$. Then the $\mathrm{GIG}(a, b, p)$ belongs to $\mathfrak{S}_{a^2/2}$ (same as the IG distribution but it has conditions on the other parameters) only if $a > 0, b > 0, p < 0$. The $\mathrm{IG}(a, b)$ is a special case of $\mathrm{GIG}(a, b, p)$ when $p = -1/2$.

Some distributions include generalized hyperbolic (GH) distribution (Eberlein and Hammerstein, 2003) are convolution equivalent when the parameters are restrained to some region (the NIG distribution is a subclass of the GH distribution). Also for example - the $\mathrm{CGMY}(C, G, M, Y)$ proposed by Carr, Geman, Madan and Yor (2002) distribution, is defined on $C, G, M > 0, Y \le 2$. It belongs to $\mathfrak{S}_M$ when $0 < Y < 2$. Please refer to Schoutens (2003) for an overview of the above distributions. The proof of convolution equivalence is not direct. It will be demonstrated when we introduce the concept of semi-heavy tail in the next chapter.

## 2.2 Scaling Rule under Convolution Equivalence

In this section, the newly derived scaling rule is based on the assumption of a specific distribution class - *convolution equivalent class.*

Under the assumptions of convolution equivalence, applying the definition of VaR and Theorem 2.4 we obtain a scaling theorem for VaR of multi-period from VaR of single-period.

**Theorem 2.3.** *Assume $G(x)$ is the distribution of return of a given asset. Let $G(x) \in \mathfrak{S}_\gamma$. Then the VaR of $n$-days ($VaR_p^n$) can be asymptotically obtained from the VaR of 1-days ($VaR_p^1$) via the following scaling rule*

$$VaR_p^n \sim VaR_p^1 + \frac{1}{\gamma}[\log n + (n-1)\log M], \text{ as } p \to 1. \quad (2.6)$$

**Proof:** Let $X \sim G$ and $\{X_t\}, t = 1, 2, \ldots$ be i.i.d. observations of log-returns from $G$. Then

$$1 - p = P(X > VaR_p^1) = \overline{G}(VaR_p^1).$$

when $p \to 1, VaR \to \infty$. From definition of VaR, for a fixed suitable number $-\infty < y < +\infty$, we have

$$\overline{G}(VaR_p^1) \sim \exp[\gamma(y - VaR_p^1)]\overline{G}(y), \text{ as } VaR_p^1 \to \infty,$$

$$1 - p = e^{\gamma(y - VaR_p^1)}\overline{G}(y)(1 + o(1)). \quad (2.7)$$

Assume distributions of returns of $n$ days $X^{(n)} = \sum_{i=1}^n X_i$ are the $n - fold$ convolution of $G$, $G^{*n}$.

$$1 - p = P(X^{(n)} > VaR_p^n) = \overline{G^{*n}}(VaR_p^n).$$

It is known that when $p \to 1, VaR \to \infty$.

Using Theorem 2.4, $\overline{G^{*n}} \sim nM^{n-1}\overline{G}$, for large $x$,

$$P(X^{(n)} > VaR_p^n) = nM^{n-1}\overline{G}(VaR_p^n)(1 + o(1)).$$

Also from the definition of VaR, for fixed $-\infty < y < +\infty$,

$$\overline{G}(VaR_p^n) \sim \exp[\gamma(y - VaR_p^n)]\overline{G}(y),$$

$$1 - p = nM^{n-1}e^{\gamma(y-VaR_p^n)}\overline{G}(y)(1 + o(1))^2. \tag{2.8}$$

Comparing (2.7) and (2.8), we obtain $e^{\gamma(VaR_p^n - VaR_p^1)} \sim nM^{n-1}$, as $p \to 0$. Thus

$$\gamma(VaR_p^n - VaR_p^1) \sim \log n + (n-1)\log M, \ p \to 1,$$

Equation (2.6) is established and the proof is complete.

We know that the inverse Gaussian (IG) distribution satisfies the condition of convolution equivalence. Since the IG distribution just has support on the positive real line, it has not been widely applied in the model of market risk. We introduce another distribution which may have greater potential. That distribution is the NEF-GHS distribution, when restricting its parameter in some region, it belongs to the convolution equivalent class $\mathfrak{S}_\gamma$. Consider the respective scaling rules for the NEF-GHS distribution under convolution equivalence.

**Example 2.2.** *The NEF-GHS scaling rule*

The *Natural Exponential Family - Generalized Hyperbolic Secant* (NEF-GHS) distribution, which was originally introduced by Morris (1982) in the context of *natural exponential families* (NEF) with specific *quadratic variance functions* (QVF). Density of the NEF-GHS distribution is of the form (without loss of generality, location and scale parameter are not introduced here).

$$f(x; \lambda, \theta) = \exp\{\lambda x - \psi(\lambda, \theta)\} \cdot \zeta(x, \lambda).$$

In the case of NEF-GHS distribution, $\psi(\lambda, \theta) = -\lambda \log(\cos(\theta))$ and $\zeta(x, \lambda)$ is the p.d.f. of *generalized hyperbolic distribution*

*(GHS)* (a detailed discussion can be found in Baten (1934), Harkness and Harkness (1968) and Jørgensen (1997)). A GHS distribution is the $\lambda$-th convolution of a distribution called *hyperbolic secant (HS) distribution*, which is symmetric and has leptokurtic properties with p.d.f. $f_{HS} = \frac{1}{2\cos(\pi x/2)}$, $x \in \mathbb{R}$. GHS is still symmetric and has higher leptokurtic flexibility but cannot take skewness into account while NEF-GHS can. This NEF-GHS distribution has most of the good properties of the simple hyperbolic secant (HS) distribution (it is a more peaked than the normal distribution).

In the next chapter, as an example for proving the property of semi-heavy tailness, we will show how to derive the semi-heavy tail property of the NEF-GHS distribution. Here let us just write down the result from Example 3.1 (see Chapter 3), which says when $\theta < 0$, the NEF-GHS belongs to the convolution equivalent class $\mathfrak{S}_{-\theta}$. Since the moment generating function (m.g.f.) is $M(u) = \exp\{-\lambda \log[\cos(u) - \beta \sin(u)]\}$, thus the respective NEF-GHS scaling rule for multi-period VaR can be represented as

$$VaR_p^n \sim VaR_p^1 - \frac{1}{\theta}[\log n + (n-1)(-\lambda \log(\cos\theta + \beta\sin\theta))], \text{ as } p \to 1.$$

About scaling theorem - Theorem 2.2, people can easily obtain an asymptotic multi-period VaR when the single-period VaR is given. If $\gamma$ is known, this asymptotic scaling model can be reasonable and especially good for higher quantile VaR 99% and multi-period VaR with long holding periods. This rule is based on the assumption of convolution equivalent tailed distribution, which can be called *convolution equivalent rule* (CE rule). We should notice that it is an addition rule compared with traditional *square-root-of-time rule* (the SQRT rule is a multiplication rule). Although distributions with convolution equivalent tail possess very excellent properties, it is still not easy for people to apply the CE rule because the estimation for the parameter $\gamma$ seems a tough task for practitioners. For application of the CE rule, we first define a new class of distributions called - *semi-heavy tailed distribution class*, which is closely related to the convolution equivalent class. Some distributions with semi-heavy tail are also convolution equivalent.

□ **End of chapter.**

# Chapter 3

# Semi-Heavy Tailed Scaling Calculation

**Summary**

Many empirical findings assert that, the returns of most financial assets exhibit certain semi-heavy tails. One of the popular models - normal inverse Gaussian (NIG) distribution possesses the semi-heavy tails. In this chapter, given the single-period VaR, we will conduct a numerical scaling calculation for multi-period VaR, based on the assumption of semi-heavy tailed distribution which is convolution equivalent. In fact, the numerical calculation rule (SH rule) can be a very useful scaling method to construct long-term internal model for financial risk managers. Furthermore, a semiparametric estimation (SP) for empirically estimating the parameters of the semi-heavy tail is derived. Combining with its scaling rule, we provide a systematic multi-period VaR model (SP-SH rule). We finally take the NIG distribution as an example (NIG-SH rule) to show how to write down the parametric semi-heavy tail rule.

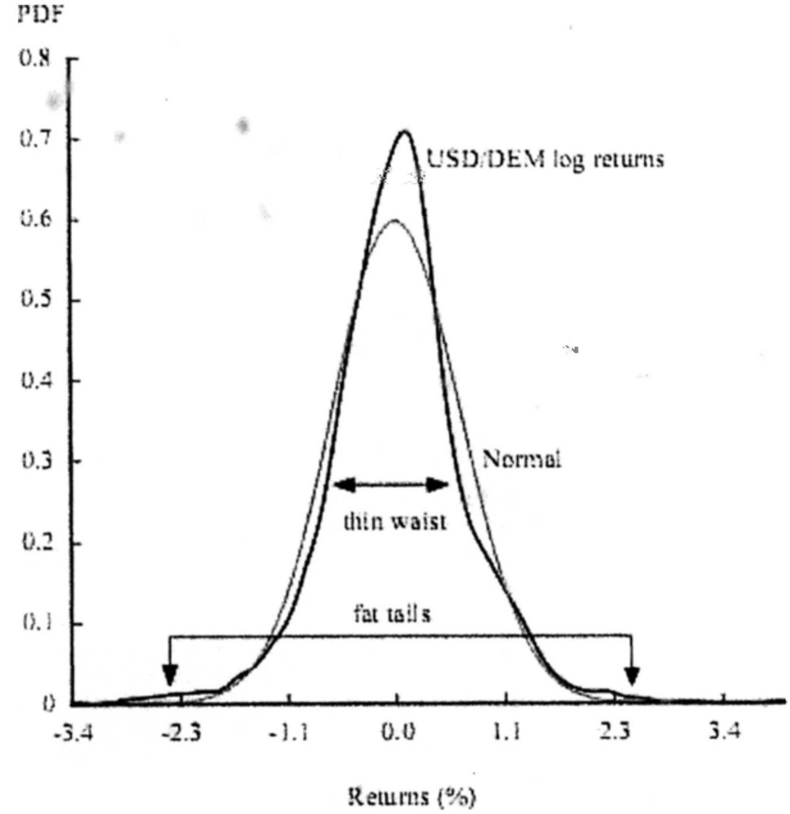


Fig. 3.1: The real USD/DEM data.

The empirical properties of real financial data have been frequently discussed. Many data sets exhibit heavy tails, skewness and high kurtosis. Fig. 3.1 shows that the real foreign exchange rate data DEM/USD possesses fat tail and thin waist while the fitting with Gaussian distribution would become inadequate. The fat tail phenomenon (also called extreme events), tends to occur more frequently than the normal distribution would predict. Consequently, greater emphasis has been placed on using distributions with fatter tails that give a larger weighting to extreme events. Literatures on heavy tails concentrate on power-law tail (Newman, 2005; Clauset *et al.*, 2009) such as Pareto distribution, stable distribution, and Extreme Value distributions. A review paper can be referred to Bradley and Taqqu (2001). Financial data is not extremely heavy as power-law tail does, however, there is another group of research studying exponential-like tail. Heyde and Kou (2004) analyzed the controversy between the two types of tails. Meerschaert and Roy (2010) proposed an exponential-like distribution, and called it - tempering Pareto distribution. In other papers (Fenner *et al.*, 2005; Clauset *et al.*, 2009), this distribution is called the power law with exponential cutoff. We follow Barndorff-Nielsen (1998)'s name - *semi-heavy tail.* The semi-heavy tailed distribution may be the most feasible heavy-tailed distribution in finance for modeling the log-return of asset prices. It is well known that the distribution for asset returns have semi-heavy tails, i.e., actual kurtosis is higher than the kurtosis of the normal distribution. But semi-heavy tail is more than the NIG distribution. Actually, there are still other properties of semi-heavy tail stimulating further study.

In this chapter, we will deduce the scaling properties based on the semi-heavy tailed distribution class.

## 3.1 The Semi-heavy Scaling Approach

**Definition 3.1.** A distribution function $G$ is said to have a *semi-heavy tail* if its probability density function (p.d.f.) $g$ satisfies

$$g(x) \sim Cx^{-\rho}e^{-\gamma x} \quad \text{as } x \to \infty, \text{ for constant } C, \rho \in \mathbb{R} \textit{ and } \gamma > 0. \tag{3.1}$$

The proofs of semi-heavy tailed properties of certain distributions are very direct. The idea is just to write the density function into its equivalent forms. Then compare it with the representation form of semi-heavy tailed distribution to see whether that distribution complies with the definition. Consider the proof of the semi-heavy tail property by taking the NEF-GHS distribution as an example.

**Example 3.1.** *The proof of the semi-heavy tail property*
The full density of NEF-GHS($\lambda, \theta$) (without loss of generality, location and scale parameter are not introduced here) is

$$f(x; \lambda, \theta) = \frac{2^{\lambda-2}}{\pi\Gamma(\lambda)}|\Gamma(\frac{\lambda + ix}{2})|^2 \exp\{\theta x + \lambda \log(\cos(\theta))\},$$

where $\lambda > 0, |\theta| < \pi/2$. It can be shown that the NEF-GHS distribution reduces to GHS for $\theta = 0$ (bell-shaped with exponential decreasing tail) and reduces to HS for $\theta = 0$ and $\lambda = 1$. Furthermore, as $\beta = \tan\theta \to \infty$, the NEF-GHS distribution can be approximated by a Gamma distribution. Thus, for $\beta \in (0, \infty)$, the distribution is a compromise between the bell-shaped distribution and Gamma distribution. For the convenience of deriving the semi-heaviness of the distribution tail, we use the product form of the gamma function.

$$f(x; \lambda, \theta) = \frac{2^{\lambda-2}}{\Gamma(\lambda)} \prod_{j=0}^{\infty} \{1+x^2/(\lambda+2j)^2\}^{-1} \exp\{\theta x+\lambda \log(\cos(\theta))\}.$$

To write $f(x)$ into the form of (3.1), the only problem is to consider the product part, $g(x) = \prod_{j=0}^{\infty}\{1 + x^2/(\lambda+2j)^2\}^{-1}$. It is seen that

$$g'(x) = g(x) \cdot (-2x) \cdot \sum_{j=0}^{\infty} \frac{1}{(\lambda+2j)^2 + x^2}.$$

Because $\frac{1}{(\lambda+2j)^2+x^2} \leq \frac{1}{(\lambda+2j)^2}$, $\forall\ x$, and the series $\sum_{j=0}^{\infty} \frac{1}{(\lambda+2j)^2}$ is convergent, according to Weierstrass theorem, the series $\sum_{j=0}^{\infty} \frac{1}{(\lambda+2j)^2+x^2}$ is uniformly convergent. That is, $\forall \epsilon > 0, \exists N_0$ s.t. $\sum_{j=N_0+1}^{\infty} \frac{1}{(\lambda+2j)^2+x^2} \leqslant \epsilon$. Therefore, for all $x$ when $\epsilon$ is small enough (or $N_0$ is large enough.), the parts of the series after $N_0$ ($> 0$) has small influence to the value of infinite summation and can be ignored so that

$$\frac{g'(x)}{g(x)} \doteq -\sum_{j=0}^{N_0} \frac{2x}{(\lambda+2j)^2 + x^2}.$$

Then integrate both sides on the interval $[0, x]$,

$$\log[g(x)] \doteq -\sum_{j=0}^{N_0}\{\log[(\lambda+2j)^2 + x^2] + 2\log(\lambda+2j)\},$$

Thus,

$$g(x) \doteq \prod_{j=0}^{N_0} \frac{(\lambda+2j)^2}{(\lambda+2j)^2 + x^2},$$

which means $g(x) \sim x^{-2N_0}, x \to \infty$, where $N_0 > 0$ is a large number such that

$$f(x) \sim C_0 \cdot x^{-2N_0} \exp(\theta x),\ x \to \infty,$$

$C_0$ is constant. Finally, it is proved that the NEF-GHS distribution has approximately semi-heavy tail.

After the introduction of how to prove the semi-heavy tail property of a general distribution, we now begin to clarify the relationship between the semi-heavy tailed distribution class, the exponential class and the convolution equivalent distribution class by the following theorem.

