arXiv:2301.08856v1 [math.ST] 21 Jan 2023

# 论极值依赖与伴随变量之间的关系

BY AMIR KHORRAMI CHOKAMI[1,a], MARIE KRATZ[2,b]

[1]*ESOMAS*系，都灵大学，意大利都灵，a*amir.khorramichokami@unito.it*

[2]*ESSEC*商学院，*CREAR*，法国，b*kratz@essec.edu*

近年来，伴随变量在选择程序中的应用使其研究重新受到关注。例如，与简单随机抽样相比，分组排序抽样利用伴随变量能够提高效率并降低成本。与此同时，寻找能够丰富描述多个时间序列间极值依赖的新方法的研究迅速增多，这既是因为它具有诸多实际应用，也因为缺乏适当的模型来评估这种依赖关系。本文旨在探讨采用伴随变量方法时的极值依赖。我们研究了向量 $(X,Y)$ 的极值依赖如何影响伴随统计量子集上的最大值的渐近行为。此外，通过讨论各种条件和结果，我们考察了对 $X$ 和 $Y$ 的边缘分布进行变换会如何影响极限的退化。

1. 引言。设 $(X_i,Y_i)$ $i=1,\ldots,n$ 是一列具有共同累积分布函数（cdf）func- tion $F(x,y)$ 的独立同分布（i.i.d.）二元随机向量。我们将次序统计量的伴随变量定义如下。将 $X$ 变量按升序排列，从而得到变量 $X$ 的次序统计量 $X_{(1)}\leq X_{(2)}\leq\cdots\leq X_{(n)}$。然后，与第 $i$ 个次序统计量 $X_{(i)}$ 相对应的 $Y$ 变量称为第 $i$ 个次序统计量的伴随变量，记为 $Y_{[i]}$。例如，若 $X$ 变量的最大值为 $X_{(n)}\equiv X_3$，则其伴随变量为 $Y_{[n]}\equiv Y_3$。次序统计量的伴随变量在实际应用中颇受关注：值得一提的是，它们可用于估计受II型删失影响的多元数据集的参数，也可用于开发选择程序（在这类程序中，测量与感兴趣度量相关联的特征更为便利），以及用于分组排序抽样方案。尤其是后者，适用于测量感兴趣变量耗时或成本高昂、但易于对一组与目标变量相关联的观测值进行排序的情形。在这些情况下，与简单随机抽样相比，分组排序抽样是更好的选择，在效率和成本降低方面均是如此，这也使伴随变量研究重新受到关注。我们列举了奠基性研究 Dell和Clutter（1972）、Stokes和Sager（1988），以及较近的 Wang等（2006）、Shao（2010）、Balci、Akkaya和Ulgen（2013）、Wang、Lim和Stokes（2016）、Wang、Ahn和Lim（2017）、Zamanzade和Wang（2018）、Ozturk（2019）等。

我们的研究动机是探讨在选择伴随变量方法时的极值依赖。纵观相关文献，这一主题在20世纪90年代已有研究。值得一提的是开创性论文Nagaraja 和 David（1994年），其中给定一个数 $k$，作者研究伴随变量最大值的分布 $V_1=\max\left(Y_{[n]},\ldots,Y_{[n-k+1]}\right)$。文中分别针对两种情形给出了有限样本和渐近结果： $k$ 固定整数（极端情形）以及 $k=[np]$, $p\in$ （0,1)

*MSC2020* 学科分类：60G70, 62E20, 62G30.关键词和短语：渐近定理、伴随变量、Copula、（尾部）依赖、极值、高斯、逻辑分布、最大值、次序统计量、帕累托、缓变/正则变函数、尾等价、弱收敛。

（分位数情形）。由此自然引出一个问题：与最大值有多接近$Y_{(n)}$? 这是实际应用中的一个基本问题，Joshi 和 Nagaraja （1995）对此进行了研究。在他们的工作中，作者提出将这个问题与两个伴随变量最大值的联合分布研究联系起来。更确切地说，给定一个二元随机向量的独立同分布序列，其母体随机向量为（$X,Y$），定义，对于 $k \geq 1$,

$$V_1 = \max\left(Y_{[n]}, \ldots, Y_{[n-k+1]}\right) \quad \text{and} \quad V_2 = \max\left(Y_{[n-k]}, \ldots, Y_{[1]}\right). \tag{1}$$

Joshi 和 Nagaraja (1995) 引入了随机变量 $W_k := V_1/Y_{(n)}$，并证明其累积分布函数满足

$$\mathbb{P}(W_k \leq w) = \mathbb{P}(V_1 \leq wV_2, V_2 > 0),$$

为此，他们需要计算$(V_1,\ V_2)$的联合分布。
在简要介绍他们的结果之前，我们先引入一些将在全文中使用的记号。

设 $G(z;\mu,\sigma,\xi)$ 表示一个广义极值分布（广义极值分布），其中 $\mu,\xi \in \mathbb{R}, \sigma > 0$。若某分布函数 $F$ 属于 $G$ 的最大吸引域（最大吸引域），记作 $F \in \mathscr{D}(G)$，则存在常数标准化序列 $a_n > 0$ 和 $b_n \in \mathbb{R}$，使得当 $n \to \infty$ 时，对于任意 $z \in \Gamma_G = \{z:\ \xi(z-\mu)/\sigma > 0\}$，有 $F^n(a_n z + b_n) \to G(z)$。最大吸引域的结构由 Fisher and Tippett (1928)、Gnedenko (1943) 等人研究；近年来，Leonetti 和 Khorrami Chokami (2022) 从另一视角进行了研究。分布 $F$ 属于某个广义极值分布 $G$ 的最大吸引域的充分条件称为冯·米塞斯条件：详尽说明请参见（Resnick, 1987，命题 1.15–1.17）；为使本文自成一体，我们在附录 C 中简要回顾这些条件。

