第 14 章　金融模型的概率分布

James B. McDonald

1. 引言

本文回顾了已经用于和可以用于解决金融领域问题的概率分布，并考察了这些分布在几个方面的应用。从纯统计的角度看，金融数据提供了来源丰富、具有各种分布特征的变量，从正态分布变量到具有不同偏度和峰度的变量。虽然正态分布或对数正态分布可能对许多金融序列提供了足够的代表性，但是其他序列却无法这样方便地建模。本文回顾正态分布、对数正态分布和稳定帕累托分布之外的其它几个重要的分布。

金融数据对个体投资者、公司计划人员、政治家和政府政策制订者来说都是非常重要的。金融数据总是在不断地变化着，这从每天的股票价格、利率、外汇汇率和黄金价格中可以明显地观察到。这些数据大多具有高度的不确定性，而这些变化则可能产生潜在的巨额收益或损失。

全世界的金融市场和证券交易所都在进行着股票、货币、商品和许多其他货物的交易。各种金融工具和交易都是可能存在的。即期市场通常方便了商品和金融工具所有权的即时转让。期货市场则方便了在未来某一特定时间以特定价格交易货物。期权给持有者以事先协议好的价格进行现货或者期货交易的权利。但是，这种权利并非一定要执行。股票、外汇、金属和商品都存在期权交易。所有这些都具有高度的不确定性。

芝加哥大学的证券价格研究中心（CRSP）是美国股票价格和收益数据的最大来源。这一数据库包括从 1962 年以来在纽约和美国证券交易所挂牌的每一只普通股的日收益率资料。CRSP 数据库也包括部分场外收益和 1926 年以后的月收益率数据。期货价格数据可以从哥伦比亚大学的期货市场研究中心得到，（参阅 Taylor（1926），p.26）。期货行业协会，一个非盈利教育机构，已经编制了一个数据库，这一数据库对那些研究期货和相关期权市场的人来说是很有用的，它包括了外汇和商品的数据资料。PACAP 数据库包括了亚洲市场的数据。

本文回顾了可以用来对金融资产收益分布进行建模的各种概率分布。第 2 节讨论了正态分布、学生 t 分布、对数正态分布、稳定分布、皮尔逊分布族和另外三个概率分布族。第 3 节考察了这些分布在描述收益分布、随机优势和期权定价中的应用。第 3 节是结论，讨论如何应用概率分布族以提供股票β值的部分适应（partially adaptive）估计量。

2. 可供选择的模型

2. 1. 一些背景知识

对金融工具收益建模通常有两种方法。第一种是描述生成价格的基本随机过程；第二种是指定一种能很好拟合经验数据的统计分布。本文将讨论可以用来描述收益的模型，但不研究基本随机过程；然而，一些模型具有结构的解释。

假设 P_t 表示交易日 t 某种金融工具的名义价格。进一步令 d_t 表示当天派发的股利（如果有的话）可以得到独立于计价单位的两种收益率的定义

$$y_t = \left(P_t + d_t\right) / P_{t-1} \qquad 0 < y_t \text{ 和}$$

$$z_t = \ln\left(y_t\right)$$

$$= \ln(P_t + d_t) - \ln(P_{t-1}), -\infty < z_t < \infty,$$

其中$(y_t - 1)$称为单收益率，而z_t称为复合收益率。由于当ε很小时，$\ln(1+\varepsilon)$的值很接近于ε，因此，基于y_t（或$y_t - 1$）的实证研究和基于z_t的研究通常能得到相同的结论。统计模型对两种数据形式都将给予讨论，Y 是正值而 Z 是任意实数值。例如，如果随机变量 Y 服从对数正态分布，那么$Z = \ln(Y)$将服从正态分布。

2.2. 基本概念和定义

假设 F(s)表示随机变量 S 的累积分布函数。在金融数据分析中经常涉及 S 的前四阶矩，令μ_i表示关于均值（μ）的i阶矩：

$$\mu_i = \mathrm{E}_F (S - \mu)^i = \int_{-\infty}^{\infty} (s - \mu)^i \, dF(s) \tag{2.1}$$

其中μ_2是方差；而通常用来衡量偏度的（$\sqrt{\beta_1}$）以及衡量峰度的（β_2）的定义式如下

$$\gamma_1 = \sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} \tag{2.2a}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \tag{2.2b}$$

对称分布的特征由γ_1=0 刻画。β_2是尾部厚度和尖度的衡量标准。γ_2=β_2—3 称为超峰度。当β_2<，=，或>3 时，称分布是低峰态、常峰态或者尖峰态（Stuart 和 Ord 1987，p.107）。尖峰态分布比正态分布更尖，尾部更厚。

正态随机变量的*正规化不完全矩或矩分布*的定义如下

$$\Phi(y;h) = \frac{\int_{-\infty}^{y} s^h f(s) ds}{E(s^h)} \tag{2.3}$$

$\Phi(y,0)$仅仅是累积分布，给出了$S \le y$的概率。$\Phi(y,1)$表示全部S中符合$S \le y$的部分。每一个$\Phi(y,h)$都具有累积分布的性质（是 y 的非减函数且当$y \to \infty$趋近于 1）—因此命名为*矩分布*。在期权定价和随机优势的讨论中将用到$\Phi(y,0)$和$\Phi(y,1)$。

在 2.3 和 2.4 节，我们转而讨论具体的概率密度函数。

2.3. 若干统计分布：正态分布，学生 t 分布，对数正态分布

正态分布、学生 t 分布、以及正态对数分布在金融文献中已经得到广泛的应用。简要回顾一

下这些重要分布的一些重要定义和性质。

*正态分布*函数用概率密度函数（pdf）定义

$$N(z;\mu,\sigma)=\frac{e^{-(z-\mu)^2/2\sigma^2}}{\sqrt{2\pi}\sigma},-\infty<z<\infty \tag{2.4}$$

正态分布是对称的（γ_1=0）且γ_2=β_2—3；它能够很好地拟合许多金融时间序列。然而，在金融收益数据中常常可以观测到显著高于正态分布的峰度（$\beta_2>3$）。

*学生 t 分布*关于原点对称且峰度等于$3+6/(v+3)$，其中 v 是“自由度”参数，而且允许比正态分布有更厚的尾部。其相应的概率密度函数，具有任意尺度系数（σ），定义为

$$\mathrm{T}(z;v,\sigma)=\frac{1}{\sqrt{\sigma}B(1/2,v/2)(1+2z^2/v\sigma^2)^{v+1/2}} \tag{2.5}$$

其中$B(,)$表示β分布（在附录 A 中给出定义）。当$h<v$时，对应于等式（2.5）的 h 阶矩（h 为偶数）为

$$\mathrm{E}_T(Z^h)=\frac{\sigma^h(v/2)^{h/2}B\left(\frac{h+1}{2},\frac{v-h}{2}\right)}{B\left(\frac{1}{2},\frac{v}{2}\right)} \tag{2.6}$$

当v变得无限大时，(2.5）近似于正态分布$\mathrm{N}(z;\mu=0,\sigma)$。Blattberg 和 Gonedes（1974）、Blattberg 和 Sargent（1971）在他们的金融论文中都应用了学生 t 分布。

Taylor（1986，p.44）发现许多收益分布不仅是肥尾，而且还呈现出正偏度。虽然学生 t 分布可以解释峰度的问题，但它不能用于偏斜数据的建模。*对数正态分布*$\mathrm{LN}(y;\mu,\sigma)$在金融中也得到广泛的应用，它的定义如下

$$\mathrm{LN}(y;\mu,\sigma)=\frac{e^{-(\ln(y)-\mu)^2/2\sigma^2}}{y\sigma\sqrt{2\pi}},\quad 0<y \tag{2.7}$$

对数正态分布的均值和方差分别为

$$\mathrm{E}(Y)=e^{\mu+\sigma^2/2} \tag{2.8a}$$

$$\mathrm{var}(Y)=\eta^2e^{2\mu+\sigma^2}\text{ 其中 }\eta^2=e^{\sigma^2}-1 \tag{2.8b}$$

Aitchison 和 Brown（1969，p.8）报告了对数正态分布偏度和峰度的表达式，$\gamma_1=\eta(\eta^2+3)$以及$\beta_2=\eta^8+6\eta^6+15\eta^4+16\mu^2+3$。这样，$\gamma_1$为正且随着参数$\sigma$的增大而增大。峰度大于 3，也随着$\sigma$的增大而增大。注意当$\sigma$取很小的值，偏度和峰度分别接近于 0 和 3。对数正态分布的累积分布函数为

$$\overline{\mathrm{LN}}(y;\mu,\sigma)=\frac{1}{2}+\frac{(\ln(y)-\mu)}{\sqrt{2\pi}\sigma}{}_1F_1\left[\frac{1}{2};\frac{3}{2};\frac{(\ln(y)-\mu)^2}{2\sigma^2}\right] \tag{2.9}$$

其中${}_1F_1[\ ]$表示附录 A 定义的合流超几何（confluent hypergeometric）序列。正态分布和对数正态分布的参数估计相对简单。参数估计的简便和分布的理论基础促进了它们在金融中的应用。虽然正态分布和对数正态分布在许多情况下能够提供合适的描述性模型，遗憾的是仍有许多数据无法用这些相对容易处理的模型来精确建模。解决这一问题有两个办法：从灵活参数分布族中选定模型，或者是运用半参数模型。本文将集中讨论灵活参数分布函数的使用问题。

2.4. 若干统计分布族

既然一些金融数据序列无法用正态分布、对数正态分布或学生 t 分布准确建模，所以通常需要使用一些更灵活的分布。这些分布包括稳定（stable）分布族、皮尔逊（Pearson）分布族、广义β（Generalized beta）分布族、第二类指数广义β（exponential Generalized beta of the Second kind）分布族和广义 t (Generalized t)分布族。每一分布族都包含了许多普通分布作为特例。这样，研究者可以检验相对于其特例，一个更一般形式的分布是否能显著地增进拟合。

稳定分布

关于股票收益正态性假设的再检验通常归功于 Mandlebrot（1963）。他发现价格变化的经验分布常比正态分布更尖峰和长尾。Mandlebrot（1963）研究了稳定分布族，稳定分布族可以用其特征函数的对数加以定义

$$K(t)=\ln C(t)=i\delta t-\gamma|t|^{\alpha}\left[1+i\beta\left(\frac{t}{|t|}\right)\tan\left(\frac{\alpha\pi}{2}\right)\right] \tag{2.10}$$

当$\beta=0$，基本密度函数是对称的，而且δ是中位数。当$\beta<0$，密度函数左偏，当$\beta>0$，密度函数右偏。参数α，称为稳定分布族的*特征指数*，取值范围是[1,2]，当α等于 1 或者 2（且$\beta=0$）分别是柯西（Cauchy）分布和正态分布。它们是这一范围内 pdf 具有已知闭合形(closed-form)表达式的唯一两个分布。为了具有有限均值，α的取值必须在$(1,2]$内。当$\alpha<2$，方差没有意义。Fama 和 Roll（1968）证明了随着α值的减小，分布尾部的厚度将增大。他们还略述了一种估计α的方法，并且用 Bergstrom 级数展开给出了其他分布的表达式。稳定分布族在加法下具有闭合性，也就是说，独立同分布的稳定分布变量之和仍然服从稳定分布。

