Advances in Applied Mathematics 应用数学进展, 2024, 13(10), 4674-4682
Published Online October 2024 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2024.1310448

一类可交换随机序列极值次序统计量的极限分布及其应用

陶　颖[1,2*]

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

收稿日期：2024年9月25日；录用日期：2024年10月17日；发布日期：2024年10月28日

摘　要

设 $X=\{X_n,n\ge 1\}$ 是一列可交换的随机变量，并假设 X 中仅仅有部分随机变量能够被观测到。在随机缺失情形下，本文证明了完全样本极值次序统计量与非完全样本极值次序统计量的联合极限分布，并用所得结果研究了阿基米德*Copula*相依结构下完全样本极值次序统计量与非完全样本极值次序统计量的渐近关系。

关键词

可交换随机变量，阿基米德*Copula*，极值次序统计量，随机缺失

The Joint Asymptotic Distributions of Extreme Order Statistics of Complete and Incomplete Samples from Exchangeable Variables and Their Applications

Ying Tao[1,2*]

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

Received: Sep. 25th, 2024; accepted: Oct. 17th, 2024; published: Oct. 28th, 2024

Abstract

Let $X=\{X_n,n\ge 1\}$ be a sequence of exchangeable variables and suppose that only parts of them

*通讯作者。

文章引用: 陶颖. 一类可交换随机序列极值次序统计量的极限分布及其应用[J]. 应用数学进展, 2024, 13(10): 4674-4682. DOI: 10.12677/aam.2024.1310448

can be observed. In this paper, we derived the joint asymptotic distributions of extreme order statistics of complete and incomplete samples under conditional independence. We also investigate the joint asymptotic relation between extreme order statistics of complete and incomplete samples under Archimedean copulas.

Keywords

Exchangeable Variables, Archimedean Copulas, Extreme Order Statistics, Random Missing

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 $a_n>0, b_n\in\mathbb{R}$ 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

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

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

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

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，研究完全样本极值和非完全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 $\widetilde{M_n}=\max\{X_k, \varepsilon_k=1, 1\leq k\leq n\}$。$S_n=\sum_{k=1}^n \varepsilon_k$ 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，文献[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

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

其中

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

同时文献[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[3])下研究了完全样本和非完全样本极值之间的渐近关系，文献[4]证明了(2)式仍然成立，其中

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

近年来，该问题逐渐成为极值领域的一个研究热点，文献[3] [5]-[7]将上述问题推广到了高斯情形；文献[8] [9]则考虑了自回归过程和线性过程情形；文献[10] [11]考虑了随机场情形；其他相关研究见[12]-[15]。

上述的研究中考虑了一些相依情形，如[2] [4]在 $D\left(u_n,v_n\right)$ 和 $D'\left(u_n\right)$ 条件下考虑该问题，[3] [5]-[7]则在相依高斯背景下考虑了该问题，[8] [9]在自回归过程和线性过程下考虑了该问题。除[3] [5]-[7]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。

本文将研究一类可交换随机序列完全样本和非完全样本极值次序统计量之间的渐近关系。专著[16]中 3.6 节详细地陈述了可交换随机变量次序统计量的渐近性质。受文献[17]的启发，本文将使用拉普拉斯变换来描述随机变量之间的可交换性，该拉普拉斯变换与阿基米德 *Copula* 有着密切的关系。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* 常常被用来描述随机变量之间的尺度不变的依赖性。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，*Copula* 已经证明了它们在构建适当的多元模型方面是非常有效的。关于 *Copula* 的相关理论介绍，可参考专著[18]。

2. 主要结论及其证明

在本节中，设 $\boldsymbol{X}=\left\{X_n,n\ge1\right\}$ 是一列随机变量序列，并设 $\varepsilon=\left\{\varepsilon_n,n\ge1\right\}$ 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 $\boldsymbol{X}=\left\{X_n,n\ge1\right\}$ 独立。令 $S_n=\sum_{k=1}^{n}\varepsilon_k$。记 $M_n=\max\left\{X_k,k=1,2,\cdots,n\right\}$，$\widetilde{M_n}=\max\left\{X_k,\varepsilon_k=1,1\le k\le n\right\}$。对于任给定的正整数 k,l,m，令 $\widetilde{M_n}^{(k)},\widehat{M_n}^{(l)},M_n^{(m)}$ 分别表示 $\left\{X_j:\varepsilon_i=1,1\le j\le n\right\}$，$\left\{X_j:\varepsilon_i=0,1\le j\le n\right\}$，$\left\{X_j:1\le j\le n\right\}$ 中的第 k,l 和 m 个最大值。