若满足 $x > 0$，则称可测函数 $h\colon \mathbb{R}_+ \mapsto \mathbb{R}_+$ 在无穷远处以指数 $a$ 正则变动（记作$h \in \mathrm{RV}_a$）

$$\lim_{t\to\infty} \frac{h(tx)}{h(t)} = x^a.$$

The 指数 $a$ 称为变动指数。当 $a$ 为0时， $h$ 称为缓变（SV） .
随机向量 的上尾依赖系数 $(X,Y)$ 的 Copula 为 $C$，边缘分布 $F_X$ 和 $F_Y$分别为，定义如下（Coles、Heffernan 和 Tawn (1999)

$$\lambda_u := \lim_{\alpha\to 1} \mathbb{P}\big(Y > F_Y^{-1}(\alpha) \mid X > F_X^{-1}(\alpha)\big) = \lim_{\alpha\to 1} \frac{1 - 2\alpha - C(\alpha,\alpha)}{1-\alpha}, \tag{2}$$

$F^{-1}$ 表示累积分布函数 $F$的广义逆函数。
注意， $\lambda_u \in [0,1]$，其中 $\lambda_u = 0$ 表示 $X$ 与 $Y$之间的极值独立 .
我们还引入以下记号，借自Joshi 和 Nagaraja （1995）。给定向量$(X,Y)$，定义

$$F_1(y \mid x) := \mathbb{P}(Y \leq y \mid X > x), \tag{3}$$

$$F_2(y \mid x) := \mathbb{P}(Y \leq y \mid X \leq x), \tag{4}$$

$$F_3(y \mid x) := \mathbb{P}(Y \leq y \mid X = x). \tag{5}$$

条件分布 $F_1$， $F_2$ 和 $F_3$ 在计算次序统计量的伴随变量分布及其变换的分布时发挥着基础作用。

定理 2，见*Joshi 和 Nagaraja（1995）*：假设 $F_X$ 满足冯·米塞斯条件之一，并且对所有 $x$和 $y$，假设 $F_X \in \mathscr{D}(G_X)$，$F_Y \in \mathscr{D}(G_Y)$，以及

$$n\mathbb{P}(X > a_n x + b_n, Y > c_n y + d_n) \underset{n\to\infty}{\to} 0, \quad \forall x, y \in \mathbb{R}, \tag{6}$$

其中，标准化常数$(a_n)_n$、$(b_n)_n$、$(c_n)_n$ 和$(d_n)_n$ 满足

$$F_3(c_n y + d_n \mid a_n x + b_n) \to 1 \qquad as\ n \to \infty. \tag{7}$$

此外，假设存在常数 $\tilde{a}_n > 0$ 和 $b_n \in \mathbb{R}$，使得

$$F_1(\tilde{a}_n y + \tilde{b}_n \mid a_n x + b_n) \underset{n\to\infty}{\to} H_1(x, y). \tag{8}$$

那么，$(V_1, V_2)$ 的累积分布函数满足

$$F_{(V_1,V_2)}(\tilde{a}_n v_1 + \tilde{b}_n, c_n v_2 + d_n) \underset{n\to\infty}{\to} H(v_1)G_Y(v_2), \quad \forall v_1, v_2 \in \mathbb{R}, \tag{9}$$

其中 $H(v_1) = (k!)^{-1} \int H_1^k(x, v_1)(-\log G_X(x))^k g_X(x)\,dx$。

我们首先通过考察条件 (7) 来展开分析。我们先在引理 1 中证明 (6) 与 (7) 之间的关系，证明见附录A。

引理1。设 $x$，$y \in \mathbb{R}$。那么，我们有

$$\begin{aligned} & n\mathbb{P}(X > a_n x + b_n, Y > c_n y + d_n) \underset{n\to\infty}{\to} 0 \\ \Leftrightarrow \quad & F_3(c_n y + d_n \mid a_n z + b_n) \underset{n\to\infty}{\to} 1, \forall z \geq x. \end{aligned} \tag{10}$$

现在可以看出，7）可以直接用上尾依赖系数来理解，即：

命题 1. 设二元随机向量 $(X, Y)$ 的 *Copula* 为 $C$。则，

$$F_3(c_n y + d_n \mid a_n x + b_n) \underset{n\to\infty}{\to} 1, \quad \forall x \geq y \quad \Leftrightarrow \quad \lambda_u = 0$$

其中，归一化常数的选取取决于 $X$ 和 $Y$ 所属的最大吸引域。

注意，命题 1 只是以另一种方式表述了 Resnick，1987，第 296 页中的命题 5.27，其内容为

$$\lambda_u = 0 \quad \Leftrightarrow \quad n\mathbb{P}(X > a_n x + b_n, Y > c_n y + d_n) \underset{n\to\infty}{\to} 0.$$

利用引理1 和命题 1，现在可以将 Joshi 和 Nagaraja(1995) 中的定理 2 改写为关于 $\lambda_u$ 的形式。推论 1 给出了这一改写，有助于阐明 $(V_1, V_2)$ 的渐近独立与 $(X, Y)$的渐近独立之间的联系。

推论 *1.* 假设 $F_X$ 满足冯·米塞斯条件之一，并且对于所有 $x$，$y$，$F_X \in \mathscr{D}(G_X)$ 和 $F_Y \in \mathscr{D}(G_Y)$。此外，假设 $X$ 与 $Y$ 之间的上尾依赖系数为 *0*：$\lambda_u = 0$。若存在常数 $\tilde{a}_n > 0$，$b_n$，使得 $F_1(\tilde{a}_n y + b_n \mid a_n x + b_n) \underset{n\to\infty}{\to} H(x, y)$，则 $(V_1, V_2)$ 的累积分布函数 $F_{(V_1,V_2)}$ 满足 *(9)*。

在实际应用中，假设 $X$ 与 $Y$渐近独立通常过于严格，尤其是在这个高度互联的世界里。例如，过去15年间，我们经历了一场重大的全球金融危机，随后又遭遇了一场大流行病，同时网络攻击也日益增多。系统性风险的存在，以及极值之间的强依赖关系，是这些事件的共同特征（例如Dacorogna 和 Kratz(2015年)）。因此，研究这种（渐近）依赖关系至关重要。这是下一节的主要目标。

**2. 伴随统计量子集上的最大值的联合渐近**。我们旨在研究伴随变量$(V_1,\ V_2)$ 在 $X$ 和 $Y$可能呈现某种渐近依赖关系时的极限分布。为此，我们利用Ledford 和 Tawn(1998年）中给出的联合尾部模型表示，随后陈述主要结果（见第2.2小节），并讨论归一化常数的选择，以避免极限退化。

