跨品种期货套利交易最优保证金比率设计
——基于 Copula 函数及极值理论的研究

赵成珍，宋锦玲

(中央财经大学金融学院，北京 100081)

摘 要：文章主要研究期货交易中最优风险保证金的比率设定问题，通过考虑变动保证金的设定来弥补目前固定保证金不足的问题。在期货套利交易保证金问题上，文章给出了商品期货套利交易保证金的测定理论。在测定的过程中，引入了收益率的非对称 Laplace 分布和 GARCH-T 函数分布，在 Gumbel Copula 函数的基础上，应用蒙特卡洛模拟算法对两种商品的套利交易的保证金问题给出了测定，并以豆油和豆粕期货为例，选取时间序列进行了实证研究，得出结论：套利交易的保证金水平在理论上小于非套利交易下两单独品种保证金的收取之和。此外再结合极值理论，在改进 VAR 的基础上，得出在极端情况下的套利交易的最优保证金比率设计。

关键词：期货交易；套利交易；套利品种；金融收益率

中图分类号：F724.5　　**文献标识码：**A　　**文章编号：**1004-292X(2014)12-0016-04

The Optimal Margin Ratio Design of Across Varieties Futures Arbitrage Trading

ZHAO Cheng-zhen，SONG Jin-ling

(School of Finance，Central University of Finance and Economics，Beijing 100081，China)

Abstract: This paper mainly studies the setting problem of the optimal risk trading margin ratio in futures trading，by changing the variation margin setting to make up the insufficient of the fixed deposit. As for the futures arbitrage trading margin issue，this paper first presents the determination theory of commodity futures arbitrage trading margin. In the process of determination，the yield of asymmetric Laplace distribution and GARCH-T function distribution are introduced. On the base of the Gumbel Copulas function，the margin ratio determination of the two goods by the Monte Carlo Simulation Algorithm is provided. The paper takes soybean oil and soybean meal futures as an example and does empirical research on the time series of them and draws the conclusion: in theory，the arbitrage trading margin level is less than the sum of two separate varieties deposit charge in the carry trade. In addition，the paper applies the extreme value theory and concludes that in extreme cases of arbitrage，the optimal margin ratio design on the basis of the improvement of VAR.

Key words: Futures Trading；Arbitrage Trading；Arbitrage varieties；Yield of finance

20 世纪 80 年代以来，全球多次爆发金融危机，金融机构的风险控制越来越得到学术界和业界的重视。始于 2007 年的美国次贷危机席卷全球，发达国家的许多百年金融老店纷纷倒闭转型。究其原因，有金融产品的因素，但是更重要的是金融风险没有得到合理的控制。期货公司施行的交易保证金制度在期货交易中起到了杠杆作用，但同时也大幅放大了期货公司的风险。因此合理设置保证金水平至关重要。保证金设置水平的高低直接影响了市场的流动性，进而影响市场的效率。如果期货公司为降低投资者的违约风险而设置偏高的保证金，将会导致交易成本上升，市场流动性以及运行效率会随之降低；反之，若期货公司设置偏低的保证金水平，交易成本的下降使得市场参与主体增多，期货交易市场的流动性也会变大，但交易价格的波动性和投资者的违约风险也都会相应上升，期货市场的效率也因此而受到损害。此外，如何衡量在极端条件下保证金的合理水平，即引入极值条件下的保证金测算水平也是一个具有重要理论和实践意义的话题。

国际上的期货交易所主要采用的是弹性交易保证金制度，这种制度依据期货交易价格的波动性以及其他相关因素等，运

收稿日期：2014-08-18

作者简介：赵成珍 (1983-)，男，山东滨州人，博士研究生，研究方向：金融风险管理；
宋锦玲 (1983-)，女，河北承德人，博士研究生，研究方向：金融机构与金融市场。

