上证综合指数 VaR 的度量

陈守东　王鲁非

摘　要　本文应用方差—协方差法及历史模拟法计算了上海证券交易所综合指数的 VaR 值。对计算结果进行 Kupiec 似然比检验后可以看出 GARCH 模型得出的上证指数收益率 VaR 值最有效，其他方法由于分布假设及所用方法的缺陷，使得计算结果不能或只能在少数置信水平下才能通过检验。

关键词　上证指数　方差—协方差法　VaR 值　Kupiec 似然比检验

一、统计描述

我们选取了从 1997 年 5 月 21 日至 2000 年 6 月 27 日的上证综指 750 个日收盘收益率数据（几何收益率）。在这 3 年左右的时间里，股指由 1354.88 点升至 1942.89 点，平均年增长率为 12.76%，下表为数据的一些描述统计量

观测值个数	750
均值	0.0604%
中位数	0.0396%
最大值	8.6656%
最小值	-8.7268%
标准差	1.6779%
偏度	-0.111442
峭度	7.037

从数据的描述统计量中我们可以看出平均收益率接近于 0，或者说与其标准差相比可以忽略不计。收益率的最大值与最小值的绝对值都小于 10%，这种情形发生在我国股票市场实行涨跌停板制度之后。

从数据的相关图我们可以看出，数据直到滞后 8 阶都不存在自相关（表 1），对数据进行滞后 4 阶、带趋势项及截距的单位根检验，结果也表明数据平稳（表 2）。

表 1：数据的自相关图

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.028	0.028	1.5888	0.207
		2	0.027	0.026	3.0686	0.216
		3	0.055	0.054	9.1723	0.027
		4	0.045	0.042	13.297	0.010
		5	0.041	0.037	16.724	0.005
		6	-0.049	-0.056	21.540	0.001
		7	-0.003	-0.007	21.560	0.003
		8	-0.042	-0.046	25.077	0.002

表 2：数据单位根检验的结果

ADF Test Statistic	-17.67702	1% Critical Value*	-2.5668
		5% Critical Value	-1.9395
		10% Critical Value	-1.6157

但从图 1 中我们可以看出数据存在波动率聚类现象，即收益率的波动率（volatility）不仅随时间 t 变化，而且常在某一时段中连续出现偏高或偏低的情形。在当今的金融市场中，金融工具的波动率聚类现象较为普遍，在这方面有许多的研究成果。一般认为，刻划这种现象的最有力的工具就是 GARCH（generalized autoregression conditional heteroskedasticity）。

本文得到国家社科基金 99BJY063 项目、教育部数量经济研究中心重大项目《我国资本市场的结构优化与风险控制》的资助。

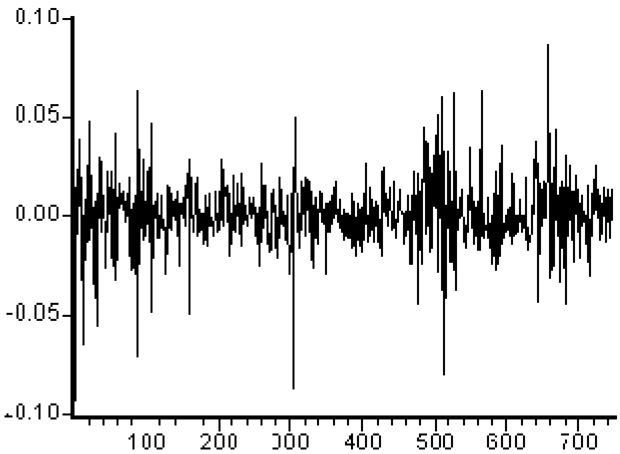

图 1：上证综指收益率

二、检验方法

本文采用 Kupiec 于 1995 年提出的似然比检验法来验证模型的有效性。在 Kupiec 检验中，设 N 为检验样本中损失高于 VaR 值的次数，T 为检验样本中总的数据个数。损失超出 VaR 值的次数服从二项分布，即 $N \sim B(T, p)$，其中 $p = 1-c$ （其中 c 为所用模型采用的置信水平,）因此 N/T 应等于概率 p 。

Kupiec 检验的原假设和备选假设为：

$$H_0 : N/T = p$$
$$H_1 : N/T \neq p$$

对应的似然比统计量为：

$$LR = 2\left[Log\left(\left(\frac{N}{T}\right)^N\left(1-\frac{N}{T}\right)^{T-N}\right) - Log\left(p^N(1-p)^{T-N}\right)\right]$$

该似然比统计量在原假设下服从自由度为 1 的 χ^2 分布，在一定的显著性水平如 95%下，我们可以构造相应的接受域。不同置信水平下的接受域如表 3 所示。显然，置信水平越高，检验越困难，当检验样本中数据个数很少时更是如此。