现在，我们将阐述主要结论。

定理 2.1. 设 F 和 $M_{\Theta}\left(\theta\right)$ 均是一维分布函数，且 $M_{\Theta}\left(0\right)=0$。记 ψ 是 $M_{\Theta}\left(\theta\right)$ 的拉普拉斯变换，且 ψ^{-1} 是 ψ 的反函数。记 $\psi^{(k)}$ 为 ψ 的 k 阶导数。假设随机向量 $\left(X_1,X_2,\cdots,X_n\right)$ 有如下的联合分布函数:

$$H\left(x_1,\cdots,x_n\right)=\int_0^{+\infty}\prod_{i=1}^{n}\exp\left\{-\theta\psi^{-1}\left(F\left(x_i\right)\right)\right\}\mathrm{d}M_{\Theta}\left(\theta\right). \tag{3}$$

假设 $\exp\left(-\psi^{-1}\circ F\right)\in D\left(G\right)$，即存在常数序列 $c_n>0, d_n\in\mathbb{R}$ 使得对 $G\left(x\right)$ 的任意连续点处有

$$\lim_{n\to\infty}\exp\left(-n\psi^{-1}\circ F\left(c_nx+d_n\right)\right)=G\left(x\right). \tag{4}$$

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

$$\frac{S_n}{n}\xrightarrow{p}\lambda, \tag{5}$$

其中随机变量 $\lambda\in(0,1)$ 几乎处处成立。则

1) 对任意的 $x,y\in\mathbb{R}$，有

$$\begin{aligned}&\lim_{n\to\infty}P\left(\widetilde{M_n}^{(k)}\le c_nx+d_n,\widehat{M_n}^{(l)}\le c_ny+d_n\right)\\&=\sum_{r=0}^{k-1}\sum_{s=0}^{l-1}\frac{(\log G(x))^r(\log G(y))^s}{r!s!}\\&\quad\times E_\lambda\left\{\lambda^r(1-\lambda)^s\psi^{(r+s)}\left(\log G^{-\lambda}(x)+\log G^{-(1-\lambda)}(y)\right)\right\};\end{aligned} \tag{6}$$

2) 当 $x<z\in\mathbb{R}$ 时，有

$$\begin{aligned}&\lim_{n\to\infty}P\left(\widetilde{M_n}^{(k)}\le c_nx+d_n,M_n^{(m)}\le c_nz+d_n\right)\\&=\sum_{t=0}^{m-1}\sum_{i=0}^{m-t-1}\sum_{j=i}^{k-1}\frac{(\log G(z))^t}{t!}\frac{(\log G(z))^i}{i!}\frac{(\log G(x)-\log G(z))^{j-i}}{(j-i)!}\\&\quad\times E_\lambda\left\{\lambda^j(1-\lambda)^t\psi^{(t+j)}\left(\log G^{-\lambda}(x)+\log G^{-(1-\lambda)}(z)\right)\right\};\end{aligned} \tag{7}$$

当 $x\ge z\in\mathbb{R}$ 时，有

$$\begin{aligned}&\lim_{n\to\infty}P\left(\widetilde{M_n}^{(k)}\le c_nx+d_n,M_n^{(m)}\le c_nz+d_n\right)\\&=\sum_{t=0}^{m-1}\sum_{i=0}^{m-t-1}\sum_{j=0}^{\min\{i,k-1\}}\frac{(\log G(z))^t}{t!}\frac{(\log G(x))^j}{j!}\frac{(\log G(z)-\log G(x))^{i-j}}{(i-j)!}\\&\quad\times E_\lambda\left\{\lambda^i(1-\lambda)^t\psi^{(t+i)}\left(-\log G(z)\right)\right\};\end{aligned} \tag{8}$$

