

Dependence Analysis of CPI and PPI Based on Copula Function

Ying Cui, He Li, Lili Wei*

School of Mathematics and Statistics, Ningxia University, Yinchuan Ningxia
Email: weill866@163.com

Received: Mar. 11th, 2019; accepted: Mar. 26th, 2019; published: Apr. 2nd, 2019

Abstract

CPI and PPI are core indexes in price systems, measure and analysis of the dependence between CPI and PPI can help us to better understand the economic development trend, which provides a reliable basis and suggestion for the government in making economic policy. Considering the essential nonlinearity of CPI and PPI, copula function with perfect performance is selected to study the dependence between CPI and PPI in this paper. Using the national monthly CPI and PPI data from January 2015 to December 2016 to research dependence, the results show that Gaussian Copula is better in depiction of rank correlation, and Gumble Copula is superior to other Copula functions in tail correlation and degree of fit with original data.

Keywords

Copula, CPI, PPI, Dependence

基于Copula函数的CPI与PPI相关性分析

崔英, 李贺, 魏立力*

宁夏大学数学统计学院, 宁夏 银川
Email: weill866@163.com

收稿日期: 2019年3月11日; 录用日期: 2019年3月26日; 发布日期: 2019年4月2日

摘要

CPI与PPI是价格体系中的两大核心指标, 两者之间相关关系的度量及分析, 有利于定量刻画经济发展趋势, 为政府制定经济政策提供依据和建议。考虑到CPI与PPI数据的本质非线性关系, 选取表现能力更强

*通讯作者。

的Copula函数来研究CPI与PPI之间的相依性。采用2011年1月至2016年12月全国月度同比CPI与PPI数据进行相关性研究,结果表明Gaussian Copula在刻画秩相关性上效果较好,而在尾部相关性以及与原始数据的拟合程度上,Gumble Copula要优于其他Copula。

关键词

Copula, CPI, PPI, 相关性

Copyright © 2019 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



1. 引言

CPI 与 PPI 作为我国价格体系中的两大核心指标,从构成上看既有相交也有相离。当构成的商品和服务价格发生变化时,会引起 CPI 与 PPI 的关联性波动,可能是同向的,也可能是反向的。对两者之间相关关系进行度量及分析,有利于我们定量刻画经济过热或紧缩的发展趋势,为政府制定经济政策提供可靠的依据和建议。考虑到 CPI 与 PPI 数据的本质非线性相关,Pearson 相关系数是不合适的,取而代之的是选用能够更加精细、全面、灵活刻画相关关系的 Copula 函数来研究 CPI 与 PPI 之间的相依性。

我国学者对 Copula 函数的研究始于文献[1],该文献对 Copula 理论应用在金融领域的可行性进行分析,并在我国股票市场进行了应用研究。文献[2]对金融时间序列时变的相依性结构用 Copula 理论进行刻画。文献[3]利用 Pair Copula-GARCH 模型研究了我国商业银行的经营风险。文献[4]利用贝叶斯估计方法,通过对沪深股市的实证研究,发现基于贝叶斯方法的 Copula 相关结构具有时变性。文献[5]建立了信用风险相关性度量的 Copula 模型,对信用风险的相关性变化进行描述。

本文利用 Copula 函数对 2011 年 1 月至 2016 年 12 月全国月度同比 CPI 与 PPI 数据的相关性进行分析。相比之前的研究,在样本数据的采集过程中,主要考虑金融危机爆发后,经济逐渐恢复平稳的 CPI 与 PPI 数据,这样更有利于反映当下的经济发展形势。其次,由于金融数据具有尖峰厚尾的性质,所以通过非参数估计法来确定 CPI 与 PPI 各自的边缘分布而不是指定拟合的边缘分布。在选择 Copula 函数时,将选择范围从椭圆族扩大到 Archimedean 族,以便于更精细地刻画 CPI 与 PPI 之间的相关关系。最后,利用经验 Copula 来代替 χ^2 检验统计量对数据进行拟合优度检验,这也更加符合实际观测数据。

