



中国股市相依结构 的统计研究

孙志宾 编著

中国建材工业出版社

中国股市相依结构的统计研究

孙志宾 编著

中国建材工业出版社

图书在版编目 (CIP) 数据

中国股市相依结构的统计研究/孙志宾编著.
—北京: 中国建材工业出版社, 2009.12

ISBN 978-7-80227-582-9

I. ①中… II. ①孙… III. ①股票-资本市场-研究-中国
IV. F832.51

中国版本图书馆 CIP 数据核字 (2009) 第 220088 号

内 容 简 介

本书研究的主要问题是中國股市相依结构的统计研究。研究的主要目的是研究中国股市股票收益率相依结构的 Copula 模型理论, 利用非参数估计和极大似然估计方法, 将图形分析方法、拟合优度检验、蒙特卡洛方法、EM 算法等引入模型的估计中, 提高了模型参数估计的精度, 增强了估计的稳健性。

中国股市相依结构的统计研究

孙志宾 编著

出版发行: 中国建材工业出版社

地 址: 北京市西城区车公庄大街 6 号

邮 编: 100044

经 销: 全国各地新华书店

印 刷: 石油工业出版社印刷厂

开 本: 787mm×1092mm 1/16

印 张: 8.75

字 数 224 千字

版 次: 2009 年 12 月第 1 版

印 次: 2009 年 12 月第 1 次

书 号: ISBN 978-7-80227-582-9

定 价: 20.00 元

本社网址: www.jccbs.com.cn

本书如出现印装质量问题, 由我社发行部负责调换。联系电话: (010) 88386906

前 言

在市场经济和政策调控条件下，股票市场经常出现大起大落现象，股票价格的剧烈波动是股票市场最显著的特征之一，因此作为监管机构对股市中的波动更为关心，一类股票指数在市场波动过程中对另一类股票指数是否有较大的影响，如果有它是如何来影响的。股市中一只股票的大起大落会不会影响到另外一支相关联的股票的波动？一只股票的波动会不会对整个股市的波动有影响，如果有，到底会有什么样的影响？影响会有多大？要回答这些问题，就必须对股市相依结构做深入的分析研究，从这个意义上讲，研究股市相依结构具有重要的理论和现实意义。

在国际上有关 Copula 理论的应用近年来得到迅速发展，但是在国内的研究才刚刚起步。从目前国际有关 Copula 理论的应用发展动态来看，虽然 Copula 理论在金融和保险市场有广泛的应用，但是在股票市场上没有一个完整的分析股市相依结构的框架，本书正是在基于这种分析的基础上，结合中国新兴股票市场刚刚起步且很多地方还未完善的特点，试图构建中国股票的相依结构的研究框架。本书针对中国股市的特点，详细、系统地梳理了在中国股市中测定、选择和模拟 Copula 函数的方法，同时也给出了对股市的对数收益率如何使用不同的二元 Copulas 来生成有效的蒙特卡洛模拟，最后则给出了度量股票相依风险的方法。

本书主要面向高校教师、科技人员、在校研究生和本科生，也可以供金融、统计、数学等专业技术人员参考。

由于本人水平有限，加之时间仓促，书中错误和不妥之处在所难免，恳切希望读者不吝指正。

孙志宾

2009.12

目 录

第一章 引言	1
§1.1 股市相依结构的研究意义	1
§1.2 论文的主要研究问题及相应研究现状	3
§1.3 本书研究的主要内容	6
§1.4 本书研究的主要创新	7
第二章 中国股市相依结构的理论度量	9
§2.1 Copula 理论及相依性度量	9
§2.1.1 Copula 的定义及基本特征	9
§2.1.2 Copula 的重要性质	11
§2.1.3 Copula 的估计	19
§2.2 常见的Copulas族及其性质	21
§2.2.1 椭圆Copulas	21
§2.2.2 Archimedean Copulas	22
§2.3 Copula的测定——中国股市相依结构的实证分析	25
§2.3.1 给定Copula的参数估计	25
§2.3.2 Copula 函数的恰当选择	27
§2.3.3 模拟算法	28
§2.3.4 计算及模拟结果	30
§2.4 本章小结	34
第三章 中国股市收益率的相依结构的研究(I)	35
§3.1 描述股市收益率的几种分布函数	36
§3.1.1 正态分布模型	36
§3.1.2 稳定paretian分布模型	36
§3.1.3 混合分布模型	37
§3.1.4 广义分布模型	38
§3.2 收益率分布函数的确定与估计	40
§3.2.1 收益率分布函数的估计方法	40
§3.2.2 分布函数参数的估计	41
§3.2.3 分布函数的确定	43
§3.3 对中国股市收益率数据的Copulas函数的拟合	44
§3.3.1 识别Copulas函数的方法	44
§3.3.2 识别Copulas函数的图形分析方法	47
§3.3.3 确定最优Copulas函数的方法	50
§3.3.4 确定最优Copula函数时可能遇到的问题	53
§3.4 中国股市主要股票指数的实证分析	54
§3.4.1 α 的估计	54
§3.4.2 上证指数与深证指数的实证分析	54
§3.4.3 深圳成分A股与深圳成股B股指数的实证分析	59
§3.4.4 商业指数与地产指数的实证分析	63
§3.4.5 工业指数与地产指数的实证分析	67
§3.4.6 工业指数与商业指数的实证分析	71