Officer（1972）发现稳定分布为股票月收益率提供了一个合理的模型。然而，他发现α的估计值对日收益率之和的个数很敏感；这就引发了对闭合性和稳定分布适当性的怀疑。Hagerman（1978）同样研究了α的估计值，发现α的估计值趋向于从日收益率的 1.5 增加到 35 天收益率的 1.9；因此他不仅对闭合性提出疑问，还为极限正态分布提供了部分证据，尤其是对于月或者更长时期。既然股票收益率分布相对于正态分布倾向于肥尾，所以可以用对称稳定分布族来建模。Akgiray 和 Booth（1988）研究了 200 只普通股票分布的尾部。这 200 只股票选自 1000 只交易最活跃的股票。他们发现在经验分布和拟合分布之间存在显著差别。Lau，Lau 和 Wingender（1990）发现基于稳定分布族的四阶矩和六阶矩的估计值的实证表现和观测的股票收益的经验特征通常是不一致的。可以参阅 Blattberg 和 Gonedes（1979）中另一个例子。

皮尔逊分布族

皮尔逊分布族为那些不能用正态分布或对数正态分布准确建模的收益分布，提供了建模的另一种途径。著名的皮尔逊（1985，1901，1916）分布族用微分方程的解来定义

$$\Psi(s)=\frac{d\ln(f(s))}{ds}=\frac{(s-a)}{b_0+b_1 s+b_2 s^2} \tag{2.11}$$

分母有两个实根，不是（1）具有相同符号的实数，（2）具有不同符号的实数，就是（3）虚数。Elderton 和 Johnson（1969）、Kendall 和 Stuart（1969）及 Ord（1972）都讨论了皮尔逊分布族的性质。第一和第二类β分布、伽玛（Gamma）分布、学生 t 分布和正态分布都是皮尔逊分布族的特例或极限情况。皮尔逊分布族的特殊成员可以通过分析β_1和β_2的值或运用卡帕准则（kappa criterion）选出来，卡帕准则的定义如下

$$\kappa=\frac{b_1^2}{4b_0 b_2}=\frac{\beta_1(\beta_2+3)^2}{4(2\beta_2-3\beta_1-6)(4\beta_2-3\beta_1)} \tag{2.12}$$

例如，当$\kappa=\beta_1=0$（且$\beta_2=3$）时得到正态分布，而$\kappa=1$得到逆伽玛（inverse gamma）分布。Ord（1972，pp.8－9）提出皮尔逊分布族的一些扩展形式，即定义式微分方程（2.11）中的分子和分母可以是任意次数的多项式（Padè 逼近）。

皮尔逊分布族成员的许多估计方法已经得到研究。Pearson 用矩法来拟合数据的概率密度函数。除正态概率密度函数以外，矩法估计量对皮尔逊分布族是无效的。最大似然估计可以得到有效估计量。基于β_1、β_2或κ的矩法或最大似然估计量对分布分类，必须考虑样本变差。Ord（1972）引证了一些研究，他指出将这些方法运用于分组数据时进行分组校正的重要性。Hirschberg、Mazumdar、Slottje 和 Zhang（1992）将卡帕准则应用于股票收益分布的模型识别问题。Lau、Wingender 和 Lau（1989）发现偏度系数的准确估计需要非常大的样本。因此在分析中需要考虑卡帕准则的样本变差。

许多作者认为决定价格变化的分布并不需要有常数方差。如果收益以方差为条件，具有十分确定的概率密度函数而且随机方差服从已知分布，那么对应的收益分布被称为具有随机波动性或非均匀性，而且可以表示成*混合分布*的形式。后面将较为详细地讨论混合分布。然而，金融混合分布的两个早期例子是 Praetz（1972）和 Clark（1973）研究的，他们都假设收益以方差为条件，服从正态分布。Clark（1973）假设方差服从对数正态分布，从而导致观测收益的肥尾分布。Praetz（1972）也假设方差是随机的且服从逆伽玛分布。这一混合使得观测收益服从学生 t 分布。人们已注意到，学生 t 分布允许比正态分布有更厚的尾部，而且正态分布是它的特例。Blattberg 和 Gonedes（1974）用学生 t 分布对收益分布建模，发现它比稳定分布族有优势。

现在讨论允许混合解释的三个分布族，这样可以适应各种肥尾并准许非对称的情况。这些分布是第二类广义贝塔（GB2）分布、广义 t（GT）分布和第二类指数广义贝塔（EGB2）分布。

第二类广义贝塔分布

第二类广义贝塔分布（GB2）[1]用概率密度函数加以定义

[1] GB2 来自于广义贝塔分布（GB），定义如下

$$\mathrm{GB}(y;a,b,c,p,q)=\frac{|a|y^{ap-1}\left(1-(1-c)(y/b)^a\right)^{q-1}}{b^{ap}B(p,q)\left(1+c(y/b)^a\right)^{p+q}}\text{当}0<y^a<\frac{b^a}{1-c}$$

令 GB 的 c＝1 就得到 GB2。这一特例在研究收益分布时似乎很重要。然而，c＝0 则得到第一类广义贝塔分布，它在金融和经济模型中也有其他方面重要的应用。详细请参阅 McDonald 和 Xu（1995）。

$$\mathrm{GB2}(y;a,b,p,q)=\frac{|a|y^{ap-1}}{b^{ap}B(p,q)\left(1+(y/b)^a\right)^{p+q}},y\geq 0 \tag{2.13}$$

其中参数 b，p，和 q 都为正。GB2 分布被 Kalbfleisch 和 Prentice（1980）称为广义 F 分布，而它的修正形式（非零阈值）被 Arnold（1983）称为 Feller-Pareto 分布。GB2 中的$\Psi(y)$函数定义如下

$$\Psi(y)=\frac{d\ln f(y)}{dy}=\frac{ap-1-(aq+1)(y/b)^a}{y\left(1+(y/b)^a\right)} \tag{2.14}$$

它既不包括皮尔逊分布族（2.11）中的$\Psi(s)$，也不作为$\Psi(s)$特例而被它所包括。参数a,b,p和q以一种复杂的方式决定密度函数的形状和位置。Y 的 h 阶矩的定义如下

$$\mathrm{E}_{\mathrm{GB2}}\left(Y^h\right)=\frac{b^hB(p+h/a,q-h/a)}{B(p,q)} \tag{2.15}$$

当$-p<h/a<q$，且允许分析无限方差的情况。参数b纯粹是尺度参数而且依赖于度量单位。

一般来说，a或q的值*越大*，密度函数的尾部*越薄*。事实上，对大的参数a值，对应的 GB2 密度函数在参数b附近有集中的概率。注意到对于大的a值，均值近似为b而方差接近于 0 可以证实这一特性。参数 p和q 的相对值在确定偏度大小和允许正偏斜或负偏斜上具有重要的作用。这种情况与总是正偏斜的分布，例如对数正态分布是相反的。

GB2 的累积分布定义如下

$$\overline{\mathrm{GB2}}(y;a,b,p,q)=\left(z^p\right)_2F_1[p,1-q;p+1;z]/pB(p,q) \tag{2.16}$$

其中$z=\left[(y/b)^a/\left(1+(y/b)^a\right)\right]$和${}_2F_1[\ ]$是超几何序列（在附录 A 中给出定义）。

GB2 中的四个参数使得它具有很大的灵活性，而且作为特例或者极限情况嵌入许多重要的统计分布。它们包括第二类贝塔分布（$\mathrm{B2}=\mathrm{GB2}(y;a=1,b,p,q)$），三型伯尔（Burr type3）分布（$\mathrm{BR3}=\mathrm{GB2}(y;a,b,p,q=1)$），十二型伯尔（Burr type 12）分布（$\mathrm{BR12}=\mathrm{GB2}(y;a,b,p=1,q)$），以及广义伽玛（Generalized gamma）分布（GG）

$$\mathrm{GG}(y;a,\beta,p)=\frac{|a|y^{ap-1}e^{-(y/\beta)^a}}{\beta^{ap}\Gamma(p)}\qquad 0<y \tag{2.17}$$

它是 GB2 的一种极限情况

$$\mathrm{GG}(y;a,\beta,p)=\mathrm{Limit}_{q\to\infty}\mathrm{GB2}\left(y;a,\beta q^{1/a},p,q\right)$$

广义伽玛分布包括伽玛分布（$\mathrm{GA}=\mathrm{GG}(y;a=1,\beta,p)$），威布尔（Weibull）分布

（$W = \mathrm{GG}(y;a,\beta,p=1)$），和对数正态分布

$$\mathrm{LN}(y;\mu,\sigma) = \mathrm{Limit}_{a\to 0}\mathrm{GG}\left(y;a,\beta=\left(\sigma^2 a^2\right)^{1/a},p(a\mu+1)/\beta^a\right)$$

广义伽玛分布的 h 阶矩（$h/a<p$）为

$$\mathrm{E}_{\mathrm{GG}}\left(Y^h\right) = \frac{\beta^h\Gamma(p-h/a)}{\Gamma(p)} \tag{2.18}$$

当参数 a 取负值时，得到逆广义伽玛（Inverse Generalized Gamma）（IGG）分布，它源于随机波动性和非均匀性的模型。

广义伽玛分布的累积分布函数为

$$\overline{\mathrm{GG}}(y;a,\beta,p) = \frac{e^{-(y/\beta)^a}(y/\beta)^{ap}}{\Gamma(p+1)}{}_1F_1\left[1;p+1;(y/\beta)^a\right] \tag{2.19}$$

费雪 F（Fisher's F）分布、洛马克斯(Lomax)分布、贵斯克（Fisk）分布、半正态分布、半学生 t 分布、卡方分布、和雷利(Rayleigh)分布都是 GB2 的特例。利用从 McDonald（1984）或 McDonald 和 Xu（1995）的分布树可以看出它们之间的相互关系。

GB2 可以由广义伽玛分布和具有如逆广义伽玛分布的随机分布的尺度参数混合生成

$$\mathrm{GB2}(y;a,b,p,q) = \int_0^\infty \mathrm{GG}(y;a,s,p)\mathrm{IGG}(s;a,b,q)ds \tag{2.20}$$

等式（2.20）允许贝叶斯解释、非均匀性或随机波动性模型以及测量误差的某些形式。在不可观测的非均匀性模型中，第一个分布可以看成是子总体的结构分布，第二个分布代表尺度参数 s 的混合分布。根据 McDonald 和 Butler（1987）的研究，在 q 按照 $\mathrm{Limit}_{q\to\infty}\mathrm{GG}\left(s;a,q^{1/a}b,q\right)$ 增大的情况下，当 $s=b$ 时混合分布趋近于退化分布；从而对应的 GB2 分布趋近于 GG 分布。在金融模型的背景下，广义伽玛分布可以作为收益分布，以假定为服从逆广义伽玛分布的尺度参数为条件。对这一混合分布的解释为 GB2 作为收益模型提供了结构性的解释（随机波动性）。

广义 T 分布

广义 T（GT）分布有对称的三参数概率密度函数，它可以用来对各种峰度水平的收益 $z_t = \ln(P_t + d_t) - \ln(P_t)$ 建模，以概率密度函数加以定义

$$\mathrm{GT}(z;\sigma,p,q) = \frac{p}{2\sigma q^{1/p}B(1/p,q)\left(1+|z|^p/q\sigma^p\right)^{q+1/p}} \tag{2.21}$$

其中 $-\infty<z<\infty$，参数 $\sigma, p,$ 和 q 都是正值。GT 分布是由 McDonald 和 Newey（1988）引入金融文献的，Box-Tiao（BT）分布是它的一种极限情形

$$\mathrm{BT}(z;\sigma,p) = \mathrm{Limit}_{q\to\infty}\mathrm{GT}(z;\sigma,p,q) = \frac{pe^{-(|z|/\sigma)^p}}{2\sigma\Gamma(1/p)} \tag{2.22}$$