2. Copula 理论及其相关性度量

1959 年,Sklar 初次提出 Copula 理论,并指出可以将 N 维联合分布函数分解为 N 个边缘分布函数以及一个 Copula 函数,这个 Copula 函数可以细致、全面地刻画随机变量之间的相关结构。由于 Copula 函数具有将联合分布与边缘分布结合在一起的性质,有人也把 Copula 函数称之为“连接函数”。本节介绍 Copula 函数的基本概念及一系列基于 Copula 函数的相关性度量。

2.1. Copula 函数的概念与性质

这里我们给出有关 Copula 函数的基本概念和有关性质[6]。

定义 1: 称满足以下条件的函数 $C(u, v)$ 为二元 Copula 函数:

- ① 定义域为 $[0, 1] \times [0, 1]$;

② $C(u, v)$ 有零基面, 即 $C(u, 0) = C(0, v) = 0$; 并且是二维递增的, 即对于任意 $0 \leq u_1 \leq u_2 \leq 1$, $0 \leq v_1 \leq v_2 \leq 1$, 有

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0;$$

③ 对于任意 $u, v \in [0, 1]$, 满足 $C(u, 1) = u$, $C(1, v) = v$ 。

由定义可知: $C(u, v)$ 是一个边缘分布为 $U(0, 1)$ 的二元联合分布函数, 且对于定义域内的任意一点 (u, v) 均有: $0 \leq C(u, v) \leq 1$ 。

定理 1: (二元分布的 Sklar 定理) 令 $H(x, y)$ 为具有边缘分布 $F(x)$ 和 $G(y)$ 的二元联合分布函数, 则存在一个 Copula 函数 $C(u, v)$, 满足

$$H(x, y) = C[F(x), G(y)] \quad (1)$$

若 $F(x), G(y)$ 是连续函数, 则 $C(u, v)$ 是唯一确定的; 反之, 若 $F(x), G(y)$ 为一元分布函数, $C(u, v)$ 是一个 Copula 函数, 由(1)式确定的 $H(x, y)$ 是具有边缘分布 $F(x), G(y)$ 的二元联合分布函数。

二元 Copula 函数有如下基本性质。

① $C(u, v)$ 关于每一个变量都是单调非减的, 也就是说, 如果保持一个变量不变, 那么 $C(u, v)$ 的值也将会随着另一个变量的增大而增大。

② 对任意的 $u_1, u_2, v_1, v_2 \in [0, 1]$, 有 $|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|$ 。

③ 对任意的 $u, v \in [0, 1]$, 令 $C^-(u, v) = \max(u + v - 1, 0)$, $C^+(u, v) = \min(u, v)$, 则有

$$\max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v)$$

称 $C^+(u, v)$ 和 $C^-(u, v)$ 分别为 Fréchet 上界和下界。

④ 若 U, V 独立, 且服从 $U(0, 1)$, 则 $C(u, v) = uv$, 记 $C^\perp(u, v) = uv$ 。

⑤ 对变量 X, Y 做严格递增的连续变换 $h_1(X), h_2(Y)$, 有 $C_{X, Y} = C_{h_1(X), h_2(Y)}$, 其中 $C_{X, Y}$ 表示连接 X, Y 的 Copula 函数; $C_{h_1(X), h_2(Y)}$ 表示连接 $h_1(X), h_2(Y)$ 的 Copula 函数。

2.2. 常用的 Copula 函数

Copula 分布作为一类连接函数, 包含很多分布族, 其中椭圆 Copula 函数族和 Archimedean Copula 函数族是最为常见的两个分布族。椭圆 Copula 函数族中主要有 Gaussian Copula 函数和 t-Copula 函数, 而 Archimedean Copula 函数族中主要有 Gumble Copula、Glayton Copula 和 Frank Copula 函数, 由于这些 Copula 具有厚尾的特征而在金融领域得到广泛应用。