§3.4.7 工业指数与上证30指数的实证分析	76
§3.5 Copula图形分析应用及蒙特卡洛模拟	81
§3.5.1 Archimedean Copula 里的第1、5、13族的密度曲线和等高线图	81
§3.5.2 实证分析综合小结	83
§3.5.3 蒙特卡洛模拟	84
§3.5.4 Copula 在投资组合风险中的应用	85
§3.6 本章小结	87
第四章 中国股市收益率的相依结构研究 (II)	89
§4.1 描述中国股市收益率的相依结构的混合Copula模型	90
§4.2 混合Copula模型的估计方法	91
§4.2.1 EM算法简介	91
§4.2.2 EM算法应用	92
§4.3 实证分析: 中国股市主要股票指数收益率的相依结构	95
§4.4 本章小结	102
第五章 中国股市股价相依风险的测度研究	104
§5.1 金融风险的管理标准—VaR及CVaR	105
§5.1.1 一致性风险度量	105
§5.1.2 传统的风险度量指标	105
§5.1.3 VaR与CVaR概述	107
§5.2 相依风险函数VaR测度的最优界	110
§5.2.1 相依风险函数VaR测度的简介	110
§5.2.2 两个基本定义	111
§5.2.3 相依风险函数的分布界	112
§5.2.4 相依风险函数的分布界的计算方法与步骤	115
§5.3 相依风险函数尾部相依的风险度量	118
§5.3.1 尾部相关性的定义	118
§5.3.2 尾部相依系数的计算	119
§5.3.2.1 Archimedean Copulas 分布族的尾部相依系数的计算	119
§5.3.2.2 椭圆Copulas分布族的尾部相依系数的计算	120
§5.3.3 尾部相依系数(TDC)的估计	121
§5.4 中国股市相依风险的实证分析	122
§5.4.1 中国股市资产投资组合的选择	122
§5.4.2 中国股市中Copula模型在投资组合风险管理方面的应用	123
§5.4.3 中国股市收益率的尾部相依系数的计算	123
§5.4.4 中国股市股票最佳投资组合实施的程序	126
§5.5 本章小结	127
参考文献	128

第一章 引言

在本章中，我们将首先剖析研究股市相依结构的理论与现实意义，再结合这个领域已有的国内外研究成果，阐述本书的研究问题，最后给出本书每章的主要研究内容和本书研究创新之处。

§1.1 中国股市相依结构的研究意义

经过十多年的发展，我国的证券市场已经取得了举世瞩目的成就，证券市场与经济生活的联系越来越紧密，在国民经济中有着举足轻重的地位，但其中存在的一些问题也日益受到人们的关注，特别是加入 WTO 大家庭之后，对于市场的进一步规范，从而防止泡沫的进一步扩大与市场风险的发生都有着更为迫切的需求。国际金融大师索罗斯的哲学观点：如果发生了甲事件，那么乙事件也将随之发生，而丙事件又将因乙事件而产生。即世界是普遍联系的世界。本文受这种观点的启发，主要研究中国股市的相依结构问题。

在市场经济条件下，股票市场经常出现大起大落现象，股票价格的剧烈波动是股票市场最显著的特征之一，因此作为监管机构对股市中的波动更为关心，一类股票指数在市场波动过程中对另一类股票指数是否有较大的影响，如果有它是如何来影响的。股市中一只股票的大起大落会不会影响到另外一只相关联的股票的波动？一只股票的波动会不会对整个股市的波动有影响，如果有，到底会有什么样的影响？影响会有多大，要回答这些问题，就必须对股市相依结构做深入的分析研究，从这个意义上讲，研究股市相依结构具有重要的理论和现实意义。

一、理论意义

以往人们对金融市场中两（多）个资产的相关性研究主要使用经典的 Pearson 线性相关系数来刻画，近来人们渐渐认识到它的不足，第一，经典的线性相关系数的本质是假定两个资产之间的联合分布为正态分布，而实践证明金融市场上资产的联合分布往往不是正态分布。第二，经典的相关系数是从线性的角度去刻画两（多）个资产之间的关系，而实际中的数据表明金融资产之间的相依结构往往是非线性关系的。第三，经典的相关系数从平均的角度来度量资产之间的关联程度。而在现实生活中我们更关心资产的极端值，即金融资产出现剧烈波动的过程。基于这三个方面原因，一种全新的度量相关性的工具 Copula 也随之产生。Copula 一词可以译成“相依函数”或“连接函数”，它是把多维随机变量的联合分布与其一维边际分布连接起来的函数，从理论上说，（Sklar 定理）它可以准确地来刻画金融市场中两（多）个金融资产的相依结构，这使我们对变量之间，相关性的认识达到了一个新的高度。