表 3：Kupiec 检验的接受域

	检验样本数据个数			
左尾概率	250	500	750	1000
5.00%	$7 \le N \le 19$	$17 \le N \le 35$	$27 \le N \le 49$	$38 \le N \le 64$
1.00%	$1 \le N \le 6$	$2 \le N \le 9$	$3 \le N \le 13$	$5 \le N \le 16$
0.50%	$0 \le N \le 4$	$1 \le N \le 6$	$1 \le N \le 8$	$2 \le N \le 9$

显著性水平为 5%

对于 Kupiec 似然比检验，当检验样本的个数为 250、置信水平为 Basle 委员会要求的 99%时，显著性水平为 5%的单边临界值为 6，当损失值超出 VaR 值的次数大于 6 时则拒绝模型。

三、方差——协方差法

我们将应用不同的模型来度量上证综合指数的 VaR 值，这些方法虽然各有不同，但它们都需要对指数

收益率的分布做出假设。在这些方法中，VaR 值的大小都与指数收益率的标准差有关，并可以看出，收益率的波动越大，VaR 值也就越大。

1、正态假设

设上证综指收益率 r_t 服从正态分布：

$$r_t \sim N\left(\mu, \sigma^2\right)$$

则相应的 VaR 值为

$$VaR = -W\left(e^{\mu+\sigma\Phi^{-1}(p)} - 1\right)$$

其中 $\Phi^{-1}(p) = \dfrac{r^* - \mu}{\sigma}$，$\Phi(\cdot)$表示标准正态分布的累计分布函数，$r^*$ 即为给定持有期、一定置信水平下的最低收益率。VaR 表达式中的 μ 和 σ 都可估计，应用前 500 个数据作估计样本对这两个参数的估计结果和将参数估计值代入 VaR 表达式得到的上证综指收益率 VaR 的估计值（为方便起见，W 设为 100）如下：

参数	估计值
μ	-0.0929%
σ	1.6187%

尾部概率	VaR 估计值	损失超出 VaR 估计值的个数
5.00%	2.64	20*
1.00%	3.70	8*
0.50%	4.09	7*

有*号的数字表示检验样本中实际损失超出 VaR 估计值的次数落在接受域外的次数

由上表我们可以清楚地看出在正态假设前提下得出的 VaR 估计值在各尾部概率下都被拒绝，且检验数据中损失超出 VaR 估计值的个数总是大于接受域的上界，由此得到正态假设前提下得出的 VaR 估计值偏小，即有一部分损失并未包含在 VaR 估计值之内。

上面的计算说明正态假设并不能很好地作为上证指数收益率所服从的分布，收益率所服从的分布应具有比正态分布更宽的尾部，因此在下面我们将应用 t 分布假设。这样做是因为 t 分布具有较宽的尾部，也就是说相对于正态分布，t 分布的尾部可以包含更多极端事件（较大的意外损失）。

2、t 分布假设

从分布图上我们可以看到，t 分布的分布图相对于正态分布更为突起，尾部也更宽。t 分布的概率分布主要参数有三个：均值 μ、比例系数 γ 和自由度 v。一个服从 t 分布的随机变量均值为 μ（$v > 1$），方差为 $v\gamma^2/(v-2)$（$v > 2$）。v 越大，t 分布越接近正态分布；反之，v 越小，分布的尾部越宽。

我们假设上证指数收益率 r_t 服从 t 分布，其对数似然方程为：

$$L = T\left[\log\Gamma\left(\frac{v+1}{2}\right) - \log\Gamma\left(\frac{v}{2}\right) - \frac{1}{2}\log \pi v - \log\gamma\right]$$
$$-\frac{v+1}{2}\sum_{t=1}^{T}\log\left(1+\frac{(r_t-\mu)^2}{\gamma\sqrt{v}}\right)$$

对于这个方程，我们应用极大似然估计求得 μ、γ 和 v 的参数值如下，其对数似然均值为 2.80858。

参数	估计值	标准差	左尾概率	VaR 估计值	损失超出 VaR 估计值的个数
μ	0.0517%	0.0574%	5.00%	2.42	25*
γ	1.0647%	0.0632%	1.00%	4.67	5
v	3.36076	0.5665	0.5%	5.98	4

在 t 分布假设下 VaR 的表达式为

$$VaR = -W\left(e^{\mu+\gamma F_v^{-1}(p)} - 1\right)$$

式中 $F_v(\cdot)$ 为标准 t 分布的累积分布函数。将 VaR 的估计值与检验样本数据比较得到的结果列于上表中。从表中可以看出 t 分布假设下的 VaR 的估计值只有当 $p = 0.05$ 时被拒绝，而在 p 取其它值时都落在接受域内，说明 t 分布假设较正态分布假设有效，这主要是由于 t 分布所具有的宽尾特性使得分布所能够包含的发生意外损失的情形进一步增多。