3) 当 $y<z\in\mathbb{R}$ 时，有

$$\begin{aligned}&\lim_{n\to\infty}P\left(\widehat{M_n}^{(l)}\le c_ny+d_n,M_n^{(m)}\le c_nz+d_n\right)\\&=\sum_{t=0}^{m-1}\sum_{i=0}^{m-t-1}\sum_{j=i}^{l-1}\frac{(\log G(z))^t}{t!}\frac{(\log G(z))^i}{i!}\frac{(\log G(y)-\log G(z))^{j-i}}{(j-i)!}\\&\quad\times E_\lambda\left\{\lambda^t(1-\lambda)^j\psi^{(t+j)}\left(\log G^{-\lambda}(z)+\log G^{-(1-\lambda)}(y)\right)\right\};\end{aligned} \tag{9}$$

当 $y\ge z\in\mathbb{R}$ 时，有

$$\begin{aligned}&\lim_{n\to\infty}P\left(\widehat{M_n}^{(l)}\le c_ny+d_n,M_n^{(m)}\le c_nz+d_n\right)\\&=\sum_{t=0}^{m-1}\sum_{i=0}^{m-t-1}\sum_{j=0}^{\min\{i,l-1\}}\frac{(\log G(z))^t}{t!}\frac{(\log G(y))^j}{j}\frac{(\log G(z)-\log G(y))^{i-j}}{(i-j)!}\\&\quad\times E_\lambda\left\{\lambda^t(1-\lambda)^i\psi^{(t+i)}\left(-\log G(z)\right)\right\}.\end{aligned} \tag{10}$$

注记 2.2. (i) (3)式定义了一个 n 维可交换的随机向量。当给定条件 $\Theta=\theta$ 时，他们是独立同分布的。由于独立同分布序列能够被扩展到无穷维序列，因此向量 $(X_1,X_2,\cdots,X_n)$ 事实上是无限维可交换随机变量序列的一部分；

(ii) 拉普拉斯变换是严格单调递减的，并且有 $\psi(0)=1, \psi(\infty)=0$；

(iii) 可以借助阿基米德 copula 来描述(3)式所定义的相依结构，本文第三节给出了定理 2.1 在阿基米德 copula 结构下的应用；

(iv) 对(3)式定义的可交换的相依结构，文献[17]给出相应极值次序统计量的极限分布。

证明：为了方便，记 $u_n(x)=c_n x+d_n$。由文献[19]可知，如果随机序列 $Y=\{Y_n, n\ge 1\}$ 的边际分布函数 $F_X \in D(G)$，并且 $Y=\{Y_n, n\ge 1\}$ 满足弱相依条件 $D(u_n, v_n)$ 和局部相依 $D'(u_n)$ 时，任意的 $x, y\in\mathbb{R}$，有

$$\begin{aligned}&\lim_{n\to\infty} P\left(\widetilde{M_n}^{(k)}(Y)\le u_n(x), \widehat{M_n}^{(l)}(Y)\le u_n(y)\right)\\&=\sum_{r=0}^{k-1}\sum_{s=0}^{l-1} E_\lambda\left\{G^\lambda(x)\frac{\left(-\log G^\lambda(x)\right)^r}{r!}G^{1-\lambda}(y)\frac{\left(-\log G^{1-\lambda}(y)\right)^s}{s!}\right\}.\end{aligned} \tag{11}$$

由定理 2.1 的条件(3)可知，在给定 Θ 的条件下，$X_1,\cdots,X_n$ 是相互独立的随机变量，并且其边际分布函数为 $\exp\left(-\Theta\psi^{-1}\circ F\right)$。注意到 $\exp\left(-\psi^{-1}\circ F\right)\in D(G)$，进而有 $\exp\left(-\Theta\psi^{-1}\circ F\right)\in D\left(G^\Theta\right)$。这意味着在 Θ 给定的条件下，$X_1,\cdots,X_n$ 满足经典的 Fisher-Tippet 定理，并且有着与(1)式一样的正规化常数 a_n, b_n，且此时极限分布函数为 G^Θ。详细情形可参考[20]定理 3.2.3 和[16]第 3.6 节。因此，利用(11)式，有

$$\begin{aligned}&\lim_{n\to\infty} P\left(\widetilde{M_n}^{(k)}\le u_n(x), \widehat{M_n}^{(l)}\le u_n(y)\mid\Theta\right)\\&=\sum_{r=0}^{k-1}\sum_{s=0}^{l-1} E_\lambda\left\{G^{\Theta\lambda}(x)\frac{\left(-\log G^{\Theta\lambda}(x)\right)^r}{r!}G^{\Theta(1-\lambda)}(y)\frac{\left(-\log G^{\Theta(1-\lambda)}(y)\right)^s}{s!}\right\},\end{aligned} \tag{12}$$