另外，从 Copula 出发衍生出来的一系列相关性指标，例如 Kendall τ ，Spearman ρ ，以及尾部相关性指标等，都精确地刻画了金融市场中资产之间的相依性度量，这些理论的发展使我们对金融市场中资产相依关系的认识有着重要的理论意义。

另外，Copula 出现，也为金融风险分析提出了一个新的理论课题，我们知道，金融市场上的风险是复杂的，金融风险分析技术的一个重要的方面，就是将风险进行分解，例如一支股票的风险可以分解为市场风险、板块风险、个股风险与随机扰动风险，市场风险是由在股市上交易的全部股票价格波动形成的，因此它是由联合分布来决定的，我们如何刻画、分析市场上这些联合或相依风险，这是一个很重要的理论问题，随着这些问题的解决，对我们正解地认识金融市场从而更为有效地规避风险，具有重要的理论意义。

二、现实意义

在中国证券市场上，比如中国股市中，一只股票的大起在落会不会影响其它股票的波动，进而会不会影响到整个股市的波动，如要有会有多大的影响？要回答这些问题，就必须考虑中国股市各种股票之间的相依结构的问题，而这些相依结构的刻画，就必然用到 Copula 及其衍生出来的一系列相依性度量的指标，因为前面我们说过，经典的 Pearson 线性相关系数用来刻画各种股票（指数）之间的相关性有它致命的缺点：假定股票的边际分布为正态，或者从平均的和线性相关的角度来考虑股票之间的相关，而实际上我们知道，金融资产（股票）的收益数据往往是厚尾的，它们之间有时根本没有线性关系（或表现为弱相关），而在股市上我们更关心股票的剧烈波动（大起大落）阶段，这时若用经典的 Pearson 相关系数来度量实际上没有什么意义，因此，我们需要寻找一种新的度量股市相依结构的工具— Copula. Copula 的出现对于我们精确地认识股市相依结构，具有最直接的现实意义。

研究股市相依结构的另一个重要的现实意义体现在风险管理领域。直观地说，风险产生于现实对未来预期的目标收益的偏离，特别是向下偏离（downside risk），风险管理就是权衡目标收益与相应的风险之间的关系并采取必要的行动计划（包括放弃投资）。而金融市场中的风险是复杂的，有时是相互依存的，金融风险分析技术的一个重要方面，就是将风险进行分解，如前所述，一支股票的风险可以分解为市场风险、板块风险、个股风险与随机扰动风险，因此我们可以把市场风险看成联合分布，而把其它的风险看成边际分布。通过 Copula 函数，我们可以把复杂的市场风险分解为容易控制的边际风险，另外，两（多）支股票的联合（相依）风险度量也是股市里经常遇到的问题，例如，我们购买几支股票的投资组合，我们用常用的风险度量工具 VaR 来度量，则这几只股票的联合风险（或相依风险）如何来度量，要回答这些问题就必须用到 Copula 函数，因此可以说，在金融风险管理过程中引入 Copula 函数，以及 Copula 函数引入到金融风险的分解，对我们更好地认识股市，更有效的在股市中进行风险管理，具有重要的现实意义。

§1.2 论文的主要研究问题及相应的研究现状。

从上一节的研究意义中可以看出,相依结构的研究是一个非常宽泛和开放的研究领域,我们这里主要从 Copula 函数出发来研究相依结构,这样既可以展现本书研究主要内容的内在逻辑关系,也可以显示本书研究的创新点及其意义。Copula 函数是本书研究的中心内容,贯穿于全书的研究始末, Copula 是把多维随机变量的联合分布用其一维边际分布连接起的函数,最早 Copula 理论源于 1959 年的 Sklar 定理,它的主要含义是随机变量的多维分布函数可以用其连续的一维边际分布函数来唯一表达。

一. Copula 的国内外研究现状

1. Copula 的国际研究现状

在国际上, Copula 的研究现状主要基于下面几个方面:

(1).Copula 在金融市场上是一个有力的工具。一般地,统计模拟问题能被分成两部分:第一是边际分布的鉴别;第二是为了更好地表达相依结构来定义恰当的 Copula. 人们以前普遍采用正态分布来模拟资产收益的问题,因为正态分布容易计算,但是资产收益往往是厚尾的,这样用正态分布的假定就不太合适了,另外,基于风险度量的各种金融资产分配问题也越来越重要,例如银行资产,这主要包括信用风险、市场风险和利率风险,而不同风险的联合分布的模拟是很重要的问题。这两个难题(正态分布的假定问题与联合分布的模拟)能通过 Copula 函数来解决,在 1999 年以前, Copula 理论在金融中还几乎没有应用,而 1999 年以后,随着人们对金融市场的认识, Copula 理论在金融中的应用如雨后春笋,很快应用到了金融市场的各个研究领域,具体文献参看 Rockinger, M. and E. Jondeau (2002).