BT 分布是对称的，也被称为*幂指数分布*。正态分布是 BT 分布当 p＝2 时的一个特例。二重指数分布或拉普拉斯分布及学生 t 分布（自由度为 v 且没有单式（unitary）方差）是 BT 和 GT 分布如下形式的特例：

$$\text{Laplace}(z;\sigma)=BT(z;\sigma,p=1)=\frac{e^{-(|z|/\sigma)}}{2\sigma}$$

$$T(z;v,\sigma)=GT\left(z;\sigma,p=2,q=\frac{v}{2}\right). \qquad (2.23a－b)$$

GT 和 BT 分布的 h 阶矩（h 是偶数）为

$$\mathrm{E}_{\mathrm{GT}}\left(Z^h\right)=\frac{\sigma^h q^{h/p}\Gamma((1+h)/p)\Gamma(q-h/p)}{\Gamma(1/p)\Gamma(q)}$$

$$\mathrm{E}_{\mathrm{BT}}\left(\mathrm{z}^{\mathrm{h}}\right)=\frac{\sigma^h\Gamma((1+h)/p)}{\Gamma(1/p)} \qquad (2.24a－b)$$

BT 分布的各阶有限矩都存在；而 GT 的 h 阶矩仅当 $h<qh$ 时才存在。柯西分布是当 p＝2 和 q＝1/2 时 GT 分布的特例，它不存在有限整数矩。

GT 分布是对称的，能够适应比正态分布更厚或者更薄的尾部。GT 分布也为回归和时间序列模型的“稳健”或部分适应估计提供了基础。后面将考虑这些分布的应用。GT 分布可以解释成具有尺度参数（σ）的 BT 分布的一种混合尺度参数（a）具有逆广义伽玛（IGG）分布：

$$\mathrm{GT}(z;\sigma,p,q)=\int_0^\infty \mathrm{BT}(\mathrm{z;s,p})\mathrm{IGG}(s;p,\sigma,q)ds \qquad (2.25)$$

这个结果是学生 t 分布对应于具有逆伽玛分布尺度参数的正态分布的推广，Praetz（1972)。

第二类指数广义贝塔分布

虽然 GT 分布的尾部灵活性很重要，但许多收益分布还是偏斜的。实值随机变量的另一种允许偏斜和尖峰态的分布就是第二类指数广义贝塔分布（EGB2)，它的概率密度函数定义如下

$$\mathrm{EGB2}(z;\delta,\sigma,p,q)=\frac{e^{p(z-\delta)/\sigma}}{|\sigma|B(p,q)\left(1+e^{(z-\delta)/\sigma}\right)^{p+q}} \qquad (2.26)$$

$$-\infty<z<\infty$$

由于 EGB2 和 GB2 可以通过对数变换联系起来，因此 EGB2 的许多特例很容易被确定。这些分布中有几种在统计文献中特别重要。指数广义伽玛分布可以定义为

$$\mathrm{EGG}(\mathrm{z};\delta,\sigma,p)=\mathrm{Limit}_{\mathrm{q}\to\infty}\mathrm{EGB2}(\mathrm{Z};\delta*=\sigma\ln\mathrm{q}+\delta,\mathrm{p,q})$$

$$=\frac{e^{p(z-\delta)/\sigma}e^{-e^{(z-\delta)/\sigma}}}{|\sigma|\Gamma(p)} \qquad (2.27)$$

在 Johnson 和 Kotz（1970，Vol.2）及 Patil 等（1984）看来，$\sigma>0$ 的 EGB2 和 EGG 仅仅是广义逻辑斯谛（Generalized Logistic）分布和龚伯茨（Gompertz）分布的另一种表示。当 $p=q$ 时，EGB2 就是广义冈贝尔(Generalized Gumbell)分布。EBR3 是二型伯尔分布；指数威布尔分布通常称为*极值一型分布*。EGB2 和 EGG 的前四阶矩在表 1 中给出（详细请参阅 McDonald 和 Xu，1995）

其中，δ 是位置参数，σ 是尺度参数，而 p 和 q 是形状参数。改变 σ 的符号将改变偏斜的方向。当 $p=q$ 时 EGB2 是对称的。峰度（μ_4/μ_2^2）大于或者等于 3。正态分布作为一种极限情况包含于 EGB2 之中，EGB2 可以用来描述回归误差、时间序列、或其他要考虑到偏离正态性的模型。EGB2 为边界影响函数（bounded influence function）的部分适应估计提供了基础。

表 1
EGB2 和 EGG 的矩

矩	EGB2	EGG
均值（μ）	$\delta+\sigma\left[\Psi(p)-\Psi(q)\right]$	$\delta+\sigma\Psi(p)$
方差（μ_2）	$\sigma^2\left[\Psi'(p)+\Psi'(q)\right]$	$\sigma^2\Psi'(p)$
偏度（μ_3）	$\sigma^3\left[\Psi''(p)-\Psi''(q)\right]$	$\sigma^3\Psi''(p)$
峰度（$\mu_4-3\mu_2^2$）	$\sigma^4\left[\Psi'''(p)+\Psi'''(q)\right]$	$\sigma^4\Psi'''(p)$

μ_i 表示关于均值的 i 阶矩，Ψ(s)表示伽玛函数的导数[dlnΓ(s)]/ds。

（详细请参阅 McDonald 和 Xu，1995）

EGB2 有如下混合解释：

$$\mathrm{EGB2}(z;\delta,\sigma,p,q)=\int_0^\infty \mathrm{GG}\left(e^z;\frac{1}{\sigma},s,p\right)e^z\,\mathrm{IGG}\left(s;\frac{1}{\sigma},e^{\delta},q\right)ds \qquad (2.28)$$

估计

GB2、GT 或 EGB2 分布族中未知参数的最大似然估计要求非线性最优化。这些估计量是渐近有效和渐近正态的。现在考察这些分布在金融文献中的应用。

3. 在金融中的应用

我们转而研究第 2 节讨论的分布在四个方面的应用：股票收益的分布、随机优势、期权定价、股票β值的部分适应估计。

3. 1. 证券价格收益的分布

在金融文献中，对证券收益分布建模有两种基本的方法。第一种从设定证券价格生成的基本随机过程着手。第二种是实证的方法，即确定一个能够合理、准确拟合观测收益率的统计分布函数。实际数据分布往往比正态分布和对数正态分布具有更厚的尾部和更尖的峰态。正如前面提到的那样，这些观测资料导致对对称－稳定分布以及其他分布的研究。一个普遍的假设是，证券价格的分布是一种混合分布。例如，将收益的对数正态分布和反映波动率的逆伽玛分布进行混合可以得到一个分布，它的峰度比对数正态分布更符合观测值的峰度。这个特殊的混合分布，就是*对数－t* 分布，对数正态分布是它的一种极限情况。前面已经提到学生 t 分布是正态分布和σ的逆伽玛分布的混合结果。

在前面一节中，GB2 被表示成广义伽玛分布和尺度参数的逆广义伽玛分布（波动率）的混合：

$$\mathrm{GB2}(y;a,b,p,q)=\int_0^\infty \mathrm{GG}(y;a,s,p)\mathrm{IGG}(s;b,q)ds \qquad (3.1)$$

（3.1）中的 $\mathrm{GG}(y;a,s,p)$ 分布可以解释为给定 s 的收益条件分布，其中假定 s 服从指定的逆广义伽玛分布。由于对数正态分布是广义伽玛分布的一种极限情况，Praetz（1972）研究发现，GB2 推广了对数正态分布和伽玛分布的混合。回想当参数 q 趋于无穷大时，（3.1）中的 IGG 分布接近于退化概率密度函数；这样 GB2 允许、但不必隐含随机波动率模型。

此外，GB2 具有直到 aq 的有限阶矩。$aq<2$ 的 GB2 分布不具有有限方差。

Bookstaber 和 McDonald（1987）研究了 21 只随机选取的股票，从 1981 年 12 月 30 日开始，500 天的日收益率（$Y_t=\left(P_t+d_t\right)/P_{t-1}$）的分布。两倍的最大似然值之差（$\mathrm{LR}=\left(\ell_{GB2}-\ell_{\mathrm{LN}}\right)$）为零假设 H_0：$\mathrm{GB2}=\mathrm{LN}$ 的似然比检验提供了基础。使用基于 $\chi^2(2)$ 的临界值是统计显著性的保守检验（conservative test）。Bookstaber 和 McDonald（1987）发现 21 只股票中的 19 只都超过了 0.995 置信水平的临界值 10.6。相对于对数正态分布，更具灵活性的 GB2 分布提供了具有更高统计显著性的拟合。

本文也作了一项独立研究，调查了随机选取的 45 家公司从 1988 年 1 月到 1992 年 12 月的 60 个月包括股利在内的股票收益率。附录 B 所列出了所选取的 45 家公司。每一数据集都用最大似然法对几种分布进行拟合。在对零假设 H_0：$\mathrm{GB2}=\mathrm{LN}$ 的检验中，45 家公司中只有 10 家公司的 LR 值大于 5.99（95％的置信水平），而只有 6 家公司的 LR 值大于 10.6。这些结果进一步证实了先前研究的发现，即较长时期的收益分布比短时期的收益分布更接近于对数正态分布（正态分布）。

表 2 和表 3 报告了其中一家公司及纽约股票交易所的估计结果。表 2 列出对 Ampco-Pittsburgh 公司（AMPCO）的收益率数据用 MLE 估计 GB2、BR12、GA 和 LN 分布的结果，包括参数估计值，矩估计值（对应估计的参数），和最大对数似然值（ℓ）。表 2 中第五行到第八行列出的均值、方差、偏度和峰度是将估计的参数值代入矩的理论公式计算出来的，例如 GB2 的（2.15）。列在表底部的矩估计值是由样本矩得到。估计的两参数 LN 分布能很好模拟样本的均值和方差，但是却没有反映样本偏度和峰度的灵活性。GB2 分布附加的两参数对偏度和峰度的模拟具有统计显著的灵活性。注意这些结果是基于最大似然估计法而不是矩法。有意思的是三参数 BR12 和 GB2 的结果非常相似。BR12 是具有闭合型累积分布的三参数分布。

运用同样的四种统计分布对价值加权纽约股票交易所指数（VWNYSE）的月收益率数据进行拟合。结果列于表 3。相应的 LR 在常规显著性水平上不具有统计显著性；然而，零假设 H_0：$\mathrm{GB2}=\mathrm{LN}$ 包括了参数空间边界上的参数。这引发了基于渐近 $\chi^2(2)$ 的推断准确性问题。AMPCO 和 VWNYSE 的数据在附录 B 中给出。

表 2

AMPCO-Pittsburgh 公司 的月收益率分布估计（1988 年 1 月-1992 年 12 月）

	GB2	BR12	GA	LN
a(μ)	29.34	24.97	1.000	(-.004276)
b(σ)	.9642	.9583	.009592	(.09625)
p	.7726	1.000	104.3	N/A
q	.4977	.6006	N/A	N/A
均值	1.0001	1.0002	1.0005	1.0004
方差	.0092	.0091	.0096	.0093
偏度	1.184	1.164	.196	.290
峰度	7.505	7.164	3.06	3.15
ℓ	60.3	60.2	54.4	55.6

