DOI:10.13546/j.cnki.tjyjc.2015.19.044

基于Credit Risk+模型的互联网金融信用风险估计

李　琦,曹国华

(重庆大学 经济与工商管理学院,重庆 400044)

摘　要: 文章基于CreditRisk+模型,使用互联网信贷平台四个行业的贷款数据,在不同置信水平下,对互联网金融的信用风险水平进行比较分析。结果表明,行业风险因子间协方差相等时,复合伽玛CreditRisk+模型和多元系统风险Credit Risk+模型计算结果几乎一致,与CSFB CreditRisk+模型和两阶段CreditRisk+模型相比能更好地反映贷款组合的非预期损失。行业风险因子间协方差不等时,多元系统风险CreditRisk+模型能克服其他CreditRisk+模型的缺陷,综合考量系统风险和行业风险的影响,能更好地估计贷款组合的信用风险水平。

关键词: 互联网金融;信贷平台;信用风险;Credit Risk+模型

中图分类号: F832.4　　**文献标识码:** A　　**文章编号:** 1002-6487(2015)19-0164-03

0　引言

互联网金融作为金融交易渠道的创新,是指利用大数据、社交网络、云计算和搜索引擎等互联网技术,实现资金融通、支付和信息中介功能的新兴金融模式。虽实现了模式创新,但互联网金融仍规避不了经营风险,互联网信贷信用风险日益凸显。我国互联网金融目前尚处于初级阶段,信贷业务呈"资金集聚、技术集聚、人才集聚、风险集聚"状态。互联网金融信用风险不仅会使投资者失去投资信心,还会损害到投资者和信贷平台的利益。因此,加强对互联网金融信用风险的管理已成为其稳健发展的重中之重。

1　理论框架

1.1　风险度量的VaR方法

风险度量实质是把一个代表风险的随机变量转化为一个实际值的过程,选择合适的度量函数是其核心问题。一般这一过程可刻画为:$r=\rho(X)$,X 表示随机损失,r 为风险度量值。标准差、方差、VaR等都是常见的风险度量方法。其中VaR方法求得的是在一定置信水平下的最大可能损失。在此过程中,假设风险的累积分布函数为 $F(x)$,可把 VaR(α) 表述成 $\Pr ob\{X \le VaR(a)\} = a$ 或 $VaR(a)=Min\{x|F(x)\ge a\}$。

1.2　CSFB CreditRisk+模型

CSFB模型假设期间内违约概率不变,贷款组合中每笔贷款违约概率很小。行业风险因子 $\gamma=(\gamma_1,\gamma_2,\cdots,\gamma_k)$ 相互独立,服从均值为1、方差为 σ_k^2 的伽玛分布,α_k 和 β_k 为形参数和规模参数。受其影响,假设债务人为A,其无条件违约概率 p_A,则其违约概率即为:$p_A(\gamma)=p_A\sum_{k=1}^{K} g_k^A\gamma_k$,其中,$\sum_{k=1}^{K} g_k^A=1$,$k=1,2,\cdots,K$。进一步地,得到该债务人违约概率生成函数:

$$G_A(z|\gamma)=1-p_A(\gamma)+p_A(\gamma)z^{v_A}=1+p_A(\gamma)(z^{v_A}-1)\approx e^{P_A(\gamma)(z^{v_A}-1)} \tag{1}$$

考虑到不同债务人条件独立,则得到贷款组合的概率生成函数:

$$\begin{aligned} G(z|\gamma)&=\prod_A G_A(z)=\prod_A e^{P_A(\gamma)(z^{v_A}-1)}=e^{\sum_A p_A(\gamma)(z^{v_A}-1)} \\ &=e^{\sum_A p_A\sum_{k=1}^{K} g_k^A\gamma_k(z^{v_A}-1)}=e^{\sum_{k=1}^{K}\gamma_K P_k(z)} \end{aligned} \tag{2}$$

其中,$P_k(z)=\sum_A g_k^A p_A(Z^{v_A}-1)$。进一步可得:

$$G(z)=\prod_{k=1}^{K}\frac{1}{(1-\beta_k P_k(z))^{a_k}}=e^{-\sum_{k=1}^{K} a_k\ln(1-\beta_k P_k(z))} \tag{3}$$