(2).Copula 在风险管理、投资组合、资产定价等方面的应用。

我们构建确定的 Copula 模型用来描述 Copula 的各种相关结构,这一方法被广泛地用在金融市场上的风险管理、投资组合的选择、资产定价等方面。具体文献参看 Embrechts, P., A. Mcneil and D. Straumann (2002), Forbes K. and R. Rigobon (2002). 近来信用衍生产品成为转移和规避信用风险的主要工具,信用衍生产品市场飞速发展,我们应用 Copula 模型来模拟和对信用衍生产品定价,其模型的相依结构采用 t-Copula. 对于停时风险模型中动态停时相关的刻画主要用停时 Copula 来刻画,具体文献参看 Glasserman, P. Heidelberger, and P. Shahabuddin, 2002. 其中提出一套模拟复合相依风险的工具,各种相依结构是由 Copula 模型来生成的。这些相依结构通常由联合分布函数和联合特征函数来鉴别。 Copula 技术还可以定价不易计算的期权及衍生产品, Copula 函数能把期权和部分停时分成特定的边际分布与相依结构两部分,此风险能用 Copula 函数来计算,这可以推广应用到完全市场模型,具体文献参看 Fermanian, J-D. and Scaillet, O. (2002).

在金融风险管理中刻画金融资产收益的联合分布是一个很重要的问题,一般说金融资产的收益的分布都是厚尾分布,其联合分布函数可以通过 Copula

函数来刻画。对一个特定组合的风险 (VaR) 值, 可以从合适的联合分布通过使用 Monte Carlo 模拟来计算。用 Copula 的方法构建数据的相依模型是要经过选择的, 当相依结构固定时, 风险越大的资产是否会导致风险大的投资组合, 回答是否定的, 这方面的文章可以参看 Luisa Tibiletti(2002)。另外, 在模拟信用事件的到达时, 总是假定指数到达时间, 由 Copula 技术就有许多不同的方法来构建联合信用事件时间。指出多元指数 Copula (从限制多元二维模型的观点上) 是一个合适的选择。Copula 函数的方法也可以用来鉴别生存时间的联合分布, 可以参看 Li, D. X. (2000)。

对于极端事件期间的投资组合的风险管理的方法是把多元极值分布分成两部分, 即每个资产的边际分布和它们的相依函数, 同时我们可以把压力试验应用到市场指数的投资组合中。这些问题都可以通过 Copula 函数来解决。另外, 用 Copula 来构建多元生存模型, 生存 Copula 的性质和相依性度量考虑了对冲风险的问题, 同时给出了信用生存时间对信用投资组合的风险度量 and 信用衍生产品的定价。主要参考文献是 Genest, C., van der Goorbergh, R. and Werker, B. (2003), Di Clemente, A., C. Romano (2003b)。

(3). 关于 Copula 在金融时序中模型的相关性度量问题。在金融时序等计量经济模型中我们到底采用什么样的相依结构, 这是一个比较新的研究领域, 主要参考文献是 Patton, A.J. (2002a), Patton, A.J. (2002b), Patton, A.J. (2004), Patton, A.J. (2004)。当一个新的计量经济工具出现时就会出现两个重要问题是: 第一 “我们怎么使用它” 第二 “它有什么好处”。在文献 [12] 回答了第一个问题, 这一篇里提供了一个 Copula 模型的两阶段最大似然估计量, 在每个边际分布里允许有不等的参数, 文章 Patton, A.J. (2002b)^[13], Patton, A.J. (2004)。提供了一个估计量的应用, 证明如何来生成、估计和计算时间序列 Copula 模型, 如何使用一个 GARCH 模型鉴别对时间变化条件的相依, 不对称相关不可能象线性相关那样用经典的线性相关系数来度量, 通过运用非线性相依 — Copula 这个新工具, 发展了时间序列在经济学里如汇率市场和股票市场应用的理论。在文章 Patton, A.J. (2004)^[15] 中刻画了股票市场表现出在下降阶段比上升阶段具有更强的相关性。

2. Copula 的国内研究现状

在国内, Copula 的研究主要基于下面几篇文章。国内第一篇关于 Copula 的文章是张尧庭 2002 年发表在统计研究第四期上的《连接函数 (Copula) 技术与金融风险分析》。文章系统地介绍了 Copula 的含义, 性质以及在金融风险分析中的应用。文章明确指出: Copula 是把简单的 (单个的) 与复杂的 (系统的) 风险连接起来的工具。比如在股市中 Copula 可以描述两只 (多只) 股票 (指数) 的联合分布, 进而可以计算它们之间的联合风险。国内第二篇关于 Copula 的文章是张尧庭 2002 年发表在统计研究第九期上的《我们应该选用什么样的相关性指标》。文章指出经典的 Pearson 线性相关系数用来刻画各种金融资产之间的相关性有它致命的缺点: 不仅从平均的角度而且从线性相关的角度来考虑金融资产之间的相关, 而实际上我们知道, 金融资产的收益数据往往是厚尾的, 它们之间有时根本没有线性关系 (或表现为弱