用风险计量模型和综合因素分析的方法对交易保证金比例进行不定期调整，比如美国芝加哥期货交易所(CBOT)、美国芝加哥商业交易所(CME)、英国伦敦国际金融期货交易所(LIFFE)以及香港交易所(HKEX)。目前，国外已开展了动态保证金比率的研究。由于我国期货市场起步较晚，我国主要采用静态的保证金比率制度。依据经验，一般以三个涨跌停板为收取标准。另外，在保证金的比率设置上采用的计量方法主要是 VaR 方法。我国目前期货市场的交易保证金比率设置方法主要停留在原始的 VaR 阶段，风险计量效果不令人满意，因为存在以下几个问题：第一， VaR 方法自身存在一些问题；第二，收益率的正态分布假设不符合现有金融数据的尖峰厚尾的特征；第三，期货价格波动率的聚类特征及其时变性没有得到体现。

因为套利交易风险相对较少，期货公司会收取较少的保证金，但对于具体收取多少保证金这个问题，到目前为止国内外尚无人给出一个科学合理的推理。国内外期货公司主要靠经验收取的模式，难以得到客户的信服。现有关于期货最优保证金比率研究主要是应用国外的股指期货，但是由于各个国家的金融环境和风险因素不同，在实证过程中就会容易造成收益函数的拟合分布呈现出各自的特点。不同市场的分布函数不同，因此不能直接套搬套用。此外，国内相关研究主要采用沪深指数进行实证分析，这在实际操作中会出现一些问题，比如如何选用具体的股票来同指数匹配，如果按照沪深指数的实际成份股来选择会造成样本太多，从而不能分析两两之间套利的问题。

文章主要研究期货交易中最优风险保证金的比率设定问题，通过研究变动保证金比率的设定来弥补目前固定保证金比率的不足。文章借鉴 SPAN 系统，并对现有模型予以改良，得出合理的风险控制保证金比率模型。文章余下部分安排如下：第二部分给出如何在 Copula 函数及极值理论的框架下对期货的套利交易保证金给出模型设计；第三部分是对模型的实证分析；第四部分是文章的结论。

一、模型设计

1. 套利交易保证金

套利交易的关键是套利品种之间的相关性。在套利交易中，一般进行多空的反向操作，因此要求两品种之间的相关性要高。在实际操作中，期货市场品种可分为金属板块、能源化工板块和农产品板块。文章在组合的风险测度以及风险的相关性方面摒弃传统的线性相关，通过引进 Copula 函数来度量板块内的相关性，从而为整体板块的保证金比率测定打下基础。

在组合的套利保证金方面，主要采用 Var 或 CVar 的保证金测度方法。文章模拟套利交易采用一手品种 A 对冲一手品种 B，由于是多空双方同时在做，因此组合的风险度量为：

$$f=\mathrm{Var}_a+\mathrm{Var}_b$$

由于在计算 Var 的过程中，金融收益率的尖峰厚尾特征及非对称分布的特征，因此文章引进了边缘分布为非对称 Laplace 收益率分布与边缘分布为 GARCH-T 收益率函数分布。

2. 非对称 Laplace 边缘分布及 GARCH-T 分布的 Copula 模拟

(1) Gumbel Copula 函数。一般称如下形式的 Copula 函数为 Gumbel Copula 函数：

$$C(u,v)=\exp\left\{-[(-\ln u)^{\theta}+(-\ln v)^{\theta}]^{\frac{1}{\theta}}\right\},\ \theta\geq 1 \tag{1}$$

Gumbel Copula 函数的一个特点是，它对变量在分布上尾处的变化非常敏感，因此能够快速捕捉到上尾的相关变化，该函数可以描述具有上尾相关特征的金融变量之间的相关关系，其密度函数具有非对称性，密度分布呈“J”字型，即上尾高，下尾低。参数 θ 描述了变量的相关程度，当 θ=1 时，变量之间的关系是独立的；当 θ→∞时，变量之间完全相关。

(2) 非对称 Laplace 分布。金融资产的收益序列分布呈现尖峰厚尾的特征，因此，传统的正态分布不能拟合或拟合的误差较大。文章采用非对称 Laplace 分布来拟合金融资产收益序列的情况。如果 X 是一个随机变量，它的分布形式如式(2)，我们就称随机变量 X 服从非对称 Laplace 分布(Asymmetric Laplace Distribution)，记为 $X\sim AL(\mu,\sigma,p)$。

$$F(x/\mu,\sigma,p)=\begin{cases} p\exp\left(\dfrac{k}{p\sigma}(x-\mu)\right), & x\leq\mu \\ 1-(1-p)\exp\left(-\dfrac{k}{(1-p)\sigma}(x-\mu)\right), & x>\mu \end{cases} \tag{2}$$