**Theorem 3.1.** *If $G$ is a semi-heavy tail distribution with p.d.f. $g$ given in (3.1), then we have the following statements:*

*(a)* $G \in \mathfrak{L}_\gamma$. *(b) If* $\rho > 1$, $G \in \mathfrak{S}_\gamma$.

**Proof:** For $g(x)$ satisfying (3.1), integrating both sides then

$$\overline{G}(x) = \int_x^\infty f(t)\ \mathrm{d}t \sim \frac{C}{\gamma} x^{-\rho} e^{-\gamma x} \sim \frac{f(x)}{\gamma}, \text{ as } x \to \infty.$$

(a) It is easy to prove that

$$\lim_{x\to\infty} \frac{\overline{G}(x-y)}{\overline{G}(x)} = \lim_{x\to\infty} \frac{\frac{C}{\gamma}(x-y)^{-\rho} e^{-\gamma(x-y)}}{\frac{C}{\gamma} x^{-\rho} e^{-\gamma x}} = e^{\gamma y},$$

which means $G \in \mathfrak{L}_\gamma$.

(b) Using Lemma 2.3 in Pakes (2004), $\overline{G}(x) \sim \frac{C}{\gamma} x^{-\rho} e^{-\gamma x}$ satisfies the condition that $\overline{G}(x) = x^{-\delta} L(x) e^{-\gamma x - c x^\omega}$ where $\gamma,\ c \geq 0$, $\omega < 1$ and $L(x)$ is slowly varying. When $c = 0$, we have $\delta > 1$, then $G \in \mathfrak{S}_\gamma$. Here, $c = 0$ if $\rho > 1$, it shows that $\lim_{x\to\infty} \frac{\overline{G * G}(x)}{\overline{G}(x)}$ exists and finite which is equal to $2\int e^{\gamma x}\ dG(x)$. For semi-heavy tailed c.d.f. with $\rho \leq 1$, it has $\int_1^\infty e^{\gamma x}\ dG(x) = \infty$ and the limit for (2) cannot exist.
Hence, if $\rho > 1$, $G \in \mathfrak{S}_\gamma$. This completes the proof.

Corollary 2.9 of Albin & Sundén (2009) gives the representation form of the semi-heavy tail distribution in the *exponential class*. That is, the distribution $G$ has semi-heavy tails satisfy (3.1) if and only if the c.d.f. can be represented as

$$G(x) = 1 - \exp\{-\int_{-\infty}^{x} c(t)\ \mathrm{d}t\}\ for\ x \in \mathbb{R}, \tag{3.2}$$

for some measurable function $c \geq 0$ that satisfies $\lim_{t\to\infty} c(t) = \gamma$, and

$$\lim_{x\to\infty} \int_{-\infty}^{x} (c(t) - \frac{\rho}{x} I_{[1,\infty)}(t) - \gamma I_{[0,\infty)}(t))\ dt = \log(\frac{C}{\gamma}). \quad (3.3)$$

Thus (3.2) and (3.3) may contribute to the proof of the following theorem.

Adding one more semi-heavy tail condition into the distribution of log-return of a given asset in Theorem 2.3, we have the following theorem

**Theorem 3.2.** *If the distribution of return of a given asset G has a semi-heavy tail distribution with p.d.f. g given in (3.1), where $\rho > 1$, then the VaR of n-days ($VaR_p^n$) and the VaR of 1-days ($VaR_p^1$) are related by*

$$\rho \left[\log\left(\frac{VaR_p^n}{VaR_p^1}\right)\right] + \gamma(VaR_p^n - VaR_p^1) \sim \log n + (n-1)\log M, . \quad (3.4)$$

*as $p \to 1$, where $M = M_G(\gamma) = \int e^{\gamma x}\, dG(x)$*

**Proof:** Because $\overline{G}(x) = \int_x^\infty f(t)dt \sim \frac{C}{\gamma} x^{-\rho} e^{-\gamma x}$, like the proof of Theorem 4.2, as $p \to 1, VaR \to \infty$

$$1 - p = \overline{G}(VaR_p^1) = \frac{C}{\gamma}(VaR_p^1)^{-\rho} e^{-\gamma VaR_p^1}(1 + o(1)).$$

and

$$1 - p = \overline{G}(VaR_p^n) = nM^{n-1}\frac{C}{\gamma}(VaR_p^n)^{-\rho} e^{-\gamma VaR_p^n}(1 + o(1)).$$

Finally, from the above two fomulas, the relationship (3.4) holds. Note that (3.2) and (3.3) can be also used for the proof of this theorem.

From (3.4), given $\rho, \gamma$ and the VaR of 1-day ($VaR_p^1$), an asymptotic value of the VaR of $n$-days ($VaR_p^n$) can be obtained from numerical calculation. As $p \to 1$, it becomes just a root finding problem.

$$VaR_p^n \approx \text{solve}\{\rho \log(VaR_p^n) + \gamma VaR_p^n = \tilde{C}\},$$

where $\tilde{C} = \log n + (n-1)\log M + \rho \log(VaR_p^1) + \gamma VaR_p^1$.

**Example 3.2.** *The Variance-Gamma distribution is semi-heavy tailed but not convolution equivalent.*

The *Variance-Gamma* (VG) distribution (which is called the generalized asymmetric Laplace distribution by Kotz *et al.* (2001)) is one of the subclasses of the *generalized hyperbolic* (GH) distribution (Prause, 1999). The VG distribution is closed under convolution. (Among the GH distributions, only the VG distribution and the NIG distribution have this property.) Details about the GH distribution for the application of risk management are discussed in McNeil *et al.* (2005).

The density of the VG$(\theta, \nu, \mu, \sigma)$ distribution is

$$\frac{2\exp(\theta(x-\mu)/\sigma^2)}{\nu^{1/\nu}\sqrt{2\pi}\sigma\Gamma(1/\nu)}\left(\frac{|x-\mu|}{\sqrt{2\theta^2/\nu+\sigma^2}}\right)^{1/\nu-1/2} \times$$

$$K_{1/\nu-1/2}(\sigma^{-2}|x-\mu|\sqrt{2\sigma^2\nu+\theta^2})$$

where $\mu, \theta \in \mathbf{R}$ and $\nu, \sigma > 0$.

Let $\alpha = \sqrt{2\sigma^2\nu+\theta^2}/\sigma^2$ and $\beta = \theta/\sigma^2$. The tail decreases as

$$x^{1/\nu-1}e^{-(\alpha-\beta)x}, \text{ as } x \to +\infty.$$

Since $\nu > 0$ s.t. $1 - 1/\nu < 1$, the VG$(\theta, \nu, \mu, \sigma)$ belongs to $\mathfrak{L}_\gamma$ but not in the class $\mathfrak{S}_\gamma$. So the VG distribution does not have the scaling rule given in (3.4).

The rule in Theorem 3.2 is called the *semi-heavy tail rule* (SH rule). When the parameters $\rho$ and $\gamma$ are estimated from the financial data, people can easily employ the SH rule to obtain an accurate multi-period VaR by scaling the single-period VaR. Because conditions for the CE rule seem obviously looser, we can also use the estimated $\gamma$ to apply to the convolution equivalent rule (CE rule, in Chapter 2). The results by using the CE rule might be much more conservative and inaccurate comparing with the results obtained from the SH rule calculation. The CE rule can only be recommended as an external rule for regulators (conservative and robust values are needed for them to designate regulatory capital). However, the SH rule could be employed as an internal model which can be considered as an alternative to the SQRT rule for risk managers.

In the next section, based on the semi-heavy tail assumption, we will derive a semiparametric estimation for single-period VaR. Furthermore, given the estimator of the single-period VaR, we can combine the semi-heavy rule (SH rule) to achieve the goal of better multi-period risk management.

## 3.2 Semiparametric Estimation

Suppose the upper tail under consideration follows the exponentially tempered power law distribution (classical semi-heavy tail distribution). The tempered Pareto distribution is widely applicable in finance, physics and communication networks. It is also named with power-law distribution with exponential cut-off. That is, for samples greater than a positive value $x_0 > 0$, $X_1, X_2, \ldots, X_n$ come from the distribution with tail function

$$\overline{G}(x) = P(X_1 > x) = Cx^{-\rho}e^{-\gamma x}, \ x \geq x_0,$$

where $\rho$, $\gamma$ are unknown parameters and $C = x_0^{\rho}e^{\gamma x_0}$. Naturally the density is

$$g(x) = Cx^{-\rho-1}e^{-\gamma x}(\rho + \gamma x) \sim Cx^{-\rho}e^{-\gamma x}, \ \text{as } x \to \infty,$$

which means that this distribution is semi-heavy tail with parameters $\rho$ and $\gamma$.

Let the decreasing order statistics of $X_1, X_2, \ldots, X_n$ denoted by $X_{(1)} \geq X_{(2)} \geq \ldots \geq X_{(k)} \geq \ldots \geq X_{(n)}$. Similar to the approach of Hill estimator (1970) for power-law (Pareto) distribution, Meerschaert, Roy and Shao (2010) developed the tail estimates of this tempered Pareto distribution, using a conditional maximum likelihood approach based on the first $1 \leq k \leq \{i : \inf_i X_{(i)} > 0\}$ upper order statistics. In their paper, given $X_{(k+1)} \leq d_x < X_{(k)}$, the conditional MLEs $\hat{\rho}$ and $\hat{\gamma}$ satisfy the normal equations

$$\sum_{i=1}^{k}(\log d_x - \log X_{(i)}) + \sum_{i=1}^{k} \frac{1}{\hat{\rho} + \hat{\gamma} X_{(i)}} = 0,$$

$$\sum_{i=1}^{k}(d_x - X_{(i)}) + \sum_{i=1}^{k} \frac{X_{(i)}}{\hat{\rho} + \hat{\gamma} X_{(i)}} = 0,$$

where $\hat{C} = \frac{k}{n} d_x^{\hat{\rho}} e^{\hat{\gamma} d_x}$. They also proved that if the above system of normal equations have a solution with $\hat{\rho} > 0$ and $\hat{\gamma} > 0$, then it is the unique conditional MLE. They also established the existence, consistency, and asymptotic normality of the unconditional MLE (the parameter estimates based on the whole data set). This estimation approach shares many excellent properties.

To apply the semi-heavy tail scaling rule, we are interested in the solution with $\hat{\rho} > 1$ and $\hat{\gamma} \geq 0$. Define

$$T_1 := \sum_{i=1}^{k}(\log X_{(i)} - \log d_x), \; T_2 := \sum_{i=1}^{k}(X_{(i)} - d_x)$$

two positive quantities. Eliminate the significance of $\gamma$ firstly from the above two normal equations and focus on the estimate of $\rho$. For $n \geq 1$ and $1 \leq k \leq n$, the estimation with $\hat{\rho} > 1$ and

$\hat{\gamma} > 0$ is finding the root of the equation $\sum_{i=1}^{k} \frac{X_{(i)}}{kX_{(i)}+\hat{\rho}(T_2-T_1X_{(i)})} = 1$, where $\hat{\rho}$ belongs to the interval $[1, k/T_1]$, and $\hat{\gamma} = (k-\hat{\rho}T_1)/T_2$. After these steps, we can get the estimators $\hat{\rho}$ and $\hat{\gamma}$.

Let $k = n$ and $d_x = X_{(1)} = \hat{x_0}$. The unconditional MLE of parameters can be obtained from the following equations

$$\sum_{i=1}^{n} \frac{1}{\hat{\rho} + \hat{\gamma}X_i} = \sum_{i=1}^{n} \log \frac{X_i}{\hat{x_0}}, \tag{3.5}$$

$$\sum_{i=1}^{n} \frac{X_i}{\hat{\rho} + \hat{\gamma}X_i} = \sum_{i=1}^{n} (X_i - \hat{x_0}). \tag{3.6}$$

Also, $\hat{C} = \hat{x_0}^{\hat{\rho}} e^{\hat{\gamma}\hat{x_0}}$. Here are the inference of the unconditional MLE $(\hat{\rho},\ \hat{\gamma})$,

**Theorem 3.3.** *(1) As $n \to \infty$, the probability that the equations (3.5) and (3.6) have a unique solution $(\hat{\rho},\ \hat{\gamma})$, converges to 1. Also, the MLE $(\hat{\rho},\ \hat{\gamma})$ are consistent estimators of $(\rho, \gamma)$.*

*(2) $(\hat{\rho}, \hat{\gamma})$ are asymptotically joint Gaussian distribution with the asymptotic mean $(\rho, \gamma)$ and the asymptotic variance-covariance matrix $\frac{1}{n}\mathbf{M}^{-1}$ where*

$$\mathbf{M} = \begin{pmatrix} E((\rho+\gamma X_1)^{-2}) & E(X_1(\rho+\gamma X_1)^{-2}) \\ E(X_1(\rho+\gamma X_1)^{-2}) & E(X_1^2(\rho+\gamma X_1)^{-2}) \end{pmatrix}$$

This theorem forms the basis to conduct the statistical inference of the following semiparametric esimation of VaR.

The choice $k$ are of crucial importance. It should be noted that the estimation for parameters needs large data set. If people have enough confidence about the assumption of distribution form in the whole support, not only the tail part, setting $k$ equals to $n$ would be an acceptable choice.

Now consider the estimation of VaR. Just like the approach of Danielsson and Vries (1998;2000), for $x > X_{(k+1)}$,

$$\frac{\overline{G}(x)}{\overline{G}(X_{(k+1)})} = (\frac{x}{X_{(k+1)}})^{-\rho} \exp[-\gamma(x - X_{(k+1)})].$$

From the empirical distribution estimator $\hat{\overline{G}}(X_{(k+1)}) = k/n$ and estimators $\hat{\rho}, \hat{\gamma}$, it suggested the following estimation for the distribution

$$\hat{G}(x) = 1 - \frac{k}{n}(\frac{x}{X_{(k+1)}})^{-\hat{\rho}} \exp[-\hat{\gamma}(x - X_{(k+1)})].$$

To express $x$ in terms of $\hat{G}(x)$, we have to invert the above function. Fix $p = \hat{G}(VaR_p)$, the Value-at-Risk can be numerically solved as a root-finding procedure. That is, to solve the following equation

$$VaR_p \approx \text{solve}\{\hat{\rho} \log VaR_p + \hat{\gamma} VaR_p = \breve{C}\}$$

where $\breve{C} = \hat{\rho} \log X_{(k+1)} + \hat{\gamma} X_{(k+1)} + \log(k) - \log[n(1-p)]$. Theorem 3.3 can be employed for further statistical inference for the estimated semiparametric VaR.

Based on this numerically obtained single-period estimation $\widehat{VaR}_p^1$, also on the $\hat{\rho}, \hat{\gamma}$ that have already been obtained from semi-parametric estimation, we can continuously employ multi-period VaR scaling rules (2.6) and (3.4) if $\hat{\rho} > 1$ (this condition can be satisfied by controlling the process of the estimation). Replace $M_\gamma$ by its empirical estimation $\hat{M} = \frac{1}{n}\sum_{i=1}^{n} e^{\hat{\gamma} X_i}$ to obtain the scaling result for regulators and risk managers.

(A) Semiparametric CE rule (SP-CE rule) for external regulation.

$$VaR_p^n \sim \widehat{VaR_p^1} + \frac{1}{\hat{\gamma}}[\log n + (n-1)\log \hat{M}], \text{ as } p \to 1. \quad (3.7)$$

(B) Semiparametric SH rule (SP-SH rule) for internal supervision.

$$\hat{\rho}\left[\log\left(\frac{\widehat{VaR_p^n}}{\widehat{VaR_p^1}}\right)\right] + \hat{\gamma}(VaR_p^n - \widehat{VaR_p^1}) \sim \log n + (n-1)\log \hat{M}, \quad . \tag{3.8}$$

as $p \longrightarrow 1$

In employing specific full parametric model to fit financial data, we need a more accurate approximation approach for multi-period VaR forecasting. In the following section, we will introduce the normal inverse Gaussian (NIG) distribution modeling and the specific scaling rules. Its saddlepoint approximation to multi-period VaR will be dealt with in the next chapter.