3、混合正态假设

从收益率的数据中我们可以看出收益率在大多数情形下波动较为平缓，总是围绕其均值小幅波动，但在某些时期收益率的波动幅度则较大，对于这种情况，我们可以设想收益率服从两个正态分布，波动较小时服从的正态分布方差较小，而波动较大时服从的正态分布方差较大。这时收益率服从的就是一种混合正态分布。

设上证指数收益率以概率 $1-\lambda$ 服从方差为 σ^2 的正态分布，以概率 λ 服从方差为 τ^2 的正态分布。其中 $\tau^2 = \sigma^2 + \delta^2 > \sigma^2$，因此收益率在服从方差较大的分布时会发生较大的波动。尽管两种分布的方差不同，但二者的均值都为 μ，这时对数似然方程为：

$$L = \sum_{t=1}^{T} \log\left\{\frac{1-\lambda}{\sigma}\varphi\left(\frac{r_t-\mu}{\sigma}\right) + \frac{\lambda}{\tau}\varphi\left(\frac{r_t-\mu}{\tau}\right)\right\}$$

其中 $\varphi(\cdot)$ 为标准正态分布的密度函数。由这个方程即可估计出相应的参数值，其对数似然均值为 2.8108。

参数	估计值	标准差	左尾概率	VaR 估计值	损失超出 VaR 估计值的个数
μ	0.048%	0.0581%	5.00%	2.56	23*
σ	1.123%	0.0571%	1.00%	3.61	8*
τ	3.460%	0.475%	0.5%	3.98	7*
λ	0.127	0.0394			

由这些参数的估计值即可求出 VaR 的估计值及检验结果。

混合正态假设还是以正态假设为基础，由检验结果可以看出基于该假设的 VaR 的估计值偏小，并不能准确的度量指数风险。

4、GARCH 模型

上面三种假设的缺点就在于并没有对波动率聚类现象提出合理的解释。对于这种金融领域的一般现象，本文中我们应用 GARCH 模型。在 GARCH 模型中，我们定义误差项 η_{t+1}，其条件均值为 0。t 时刻之前的

信息由$\Re_t$表示。在本文中η_{t+1}表示指数收益率的误差项。

为了说明波动率聚类现象，GARCH 模型假设误差项的条件方差$\sigma_t^2 = Var[\eta_{t+1} \mid \Re_t]$。一般情况下，GARCH（p,q）可由下式表示

$$\sigma_t^2 = \omega + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{q} \alpha_i \eta_{t-i}^2 \qquad (1)$$

（1）式包含 p 阶条件方差，q 阶误差项的平方。这里我们令误差项为收益率对其均值的偏差，即

$$r_{t+1} = \mu + \eta_{t+1}$$

由于σ_t^2为收益率的条件方差，我们可以将η_{t+1}表示为$\sigma_t \varepsilon_{t+1}$的形式，其中$\varepsilon_{t+1}$均值为 0，方差为 1。

为了便于计算，GARCH（p,q）模型的滞后阶数一般取 1。这里我们应用 GARCH（1,1）模型如下

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \alpha \eta_{t-1}^2 \qquad (2)$$

ϖ、α和β都应大于等于 0。（2）式很好地解释了波动率聚类现象。

如果当期的市场波动率很大，那么下一期的市场波动就应较大；反之，如果收益与其均值差异不大时，当期波动率较低，则下期的波动率也较低。当$\alpha + \beta < 1$时，受到冲击后条件方差随时间逐渐向其均值回归；$\alpha + \beta = 1$时，冲击将对条件方差产生持久的影响。

为了对参数进行极大似然估计，我们必须对误差项服从的概率分布做出假设。这里假设误差项η_{t+1}服从 t 分布，其条件对数似然比函数为

$$l_t(\eta_{t+1}) = \log\Gamma\left(\frac{v+1}{2}\right) - \log\Gamma\left(\frac{v}{2}\right) - \frac{1}{2}\log\pi(v-2) - \log\sigma_t - \frac{v+1}{2}\log\left(1+\left(\frac{\eta_{t+1}}{\sigma_t\sqrt{v-2}}\right)\right)$$

对指数收益率应用 GARCH（1,1）模型误差项的 t 分布假设后可以得到参数的极大似然估计结果如下，其对数似然均值为 2.8547。

参数	估计值	标准差	左尾概率	VaR 估计值	损失超出 VaR 估计值的个数
μ	0.0656%	0.0553%	5.00%	3.33	12
ω	0.0038%	0.0015%	1.00%	4.69	5
α	0.2035	0.0698	0.10%	5.18	4
β	0.6506	0.0935			
v	5.0790	1.1275			