2.1. *LT* 联合尾部模型。 首先，我们回顾二元情形下与缓变概念相关的一些定义。这些定义将用于定义联合尾部模型。

定义 1. 满足以下条件的函数 $\mathcal{L}$

$$\lim_{n\to\infty}\frac{\mathcal{L}(nx,ny)}{\mathcal{L}(n,n)}=r(x,y),\qquad \text{with } r(ax,ay)=r(x,y), \tag{11}$$

对所有 $a>0$ 和 $(x,\ y)\in\mathbb{R}_+^2$，称为二元缓变 （BSV）函数。

可以证明，存在一个单变量函数 $r_*$，使得

$$r(x,y)=r_*(w),\quad \text{with}\quad w=\frac{x}{x+y}. \tag{12}$$

定义 2. 对于一个 BSV 函数 $\mathcal{L}$，若函数 $r_*$ 在式（12）中定义，且满足 s

$$\frac{r_*(w)}{r_*(1-w)} \text{ is SV at } w=0 \text{ and } w=1,$$

称为准对称。

现在介绍联合尾部模型，即 LT 联合尾部模型，由Ledford 和 Tawn （1998年）提出，用于研究最大值的伴随统计量的渐近分布，$Y_{[n]}$，即边缘分布的极限分布 $V_1$（情形 $k=1$），并计算$\mathbb{P}\big(Y_{(n)}=Y_{[n]}\big)$。

设（$X,\ Y$）为具有单位 Fréchet 边缘分布的二元随机向量（即 $F_X$（$x$）$=\mathrm{e}^{-1/x},\ x>0$）及联合分布函数 $F$（$x,\ y$）。假设存在函数 $c$（$t$）和$\psi$（$x,\ y$），使得联合生存函数

$$\bar{F}(x,y)=1-\mathrm{e}^{-1/x}-\mathrm{e}^{-1/y}+F(x,y)$$

满足

$$\lim_{t\to\infty}\frac{t\bar{F}(tx,ty)}{c(t)}=\psi(x,y) \text{ for all } (x,y)\in\mathbb{R}_+^2. \tag{13}$$

另外，假设存在 $\gamma\in\mathbb{R}$ 使得

$$\begin{cases}\bar{F}(0,t)/\bar{F}(t,0) & \text{is RV}_\gamma \text{ at infinity and}\\ \bar{F}(t,0)/\bar{F}(0,t) & \text{is RV}_{-\gamma} \text{ at infinity.}\end{cases} \tag{14}$$

则有如下结论。

定义 3（LT 联合尾部模型）。设（$X,\ Y$）具有单位 Fréchet 边缘分布。令$\eta\in$（$0,1]$为用于确定尾部依赖系数的衰减速率 $F^{-1}$（$t,\ t$），当 $t\to\infty$时。在假设（13）和（14）下，（$X,\ Y$）的联合生存分布为

$$\mathbb{P}(X>x,Y>y)=\mathcal{L}(x,y)x^{-\alpha}y^{-\beta},\quad \forall x,y, \tag{15}$$

其中 $\alpha$、$\beta>0$、$\alpha+\beta=\eta^{-1}$ 以及函数 $\mathcal{L}$（$x,\ y$） 是一个准对称 BSV 函数（见定义2）。

系数 $\eta$ 描述了 $X$ 与 $Y$之间的极限依赖类型，而函数 $\mathcal{L}$ 则描述了在给定 $\eta$特定值时这种依赖关系的相对强度；有关进一步说明，参见 Ledford 和 Tawn（1998）以及 Heffernan（2000）。正如 de Haan 和 Zhou（2011）所指出的，$\eta=1/2$ 对应于 $X$ 与 $Y$相互独立的情况，而 $\eta>1/2$ 的情况则表示（$X$，$Y$）的极值之间存在正相关。$\eta$ 低于 $1/2$ 的取值表示 $X$ 与 $Y$ 的极值之间存在负相关。完全负相关和完全正相关这两种边界情况分别对应于 $\eta\to 0$ 和 $\eta=1$，其中 $\mathcal{L}(\cdot)=1$。Coles、Heffernan 和 Tawn（1999）引入了一种基本的依赖关系度量：由于 $\bar{}$ $\bar{}$生存 copula $C$由 $C(u,\ v)=1-u-v+C(u,\ v)$给出，因此他们考虑了以下指标来描述上尾依赖：

$$\bar{\chi}:=\lim_{u\to 1}\frac{2\log(1-u)}{\log \bar{C}(u,u)}-1=2\eta-1 \tag{16}$$

$$\lambda_u:=2-\lim_{u\to 1}\frac{\log C(u,u)}{\log u}=\begin{cases}c & \text{if } \bar{\chi}=1,\ \mathcal{L}(t)\to c>0,\ \text{as } t\to\infty,\\ 0 & \text{if } \bar{\chi}=1,\ \mathcal{L}(t)\to 0,\ \text{as } t\to\infty,\\ 0 & \text{if } \bar{\chi}<1.\end{cases} \tag{17}$$

Since $\bar{\chi}=1$ 当且仅当 $\eta=1$时，我们也得到了 $\lambda_u$ 与 $\eta$ 之间的联系。换言之，

- $X$ 与 $Y$在 $\alpha+\beta=1$ 与 $\mathcal{L}(n,\ n)\underset{n\to\infty}{\to}0$时，或在 $\alpha+\beta\geq 1$ 时渐近独立（此时对 $\mathcal{L}(n,\ n)$的极限 $n\to\infty$ 没有任何条件）；
- $X$与 $Y$在 $\alpha+\beta=1$与 $\mathcal{L}(n,\ n)\underset{n\to\infty}{\to}c>0$时渐近依赖。请注意，在这种情况下，$c=\lambda_u$。

这在考察对Joshi 和 Nagaraja(1995)中定理 2 的推广，并放宽条件(7)时，证明很有用。

为简便起见，我们用记号 $C_{\not\perp}$ 表示渐近依赖情形，即 $\alpha+\beta=1$ 与 $\mathcal{L}(n,\ n)\underset{n\to\infty}{\to}\lambda_u>0$。因此，指示函数 $\mathbb{1}_{C_{\not\perp}}=1$ 指出了渐近依赖的存在。