即，我们获得了关于 Θ 的逐点收敛。注意到上述概率是一致有界的，并且在 $[0,\infty)$ 上关于 $\mathrm{d}M_\Theta(\theta)$ 是可积的。因此，根据控制收敛定理和(12)式，有

$$\begin{aligned}&\lim_{n\to\infty} P\left(\widetilde{M_n}^{(k)}\le u_n(x), \widehat{M_n}^{(l)}\le u_n(y)\mid\Theta\right)\\&=\lim_{n\to\infty}\int_0^{+\infty} P\left(\widetilde{M_n}^{(k)}\le u_n(x), \widehat{M_n}^{(l)}\le u_n(y)\mid\Theta\right)\mathrm{d}M_\Theta(\theta)\\&=\int_0^{+\infty}\lim_{n\to\infty} P\left(\widetilde{M_n}^{(k)}\le u_n(x), \widehat{M_n}^{(l)}\le u_n(y)\mid\Theta\right)\mathrm{d}M_\Theta(\theta)\\&=\int_0^{+\infty}\sum_{r=0}^{k-1}\sum_{s=0}^{l-1} E_\lambda\left\{G^{\theta\lambda}(x)\frac{\left(-\log G^{\theta\lambda}(x)\right)^r}{r!}G^{\theta(1-\lambda)}(y)\frac{\left(-\log G^{\theta(1-\lambda)}(y)\right)^s}{s!}\right\}\mathrm{d}M_\Theta(\theta)\\&=\sum_{r=0}^{k-1}\frac{\left(\log G(x)\right)^r}{r!}\sum_{s=0}^{l-1}\frac{\left(\log G(y)\right)^s}{s!}\\&\quad\times E_\lambda\left\{\int_0^{+\infty}\lambda^r(1-\lambda)^s(-\theta)^{(r+s)}\exp\left(-\theta\left[\log G^{-\lambda}(x)+\log G^{-(1-\lambda)}(y)\right]\right)\mathrm{d}M_\Theta(\theta)\right\}.\end{aligned} \tag{13}$$

根据定理 3.1 条件，ψ 是 $M_\Theta(\theta)$ 的拉普拉斯变换，即

$$\psi(t)=\int_0^{+\infty}\exp(-t\theta)\mathrm{d}M_\Theta(\theta)$$

进而有

$$\psi^{(k)}(t)=\int_0^{+\infty}(-\theta)^k\exp(-t\theta)\mathrm{d}M_\Theta(\theta)$$

带入(13)式可得

$$\begin{aligned}&\lim_{n\to\infty}P\left(\widetilde{M_n}^{(k)}\le u_n(x),\widehat{M_n}^{(l)}\le u_n(y)\right)\\&=\sum_{r=0}^{k-1}\frac{(\log G(x))^r}{r!}\sum_{s=0}^{l-1}\frac{(\log G(y))^s}{s!}\\&\times E_\lambda\left\{\lambda^r(1-\lambda)^s\psi^{(r+s)}\left(\log G^{-\lambda}(x)+\log G^{-(1-\lambda)}(y)\right)\right\}.\end{aligned}$$

为了完成(6)式的证明，我们还需要证明当$(x,y)\to(\infty,\infty)$时，

$$\lim_{n\to\infty}P\left(\widetilde{M_n}^{(k)}\le u_n(x),\widehat{M_n}^{(l)}\le u_n(y)\right)=1$$