$\alpha + \beta \prec 1$，因此该 GARCH（1,1）模型收敛。将各参数的估计值代入 VaR 的表达式，可得 VaR 的不同时期的估计值，如表所示。由表可见，误差项服从 t 分布的 GARCH（1,1）模型对上证综指收益率 VaR 值

的估计最为有效。

此外，我们也可以使用历史模拟法来计算 VaR 值。这种方法不需要对收益率服从的分布做出假设并且是一种非参数法，之所以如此是因为历史模拟法主要应用收益率的经验分布来计算其 VaR 值。

历史模拟法的主要步骤是：首先，将估计样本分为多个相同长度的子样本。令子样本长度为n，估计样本长度为T，在本文中取$n=250$，则我们一共可以得到$T-n+1=251$个子样本。在这 251 个子样本中，每相邻两个样本中的数据只有一个不同。其次，求出每个子样本的 p 阶分位数$R_t^{\,p}$。由此得出每个子样本的 VaR 估计值：

$$\hat{V}aR_t = -WR_t^{\,p}$$

最后，求出这些子样本 VaR 估计值的均值，并将这一均值作为最终的 VaR 估计值。

当$n=250$时，置信水平为 99%的子样本的 VaR 估计值位于子样本的第二与第三个最大损失之间，我们可用差值法来求得这一估计值。另外，由上面的叙述我们知道，历史模拟法赋予样本内各数据相同的权重，这样一来有可能使得当一数据被新的子样本舍弃时得出的子样本 VaR 估计值变动较为剧烈。同时，对于子样本长度n的选取也存在争论，n如果很小，那么子样本的 VaR 估计值对于市场的意外波动过于敏感；反之，当n很大时，子样本的 VaR 估计值就有可能包含与目前市场毫无关系的旧数据。

对上证指数收益率应用历史模拟法得出的 VaR 估计结果如下：

左尾概率	VaR 估计值	损失超出 VaR 估计值的个数
5.00%	2.08	31*
1.00%	4.01	7*
0.10%	5.75	4

四、各种方法的比较

在前面应用方差——协方差法及历史模拟法计算上证综指收益率 VaR 估计值的过程中，我们共使用了 5 种具体方法。下面我们就对这 5 种方法得出的结果加以比较并进一步分析各种方法产生不同结果的原因。

表 4：各种方法的结果比较

左尾概率	5.00%	1.00%	0.50%
理论的失败次数	12.5	2.5	1.25
方差——协方差法			
正态假设	20*	8*	7*
T 分布假设	25*	5	4
混合正态	23*	8*	7*
GARCH	12	5	4
历史模拟法	31*	7*	4

从表 4 中可以看出，正态假设及混和正态假设得到的结果最差，这是由于这两种假设都立足于收益率服从的分布是正态分布，这样一来就使得分布的尾部并不能充分包含发生意外损失的情形，因此这两种假设下得出的 VaR 估计值偏小。由检验结果也可以看出实际损失超出 VaR 估计值的次数都大于接受域的上界。

t 分布所具有的宽尾特征使得基于 t 分布假设的 VaR 估计值较基于正态假设及混和正态假设得到的 VaR 估计值更为有效。

从检验的结果看，应用 t 分布假设时得出的 VaR 估计值只是在置信水平为 95%时未能通过检验，而在其它两种置信水平下都通过了检验，这一点比应用正态假设及混和正态假设时有了很大提高。基于历史模拟法的 VaR 估计值只是在置信水平为 99.9%时才通过了检验，对于另两种置信水平则无效，说明这种方法本身的有效性较低。产生这种情况的主要原因是收益率的历史分布不能很好地代替其真实分布。在这五种方法中，基于误差项服从 t 分布的 GARCH 假设得出的 VaR 估计值最为有效。从检验结果可知，应用该假设得出的 VaR 估计值不仅在各置信水平下完全通过检验，而且实际的失败次数与其它方法得出的结果相比也最接近于理论上的失败次数。通过上面的分析，我们完全有理由认为上证综指收益率适用于误差项服从 t 分布的 GARCH 过程。

参 考 文 献

1、Philippe Jorion 《Value at Risk: The New Benchmark for Controlling Market Risk》，The McGraw-Hill Companies，Inc.1997.

2、Kevin Dowd 《Beyond Value at Risk：The New Science of Risk Management》，Chichester, England；Wiley c1998.

3、Kupiec 《Techniques for Verifying the Accuracy of Risk Measurement Models》，《Journal of Derivatives》，1955,2.

4、范英，《VaR 方法及其在股市风险分析中的应用初探》，《中国管理科学》，第 8 卷第 3 期，2000 年 9 月。

（作者陈守东，吉林大学数量经济研究中心、吉林大学商学院教授，王鲁非，吉林大学商学院）