二元 Gaussian Copula 的分布函数为

$$C^{Ga}(u, v; \rho) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{s^2 + t^2 - 2\rho st}{2(1-\rho^2)}\right] ds dt, \quad (2)$$

其中, $\Phi^{-1}(\cdot)$ 是标准正态分布函数 $\Phi(\cdot)$ 的逆函数。

自由度为 k 的二元 t-Copula 的分布函数为

$$C^t(u, v; \rho, k) = \int_{-\infty}^{t_k^{-1}(u)} \int_{-\infty}^{t_k^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \left[1 + \frac{s^2 + t^2 - 2\rho st}{k(1-\rho^2)}\right]^{-(k+2)/2} ds dt, \quad (3)$$

其中, $t_k^{-1}(\cdot)$ 是自由度为 k 的标准 t 分布函数 $t_k(\cdot)$ 的逆函数。

二元 Gumble Copula、Clayton Copula、Frank Copula 的分布函数分别为

$$C^{Gu}(u, v; \alpha) = \exp \left\{ - \left[(-\ln u)^\alpha + (-\ln v)^\alpha \right]^{\frac{1}{\alpha}} \right\};$$

$$C^{Cl}(u, v; \alpha) = (u^{-\alpha} + v^{-\alpha} - 1)^{-\frac{1}{\alpha}}, 0 \leq u, v \leq 1;$$

$$C^F(u, v; \alpha) = -\frac{1}{\alpha} \ln \left(1 + \frac{(1 - e^{-\alpha u})(1 - e^{-\alpha v})}{1 - e^{-\alpha}} \right), 0 \leq u, v \leq 1.$$
(4)

2.3. 基于 Copula 函数的相关性度量

Copula 函数作为刻画变量间相依性结构的工具, 在度量具有非线性关系的变量之间的相依性结构时具有明显的优势, 因此产生了一系列基于 Copula 函数的相关性度量。基于 Copula 函数的度量包括 Kendall's τ , Spearman's ρ_S , 尾部相关系数 λ 等。

1) Kendall 秩相关系数 τ

令 (X, Y) 和 (X', Y') 是独立同分布的随机变量, 若 $(X - X')(Y - Y') > 0$, 称 (X, Y) 和 (X', Y') 是一致的; 若 $(X - X')(Y - Y') < 0$ 则称 (X, Y) 和 (X', Y') 是不一致的。

Kendall's 秩相关系数 τ 的一般形式为:

$$\begin{aligned} \tau &= P[(X - X')(Y - Y') > 0] - P[(X - X')(Y - Y') < 0] \\ &= 4P(X > X', Y > Y') - 1 \end{aligned}$$

若随机变量 X, Y 的边缘分布分别为 $F(x), G(y)$, 相应的 Copula 函数为 $C(u, v)$, 则 Kendall's 秩相关系数 τ 可由相应的 Copula 函数 $C(u, v)$ 给出:

$$\tau = 4 \int_{[0,1]^2} C(u, v) dC(u, v) - 1.$$

2) Spearman's 秩相关系数 ρ_S

设 $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)$ 独立同分布的随机变量, Spearman's 秩相关系数 ρ_S 的一般形式为:

$$\rho_S = 3 \{ P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_3)(Y_1 - Y_3) < 0] \}$$

若随机变量 X, Y 的边缘分布分别为 $F(x), G(y)$, 相应的 Copula 函数为 $C(u, v)$, 则 Spearman's 秩相关系数 ρ_S 可由相应的 Copula 函数 $C(u, v)$ 给出:

$$\rho_S = 12 \int_{[0,1]^2} C(u, v) dudv - 3$$

3) 尾部相关系数 λ

尾部相关性是人们在金融风险中比较关心的, 包括上尾相关和下尾相关。令 X, Y 为连续的随机变量, 具有边缘分布 $F(x), G(y)$ 和 Copula 函数为 $C(u, v)$, 如果