其中，$k=\sqrt{p^2+(1-p)^2}$，μ 表示位置参数，σ 表示标准差，p 表示形状参数 (其中，$0<p<1$)，控制随机变量分布的偏度和峰度，p 的不同取值使得偏度为正或为负。当 $p<0.5$时，密度函数向右偏，即偏度为正；当 $p=0.5$ 时，密度函数是对称的，也就是传统的 Laplace 分布；而当 $p>0.5$ 时，密度函数向左偏，即偏度为负。

(3) GARCH-T 函数分布。GARCH 模型的数学表达式如下所示：

$$\begin{cases} \varepsilon_t=h_t^{\frac{1}{2}}\xi_t \\ h_t=\alpha_0+\sum_{i=1}^{q}\alpha_i\varepsilon_{t-1}^2+\sum_{i=1}^{p}\beta_i h_{t-1}^2+\xi_t \\ \xi_t\propto t_v \end{cases} \tag{3}$$

(4) 具有非对称 Laplace 边缘分布以及 GARCH-T 分布 Copula 函数的蒙特卡洛模拟。具体通过以下步骤完成：

Step1：产生两个(0,1)区间上的均匀分布随机数 u,w；

Step2：令 $x=F_1^{-1}(u)$，$F_1(\cdot)$为品种 1 的的边际分布函数；

Step3：对给定的 Copula 函数求得第二个序列在均匀分布上的随机数 $v=C_u^{-1}(w)$，文章将按照此方法来产生在 Gumbel Copula 函数相依结构下的另一模拟值；

Step4：根据 Step3 中计算的随机数 v 可以求出品种 2 的随机数 $y=F_2^{-1}(v)$，F_2 为品种 2 变量的边际分布函数；

Step5：将 Step2 和 Step4 中产生的随机数(x,y)，计算 x,y 的套利组合；

Step6：将 Step1～Step5 重复做 N 次就可以计算出不同显著性水平 α 下的 VaR 值。

此处的组合采用了等权重交易，文章模拟套利交易采用一手豆油对冲一手豆粕。由于是多空双方同时在做，因此组合的风险度量为：$f=\mathrm{Var}_{rdy}+\mathrm{Var}_{rdp}$。

3. 极值理论及其在期货套利交易中的应用

极值理论是测量极端市场条件下风险损失的一种方法，它的优势是具有超越样本数据的估计能力，并可以准确地描述分布尾部的分位数。极值理论主要包括两类模型：BMM 模型(Block Maxima Method)和 POT 模型(Peaks Over Threshold)。由 Pickands-Balkama-de Haan 定理可知，对于充分大的阈值 u，超额数的分布函数可以用广义帕累托分布近似拟合。

(1) 广义帕累托分布(GPD)：

$$G_{\xi,\beta}(x)=\begin{cases}1-\left(1+\dfrac{\xi x}{\beta}\right)^{-\frac{1}{\beta}}，\xi\neq 0\\ 1-\exp\left(-\dfrac{x}{\beta}\right)，\xi=0\end{cases}\tag{4}$$

其中，ξ 是重要的形状参数，而 β 是分布的尺度参数，且 $\beta>0$。当 $\xi\geq 0$ 时，$x\geq 0$；当 $\xi\leq 0$ 时，$0\leq x\leq\frac{-\beta}{\xi}$。$\xi\geq 0$ 时广义帕累托分布是厚尾的，这种情形是与风险测量最相关的，也是最需要测量的情形。

