Advances in Applied Mathematics 应用数学进展, 2023, 12(11), 4824-4833
Published Online November 2023 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2023.1211476

Hans汉斯

*Copulas*相依结构下完全和非完全样本极值的联合渐近分布

方 圆[1,2]

[1]浙江师范大学数学科学学院，金华 浙江
[2]嘉兴学院数据科学学院，嘉兴 浙江

收稿日期：2023年10月27日；录用日期：2023年11月21日；发布日期：2023年11月27日

摘 要

在日常生活中，观察某一组数据时，各种偶然因素或不可避免的因素都会导致数据出现随机丢失。通过研究满足阿基米德*Copula*相依结构极值的联合分布，可以减少或避免极端事件发生时所带来的损失。本文研究了随机缺失情形下，随机序列满足阿基米德*Copula*相依结构时完全样本与非完全样本极值的联合渐近分布，并给出几种典型的示例来说明主要结论。这不仅在理论中具有重要意义，在实际生活中也具有一定的意义。

关键词

随机缺失，*Copula*，极值

Asymptotic Distributions of Extremes of Complete and Incomplete Samples under *Copulas* Dependent

Yuan Fang[1,2]

[1]Department of Mathematic, Zhejiang Normal University, Jinhua Zhejiang
[2]College of Data Science, Jiaxing University, Jiaxing Zhejiang

Received: Oct. 27th, 2023; accepted: Nov. 21st, 2023; published: Nov. 27th, 2023

Abstract

In daily life, when observing a certain set of data, various accidental or unavoidable factors will

文章引用：方圆. *Copulas* 相依结构下完全和非完全样本极值的联合渐近分布[J]. 应用数学进展, 2023, 12(11): 4824-4833. DOI: 10.12677/aam.2023.1211476

lead to random missing. By studying the joint distribution of the extreme values of the Archimedean Copula dependence structures, the losses caused by extreme events can be reduced or avoided. This paper studies the asymptotic distributions of extremes of complete and incomplete samples under random missing. Some examples are given to illustrate the main results. This is not only of great significance in theory, but also has certain significance in real life.

Keywords

Random Missing, Copula, Extremes

1. 引言

经典的极值理论是研究社会生活中极端现象的一门学科，其主要研究一列随机序列极大值的相关极限理论。设 $\boldsymbol{X}=\{X_n,n\geq 1\}$ 是一列独立同分布的实值随机序列，具有边际分布函数 $F(x)$。若存在常数序列 $a_n>0, b_n\in R$ 和非退化分布函数 $G(x)$ 使得对 $G(x)$ 的任意连续点处有

$$\lim_{n\to\infty} F^n(a_n x+b_n)=G(x), \tag{1}$$

则称分布函数 F 属于非退化函数 $G(x)$ 的最大吸引场，简记为 $F\in D(G)$。注意到 $G(x)$ 必为如下三大极值分布

$$\text{Gumbel}:\Lambda(x)=\exp\left(-\mathrm{e}^{-x}\right),\ -\infty<x<+\infty;$$

$$\text{Frechet}:\Phi_\alpha(x)=\begin{cases}0, & x\leq 0,\\ \exp\left(-x^{-\alpha}\right), & x>0,\end{cases}\quad \alpha>0;$$

$$\text{Weibull}:\Psi_\alpha(x)=\begin{cases}\exp\left(-(-x)^{-\alpha}\right), & x\leq 0,\\ 1, & x>0,\end{cases}\quad \alpha>0.$$

上述的经典结果及其相关推广见专著[1]。

在实际应用中，有些数据可能因不同的原因以一种非常不规则的方式丢失。在许多领域(例如金融、水文、气象等)，不同的研究者可能对不同的观测频率的样本感兴趣。在这些情况下，研究完整样本极值和非完整样本的极值的渐近理论以及它们之间的渐近关系变得很重要。

假设 $\boldsymbol{X}$ 中仅有部分随机变量能够被观测到。令 $\boldsymbol{\varepsilon}=\{\varepsilon_n,n\geq 1\}$ 为一列伯努利序列，表示随机变量 X_n 被观测到的事件的指标且与 $\boldsymbol{X}=\{X_n,n\geq 1\}$ 独立。在 $\boldsymbol{X}$ 的前 n 个样本中，记 $M_n=\max\{X_k,k=1,2,\cdots,n\}$ 和随机缺失序列的最大值为 $\tilde{M}_n=\max\{X_k,\varepsilon_k=1,1\leq k\leq n\}$。$S_n=\sum_{k=1}^{n}\varepsilon_k$ 并假设其满足，当 $n\to\infty$ 时

$$\frac{S_n}{n}\xrightarrow{p}\lambda,$$

其中 λ 为随机或非随机变量。

当 λ 为一个常数时，文[2]研究了独立同分布随机序列完全样本和非完全样本极值之间的渐近关系，

获得了如下结论：对任意 $x<y\in\mathbb{R}$ 有

$$\lim_{n\to\infty}P\left(a_n\left(\tilde{M}_n-b_n\right)\le x, a_n\left(M_n-b_n\right)\le y\right)=H\left(x,y,\lambda\right), \tag{2}$$

其中

$$H\left(x,y,\lambda\right)=G^{\lambda}\left(x\right)G^{1-\lambda}\left(y\right).$$