因 $\alpha_k=\frac{1}{\sigma_k^2}$,$\beta_k=\sigma_k^2$,则得:

$$G(z)=e^{-\sum_{k=1}^{K}\frac{1}{\sigma_k^2}\ln(1-\sigma_k^2 P_k(z))} \tag{4}$$

1.3　复合伽玛CreditRisk+模型

在CSFB模型基础上,复合伽玛模型引入随机变量 γ_0,其服从均值为1、方差为 σ^2 的伽玛分布。行业风险因子因 γ_0 不再相互独立,其形参数为 $\alpha_k=\gamma_0\overline{a_k}$,$\overline{a_k}$ 为常数。债务人A违约概率同CSFB CreditRisk+模型,其违约损失概率生成函数为:

$$G_A(z|\gamma_0,\gamma_1,\cdots,\gamma_K)=e^{p_A\sum_{k=1}^{K} g_k^A\gamma_k(z^{v_A}-1)} \tag{5}$$

由式(5)进一步可得到:

基金项目: 重庆大学金融实验项目(2013JGSYJX005)

作者简介: 李　琦(1981-),女,重庆人,博士研究生,研究方向:金融风险。
曹国华(1967-),男,安徽宣城人,教授,博士生导师,研究方向:金融学。

$$G(z|\gamma_0,\gamma_1,\cdots,\gamma_K)=\prod_A G_A(z|\gamma_0,\gamma_1,\cdots,\gamma_K)$$
$$=\prod_A e^{p_A\sum_{k=1}^{K}g_k^A\gamma_k(z^{v_A}-1)}=e^{\sum_{k=1}^{K}\gamma_k P_k(z)} \tag{6}$$

由式(6)和 $\alpha_k=\frac{\gamma_0}{\beta_k}$，可得：

$$G(z|\gamma_0)=\prod_{k=1}^{K}\frac{1}{(1-\beta_k P_k(z))^{\frac{\gamma_0}{\beta_k}}}=e^{\gamma_0[-\sum_{k=1}^{K}\frac{1}{\beta_k}\ln(1-\beta_k P_k(z))]} \tag{7}$$

因 $\alpha=\frac{1}{\sigma^2}$，$\beta=\sigma^2$，可得：

$$G(z)=e^{-\frac{1}{\sigma^2}\ln(1+\sigma^2(\sum_{k=1}^{K}\frac{1}{\beta_k}\ln(1-\beta_k P_k(z))))} \tag{8}$$

且存在 $\sigma_{kl}=\begin{cases}\sigma^2, k\neq l\\ \beta_k+\sigma^2, k=l\end{cases}$ 成立，当 $k\neq l$ 时，$\sigma_{kl}=\sigma^2$，行业风险因子协方差等于 γ_0 的方差 σ^2；而当 $k=l$ 时，则有 $\sigma_{kk}>\sigma^2$。

1.4　两阶段CreditRisk+模型

在两阶段模型中，考虑了系统风险因素，系统风险因子设定为 $Y_1,Y_2,\cdots,Y_N$，Y_i 服从均值为1，方差为 δ_i^2 的伽玛分布。行业风险因子用系统风险因子表示成：$\gamma_k=b_{k1}Y_1+b_{k2}Y_2+\cdots+b_{kN}Y_N$。其中 $k=1,2,\cdots,K$，b_{ki} 之和等于1。债务人A违约概率同上，由式(2)可得：

$$G(z|Y_1,Y_2,\cdots,Y_N)=e^{\sum_{k=1}^{K}\gamma_k P_k(z)}=e^{\sum_{k=1}^{K}\sum_{i=1}^{N}b_{ki}Y_iP_k(z)}$$
$$=e^{\sum_{k=1}^{K}b_{ki}P_k(z)\sum_{i=1}^{N}Y_i}=e^{\sum_{i=1}^{N}Y_i\sum_{k=1}^{K}b_{ki}P_k(z)} \tag{9}$$

进一步可算得：

$$G(z)=\prod_{i=1}^{N}\frac{1}{(1-\beta_i\sum_{k=1}^{K}b_{ki}P_k(z))^{a_i}}=e^{-\sum_{i=1}^{N}\frac{1}{\delta_i^2}\ln(1-\delta_i^2\sum_{k=1}^{K}b_{ki}P_k(z))} \tag{10}$$