相关),例如在股市上我们更关心股票(指数)的剧烈波动(大起大落)阶段,这时若来用经典的 Pearson 线性相关系数来度量实际上没有什么意义,因此,我们需要寻找一种新的度量股市相依结构的工具— Copula. 文章同时指出从 Copula 函数出发引出的尾部相关性及变化协调引出的相关性等统计指标比经典的 Pearson 线性相关系数更好地刻画了金融资产之间的相关性度量. 国内第三篇关于 Copula 的文章是崔嵬、张尧庭等 2003 年发表在统计研究第六期上的《如何选择度量金融风险的指标》. 文章指出在计算非线性相关投资组合资产的风险值(例如 VaR)时,究竟采用什么样的指标更合适以及传统计算 VaR 的方法在现实的金融市场上实际上有什么不足. 明确指出 Copula 函数的方法在金融市场上风险的计算要优于其它的方法.

综合上述,在国际上有关 Copula 理论的应用近年来得到迅速发展,但是在国内的研究才刚刚起步. 从近年国际有关 Copula 的应用发展动态来看,虽然 Copula 理论在金融与保险市场有广泛的应用,但是在股票市场上没有一个完整的分析股市相依结构的框架,本书正是在基于这种分析的基础上,结合中国新兴股票市场刚刚起步且很多地方还有完善的特点,试图构建中国股市的相依结构的研究框架.

二. Copula 的四个问题

设 x_1^* , x_2^* 表示两个风险资产的收益, $H^*(x_1^*, x_2^*)$ 表示资产 x_1^* , x_2^* 的联合分布. 我们有以下四个问题.

(1) 已知两种资产收益的边际分布,如何去找到它们的联合分布,也就是其相依结构.

(2) 当 x_1^* , x_2^* 给定时,如何确定 $H^*(x_1^*, x_2^*)$ 的值.

(3) 当给定联合概率值 $H^*(x_1^*, x_2^*) = \gamma$, 我们如何来确定 (x_1^*, x_2^*) .

(4) 对给定的 u , 如何计算尾部相关系数 $P(x_1^* > u | x_2^* > v)$.

我们的研究始终围绕着这四个问题来展开. 以上四个问题可以通过 Copula 函数来解决,系统介绍 Copula 的著作可以参考 Nelsen, R. B. 《An Introduction to Copulas》, 在此文献中,作者给出了各种各样的 Copulas 来描述二(多)元资产的相依结构,其中包括椭圆 Copula 族,阿基米德 Copula 族等,这些 Copula 族很好在刻画了金融市场上尖峰厚尾特性的数据的相依结构.

三、本书主要研究的四个问题与相应的研究现状

从上面的分析可以看出,针对金融市场相依结构的 Copula 分析是个非常宽阔的研究领域,本文着重就上述的四个要点,结合中国股票市场,研究下面四个方面的问题. 应该说这四个问题是这个研究领域的核心内容,可以构成一个较为完整的框架体系,也是目前该领域的研究热点.

1. 在股市数据的边际分布假定为正态时,如何在中国股市测定、选择和模拟最适合已知数据的 Copula 函数. 最早研究股市的相依结构的文章是

Claudio Romano (2002), 他在意大利股票市场上首次使用 Copula 函数来描述不同股票 (指数) 间的相依结构, 不过他是在正态边际的假设下, 给出了不同类型的 Copula 函数来描述股票的相依结构。

本书针对中国股市的特点, 详细系统地梳理了在中国股市中测定、选择和模拟 Copula 函数的方法, 同时也给出了对股市中给定的对数收益率数据如何使用不同的二元 Copulas 函数来生成有效的蒙特卡洛模拟。

2. 对股市数据的边际分布不做正态假定时, 针对中国股市的特点如何在股市中测定、选择和模拟最适合中国股市数据的 Copula 。

当我们将中国股市股票数据不做正态假定时, 我们采用 Genest 和 Rivest (1993) 提出的非参数方法和极大似然的方法等实现了最优 Copula 函数的选取。

3. 构造了一个混合 Copula 模型 (作者将它命名为 A-1-5-13 混合 Copula 模型) 来刻画中国股市的相依结构, 并且给出了这种多参数模型的具体估计和检验方法。

通过对中国股市的实证分析, 我们得到了 Archimedean Copula 族里的第 1, 5, 13 族都可以用来描述中国股市的相依结构, 为了更好地刻画中国股市的相依结构, 我们构造了 A-1-5-13 混合 Copula 模型来刻画中国股市的相依结构, 并且给出了这种多参数模型的具体估计和检验方法。

4. 如何分析计算中国股市的相依风险。

我们系统归纳了计算股市的投资组合相依风险的最优界及具体计算相依风险的一种方法, 同时也总结出描述股市相依结构的 Copula 函数的尾部相依系数的计算和估计方法, 这使得我们对股市中风险的度量有了更新的认识。这样我们不仅可以计算单个股票的风险, 同时还可以计算股票的相依风险, 达到同时对两组 (只) 股票进行控制的风险管理, 在此基础上, 给出了中国股市股票最佳投资组合实施的方案和程序。