同时文[2]也研究了一类平稳相依情形，即在极值理论领域一类非常经典的相依条件 $D\left(u_n,v_n\right)$ 和 $D'\left(u_n\right)$ (其定义详见专著[1])下研究了完全样本和非完全样本极值之间的渐近关系，证明了(2)仍然成立。

从那时起，该问题逐渐成为极值领域的一个研究热点，并被推广到一些其他随机情形。该问题在平稳高斯情形下的推广见文[3] [4] [5] [6]，在自回归过程和线性过程下的推广见文[7] [8]，在随机场下的推广见文[9] [10]，在几乎处处极限定理方面的推广见文[11] [12] [13]，在其他相关情形下的推广见文[14] [15]及其参考文献。

上述的研究中考虑了一些相依情形，如[2] [14]在 $D\left(u_n,v_n\right)$ 和 $D'\left(u_n\right)$ 条件下考虑该问题，[3] [4] [5] [6]则在相依高斯背景下考虑了该问题，[7] [8]在自回归过程和线性过程下考虑了该问题。除[3] [4] [5] [6]中的高斯情形之外，其他研究中考虑的相依性都很弱的，其不影响极限分布 $H\left(x,y,\lambda\right)$ 的形式。因此非常有必要在强相依背景下研究非高斯随机序列完全样本和非完全样本极值之间的渐近关系。

本文将探讨随机序列满足一类 *Copula* 相依结构时，完全样本和非完全样本极值之间的渐近关系。在过去的几年里，与 *Copula* 相关的主题引起了人们的极大兴趣。*Copula* 被用来描述随机变量之间的尺度不变依赖关系。对这种随机依赖结构的理解在概率论的所有领域中都变得非常重要。特别是在精算领域和金融领域的现代风险管理和压力测试方面，*Copula* 已经证明了它们在构建适当的多元模型方面的有效性。关于 *Copula* 的相关理论介绍，可参考专著[16]。

2. 多维 *Copula*

在本节中，我们将介绍一些关于 *Copula* 的定义，性质及相关定理，它们均来自专著[16]。*Copula* 是将多维分布函数与其边缘分布函数连接在一起的一种多元函数，自变量是边缘分布函数。它清晰反映两个边缘分布函数之间联系的结构，可以通过了解这些结构是如何影响二维分布函数的各种性质。

定义 2.1. (*Copula*) 令 $d\ge 2$。一个 d 维 *Copula* 函数是一个定义在 $[0,1]^d$ 上的 d 维分布函数，其边际分布函数服从 $(0,1)$ 上的均匀分布。

因此，对给定的 *Copula* 函数 C 和边际分布函数 $F_1,\cdots,F_d$ 有

$$F\left(x_1,\cdots,x_d\right)=C\left(F_1\left(x_1\right),\cdots,F_d\left(x_d\right)\right) \tag{3}$$

是一个分布函数。反之，对给定的具有边际分布函数 $F_1,\cdots,F_d$ 的多维分布函数 F，存在一个满足(3)的 *Copula* 函数 C。这个 *Copula* 函数 C 是不唯一的，如果 $F_1,\cdots,F_d$ 是连续函数，则有

$$C\left(x_1,\cdots,x_d\right)=F\left(F_1^{-1}\left(x_1\right),\cdots,F_d^{-1}\left(x_d\right)\right), \tag{4}$$

其中 F_i^{-1} 表示分布函数 F_i 的广义逆函数。(3)和(4)的证明见专著[16]。

常见的 *Copula* 有 *commonotonic Copula*

$$C\left(x_1,\cdots,x_d\right)=\min\left\{x_1,\cdots,x_d\right\}$$

和独立 *Copula*

$$C\left(x_1,\cdots,x_d\right)=x_1\cdot\cdots\cdot x_d.$$

本文主要考虑 *Archimedean Copula*。

定义 2.2. 令 $d \ge 2$。设 $\psi:[0,1]\to[0,\infty]$ 是严格递减的、凸的函数，并且使得 $\psi(0)=\infty, \psi(1)=0$ 对于 $x_i \in [0,1], i=1,\cdots,d$ 有

$$C_n^{\psi}(x_1,\cdots,x_d)=\psi^{-1}\left(\sum_{i=1}^{d}\psi(x_i)\right),$$

其中，ψ 称为 C_n^{ψ} 的生成元。

定义 2.3. 一个定义在 I 上函数 g 被称为在 I 上完全单调的，如果它是连续的并且具有交替符号的所有阶导数，即 $(-1)^k \frac{\mathrm{d}^k}{\mathrm{d}x^k} g(x) \ge 0$，对所有 $k \ge 0$ 和所有 $x \in I$ 成立。

定理 2.4. 对所有 $d \ge 2$，C_d^{ψ} 是一个 *Copula* 当且仅当生成元 ψ 具有逆函数 ψ^{-1} 并且 ψ^{-1} 在 $[0,\infty]$ 上完全单调。

定义 2.5. (*Archimedean Copula*) 如果 ψ^{-1} 在 $[0,\infty]$ 上完全单调，则称 C_d^{ψ} 是 *Archimedean Copula*。