计算出行业风险因子的方差：

$$\mathrm{var}[\gamma_k]=\mathrm{var}[b_{k1}Y_1+b_{k2}Y_2+\cdots+b_{kN}Y_N]=b_{k1}^2\delta_1^2+b_{k2}^2\delta_2^2+\cdots+b_{kN}^2\delta_N^2 \tag{11}$$

和行业风险因子间的协方差：

$$\mathrm{Cov}(\gamma_k,\gamma_l)=E(\gamma_k\cdot\gamma_l)-E(\gamma_k)\cdot E(\gamma_l)=E(\gamma_k\cdot\gamma_l)-1=\sum_{i=1}^{N}b_{ki}b_{li}\delta_i^2 \tag{12}$$

因 $\gamma_k=b_{k1}Y_1+b_{k2}Y_2+\cdots+b_{kN}Y_N$，且 b_{ki} 之和等于1。可得：

$$p_A(Y_1,Y_2,\cdots,Y_N)=p_A\sum_{k=1}^{K}g_k^A(\sum_{i=1}^{N}b_{ki}Y_i)=p_A\sum_{i=1}^{N}Y_i(\sum_{k=1}^{K}g_k^Ab_{ki})=p_A\sum_{i=1}^{N}c_i^AY_i \tag{13}$$

其中，c_i^A 为系统风险因子对债务人的影响权重，且 $\sum_{i=1}^{N}c_i^A=1$，$c_i^A=\sum_{k=1}^{K}g_k^Ab_{ki}$。

1.5　多元系统风险Credit Risk+模型

系统风险因子 $Y_1,Y_2,\cdots,Y_N$ 设定同上文，α_i 和 β_i 为其形参数和规模参数。行业风险因子形参数为 $\alpha_k=(b_{k1}Y_1+b_{k2}Y_2+\cdots+b_{kN}Y_N)\overline{\alpha_k}$，$\overline{\alpha_k}$ 为常数，且 b_{ki} 之和等于1，$k=1,2,\cdots,K$。行业风险因子的期望为1，得：

$$\alpha_k=\frac{b_{k1}Y_1+b_{k2}Y_2+\cdots+b_{kN}Y_N}{\beta_k} \tag{14}$$

将式(14)代入式(7)，且 $\alpha_i=\frac{1}{\delta_i^2}$，$\beta_i=\delta_i^2$，可得：

$$G(z)=\prod_{i=1}^{N}\frac{1}{(1-\beta_iA_i(z))^{\alpha_i}}=e^{-\sum_{i=1}^{N}\alpha_i\ln(1-\beta_iA_i(z))}=e^{-\sum_{i=1}^{N}\frac{1}{\delta_i^2}\ln(1-\delta_i^2A_i(z))} \tag{15}$$

其中 $A_i(z)=-\sum_{k=1}^{K}\frac{b_{ki}}{\beta_k}\ln(1-\beta_kP_k(z))$。由式(14)还可得：

$$\mathrm{var}[\gamma_k]=\beta_kE_{Y_1,Y_2,\cdots,Y_N}[\mathrm{var}_{Y_1,Y_2,\cdots,Y_N}[\gamma_k|Y_1,Y_2,\cdots,Y_N]]+\mathrm{var}_{Y_1,Y_2,\cdots,Y_N}[E[\gamma_k|Y_1,Y_2,\cdots,Y_N]]=\beta_k\sum_{i=1}^{N}b_{ki}+\mathrm{var}(\sum_{i=1}^{N}b_{ki}Y_i)$$
$$=\beta_k+\sum_{i=1}^{N}b_{ki}^2\delta_i^2 \tag{16}$$

式(16)表明，行业风险因子方差受 δ_i^2 和 β_k 共同作用。当 δ_i^2 为0时，$\mathrm{var}[\gamma_k]=\beta_k$，行业风险因子服从均值为1、方差为 β_k 的伽玛分布，模型就等同于CSFB CreditRisk+模型。令 $k\neq l$ 时，得行业风险因子的协方差为：

$$\mathrm{Cov}(\gamma_k,\gamma_l)=E(\gamma_k,\gamma_l)-E(\gamma_k)\cdot E(\gamma_l)$$
$$=E_{Y_1,Y_2,\cdots,Y_N}[E_{Y_1,Y_2,\cdots,Y_N,k,l}[\gamma_k\cdot\gamma_l|Y_1,Y_2,\cdots,Y_N]]-1$$
$$=E_{Y_1,Y_2,\cdots,Y_N}[E_{Y_1,Y_2,\cdots,Y_N,k}[\gamma_k|Y_1,Y_2,\cdots,Y_N]\times E_{Y_1,Y_2,\cdots,Y_N,l}[\gamma_l|Y_1,Y_2,\cdots,Y_N]]-1=\sum_{i=1}^{N}b_{ki}b_{li}\delta_i^2 \tag{17}$$