利用文献[17]中定理 1.2 可得

$$\begin{aligned}&\lim_{n\to\infty}P\left(\widetilde{M_n}^{(k)}\le u_n(x),\widehat{M_n}^{(l)}\le u_n(y)\right)\\&\ge\lim_{n\to\infty}P\left(\widetilde{M_n}\le u_n(x),\widehat{M_n}\le u_n(y)\right)\\&\ge\lim_{n\to\infty}P\left(\widetilde{M_n}\le u_n(\min\{x,y\}),\widehat{M_n}\le u_n(\min\{x,y\})\right)\\&=\lim_{n\to\infty}P\left(M_n\le u_n(\min\{x,y\})\right)\\&=\psi\left(-\log G(\min\{x,y\})\right).\end{aligned}$$

注意到，当$(x,y)\to(\infty,\infty)$时，显然有$\psi\left(-\log G(\min\{x,y\})\right)\to 1$。定理 3.1 的 1)证毕。定理 3.1 的 2)与 3)的证明是类似的，故略去。

3. 应用

在本节中，我们将应用定理 2.1 来探讨 Copula 相依结构下完全样本极值次序统计量与非完全样本极值次序统计量之间的渐近关系。首先，介绍一些关于 Copula 的定义，性质及相关定理，它们均来自专著[18]。Copula 是一个自变量为边缘分布函数的多元函数，它能够将多维分布函数与其边缘分布函数进行有效的连接，从而清晰地反映两个边缘分布函数之间联系的结构。

定义 3.1. (*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。

下面的 Sklar 定理是一个研究 *Copula* 的重要工具，其证明见专著[18]。

命题 3.1. 对给定的 *Copula* 函数 C 和边际分布函数 $F_1,\cdots,F_d$，有

$$F(x_1,\cdots,x_d)=C\left(F_1(x_1),\cdots,F_d(x_d)\right)\tag{14}$$

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(14)式的 *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

$$C(x_1,\cdots,x_d)=F\left(F_1^{-1}(x_1),\cdots,F_d^{-1}(x_d)\right),\tag{15}$$

其中 F_i^{-1} 表示分布函数 F_i 的广义逆函数。

本文主要考虑阿基米得 *Copula*，在介绍其定义之前，我们先给出几个辅助结果。

定义 3.2。令$d\ge 2$。设$\psi^{-1}:[0,1]\to[0,\infty]$是严格递减的、凸的函数，并且使得$\psi^{-1}(0)=\infty$，$\psi^{-1}(1)=0$

对于 $x_i \in [0,1], i=1,\cdots,d$ 有

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

其中，ψ^{-1} 称为 C_d^{ψ} 的生成元。

定义 3.3. 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。

命题 3.2. 对所有 $d \ge 2$，C_d^{ψ} 是一个 *Copula* 当且仅当生成元 ψ^{-1} 具有逆函数 ψ 并且 ψ 在 $[0,\infty]$ 上完全单调.

定义 3.4. (阿基米德 Copula)如果 ψ 在 $[0,\infty]$ 上完全单调，则称 C_d^{ψ} 是阿基米德 *Copula*。

命题 3.3. 设 $\boldsymbol{X}=\{X_n, n\ge 1\}$ 是一列同分布的随机变量，如果对任意的 $n \ge 2$，随机向量 $(X_1,\cdots,X_n)$ 具有阿基米德 Copula C_n^{ψ}，则称随机序列 $\boldsymbol{X}=\{X_n, n\ge 1\}$ 是阿基米德 Copula 相依的。

当随机变量序列是阿基米德 Copula 相依的，如果其共同分布函数 F_X 满足 $\exp\left(-\psi^{-1}\circ F_X\right)\in D(G)$，文献[21]研究了该随机序列极值的极限分布问题，并探讨了其结果在精算中的应用。如果其共同分布函数 F_X 满足 $F_X \in D(G)$，文献[22]-[24]则获得了该随机序列极值以及极值次序统计量的分布极限定理和几乎处处中心极限定理。利用定理 2.2，我们获得了阿基米德 Copula 相依随机序列完全样本极值次序统计量与非完全样本极值次序统计量之间的联合极限分布。

定理 3.1. 设 $\boldsymbol{X}=\{X_n, n\ge 1\}$ 是一列同分布的阿基米德 Copula 相依的随机变量，具有连续的边际分布函数 $F_X(x)$.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。假设 $S_n=\sum_{k=1}^{n}\varepsilon_k$ 满足，当 $n\to\infty$ 时，

$$\frac{S_n}{n}\xrightarrow{p}\lambda, \tag{16}$$

其中随机变量 $\lambda\in(0,1)$ 几乎处处成立。则存在常数序列 $c_n>0, d_n\in\mathbb{R}$ 使得(6)~(10)式均成立。

注记 3.1. 文献[25]考虑了一列同分布的阿基米德 Copula 相依的随机变量最大值与其非完全样本情形下最大值的联合极限分布，同时本文中条件(16)中的 λ 为随机变量，而文献[25]中 λ 为常数，因此，定理 3.1 推广了文献[25]的主要结论.