Archimedean Copula 在实践中很有趣，通常它们只有一个参数，进而它们很容易构造。下面介绍几类常见的 *Archimedean Copula*。

定义 2.6. (*Gumbel Copula*) 若 $C_d^{Gu,\alpha}(x_1,x_2,\cdots,x_d)$ 具有如下形式

$$C_d^{Gu,\alpha}(x_1,x_2,\cdots,x_d)=\exp\left\{-\left[\sum_{i=1}^{d}(-\log x_i)^{\alpha}\right]\right\}$$

则称其为 *Gumbel Copula*，其生成元 $\psi(t)=(-\log(t))^{\alpha}, \alpha \ge 1$。

定义 2.7. (*Clayton copula*) 若 $C_d^{Cl,\alpha}(x_1,x_2,\cdots,x_d)$ 具有如下形式

$$C_d^{Cl,\alpha}(x_1,x_2,\cdots,x_d)=\left(x_1^{-\alpha}+\cdots+x_d^{-\alpha}-d+1\right)^{-\frac{1}{\alpha}}$$

则称其为 *Clayton copula*，其生成元 $\psi(t)=t^{-\alpha}-1, \alpha \ge 0$。

文[17]研究了满足上述 *Archimedean Copula* 结构的随机序列极值的极限分布问题，文[18] [19]则在 *Archimedean Copula* 结构下获得了随机序列极值的几乎处处中心极限定理。

3. 主要结论及其证明

在本节中，设 $\boldsymbol{X}=\{X_n, n\ge 1\}$ 是一列随机变量序列，并设 $\varepsilon=\{\varepsilon_n, n\ge 1\}$ 为一列伯努利随机序列，其中 ε_n 表示随机变量 X_n 被观测到的事件的指标且与 $X=\{X_n, n\ge 1\}$ 独立。令 $S_n=\sum_{k=1}^{n}\varepsilon_k$。记 $M_n=\max\{X_k, k=1,2,\cdots,n\}$，$\tilde{M}_n=\max\{X_k, \varepsilon_k=1, 1\le k\le n\}$。

现在，我们将阐述主要结论。

定理 3.1. 设 $X=\{X_n, n\ge 1\}$ 是一列同分布的随机变量序列具有连续的边际分布函数 $F(x)$，满足以下条件：

(i) 对所有的 $n \ge 1$，随机向量 $(X_1,X_2,\cdots,X_n)$ 具有 *Archimedean Copula* C_n^{ψ} 结构；

(ii) $\exp(-\psi\circ F)\in D(G)$，即存在常数序列 $c_n>0, d_n\in\mathbb{R}$ 使得对 $G(x)$ 的任意连续点处有

$$\lim_{n\to\infty}\exp\left(-n\psi\circ F(c_n x+d_n)\right)=G(x). \tag{5}$$

进一步假设 $S_n=\sum_{k=1}^{n}\varepsilon_k$ 满足

$$\frac{S_n}{n}\xrightarrow{p}\lambda, \tag{6}$$

其中 λ 为一个常数。则对 $x<y\in\mathbb{R}$，有

$$\lim_{n\to\infty}P\left(\tilde{M}_n\le c_nx+d_n,M_n\le c_ny+d_n\right)=\psi^{-1}\left(-\log\left(G^{\lambda}(x)G^{1-\lambda}(y)\right)\right). \tag{7}$$

注记 3.2. (i). 记 $\hat{M}_n=\max\left\{X_k,\varepsilon_k=0,1\le k\le n\right\}$，由定理易得：对 $x,y\in\mathbb{R}$，有

$$\lim_{n\to\infty}P\left(\tilde{M}_n\le c_nx+d_n,\hat{M}_n\le c_ny+d_n\right)=\psi^{-1}\left(-\log\left(G^{\lambda}(x)G^{1-\lambda}(y)\right)\right). \tag{8}$$

(ii). 结果(2)表明在弱相依情形下，非完全样本极值与完全样本极值之间是渐近独立；而定理 3.1 表明非完全样本极值与完全样本极值之间的渐近关系取决于生成元 ψ。对于 *Gumbel copula* $\psi(t)=(-\log t)^{\alpha}$，如果选取 $\alpha=1$，则非完全样本极值与完全样本极值之间是渐近独立的，否者是渐近相依的。详细情况可见下节中的例子。

(iii). 在实际应用中，我们除了对极大值感兴趣外，我们也常常对极值次序统计量感兴趣，因此，在 *Copula* 结构下考虑完全样本极值次序统计量和非完全样本极值次序统计量的渐近关系也是有意义的工作。但是，我们的证明方法对极值次序统计量不成立。

(iv). 将(6)中的常数 λ 推广到随机变量情形是非常有意义的工作，相关研究可见[2] [4] [14]。但是本文的方法对这种情况不成立。这个问题与(iii)中提到的问题将在另外一篇文章被解决。

(v). 文[17]给出了 $\exp(-\psi\circ F)\in D(G)$ 的充分必要条件，同时给出了选择常数列 c_n 和 d_n 的方法，详见文[17]中定理 4.4 和推论 4.5。