2.2. 主要结果。 下面陈述关于伴随变量最大值的联合渐近行为的主要结果，并允许 $X$ 和 $Y$ 存在渐近依赖：

定理 1. 设 $(X,Y)$ 是一个具有单位 *Fréchet* 边缘分布和上尾依赖系数 $\lambda_u$ 的二元随机向量。假设 $(X,Y)$ 服从定义 *3* 中给出的 *LT* 联合尾部模型，其生存累积分布函数为 (15)，且 *BSV* 函数 $\mathcal{L}$ 满足

$$\mathcal{L}(nx,\tilde{a}_n y+\tilde{b}_n)\underset{n\to\infty}{\to}\tilde{c}(x,y)\quad \textit{where} \tag{18}$$

$$\tilde{c}(x,y)\,x^{1-\alpha}<y^{\beta},\ \tilde{c}(x,y)\underset{y\to 0}{\sim}y^{-\beta},\ \tilde{c}(x,y)=o\left(y^{-\beta}\right)\ \textit{for}\ y\to\infty,$$

$$\textit{with}\quad \tilde{a}_n=\mathcal{O}\left(n^{\frac{1-\alpha}{\beta}}\right)\quad \textit{and}\quad \tilde{b}_n=O\left(\tilde{a}_n\right)\ \textit{or}\ o\left(\tilde{a}_n\right).$$

*T*那么，伴随变量最大值 $(V_1,V_2)$ 在 (1) 中定义，其联合分布满足：

$$F_{(V_1,V_2)}(\tilde{a}_n v_1+\tilde{b}_n,\,nv_2)\xrightarrow[n\to\infty]{}\int_0^{+\infty}H_1^k(v_1\mid x)\,H_2(v_2\mid x)\,\frac{x^{-k-2}}{k!}\,\mathrm{e}^{-1/x}\,dx,\quad \forall v_1,v_2\in\mathbb{R}, \tag{19}$$

其中 $H_1(y\mid x)$：$=1-\tilde{c}(x,y)\,x^{1-\alpha}\,y^{-\beta}$以及 $H_2(y\mid x)$分别由下文 (20) 和 (21) 中给出的极限的乘积定义，其中 $r(x,y)=\lambda_u^{-1}\tilde{c}(x,y)$。

为证明定理1，我们需要评估式（3）、（4）和（5）中分别定义的三个条件分布 、$F_2$ 和 $F_3$ 的渐近行为，并对每个分布进行适当变换，以获得非退化的极限分布。以下引理将给出这一结果。

引理 2. 假设 $LT$ 联合尾部模型 (15) 成立，并将 $r$ 称为与 $\mathcal{L}$相对应的极限函数。那么，当 $n \to \infty$ 时，

$$F_2^n(ny \mid nx) \to \exp\left\{-\frac{1}{y}\left(1-\lambda_u r(x,y)\left(\frac{y}{x}\right)^{\alpha}\mathbb{1}_{C_{\not\perp}}\right)\right\}, \tag{20}$$

$$F_3(ny \mid nx) \to 1+\lambda_u x^{-\alpha+2}y^{\alpha-1}\left(\frac{\partial}{\partial x}r(x,y)-\frac{\alpha}{x}r(x,y)\right)\mathbb{1}_{C_{\not\perp}}. \tag{21}$$

此外，在条件(18)下，有

$$F_1(\tilde{a}_n y+\tilde{b}_n \mid nx) \underset{n\to\infty}{\to} \tilde{c}(x,y). \tag{22}$$

注意，在（ $\mathbb{1}_{C_{\not\perp}}=0$20）和（21）中取 极限 $\lim_{n\to\infty} F_2^n(ny \mid nx) = \mathrm{e}^{-1/y}$ 以及极限 $\lim_{n\to\infty} F_3(ny \mid nx) = 1$，即可得到具有渐近独立性的 Copula 的特征形式，Joshi 和 Nagaraja (1995)对此进行了讨论。定理 1 和引理 2 的证明见附录 A。

值得注意的是，条件（18）是确保 $H_1(v_1 \mid x)$ 对于$v_1 \in (0,+\infty)$是一个正规的累积分布函数所必需的。此外，定理 1 明确说明了˜归一化常数 $\tilde{a}_n$ 和 $b_n$的选取方式。事实上，这是一个微妙的问题，因此我们将对此进行全面讨论：下一小节将解释条件（18），而第 4 节将讨论高斯二元分布这一典型例子，在该例中条件（18）不成立。

˜2.3. 关于 $\tilde{a}_n$ 和 $b_n$ 的选择的讨论。 下文将仔细推导哪些˜是常数 $\tilde{a}_n$ 和 $b_n$仅有的可能渐近行为。我们着重强调本小节的重要性，因为在渐近理论中，找到能产生非退化极限的合适常数通常是一个难题。

˜首先注意到， $F_1(\tilde{a}_n y+b_n \mid nx)$ 的极限很大程度上取决于˜$\mathcal{L}(nx,\tilde{a}_n y+b_n)$的极限。事实上，我们可以写成

$$\begin{aligned} F_1(\tilde{a}_n y+\tilde{b}_n \mid nx) &= 1-\frac{\mathcal{L}(nx,\tilde{a}_n y+\tilde{b}_n)(nx)^{-\alpha}(\tilde{a}_n y+\tilde{b}_n)^{-\beta}}{1-\mathrm{e}^{-1/(nx)}} \\ &\underset{n\to\infty}{\sim} 1-\frac{\mathcal{L}(nx,\tilde{a}_n y+\tilde{b}_n)}{n^{\alpha-1}x^{\alpha-1}(\tilde{a}_n y+\tilde{b}_n)^{\beta}} \end{aligned} \tag{23}$$

利用近似式 $1-\mathrm{e}^{-1/(nx)} \underset{n\to\infty}{\sim} (nx)^{-1}$。

˜下文将讨论应如何选择序列 $\tilde{a}_n$ $b_n$ $F_1$ $\tilde{a}_n y+$，使得$(b_n \mid nx) \underset{n\to\infty}{\to} 1$，即 $V_1$ 不退化。