N/A—不适用

样本矩：（均值，方差，偏度，峰度）＝（1.0005，.0105，1.73，9.13）

表 3

VMNYSE 的月收益率分布估计（1988 年 1 月-1992 年 12 月）

	GB2	BR12	GA	LN
a(μ)	118.6	53.09	1.000	(.01106)
b(σ)	1.013	1.010	.001239	(.03501)
p	.3464	1.000	816.8	N/A
q	.3672	.9721	N/A	N/A
均值	1.012	1.012	1.012	1.012
方差	.0013	.0013	.0013	.0013
偏度	.129	.198	.0700	.1051
峰度	5.39	4.31	3.01	3.02
l	116.8	116.5	115.3	115.3

样本矩：（均值，方差，偏度，峰度）＝（1.012，.0012，.0511，3.79）

3.2. 随机优势

本节将讨论对不同收益分布进行比较的方法以及概率密度函数在这一重要问题中的一些应用。首先回顾均值-方差等级、一阶和二阶随机优势的概念。这些等级之间的关系和期望效用给出了最优性的概念。一些概率密度函数的参数约束导致随机优势的情况将得到讨论。最后，将讨论洛伦茨优势（Lorenz dominance）和均值-基尼优势（Mean－Gini dominance）的概念，以及它们与随机优势之间的关系。

均值-方差和随机优势

令 F_1 和 F_2 表示对应于两种不同资产 X_1 和 X_2 的累积收益分布。进一步，假设 μ_i 和 σ_i^2 分别表示 X_i 的均值和方差。Markowitz（1959）和 Tobin（1958）提出给分布划分等级的*均值-方差*（MV）准则。根据均值-方差准则，分布 F_1 被称为优于(更可取)分布 F_2

MV：

$F_1 >_{MV} F_2$ 或 $X_1 >_{MV} X_2$ 当且仅当：

$$\mu_1 \geq \mu_2 \text{ 且} \qquad \sigma_1^2 \leq \sigma_2^2 \tag{3.2}$$

其中至少有一个是严格不等式。均值-方差准则将备择集分为“容许或有效”集（S_{MV}）和“非容许或非有效”集。将具有比原始资产集合中的成员更小的均值和更大的方差的资产剔除就会得到容许集。这样，非容许集不包含任何比容许集有更大均值和更小方差的资产。

从表 2 和表 3 的数字可以看出，$\text{VWNYSE} >_{MV} \text{AMPCO}$。45 家随机选取的公司相应的均值-方差有效集包括 Aileem、Atlantic Energy、General Public Utilities、NUCOR、Union Pacific 和 Walgreen。

一阶和二阶随机优势的概念为划分分布等级提供了可供选择的决策规则。分布 F_1 被称为*一阶*

随机优于（FSD）分布 F_2

FSD：

$$\boxed{\begin{array}{l} F_1 >_{\text{FSD}} F_2 \text{ 当且仅当：} \\ F_1(x) \le F_2(x) \text{ 对所有 } x\text{，} -\infty < x < \infty\text{，，} \\ \text{且} \\ \text{对某个 } x_0\text{，有 } F_1(x_0) < F_2(x_0) \end{array}} \tag{3.3}$$

因而，$F_1 >_{\text{FSD}} F_2$ 要求 F_1 永不能位于 F_2 之上而且在某处位于 F_2 之下。随之得出结论，FSD 必要但非充分的条件，是最佳资产的均值（如果存在）至少和优势资产的均值一样大。对应的有效集记为 S_{FSD}，它不必等于 S_{MV}。

分布 F_1 被称为*二阶随机优于*（SSD）分布 F_2，记

SSD：

$$\boxed{\begin{array}{l} F_1 >_{\text{SSD}} F_2 \text{ 当且仅当：} \\ \int_{-\infty}^{x} F_1(t)\,dt \le \int_{-\infty}^{x} F_2(t)\,dt \text{ 或} \\ \text{对所有 } x\text{，} -\infty < x < \infty\text{，有 } \int_{-\infty}^{x} \left[F_1(t) - F_2(t)\right] dt \le 0 \end{array}} \tag{3.4}$$

其中至少对某一 x 是严格不等式。$F_1 >_{\text{SSD}} F_2$ 要求 F_1 的积分不能位于 F_2 的积分之上而且在某处位于 F_2 积分之下。与 FSD 相反的是，SSD 允许 F_1 和 F_2 在多处相交，只要负区域（当 $F_1 > F_2$）的绝对值小于累积正区域（当 $F_2 > F_1$）。一阶随机优势包含了二阶随机优势。因此 $\text{S}_{\text{SSD}} \subset \text{S}_{\text{FSD}}$。再次注意一下对应于 MV、FSD 和 SSD 的容许集并不需要相同，因此可能导致不同的决策结果。期望效用的概念给解决这些差别提供了一种方法。

期望效用和最优性

Von Neumann 和 Morgenstern（1953）证明了期望效用可以作为不确定条件下决策的基础。因此如果 $U(x)$ 表示效用函数，分布可以按照期望效用进行分级。

$$\text{E}_i(Y) = \int U(Y)\,dF_i(Y) \tag{3.5}$$

显然，基于期望效用的等级划分依赖于对效用函数所作的假设，且可能不同于 MV、FSD 或 SSD 准则。最优有效集是一系列使得期望效用（对应于效用函数的不同假设）最大化的分布（或资产）的集合。因此，在某些限制假定下，S_{SSD} 和 S_{MV} 可以是最优的。

Tobin（1958）及 Hanoch 和 Lévy（1969）发现当效用函数是二次函数或收益分布是正态分布时，均值-方差准则是有效的（均值－方差容许集 Smv 是最优的）。Pratt（1964）和 Arrow（1965）

讨论了二次效用函数的局限性（递增的绝对风险厌恶）。再者，关于收益分布正态性的假设排除了偏度和尖峰态，而这是许多收益分布具有的特征。

Quirk 和 Saposnik（1962）、Fishburn（1964）、Hanoch 和 Lévy（1969）证明了当且仅当效用函数是非减函数时 FSD 是最优的。这是根据下式

$$\mathrm{E}_{F_1}U(X)-\mathrm{E}_{F_2}U(X)=\int_{-\infty}^{\infty}[F_2(t)-F_1(t)]dU(t) \tag{3.6}$$

得出的。已经证明当效用函数是非减的凹函数时，SSD 提供最优的等级，详细请参阅 Hanoach 和 Lévy（1969）。

随机优势和参数族

Ali（1975）研究了当分布属于各种参数分布族时的随机优势。Ali 使用 Lehmann（1959）报告的关于单调似然比的结果，为对应于 FSD 和 SSD 不同分布族的参数空间划分子集。他研究了伽玛、贝塔、t、F、χ^2 和对数正态分布族。作为一个例子，考虑伽玛密度函数

$$\mathrm{GA}(Y;\beta,p)=\mathrm{GG}(Y;a=1,\beta,p)=\frac{Y^{p-1}e^{-y/\beta}}{\beta^{p}\Gamma(p)} \tag{3.7}$$

Ali（1975）发现

$$\boxed{\begin{array}{c}\mathrm{GA}(Y;\beta_1,p_1)\succ_{\mathrm{FSD}}\mathrm{GA}(y;\beta_2,p_2)\text{ 当且仅当}\\ \beta_2\le\beta_1\text{ 且 }p_2\le p_1\end{array}} \tag{3.8}$$

其中至少有一个是严格不等式。因此，判定伽玛分布族中的一个分布是否优于另一个分布，只要比较参数值[2]。但是这不便于比较两个来自不同族的分布。因为 GB2 嵌入了伽玛和贝塔分布族，在尝试寻找对来自不同族分布进行比较的简便方法，也可以考虑用同一方法。

运用 Lehmann（1959）提出的方法论，首先要计算似然比：$\mathrm{LR}(\mathrm{y};\Theta_1,\Theta_2)=\ln f(y;\Theta_1)-\ln f(y;\Theta_2)$。如果当 $\Theta_1>\Theta_2$ 时 $d\mathrm{LR}(y;\Theta_1,\Theta_2)/dy$ 是单调非减的，那么有 $F_{\Theta 1}\succ_{\mathrm{FSD}}F_{\Theta 2}$。进一步举例说明，广义伽玛分布的对数似然比的导数如下

$$\frac{dLR_{GG}}{dy}=\frac{a_1p_1-a_2p_2}{y}+\frac{a_2}{y}\left(\frac{y}{\beta_2}\right)^{a2}-\frac{a_1}{y}\left(\frac{y}{\beta_1}\right)^{a} \tag{3.9}$$

参数 p 和 β 值的增加看来将导致对应于较大参数值的一阶随机优势。这对于任意 a 都成立。这证实了先前列举的伽玛分布的一些结论。当参数 p 或 β 的值增加，而另一个参数的值则减少，二者综合，参数 a 变化的影响不清楚。同样，第二类广义贝塔分布的对数似然比的导数可以写成

[2] Pope 和 Zimer（1984）研究了有效性功效（power）检验中抽样变差在估计的均值、方差和参数值中的作用。

$$\frac{dLR_{\mathrm{GB2}}}{dy}=\frac{a_1p_1-a_2p_2}{y}+\frac{a_2\left(p_2+q_2\right)}{y}\left[\frac{1}{1+\left(b_2/y\right)^{a_2}}\right]-\frac{a_1\left(p_1+q_1\right)}{y}\left[\frac{1}{1+\left(b_1/y\right)^{a_1}}\right] \quad (3.10)$$

随机优势和洛伦茨优势[3]

Atkinson（1970）指出，随机优势准则可以根据洛伦茨曲线重新构造，洛伦茨曲线在经济文献中用于比较收入分布。一个收入分布的洛伦茨曲线，描绘的是人口不同的百分比所相应拥有的全部收入的百分比。洛伦茨曲线因而是不完全矩（$\phi(y;0),\phi(y;1)$）标绘的图，其中$\phi(y;0)$ 表示收入低于 y 的人口比例，而$\phi(y;1)$是收入低于 y 的这部分人口所拥有的收入占全部收入的比例。Atkinson（1970）证明了对两个具有相同均值的分布而言，$F_1>_{SSD}F_2$意味着F_1的洛伦茨曲线位于F_2的洛伦茨曲线之上。有关洛伦茨优势的文献采用这样的定义，即F_2 *洛伦茨优于* F_1

L：

$F_2>_{\mathrm{L}}F_1$ 当且仅当
F_1的洛伦茨曲线位于F_2的洛伦茨曲线之上

（3.11a）

考虑逆洛伦茨的等级划分

IL：

$F_1<_{\mathrm{IL}}F_2$ 当且仅当
F_1的洛伦茨曲线位于F_2的洛伦茨曲线之上

（3.11b）

有助于提醒我们 L 等级划分的方向和 SSD、FSD 等相反。此外，Atkinson（1970）指出对具有相同均值的分布，$F_1>_{SSD}F_2$等价于洛伦茨优势或逆洛伦茨优势（$F_1<_{\mathrm{IL}}F_2$）。在这种情况下，不相交的洛伦茨曲线的等级，除是非减凹函数之外，独立于社会福利函数的形式。在交叉的洛伦茨曲线的情形，不同的福利函数会导致不同的等级。对于不同均值，

$$\boxed{\mu_1\geq\mu_2 \text{ 且 } F_1>_{\mathrm{IL}}F_2 \text{ 意味着 } F_1>_{SSD}F_2} \quad (3.12)$$

某些分布，例如伽玛分布、帕累托分布和对数正态分布，不允许交叉的洛伦茨曲线；等级特征由单一的形状参数刻画。其他分布，例如伯尔分布或者广义伽玛分布，允许洛伦茨曲线交叉，就要求更复杂参数约束来刻画洛伦茨优势的特征。这些结果的一部分将会述评。