$$\lim_{u \rightarrow 1^-} \bar{C}(u, u)/(1-u) = \lambda_U$$

存在, $\lambda_U \in (0, 1]$, $\bar{C}(u, v) = 1 - u - v + C(u, v)$ 为生存 Copula 函数, 则称 X, Y 上尾相关; $\lambda_U = 0$ 时, 称 X, Y 在分布上尾渐近独立。同样地, 如果

$$\lim_{u \rightarrow 0^+} C(u, u)/u = \lambda_L$$

存在, $\lambda_L \in (0, 1]$ 时, 则称 X, Y 下尾相关; $\lambda_L = 0$ 时, 称 X, Y 在分布下尾渐近独立。我们把 λ_U, λ_L 统称为

尾部相关系数, 且 $\lambda_U, \lambda_L \geq 0$ 。

3. CPI 与 PPI 相关性的实例分析

3.1. 边缘分布的确定

采用 2011 年 1 月至 2016 年 12 月全国月度同比 CPI(x_i) 与 PPI(y_i) 数据。令 X, Y 分别表示 CPI 与 PPI, 为了确定 X, Y 的分布类型, 首先分别做出 CPI 与 PPI 的频率直方图, 如图 1 所示。

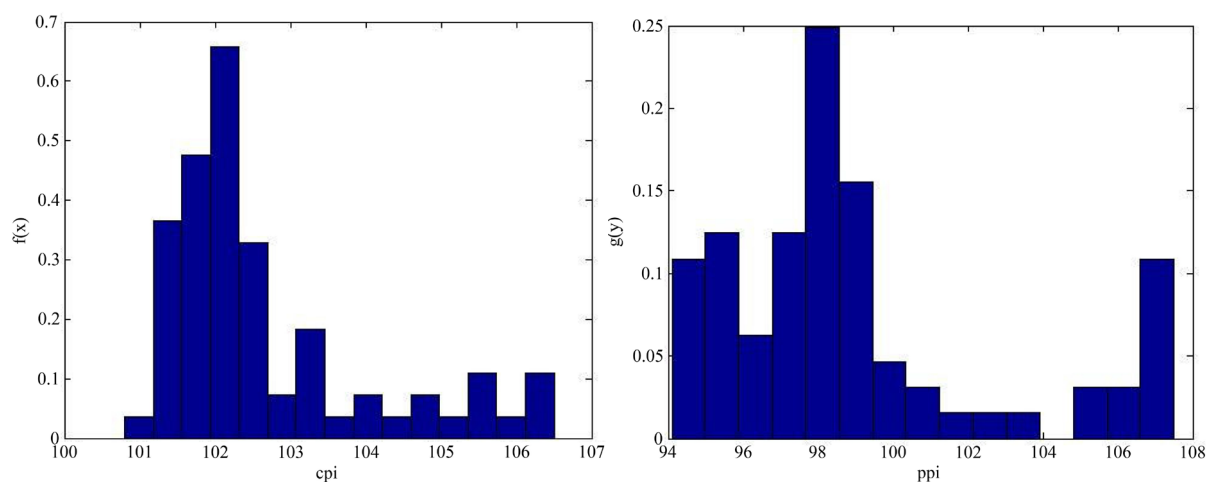


Figure 1. Curve: CPI and PPI frequency histogram

图 1. CPI 与 PPI 的频率直方图

表 1 给出了 CPI 与 PPI 的描述性统计量的值。通过图 1 和表 1 可以看出, CPI 与 PPI 均呈现出非正态分布的特征, 初步可以判断这两个变量不服从正态分布。

Table 1. Descriptive statistics of CPI and PPI

表 1. CPI 与 PPI 的描述性统计量

变量	最大值	最小值	均值	标准差	偏度	峰度
CPI	106.5	100.8	102.7	1.40	1.3329	3.7722
PPI	107.5	94.1	99.0	3.81	1.0351	3.1263