G给定一对 $(x,y)$，我们研究 (23) 的极限，考察其对 $\mathcal{L}$ 的依赖，具体而言 y:

˜(i) 假设 $\mathcal{L}(nx,\ \tilde{a}_n y+b_n) \to 0$ 与 $n\to\infty$ 相同。若要使 $F_1(\tilde{a}_n y+b_n \mid nx) \not\to 1$成立，我们需要˜$\mathcal{L}(nx,\ \tilde{a}_n y+b_n)\, n^{\alpha-1}x^{\alpha-1}(\tilde{a}_n y+\tilde{b}_n)_{\beta} \not\to 0$。但是，由于我们处于 $\mathcal{L}(nx,\ \tilde{a}_n y+\tilde{b}_n) \underset{n\to\infty}{\to} 0$ 的情形，这意味着以下两者之一：

- $n^{\alpha-1}x^{\alpha-1}(\tilde{a}_n y+\tilde{b}_n)^{\beta} \underset{n\to\infty}{\to} 0$ 以 $\mathcal{L}$ 的相同速率变化；这不可能，因为模型假设（$\mathcal{L}\sim$ 缓变），或者
  - $n^{\alpha-1}x^{\alpha-1}(\tilde{a}_n y+\tilde{b}_n)^{\beta} \underset{n\to\infty}{\to} 0$ 比 $\mathcal{L}$更快，这不可能，因为我们会有 $\frac{\mathcal{L}(nx,\ \tilde{a}_n y+b_n)}{n^{\alpha-1}x^{\alpha-1}(\tilde{a}_n y+\tilde{b}_n)^{\beta}} \underset{n\to\infty}{\to} \infty$，这否定了 $F_1$ 是累积分布函数这一点。因此，情形(i) 不可能成立。

(ii) 假设 $\mathcal{L}(nx,\ \tilde{a}_n y+\tilde{b}_n)\to\infty$ 当 $n\to\infty$时。因此，由于 $n\to\infty$ $n^{\alpha-1}x^{\alpha-1}\ \tilde{a}_n y+\tilde{b}_n\ ^{\beta}$，项 ()趋于无穷的速度快于 $\mathcal{L}(nx,\ \tilde{a}_n y+b_n)$，由此可得 $F_1(\tilde{a}_n y+b_n\mid nx) \underset{n\to\infty}{\to} 1$。因此，这种情形 (ii) 也必须排除。

(iii) 前述推理意味着，当 $n\to\infty$ 时，我们只能有：

$$\mathcal{L}(nx,\tilde{a}_n y+\tilde{b}_n) \underset{n\to\infty}{\to} \tilde{c}(x,y)\in(0,\infty).$$

$\tilde{a}_n$ 和 $b_n$ 需要满足哪些条件？这里，我们必须区分 $\tilde{c}(x,\ y)$是否依赖于 $y$。

(A) 假设 $\tilde{c}(x,\ y)$ 依赖于 $y$。那么，$\frac{\mathcal{L}(nx,\ \tilde{a}_n y+b_n)}{n^{\alpha-1}x^{\alpha-1}(\tilde{a}_n y+\tilde{b}_n)} \underset{n\to\infty}{\sim} \tilde{c}(x,\ y)\, n^{\alpha-1}x^{\alpha-1}(\tilde{a}_n y+\tilde{b}_n)^{\beta}$，因此我们必须考察.....的相对行为，以及 $\tilde{a}_n$ $\tilde{b}_n$.

(a) 如果 $\tilde{a}_n=o(\tilde{b}_n)$，那么 $\tilde{a}_n y+\tilde{b}_n \underset{n\to\infty}{\sim} \tilde{b}_n$，这意味着 $\frac{\tilde{c}(x,y)}{n^{\alpha-1}x^{\alpha-1}\tilde{b}_n^{\beta}}$，因此，为了保持与 $y$ 的依赖关系，我们需要 $n^{\alpha-1}\tilde{b}_n^{\beta} \not\to 0,\infty$。这个条件显然取决于 $\alpha$ 的取值：令 $\gamma=(1-\alpha)/\beta$，则 • 若 $0<\alpha<1$，$n^{\alpha-1} \underset{n\to\infty}{\to} 0$，唯一的可能性是 $\tilde{b}_n=\mathcal{O}(n^{\gamma})$，其中 $\gamma\in(0,1)$，因为 $\alpha+\beta\geq 1$。

• 若 $\alpha=1$，则 $b_n$ 将与 $n$ 相互独立；在这种情况下，便无法提出 $(\tilde{a}_n,\ b_n)$。

• 若 $\alpha>1$，则 $b_n=\mathcal{O}(n^{\gamma})$ 且 $\gamma<0$。然而，由于我们关注的是极值中的依赖关系，因此需要考虑

$$n\mathbb{P}(X>nx,Y>n^{\gamma}y)\leq n^{1-\gamma}\cdot n^{\gamma}\mathbb{P}(X>n^{\gamma}x,Y>n^{\gamma}y) \underset{n\to\infty}{\not\to} 0,$$

这意味着 $\gamma\in(0$<style id='4', $1)$。因此，无法考虑 $\alpha>1$ 这种情况。

(b) 若 $\tilde{b}_n=o(\tilde{a}_n)$ 或者 $\tilde{b}_n=\mathcal{O}(\tilde{a}_n)$，则 $\tilde{a}_n y+\tilde{b}_n \underset{n\to\infty}{\sim} \tilde{a}_n y$ 且