洛伦茨优势：十二型伯尔分布

[3] Shorrocks（1983）和 Kakwani（1984）建立了广义洛伦茨曲线，它考虑了等级分布的不同均值。广义曲线按照分布均值等比例扩大洛伦茨曲线得到。广义洛伦茨优势等价于依照 S－凹社会福利函数优先选择。在广义洛伦茨优势和二阶随机优势之间存在对偶性。Bishop、Chakraborti 和 Thistle（1989）概括广义洛伦茨曲线的无分布推断过程。

Wilfing 和 Kramer（1993）发现刻画十二型伯尔分布 $\mathrm{GB2}(y;a,b,p=1,q)$ 洛伦茨优势特征的参数约束：

$$\boxed{\begin{array}{l}\mathrm{GB2}(y;a_1,b_1,p=1,q_1) \succ_{\mathrm{IL}} \mathrm{GB2}(y;a_2,b_2,p=1,q_2) \\ \text{当且仅当 } a_1 \geq a_2 \text{ 且 } a_1q_1 \geq a_2q_2\end{array}} \qquad (3.13)$$

比较表 2 和表 3 中十二型伯尔分布的估计参数，隐含 $\mathrm{VWNYSE} \succ_{\mathrm{IL}} \mathrm{AMPCO}$。

洛伦茨优势：第二类广义贝塔分布

对更一般情况的 GB2，Wilfing 和 Kramer（1993）发现如下洛伦茨优势的*必要*条件：

$$\boxed{\begin{array}{l}\mathrm{GB2}(y;a_1,b_1,p_1,q_1) \succ_{\mathrm{IL}} \mathrm{GB2}(y;a_2,b_2,p_2,q_2) \text{ 意味着} \\ a_1p_1 \geq a_2p_2 \text{ 且 } a_1q_1 \geq a_2q_2\end{array}} \qquad (3.14)$$

Wilfing（1992）发现了充分条件：

$$\boxed{\begin{array}{l}a_1 \geq a_2\text{，} p_1 \geq p_2\text{，且 } q_1 \geq q_2 \text{ 意味着} \\ \mathrm{GB2}(y;a_1,b_1,p_1,q_1) \succ_{\mathrm{IL}} \mathrm{GB2}(y;a_2,b_2,p_2,q_2)\end{array}} \qquad (3.15)$$

因此，参数 a, p 或 q 的增加将导致逆洛伦茨优势。基于表 2 和表 3 报告的 GB2 参数的估计值，我们注意到 VMNYSE 洛伦茨优于 AMPCO 的必要、但非充分的条件得到满足。

洛伦茨优势：广义伽玛分布

Taille（1981，p.190）研究了具有两个形状参数的广义伽玛分布。他指出与不交叉洛伦茨曲线相关联的参数约束

$$\boxed{\begin{array}{l}\mathrm{GG}(y;a_1,b_1,p_1) \succ_{\mathrm{IL}} \mathrm{GG}(y;a_2,b_2,p_2) \text{ 当且仅当} \\ a_1 \geq a_2 \text{ 且 } a_1p_1 \geq a_2p_2\end{array}} \qquad (3.16)$$

均值-基尼优势

均值-方差排序具有公认的局限性。一个可选择的排序（涉及洛伦茨排序）是使用基尼系数。基尼系数是 45 度均等线和洛伦茨曲线之间面积的两倍，它作为不平等的数量度量由来已久，并且用作比较收益分布的一个准则。这一方法是由 Yitzhaki（1982）以及 Shalit 和 Yitzhaki（1984）的论文引入金融文献的。基尼系数的定义式为

$$G_i = \frac{1}{2\mu_i}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|s-t|\,dF_i(s)\,dF_i(t) \qquad (3.17)$$

洛伦茨优势 $F_1 >_{\mathrm{IL}} F_2$ 意味着 $G_1 < G_2$。Yitzhaki（1982）认为均值和基尼系数可以用来描述普通分布随机优势的必要条件，而运用均值-方差准则则不可能做到这一点。根据均值-基尼准则（MG），分布 F_1 被称为优于分布 F_2：

MG：

$$
\boxed{\begin{array}{l} F_1 >_{\mathrm{MG}} F_2 \text{ 当且仅当：} \\ \mu_1 \geq \mu_2 \\ G_1 \leq G_2 \end{array}} \tag{3.18}
$$

其中至少有一个是严格不等式。将均值-基尼准则应用于前面讨论的 45 只股票，得到和基于均值-方差准则*相同*的有效集，即 Aileem、Atlantic、Energy、General Public Utilities、NUCOR、Union、Pacific 和 Walgren。

Yitzhaki（1982）提出对分布划分等级的附加准则，基于下列命题：

命题 1. $\lambda_n \geq 0, n = 1,2,\cdots,$ 是 FSD 和 SSD 的必要条件。其中

$$
\lambda_n = \int \left[[1 - F_1(t)]^n - [1 - F_2(t)]^n \right] dt \tag{3.19}
$$

求 λ_1 和 λ_2 的值有

$$
\lambda_1 = \mu_1 - \mu_2 \geq 0
$$

和

$$
\lambda_2 = \mu_1(1 - G_1) - \mu_2(1 - G_2) = \int \left[[1 - F_1(t)]^2 - [1 - F_2(t)]^2 \right] dt \geq 0
$$

这些条件导致一个不同的均值-基尼（MG1）准则，根据 MG1，F_1 被称为优于 F_2：

MG1：

$$
\boxed{\begin{array}{l} F_1 >_{\mathrm{MG1}} F_2 \text{，当且仅当} \\ \mu_1 \geq \mu_2 \\ \mu_1(1-G_1) \geq \mu_2(1-G_2) \end{array}} \tag{3.20}
$$

其中至少有一个不等式。

$F_1 >_{\mathrm{MG}} F_2$ 意味着 $F_1 >_{\mathrm{MG1}} F_2$，反之则不成立。因此，对应 MG1 的有效集将包含在从 MG 准则得到的有效集。准则越弱，有效集越小。对交叉次数不超过一次的累积分布，Shalit 和 Yitzhaki（1984）认为“$>_{\mathrm{MG1}}$”（具有相同均值）是一阶和二阶优势的充分条件，且 $S_{\mathrm{MG1}} = S_{\mathrm{SSD}}$。将 MG1 应用于 45 只股票，Altantic Energy 从 MV 和 MG 有效集中被剔除。

表 4 报告了对应于正态分布、对数正态分布、伽玛分布、贝塔（一型和二型）、十二型伯尔分

布、广义伽玛分布和 GB2 分布的基尼系数的表达式。

表 4

基尼系数

分布	基尼系数
正态	$\dfrac{\sigma}{\mu\sqrt{\pi}}$
对数正态	$2\overline{\mathrm{LN}}\left(\dfrac{\sigma}{\sqrt{2}};0,1\right)-1$
伽玛	$\dfrac{\Gamma(p+1/2)}{\sqrt{\pi}\,\Gamma(p+1)}$
B_1	$\dfrac{B(p+q,1/2)B(p+1/2,1/2)}{\pi B(q,1/2)}$
B_2	$\dfrac{2B(2p,2q-1)}{pB^2(p,q)}$
$BR12$	$1-\dfrac{\Gamma(q)\Gamma(2q-1/a)}{\Gamma(q-1/a)\Gamma(2q)}$
GG	G_{GG}
GB2	G_{GB2}

其中

$$\mathrm{G}_{\mathrm{GG}}=\frac{\left[(1/p)_2F_1[1,2p+1/a;p+1;1/2]\right.}{\left[2^{2p+1/a}B(p,p+1/a)\right]}$$

$$\frac{-\left(\dfrac{1}{p+1/a}\right)_2F_1[1,2p+1/a;p+1/a+1;1/2]\left.\right]}{\left[2^{2p+1/a}B(p,p+1/a)\right]}$$

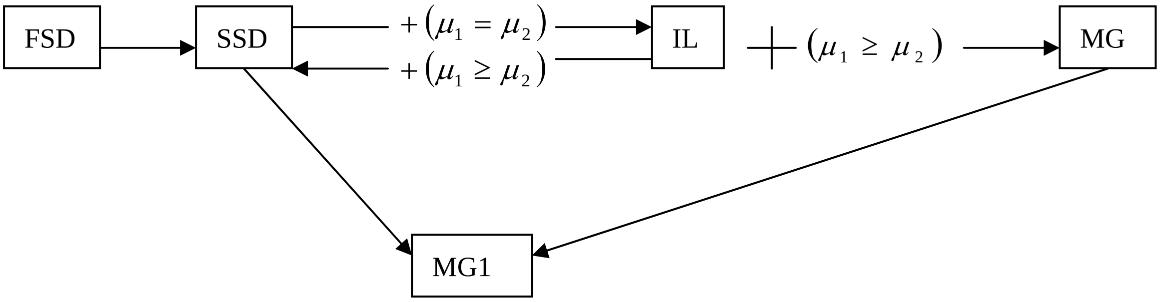
$$G_{GB2}=\frac{\left[(1/p)_3F_2[1,p+q,2p+1/a;p+1,2(p+q);1]\right.}{B(p,q)B(p,p+1/a)B(2q-1/a,2p+1/a)]}$$

$$\frac{-\left(\dfrac{1}{p+1/a}\right)\ {}_3F_2[1,p+q,2p+1/a;p+1/a+1,2(p+q);1]}{B(p,q)B(p,p+1/a)B(2q-1/a,2p+1/a)]}$$

有关这些公式的参考文献参阅 Nair（1936）、Aitkinson 和 Brown（1970）、McDonald（1984），Salem 和 Mount（1974）、Singh 和 Maddala（1976）。这些公式可以用于构造 MG 和 MG1 有效集。也可以使用基尼的非参数估计。

可选择的等级划分之间的关系

下图概括了 FSD、SSD、IL、MG 和 MG1 等级之间的一些关系：



如果累积分布最多只有一处相交且均值相等，那么 MG1 隐含着 SSD。在均值相等的条件下，SSD 和 IL 是等价的。

表 4 中的结果可以用来为不同参数族构造 MG 或 MG1 有效集。可以证明正态分布的各种有效集之间存在以下关系：

$$S_{MG1}\subset S_{MG}=S_{SSD}=S_{MV}\text{, Yitzhaki（1982）}$$

对数正态分布各种有效集之间的关系则不同，可以证明是 $S_{MG1}\subset S_{\mathrm{SSD}}\subset S_{\mathrm{MG}}=S_{MV}$。因此，对数正态分布提供了均值－方差准则和随机优势不一致的一个例子，Yitzhaki（1982）。同样可以参阅 Elton 和 Creber（1973）。

3. 3. 期权定价

Black Scholes（1973）期权定价公式已经被广泛地用于金融资产的定价。这个公式基于收益服从对数正态分布的假设，这和数据之间的吻合可能很差。解决这一问题的方法是基于广义贝塔分布来估计期权定价公式。已经指出，对数正态分布是 GB2 分布的一个极限情况，因而 GB2 允许对对数正态分布的偏离。关于 GB2 是混合分布的解释（参见（2.20））同样也说明了由于随机波动性，允许对对数正态分布的偏离。