§1.3 本书研究的主要内容:

本书将分成五章, 第一章即为文章导论部分, 第二、三、四章由简单到复杂地研究了中国股市的相依结构, 第五章给出了如何计算和分析中国股市的相依风险, 具体内容安排如下:

第二章的主要内容集中在三个方面:

一、系统地介绍了 Copula 理论, 其中包括 Copula 的定义、性质及常见的 Copula 族和 Copula 参数的一些估计方法。

二、在边际分布假定为正态前提下, 具体给出了测定、选择和模拟 Copulas 函数的程序, 其中包括 Copula 的参数估计及 Monte Carlo 模拟。

三、基于中国股市商业指数和工业指数的相依结构分析, 在边际为正态分布的假定下我们得到了正态 Copula 是描述商业指数和工业指数的最优 Copula。

第三章的工作集中在两个方面：

一、在边际分布不做正态假定的前提下给出了具体选择最优 Archimedean Copula 的三个步骤：

（一）参数估计方法

（二）图形分析方法

（三）确定最优 Copula 函数的方法—拟合优度检验。

二、通过对中国股市主要股票指数的相依结构的实证分析，得出了中国股市主要股指的相依结构，最后通过 Monte Carlo 模拟，说明得到的结果是很好的。

第四章的工作集中在两个方面：

一、按照一定的比例权重，把第三章得到的三个能较好的描述中国股市相依结构的 Copula 函数做一线性组合来刻画中国股市收益率的相依结构。且利用 EM 算法给出了混合 Copula 的参数极大似然估计。

二、通过对中国股市大盘股与小盘股的投资组合之间及主要股票指数之间的相依结构分析，我们给出了混合 Copula 的参数估计值，得到了投资组合的大、小盘股及主要股票指数之间的相依结构。

第五章的主要工作集中在三个方面：

一、简单介绍了金融风险的管理标准，

二、系统归纳了相依风险函数 VaR 度量的最优界，同时得到了中国股市不同指数或两组股票的投资组合之间的相依风险值，

三、研究了中国股市中二元资产尾部相关的一些性质，同时给出了计算和估计尾部相关系数的方法。

四、给出了中国股市股票最佳投资组合实施的程序。

§1.4 本书研究的主要创新

本书将着重从统计角度对中国股市相依结构的研究，力图建立一个较为完整且便于付诸实现的 Copula 分析、研究和实施的框架，在这个框架下对中国股市进比较全面的实证分析，并对一些技术性环节给出了一些新的解决方法，有关在国际上将 Copula 用于分析股市相依结构的文章所见不多，而在国内有关方面的具体研究尚未见到。本书很多理论和方法都是初次拿到中国股票市场来分析，同时也得到了一些不错的结果，但毕竟还有很多东西需要完善，这使得我们能更好地认识中国股市，对中国股市进行风险管理打下了良好的基础。具体而言，本书的主要技术性创新主要体现在以下四个方面。

一、本书在中国股票市场上提出了一个完整地分析中国股市相依结构进而用于风险管理的研究框架。

本书始终围绕这一框架—中国股市相依结构的研究来展开，并且通过实证分析得出对这一框架的研究结果，将对中国股市发展具有参考价值。这是本文最大的创新点。

二、系统归纳和总结了中国股市测定选择和模拟最优适合股票数据的 Copula 函数方法和程序.

首先在边际分布为正态假定下, 我们系统归纳了中国股市测定选择和模拟最优适合股票数据的 Copula 函数, 实证分析得出了中国股市商业指数与工业指数的最优 Copula 函数为正态 Copula, 这主要是由于假定边际分布为正态造成的。其次不考虑股市数据的边际分布, 在 Archimedean Copula 族里给出了在股市中测定、选择和模拟最优 Copula 函数的方法和程序, 实证分析得出第 1、5、13 族都是适合中国股市数据的 Copula 函数。一般来说, 当股票相依结构较强时, 13 族是适合中国股市的最优 Copula 函数, 最后通过蒙特卡洛模拟证明这种选取最优 Copula 函数的方法是正确可行的。

三、构造了 A-1-5-13 混合 Copula 模型来刻画中国股市的相依结构, 并且给出了这种多参数模型的具体估计和检验方法.