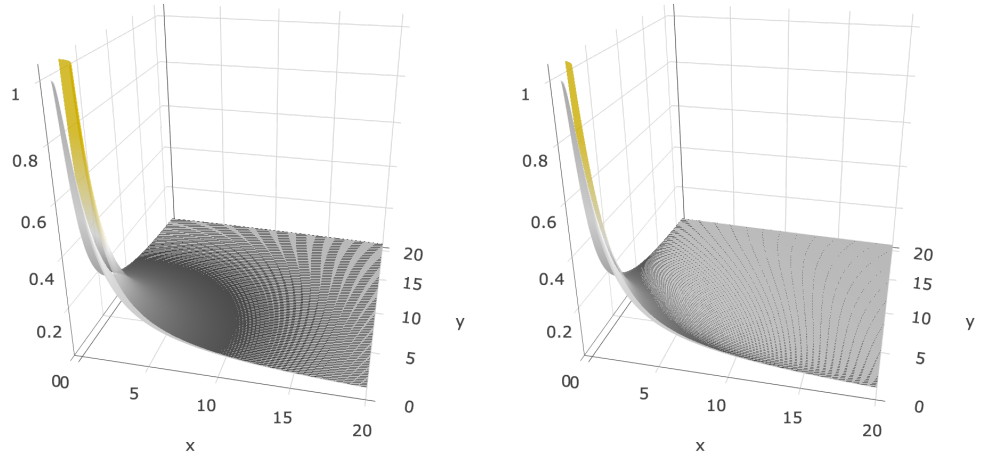
$$\frac{\tilde{c}(x,y)}{n^{\alpha-1}x^{\alpha-1}(\tilde{a}_n y+\tilde{b}_n)^{\beta}} \underset{n\to\infty}{\sim} \frac{1}{n^{\alpha-1}\tilde{a}_n^{\beta}}\,\tilde{c}(x,y)\,x^{1-\alpha}.$$

因此，我们需要考虑 $\tilde{a}_n=\mathcal{O}(n^{\gamma})$ 和 $\alpha\in(0,1)$。

(B) 假设 $\tilde{c}(x,y)$ 不依赖于 $y$。在这种情况下，

$$\frac{\mathcal{L}(nx,\tilde{a}_n y+\tilde{b}_n)}{n^{\alpha-1}x^{\alpha-1}(\tilde{a}_n y+\tilde{b}_n)^{\beta}} \underset{n\to\infty}{\sim} \frac{n^{1-\alpha}}{\max\left(\tilde{a}_n^{\beta}y^{\beta},\tilde{b}_n^{\beta}\right)}\tilde{c}x^{1-\alpha},$$

并且，由于我们不希望失去对 $y$ 的依赖关系，因此需要假设$\max\left(\tilde{a}_n^{\beta}y^{\beta},\tilde{b}_n^{\beta}\right)=\tilde{a}_n^{\beta}y^{\beta}$。此外，我们还需要要求 $n^{1-\alpha}\tilde{a}_n^{-\beta}\not\to 0$（否则$F_1(\tilde{a}_n y+b_n\mid nx) \underset{n\to\infty}{\to} 1$），这意味着 $\tilde{a}_n=\mathcal{O}(n^{\gamma})$ 且 $\alpha\in(0,1)$。




(a) 生存 *Clayton Copula* ($\theta = 2$)　　(b) *Logistic Copula* ($\gamma = 1/2$)

¯图 1：联合尾部 $F(x,\ y)$（灰色），具有单位 Fréchet 边缘分布，如式（24）所示；其近似值（黑色和金色）由公式（15）给出。两个示例中的参数 $\alpha$ 和 $\beta$ 均设为 0。5 。

¯因此，为了得到 ˜ $F_1\ \tilde{a}_n y + b_n \mid nx$ $\alpha \in ()$ 的非退化极限，我们需要假设 $(0,1)$、$\tilde{a}_n = \mathcal{O}(n^\gamma)$ 和 $b_n = o(\tilde{a}_n)$，或者 $b_n = \mathcal{O}(\tilde{a}_n)$。现在，主要结果已给出，并对其中涉及的常数作了说明，下面我们通过风险文献中的标准示例来阐释理论结果。

**3. 结果示例。** 本节考虑两个示例：采用生存 Clayton Copula 的帕累托 - Lo-max，以及双变量对称 Logistic 极值 Copula。我们应用定理1 给出的渐近近似，并将其与通过模拟得到的经验结果进行比较。选择这类示例是因为极值依赖由参数决定，因此易于控制。还请注意，对于以下示例，我们也可以通过直接计算获得相同结果（但需要耗费大量时间并进行繁琐的计算），因此可以利用所得的数学表达式进行比较。

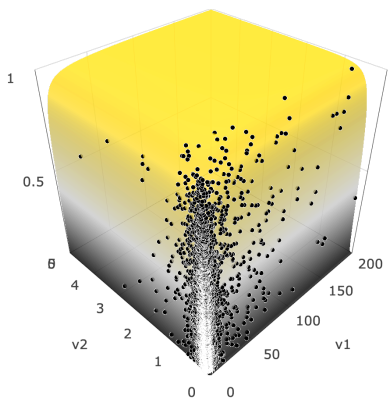
我们模拟 5000 次 100000 个二元样本的重复试验。用于计算 $k$ 的子样本维度 $V_1$ 设为 10 （附录 B 包含了其他 $k$ 取值下的模拟，以深入了解其对渐近结果的影响）。

3.1. 帕累托 - 洛马克斯边缘分布与生存 *Clayton Copula*。 设 $(X_P, Y_P)$ 为一个具有生存 Clayton$(\theta)$ Copula 和帕累托 - 洛马克斯$(\nu,1)$ 边缘分布的随机向量（下标 $P$表示帕累托边缘分布）：

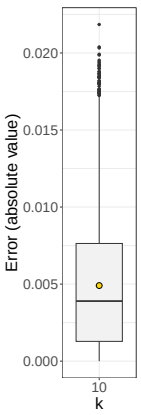
$$\bar{F}_{X_P}(x) = (1+x)^{-\nu}, \forall x > 0, \quad \bar{F}_{Y_P}(y) = (1+y)^{-\nu}, \forall y > 0, \quad \nu > 0,$$
$$\bar{F}_P(x,y) = \mathbb{P}(X_P > x, Y_P > y) = \left((1+x)^{\nu\theta} + (1+y)^{\nu\theta} - 1\right)^{-1/\theta}, \theta > 0.$$