Cox 和 Ross（1976）推导出了证券过程的累积分布函数和该证券期权均衡值之间的关系。如果在金融资产定价中可以假设风险中性，那么欧式看涨期权的均衡价格就是其到期日期望收益的现值，

$$C(S_T,T,X)=e^{-rT}\mathrm{E}[C(S_0,0)]$$

$$=e^{-rT}\int_X^{\infty}(S-X)f(S|S_T,T)dS \qquad (3.21)$$

其中 C,T,r,X 和 S_T 分别表示期权的价格、距期权到期日的时间、利率、执行价格和股票价格（距到期日的 T 个时期），Bookstaber（1987）。用正规化不完全矩可以更方便地写出该表达式，标准化不完全矩为

$$\phi(y;h)=\frac{\int_{-\infty}^{y}s^h f(s)ds}{\mathrm{E}(y^h)}$$

进一步，令

$$\bar{\phi}(y;h)=1-\phi(y;h)$$

这样，欧式看涨期权的均衡价格（3.21）可以写成

$$C(S_T, T, X) = S_T \overline{\phi}\left(\frac{X}{S_T};1\right) - e^{-rT} X \overline{\phi}\left(\frac{X}{S_T};0\right), \tag{3.22}$$

McDonald 和 Bookstaber（1991）。

将 f（）选为对数正态函数，就可以得到 Black Scholes（1973）期权定价公式，注意对数正态分布的正规化不完全矩是修正参数的对数正态累积分布函数[4]：

$$\phi_{\mathrm{LN}}(y;h) = \overline{\mathrm{LN}}\left(y;\mu + h\sigma^2, \sigma^2\right)$$

欧式看涨期权价格类似的表达式也可以从相应的 GB2 和 GG 分布得到，只需注意

$$\phi_{\mathrm{GB2}}(y;h) = \overline{\mathrm{GB2}}\left(y;a,b,p+\frac{h}{a},q-\frac{h}{a}\right)$$

$$\phi_{\mathrm{GG}}(y;h) = \overline{\mathrm{GG}}\left(y;a,\beta,p+\frac{h}{a}\right) \tag{3.23a－b}$$

Butler 和 McDonald（1989）。注意 GG 和 GB2 分布的不完全矩是 GG 和 GB2 累积分布函数（2.18）和（2.16）族的一员，因此它们具有闭合性。

McDonald 和 Bookstaber（1991）研究了基于 GB2 的欧式期权定价公式的运用，GB2 分布可以具有和对数正态分布不同的偏度和峰度。他们发现，相对于对数正态分布，峰度增大，Black-Scholes 模型会高估平价期权。而对极度实值期权，Black-Scholes 模型开始就低估它们的价格。Black-Scholes 公式的价格偏离对峰度和偏度都是敏感的。这些发现通过数字例子进一步加以说明。

例如，当$T = .25, r = .10, X = 100$和$\sigma^2 = .40$，得到 Black Scholes（BS）的价格是$13.68。对应的对数正态分布的偏度和峰度分别为 1.0007 和 4.856。现在考虑，逐渐增大峰度或减小偏度，用矩法拟合一个 GB2 分布。给出估计的 GB2 分布，期权价格可以由（3.22）和（3.23a）导出。表 5 列出了一些具有代表性的期权价格。

表 5

GB2 期权价格（$T = .25, r = .10, X = 100$和$\sigma^2 = .40$）

S_T	BS	%Δ峰度		%Δ偏度		
		50	100	－25	－50	－75
90	8.39	7.94	7.53	8.20	7.98	7.75
100	13.68	13.40	13.20	13.72	13.76	13.96
110	20.19	20.21	20.30	20.50	20.81	21.19

表中的数值反映了非正态性（对数正态性）对 Black-Scholes 定价公式精确性的影响作用。例如，如果对数正态分布能够准确地代表收益分布，那么价格为 100、执行价格为 100 的股票期权价格就是$13.68。如果收益分布的均值、方差和偏度都和该对数正态分布相同，但是峰度是 9.72（两倍的 4.86），基于 GB2 估计的期权价格就是$13.20。

Hull 和 White（1987）以及 Wiggens（1987）也研究了出现随机波动性的期权定价公式。既然 GB2 分布本身给出了混合解释，基于 GB2 的期权定价公式也就可以解释成是基于某种形式的随机波动性。

[4] Aitchison 和 Brown（1969，p.12）给出了对数正态分布的正规化不完全矩或矩分布的表达式。也可以参见（2.9）。<>

3.4.贝塔的估计：适应（adaptive）和部分适应（partial adaptive）估计、ARCH、GARCH 及其应用

回归分析是金融数据建模的重要工具。基本的线性回归模型定义为

$$Y_t = X_t\beta + \varepsilon_t \tag{3.24}$$

其中Y_t和X_t分别表示因变量的第 t 个观测值、解释变量的 1×K 向量，而β是未知的 K×1 常数向量。ε_t是随机扰动项，假定是独立同分布且具有 0 均值和常数方差：

$$\mathrm{E}(\varepsilon_t) = 0 \tag{3.25}$$

$$\mathrm{E}(\varepsilon_t^2) = \sigma_t^2 = \sigma^2$$

若假定当 n 无限增大时，$(X'X/n)$的极限是一个正定矩阵 C，其中$X' = (X_1' X_2' \ldots X_n')$，那么普通最小二乘（OLS）估计量$\hat{\beta}=(X'X)^{-1}X'Y$具有渐近分布$\left[N(\beta;\sigma^2 C/n)\right]$。如果随机扰动项是正态分布，那么最小二乘估计量就是有效估计量。然而，如果正态性假设不能满足，最小二乘估计量仍是所有线性无偏估计量中具有最小方差的估计量，但可能存在更有效的非线性估计量。众所周知，OLS 对在肥尾收益分布中经常遇到的异常值非常敏感。

在金融和统计文献中，大量对异常值不如 OLS 敏感的估计方法也已得到了研究。一种最经常运用的方法是最小绝对离差法（LAD），定义如下

$$\text{LAD:}\quad \boxed{\hat{\beta}_{\mathrm{LAD}} = \arg\min_\beta \sum_t \left|Y_t - X_t\beta\right|} \tag{3.26}$$

Basset 和 Koenker（1978）证明了，如果ε的概率密度函数$f(\varepsilon)$连续且在它的中位数处密度为正，那么这个估计量就是渐近正态的。对许多肥尾分布来说，已经证明 LAD 估计量比最小二乘估计量有效、至少是渐近有效的估计量；例如，参阅 Smith 和 Hall（1972），Kadiyala 和 Murthy（1977），及 Coursey 和 Nyquist（1983）。当随机扰动项服从拉普拉斯分布，LAD 就是最大似然估计量。Sharpe（1971）以及 Cornell 和 Dietrich（1978）用 LAD 来估计市场模型的β值。

L_p估计量，定义为

$$L_p \quad \boxed{\hat{\beta}_{Lp} = \arg\min_\beta \sum_t \left|Y_t - X_t\beta\right|^p} \tag{3.27}$$

给出了最小二乘估计量（p＝2）和 LAD 估计量（p＝1）的概括。一些关于L_p估计量的早期研究还介绍了 p 的取值问题；例如参阅 Hogg（1974）。

M 估计量是另一类可以包容可能的非正态性的估计量，定义为

$$\boxed{\hat{\beta}_{\mathrm{M}} = \arg\min_\beta \sum_t \rho\left((Y_t - X_t\beta)/\sigma\right)}$$

M：（3.28）

其中 σ 是分布的尺度估计量。函数 $\rho(\ )$ 对误差值分配权重。函数 $\Psi(\varepsilon)=\rho'(\varepsilon)$ 衡量随机扰动项在估计过程中的“影响”。当 $E(\Psi(\varepsilon))=0$ 且 $\mathrm{Var}(\Psi(\varepsilon))$ 有限，M 估计量将具有渐近正态分布。最小二乘估计量、LAD 估计量和 L_p 估计量都是 M 估计量的特例。Huber（1981）研究了另外的 M 估计量。M 估计量的关键问题是选择适当的 $\rho(\varepsilon)$ 函数。当选择 $\rho(\varepsilon)$ 为 $\{-\ln f(\varepsilon)\}$，M 估计量就是 MLE 估计量而且是有效估计量。Koenker（1982）对相关资料做精彩的综述。

因为很少能知道 $f(\varepsilon)$ 的形式，在文献中提出了两种方法。第一种方法，被称为“部分适应”，是要选择 $\rho(\varepsilon)$ 为灵活参数概率密度函数对数值的相反数，它包括了正态分布，而且允许肥尾和可能的非对称。Blattberg 和 Sargent（1971）假设稳定帕累托误差，Zeckhauser 和 Thompson（1970）基于幂指数或 BT 误差，都对部分适应过程进行了研究。另一种方法就是“完全适应”。核估计量或基于广义矩的方法是完全适应过程的例子。完全适应估计量，和基于误差实际分布的最大似然估计量一样渐近有效。但是，完全适应估计量不需要为在实践中遇到的样本容量问题展示相同的效率特性。

部分适应估计

在可能偏离正态性的情况下，BT、GT 和 EGB2 概率密度函数为估计回归模型提供了基础。BT 和 GT 是对称的分布，但允许存在不同的峰度。EGB2 的峰度变化范围不大，但允许对称和非对称的误差分布。为了阐明这些方法，先看一下从 Box-Tiao 概率密度函数（2.22）中得到的对数似然函数

$$\ell_{\mathrm{BT}}(\beta,\ \sigma,p)=n[\ln(p)-\ln(2\sigma\Gamma(1/p))]-\sum_t(|Y_t-X_t\beta|/\sigma)^p \qquad (3.29)$$

当 p＝1 或 2，分别对 β 最大化 $\ell_{\mathrm{BT}}(\ )$ 就得到 LAD 和 OLS。对 β 和 p 最大化 $\ell_{\mathrm{BT}}(\ )$ 内生出 p 的选择。肥尾误差分布倾向于 p 取小的值，而接近正态的数据倾向于 p 的估计值接近于 2。使用广义 t 分布不仅包含可以用学生 t 分布族的成员近似估计的误差分布，而且包含 Box-Tiao（幂指数族）—两者都包含正态分布。对有限的 q，Ψ_{GT} 是减函数且在估计过程中“调整”异常值。EGB2 分布族允许肥尾和非对称。对有限的 q ，Ψ_{EGB2} 是有界的，但不是减函数。

适应估计量—正态核估计量

回归参数的正态核估计量可以通过假设误差具有近似的概率密度函数

$$f(\varepsilon)=\left(\frac{1}{sN}\right)\sum_{n=1}^{N}\phi\left(\frac{\varepsilon-e_{nN}}{s}\right) \qquad (3.30)$$

而得到。其中 ϕ 和 e_{nN} 分别表示标准正态密度函数和最小二乘残差

$$e_{nN}=Y_n-X_n\hat{\beta}$$

$\hat{\beta}$ 是 β 的最小二乘估计量。S 是平滑参数。也可以引入调整参数，Hseih 和 Manski（1987）。McDonald

和 White（1993）做了一个小的蒙特卡罗模拟研究，比较了 LAD、OLS、部分适应（EGB2、BT、GT）、正态核、广义矩法估计量的有限样本特性。他们发现，由于几种非正态误差分布相对于正态误差分布来说，具有最小效率损失，适应估计量和部分适应估计量优于 OLS 和 LAD 估计量。此外，在非对称误差分布的情况下，EGB2 估计量优于其他所有估计量。

ARCH 和 GARCH 模型

金融中的许多应用都发现回归误差具有小残差群集（clusters）和大残差群集的特征，这是传统回归模型无法描述的。在这些应用中，大（小）残差倾向于跟随大（小）残差。这一经验发现提出了残差的自回归条件件异方差（ARCH）表达式，如

$$\varepsilon_t = u_t\left[\alpha_0 + \alpha_1\varepsilon_{t-1}^2\right]^{0.5} \tag{3.31}$$