下面用 jbstest、kstest 和 lillietest 函数分别对 X, Y 进行正态性检验, 结果由表 2 给出。

Table 2. Test results of CPI and PPI normality

表 2. CPI 与 PPI 的正态性检验结果

	Jarque-Bera 检验	Kolmogorov-Smirnov 检验	Lilliefors 检验
CPI	0.0025	0.0015	1.0000e-03
PPI	0.0095	0.0039	1.0000e-03

对 X, Y 进行检验的结果为: $p \leq 0.01$, 说明 X, Y 的分布非正态, 在已知的分布族中也难以找到适合者。因此, 在估计 X, Y 的分布时, 采用核密度估计的方法。图 2 给出了 CPI 与 PPI 的核分布估计图。

3.2. Copula 函数的选取及参数估计

在确定了 X 的边缘分布 $F(x)$ 和 Y 的边缘分布 $G(y)$ 之后, 根据 $u_i = F(x_i), v_i = G(y_i)$ 的直方图选取

Copula 函数。图 3 给出了 CPI 与 PPI 的二维频率直方图，即 Copula 密度函数的估计。

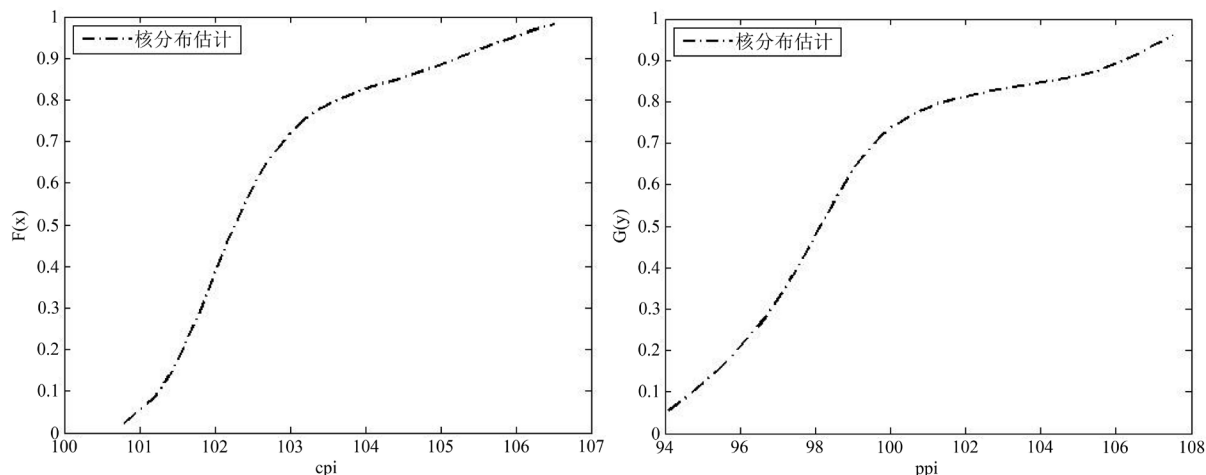


Figure 2. Curve: The core distribution of CPI and PPI is estimated
图 2. CPI 与 PPI 的核分布估计图