定理 3.1 的证明： 由全概率公式，(3)和(4)式可得

$$\begin{aligned}
&P\left(\tilde{M}_n\le c_nx+d_n,M_n\le c_ny+d_n\right)\\
&=\sum_{k=0}^{n}P(S_n=k)P\left(\tilde{M}_n\le c_nx+d_n,M_n\le c_ny+d_n\mid S_n=k\right)\\
&=\sum_{k=0}^{n}P(S_n=k)P\left(X_1\le c_nx+d_n,\cdots,X_k\le c_nx+d_n,X_{k+1}\le c_ny+d_n,\cdots,X_n\le c_ny+d_n\right)\\
&=\sum_{k=0}^{n}P(S_n=k)P\big(F(X_1)\le F(c_nx+d_n),\cdots,F(X_k)\le F(c_nx+d_n),\\
&\qquad\qquad F(X_{k+1})\le F(c_ny+d_n),\cdots,F(X_n)\le F(c_ny+d_n)\big)\\
&=\sum_{k=0}^{n}P(S_n=k)C_n^{\psi}\left(F(c_nx+d_n),\cdots,F(c_nx+d_n),F(c_ny+d_n),\cdots,F(c_ny+d_n)\right)\\
&=\sum_{k=0}^{n}P(S_n=k)\psi^{-1}\left(k\psi\left(F(c_nx+d_n)\right)+(n-k)\psi\left(F(c_ny+d_n)\right)\right)\\
&=\sum_{k=0}^{n}P(S_n=k)\psi^{-1}\left(-\log\left(H^k(c_nx+d_n)H^{n-k}(c_ny+d_n)\right)\right),
\end{aligned} \tag{9}$$

其中，$H(x)=\exp\left(-\psi\circ F(x)\right)$。设 $0<\varepsilon<\lambda$，则可将(9)右端改写为

$$\begin{aligned}
&\sum_{k=0}^{n}P(S_n=k)\psi^{-1}\left(-\log\left(H^k(c_nx+d_n)H^{n-k}(c_ny+d_n)\right)\right)\\
&=\sum_{k:\left|\frac{k}{n}-\lambda\right|>\varepsilon}P(S_n=k)\psi^{-1}\left(-\log\left(H^k(c_nx+d_n)H^{n-k}(c_ny+d_n)\right)\right)\\
&\quad+\sum_{k:\left|\frac{k}{n}-\lambda\right|\le\varepsilon}P(S_n=k)\psi^{-1}\left(-\log\left(H^k(c_nx+d_n)H^{n-k}(c_ny+d_n)\right)\right)\\
&=\Sigma_1+\Sigma_2
\end{aligned} \tag{10}$$

对第一项Σ_1，注意到对任意$x \in [0,\infty)$，$\left|\psi^{-1}(x)\right| \le 1$，进而利用条件(6)可得，当$n \to \infty$时

$$\Sigma_1 \le \sum_{k:\left|\frac{k}{n}-\lambda\right|>\varepsilon} P(S_n = k) \to 0. \tag{11}$$

对第二项Σ_2，利用ψ^{-1}的单调性可得

$$\Sigma_2 \le \psi^{-1}\left(-\log\left(H^{n(\lambda-\varepsilon)}(c_n x + d_n) H^{n-n(\lambda+\varepsilon)}(c_n y + d_n)\right)\right) \sum_{k:\left|\frac{k}{n}-\lambda\right|\le\varepsilon} P(S_n = k) \tag{12}$$

和

$$\Sigma_2 \le \psi^{-1}\left(-\log\left(H^{n(\lambda+\varepsilon)}(c_n x + d_n) H^{n-n(\lambda-\varepsilon)}(c_n y + d_n)\right)\right) \sum_{k:\left|\frac{k}{n}-\lambda\right|\le\varepsilon} P(S_n = k). \tag{13}$$

注意到由(5)可知，$H \in D(G)$，进而结合(6)及(9)-(13)式，对于任意$\varepsilon \in (0,\lambda)$，有

$$\limsup_{n\to\infty} P\left(\tilde{M}_n \le c_n x + d_n, M_n \le c_n y + d_n\right) \le \psi^{-1}\left(-\log\left(G^{\lambda-\varepsilon}(x) G^{1-\lambda-\varepsilon}(y)\right)\right), \tag{14}$$

$$\liminf_{n\to\infty} P\left(\tilde{M}_n \le c_n x + d_n, M_n \le c_n y + d_n\right) \ge \psi^{-1}\left(-\log\left(G^{\lambda+\varepsilon}(x) G^{1-\lambda+\varepsilon}(y)\right)\right). \tag{15}$$

最后，令$\varepsilon \to 0$，则可得

$$\limsup_{n\to\infty} P\left(\tilde{M}_n \le c_n x + d_n, M_n \le c_n y + d_n\right) \le \psi^{-1}\left(-\log\left(G^{\lambda}(x) G^{1-\lambda}(y)\right)\right), \tag{16}$$

$$\liminf_{n\to\infty} P\left(\tilde{M}_n \le c_n x + d_n, M_n \le c_n y + d_n\right) \ge \psi^{-1}\left(-\log\left(G^{\lambda}(x) G^{1-\lambda}(y)\right)\right). \tag{17}$$