其中 u_t 独立且同 N[0,1]分布。Engle（1982）指出，当 $\alpha_1 < 1$，有

$$\operatorname{Var}\left[\varepsilon_t \middle| \varepsilon_{t-1}\right] = \sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 \qquad \text{（3.32a－b）}$$

$$\operatorname{Var}\left[\varepsilon_t\right] = \alpha_0 / (1-\alpha_1)$$

这个模型（3.31）称为一阶 ARCH 模型，ARCH（1）。

当残差是 ARCH（1），即使残差是非正态的，OLS 估计量仍然是 β 的最佳线性无偏估计量；然而，它们在非线性估计量中将不是有效的。 p 阶 ARCH 模型，定义为

$$\text{ARCH}(p)\text{:}\quad \sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \ldots + \alpha_p\varepsilon_{t-p}^2 \tag{3.33}$$

Bollerslev（1986）提出了广义 ARCH（GARCH）模型，定义为

$$\begin{aligned}\text{GARCH}(p,q)\text{:}\quad \sigma_t^2 &= \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \ldots \\ &+ \alpha_p\varepsilon_{t-p}^2 + \delta_1\sigma_{t-1}^2 + \ldots + \delta_q\sigma_{t-q}^2\end{aligned} \tag{3.34}$$

GARCH 设定允许许多模型有简约参数化方法；它要求有高阶的 ARCH 模型。GARCH 公式允许方差以比 ARCH 模型更一般的方式随时间变化。Bollerslev 报告了 GARCH（1,1）模型的 0 直到 12 阶的矩稳定性条件。Greene（1993）对 ARCH 和 GARCH 模型作了全面的述评。Bollerslev、Chou 和 Kroner（1992）对理论和实证应用进行了详尽的综述。Nelson（1991）在他的 ARCH 和 GARCH 模型应用中，用 BT 作为灵活参数模型。EGB2 和 GT 公式会增加其灵活性。

部分适应和完全适应方法可以和 ARCH 及 GARCH 模型结合起来，解释某些实证金融应用中发现的非正态性（偏度或尖峰态误差分布）和群集现象。

市场模型的一个应用：（AMPCO）

使用 3.1 节给出的月收益率数据，估计股票的β值。因变量是 $Y = \ln\left((P_t + d_t)/P_{t-1}\right) - r_t$，其中 P_t 和 d_t 表示 AMPCO 在时期 t 的股票价格和股利，而 r_t 表示 30 日国库券的月收益率（无风险利率的代理变量）。自变量的构造是 X＝用价值加权的纽约股票交易所指数计算的月收益率减去无风险利率的对数。最小二乘估计的结果是

$$
\begin{aligned}
&\hat{Y} = -0.169 + 1.085X \\
&(t) \quad (-1.44) \quad (3.4) \\
&R^2 = .166 \\
&DW = 1.56 \\
&\text{Log}-\text{likelihood} = \ell = 60.62 \\
&\text{偏度}═1.56 \\
&\text{峰度}═8.7
\end{aligned}
$$

偏度和峰度的值表明误差服从正态分布的假设存在问题。这可以通过 Jarque-Bera 检验，也可以使用 6 组拟合优度检验得到证实。这一模型对误差使用 LAD、BT、GT、EGB2 和核设定重新进行了估计。表 6 报告了估计的结果。

表 6

估计β：AMPCO－月收益率

市场模型：$Y_t = \alpha + \beta X_t + \varepsilon_t$

（1988 年 1 月—1992 年 12 月）

	OLS	LAD	BT	GT	EGB2	核
α	-.0169	-.0186	-.0187	-.024	-.016	-.0193
β	1.085	1.176	1.187	.878	.993	1.149
p	2.000	1.000	1.11	6303.4	.984	
q	∞		∞	.0003	.552	
R^2	.166	.166	.165	.160	.165	.166
	60.6	65.3	65.4	69.1	66.9	－

BT、GT 和 EGB2 设定在对数似然值上提供了相对于正态性假设（即用最小二乘法）具有统计显著性的改进。在β值的估计上存在相当大的差异。只有 EGB2 和核估计量允许偏斜的误差分布。这些估计量的性质还需要进一步的研究。在 Butler 等（1990）及 McDonald 和 Nelson（1993）中可以发现两个部分适应估计（不是核估计）的应用。

我们估计了 45 家随机选取的公司每一家的β值。没有一家公司的误差项表现出严重的 ARCH 行为。这种行为在周收益率和日收益率中更可能观测到。

3. 5. 其他应用

灵活参数概率分布族的这些应用，仅仅是灵活参数分布具有广泛潜在用途的一个启示。在金融中的其他应用包括：具有 ARCH 或 GARCH 成分的 ARIMA 预测模型、定性反应模型和经济周期持续时间模型。这些模型很容易进行估计。另一个经常的应用是使基础分布的参数成为外生变量的可估计函数。这使得对所关注的分布的预期变动建模成为可能。

附录 A：特殊函数

本节将回顾论文主要部分中讨论的一些函数和符号。对这一领域额外的相关资料有兴趣的读者可以参考 Abramowitz 和 Stegun（1964）、Luke（1969）、Rainville（1960）和 Sneddon（1961），他们提供了有用的文献。

伽玛函数，$\Gamma(z)$，定义为

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt \qquad \text{（A.1）}$$

对实数(z)>0。对（A.1）分部积分得到递归关系

$$\Gamma(z)=(z-1)\Gamma(z-1) \tag{A.2}$$

Rainville（1960）推导出两个有用的结论是

$$\Gamma(.5)=\sqrt{\pi} \tag{A.3}$$

和

$$\Gamma(z)\to e^{-z}z^{z-.5}(2\pi)^{.5}\text{当}z\to\infty, \tag{A.4}$$

第二个结论被称为斯特林（stirling）逼近。

贝塔函数，$B(p,q)$，定义为

$$B(p,q)=\int_0^{\infty}t^{p-1}(1-t)^{q-1}dt \tag{A.5}$$

$$=\int_0^{\infty}\frac{t^{p-1}}{(1+t)^{p+q}}dt$$

其中 p 和 q 都为正。$B(p,q)$也可以用伽玛函数表示为

$$B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \tag{A.6}$$

文中讨论的累积分布函数可以用超几何序列表示，这种表示法通过 pochammer 符号可以得到简化

$$\begin{aligned}(a)_n&=(a)(a+1)(a+2)\ldots(a+n-1)\\&=\frac{\Gamma(a+n)}{\Gamma(a)},\text{当}1\le n\\&=1,\text{当}n=0\end{aligned} \tag{A.7}$$

广义超几何序列的定义如下

$${}_pF_q[a_1,a_2,\ldots,a_p;b_1,b_2,\ldots,b_q;x]=\sum_{i=0}^{\infty}\frac{(a_1)_i(a_2)_i\ldots(a_p)_i x^i}{(b_1)_i(b_2)_i\ldots(b_q)_i i!} \tag{A.8}$$

广义超几何序列有两个重要的特例，一个是合流超几何序列（p＝q＝1）

$${}_1F_1[a_1;b_1;x]=\sum_{i=0}^{\infty}\frac{(a_1)_i x^i}{(b_1)_i i!} \tag{A.9}$$

一个是超几何序列（p＝2，q＝1）

$${}_2F_1[a_1,a_2;b_1;x]=\sum_{i=0}^{\infty}\frac{(a_1)_i(a_2)_i x^i}{(b_1)_i i!} \tag{A.10}$$

作为说明这些函数灵活性的例子，指数函数 e^x 和二次展开式 $(1-x)^n$ 可以表达成广义超几何序列的特例

$$e^{x} = {}_1F_1\left[a;a;x\right]\text{和}\left(1-x\right)^{n} = {}_1F_0\left[-n;x\right]$$

论文中讨论的许多随机变量的累积分布函数，可以用不完全伽玛和不完全贝塔函数来表示。Rainville（1960，p.127）将不完全伽玛函数的定义为

$$\begin{aligned}\gamma_x\left(p\right) &= \int_0^x e^{-t} t^{p-1} dt \\ &= \left(\frac{x^p}{p}\right) {}_1F_1\left[p, p+1; -x\right]\end{aligned} \qquad \text{（A.11）}$$

Luke（1969，Vol 2，P.178）对不完全贝塔函数的定义是

$$B_x\left(p,q\right) = \int_0^x t^{p-1}\left(1-t\right)^{q-1} dt = \frac{x^p}{p} {}_2F_1\left[p, 1-q; p+1; x\right] \qquad \text{（A.12）}$$

附录 B

数据资料：

1. 选取的公司

1.Aileen Inc.
2.Aluminum Company Amer
3.American Home Products Corp.
4.Ampco-Pittsburg Corp.
5.Armatron International Inc.
6.Atlantic Energy Inc.N.J.
7.Becton Dickinson & Co.
8.Bethlehem Corp.
9.Brascan Ltd.
10.Brown Forman Inc.
11.Caterpillar Inc.
12.Cleveland Cliffs Inc.
13.Coastal Corp.
14.Cominco Ltd.
15.Crowley Milner & Co.
16.Curtiss Wright Corp.
17.Dole Food Co.
18.FPL Group Inc.
19.General Public Utils Corp.
20.Hapmpton Industries Inc.
21.Hershey Foods Corp.
22.KATV Industries Inc.
23.LVI Group Inc.
24.MEI Diversified Inc.
25.Manville Corp.
26.Masco Corp.
27.Mesabi Trust
28.Minnesota Power & Light
29.Nevada Power Co.
30.Nucor Corp.
31.Oneida Ltd.
32.Perkin Elmer Corp.
33.Proler International Corp.
34.Quaker State Corp.
35.Quantum Chemical
36.Rockwell International Corp.
37.Russell Corp.
38.Ryder Systems Inc.
39.SPS Technologies Inc.
40.Speed O Print Business Mach.
41.Thomas Industries Inc.
42.Union Pacific Corp.
43.Walgreen Co.
44.Wheeling Pittsburgh Corp.
45.Witco. Corp.

数据

AMPC	VMNYSE	短期国库券

1.064220	1.046050	1.002942
1.017241	1.048949	1.004556
0.932203	0.975659	1.004407
1.041818	1.010124	1.004616
0.921053	1.005238	1.005053
1.000000	1.048774	1.004853
1.024762	0.994076	1.005072
0.934579	0.971946	1.005938
1.050000	1.038680	1.006167
0.977143	1.023418	1.006101
0.990196	0.985484	1.005662
1.075248	1.018858	1.006341
1.074074	1.067892	1.005514
1.017241	0.980880	1.006131
1.050847	1.021679	1.006706
0.948387	1.047366	1.006748
1.042735	1.038815	1.007873
0.950820	0.997407	1.007093
0.970690	1.083675	1.006955
1.080357	1.020178	1.007392
0.958678	0.996213	1.006545
0.936207	0.972320	1.006765
0.953704	1.020250	1.006866
0.899029	1.021427	1.006069
0.858696	0.932313	1.005670
0.924051	1.013580	1.005679
1.041096	1.023607	1.006441
1.021053	0.973387	1.006873
0.948052	1.089550	1.006771
0.917808	0.994140	1.006251
0.994030	0.996466	1.006771
0.924242	0.912555	1.006572
0.786885	0.951331	0.005984
0.887500	0.992493	1.006818
1.119048	1.063259	1.005651
1.119149	1.028244	1.005989
1.038462	1.042467	1.005177
1.240741	1.072492	1.004767
1.074627	1.024085	1.004391
0.952778	1.002769	1.005335
0.955882	1.040264	1.004721
0.892308	0.957727	1.004171
0.975862	1.045763	1.004884

0.875000	1.024660	1.004610
1.244898	0.986335	1.004558
1.059016	1.014917	1.004246
0.937500	0.962219	1.003915
1.033333	1.106464	1.003792
1.041935	0.988254	1.003391
1.046875	1.011946	1.002828
0.955224	0.981095	1.003376
0.978125	1.023507	1.003249
0.919355	1.005404	1.002758
0.982456	0.984084	1.003201
1.028571	1.041030	1.003077
0.842105	0.980654	1.002605
1.000000	1.009555	1.002573
0.991667	1.006615	0.002286
1.042553	1.034350	1.002346
1.469388	1.014859	1.002823

致谢

作者在此向 Darin Clay 和 Julia Sunny 在研究上给予的资助以及 Scott Carson 和 Grant McQueen 对本论文初稿的评论表示感谢。

参考文献

Abramowitz M. and I. A. Stagun (1964). Hand book of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards, Applied Mathematics Series No. 55, Washington, D.C.