## 3.3 Normal Inverse Gaussian Distribution

A prominent subset of the *generalized hyperbolic* (GH) distribution (Prause, 1999) is the *normal inverse gaussian* (NIG) distribution, which have four parameters: $\alpha$ represents the tail heaviness, $\beta$ represents the asymmetry parameter, $\mu$ represents the location parameter and $\delta$ represents the scale parameter. The NIG distribution are considered as a benign distribution that fitted the financial data. We choose the NIG distribution here, because it is a full parametric distribution with totally semi-heavy tail and belongs to the convolution equivalent class (no need of parameter constraints).

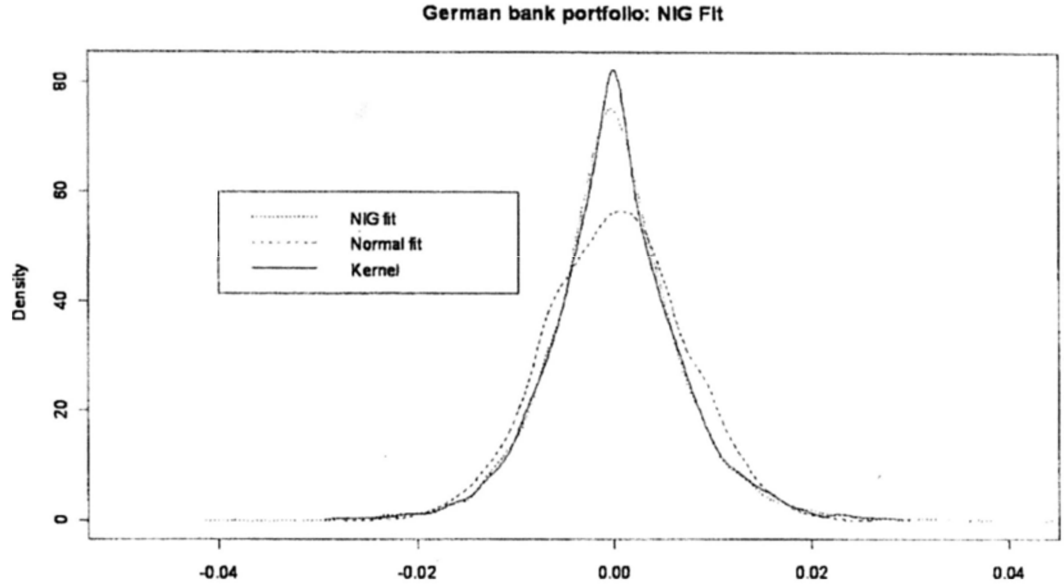


Fig. 3.2: German Bank Portfolio data: The NIG fit v.s. Normal fit.

The empirical experience suggest an excellent fit of the NIG distribution to the financial data. Fig. 3.2 displays the NIG fit with the German Bank Portfolio data. Comparing with Normal (Gaussian) fit, it is found the NIG is more feasible for describing real financial data than the Gaussian distributions.

The NIG distribution is able to model symmetric and asymmetric distributions with possibly long tails in both directions. Its tail behavior is often classified as semi-heavy, i.e. the tails are lighter than those of non Gaussian power laws, but much heavier than Gaussian. When $\mu = 0, \delta = 1$, Barndorff-Nielsen (1997) introduced and derived the semi-heavy tail property of the standard NIG distribution.

$$f_{NIG}(x) = \frac{\alpha}{\pi} \exp(\sqrt{\alpha^2 - \beta^2} + \beta x) \frac{1}{\sqrt{(1 + x^2)}} K_1(\alpha \sqrt{1 + x^2}),$$

where $K_1$ is the modified Bessel function of the second kind with index 1.

The Bessel functions $K_\lambda, \lambda \in R$,

$$K_\lambda(\omega) = \frac{1}{2}\int_0^\infty x^{\lambda-1} e^{\frac{1}{2}\omega(x+x^{-1})} \mathrm{d}x, \qquad \omega > 0.$$

satisfies the relations

$$\begin{aligned}
K_{-\lambda}(\omega) &= K_\lambda(\omega), \\
K_{\lambda+1}(\omega) &= 2\lambda/\omega K_\lambda(\omega) + K_{\lambda-1}(\omega), \\
K_\lambda^{'}(\omega) &= -\frac{1}{2}(K_{\lambda+1}(\omega) + K_{\lambda-1}(\omega)), \\
K_{\pm\frac{1}{2}}(t) &= \sqrt{\frac{\pi}{2}} t^{-\frac{1}{2}} e^{-t}.
\end{aligned}$$

Using $K_\gamma(x) \sim \sqrt{\frac{\pi}{2}} x^{-\frac{1}{2}} e^{-x}$ as $x \to \infty$,

$$f_{NIG}(x) \sim \frac{\alpha}{\pi} \exp(\sqrt{\alpha^2 - \beta^2} + \beta x) \frac{1}{\sqrt{(1+x^2)}} \sqrt{\frac{\pi}{2}} x^{-\frac{1}{2}} e^{-x}, \quad \text{as } x \to \infty.$$

Finally, we obtain that $f$ is represented as

$$f_{NIG}(x) \sim Constant\ x^{-\frac{3}{2}} \exp\{-(\alpha - \beta)x\}, \quad \text{as } x \to \infty.$$

For $\rho = -\frac{3}{2} < -1$, the NIG distribution has the semi-heavy tail property and belongs to the class $\mathfrak{S}_{\alpha-\beta}$. It satisfies the condition of Theorem 2.3 and Theorem 3.2, so that we can employ the respective scaling rules (CE rule and SH rule).

**NIG Scaling Rules**

If the log-return of an asset has a NIG distribution, we can use Theorem 3.1 to calculate its multi-period VaR from scaling the single-period VaR. The moment generating function (m.g.f.) of the NIG$(\alpha, \beta, 1, 0)$ is

$$M(u) = \exp[\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + u)^2}].$$

The specific convolution equivalent scaling rule (CE rule) for the standard NIG distribution ($\mu = 0, \delta = 1$) then becomes

$$VaR_p^n \sim VaR_p^1 + \frac{1}{\alpha - \beta}[\log n + (n-1)\sqrt{\alpha^2 - \beta^2}], \text{ as } p \to 1, \tag{3.9}$$

where $\alpha$ and $\beta$ are the parameters of NIG distribution.

Let $\bar{\alpha} = \delta\alpha$ and $\bar{\beta} = \delta\beta$, which are called invariant parameters. This feature of parametrization NIG$(\bar{\alpha}, \bar{\beta}, \mu, \sigma)$ has density

$$f(x) = \frac{\bar{\alpha}}{\pi\delta}\exp(\sqrt{\bar{\alpha}^2 - \bar{\beta}^2} + \bar{\beta}\frac{x-\mu}{\delta})\frac{K_1(\bar{\alpha}\sqrt{1 + (\frac{x-\mu}{\delta})^2})}{\sqrt{(1 + (\frac{x-\mu}{\delta})^2)}}.$$

It is a location-scale family. Through the location-scale transformation, we have

$$X \sim \text{NIG}(\bar{\alpha}, \bar{\beta}, \mu, \sigma) \Leftrightarrow \frac{X - \mu}{\delta} \sim \text{NIG}(\bar{\alpha}, \bar{\beta}, 0, 1).$$

So after the location-scale transformation, the scaling rule for the NIG distribution with NIG$(\alpha, \beta, \delta, \mu)$ is finally simplified as

$$VaR_p^n \sim VaR_p^1 + \frac{\log n + (n-1)[\delta\sqrt{\alpha^2 - \beta^2} + \mu(\alpha - \beta)]}{\alpha - \beta}, \tag{3.10}$$

as $p \to 1$. where $\mu$, $\delta$, $\alpha$ and $\beta$ are the parameters of the NIG distribution. This can be denoted by the *NIG-CE rule* for external regulation.

By using scaling method of (3.4), from (3.10), the less conservative and more accurate semi-heavy scaling rule can be obtained by

$$\frac{3}{2}\log VaR_p^n + (\alpha - \beta)VaR_p^n = \check{C}, \text{ as } p \to 1, \tag{3.11}$$

where

$$\breve{C} = \frac{3}{2}\log VaR_p^1 + (\alpha-\beta)VaR_p^1 + \log n + (n-1)[\delta\sqrt{\alpha^2-\beta^2} + \mu(\alpha-\beta)].$$

Similarly to the semiparametric SH rule (SP-SH rule), this rule can be denoted by the *NIG-SH rule* of internal models for risk managers.

**Estimations of the NIG distribution**

As is konwn, the m.g.f. of the NIG$(\alpha, \beta, \delta, \mu)$ is as

$$M(u) = \exp\{\delta[\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+u)^2}]\}.$$

The mean of the NIG-distribution is $\mu + (\delta\beta)/\gamma$ and the variance is $\delta\alpha^2/\gamma^3$. The skewness is $3\delta\alpha^2\beta\gamma^{-5}$ and the kurtosis is $3\delta\alpha^2(\alpha^2+4\beta^2)\gamma^{-7}$. These statistical quantities can be used for the estimation of parameters, which is called the Method of Moment (MoM) estimation approach. But a more efficient estimator is the Maximum Likelihood estimation (MLE). The log-likelihood function for the observations $\tilde{x} = (x_1, \ldots, x_n)$ which follows the $NIG(\alpha, \beta, \delta, \mu)$ is constructed as

$$L(\tilde{x}|\alpha, \beta, \delta, \mu) = \sum_{i=1}^{n} \log f_{NIG}(x_i|\alpha, \beta, \delta, \mu).$$

After fitting the NIG$(\alpha, \beta, \delta, \mu)$ with real financial data, people can get all estimations of parameters denoted by $\hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu}$. Then given the single-period VaR, we can explicitly employ the specific asymptotic scaling approaches like (3.10), (3.11) and saddlepoint approximation approach proposed in the next chapter for multi-period VaR calculation. Under the model with the NIG distributional assumption, the single-period VaR at level $p$ is the $p$-quantile of NIG$(\hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu})$. We can get the multi-period VaR adopting NIG-CE rule for regulators and NIG-SH rule for internal managers.

In the following chapter, we employ the saddle-point approximations technique for constructing a fast and accurate analytical approximation to the tail of the distribution of asset returns. We show how to obtain an accurate VaR without recoursing to Monte Carlo simulations.

---

□ **End of chapter.**

# Chapter 4

# Saddlepoint Approximation to VaR

**Summary**

Saddlepoint approximation is gradually emerging as a powerful tool for obtaining an accurate expression for density and distribution function. This chapter reviews the approach of saddlepoint approximation. Then under the specific distributional assumption - NIG distribution, through approximating its tail probability, we numerically conduct a multi-period VaR claculation. This direct approximation approach of multi-period risk provides banking supervisors a fast and accurate way to quantify market risk.

Saddlepoint approximation is known to give excellent approximation to the p.d.f. and the tail probability of a distribution. In financial risk management, sallepoint approximation have been used in the modeling of credit risk (Martin *et al.*, 2001; Huang *et al.*, 2007). For the computation of VaR, Feuerverger and Wong (2000) devised a saddlepoint approximation approach to calculate VaR under multivariate normal or non-normal portfolios with large complex risk factos. Broda and Paolella (2009) proposed a model for the calculation of portfolio VaR based on GO-GARCH model, independent component analysis (ICA) and generalized hyperbolic (GH) distributional assumption. Tian and Chan (2010) applied saddlepoint approximation technique to the conditional heteroscedastic model combined with quasi-residual modeling of volatility. In this chapter, we first give a review on this topic. For further details, please refer to the books Jensen (1994) and Butler (2007). Then based on the preliminary theory, we derived the multi-period estimation of VaR under the NIG distributional assumption.

## 4.1 Theoretical Preliminaries

The saddlepoint approximation approach, introduced by Daniel (1954; 1987), is an accurate and fast device especially for tail probability and density estimation. It is not surprising to see that during the past two decades, research in the area has vastly increased. The detailed theoretical review on saddlepoint approximation can be referred to the book written by Jensen (1994). For the application of the saddlepoint approximation techniques, see Guotis and Casella (1999), Huzurbazar (2006) and Butler (2007).

Recently, the saddlepoint approximation has been a powerful tool in higher-order asymptotic approximation and therefore it has wide applications in many scientific areas. Derivations and implementations of it rely on tools such as exponential tilting, Edgeworth expansions, Hermite polynomials, complex integration and other advanced notions.

- **From Taylor to Laplace**

Perhaps the simplest way to approximate a positive function $f(x)$ is to use the first few terms of its Taylor series expansion. We will use that idea, not for $f(x)$ itself but for $h(x) \equiv \log f(x)$. Then

$$f(x) \approx \exp\left\{h(x_0) + (x - x_0)h'(x_0) + \frac{(x - x_0)^2}{2!}h''(x_0)\right\}.$$

The above approximation simplifies if we choose $x_0 = \hat{x}$, where $h'(\hat{x}) = 0$. We have

$$f(x) \approx \exp\left\{h(\hat{x}) + \frac{(x - \hat{x})^2}{2!}h''(\hat{x})\right\}.$$

Hence, we have

$$\begin{aligned} \int f(x)\mathrm{d}x &\approx \int \exp\left\{h(\hat{x}) + \frac{(x - \hat{x})^2}{2!}h''(\hat{x})\right\}\mathrm{d}x \\ &\approx \exp\{h(\hat{x})\}\left(-\frac{2\pi}{h''(\hat{x})}\right)^{1/2}. \end{aligned}$$

The above approach is called the *Laplace approximation.*

The next, perhaps the most natural step is to consider the extended case. Rewrite the function $f$ as

$$f(x) = \int m(x, t)\,\mathrm{d}t$$

for some positive $m(x, t)$. This is always possible. By defining $\kappa(x, t) = \log m(x, t)$, we consider the Laplace approximation of the integral of $\exp \kappa(x, t)$ with respect to the variable $t$. For any fixed $x$, we have

$$\begin{aligned} f(x) &\approx \int \exp\left\{\kappa(x, \hat{t}(x)) + \frac{(t - \hat{t}(x))^2}{2!}\frac{\partial^2\kappa(x, t)}{\partial t^2}\Big|_{\hat{t}(x)}\right\}\mathrm{d}t \\ &= \exp\{\kappa(x, \hat{t}(x))\}\left(-\frac{2\pi}{\frac{\partial^2\kappa(x,t)}{\partial t^2}\big|_{\hat{t}(x)}}\right)^{1/2}, \end{aligned}$$

where, for each $x$, $\hat{t}(x)$ satisfies $\partial\kappa(x, t)/\partial t = 0$ and $\frac{\partial^2\kappa(x,t)}{\partial t^2} < 0$, and hence maximizes $\kappa(x, t)$.

- **Derivation of saddlepoint approximation by using the inverse Fourier transform**

For a probability density function (p.d.f.) $f(x)$, its *moment generating function (m.g.f)* is defined as

$$M_X(t) = \mathbb{E}\exp(t \cdot X) = \int_{-\infty}^{+\infty} \exp(tx)f(x)\, \mathrm{d}x,$$

By using the inversion formula, we have

$$\begin{aligned} f(x) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} M_X(it)\exp(-itx)\, \mathrm{d}t \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \exp\{K_X(it) - itx\}\mathrm{d}t, \end{aligned}$$

where $i = \sqrt{-1}$, $K_X(t) = \log M_X(t)$ is called the *the cumulant generating function* (c.g.f.) and $M_X(it)$ is called the *characteristic function* (c.f.).

Make a change of variable $t' = it$, then for $\tau$ in a neighborhood of zero,

$$\begin{aligned} f(x) &= \frac{1}{2\pi i}\int_{\tau-\infty}^{\tau+\infty} \exp\{K_X(t) - tx\}dt \\ &\approx \frac{1}{2\pi i}\int_{\tau-\infty}^{\tau+\infty} \exp\{K_X(\hat{t}(x)) - \hat{t}(x)x + \frac{(t-\hat{t}(x))^2}{2}K''_X(\hat{t}(x))\}dt \\ &\approx \left(\frac{1}{2\pi K''_X(\hat{t}(x))}\right)^{1/2} \exp\{K_X(\hat{t}(x)) - \hat{t}(x)x\}. \end{aligned}$$

where $\hat{t}(x)$ satisfies $K'_X(t) = x$. Viewed as a point in the complex plane, $\hat{t}(x)$ is neither a maximum nor a minimum but a *saddlepoint* of $K_X(t) - tx$.