证明： 由文献[26]可知，对每个阿基米德 Copula C_n^{ψ}，都存在一维分布函数 F 和一个正值随机变量 Θ 具有拉普拉斯变换 $\psi(t)=E(\exp(-t\Theta))$，使得满足阿基米德 Copula C_n^{ψ} 结构的随机变量 $(X_1,X_2,\cdots,X_n)$ 关于 Θ 条件独立，即其联合分布函数可以表示为(3)式。选择常数序列 $c_n>0, d_n\in\mathbb{R}$ 使得 $\exp\left(-\psi^{-1}\circ F\right)\in D(G)$，其中 ψ^{-1} 是 ψ 的逆函数。进而，定理 2.1 的条件都成立，因此，由定理 2.1 立刻可证定理 3.1 成立。

本文研究了可交换的随机变量在随机缺失情形下的极值次序统计量的极限分布问题，得到了完全样本极值次序统计量与非完全样本极值次序统计量的联合极限分布，并用所得结果研究了阿基米德 Copula 相依结构下完全样本极值次序统计量与非完全样本极值次序统计量的渐近关系，获得了它们的联合极限分布。本文的创新性主要体现在如下两个方面：其一，将文献[25]中条件 $\frac{S_n}{n}\xrightarrow{p}\lambda$ 中常数 λ 推广到随机变量情形；其二，将文献[25]中完全样本极大值与非完全样本极大值的联合极限分布推广到了极值次序统计量情形。

在未来的研究中，我将考虑更一般的 Copula 相依结构情形下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，同时，通过解决上述所提到的问题，将所得结论应用到缺失数据的极值统

计分析中。

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.

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

[5] 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

[6] 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

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

[8] 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

[9] 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

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

[11] Zheng, S. and Tan, Z. (2023) On the Maxima of Non Stationary Random Fields Subject to Missing Observations. *Communications in Statistics-Theory and Methods*, **53**, 6339-6361. https://doi.org/10.1080/03610926.2023.2244098

[12] 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

[13] Tan, Z. and Wang, Y. (2011) Some Asymptotic Results on Extremes of Incomplete Samples. *Extremes*, **15**, 319-332. https://doi.org/10.1007/s10687-011-0140-z

[14] 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.

[15] Tong, B. and Peng, Z.X. (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

[16] Galambos, J. (1987) The Asymptotic Theory of Extreme Order Statistics. 2nd Edition, Krieger.

[17] Wüthrich, M.V. (2005) Limit Distributions of Upper Order Statistics for Families of Multivariate Distributions. *Extremes*, **8**, 339-344. https://doi.org/10.1007/s10687-006-0007-x

[18] Nelsen, R.B. (1999) An Introduction to Copulas. Springer.

[19] 刘慧燕, 谭中权. 随机缺失情形下平稳随机场超过数点过程的极限性质[J]. 数学学报(中文版), 1-20. https://link.cnki.net/urlid/11.2038.O1.20240407.1333.004, 2024-08-15.

[20] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) Modelling Extremal Events for Insurance and Finance. Springer.

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

[22] 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

[23] Dudziński, M. and Furmańczyk, K. (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

[24] Dudziński, M. and Furmańczyk, K. (2018) Application of Copulas in the Proof of the Almost Sure Central Limit Theorem for the *k*th Largest Maxima of Some Random Variables. *Applicationes Mathematicae*, **45**, 31-51. https://doi.org/10.4064/am2340-10-2017

[25] 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.

[26] Marshall, A.W. and Olkin, I. (1988) Families of Multivariate Distributions. *Journal of the American Statistical Association*, **83**, 834-841. https://doi.org/10.1080/01621459.1988.10478671