Aitchisoa, J. and J. A. C. Brown (1969). The Lognormal Distribution with Special References to Its Uses in Economics. Cambridge University press, Cambridge.

Akgiray, V. and G. G. Booth (1988). The stable-law model of stock returns .J. Business Econorn. Statist. 6(1), 51-57.

Ali, M. M. (1975). Stochastic dominance and portfolio analysis. J. Financ. Econom. 2, 205-229.

Arnold, B. (1983). Pareto Distributions. International Cooperative, Burtonsvillc, MD.

Arrow, J. K. (1965). Aspects of the Theory of Risk Bearing. Helsinki.

Atldnson, A. B. (1970). On the measurement of inequality. J. Econom. Theory 2, 24~63.

Basset, G. and R. Koenker (1978). Asymptotic theory of least absolute error regression,J.. Amer.-Statis. Assoc. 73, 618-622.

Bishop, J. A., S. Chakraborti and P. D. Thistle (1989). Asymptotically distribution free statistical inference for generalized Lorenz curves. Rev. Econom. Statist. 71, 725-727.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. J. Politic. Econom. 81, 637-659.

Blattberg, R. C. and N. J. Gonedes (1974). A comparison of the stable and student distributions as

statistical models
for stock prices. J. Business 47, 244-280.
Blattberg, R. and T. Sargent (1971). Regression with non-Gaussian disturbances: Some sampling results. Econometrics 39, 501-510.
Bollerslev, T. (1986). Generalized autoregresaive conditional heteroscedasticity. J. Econometrics 31, 307-327.
Bollerslev, T., R. Y. Chou and K. F. Kroner (1992). ARCH modeling in finance. J. Econometrics 52, 5-59.
Bookstaber, R. M. (1987). Option Pricing and Investment Strategies. Probus Publishing Co., Chicago.
Bookstaber, R. M. and J. B. McDonald (1987). A general distribution for describing security price returns. J.
Business 60, 401-424.
Butler, R. J. and 3. B. McDonald (1989). Using incomplete moments to measure inequality. J. Econometrics 42, 109-
119.
Butler, R. J., J. B. McDonald, R. Nelson, and S. White (1990). Partially adaptive estimation of regression models.
Rev. Econom. Statist. 72, 321-327.
Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices.
Econometrics 41, 135-155.
Cornell, D. and J. K. Dietrich (1978). Mean-absolute-deviation versus least squares regression esti-mation of beta
coefficients. J. Financ. Quant. Anal. 13, 123-131.
Coursey, D. and H. Nyquist (1983). On least absolute error estimation with linear regression models with dependent
stable residuals. Rev. Econom. Statist.65, 687-692.
Cox, J. C. and S. A. Ross (1976). The valuation of options for alternative stochastic processes. Financ. Econom. 3,
145-166.
Elderton, Sir W. P. and N. L. Johnson (1969). Systems of Frequency Curves. Cambridge University Press, London.
Elton, E. J. and M. J. Gruber (1974). Portfolio theory when investment relatives are lognormally distributed. J.
Finance 29, 1265-1273.
Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom
inflations. Econometrics 50, 987-1008.
Fuma, E. F. and R. Roll (1968). Some properties for symmetric stable distributions. J. Amer. Statist.
Assoc. 63, 817-836.
Fishburn, P. C. (1964). Decision and Value Theory. Wiley, New York.
Greene, W. H. (1993). Econometric Analysis. Macmillan, New York.
Hagerman, R. L. (1978). More evidence on the distribution of security returns. J. Finance 33, 1213-1221.
Hanoch, G. and H. Levy (1969). The efficiency analysis of choices involving risk. Rev. Econom. Stud.

36, 335-346.

Hirsehberg, J., S. Mazumdar, D. Slottje and G. Zhang (1992). Analyzing functional forms of stock returns. J. Appl. Financ. Econom. 2(4), 221-227.

Hogg, R. V. (1974). Adaptive robust procedures: A partial review and some suggestions for future applications and theory, d. Amer.Statist. Assoc. 69, 909-927.

Hsieh, D. A. and C. F. Manski (1987). Monte Carlo evidence on adaptive maximum likelihood estimation of a regression. Ann. Statist. 15, 541-551.

Huber, P. J. (1981). Robust Statistics. Wiley, New York.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. J. Finance. 52, 281-300.

Johnson, N. L. and S. Kotz (1970). Continuous Univariate Distributions. Vol. 2. Wiley, New York.

Kadiyala, K. R. and K. S. R. Murthy (1977). Estimation of regression equation with Cauchy disturbances. Canad. J. Statist. Section C: Applications. 5, 111-120.

Kakwani, N. C. (1984). Welfare Rankings of Income Distributions. Advances in Econometrics, 3. Edited by R. L. Basmann and G. F. Rhodes. Greenwich, Conn., JAI Press.

Kalbfleisch, J. D. and R. L. Prentice (1980). The Statistical Analysis of Failure Time Data Wiley, New York.

Kendall, M. G. and A. Stuart (1969, 1967). The Advanced Theory of Statistics, Vol. I and II. Griffin, London.

Koenker, R. (1982). Robust methods in econometrics. Econometric Rev. I, 213~55.

Lau, A. H., H. Lau and J. R. Wingender (1990). The distribution of stock returns: New evidence against the stable model. J. Business Econom. Statist. 8, 217-223.

Lau, H., J. R. Wingender and A. H. Lau (1989). On estimating skewness in stock returns. Mgmt. Sci. 35(9), 1139-1142.

Lehmann, E. L. (1959). Testing Statistical Hypotheses. Wiley, New York. 74-75.

Luke, Y. L. (1969), The Special Functions and their Approximations. Vol. I and II. Academic Press, New York.

Mandelbrot, B. (1963). The variation of certain speculative prices. J. Business: 36, 394 419.

Markowitz, H. M. (1959). Portfolio Selection. Wiley, New York.

McDonald, J. B. (1984). Some generalized functions for the size distribution of income. Econometrics. 52, 647-663.

McDonald, J. B. and R. M. Bookstaber (1991). Option pricing for generalized distributions. Communications in Statistics.' Theory and Methods. 20(12), 4053-4068.

McDonald,J. B. and R. J. Butler (1987). Some generalized mixture distributions with an application to unemployment duration. Rev. Econom. Statist. 69, 232-240.

McDonald, J. B. and R. Nelson (1993). Beta estimation in the market model: Skewness and Leptokurtosis. Comm. Statist. 22:10

McDonald, J. B. and W. K. Newey (1988). Partially adaptive estimation of regression models via the generalized T distribution. Econometric Rev. 12, 103-124.

McDonald, J. B. and S. B. White (1993). A comparison of some robust, adaptive, and partially adaptive estimators of regression models. Econometric Rev. 12, 103-124.

McDonald, J. B. and Y. J. Xu (1995). A generalization of the beta distribution with applications. J. Econometrics, 66, 133-152.

Nair, U. S. (1936). The standard error of Gini's mean difference. Biometrika. 38, 428-36.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrics. 59, 347-370.

Officer, R. R. (1972). The distribution of stock returns. J. Amer. Statist. Assoc. 67, 807--812.

Ord, J. K. (1972). Families of Frequency Distributions. Griffin, London.

Patil, G. P., M. T. Boswell, and M. V. Ratnaparkhi (1984). Dictionary and Classified Bibliography of Statistical Distributions in Scientific Work. International Cooperative Publishing, Burtousville, MD.

Pearson, K. (1895). Memoir on skew variation in homogeneous materials. Phil. Trans. Roy. Soc.. A. 186 343-414.

Pearson, K. (1901). Supplement to a memoir on skew variation. Phil. Trans. Roy. Soc.. A. 197, 443-459.

Pearson, K. (1916). Second supplement to a memoir on skew variation. Phil. Trans. Roy. Soc.. A. 216, 429-457.

Pope, R. D. and R. F. Ziemer (1984). Stochastic efficiency, normality, and sampling errors in agricultural risk analysis. Amer. J. Agri. Econom. 66, 31-40.

Praetz, P. D. (1972). The distribution of share price charges. J. Business. 45, 49-55.

Pratt, J. W. (1964). Risk Aversion in the Small and Large. Econometrica. 122-136.

Quirk, J. P. and R. Saposnik (1962). Admissibility and measurable utility functions. Rev. Econom. Stud.

Rainville, E. D. (1960). Special Functions. MacMillan, New York.

Salem A, B. and T. D. Mount (1974), A convenient descriptive model of income distribution: The gamma density. 42, 1115-1127.

Shalit, H. and S. Yitzhaki (1984). Mean-Gini, portfolio theory, and the pricing of risky assets. J. Finance 39, I449-1468.

Sharpe, W. F. (1971). Mean-absolute-deviation characteristic Lines for securities and portfolios. Mgmt. Sci. 18, 1-13.

Shorrocks, A. F. (1983). Ranking income distributions. Economica 3-17.

Singh, S. K. andG. S. Maddala (1976). A function for the size distribution of incomes. Econometrica 44, 963-973.

Sneddon, I. N. (1961). Special Functions of Mathematical Physics and Chemistry, 2nd ed., Interscience Publishers, Edinburgh.

Smith, V. K. and T. W. Hall (1972). A comparison of maximum likelihood versus BLUE estimators. Rev. Econom. Statist. 54, 186-190.

Stuart, A. and J. K. Ord (1987). Kendall's Advanced Theory of Statistics, Vol, 1. Oxford Press, New York.

Taillie, C. (1981). Lorenz ordering within the generalized gamma family of income distributions. Statistical Distributions in Scientific Work. 6, 181-192.

Taylor, S. (1986). Modeling Financial Time Series. Wiley, New York.

Tobin, J. (1958). Liquidity preference as behavior towards risk. Rev. Econorn. Stud. 25, 65-68.

Von Neumann, J. and O. Morgenstern (1953). Theory of Games and Economic Behavior. 3rd ed., Princeton Press, Princeton.

Wiggens, J. B. (1987). Option values under stochastic volatility. J. Financ. Econom. 19, 351-372.
Wilfing, B. (1992). A sufficient condition for Lorenz domination of generalized beta income distributions of the second kind. University of Dortmund, Mimeo.
Wilting, B. and W. Kramer (1993). The Lorenz-ordering of Singh-Maddala income distributions. Econom. Lett. 43, 53-57.
Yitzhaki, S. (1982). Stochastic dominance, mean variance and Gini's mean difference. Amer. Econorn. Rev. 72, 178-85.
Zeckhauser, R. and M. Thompson (1970). Linear regression with non-normal error terms. Rev. Econorn Statist. 52, 280-286.