An approximation to the cumulative function (c.d.f.) of $X$ is equally straightforward,

$$F_X(x) = \Phi(r) + \phi(r)\left\{\frac{1}{r} - \frac{1}{q}\right\},$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal distribution and density function, respectively. $r = \text{sgn}(\hat{t}(x))\left[2\{\hat{t}(x)x - K_X(\hat{t}(x))\}\right]^{\frac{1}{2}}$ and $q = \hat{t}(x)\{K''_X(\hat{t}(x))\}^{\frac{1}{2}}$, see Lugannani and Rice (1980).

For lattice variables, $\hat{t}(x)$ is the solution to $K'(t) = x - \Delta/2$, where $\Delta$ is a constant by which the possible values of components of $X$ are separated. Without loss of generality, $\Delta = 1$ is employed here. And $q = 2\sinh(\hat{t}(x)\{K''(\hat{t}(x))\}^{1/2})$

The tail probability is approximated as a contour integral involving the moment generating function: (let $\widehat{t}(x) = s_x$)

$$
\begin{aligned}
& P(X > x) \\
= \; & \frac{1}{2\pi i}\int_{-i\infty,(0+)}^{+i\infty} \frac{\exp(K_X(t)) - tx}{t} dt \\
\approx \; & \begin{cases} \exp\left(a(x)\right)\Phi\left(-b(x)\right), & x > \mathbb{E}(X), \\ 1/2, & x = \mathbb{E}(X), \\ 1 - \exp\left(a(x)\right)\Phi\left(-b(x)\right), & x < \mathbb{E}(X), \end{cases}
\end{aligned}
$$

where $a(x) = K(s_x) - s_x x + \frac{1}{2}s_x^2 K''_X(s_x)$ and $b(x) = \sqrt{s_x^2 K''(s_x)}$.

- **Derivation of saddlepoint approximation by using the Edgeworth Expansions**

An alternative derivation of saddlepoint approximation is based on Edgeworth expansions. It provides a more precise picture of how the greater accuracy can be achieved.

Let $X_1, \ldots, X_n$ be iid with density $f$ with mean $\mu$ and finite variance $\sigma^2$. Let $\overline{X} = (X_1 + \cdots + X_n)/n$. A useful form of an Edgeworth expansion can be written

$$P\{n^{1/2}(\overline{X} - \mu)/\sigma \le w\} = \Phi(w) + \phi(w)\left[\frac{-1}{6\sqrt{n}}\kappa(w^2 - 1) + O\left(\frac{1}{n}\right)\right]$$

where $\Phi$ and $\phi$ are the distribution and density function of a standard normal and $\kappa = E(X_1 - \mu)^3$ is the skewness.

Make the transformation $\overline{x} = \sigma w + \mu$, we obtain the approximation to the density of $\overline{X}$ as

$$f_{\overline{X}}(\overline{x}) = \frac{\sqrt{n}}{\sigma}\phi\left(\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}\right) \times \left[1 + \frac{\kappa}{6\sqrt{n}}\left\{\left(\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}\right)^3 - 3\left(\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}\right)\right\} + O\left(\frac{1}{n}\right)\right].$$

Ignoring the terms in braces produces the usual normal approximation, which is accurate to $O(\frac{1}{\sqrt{n}})$. If we are using values of $\overline{x}$ near $\mu$, then the value of the above expression in braces is close to zero, and the approximation will then be accurate to $O(\frac{1}{n})$. The trick of saddlepoint approximation is to make this always the case.

- **Exponential tilting**

If $X_1, \ldots, X_n$ are iid from $f(x|\tau) = exp\{\tau x - K(\tau)\}f(x)$, then

$$\begin{aligned} & f_{\overline{X}}(\overline{x}|\tau) \\ = {} & \exp\{-n[\tau\overline{x} - K(\tau)]\}f_{\overline{X}}(\overline{x}), \\ = {} & \exp\{-n[\tau\overline{x} - K(\tau)]\}\frac{\sqrt{n}}{\sigma_\tau}\phi\left(\frac{\overline{x}-\mu_\tau}{\sigma_\tau/\sqrt{n}}\right) \times \Bigg[1 + \frac{\kappa}{6\sqrt{n}} \\ & \left\{\left(\frac{\overline{x}-\mu_\tau}{\sigma_\tau/\sqrt{n}}\right)^3 - 3\left(\frac{\overline{x}-\mu_\tau}{\sigma_\tau/\sqrt{n}}\right)\right\} + O\left(\frac{1}{n}\right)\Bigg]. \end{aligned}$$

Given $\overline{x}$, we choose $\tau$ so that $\mu_\tau = \overline{x}$, which is equivalent to choosing $\tau$ so that $K'(\tau) = \overline{x}$, the familiar saddlepoint equation. Denote this value by $\widehat{\tau}$, we get the approximation

$$f_{\overline{X}}(\overline{x}) = \exp\{-n[\widehat{\tau}\overline{x} - K(\widehat{\tau})]\}\frac{\sqrt{n}}{\sigma_\tau}\phi(0) \times \left[1 + O\left(\frac{1}{n}\right)\right],$$

where $\mu_\tau, \widehat{\sigma}^2$ and $\kappa_\tau$ are the mean, variance, and skewness of $f(\cdot|\tau)$, respectively. Note that $\sigma_\tau^2 = K''(\tau)$, we have

$$f_{\overline{X}}(\overline{x}) = \left(\frac{n}{2\pi K''_X(\widehat{\tau})}\right)^{1/2} \exp\{-n[\widehat{\tau}\overline{x} - K(\widehat{\tau})]\} \times \left[1 + O\left(\frac{1}{n}\right)\right].$$

**Example 4.1.** *Example: saddlepoint approximation to NEF-GHS distribution*

Assume a series of returns of a given asset follow the NEF-GHS distribution. The i.i.d. return is denoted by $X_i$ then $X_i \overset{iid}{\sim}$ NEF-GHS$(\mu, \delta, \theta, \lambda), i = 1, 2, \ldots$. The NEF-GHS has the excellent property of closure under convolution. Then distribution for $X^{(n)} = \sum_{i=1}^{n} X_i$ is NEF-GHS$(\mu, \delta, \theta, n\lambda)$, the multi-period VaR can be easyly derived as follows:

The *moment generating function* (m.g.f.) of $X^{(n)}$ is

$$M(u) = \exp\{\mu u - n\lambda \log[\cos(\delta u) - \beta \sin(\delta u)]\},$$

where $\beta \equiv \tan\theta$. The *cumulant generating function* (c.g.f.) is

$$K(u) = \log(M(u)) = \mu u - n\lambda \log[\cos(\delta u) - \beta \sin(\delta u)].$$

A saddlepoint $s$ is obtained by solving $K'(\hat{t}) = x$, thus

$$s = \hat{t} = \frac{1}{\delta} \arctan\left[\frac{x - \mu - n\lambda\delta\beta}{n\lambda\delta + (x-\mu)\beta}\right]. \tag{4.1}$$

Evaluate the $p$-th quantile $q_p$ of the distribution by using Lugannani & Rice formula (1980) as follows

$$p = P(X^{(n)} > t) = \begin{cases} \exp\{a(t)\}\,\Phi(-), & t > \mathbb{E}(X^{(n)}), \\ 1/2, & t = \mathbb{E}(X^{(n)}), \\ 1 - \exp\{a(t)\}\,\Phi\left(-\sqrt{\hat{t}^2 K''(\hat{t})}\right), & t < \mathbb{E}(X^{(n)}), \end{cases} \tag{4.2}$$

where $a(t) = K(\hat{t}) - \hat{t}t + \frac{1}{2}\hat{t}^2 K''(\hat{t})$, $b(t) = \sqrt{\hat{t}^2 K''(\hat{t})}$and $\Phi(\cdot)$

denotes the c.d.f. of normal (Gaussian) distribution. In (4.2),

$$\begin{aligned}
\mathbb{E}(X^{(n)}) &= \mu + \lambda\delta\beta, \\
K(u) &= \mu u - n\lambda \log[\cos(\delta u) - \beta \sin(\delta u)], \\
K'(u) &= \mu + \frac{n\lambda\delta(\tan(\delta u) + \beta)}{1 - \beta \tan(\delta u)}, \\
K''(u) &= n\lambda\delta^2 \left[1 + \frac{(\tan(\delta u) + \beta)^2}{(1 - \beta \tan(\delta u))^2}\right].
\end{aligned}$$

Equation (4.1) can be updated into (4.2) and we can get the tail probability. From the tail probability, the quantile can be numerically calculated. A similar procedure will be discussed in the next section, under the specific assumption of the NIG distribution.

It seems that the representation of the saddlepoint approximation is complex. Actually, the saddlepoint approximation method can be very easily calculated by a desk calculator. Wang (1995) derived two simple methods to calculate the one-step quantile approximation via saddlepoint techniques. They can also be used to obtain approximate quantile estimate. Here we take the first approach in that paper to calculate quantile for comparison. Also, take the NEF-GHS distribution as an example. In addition to $K'(\cdot)$, $K''(\cdot)$, and $K'''(\cdot)$, it should be noted that other quantities are needed.

$$K'''(u) = \frac{2n\lambda\delta^3 [\tan(\delta u) + \beta]}{1 - \beta \tan(\delta u)} \left[1 + \frac{(\tan(\delta u) + \beta)^2}{(1 - \beta \tan(\delta u))^2}\right],$$

such that

$$\begin{aligned}
K(\hat{t}) =& \frac{\mu}{\delta} \arctan\left[\frac{x - \mu - n\lambda\delta\beta}{n\lambda\delta + (x - \mu)\beta}\right] - n\lambda \\
& \log\left[sgn(n\lambda\delta + (x - \mu)\beta) \frac{n\lambda\delta(1 + \beta^2)}{\sqrt{(\beta^2 + 1)\left[(x - \mu)^2 + n^2\lambda^2\delta^2\right]}}\right],
\end{aligned}$$

$$K''(\hat{t}) = n\lambda\delta^2 + (x - \mu)\delta.$$

Define

$$r = \omega - \frac{1}{\omega}\log(\frac{\omega}{z}),$$

where

$$\omega \equiv sgn(\hat{t})\sqrt{2[\hat{t}x - K(\hat{t})]},$$

$$z \equiv \hat{t}\sqrt{K''(\hat{t})}.$$

Because $\hat{t}$, $r$ are both functions of $x$, when the parameters $\lambda$ and $\beta$ are known or estimated, $\hat{t}$, $r$ are denoted by $\hat{t}(x)$, $r(x)$ respectively. Then

$$q_0 = \sqrt{n}\Phi^{-1}(\alpha)\sigma + \frac{[\Phi^{-1}(\alpha)]^2 - r(n\Phi^{-1}(\alpha)\sigma)^2}{2\hat{t}(n\Phi^{-1}(\alpha)\sigma)},$$

$$\Delta_q = \frac{[\Phi^{-1}(\alpha)]^2 - r(q_0)^2}{2\hat{t}(q_0)},$$

where $\sigma = \lambda(1 + \beta^2)$.

The one-step $\alpha$-quantile can be approximated by $x_\alpha \approx q_1 = q_0 + \Delta q_0$. If a more accurate approximation is needed, more iterations can be implemented by using the above $\Delta$ with updated $q_i$. This approach is a heuristic approximation illustrating the convenience of saddlepoint-based method.

**Example 4.2.** *Accuracy of the saddlepoint approximations*

In this example, we display in Figure 4.1 the loss exceedance curve (VaR against tail probability) for the given asset which is assumed to follow a GH distribution with $\lambda = 1, \alpha = 1, \beta = 0.3, \delta = 1, \mu = 0$. ($\lambda = 1$, it is a special subclass - hyperbolic (HYP) distribution.) In this example, the true return distribution (HYP) are compared with the saddlepoint approximations and normal approximations.

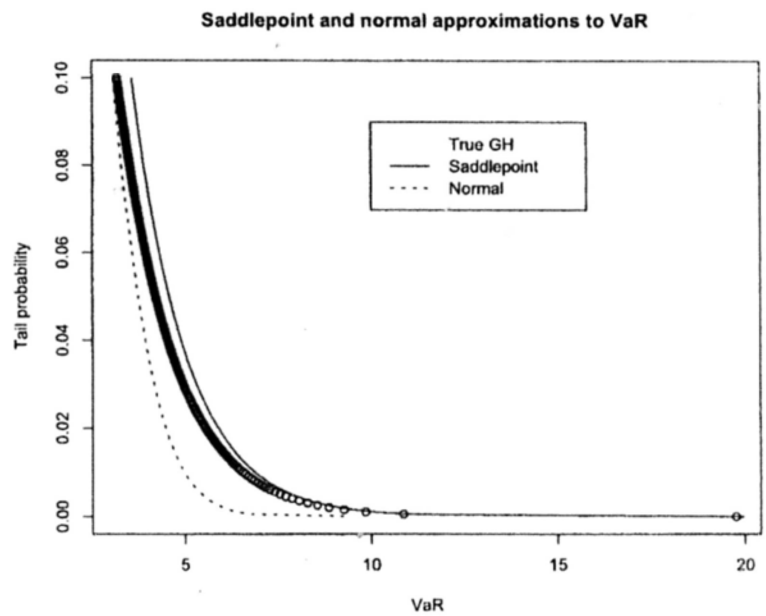


Fig. 4.1: Accuracy of the saddlepoint approximations.

*Fig. 4.1 The dotted line is the true hyperbolic distribution with the parameters* $\lambda = 1, \alpha = 1, \beta = 0.3, \delta = 1, \mu = 0$. *The solid line represents the saddlepoint approximations with the sample size* $n = 1$. *and dashed line corresponds to the normal approximations with the simple size* $n = 200$.

## 4.2 NIG Multi-period VaR Approximation

Several techniques proposed in the literature for calculating VaR usually rely on the use of Monte Carlo simulation, which has ponderous burden in calculation and hence takes a long time. The aim of this section is to perform statistical calculation of VaR based on saddlepoint approximations method that gives an analytical technique for a rapid and accurate construction of the return distribution. For single-period VaR approximation, the approach proposed by Tian and Chan (2008) considered the conditional heteroscedastic model, which is appropriate when the return distribution of the asset is assumed to be normal inverse Gaussian (NIG) distribution.