故(7)得证。

4. 例子

在本节中，我们将给出满足定理 3.1 条件的几种类型的例子。

4.1. Weibull 吸引场情形

首先，我们给出$\exp(-\psi \circ F)$被吸引到 *Weibull* 分布的情形。

例 4.1 假设随机变量$X_1, X_2, \cdots, X_n$满足定理 3.1 中的条件。进一步假设其服从区间$(0,1)$上的均匀分布。

(i). 在 *Gumbel copula* 条件下，其中$\psi(t) = (-\log t)^{\alpha}, \alpha \ge 1$，则对$0 < x < y < \infty$有

$$\lim_{n\to\infty} P\left(\tilde{M}_n \le 1 - n^{-1/\alpha} x, M_n \le 1 - n^{-1/\alpha} y\right) = \exp\left\{-\left(\lambda x^{\alpha} + (1-\lambda) y^{\alpha}\right)^{\frac{1}{\alpha}}\right\}.$$

(ii). 在 *Clayton copula* 条件下，其中$\psi(t) = t^{-\alpha} - 1, \alpha > 0$，则对$0 < x < y < \infty$有

$$\lim_{n\to\infty} P\left(\tilde{M}_n \le 1 - (\alpha n)^{-1} x, M_n \le 1 - (\alpha n)^{-1} y\right) = \left(1 + \lambda x + (1-\lambda) y\right)^{-\frac{1}{\alpha}}.$$

证明: (i). 考虑$t > 0$，注意到$\psi(t) = (-\log t)^{\alpha}, \alpha \ge 1$以及$(0,1)$上均匀分布的上尾端点为 1，则

$$\lim_{x\to\infty} \frac{\psi \circ F\left(1-(xt)^{-1}\right)}{\psi \circ F\left(1-x^{-1}\right)} = \lim_{x\to\infty} \frac{\psi\left(1-(xt)^{-1}\right)}{\psi\left(1-x^{-1}\right)} = \lim_{x\to\infty} \frac{\left(-\log\left(1-(xt)^{-1}\right)\right)^{\alpha}}{\left(-\log\left(1-x^{-1}\right)\right)^{\alpha}} = \lim_{x\to\infty} \frac{(xt)^{-\alpha}}{x^{-\alpha}} = t^{-\alpha}.$$

因此，由文[17]中定理 4.4 可知$\exp(-\psi \circ F) \in D(\Psi_{\alpha})$，进而由文[17]中推论 4.5 可知其正则化常数$c_n = n^{-1/\alpha}, d_n = 1$代入(7)可证(i)成立。

(ii). 考虑$t>0$，注意到$\psi(t)=t^{-\alpha}-1,\alpha>0$，则

$$\lim_{x\to\infty}\frac{\psi\circ F\left(1-(xt)^{-1}\right)}{\psi\circ F\left(1-x^{-1}\right)}=\lim_{x\to\infty}\frac{\psi\left(1-(xt)^{-1}\right)}{\psi\left(1-x^{-1}\right)}=\lim_{x\to\infty}\frac{\left(1-(xt)^{-1}\right)^{-\alpha}-1}{\left(1-x^{-1}\right)^{-\alpha}-1}=\lim_{x\to\infty}\frac{\alpha(xt)^{-1}}{\alpha x^{-1}}=t^{-1}.$$

因此，由文[17]中定理 4.4 可知$\exp(-\psi\circ F)\in D(\Psi_1)$，进而由文[17]中推论 4.5 可知其正则化常数$c_n=(\alpha n)^{-1},d_n=1$。代入(7)可证(ii)成立。

4.2. *Fréchet* 吸引场情形

其次，我们给出$\exp(-\psi\circ F)$被吸引到 *Fréchet* 分布的情形。

例 4.2 假设随机变量$X_1,X_2,\cdots,X_n$满足定理 3.1 中的条件。进一步假设其边际分布为 *Pareto* 分布，即对于$K,\beta>0$，当$x\to\infty$时，有$1-F(x)\sim Kx^{-\beta}$。

(i). 在 *Gumbel copula* 条件下，其中$\psi(t)=(-\log t)^{\alpha},\alpha\ge1$，则对$0<x<y<\infty$有

$$\lim_{n\to\infty}P\left(\tilde{M}_n\le K^{\frac{1}{\beta}}n^{\frac{1}{\alpha\beta}}x,M_n\le K^{\frac{1}{\beta}}n^{\frac{1}{\alpha\beta}}y\right)=\exp\left\{-\left(\lambda x^{-\alpha\beta}+(1-\lambda)y^{-\alpha\beta}\right)^{\frac{1}{\alpha}}\right\}.$$

(ii). 在 *Clayton copula* 条件下，其中$\psi(t)=t^{-\alpha}-1,\alpha>0$，则对$0<x<y<\infty$有

$$\lim_{n\to\infty}P\left(\tilde{M}_n\le(\alpha Kn)^{\frac{1}{\beta}}x,M_n\le(\alpha Kn)^{\frac{1}{\beta}}y\right)=\left(1+\lambda x^{-\beta}+(1-\lambda)y^{-\beta}\right)^{-\frac{1}{\alpha}}.$$

证明: (i). 由文[17]中(5.11)可知$\exp(-\psi\circ F)\in D(\Phi_{\alpha\beta})$且其正则化常数$c_n=K^{\frac{1}{\beta}}n^{\frac{1}{\alpha\beta}},d_n=0$。代入(7)可证(i)成立。