首先，我们将帕累托边缘分布转换为单位 Fréchet 边缘分布，以符合定理 1中给出的模型假设。我们可以写成

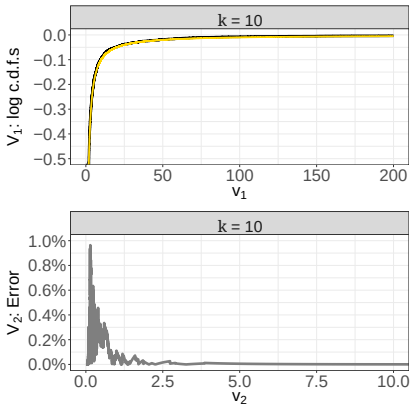
$$F(x,y) = F_P\left(F_{X_P}^{-1}\left(\mathrm{e}^{-1/x}\right), F_{Y_P}^{-1}\left(\mathrm{e}^{-1/y}\right)\right),$$




(a) 二元分布 $(V_1, V_2)$。




(b)误差评估——价。




(c)边缘分布诊断。

图 2：生存 Clayton$(\theta = 2)$，其中 $k = 10$。面板 (a) 显示 $V_1$ 的经验单变量对数累积分布函数（黑色），以及通过对渐近联合累积分布函数（19）进行数值积分而得到的渐近理论对数值（金色）。面板 (b) 显示 $V_2$ 的经验累积分布函数与理论累积分布函数之间误差的绝对值。面板 (c) 上方的图显示 $V_1$ 的经验单变量累积分布函数（黑色），以及通过对渐近联合累积分布函数（19）进行数值积分而得到的对数尺度下的渐近理论值（金色）。面板 (c) 下方的图表示 $L_1$ 边缘分布的经验累积分布函数与理论累积分布函数之间的误差。

其中，对于帕累托分布，$F_{X_P}^{-1}(q) = F_{Y_P}^{-1}(q) = (1-q)^{-1/\nu} - 1$。我们得到

$$\bar{F}(x,y) = \left( \left(1 - \mathrm{e}^{-1/x}\right)^{-\theta} + \left(1 - \mathrm{e}^{-1/y}\right)^{-\theta} - 1 \right)^{-1/\theta}. \tag{24}$$

现在注意，联合生存函数（24）对于 $x > x_0 > 0$ 和 $y > y_0 > 0$，可以改写为

$$\begin{aligned} \bar{F}(x,y) &\sim \left(x^\theta + y^\theta - 1\right)^{-1/\theta} = \left(x^\theta + y^\theta - 1\right)^{-1/\theta} (xy)^{1/2} (xy)^{-1/2} \\ &= \mathcal{L}(x,y) x^{-1/2} y^{-1/2}, \end{aligned}$$

其中 $\mathcal{L}(x,y) = \left(x^\theta + y^\theta - 1\right)^{-1/\theta} (xy)^{1/2}$ 并且显然 $\alpha = \beta = 1/2$。注意，

$$\mathcal{L}(n,n) = \left(2n^\theta - 1\right)^{-1/\theta} n \to 2^{-1/\theta} \quad \text{as } n \to \infty.$$

这与上尾依赖系数的值一致，因为

$$\lambda_U = \lim_{q \to 1} \mathbb{P}\left(Y > F_Y^{-1}(q) \mid F_X^{-1}(q)\right) = \lim_{q \to 1} (1-q)^{-1} \left(2(1-q)^{-\theta} - 1\right)^{-1/\theta} = 2^{-1/\theta}.$$

根据定理1，因为 $1 - \alpha = \beta$，我们有 $\tilde{a}_n = a_n = n$。因此，

$$\mathcal{L}(nx, ny) = \left(n^\theta x^\theta + n^\theta y^\theta - 1\right)^{-1/\theta} n (xy)^{1/2} \underset{n \to \infty}{\to} \left(x^\theta + y^\theta\right)^{-1/\theta} (xy)^{1/2} = \tilde{c}(x,y).$$

注意，对于每个 $a > 0$，

$$\tilde{c}(ax, ay) = \left((ax)^\theta + (ay)^\theta\right)^{-1/\theta} (a^2 xy)^{1/2} = \tilde{c}(x,y),$$

这意味着 $\tilde{c}(x,y)$ 并且 $r(x,y) := 2^{1/\theta}\tilde{c}(x,y)$ 满足（11）。此外，容易证明 $\tilde{c}(x,y)$ 满足定理 18中的条件（1）。由于该结论成立

$$\frac{\partial}{\partial x}r(x,y) = 2^{1/\theta}\frac{\partial}{\partial x}\tilde{c}(x,y) = 2^{1/\theta}x^{-1}\tilde{c}(x,y)\left(\frac{1}{2} - \left(1+\left(\frac{y}{x}\right)^{\theta}\right)^{-1}\right),$$

应用引理2 和定理1可得

$$H_1(y \mid x) = 1 - \left(1+\left(\frac{y}{x}\right)^{\theta}\right)^{-1/\theta}, \tag{25a}$$

$$H_2(y \mid x) = \left(1 - \left(1+\left(\frac{y}{x}\right)^{\theta}\right)^{-1/\theta-1}\right)\mathrm{e}^{-\frac{1}{y}+\left(1+\left(\frac{y}{x}\right)^{\theta}\right)^{-1/\theta}}. \tag{25b}$$

图2a 中的图形展示了应用定理 1 计算得到的二元曲面（积分采用数值方法计算），以及在模拟的 $V_1$ 和 $V_2$中计算得到的二元经验累积分布函数的取值。这些点位于该曲面上，表明定理1的渐近近似质量非常好。图2b和2c也能佐证这一点：它们分别展示了图2a中黑色点云与曲面之间误差绝对值的箱线图、 $V_1$边缘分布的渐近近似，以及 $V_2$中的误差。关于误差（图2c下图），我们观察到，对于较小的 $v_2$，误差很小（最大误差接近1%），并且很快收敛到0。具体而言，对于我们考虑的所有 $v_2$ 取值，误差均为0% 。

3.2. 单位 *Fréchet* 边缘分布与 *Logistic* 极值 *Copula*。 设$(X,Y)$ 为一个具有单位 Fréchet 边缘分布和二元对称 Logistic 极值依赖的随机向量，即

$$F(x,y) = \mathrm{e}^{-V(x,y)},$$

其中

$$V(x,y) = \left(x^{-1/\gamma} + y^{-1/\gamma}\right)^{\gamma},\ x,y>0,\ \gamma\in(0,1).$$