式(17)表明，模型中行业风险因子之间的协方差只受系统风险因子的作用。在分析中，一般由行业风险因子间的协方差矩阵及其自身特性，估计出无系统风险因子影响时的规模参数 $\beta_k(=\sigma^2)$，再根据协方差矩阵和式(16)和(17)估计出其他参数。

2　实证模拟及结果分析

2.1　方案设定

本文收集了一互联网信贷平台贷款数据进行模拟，整个贷款组合来自餐饮旅馆业、农业、教育产业和文化产业等行业，每个行业各有40笔贷款。四个行业意味着有四个行业风险因子，这四个行业风险因子都服从均值为1的伽玛分布。本文考虑行业风险因子协方差相等和不相等两种情况，在不同置信水平下，分别估计上文不同信用风险模型贷款组合风险VaR值，并对模型进行比较分析。

2.2　结果分析

为便于模型间的比较，首先考虑贷款组合 γ_i 间协方差相等的情况。先从复合伽玛模型入手，假设该模型中随机变量和行业风险因子的方差分别为0.01、0.04、0.16、0.36和0.64。此时可以得到 γ_i 协方差矩阵为：

$$\begin{bmatrix}0.04&0.01&0.01&0.01\\0.01&0.16&0.01&0.01\\0.01&0.01&0.36&0.01\\0.01&0.01&0.01&0.64\end{bmatrix}$$

由式 $\sigma_{kl}=\begin{cases}\sigma^2, k\neq l\\ \beta_k+\sigma^2, k=l\end{cases}$ 知规模参数 β_i 取值分别为0.03、0.15、0.35和0.63。若多元系统风险模型与复合伽玛模型的行业风险因子相关性结构相同，则其 β_i 值也可设

定为0.03、0.15、0.35和0.63。并假设存在3个系统风险因子Y_1,Y_2,Y_3，根据协方差矩阵和式(16)、(17)，可得Y_1,Y_2,Y_3方差分别为$\delta_1^2=0.0945$，$\delta_2^2=0.05$，$\delta_3^2=0.01$，并由式(14)得到以下关系：

$$\begin{cases}\alpha_1=\dfrac{0.1Y_1+0.4Y_2+0.5Y_3}{\beta_1}\\ \alpha_2=\dfrac{0.1Y_1+0.4Y_2+0.5Y_3}{\beta_2}\\ \alpha_3=\dfrac{0.1Y_1+0.4Y_2+0.5Y_3}{\beta_3}\\ \alpha_4=\dfrac{0.1Y_1+0.4Y_2+0.5Y_3}{\beta_4}\end{cases}$$

对CSFB模型而言，由于γ_i间相互独立，其协方差矩阵就是上文矩阵中对角线之外元素全为0的情形。本文借助MatLab8.0工具，在不同置信水平下，评估四个模型贷款组合的VaR，结果如表1所示。

表1 协方差相等时不同置信水平下各CreditRisk+模型的Var值

CreditRisk+模型类别	90%	95%	99%	99.5%	99.9%
CSFB模型	801	911	1034	1126	1389
复合伽玛模型	808	918	1045	1147	1456
两阶段模型	805	911	1023	1110	1305
多元系统风险模型	808	918	1045	1147	1460

从表1可以看出，当置信水平大于99%时，两阶段模型的Var值，却比同等置信水平下CSFB模型Var值要小。这可能是因为，前者计算贷款组合的违约方差使用的是加权平均的方法，风险未能被充分地反映出来。值得注意的是，与CSFB模型和两阶段模型相比，在等同的置信水平下，复合伽玛模型和多元系统风险模型的Var值都要大，结果基本一致，可认为这两个模型能更好地反应现实中贷款组合的非预期损失的风险水平。这可以从两个方面进行解释：从模型假设上来说，CSFB模型中风险因子相互独立的假设脱离了其受宏观经济变量诸多系统风险影响的现实，存在风险低估的可能；从技术角度来说，由于复合伽玛模型和多元系统风险模型在γ_i间协方差相等情况下，能将协方差矩阵纳入模型框架中，使两模型具备一定程度的完备性。