Also, when we want to calculate the multi-period VaR, saddlepoint approximation technique can be used. Assume the distribution of the return of a given asset is the NIG distribution, which has $X_i \overset{iid}{\sim} \text{NIG}(\alpha, \beta, \delta, \mu), i = 1, 2, \ldots$. The NIG distribution has benign property of convolution closure. Then distribution for $X^{(n)} = \sum_{i=1}^{n} X_i$ is NIG$(\alpha, \beta, n\delta, n\mu)$, the multi-period VaR can be easily derived by following the procedures:

The *moment generating function* (m.g.f.) of $X^{(n)}$ is

$$M(u) = \exp\{n\mu u + n\delta[\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + u)^2}]\},$$

The *cumulant generating function* (c.g.f.) is

$$K(u) = \log(M(u)) = n\mu u + n\delta[\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + u)^2}].$$

A saddlepoint $s$ is obtained by solving $K'(\hat{t}) = x$, thus we evaluate the $p$-th quantile $q_p$ of the distribution as follows

$$p = P(X^{(n)} > t) = \begin{cases} \exp\{a(t)\} \Phi(-b(t)), & t > \mathbb{E}(X^{(n)}), \\ 1/2, & t = \mathbb{E}(X^{(n)}), \\ 1 - \exp\{a(t)\} \Phi(-b(t)), & t < \mathbb{E}(X^{(n)}), \end{cases} \tag{4.3}$$

where $a(t) = K(\hat{t}) - \hat{t}t + \frac{1}{2}\hat{t}^2 K''(\hat{t})$, $b(t) = \sqrt{\hat{t}^2 K''(\hat{t})}$ and $\Phi(\cdot)$

denotes the c.d.f. of normal distribution. In (4.3),

$$
\begin{aligned}
\mathbb{E}(X^{(n)}) &= \mu + \frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}} \cdot \frac{K_{3/2}(\delta\sqrt{\alpha^2-\beta^2})}{K_{1/2}(\delta\sqrt{\alpha^2-\beta^2})}, \\
K(u) &= n\mu u + n\delta[\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+u)^2}], \\
K'(u) &= \mu + \frac{\delta(\beta+u)}{\sqrt{\alpha^2-(\beta+u)^2}} \cdot \frac{K_{3/2}(\delta\sqrt{\alpha^2-(\beta+u)^2})}{K_{1/2}(\delta\sqrt{\alpha^2-(\beta+u)^2})}, \\
K''(u) &= \frac{\delta}{\sqrt{\alpha^2-(\beta+u)^2}} \frac{K_{3/2}(\delta\sqrt{\alpha^2-(\beta+u)^2})}{K_{1/2}(\delta\sqrt{\alpha^2-(\beta+u)^2})} + \frac{\delta^2(\beta+u)^2}{\alpha^2-(\beta+u)^2}
\end{aligned}
$$

$$
\times \frac{K_{5/2}(\delta\sqrt{\alpha^2-(\beta+u)^2})}{K_{1/2}(\delta\sqrt{\alpha^2-(\beta+u)^2})} - \frac{\delta^2(\beta+u)^2}{\alpha^2-(\beta+u)^2} \frac{K_{3/2}^2(\delta\sqrt{\alpha^2-(\beta+u)^2})}{K_{1/2}^2(\delta\sqrt{\alpha^2-(\beta+u)^2})}.
$$

Once the above quantities are calculated, it is straightforward to calculate the VaR. Use the value at saddlepoint to replace $\hat{t}$ of (4.3), then adjust $t$ until the right-hand side of (4.3) equals to a given probability $p$. This step is a simple root-finding problem. Saddlepoint approximation to multi-period VaR is a direct method without scaling from single-period VaR.

Above all, we have completed the saddlepoint approximation to the multi-period VaR under the distributional assumption - the NIG distribution. In the next chapter, we conduct some data analysis to evaluate outcomes from using the scaling rules in the above two chapters and saddlepoint approximation in this chapter.

---

□ **End of chapter.**

# Chapter 5

# Data Analysis

**Summary**

We make use of the approaches proposed in the preceding chapters to data analysis. Simulation study is aimed to evaluate the CE rule, the semparametric estimation (SP) with SH rule and the saddlepoint approximation to specific distributions - the NIG distribution. Two real data sets are analyzed. Through empirical research, we explore the feasibility of these scaling methods with calculations and approximations.

In this chapter, we conduct a simple simulation study to evaluate the convolution equivalent scaling rule (CE rule), the semiparametric estimation with the semi-heavy rule (SP-SH rule) and the saddlepoint approximation to VaR. Data are simulated from the NIG distribution, which is appropriate for financial data (Venter and Jongh, 2002). Although there are many drawbacks of the SQRT rule, it is still better than many other rules. It is better than the $\alpha$-root rule (Danielsson & De Vries, 1998) because of its acceptable prudence. And it is better than Drost-Nijman rule (1993) because of its simplicity. In the section of real data example, internal rules (SP-SH rule and NIG-SH rule) is compared to the SQRT rule. The performance of the saddle-point approximation technique in the multi-period risk management is also studied.

## 5.1 Simulation

In this section, we evaluate the performance of the convolution equivalent rule (CE rule), the systematic semiparametric methodologies (SP-SH rule) and the saddle-point approximations by using Monte Carlo simulations.

**(A)** *Rule under convolution equivalent tail (CE rule).*

Data are simulated from the NIG(1,0.3,1,0) distribution and the holding period is 10 days. Fig.5.1 shows the conservative property of the convolution equivalent scaling rule comparing with the normal approximated multi-period VaR (scaling via the SQRT rule). The multi-period VaR estimated from scaling the single-period VaR by using convolution equivalent rule are bigger than the true VaR with a constant level, which can be a desirable property for external regulation. The graph of calculated VaR by using convolution equivalent scaling rule can provide reference and justification for external risk regulators to implement stricter supervision.

**(B)** *Semiparametric estimation(SP) + Rule under semi-heavy tail (SH rule).*

Simulate data from the NIG(1,-0.04,1,0) distribution to evaluate the semiparametric estimation. The multi-period holds for 10 days. The sample size for simulated data is 10,000. The theoretical value for $\alpha, \beta$ should be around the values 1.5 and 1.04 while the semiparametric estimation $\hat{\alpha}, \hat{\beta}$ are 1.6609 and 0.9142 respectively. Detailed analysis and explanation of estimation for parameters can be referred to Meerschaert *et al* (2010). The confidence level of VaR is 99%. True theoretical VaR for single-period is 2.5836 while the estimated single-period VaR is

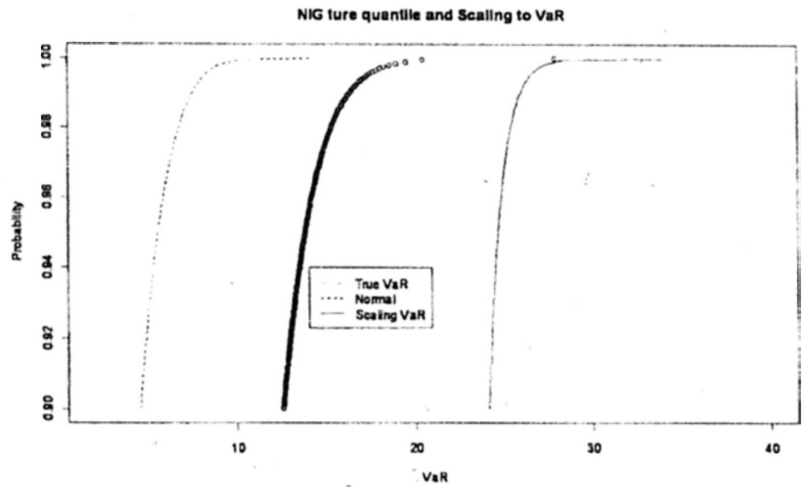


Fig. 5.1: NIG assumption with 10-day convolution equivalent rule.

2.6057. True theoretical VaR for multi-period is 7.0842 while multi-period VaR calculated from the semiparametric combined with semi-heavy scaling rule is 7.7457. The respective backtesting statistic $V^{freq} = \frac{1}{t_1-t_0+1}\sum_{t=t_0}^{t_1} I_{\{R_t \leq \hat{VaR}_p\}}$ for the true theoretical multi-period VaR and the corresponding estimated VaR is 99.3% and 99.6% respectively which shows the proposed method - semi-parametric estimation plus semi-heavy rule (denoted by SP-SH rule approach), is remarkably useful.

**(C)** *Saddlepoint approximated multi-period VaR.*

Simulate data to evaluate the saddlepoint approximation approach in the multi-period risk management. We want to test the accuracy of saddlepoint approximation to VaR under the NIG assumption model. Data are simulated from NIG(1,0.3,1,0) distribution with holding period 10 days. The tail probabilities are from 0.000001 to 0.1. The graph in Fig. 5.2 displays that the NIG saddlepoint approximation to the quantile at the tail probability are much better than the normal approximation. The approximation is a quick and accurate method that can be employed for higher quantile estimation.

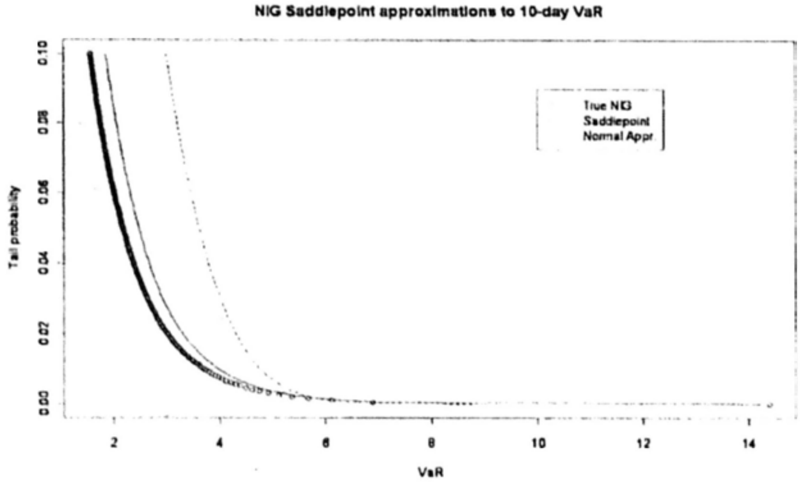


Fig. 5.2: Saddlepoint approximation to 10-day VaR.

Tab. 5.1: Descriptive statistics of Bank of America Corporation (BAC) and Citigroup, Inc. (C) daily Price.

| DESCRIPTION | n | Mean | Std | Skewness | Kurtosis |
|---|---|---|---|---|---|
| BAC | 6282 | 17.20152 | 12.68852 | 0.6977284 | -0.5884626 |
| Citigroup | 8658 | 12.31203 | 15.16061 | 1.035472 | -0.5155512 |

## 5.2 Real Example

The calculations and approximations of VaR are based on two financial time series: the stock prices of Bank of America Corporation (BAC) and Citigroup, Inc. (C). These daily price data are collected from 1986-05-29 and 1977-01-03 respectively, both to 2011-04-25. The data sets can be downloaded from the Yahoo Finance website. Each of the two series consists more than 5000 observations. Fig. 5.3 and Fig. 5.4 display the two data sets, the respective log-return plots and Q-Q plots. Table 5.1 gives the descriptive statistics of the two data sets.

From Fig. 5.3, Fig. 5.4 and Table 5.1, it is easily found the two data sets are typical financial data that exhibits fatter tail and higher kurtosis. Q-Q plots show they are distinctly different from the Gaussian distribution.

Usually, for multi-period problem, the holding period is of great importance. 10-day is the standard holding period required by the Basel Committee. It seems interesting to deal with shorter or longer holding period both for internal supervision and external regulation purpose, such as 5-day (5-day has been studied by many researchers), 30-day (30-day has been consider as the optimal holding period, see Courtois & Walter, 2010). So we try three holding periods - 10-day, 5-day and 30-day.

We want to obtain the multi-period VaR from scaling the single-period VaR on the two data sets. The choice of single-period VaR becomes our first consideration. We select three classical approaches, the traditional Gaussian-based VaR, the semi-parametric approach that we derived under tempered Pareto (semi-heavy) distribution assumption and the last one based a semi-heavy tailed parametric distribution - Normal Inverse Gaussian (NIG) distribution.

The single-period Value-at-Risk would be calculated by using three approaches:

- Normal VaR:

$$\widehat{VaR_p} = \hat{\mu} + \hat{\sigma}\Phi_p^{-1},$$

  where $\hat{\mu}$ - estimated mean and $\hat{\sigma}$ - estimated standard deviation are classical MLEs of normal distribution, $\Phi^{-1}$ is the inverse of the standard Gaussian distribution function;

- Semi-parametric VaR:

$$\widehat{VaR_p} \approx \text{solve}\{\hat{\rho}\log \widehat{VaR_p} + \hat{\gamma}\widehat{VaR_p} = \breve{C}\},$$

  where $\breve{C} = \hat{\rho}\log X_{(k+1)} + \hat{\gamma}X_{(k+1)} + \log(k) - \log[n(1-p)]$, $\hat{\rho}, \hat{\gamma}$ are the conditional MLEs, $\lceil k = n(1-p)\rceil$ and $X_{(k+1)}$

represents the order statistic. Details are given in Section 3.2;

- NIG VaR:

$$\widehat{VaR_p} = F_{NIG}^{-1}(\hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu}),$$

where $\hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu}$ are the MLEs of the NIG distribution and $F_{NIG}^{-1}$ is the quantile function of the NIG distribution.

The estimation results of all parameters are displayed as follows. For the BAC data, $\hat{\mu_1} = 0.000232, \hat{\sigma_1} = 0.026954$ in the normal assumption case, $\rho_1 = 1.436640, \gamma_1 = 6.383335$ in the case of semiparametric estimation and $\alpha_1 = 20.652314, \beta_1 = -0.165324, \delta_1 = 0.012976, \mu_1 = 0.000336$ in the NIG assumption case. For the Citygroup Data, $\hat{\mu_1} = 0.000185, \hat{\sigma_1} = 0.026584$ in the normal assumption case, $\rho_1 = 1.812778, \gamma_1 = 4.342397$ in the case of semiparametric estimation and $\alpha_1 = 22.568687, \beta_1 = 0.721413, \delta_1 = 0.013464, \mu_1 = -0.000245$ in the NIG assumption case. The plots of fitting of the NIG distribution are given in Fig. 5.5 and Fig. 5.6.

Two rules for external regulators, SP-CE rule and NIG-CE rule are employed to calculate conservative regulatory multi-period VaRs under significance level 99% and 95% with holding period 10-days, 5-days and 30-days. They are reported in Table 5.2, where the values are calculated through semi-parametric estimation and NIG fitting combined with respective correctly specified convolution equivalent rule (CE rule). Three scaling rules for internal risk managers to do multi-period risk management : SQRT rule, SP-SH rule and NIG-SH rule would be used based on each of the three single-period VaR (including correctly specified rule and misspecified rule). All calculation results of VaR with confidence level 95% and 99% have been illustrated in Table 5.3 - Table 5.6. Red values means the calculation of multi-period VaRs are calculated from the single-period Value-at-Risk by using correctly specified respective scaling rule

Tab. 5.2: 10-day Convolution equivalent rules for regulators: BAC and Citygroup data.

| Data | Period | Level | SP-CE rule | NIG-CE rule |
|---|---|---|---|---|
| BAC | 10-day | 99% | 0.319311 | 0.304337 |
| | 10-day | 95% | 0.283579 | 0.266153 |
| | 5-day | 99% | 0.232233 | 0.204997 |
| | 5-day | 95% | 0.196501 | 0.166813 |
| | 30-day | 99% | 0.520889 | 0.621285 |
| | 30-day | 95% | 0.485157 | 0.583101 |
| Citygroup | 10-day | 99% | 0.325380 | 0.302425 |
| | 10-day | 95% | 0.293507 | 0.264977 |
| | 5-day | 99% | 0.233378 | 0.202416 |
| | 5-day | 95% | 0.201505 | 0.164968 |
| | 30-day | 99% | 0.545195 | 0.625838 |
| | 30-day | 95% | 0.513322 | 0.588390 |

while black values means the calculation of multi-period VaRs are calculated from scaliing the single-period Value-at-Risk by using misspecified scaling rule. Table 5.3 and Table 5.4 display the 99% and 95% VaR of 10-day VaR, Table 5.5 and Table 5.6 display the 99% and 95% VaR of 5-day VaR while Table 5.3 and Table 5.4 display the 99% and 95% VaR of 30-day VaR. The three holding periods represent the regular supervision, shorter supervision and longer supervision, respectively.

External risk regulators always consider that multi-period risk are higher than what we actually seen. For banking supervision departments, an external multi-period risk should be a conservative indicator so that the public can take some measures in time of falling to prevent the outbreak of larger financial catastrophe. The values calculated through SP-CE rule and NIG-CE rule are displayed in Table 5.2.

Table 5.2 shows that the convolution equivalent rule for multi-period VaR would be the most prudential choice for external regulators. It may avoid a lot of potential losses caused in the long run because the larger VaR value would correspond to an increase of capital requirement to prepare for the unforeseen emergencies. The results seem more conservative than the SQRT rule. The drawback may be that the CE rule lacks risky spirit in supervising which enlarge the possibility of hampering the development of the economy. McNeil and Frey (2000) empirically justified such a result so that prudential values of VaR should be more feasible for external regulation.

The suitability of the internal models for estimating financial risks should be compared. Backtesting is recommended by the Basel Committee to estimate the model risk. The idea is to compare the $d$-day VaR with the actual observed profit or loss over the next $d$ days. Given the observed price data, firstly calculate log-returns $R_t$ (In multi-period problem, the sum of single-period log-return is the log-return with respective holding period). Then the most commonly used statistic that provides information about the quality of the internal calculation of VaR - *the frequency of exceedances.*

$$V^{freq} = \frac{1}{t_1 - t_0 + 1} \sum_{t=t_0}^{t_1} I_{\{R_t \leq V\hat{a}R_p\}}$$

This measure is used by the Basel Committee on Banking Supervision and a good estimation for VaR will lead to a value which is close to the level $p$. In this chapter, $p$ are chosen to equal to 0.95 or 0.99. The calculated multi-period values of VaR are shown in all four tables for comparing the scaling rules.