GPD 分布的拟合：确定阈值 u 之后，可以用 $\frac{n-N_u}{n}$ 作为 $F(x)$ 的估计值，N_u 是超过阈值的样本个数，n 是样本容量。利用大于阈值的样本来估计出参数 $\hat{\xi},\hat{\beta}$，则原来的方程 $F(x)=(1-F(u))G_{\xi,\beta}(x-u)+F(u)$，$x>u$ 的尾部估计变为：

$$\hat{F}(x)=(1-\frac{n-N_u}{n})G_{\hat{\xi},\hat{\beta}}(x-u)+\frac{n-N_u}{n}=1+\frac{N_u}{n}(G_{\hat{\xi},\hat{\beta}}(x-u)-1)$$

$$=1-\frac{N_u}{n}\left(1+\frac{(x-u)\hat{\xi}}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}}\tag{5}$$

(2) 极值理论中阈值的确定和参数估计。阈值 u 的确定非常关键，如果阈值 u 过高会导致超额数据的样本太小，估计参数的方差会偏高；而太小的 u 值则会产生有偏的估计量。通常有两种方法确定阈值 u：一是根据超额均值函数 MEF 的性质，即选取充分大的 u 值，使得 $X>u$ 时，MEF 是近似线性的；二是根据 Hill 图来确定。

令 $X(1)>X(2)\cdots>X(n)$表示独立同分布的有序数据；尾部指数 Hill 的统计量定义为：$H_{k,n}=\frac{1}{k}\sum_{i=1}^{k}\ln\left(\frac{X(i)}{X(k)}\right)$；Hill 图定义为点 $\{(k,H_{k,n}^{-1}),1<k<n-1\}$的集合。阈值 u 选择图形中尾部指数的稳定区域起始点的横坐标 k 所对应的数据 X_k。

(3) 基于极值分布理论的保证金比率的测算。根据 CVaR 计算的极值分布的最优保证金比率的计算：

$$C\text{VaR}_p=\text{VaR}_p+E(X-\text{VaR}_p|X>\text{VaR}_p)\tag{6}$$

其中，$E(X-\text{VaR}_p|X>\text{VaR}_p)$为给定 VaR_p 的超额均值函数 MEF(Mean Excess Function)，当 $\xi<1$ 时，GPD 的超额均值函数可表示为：

$$e(z)=E(X-z|X>z)=\frac{\beta+\xi z}{1-z}\beta+\xi z\tag{7}$$

$$\text{则：}C\text{VaR}=\text{VaR}_p+\frac{\beta+\xi(\text{VaR}_p-u)}{1-\xi}=\frac{\text{VaR}_p}{1-\xi}+\frac{\beta-\xi u}{1-\xi}\tag{8}$$

若给定概率 $P>F(u)$，通过反解 $\hat{F}(x)=1-\frac{N_u}{n}(1+\frac{(x-u)\hat{\xi}}{\hat{\beta}})$ 可以计算得到 P 分位数的估计：

$$\text{VaR}_p=\begin{cases}u+\dfrac{\hat{\beta}}{\hat{\xi}}\left(\left(\dfrac{n}{N_u}(1-p)\right)^{-\xi}-1\right)，\xi\neq 0\\ u-\beta\ln\dfrac{n}{N_u}(1-p)，\xi=0\end{cases}\tag{9}$$

$$\text{此时 } f=C\text{Var}_{rdy}+C\text{Var}_{rdp}\tag{10}$$

二、Copula 函数及极值理论在套利交易保证金中模型实证

期货品种的性质越相似，其相关性越强。目前的农产品市场，相对相似性最好的品种有豆 1 (非转基因) 和豆 2 (转基因)，豆油和豆粕 (都为大豆的加工品)，但豆 1 和豆 2 因为涉及转基因的问题，也存在一定的差异性，因此文章的实证也采用豆油和豆粕两个品种。

1. 数据的预处理

关于数据选取方面，文章选取最活跃合约的方法来构造连续期货价格合约数据。连续合约是指主力合约以及换月后主力合约，对主力合约的判断主要来源于持仓量和交易量这两个指标，并且以持仓量为先。豆油和豆粕的期货价格具有有较强的替代性，因此文章选取这两类期货产品的价格进行实证分析。选取从 2004 年 1 月 1 日到 2009 年 6 月 30 日共 1112 个交易日的数据。数据来源于富远期货交易行情软件。

$$\text{序列的收益率：} r1=\ln(b1)-\ln(b1(-1))\tag{11}$$

一些文献资料采用滞后三天的对数价格作为收益率的数据序列。但是采用滞后三天的数据虽然能对涨跌停板的作用起到一定的消除作用，但在实证研究过程中往往会放大风险。因此文章采用滞后一天的对数价格来拟合收益率。文章主要用 Eviews6.0 和 Excell 来处理和回归数据。