(ii). 考虑$t>0$，注意到$\psi(t)=t^{-\alpha}-1,\alpha>0$，则

$$\lim_{x\to\infty}\frac{\psi\circ F(xt)}{\psi\circ F(x)}=\lim_{x\to\infty}\frac{\psi\left(1-K(xt)^{-\beta}\right)}{\psi\left(1-K(x)^{-\beta}\right)}=\lim_{x\to\infty}\frac{\left(1-K(xt)^{-\beta}\right)^{-\alpha}-1}{\left(1-K(x)^{-\beta}\right)^{-\alpha}-1}=\lim_{x\to\infty}\frac{1+\alpha K(xt)^{-\beta}-1}{1+\alpha K(x)^{-\beta}-1}=t^{-\beta}.$$

因此$\exp(-\psi\circ F)\in D(\Phi_\beta)$。由文[12]中推论 4.5 可知其正则化常数$c_n=(\alpha Kn)^{\frac{1}{\beta}}$，$d_n=0$。代入(7)可证(ii)成立。

4.3. *Gumbel* 吸引场情形

最后，我们给出$\exp(-\psi\circ F)$被吸引到 *Gumbel* 分布的情形。

例 4.3 在定理 3.1 的条件下，设随机变量$X_1,X_2,\cdots,X_n$的边际分布为参数为θ的指数分布，其中$\theta>0$。

(i). 在 *Gumbel copula* 条件下，其中$\psi(t)=(-\log t)^{\alpha},\alpha\ge1$， 则对$x<y\in\mathbb{R}$有

$$\lim_{n\to\infty}P\left(\tilde{M}_n\le c_nx+d_n,M_n\le c_ny+d_n\right)=\exp\left\{-\left(\lambda e^{-x}+(1-\lambda)e^{-y}\right)^{\frac{1}{\alpha}}\right\},$$

其中，$c_n=\frac{1}{\theta\alpha},d_n=\frac{1}{\theta\alpha}\log n$。

(ii). 在 *Clayton copula* 条件下，其中$\psi(t)=t^{-\alpha}-1,\alpha>0$，则对$x<y\in\mathbb{R}$有

$$\lim_{n\to\infty} P\left(\tilde{M}_n \le c_n x + d_n, M_n \le c_n y + d_n\right) = \left(1+\lambda e^{-x} + (1-\lambda)e^{-y}\right)^{-\frac{1}{\alpha}},$$

其中，$c_n = \frac{1}{\theta}, d_n = \frac{1}{\theta}\log(n\alpha)$。

证明：(i). 由文[17]中(5.17)可知 $\exp(-\psi\circ F)\in D(\Lambda)$ 且其正则化常数 $c_n = \frac{1}{\theta\alpha}, d_n = \frac{1}{\theta\alpha}\log n$。代入(7)可证(i)成立。

(ii). 考虑 $t>0$，注意到 $\psi(t) = t^{-\alpha}-1, \alpha>0$，则

$$\begin{aligned}\lim_{x\to\infty}\frac{\psi\circ F(xt)}{\psi\circ F(x)} &= \lim_{x\to\infty}\frac{\psi\left(1-\exp(-\theta xt)\right)}{\psi\left(1-\exp(-\theta x)\right)} = \lim_{x\to\infty}\frac{\left(1-\exp(-\theta xt)\right)^{-\alpha}-1}{\left(1-\exp(-\theta x)\right)^{-\alpha}-1}\\ &= \lim_{x\to\infty}\frac{1+\alpha\exp(-\theta xt)-1}{1+\alpha\exp(-\theta x)-1} = \exp\left(-\theta x(t-1)\right)\in\{0,1,\infty\}.\end{aligned}$$

因此 $\exp(-\psi\circ F)\in D(\Lambda)$。由文[17]中推论 4.5 可知其正则化常数 $c_n = \frac{1}{\theta}, d_n = \frac{1}{\theta}\log(n\alpha)$。

例 4.4 在定理 3.1 的条件下，设随机变量 $X_1, X_2, \cdots, X_n$ 的边际分布为标准正态序列。

(i). 在 *Gumbel copula* 条件下，其中 $\psi(t) = (-\log t)^{\alpha}, \alpha\ge 1$，则有

$$\lim_{n\to\infty} P\left(\tilde{M}_n \le c_n x + d_n, M_n \le c_n y + d_n\right) = \exp\left\{-\left(\lambda e^{-x} + (1-\lambda)e^{-y}\right)^{\frac{1}{\alpha}}\right\}, \tag{18}$$

其中 $c_n = (2\alpha\log n)^{-\frac{1}{2}}$，$d_n = \left(2\alpha^{-1}\log n\right)^{\frac{1}{2}} - \frac{\alpha}{2}(2\alpha\log n)^{-\frac{1}{2}}(\log 4\pi + \log\log n - \log\alpha)$。

(ii). 在 *Clayton copula* 条件下，其中 $\psi(t) = t^{-\alpha}-1, \alpha>0$，则有

$$\lim_{n\to\infty} P\left(\tilde{M}_n \le c_n x + d_n, M_n \le c_n y + d_n\right) = \left(1+\lambda e^{-x} + (1-\lambda)e^{-y}\right)^{-\frac{1}{\alpha}}, \tag{19}$$