Table 5.3 - Table 5.6 display the VaRs calculated from three internal scaling rules: SQRT rule, SP-SH rule and NIG-SH rule. Among red values in the diagonal, the SP-SH rule seems the most excellent tool of internal risk supervision for risk managers in financial institutions because the values of $V^{freq}$ are the clos-

Tab. 5.3: 10-day 99% VaR calculation results of BAC and Citygroup data.

| Data | Methods | SQRT rule ($V^{freq}$) | SP-SH rule ($V^{freq}$) | NIG-SH rule ($V^{freq}$) |
|---|---|---|---|---|
| BAC | Normal | 0.199024 | 0.194146 | 0.206730 |
| | | (0.988042) | (0.987085) | (0.989158) |
| | Semiparametric | 0.221728 | 0.209021 | 0.2179016 |
| | | (0.990274) | (0.989318) | (0.990274) |
| | NIG | 0.236700 | 0.218490 | 0.225026 |
| | | (0.992028) | (0.990274) | (0.990593) |
| Citygroup | Normal | 0.196149 | 0.183810 | 0.208875 |
| | | (0.989709) | (0.988437) | (0.990981) |
| | Semiparametric | 0.209982 | 0.194688 | 0.219931 |
| | | (0.991096) | (0.989709) | (0.992137) |
| | NIG | 0.234398 | 0.213484 | 0.226984 |
| | | (0.993293) | (0.991559) | (0.992715) |

est to the true $p$ among all three holding periods. It seems the NIG-SH rule performs better in the higher confidence level than the SQRT rule while it is no better in the lower confidence level. For simplicity, the SQRT rule would be an acceptable choice when the confidence level is 99%, but it is not recommended when the confidence level is 95%. Overall, among all correctly specified approaches, the SP-SH rule is the best. In the misspecified case of scaling rules, which means among all non-diagonal block values, the SQRT rule has performed very well, which verified the conclusion of Kaufmann (2005). If we have obtained a better estimation of single-period VaR than normal-based estimation of VaR, the SQRT rule can perform very well in calculation of multi-period VaR, either under the semi-parametric structure or the parametric structure. The SP-SH rule also performs constantly well although some values seem small when the holding period is long. The NIG-SH rule highly overestimates

Tab. 5.4: 10-day 95% VaR calculation results of BAC and Citygroup data.

| Data | Methods | SQRT rule ($V^{freq}$) | SP-SH rule ($V^{freq}$) | NIG-SH rule ($V^{freq}$) |
|---|---|---|---|---|
| BAC | Normal | 0.140936 | 0.127627 | 0.175354 |
| | | (0.975765) | (0.969228) | (0.984375) |
| | Semiparametric | 0.108663 | 0.101551 | 0.155205 |
| | | (0.953125) | (0.945950) | (0.980389) |
| | NIG | 0.115951 | 0.107567 | 0.159997 |
| | | (0.961097) | (0.951690) | (0.981027) |
| Citygroup | Normal | 0.138860 | 0.119441 | 0.177851 |
| | | (0.975486) | (0.961725) | (0.987512) |
| | Semiparametric | 0.109147 | 0.094054 | 0.157941 |
| | | (0.953053) | (0.937442) | (0.982655) |
| | NIG | 0.115977 | 0.099897 | 0.162675 |
| | | (0.959066) | ( 0.943571) | (0.984158) |

the VaR when the confidence level is 95%.

The NIG saddlepoint approximated VaRs for two data sets are also calculated in Table 5.9. This seems also an accurate internal multi-period VaR approximation model especially when the confidence level is 99%. It is a direct calculation approach without scaling from single-period. Since the approximation is mainly on the tail, the calculated values could be more close to the true VaR in the high confidence level than in the low confidence level. The saddlepoint approximation approach is a fast and an accurate internal VaR model for multi-period risk management.

---

□ **End of chapter.**

Tab. 5.5: 5-day 99% VaR calculation results of BAC and Citygroup data.

| Data | Methods | SQRT rule ($V^{freq}$) | SP-SH rule ($V^{freq}$) | NIG-SH rule ($V^{freq}$) |
|---|---|---|---|---|
| BAC | Normal | 0.140731 | 0.142237 | 0.137024 |
| | | (0.986936) | (0.987255) | (0.985981) |
| | Semiparametric | 0.156785 | 0.154751 | 0.146948 |
| | | (0.990441) | (0.989963) | (0.988211) |
| | NIG | 0.167372 | 0.162791 | 0.153328 |
| | | (0.991556) | (0.991238) | (0.989485) |
| Citygroup | Normal | 0.138699 | 0.137265 | 0.137549 |
| | | (0.988790) | (0.988097) | (0.988212) |
| | Semiparametric | 0.148480 | 0.146101 | 0.147398 |
| | | (0.990524) | (0.990292) | (0.990292) |
| | NIG | 0.165744 | 0.161511 | 0.153731 |
| | | (0.992257) | (0.991910) | (0.991332) |

Tab. 5.6: 5-day 95% VaR calculation results of BAC and Citygroup data.

| Data | Methods | SQRT rule ($V^{freq}$) | SP-SH rule ($V^{freq}$) | NIG-SH rule ($V^{freq}$) |
|---|---|---|---|---|
| BAC | Normal | 0.099656 | 0.093899 | 0.109768 |
| | | (0.971961) | (0.967819) | (0.977218) |
| | Semiparametric | 0.076836 | 0.073829 | 0.092869 |
| | | (0.952684) | (0.946790) | (0.967182) |
| | NIG | 0.081990 | 0.078423 | 0.096837 |
| | | (0.958101) | (0.954278) | (0.970209) |
| Citygroup | Normal | 0.098189 | 0.089856 | 0.110512 |
| | | (0.971571) | (0.964059) | (0.978042) |
| | Semiparametric | 0.077179 | 0.070464 | 0.093747 |
| | | (0.948342) | (0.937709) | (0.967063) |
| | NIG | 0.082008 | 0.074914 | 0.097684 |
| | | (0.956200) | (0.945106) | (0.971339) |

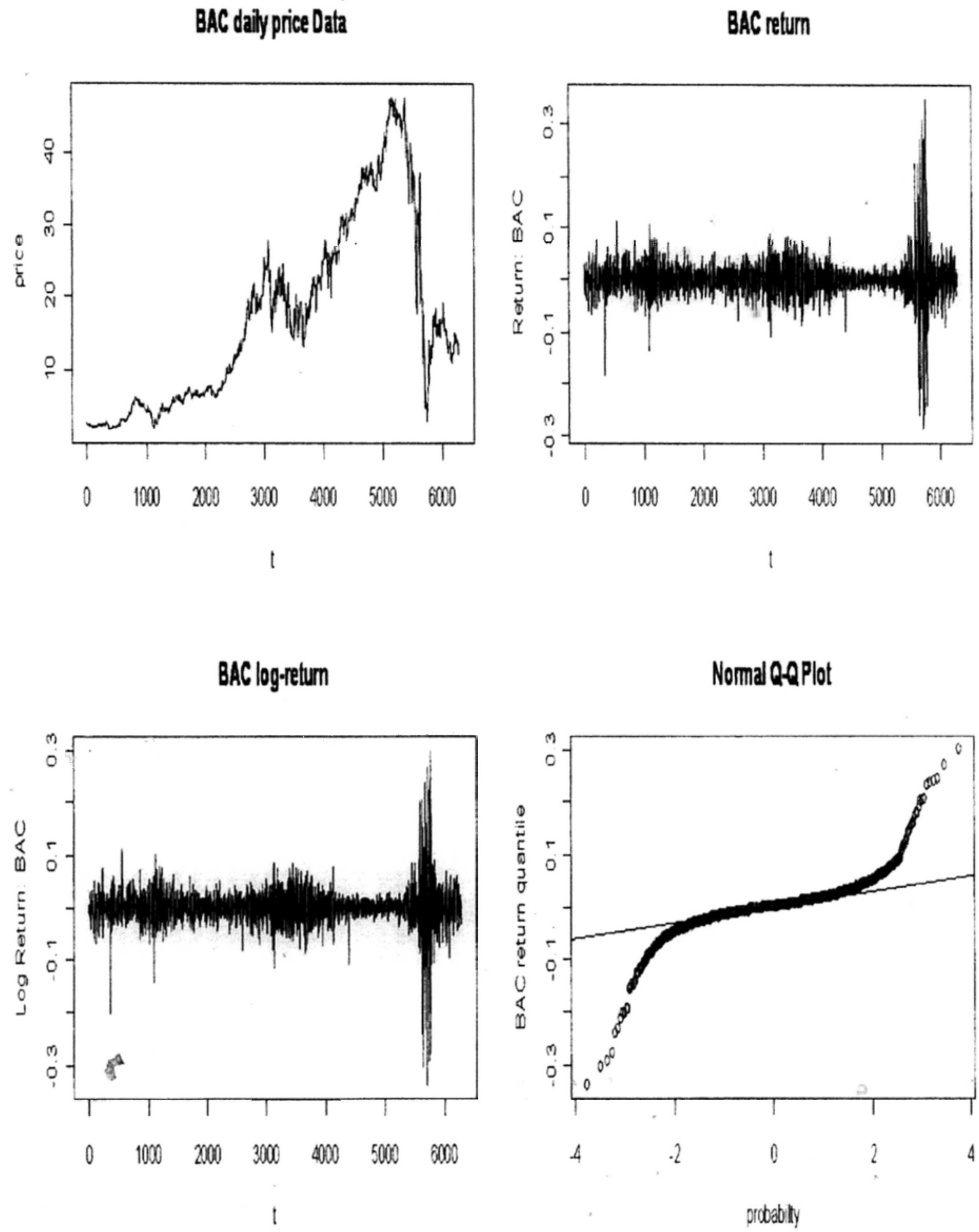


Fig. 5.3: Bank of America Corporation (BAC) data.

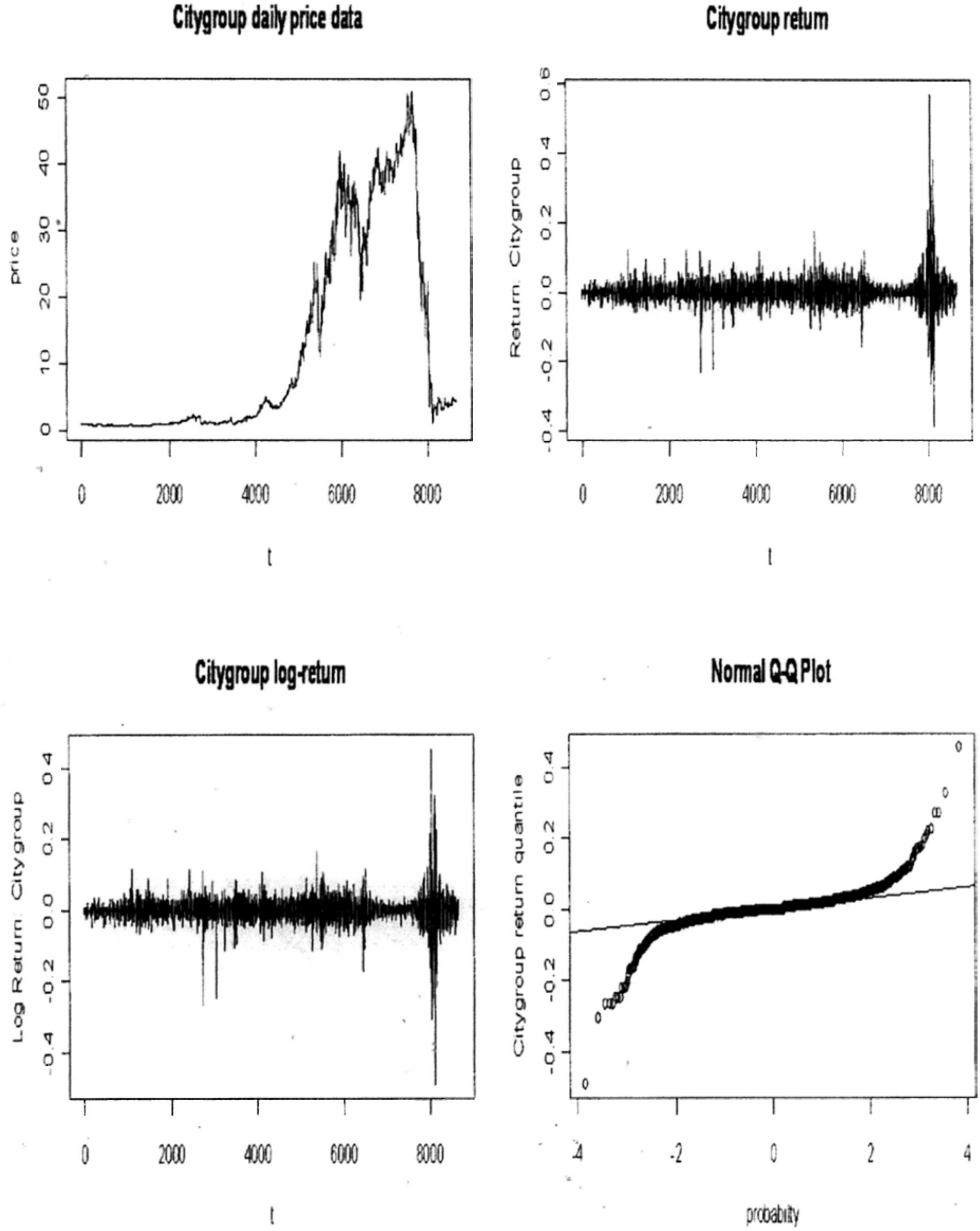


Fig. 5.4: Citigroup, Inc. (C) data.

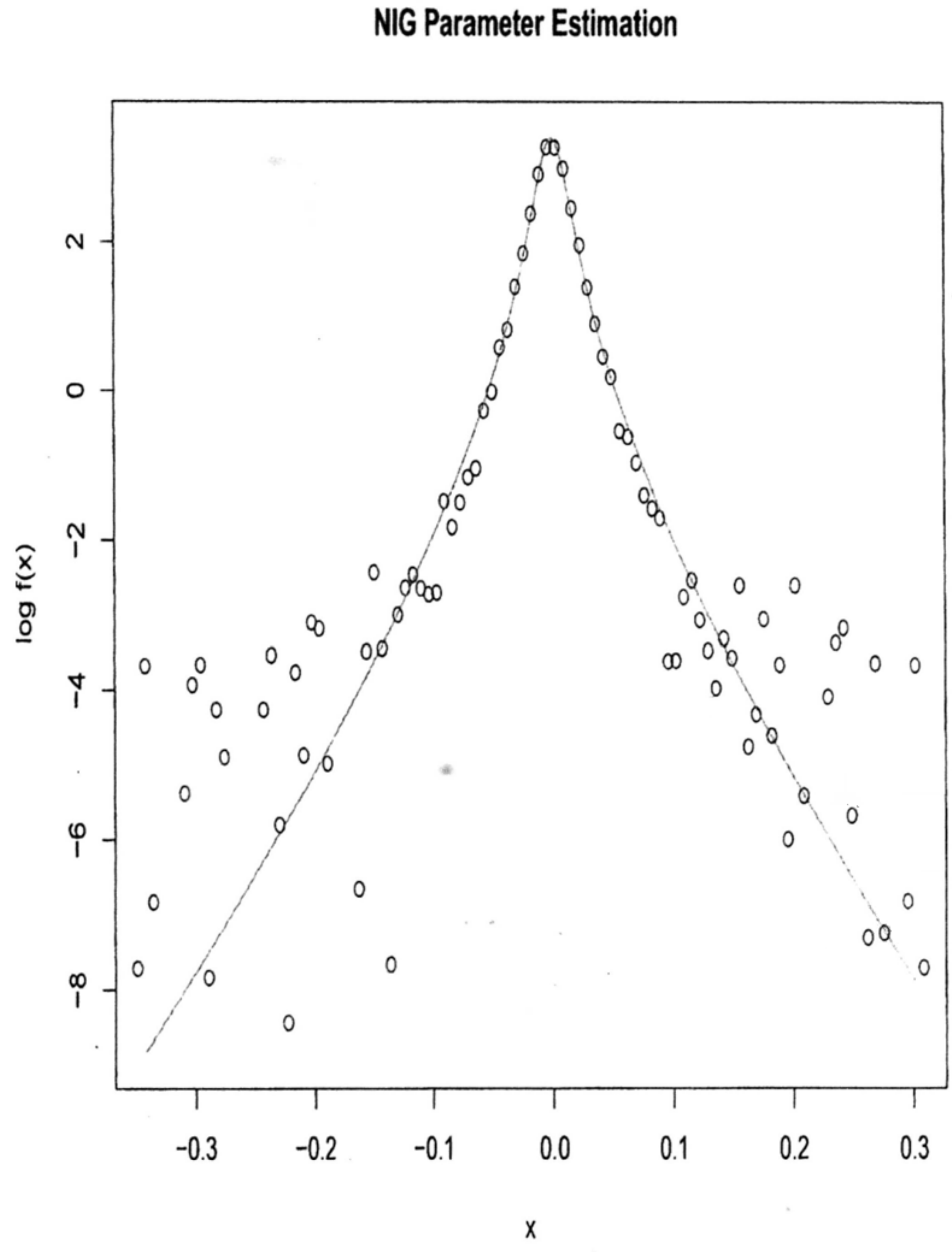


Fig. 5.5: Bank of America Corporation (BAC) data with NIG fit.