2. 非对称 Laplace 分布参数估计和 GARCH-T 函数参数估计

文章首先对豆油和豆粕的收益率的分布进行检验，表 1 是分布的经验统计数据。

表 1 豆油与豆粕分布的经验统计数据

	均值	方差	偏度	峰度	JB
豆油	0.000465	0.017555	0.5333804	4.676569	133.3352
豆粕	0.000123	0.016027	1.346684	19.79644	13395.63

通过数据的正态性检验可以看出，两者的收益明显地不符合正态。文章利用前面提到数据和收益率的处理方法，对非对称的 Laplace 分布参数进行了处理和参数估计，豆油和豆粕的非对称分布的参数估计表 2 所示：

表 2 豆油和豆粕的非对称分布的参数估计

估计参数	P	K	$1-2P$	σ	$E(x)$	μ
豆油	0.5641	0.4359	−0.1282	0.017555	0.000465	0.005627998
豆粕	0.6836	0.3164	−0.3672	0.016027	0.000123	0.018723235

从估计的值来看，结果显示是右偏的，得出的结论与实际的表 2 所得到的经

表 3 GARCH-t 模型的参数估计

GARCH-t 模型	α_0	α_1	β_1	DOF
豆粕	9.71E−07 (0.000234)	0.029218 (0.00456)	0.968104 (0.009845)	3.565415 (0.0000)
豆油	7.20E−07 (0.3407)	0.094091 (0.0000)	0.914664 (0.0000)	5.330153 (0.0000)

验统计数据相符。GARCH-T 模型的参数估计如表 3 所示，其中括号内的值为 P 值。

从回归的结果来看，除了豆油的均值外，其他都在 1%的显著水平下显著。而豆油的均值则可认为是 0。

3. Copula 函数的参数估计

由于文章的套利研究用的是两列收益序列，因此可以用一些简便的方法估计 Copula 函数的参数，文章采用 Genest and Rivest 的方法，对 τ 的估计结果如表 4 所示。

表 4　估计的 τ 结果

待估参数	τ	θ	1/θ
估计值	0.467	1.876173	0.533

4. 基于 Copula 函数的模拟

对于非对称 Laplace 边缘分布，利用前面讲到的算法，文章模拟了 10000 次，结果如表 5 所示。由于有了非对称 Laplace 分布的蒙特卡洛模拟做基础，GARCH-T 模型和上述过程基本类似，此处对于参数不再过多说明，只是给出估计的参数结果，见表 6。

表 5　非对称 Laplace 分布套利交易保证金模拟结果

	α=95%	α=99%	组合 α=95%	组合 α=99%
豆油	0.045397644	0.073960587	0.083534	0.138086832
豆粕	0.03813735	0.064126265		

表 6　GARCH-t 分布套利交易保证金模拟结果

	α=95%	α=99%	组合 α=95%	组合 α=99%
豆油	0.03833569	0.06946087	0.07823495	0.13970766
豆粕	0.03989926	0.07024679		

5. 基于极值理论保证金比率

文章结合经验分布总结出门限的选择，用极大似然估计法求得 ξ,β。由 GPD 分布函数，可求出其密度函数为：

$$f_u(y)=\frac{\xi}{\beta}\left(1+\xi\frac{y}{\beta}\right)^{-\frac{1}{\xi}-1} \tag{12}$$

当 $\xi<0$ 时，$0\leq x\leq-\frac{\beta}{\xi}$，其对数似然函数为：

$$L(\xi,\beta;\hat{x})=\begin{cases} n\ln\xi-n\ln\beta-\left(\frac{1}{\xi}+1\right)\sum_{i=1}^{n}\ln\left(1+\frac{\xi}{\beta}\hat{x}_i\right), & \xi\neq0 \\ -n\ln\beta-\frac{1}{\xi}\sum_{i=1}^{n}\ln\hat{x}_i, & \xi=0 \end{cases} \tag{13}$$