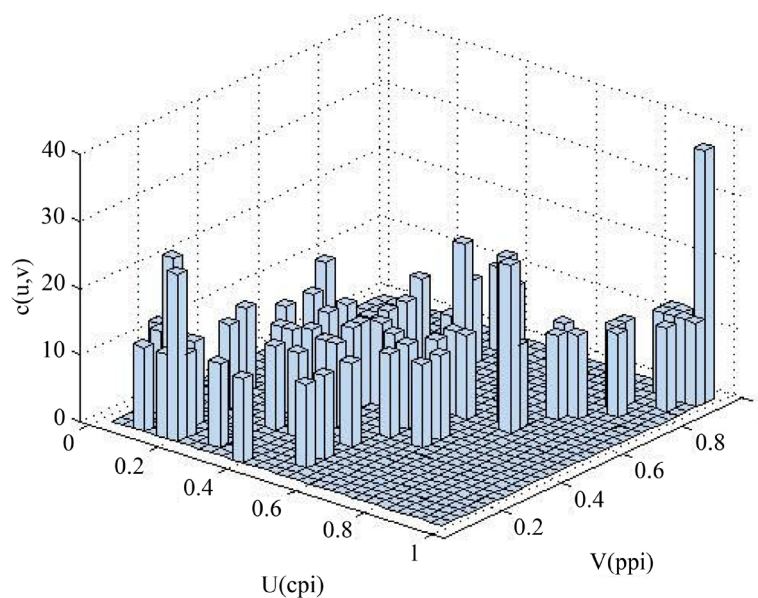


Figure 3. Curve: Binary histogram of marginal distribution of CPI and PPI
图 3. CPI 与 PPI 的边缘分布的二元直方图

在选择 Copula 函数形式时,我们考虑选择用椭圆 Copula 函数族中的 Gaussian Copula 函数和 t-Copula 函数, 或者选择 Archimedean Copula 函数族中的 Frank Copula 函数来刻画变量呈现对称尾部的特质; 用 Archimedean Copula 函数族中的 Gumble Copula 和 Glayton Copula, 或者选择混合 Copula 函数来刻画非对称的尾部特质。图 3 可以看出, $c(u,v)$ 的尾部较为对称, 所以初次尝试选用椭圆族中的 Gaussian Copula 函数或 t-Copula 函数来拟合 CPI 与 PPI 的相关关系。随后再选取 Archimedean Copula 族中的 Gumble Copula, Glayton Copula 或 Frank Copula 来拟合 CPI 与 PPI 的相关关系。Copula 模型中参数的估计值由极大似然估计得到。

估计的 $\hat{\rho}^{Ga} = 0.7958$ 代入(2)式, 得到 Gaussian Copula 函数为:

$$\hat{C}^{Ga}(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-0.7958^2}} \exp\left[-\frac{s^2 + t^2 - 2 \times 0.7958st}{2 \times (1-0.7958^2)}\right] dsdt.$$

估计的 $\hat{\rho}' = 0.8373, \hat{k} = 8.3252 \approx 8$ 代入(3)式, 得到 t-Copula 函数为:

$$\hat{C}^t(u, v) = \int_{-\infty}^{t_4^{-1}(u)} \int_{-\infty}^{t_4^{-1}(v)} \frac{1}{2\pi\sqrt{1-0.8373^2}} \left[1 + \frac{s^2 + t^2 - 2 \times 0.8373st}{4 \times (1-0.8373^2)}\right]^{\frac{8+2}{2}} dsdt.$$

估计的 $\hat{\alpha}^{Gu} = 2.7146, \hat{\alpha}^{Cl} = 1.9037, \hat{\alpha}^F = 8.2744$ 代入(4)式, 得到 Gumbel Copula、Clayton Copula、Frank Copula 函数分别为:

$$\begin{aligned} \hat{C}^{Gu}(u, v) &= \exp\left\{-\left[(-\ln u)^{2.7146} + (-\ln v)^{2.7146}\right]^{\frac{1}{2.7146}}\right\}; \\ \hat{C}^{Cl}(u, v) &= \left[u^{-1.9037} + v^{-1.9037} - 1\right]^{\frac{1}{1.9037}}; \\ \hat{C}^F(u, v) &= -\frac{1}{8.2744} \ln\left[1 + \frac{(1-e^{-8.2744u})(1-e^{-8.2744v})}{1-e^{-8.2744}}\right]. \end{aligned}$$

计算 CPI 与 PPI 原始观测数据的 τ, ρ_s 。同时, 分别计算出 Gaussian Copula、t-Copula、Gumbel Copula、Clayton Copula、Frank Copula 对应的 τ, ρ_s , 对比结果如表 3 所示。

Table 3. Kendall's rank correlation coefficient, Spearman's rank correlation coefficient and tail correlation coefficient
表 3. Kendall's 秩相关系数、Spearman's 秩相关系数、尾部相关系数