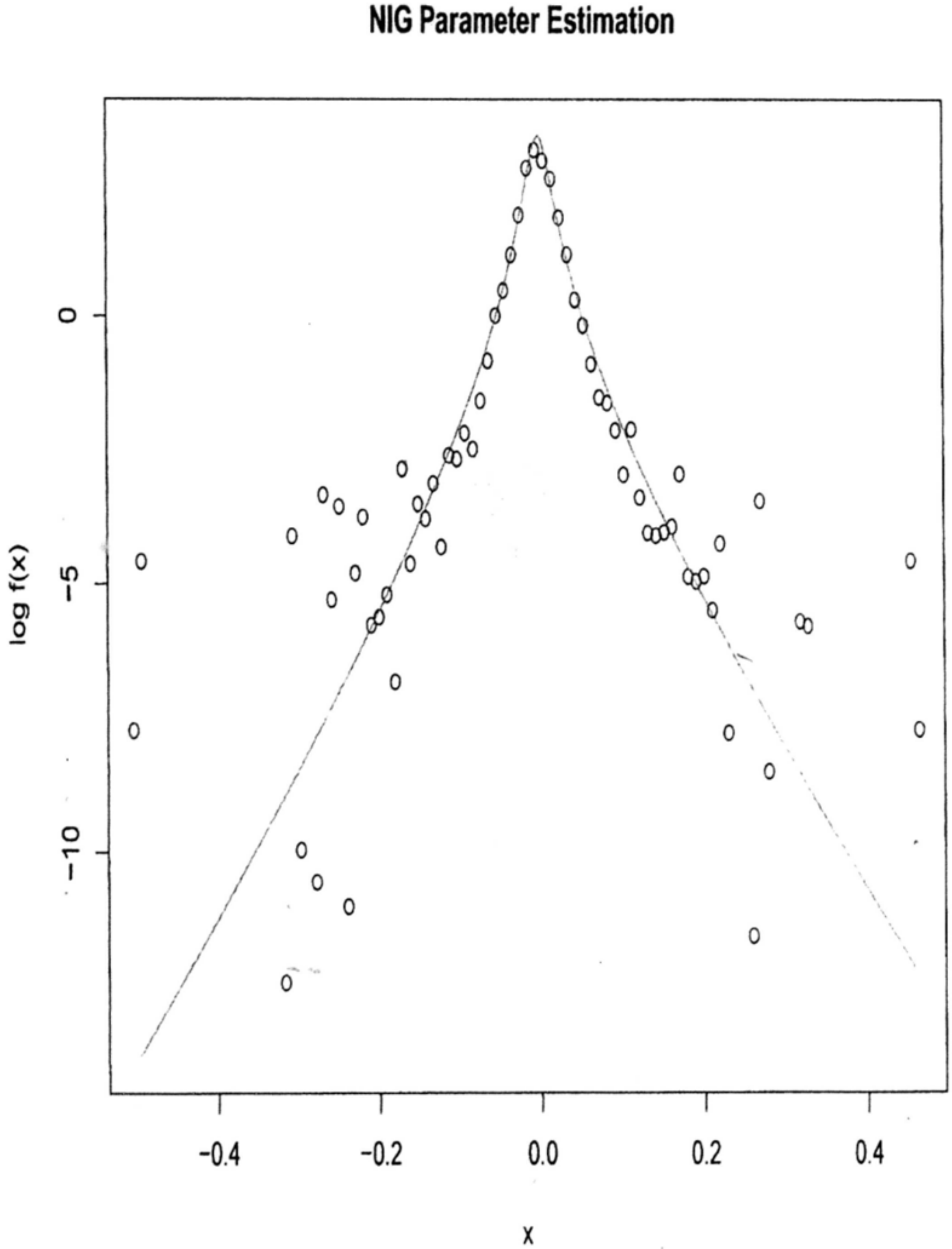


Fig. 5.6: Citigroup, Inc. (C) data with NIG fit.

Tab. 5.7: 30-day 99% VaR calculation results of BAC and Citygroup data.

| Data | Methods | SQRT rule ($V^{freq}$) | SP-SH rule ($V^{freq}$) | NIG-SH rule ($V^{freq}$) |
|---|---|---|---|---|
| BAC | Normal | 0.344720 | 0.312495 | 0.465234 |
| | | (0.992163) | (0.988644) | (0.996481) |
| | Semiparametric | 0.384044 | 0.331023 | 0.478214 |
| | | (0.994082) | (0.991363) | (0.996641) |
| | NIG | 0.409976 | 0.342677 | 0.486429 |
| | | (0.995042) | (0.992163) | (0.996961) |
| Citygroup | Normal | 0.339741 | 0.286618 | 0.475769 |
| | | (0.992466) | (0.987367) | (0.997218) |
| | Semiparametric | 0.363700 | 0.301024 | 0.488546 |
| | | (0.993857) | (0.989569) | (0.997450) |
| | NIG | 0.405989 | 0.325570 | 0.496639 |
| | | (0.996291) | (0.991423) | (0.997450) |

Tab. 5.8: 30-day 95% VaR calculation results of BAC and Citygroup data.

| Data | Methods | SQRT rule ($V^{freq}$) | SP-SH rule ($V^{freq}$) | NIG-SH rule ($V^{freq}$) |
|---|---|---|---|---|
| BAC | Normal | 0.244108 | 0.205667 | 0.428007 |
| | | (0.976807) | (0.964171) | (0.995202) |
| | Semiparametric | 0.188209 | 0.167592 | 0.403340 |
| | | (0.954255) | (0.939859) | (0.994882) |
| | NIG | 0.200833 | 0.176503 | 0.409270 |
| | | (0.961292) | (0.946897) | (0.995042) |
| Citygroup | Normal | 0.240512 | 0.186129 | 0.439197 |
| | | (0.978790) | (0.950974) | (0.996755) |
| | Semiparametric | 0.189049 | 0.147880 | 0.415023 |
| | | (0.953176) | (0.913421) | (0.996523) |
| | NIG | 0.200879 | 0.156739 | 0.420831 |
| | | (0.963027) | (0.924316) | (0.996639) |

Tab. 5.9: Saddlepoint approximated VaRs of BAC and Citygroup data.

| Data | Period | Level | VaR | $V^{freq}$ |
|---|---|---|---|---|
| BAC | 10-day | 99% | 0.207593 | 0.989318 |
| | 10-day | 95% | 0.137342 | 0.974490 |
| | 5-day | 99% | 0.156103 | 0.990282 |
| | 5-day | 95% | 0.098841 | 0.971643 |
| | 30-day | 99% | 0.340379 | 0.992003 |
| | 30-day | 95% | 0.235800 | 0.975048 |
| Citygroup | 10-day | 99% | 0.203705 | 0.990634 |
| | 10-day | 95% | 0.134751 | 0.972248 |
| | 5-day | 99% | 0.153335 | 0.990986 |
| | 5-day | 95% | 0.097265 | 0.970762 |
| | 30-day | 99% | 0.333119 | 0.991886 |
| | 30-day | 95% | 0.230178 | 0.975429 |

# Chapter 6

# Conclusion and Further Research

This thesis studied multi-period financial risk management and developed new scaling rules based on distributions belonging to the *convolution equivalent class* and the *semi-heavy tailed distribution class* - the CE rule and the SH rule, respectively. The two scaling rules have their own conditions and assumptions. Fig. 6.1 displays the applicable conditions, the relationships and the related background. And the thesis also employed a saddlepoint approximation method to obtain the multi-period VaR which provides a direct calculation of the internal model for controlling risk. Based on the assumption of semi-heavy tail, the estimation of single-period VaR by semi-parametric approach (SP) is derived and the respective scaling rules are also applied. The research found that the CE rules are suitable for conducting external risk management because the VaRs scaling by using the CE rule are conservative enough to enhance the capital requirements. However, the SH rule can be considered as an alternative to the SQRT rule. People need more accurate calculation approaches of VaR for internal use. And the SH rule offers such an internal scaling model for risk managers. The newly derived scaling rules can be divided into two groups. One is based on the semiparametric semi-heavy rule (such as the SP-CE rule

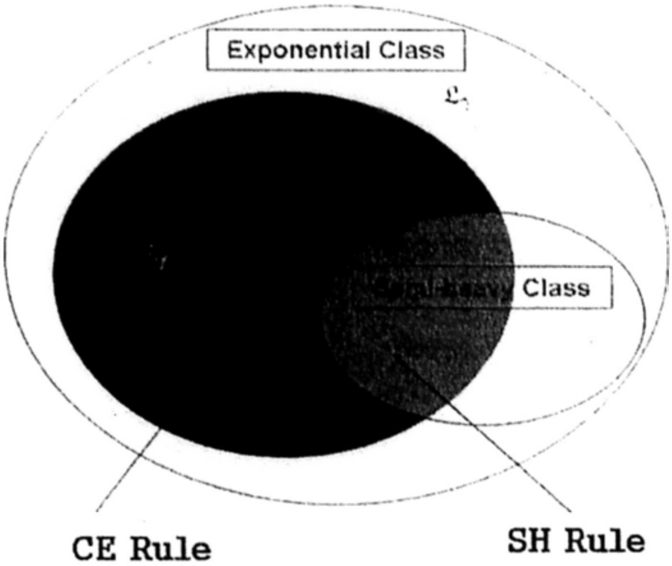


Fig. 6.1: The assumption and relationship of two scaling rules.

and SP-SH rule) and the other is based on some parametric structure that satisfies the regularity conditions of scaling rules. For example, under the normal inverse Gaussian distribution (NIG) model, there are the NIG-CE rule and the NIG-SH rule. Through real data analysis, the SP-CE rule and the NIG-CE rule provide prudential choice for external regulators. Among correctly specified internal models, the main finding is that the SP-SH rule performs the best among three internal scaling models (the SQRT rule, the SP-SH rule and the NIG-SH rule). The NIG-SH rule performs better when the quantile is high while overestimation of VaR would happen when the quantile is low. Due to the theoretical drawbacks of the SQRT rule, the semi-heavy scaling rule combined with semi-parametric estimation of single-period VaR (the SP-SH rule) are recommended to risk managers as the best choice for multi-period internal risk management.

For further research, the extensions of the CE rule and the SH rule to a continuous process may stimulate more theory on the convolution equivalent class and semi-heavy tail analysis. It is also an open question whether the conditional heteroscedastic model with semi-heavy tailed innovations can outperform the traditional Gaussian assumption or other assumption in the multi-period risk management problem. The multivariate case would be another direction for extension of this thesis.

□ **End of chapter.**

# Bibliography


[1] Ahmed, S., Filipović, D. and Svindland, G. (2008). A note on natural risk statistics. *Operations Research Letters* **36**, 662–664.

[2] Albin, J.M.P. and Sundén, M. (2009). On the asymptotic behaviour of Lévy processes, Part I : Subexponential and exponential processes. *Stochastic Processes and their Applications* **119**, 281–304.

[3] Angelidis, T., Benos, A. and Degiannakis, S. (2004). The use of GARCH models in VaR estimation. *Statistical Methodology* **1**, 105–128.

[4] Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance* **9**, 203–228.

[5] Barndorff-Nielsen, O.E. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proc. Roy. Soc. London Wer. A* **353**, 401–419.

[6] Barndorff-Nielsen, O.E. (1997). Normal inverse Gaussian processes and stochastic volatility modeling. *Scandinavian Journal of Statistics* **24**, 1–13.

[7] Barndorff-Nielsen, O.E. (1998). Processes of normal inverse Gaussian type. *Finance Stochast.* **2**, 41–68.

[8] Baten, W.D. (1934). The probability law for the sum of n independent variables,each subject to the law $1/2h\ \mathrm{sech}(\pi x/2h)$. *Bull. Amer. Math. Soc.* **40**, 284–290.

[9] Beder, T. (1995). VAR: seductive but dangerous. *Financial Analysts Journal* **51**, 12–24.

[10] Berkowitz, J., Christoffersen, P. and Pelletier, D. (2005). Evaluating Value-at-Risk models with desk-level data. *Management Science, Articles in Advance* **Jan**, 1–15.

[11] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation* . Cambridge University Press, Cambridge.

[12] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307–327.

[13] Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992). ARCH modeling in finance: a selective review of the theory and empirical evidence. *Journal of Econometrics* **52**, 5–59.

[14] Bradley, B.O. and Taqqu, M.S. (2003). Financial risk and heavy tail. In Rachev (ed.) *Handbook of Heavy-tailed Distributions in Finance* . North Holland.

[15] Braverman, M. (1997). Suprema and dojourn times of Lévy processes with exponential tails. *Stochastic Process. Appl.* **68**, 265–283.

[16] Braverman, M. and Samorodnitsky, G. (1995). Functionals of infinitely divisible stochastic processes with exponential tail. *Stochastic Process. Appl.* **56**, 207–231.

[17] Broda, S. and Paolella, M. (2009). CHICAGO: A fast and accurate method for portfolio risk calculation. *Journal of Financial Econometrics* **7**, 412–436.

[18] Brooks, C., Clare, A.D., Dalle-Mulle, J.W. and Persand, G. (2005). *Journal of Empirical Finance* **12**, 339–352.

[19] Brummelhuis, R. and Kaufmann, R. (2004). Time scaling for GARCH(1,1) and AR(1)-GARCH(1,1) processes. *Working paper* , RiskLab, ETHZ.

[20] Butler, R.W. (2007). *Saddlepoint Approximations with Applications* . Cambridge University Press, Cambridge.

[21] Cai, Z. and Wang, X. (2008). Nonparametric estimation of conditional VaR and expected shortfall. *Journal of Econometrics* **147**, 120–130.

[22] Carr, P., Geman, H., Madan, D. and Yor, M. (2002). The fine structure of asset returns: an empirical investigation. *Journal of Business* **75-2**, 305–332.

[23] Chen, S.X. and Tang, C.Y. (2005). Nonparametric inference of Value-at-Risk for dependent financial returns. *Journal of Financial Econometrics* **3-2**, 227–255.

[24] Chen,Y., Härdle, W. and Jeong, S.O. (2008), Nonparametric risk management with Generalized Hyperbolic distributions, *Journal of the American Statistical Association* **14**, 910–923.

[25] Chen,Y., Härdle, W. and Spokoiny, V. (2010), GHICA: risk analysis with GH distributions and independent components, *Journal of Empirical Finance* **17**, 255–269.

[26] Cheng, B.N., and Rachev, S.T. (1995). Multivariate stable securities in financial markets. *Manuscript*.