实证分析表明: 本人构造的这个混合 Copula 模型更精确地刻画了适合中国新兴股票市场特点的相依结构, 同时第四章也给出了一种多参数模型的估计方法—EM 算法, 最后通过检验和蒙特卡洛模拟, 证明 A-1-5-13 混合 Copula 的确能更精确地描述中国股市的相依结构。

四、给出了中国股市股票最佳投资组合实施的方案和程序

在得出中国股市的相依结构后, 我们可以计算其相依风险, 同时也得到了股市上二元资产的尾部相依系数的计算与估计, 这使得我们对股市中风险的度量有了更新的认识。在此基础上, 给出了中国股市股票最佳投资组合实施的方案和程序。实证分析表明: 我们给出的股票最佳投资组合实施的方案和程序在理论上正确, 在实际操作过程中可行。

第二章 相依结构的理论度量及对中国股市的应用

本章简要介绍了 Copulas 理论以及在正态边际分布下给出了在中国股市中测定、选择和模拟 Copulas 函数的方法，同时也给出了对股市的对数收益率如何使用不同的二元 Copulas 生成有效的蒙特卡洛模拟，本章可以分为三个部分，第一部分介绍 Copulas 理论及相依性质，主要介绍了 Copulas 的定义及 Copulas 的一些性质，随机变量的 Copulas 对其严格增变换 Copulas 具有不变性，这是一条很有用的性质，例如 Copulas 函数 $C(u_1, \dots, u_n)$ 反映了市场的结构，它与股票价格的度量单位无关，变换只是一个线性变换 $ax + b$ ，且 $a > 0$ ，它是严格单调增的变换。另外由 Copulas 函数 $C(u_1, \dots, u_n)$ 导出的相关性指标比经典的相关系数更加合乎人们的要求，因为经典的线性相关系数实际上是线性变换下不变的一种相关性指标，涉及到非线性函数的相关性，就会导出错误的结论。且经典相关系数是从平均的角度去度量相关性，这些性质都反映了经典相关系数对相关性度量的不足之处。而 Copulas 函数真正地抓住了两（多）个变量之间的相依结构，具有很好的性质。

第二部分介绍常见的 Copulas 族及其性质，主要介绍都椭圆 Copulas 族，Archimedean Copulas 族及其一些性质。

第三部分给出了 Copulas 的测定、选择和模拟 Copulas 函数的方法。在边际分布为正态分布的假定下，我们给出了中国股市的工业指数和商业指数上证指数和深证指数的相依结构的实证分析。首先给出 Copulas 的参数估计，其次要选择最优的 Copulas，最后给出了模拟 Copulas 的算法，通过上述方法和步骤我们得到最优的 Copulas 函数是正态 Copulas 函数，这主要是因为假定两只股票（指数）的边际分布为正态分布，而在中国股市实际的相依结构的选取过程中对边缘分布的选取是很重要的。

§2.1 Copula 理论及相依性度量

§2.1.1 Copula 的定义及基本特征

我们首先给出 Copula 的定义。

定义 2.1 n 维 Copula 是具有下面性质的函数 C ：

- (1) $\text{Dom}C = I^n = [0, 1]^n$;
- (2) C 是有基底的 (grounded) 和 n - 增的;
- (3) C 的边际 C_n 满足: $C_n(u) = (1, \dots, 1, u, 1, \dots, 1) = u$, 对一切 $u \in I$.

简单来说，Copula 是把多维随机变量的联合分布与其一维边际分布联系起来的函数，通常称为连接函数。假设我们有 n 个资产的投资组合，各个资产收益的边际分布函数分别为 $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$ ，其联合分布函数为 $F(x_1, x_2, \dots, x_n)$ 。在实际中，通常边际分布是容易得到的，而联合分布则

很难给出。著名的 Sklar 定理将多维随机变量的联合分布与边际分布联系起来。

定理 2.1(Sklar 定理) 令 F 为 n 维分布函数, 其连续边际分布为 $F_1, \dots, F_n$, 则存在函数 C 有下面唯一的 Copular 表达式:

$$F(x_1, \dots, x_N) = C(F_1(x_1), \dots, F_n(x_n)) \quad (2.1)$$

根据 Sklar 定理中的 (2.1) 式, 我们就可以通过 Copula 函数 C 由边际分布来构建多元联合分布, 在股市中就是去刻画多个风险资产的联合分布。很自然地想到用 Monte carlo 来模拟相依数据, 而 Copula 理论在模拟过程中起重要作用。

为了说明用 Copula 构建多元联合分布的方法, 我们考虑边际分布和 Copula 函数都是已知的简单情形, 有关边际分布和 Copula 函数的估计问题在后面的章节将作深入讨论。

显然, 任意边际分布函数 $F_i(x_i) (i = 1, 2, \dots, n)$ 值都可以看作是 $I = [0, 1]$ 上均匀分布随机变量 U_i 的取值。 F^{-1} 记作 F 的伪逆 (pseudo-inverse) 函数。例如有: $x = F^{-1}(u) = \sup \{x | F(x) \leq u\}$ 。假定边际分布 $F_1, \dots, F_n$ 是连续的, 那么由 Sklar 定理可知, 存在唯一的 Copula 函数 C , 使:

$$C(u_1, u_2, \dots, u_n) = F(x_1, x_2, \dots, x_n) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_n^{-1}(u_n))$$

因此, 只需独立地模拟 $[0, 1]$ 均匀随机变量 $U_i, i = 1, \dots, n$, 而 $(F_1^{-1}(U_1), F_2^{-1}(U_2), \dots, F_n^{-1}(U_n))'$ 的联合分布就是 F 。在上述拟合过程中, $[0, 1]$ 上均匀分布 U_i 的随机模拟是很容易实现的。