其中的 $\hat{x}_i$ 为 x 的观测值。由似然函数可推导出似然方程，从而求出 ξ,β 的最大似然估计值。对于豆粕和豆油，文章选择 0.019947283 和 0.0245012，估计的参数结果见表 7。

表 7　豆油和豆粕的 GPD 参数计算结果

	ξ	β	ML
豆油	−0.1692 (−1.1752)	−0.2225 (2.1118)	248.9
豆粕	−0.2225 (2.1118)	0.0140 (6.5620)	272.1

文章应用得到的数据计算尾部的 VaR，并在 VaR 的基础上计算得到结果（见表 8）。

通过比较可以看出，由 CVaR 方法计算的准备金大于 VaR 计算得到的准备金。前面提到了 VaR 计算的缺陷，因此文章为了安全选择高者，即在考虑极端情况下，应该将套利交易保证金得水平设置得高一些。

表 8　极值情况下套利保证金的计算结果

	95%VAR	95%CVAR	组合 α=95%
豆油	0.044412	0.044412+0.005743 =0.050155	0.104277
豆粕	0.04895	0.04895+0.006172 =0.054122	

三、结论

文章在验证了商品收益率序列不服从正态分布的情况下，引入了有偏的非对称 Laplace 分布和 GARCH-T 来拟合收益率的边际分布。考虑到商品之间的相关性，文章引入了非线性相关系数的度量。此外，在最为关键的商品套利问题的研究上，文章引入了 Copula 函数使得两商品之间的内部关系能够明确度量。在具体的保证金核算方面，还引入了蒙特卡洛算法进行模拟，文章得出的结论与经验相符。

具有内部联系的两个相关期货品种组合的风险测定并不是两个单品种的简单加减关系。从数据模拟的结果来看，在采用了 Copula 函数的基础上的组合风险较小，在 99%的置信水平下，1%左右的保证金基本上能够覆盖所有的风险，这也符合经济学原理：套利组合的收益率水平较低，组合的风险也较小。在实际操作中，期货公司收取单品种的保证金比率一般在 10%~15%之间，而对非套利的商品交易，两者的保证金比率可能会在 20%~30%之间，这大大的限制了保证金的使用效率，同时也降低了期货市场的流动性。

在测定完正常风险的基础上，文章还给出了极值理论的测定问题。从分析的结果来看，要想在较高的置信水平下消除极值现象仍需要较高的风险准备水平，较高的准备金加上正常的保证金大约在 10%左右，这解释了为何国内的期货公司的保证金比率基本在 10%以上，10%的保证金不但能覆盖正常风险所需的保证金，还可以消除极值理论带来的不利影响，而不同之处是所收取的保证金包括了保证金和风险准备金。

文章的理论和实证分析都得出结论：对于相关性较强的两商品之间进行套利，由于其套利的风险较小，因此，收取的保证金应该小于两者之和。

【参考文献】

[1] Fabozzi F J, Focardi S M, Jonas C L. On the challenges in quantitative equity management [J]. Quantitative Finance, 2008, 8(7): 649-665.

[2] K. Trianktafyllopoulos, G. Montana. Dynamic modeling of mean-reverting spreads for statistical arbitrage [J]. Compute Management Sciences 2008, Berlin: Springer-Verlag, 2009: 23-49.

[3] Lee M C. Using support vector machine with a hybrid feature selection method to the stock trend prediction [J]. Expert Systems with Applications, 2009, 36(8): 10896-10904.

[4] 刘志东，徐淼. 基于 Copula 的资产组合风险价值模拟方法 [J]. 统计与决策，2009(3): 17-20.

[5] 单国莉. 一种确定最优 Copula 的方法及应用 [J]. 山东大学学报，2005.

[6] 任仙玲，张世英. 基于核估计及多元阿基米德 Copula 的投资组合风险分析 [J]. 管理科学，2007(10): 92-97.

[7] Embrechts P, McNeil A J, Straumann D. Correlation and dependence in risk management: Properties and pitfalls [J]. Risk Management: Value at Risk and Beyond, 2002.

（责任编辑：LI）