下面来分析γ_i间的协方差不相等的情况。此时仍采用上文贷款数据，得到新的协方差矩阵：

$$\begin{bmatrix}0.04 & 0.0197 & 0.0186 & 0.073\\ 0.0197 & 0.05 & 0.0183 & 0.0735\\ 0.0186 & 0.0183 & 0.06 & 0.0786\\ 0.073 & 0.0735 & 0.0786 & 0.58\end{bmatrix}$$

在多元系统风险模型中，根据给定的规模参数值和协方差矩阵，可以得到行业风险因子的形参数如下：

$$\begin{cases}\alpha_1=\dfrac{0.7Y_1+0.2Y_2+0.2Y_3}{\beta_1}\\ \alpha_2=\dfrac{0.6Y_1+0.3Y_2+0.2Y_3}{\beta_2}\\ \alpha_3=\dfrac{0.5Y_1+0.4Y_2+0.2Y_3}{\beta_3}\\ \alpha_4=\dfrac{0.1Y_1+0.2Y_2+0.7Y_3}{\beta_4}\end{cases}$$

对于前文的四个行业的贷款组合数据，四个模型的VaR值估计如表2所示。

表2 协方差不等时不同置信水平下各CreditRisk+模型的Var值

CreditRisk+模型类别	90%	95%	99%	99.5%	99.9%
CSFB模型	745	827	914	1039	1201
复合伽玛模型	798	868	975	1056	1264
两阶段模型	753	836	903	990	1148
多元系统风险模型	803	873	994	1103	1342

由表2可看出，与协方差相等时的分析一致，在置信度大于99%时，两阶段模型计算出来的结果比CSFB模型结果要小。在不同置信水平下，与CSFB模型和两阶段模型相比，复合伽玛模型和多元系统风险模型的VaR都要大。但与协方差相等时分析不一致的是，复合伽玛模型与多元系统风险模型的VaR估计存在差异，前者比后者的结果要小，这说明协方差不等时，前者对贷款组合信用风险水平存在低估现象。这可能是因为前者在用加权平均的方法估计出的γ_0方差替代行业风险因子间的协方差时，这一过程存在误差，而多元系统风险模型能将行业风险因子协方差矩阵纳入到模型框架内。

3 结语

本文基于CreditRisk+模型框架，使用互联网信贷平台四个行业的贷款数据，在不同置信水平下，对不同模型下互联网金融的信用风险水平进行比较分析。结果表明，行业风险因子间协方差相等时，在等同的置信水平下，复合伽玛CreditRisk+模型和多元系统风险Credit Risk+模型的Var值几乎一致，都要大于CSFB CreditRisk+模型和两阶段CreditRisk+模型，因此可认为这两个模型能更好地反应现实中贷款组合的信用风险水平。当行业风险因子协方差不相等时，复合伽玛CreditRisk+模型和多元系统风险Credit Risk+模型的VaR值依然比CSFB CreditRisk+模型和两阶段CreditRisk+模型要大。但此时复合伽玛CreditRisk+模型相比多元系统风险Credit Risk+模型，对贷款组合信用风险水平却存在低估现象，多元系统风险Credit Risk+模型能更好的评估贷款组合的信用风险水平。综上所述，多元系统风险CreditRisk+模型能克服其他CreditRisk+模型的缺陷，综合考量系统风险和行业风险的影响，能更好地估计贷款组合的非预期损失，其在互联网金融信贷平台信用风险估计方面可能具有较好的适用性。

参考文献：

[1]谢清河.我国互联网金融发展问题研究[J].经济研究参考,2013,(49).

[2]孙小丽,彭龙.KMV模型在中国互联网金融中的信用风险测算研究[J].北京邮电大学学报.(社会科学版),2013,15(6).

[3]汪飞星,姚磊.聚合信用风险模型的改进和研究[J].价值工程,2013,(5).

[4]吕志华,彭建刚.CreditRisk+模型采用Poisson分布所产生的经济资本计量误差分析[J].管理评论,2011,(1).

[5]彭建刚,吕志华.基于行业特性的多元系统风险因子CreditRisk+模型[J].中国管理科学,2009,17(3).

（责任编辑/浩 天）