其中，$c_n = (2\log n)^{-\frac{1}{2}}$，$d_n = (2\log n)^{\frac{1}{2}} - \frac{1}{2}(2\log n)^{-\frac{1}{2}}\left(\log\left(4\pi\alpha^{-2}\right) + \log\log n\right)$。

证明：(i). 考虑 $t>0$，注意到 $\psi(t) = (-\log t)^{\alpha}$，则

$$\begin{aligned}\lim_{x\to\infty}\frac{\psi\circ\Phi(xt)}{\psi\circ\Phi(x)} &= \lim_{x\to\infty}\frac{\psi\left(1-\left(1-\Phi(xt)\right)\right)}{\psi\left(1-\left(1-\Phi(x)\right)\right)} = \lim_{x\to\infty}\frac{\left(1-\Phi(xt)\right)^{\alpha}}{\left(1-\Phi(x)\right)^{\alpha}}\\ &= \lim_{x\to\infty}\frac{(xt)^{-\alpha}(2\pi)^{-\alpha/2}e^{-\alpha(tx)^2/2}}{(x)^{-\alpha}(2\pi)^{-\alpha/2}e^{-\alpha x^2/2}} = \lim_{x\to\infty} t^{-\alpha}e^{-\alpha x^2\left(t^2-1\right)/2}\in\{0,1,\infty\}.\end{aligned}$$

因此，由文[17]中定理 4.4 可得 $\exp(-\psi\circ\Phi)\in D(\Lambda)$。设 $u_n = u_n(x) = c_n x + d_n$，进而有 $n\psi\circ\Phi(u_n)\sim e^{-x}$，即 $n\left(1-\Phi(u_n)\right)^{\alpha}\sim e^{-x}$。根据文[1]中(1.5.4)式可知，当 $n\to\infty$ 时

$$\frac{e^{-x}u_n^{\alpha}}{n\left(\phi(u_n)\right)^{\alpha}}\to 1. \tag{20}$$

对上式两边取对数，有

$$-\log n - x + \alpha\log u_n + \frac{\alpha}{2}\log 2\pi + \frac{\alpha u_n^2}{2}\to 0. \tag{21}$$

进而有 $\dfrac{\alpha u_n^2}{2\log n} \to 1$，故

$$\log u_n = \frac{1}{2}\left(\log 2 + \log\log n - \log\alpha\right) + o(1). \tag{22}$$

将(22)代入(21)可得

$$\frac{u_n^2}{2} = \frac{1}{\alpha}x + \frac{1}{\alpha}\log n - \frac{1}{2}\log 4\pi - \frac{1}{2}\log\log n + \frac{1}{2}\log\alpha + o(1). \tag{23}$$

则

$$u_n = \left(2\alpha^{-1}\log n\right)^{\frac{1}{2}}\left\{1 + \frac{x - \frac{\alpha}{2}\log 4\pi - \frac{\alpha}{2}\log\log n + \frac{\alpha}{2}\log\alpha}{2\log n} + o\left(\frac{\alpha}{\log n}\right)\right\}. \tag{24}$$

而 $u_n = c_n x + d_n$，故(i)得证。

(ii). 考虑 $t > 0$，注意到 $\psi(t) = t^{-\alpha} - 1, \alpha > 0$，则

$$\begin{aligned}\lim_{x\to\infty}\frac{\psi\circ\Phi(xt)}{\psi\circ\Phi(x)} &= \lim_{x\to\infty}\frac{\psi\left(1-\left(1-\Phi(xt)\right)\right)}{\psi\left(1-\left(1-\Phi(x)\right)\right)} = \lim_{x\to\infty}\frac{\left(1-\Phi(xt)\right)}{\left(1-\Phi(x)\right)} \\ &= \lim_{x\to\infty}\frac{(xt)^{-1}(2\pi)^{-1/2}\mathrm{e}^{-(tx)^2/2}}{x^{-1}(2\pi)^{-1/2}\mathrm{e}^{-x^2/2}} = \lim_{x\to\infty} t^{-1}\mathrm{e}^{-x^2\left(t^2-1\right)/2} \in \{0,1,\infty\}.\end{aligned}$$

因此，由文[17]中定理 4.4 可得 $\exp(-\psi\circ\Phi) \in D(\Lambda)$。设 $u_n = u_n(x) = c_n x + d_n$，进而有 $n\psi\circ\Phi(u_n) \sim \mathrm{e}^{-x}$，即 $n\alpha\left(1-\Phi(u_n)\right) \sim \mathrm{e}^{-x}$。余下的证明与(i)完全一样，故略去。

由于只有少数的统计工具可以用来测试依赖性结构，因此很难将满足 *Copula* 结构的数据匹配到真实数据中，此时则需通过考虑它们的渐近行为。本文研究了随机缺失情形下，随机序列满足阿基米德 *Copula* 相依结构完全样本与非完全样本极值的联合渐近分布，将满足阿基米德 *Copula* 相依结构极值的分布推广到了随机缺失的情形下。结果(2)表明在弱相依情形下，非完全样本极值与完全样本极值之间是渐近独立；而本文定理 3.1 则是在另一种相依情形下表明了非完全样本极值与完全样本极值之间的渐近关系取决于生成元 ψ。