[27] Chover, J., Ney, P. and Wainger, S. (1973a). Functions of probability measures. *J. Analyse Math* **26**, 255–302.

[28] Chover, J., Ney, P. and Wainger, S. (1973b). Degeneracy properties of subcritical branching process. *Ann. Probab.* **1**, 663–673.

[29] Christyakov, V.P. (1964). A theorem on sums of independent positive random variables and its applications to branching processes. *Theory Probab. Appl.* **9**, 640–648.

[30] Christoffersen, P. (1998). Evaluating internal forecasts. *International Economic Review* **39**, 841–862.

[31] Christoffersen, P. and Pelletier, D. (2004). Backtesting Value-at-Risk: a duration-based approach. *Journal of Financial Econometrics* **2**, 84–108.

[32] Clauset, A., Shalizi, C.R. and Newman, M.E.J. (2009). Power-law distributions in empirical data. *SIAM Review* **51-4**, 661–703.

[33] Cline, D.B.H. (1987). Convolutions of distributions with exponential and subexponential tails. *J. Austral. Math. Soc. (Series A)* **43**, 347–365.

[34] Cont, R. (2001). Empirical properties of asset returns, stylized facts and statistical issues. *Quantitative Finance* **1**, 1–14.

[35] Courtois, O.L. and Watlter, C. (2010). A study on Value-at-Risk and Lévy processes. *Conference paper* . ICA2010.

[36] Dacorogna, M., Muller, U., Pictet, O. and De Vries, C. (2001). Extremal forex returns in extremely large data sets. *Extremes.* **4**, 105–127.

[37] Daniel, H.E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Stat.* **25**, 631–650.

[38] Daniels, H. E. (1987). Tail probability approximations. *International Statistical Review* **55**, 37-48.

[39] Danielson, J. and De Vries, C. (1998). Beyond the sample: extreme quantile and probability estimation. *Discussion paper* Tinbergen Institute, No 98-016/2.

[40] Danielson, J. and De Vries, C. (2000). Value-at-Risk and extreme returns. *Annales d'Economie et de Statistique* **60**, 239-270.

[41] Danielsson, J., Hartmann, P. and De Vries, C. (1998). The cost of conservatism: Extreme returns, Value-at-Risk, and the Basel 'Multiplication Factor'. *Risk* **11**. 101 103.

[42] Danielsson, J., Jorgensen, B. N., Samorodnitsky, G., Sarma, M., and de Vries, C. G. (2005). Subadditivity Re-Examined: The Case for Value-at-Risk. *Manuscript*.

[43] Danielsson, J. and Zigrand, J. (2005). On time-scaling of risk and the square-root-of-time rule. *Working paper,* Manuscript.

[44] Diebold, F.X., Hickman, A., Inoue, A. and Schuermann, T. (1998). Converting 1-day volatility to h-day volatility: scaling by $\sqrt{h}$ is worse than you think. *Risk* **49**, 335–347.

[45] Dowd, K., Blake, D. and Cairns, A. (2003). Long-term Value-at-Risk. *Working paper* , Manuscript.

[46] Drees, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli* **9**, 617 657.

[47] Drost, F.C. and Nijman, T.E. (1993). Temporal aggregation of GARCH processes. *Econometrica* **61**, 909–927.

[48] Duffie, D. and Pan, J. (1997). An overview of Value-at-Risk. *Journal of Derivatives* **4**, 7–49.

[49] Eberlein, E. and Madan, D.B. (2009). The distribution of returns at longer horizon. *Working paper*, Manuscript.

[50] Eberlein, E. and Hammerstein, E.A.v. (2003). Generalized hyperbolic and inverse Gaussian distributions: limiting cases and approximation of processes. *Working paper*, Manuscript.

[51] Eling, M. and Tibiletti, L. (2010). Internal vs. external risk measures: How capital requirements differ in practice. *Operations Research Letters* **38**, 482–488.

[52] Embrechts, P. (1983). A property of the generalized inverse Gaussian distribution with some applications. *Journal of Probability* **20**, 537–544.

[53] Embrechts, P , Goldie, C.M. and Veraverbeke, N. (1979). Subexponentiality and infinite divisibility. *Probability Theory and Related Fields* **49**, 335–347.

[54] Embrechts, P. and Goldie, C.M. (1980). On closure and factorization properties of subexponential and related distributions. *J. Austral. Math. Soc. (Ser.A)* **29**, 243–256.

[55] Embrechts, P. and Goldie, C.M. (1982). On convolution tails. *Stochastic Processes and their Applications*, **13-3**, 263–278.

[56] Embrechts, P., Kaufmann, R. and Patie, P. (2004). Strategic long-term financial risks: single risk factors. *Working paper* , RiskLab, ETH, Zurich.

[57] Embrechts, P. Klüppelberg, C. and Mikosch, T. (1997). *Modeling Extremal Events for Insurance and Finance*. Applications of Mathematics 33, Springer.

[58] Engle, R.F. (2002). New frontiers for ARCH models. *Journal of Applied Economics* **17**, 425–446.

[59] Engle, R.F. and Manganell, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics* **22**, 367–381.

[60] Epperlein, E. and Smillie, A. (2006). Cracking VaR with kernels. *RISK* **19**, 70–74.

[61] Fan, J. and Gu, J. (2003). Semiparametric estimation of Value-at-Risk. *Econometrics Journal* **6**, 261–290.

[62] Fasen, V. (2009). Extremes of Lévy driven mixed MA processes in the class of convolution equivalent distributions. *Extremes* **12-3**, 265–296.

[63] Feller, W. (1971). *An Introduction to Probability Theory and its Applications, Vol. II, 2nd ed.*. Wiley, New York.

[64] Fenner, T., Levene, M. and Loizou, G. (2005). A stochastic evolutionary model exhibiting power-law behaviour with an exponential cutoff. *Physica A* **355**, 641–656.

[65] Feuerverger, A. and Wong, A.C.M. (2000). Computation of value-at-risk for nonlinear portfolios. *Journal of Risk* **3-1**, 37–55.

[66] Fong, G. and Vasicek, O.A. (1997). A multidimensional framework for risk analysis. *Financial Analysts Journal* **July/August**, 51–58.

[67] Foss, S. and Korshunov, D. (2007). Lower limits and equivalences for convolution tails. *The Annals of Probability* **35**, 366–383.

[68] Frain, J.C. (2008). Value-at-risk and the $\alpha$−stable distribution. *TEP working paper* **No.** 0308.

[69] Giannopoulos, K (2003). VaR modeling on long run horizons. *Automation and Remote Control* **64**, 1094–1100.

[70] Glosten, L., Jagannathan, R. and Runkle, D. (1993). Relationship between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* **48** 1779–1801.

[71] Goldberg, L.R., Miller, G and Weinstein, J. (2008). Beyond value at risk: forecasting portfolio loss at multiple horizons. *Journal of Investment Management* **6-2**, 73 98.

[72] Goldie, C.M. and Klüppelberg, C. (1998). Subexponential distributions. In Adler, Feldman and Taqqu (ed.) *A Practical Guide to Heavy Tails* . 435–459.

[73] Gourieroux, C., Laurent, J.P. and Scaillet, O. (2000). Sensitivity analysis of Values at Risk. *Journal of Empirical Finance* **7**, 225–245.

[74] Grayling, S. (1997). *VaR: Understanding and Applying Value-at-Risk* . London: Risk.

[75] Goutis, C. and Casella, G. (1999). Explaining the saddlepoint approximation. *The American Statistician* **53**, 216-"224.

[76] Hafner, C.M. (2008). Temporal aggregation of multivariate GARCH processes. *Journal of Econometrics* **142**, 467–483.

[77] Harkness, W.L. and Harkness, M.L. (1968). Generalized hyperbolic secant distributions. *J. Amer. Statist. Assoc.* **63**, 329–337.

[78] Heston, S. L. (1993). A closed-form solution for options with stochastic volatility, with applications to bond and currency options. *Review of Financial Studies* , **6**, 327–343.

[79] Heyde, C.C. and Kou, S.G. (2004). On the controversy over tailweight of distributions. *Operations Research Letters* **32**, 399–408

[80] Heyde, C.C., Kou, S.G. and Peng X.H. (2007). What is a good external risk Measure: bridging the gaps between robustness, subadditivity, and insurance risk. *Columbia University working paper*. Preprint.

[81] Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics* **19**, 1547–1569.

[82] Huang, A. Y. (2009). A value-at-risk approach with kernel estimator. *Applied Financial Economics* **19-5**, 379–395.

[83] Huang, X., Oosterlee, C.W. and Mesters, M.A.M (2007). Computation of VaR and VaR Contribution in the Vasicek portfolio credit loss model: a comparative study. *Report of the Department of Applied Mathematical Analysis* REPORT 07-06. ISSN 1389-6520.

[84] Huzurbazar, S. (2006). *Saddlepoint Approximations* In: *Encyclopedia of Statistical Science 2nd Edition.* Wiley, New York.

[85] Jensen, J.L. (1994). *Saddlepoint Approximations* . Oxford Statistical Science, Oxford.

[86] Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions, Volume 1, Second Edition.* Wiley, New York.

[87] Jørgensen, B. (1997). *The Theory of Dispersion Models* . Chapman and Hall, New York.

[88] Jorison, P. (2007). *Value-at-Risk: The New Benchmark for Managing Financial Risk* . McGraw-Hill, New York.

[89] JP Morgan, Reuters (1996). *RiskMetrics Technical Document* . 4th Edition.

[90] Khindanova, I. and Rachev, S. (2000). Value-at-Risk Recent Advances. In: Anastassiou, G (ed)*Handbook of Analytic Computational Methods in Applied Mathematics*. CRC Press. 801 885.

[91] Kaufmann, R. (2005). *Long-Term Risk Management* . ETHZ PhD thesis.

[92] Kaufmann, R. and Patic, P. (2003). Strategic long-term financial risks: the one-dimensional case. *Working paper* RiskLab, ETH, Zurich.

[93] Klüppelberg, C. (1989). Subexponential distributions and characterizations of related classes. *Prob. Theory Relat. Fields* **82**, 259 269.

[94] Kotz, S., Kozubowski, T. and Krysztof, P. (2001). *The Laplace Distribution and Generalizations.* Birkhäuser, Boston.

[95] Kupiec, P. (1995). Techniques for verifying the accuracy of risk management models. *Journal of Derivatives* **3**, 73 84.

[96] Kupiec, P. and O'Brien, J. M. (1995). The use of bank measurement models for regulatory capital purposes. *FEDS Working Paper* **No.** 95-11.

[97] Lowenstein, R. (2002). *When Genius Failed: The Rise and Fall of Long-Term Capital Management. Paperback* Fouth Estate (Harper-Collins).

[98] Lugannani, R. and Rice, S. (1980). Saddlepoint approximations for the distribution of the sum of independent random variables. *Advances in Applied Probability* **12**, 475–490.

[99] Mantegna, R. and Stanley, H. (1995). Scaling behavior of an economic index. *Nature* **376**, 46–49.

[100] Martin, R., Thompson, K. and Browne, C. (2001). Taking to the saddle. *Risk* **Jun**, 91–94.

[101] McCulloch, J.H. (1996). Financial applications of stable distributions. In Maddala and Rao (ed.) *Handbook of Statistics, Vol. 14* . Elsevier, Amstadam, 393–425.

[102] McNeil, A.J., and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance* **7 (3-4)**, 271–300.

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

[104] Meerschaert, Roy, P. and Shao, Q. (2010). Parameter estimation for exponentially tempered power law distributions. *Communications in Statistics - Theory and Methods* , to appear.

[105] Menkens, O. (2007). Value at Risk and Self-Similarity. In Miller *et al.* , (ed.) *Numerical Methods for Finance* Chapman & Hall - CRC Financial Mathematics Series, 8. CRC Press, 225–253.

[106] Mittnik, S., Rachev, S. and Schwartz, E. (2002). Value-at-risk and asset allocation with stable return distributions.

*Allgemeines Statistisches Archiv, Physica-Verlag* **Jan**, 53–67.

[107] Morris, C.N. (1982). Natural exponential families with quadratic variance functions. *The Annals of Statistics* **10**, 65–80.

[108] Neftci, S.N. (2000). Value-at-Risk calculation, extreme events, and tail estimation. *Journal of Derivatives* **7**, 23–37.

[109] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* **59**, 347–370.

[110] Newman, M.E.J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics* **46**, 323–351.

[111] Owen, A. and Tavella, D. (1997). Scrambled nets for Value-at-Risk calculations. In Grayling, S.(ed.) *VaR: Understanding and Applying Value-at-Risk.*. London: Risk, 289–297.

[112] Pakes, A.G. (2004). Convolution equivalence and infinite divisibility. *J. Appl. Probab.* **41**, 407–424.

[113] Pakes, A.G. (2007). Convolution equivalence and infinite divisibility: Corrections and corollaries. *J. Appl. Probab.* **44**, 295–305.

[114] Poon, S. and Granger, C.J.W. (2003). Forecasting volatility in financial markets: a review. *Journal of Economic Literature* **44**, 295–305.

[115] Praetz, P. (1972). The distribution of share price changes. *Journal of Business* **45**, 49–55.

[116] Prause, K. (1999). The generalized hyperbolic model: estimation, financial derivatives and risk measures. *PhD thesis,* University of Freiburg.

[117] Provizionatou, V., Markose, S. and Menkens, O. (2005). Empirical scaling rules for Value-at-Risk (VaR). *Working paper* , University of Essex

[118] Rogozin, B.A. (2000). On the constant in the definition of subexponential distributions. *Theory Prob. Appl.* **44**, 409–412.

[119] Schoutens, W. (2003). *Lévy Processes in Finance.* Wiley, New York.

[120] Seshadri, V. (1999). *The inverse Gaussian distribution. Statistical theory and applications.* Springer-Verlag, New York.

[121] Shimura, T. and Watanabe, T. (2005). Infinite divisibility and generalized subexponentiality. *Bernoulli* **11**, 445–469.

[122] Stahl, (1997). Three Cheers. *Risk* **10**, 67–69.

[123] Talor, J.W. (2008). Estimating value at risk and expected shortfall using expectiles. *Journal of Financial Economics* **9**, 1–22.

[124] Teugels, J.L. (1975). The class of subexponential distributions. *Ann. Probab.* **3**, 1000–1011.

[125] The Basel Committee on Banking Supervision (1988). Base realignments and closures. Basel Committee.

[126] The Basel Committee on Banking Supervision (1996). Amendment to the capital accord to incorporate market risk. Basel Committee.

[127] The Basel Committee on Banking Supervision (2009). Strengthening the resilience of the banking sector. Basel Committee.

[128] Tian, M. and Chan, N.H. (2010). Saddle point approximation and volatility estimation of value-at-risk. *Statistica Sinica* **20**, 1239–1256.

[129] Ventor, J.H. and Jongh, P.J. (2002). Risk estimation using the Normal Inverse Gaussian distribution. *Journal of Risk* **4**. 1–23.

[130] Wang, Y. and Wang, K. (2011). Random walks with non-convolution equivalent increments and their applications. *Journal of Mathematical Analysis and Applications* **374**, 88–105.

[131] Watanabe, T. (2008). Convolution equivalence and distributions of random sums. *Probab. Theory Related Fields* **142**, 367–397.

[132] Watanabe, T. and Yamamuro, K.(2010a). Ratio of the tail of an infinitely divisible distribution on the line to that of its lévy measure. *Electronic Journal of Probability* **15**, 44–74.

[133] Watanabe, T. and Yamamuro, K.(2010b). Local subexponentiality and sel-decomposability. *J. Theor. Probab.* **23-4**, 1039–1067.

[134] Yu, C. , Wang, Y. and Yang, Y. (2010). The closure of the convolution equivalent distribution class under convolution roots with applications to random sums. *Statistics and Probability Letters* **80**, 462–472.

[135] Yu, P.L., Li, W.K. and Jin, S. (2010). On some models for Value-at-Risk. *Econometric Reviews* **29(5-6)**, 622–641.

[136] Zakoian, J. (1994). Threshold heteroskedastic functions. *Journal of Economic Dynamics and Control* **18**, 931–955.