一个可行的办法是采用条件分布的方法来模拟。考虑 $n > 2$ 的情形, 用

$$C_i(u_1, \dots, u_i) = C(u_1, \dots, u_i, 1, \dots, 1), \quad i = 1, \dots, n$$

表示 $C(u_1, \dots, u_n)$ 的第 i 维边际分布, 记 $C_1(u_1) = u_1, C_n(u_1, \dots, u_n) = C(u_1, \dots, u_n)$ 。假定 $(U_1, \dots, U_n)' \sim C(u_1, \dots, u_n)$, 则给定前 $i-1$ 个分量时 U_i 的条件分布为

$$\begin{aligned} C_i(u_i | u_1, \dots, u_{i-1}) &= P[U_i \leq u_i | U_1 = u_1, \dots, U_{i-1} = u_{i-1}] \\ &= \frac{\partial^{i-1} C_i(u_1, \dots, u_i)}{\partial u_1 \dots \partial u_{i-1}} \bigg/ \frac{\partial^{i-1} C_{i-1}(u_1, \dots, u_{i-1})}{\partial u_1 \dots \partial u_{i-1}} \end{aligned}$$

上述条件分布可以从 $i = 1$ 开始计算, 递推得到 $C_i(u_i | u_1, \dots, u_{i-1}), i = 2, \dots, n$, 注意 $i = 1$ 时就是非条件分布 $C_1(u_1) = u_1$ 。具体模拟步骤如下:

- (1) 从 $U(0, 1)$ 中模拟一个值 u_1 ;
- (2) 从 $C_2(u_2 | u_1)$ 中模拟一个值 u_2 ;
- (3) 从 $C_3(u_3 | u_1, u_2)$ 中模拟一个值 u_3 ;

⋮

(n) 从 $C_n(u_n|u_1, \dots, u_{n-1})$ 中模拟一个值 u_n .

最后得到的 $(u_1, u_2, \dots, u_n)'$ 便具有联合分布 C 。

对于一个给定的多元分布，存在与其边际分布相联系的 Copula 函数，而且，如果边际分布函数连续，则 Copula 函数唯一 (Sklar 1959)。例如，对于 $\delta > 0$ 考虑联合分布

$$F(x, y) = \exp \left\{ - \left[\exp(-x) + \exp(-y) - (\exp(\delta x) + \exp(\delta y))^{-\frac{1}{\delta}} \right] \right\},$$

$-\infty < x, y < +\infty$. 令 $y \rightarrow \infty, x \rightarrow \infty$ ，得到边际分布为：

$$F_1(x) = \exp(-\exp(-x)), \quad F_2(y) = \exp(-\exp(-y))$$

令 $u = F_1(x), v = F_2(y)$ ，则其 Copula 为

$$C(u, v) = uv \exp \left\{ [(-\log u)^{-\delta} + (-\log v)^{-\delta}]^{-\frac{1}{\delta}} \right\}.$$

若我们考察股市中的某一投资组合，对于每个资产我们可以确定它的边际分布，然后由边际分布去估计 Copula 函数，Copula 函数包含了这些资产之间的相关信息，由此我们确定了其联合分布，因此就可以评估投资组合的联合风险。

§2.1.2 Copula 的重要性质

1. 不变性 设 $(Z_1, \dots, Z_n)'$ 为具有连续分布的随机向量，其 Copula 函数为 C ，而 $T_1, \dots, T_n$ 为严格增函数，则 $(T_1(Z_1), \dots, T_n(Z_n))'$ 也具有相同的 Copula 函数 C 。

关于不变性有以下注记：

(1) Copula 对于随机变量的严格单调增变换是不变的，由此可以导出一系列有意义的推论。例如，Copula 函数 $C(u_1, \dots, u_n)$ 反映了股票市场的相依结构，但它与股票价格的度量单位无关。因为不同的价格度量单位变换只是一个简单的线性变换 $ax + b$ ，且 $a > 0$ 时是严格单调增的变换。

(2) 由连接函数 $C(u_1, \dots, u_n)$ 导出的一些相关性描述指标，比经典的线性相关系数更加合乎人们的要求。因为通常的相关系数实际上是线性变换下不变的一种相关性指标，涉及到非线性变换的相关性，它就会导出错误的结论。例如若 $\xi \sim N(0, 1)$ ，而令 $\eta = \xi^2$ ，显然 η 依赖于 ξ 。但是计算 ξ 和 η 之间的线性相关系数： $\text{Cov}(\xi, \eta) = E\xi\eta - (E\xi)(E\eta) = E\xi^3 - (E\xi)(E\xi^2) = 0$ ，因而得到 ξ 和 η 是不相关。因此，用经典线性相关系数作为度量相关性的指标，在非线性变换之下发生改变，甚至招致误导。但是，由连接函数 C 所引