在 *Copula* 结构下，运用本文中的方法考虑完全样本极值次序统计量和非完全样本极值次序统计量的渐近关系以及将常数 λ 推广到随机变量情形是不成立的。今后我将会继续研究满足阿基米德 *Copula* 相依结构下的相关问题，解决上述所提到的问题，且会更加全面地考虑问题，更好地将理论与实践结合在一起。

基金项目

本文受浙江省自然科学基金(编号 LY18A010020)。

参考文献

[1] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983) Extremes and Related Properties of Random Sequences and Processes. Springer-Verlag, NewYork. https://doi.org/10.1007/978-1-4612-5449-2

[2] Mladenović, P. and Piterbarg, V. (2006) On Asymptotic Distribution of Maxima of Complete and Incomplete Samples from Stationary Sequences. *Stochastic Processes and Their Applications*, **116**, 1977-1991.

https://doi.org/10.1016/j.spa.2006.05.009

[3] Cao, L. and Peng, Z. (2011) Asymptotic Distributions of Maxima of Complete and Incomplete Samples from Strongly Dependent Stationary Gaussian Sequences. *Applied Mathematics Letters*, **24**, 243-247. https://doi.org/10.1016/j.aml.2010.09.012

[4] Hashorva, E., Peng, Z. and Weng, Z. (2013) On Piterbarg Theorem for Maxima of Stationary Gaussian Sequences. *Lithuanian Mathematical Journal*, **53**, 280-292. https://doi.org/10.1007/s10986-013-9208-6

[5] Peng, Z., Cao, L. and Nadarajah, S. (2010) Asymptotic Distributions of Maxima of Complete and Incomplete Samples from Multivariate Stationary Gaussian Sequences. *Journal of Multivariate Analysis*, **101**, 2641-2647. https://doi.org/10.1016/j.jmva.2010.06.016

[6] Peng, Z., Tong, J. and Weng, Z. (2019) Exceedances Point Processes in the Plane of Stationary Gaussian Sequences with Data Missing. *Statistics and Probability Letters*, **149**, 73-79. https://doi.org/10.1016/j.spl.2019.01.022

[7] Glavaš, L., Mladenović, P. and Samorodnitsky, G. (2017) Extreme Values of the Uniform Order 1 Autoregressive Processes and Missing Observations. *Extremes*, **20**, 671-690. https://doi.org/10.1007/s10687-016-0282-0

[8] Glavaš, L. and Mladenović, P. (2020) Extreme Values of Linear Processes with Heavy-Tailed Innovations and Missing Observations. *Extremes*, **23**, 547-567. https://doi.org/10.1007/s10687-020-00390-3

[9] Panga, Z. and Pereira, L. (2018) On the Maxima and Minima of Complete and Incomplete Samples from Nonstationary Random Fields. *Statistics and Probability Letters*, **137**, 124-134. https://doi.org/10.1016/j.spl.2018.01.019

[10] Zheng, S. and Tan, Z. (2023) On the Maxima of Nonstationary Random Fields Subject to Missing Observations. Communications in Statistics-Theory and Methods. https://doi.org/10.1080/03610926.2023.2244098

[11] Peng, Z., Wang, P. and Nadarajah, S. (2009) Limiting Distributions and Almost Sure Limit Theorems for the Normalized Maxima of Complete and Incomplete Samples from Gaussian Sequence. *Electronic Journal of Statistics*, **3**, 851-864. https://doi.org/10.1214/09-EJS443

[12] Dudzinski, M. (2017) Some Applications of the Archimedean Copulas in the Proof of the Almost Sure Central Limit Theorem for Ordinary Maxima. *Open Mathematics*, **15**, 1024-1034. https://doi.org/10.1515/math-2017-0085

[13] Tong, B. and Peng, Z. (2011) On Almost Sure Max-Limit Theorems of Complete and Incomplete Samples from Stationary Sequences. *Acta Mathematica Sinica* (*English Series*), **27**, 1323-1332. https://doi.org/10.1007/s10114-011-8616-y

[14] Krajka, T. (2011) The Asymptotic Behaviour of Maxima of Complete and Incomplete Samples from Statioanry Sequences. *Stochastic Processes and Their Applications*, **121**, 1705-1719. https://doi.org/10.1016/j.spa.2011.04.001

[15] 谭中权. 连续与离散时间 Gauss 次序统计过程的极值[J]. 中国科学(数学), 2018, 48(5): 623-642.

[16] Nelsen, R.B. (1999) An Introduction to Copulas. Springer, New York. https://doi.org/10.1007/978-1-4757-3076-0

[17] Wüthrich, M.V. (2004) Extreme Value Theory and Archimedean Copulas. *Scandinavian Actuarial Journal*, **104**, 211-228. https://doi.org/10.1080/03461230110106534

[18] Dudzinski, M. and Furmanczyk, K. (2017) On Some Applications of the Archimedean Copulas in the Proofs of the Almost Sure Central Limit Theorems for Certain Order Statistics. *Bulletin of the Korean Mathematical Society*, **54**, 839-874. https://doi.org/10.4134/BKMS.b160329

[19] Dudzinski, M. (2017) Some Applications of the Archimedean Copulas in the Proof of the Almost Sure Central Limit Theorem for Ordinary Maxima. *Open Mathematics*, **15**, 1024-1034. https://doi.org/10.1515/math-2017-0085