	τ	ρ_s	λ_u	λ_l
原始数据	0.5996	0.7644		
Gaussian Copula	0.5859	0.7816	0	0
t-Copula	0.6318	0.8250	0.3952	0.3952
Gumbel Copula	0.6795	0.8595	0.7076	0
Glayton Copula	0.6311	0.8157	0	0.6944
Frank Copula	0.6992	0.8866	0	0

原始数据求得的 τ, ρ_s 与五种常用 Copula 求得的 τ, ρ_s 对比发现, 在反映 CPI 和 PPI 的秩相关性上, Gaussian Copula 的 τ, ρ_s 与原始观测数据的 τ, ρ_s 接近程度会更好些, 说明 Gaussian Copula 能较好地刻画 CPI 与 PPI 之间的秩相关性。

在刻画尾部相关性时, Gumbel Copula 在分布上尾处的相关系数 $\lambda_u = 0.7076$, Clayton Copula 在分布下尾处的相关系数 $\lambda_l = 0.6944$, 效果要优于 t-Copula 的尾部相关系数 $\hat{\lambda}_u = \hat{\lambda}_l = 0.3952$ 。

3.3. Copula 函数模型的评价

在建立好模型之后, 要对 Copula 函数模型进行拟合优度的评价, 所用的方法是基于欧式平方距离准则, 选择与经验 Copula 之间距离最小的 Copula 函数作为最优。以 Gaussian Copula 为例:

$$d_{Ga}^2 = \sum_{i=1}^n \left| \hat{C}_n(u_i, v_i) - \hat{C}^{Ga}(u_i, v_i) \right|^2,$$

其中, $\hat{C}_n(u, v)$ 为经验 Copula 函数。

表 4 给出了不同 Copula 函数在基于平方欧式距离下与经验 Copula 函数的距离。

Table 4. Fitting evaluation of Copula function of CPI and PPI

表 4. CPI 与 PPI 的 Copula 函数的拟合评价

Copula 函数	参数估计	d^2
$C^{Ga}(u, v)$	$\rho = 0.7958$	0.0220
$C^t(u, v)$	$\rho = 0.8373, k = 8$	0.0221
$C^{Gu}(u, v)$	$\alpha = 2.7146$	0.0147
$C^{Cl}(u, v)$	$\alpha = 1.9037$	0.0925
$C^F(u, v)$	$\alpha = 8.2744$	0.0293

基于距离评判准则给出了最优 Copula 是与经验 Copula 之间距离最短的。因为 $d_{Gu}^2 < d_{Ga}^2 < d_t^2 < d_F^2 < d_{Cl}^2$ ，所以 Gumble Copula 更能拟合 CPI 与 PPI 的原始观测数据间的关系。

基金项目

本文受宁夏回族自治区研究生教育创新计划示范课程建设项目“数理统计引论”资助，项目编号为 YKC201603。

参考文献

- [1] 张尧庭. 连接函数(Copula)技术与金融风险分析[J]. 统计研究, 2002, 19(4): 48-51.
- [2] 韦艳华, 张世英, 孟利锋. Copula 理论在金融上的应用[J]. 西北农林科技大学学报(社会科学版), 2003, 3(5): 97-101.
- [3] 黄恩喜, 程希骏. 基于 Pair Copula-GARCH 模型的多资产组合 VaR 分析[J]. 中国科学研究生院学报, 2010, 27(4): 440-446.
- [4] 王璐, 黄登仕. 沪深股市相关性结构之谜: 基于贝叶斯 Copula 的研究[J]. 运筹与管理, 2014, 23(2): 213-219.
- [5] 罗长青, 朱慧明, 欧阳资生. 跳跃 - 扩散条件下信用风险相关性度量的变结构 Copula 模型[J]. 中国管理科学, 2014, 22(3): 1-12.
- [6] Harry, J. (2014) Dependence Modeling with Copulas. CRC Press, New York.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2251, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: sa@hanspub.org