考虑 Ledford 和 Tawn（Ledford 和 Tawn(1997）提出的由一般极值依赖函数 $V$刻画的二元联合尾部近似（15），本例中可写为

$$\mathcal{L}(x,y) = (xy)^{-1/2}(x+y) - (xy)^{1/2}V(x,y). \tag{26}$$

注意

$$\mathcal{L}(n,n) = 2 - V(1,1),$$

注意，由于 $V$是 $-1$ 阶齐次的（即 $V(nx,ny) = n^{-1}V(x,y)$）。此外， $1-\alpha=\beta$，因此 $\tilde{a}_n = a_n = n$。于是可写为

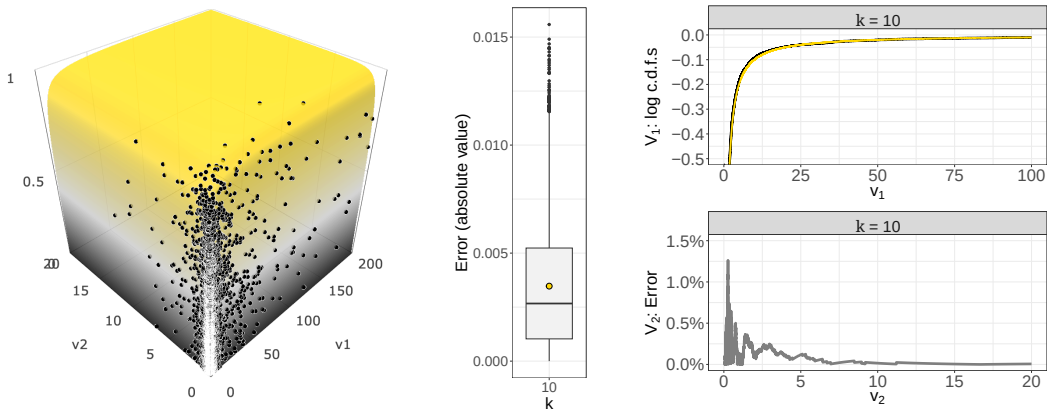
$$\begin{aligned}\mathcal{L}(nx,ny) &= n^{-1}(xy)^{-1/2}n(x+y) - n(xy)^{1/2}n^{-1}V(x,y)\\ &= (xy)^{-1/2}(x+y) - (xy)^{1/2}V(x,y) = \tilde{c}(x,y).\end{aligned}$$

注意，对于每个 $a>0$,

$$\tilde{c}(ax,ay) = \tilde{c}(x,y),$$

从而候选函数 $\mathcal{L}$ 在（26）中是一个适当的 BSV 函数。通过简单的代数运算即可证明$\tilde{c}$（$x$， $y$）满足定理1中的条件（18）。由于它成立

$$\frac{\partial}{\partial x}\tilde{c}(x,y) = \frac{(xy)^{-1/2}}{2}\left(1 - \frac{y}{x} - yV(x,y)\left(1 - 2(xV(x,y))^{-1/\gamma}\right)\right),$$




(a) 二元分布 $(V_1, V_2)$。

(b)误差评估—计。

(c)边缘分布诊断。

图 3：逻辑分布$(\gamma = 1/2)$，其中 $k = 10$。面板 (a) 显示$V_1$ 的经验单变量对数累积分布函数（黑色），以及通过对渐近联合累积分布函数（19）进行数值积分得到的渐近理论对数值（金色），从而得到边缘分布。面板 (b) 显示 $V_2$经验累积分布函数与理论累积分布函数之间的误差绝对值。面板 (c) 上图显示 $V_1$ 的经验单变量累积分布函数（黑色），以及通过对渐近联合累积分布函数（19）进行数值积分得到的渐近理论值（金色），采用对数尺度。面板 (c) 下图表示边缘经验累积分布函数与理论累积分布函数之间的 $L_1$ 误差。

应用引理2 和定理1，可得

$$H_1(y \mid x) = x\left(V(x,y) - y^{-1}\right), \tag{27a}$$

$$H_2(y \mid x) = \left(1 + \left(\frac{y}{x}\right)^{-\frac{1}{\gamma}}\right)^{\gamma-1} \mathrm{e}^{-\frac{1}{x}\left(\left(1+\left(\frac{y}{x}\right)^{-\frac{1}{\gamma}}\right)^{\gamma}-1\right)}. \tag{27b}$$

得到函数 $H_1$ 和 $H_2$后，我们现在可以通过数值积分计算联合累积分布函数（19）。

**4. 高斯情形的讨论**。迄今为止，我们的重点是求出当$(X,Y)$ 呈现渐近依赖或极值依赖时$(V_1, V_2)$ 的联合渐近累积分布函数；而我们想探究定理1 中的条件（18）是否与极值依赖假设密切相关。换言之，如果 BSV 函数 $\mathcal{L}$ 趋于0 或趋于无穷，会发生什么？因此，我们重新考察$(X,Y)$ 服从高斯分布的情形，因为众所周知，此时$(X,Y)$是渐近独立的。由于该分布在实际应用中颇受关注，且已知易于处理，文献中早已研究过这一情形。尽管在此情形下直接计算很简单，但在应用涉及 SV 函数近似的模型时却会遇到问题。事实上，即便给出了这类近似，也不会用它们来计算例如我们上文讨论过的条件分布（至少据我们所知，文献中并未如此处理）。本节旨在探究这一表面上的悖论，并提供解决这一问题的工具，从而弥合直接计算与通过渐近模型进行计算之间的差距，同时关注 BSV $\mathcal{L}$的渐近行为。

4.1. 现有文献。 我们先回顾一下关于高斯情形的现有文献。

(a) Joshi 和 Nagaraja(1995)：通过直接计算，作者得以给出若干常˜数 $\tilde{a}_n$， $b_n$ 使得 $(V_1,\ V_2)$ 的重标度分布收敛到标准正态累积分布函数的 $k$ 次幂与 Gumbel 累积分布函数的乘积。将这些常数应用于我们的情形时，在把高斯边缘分布变换为单位 Fréchet 后，定理 18 中的条件 1 并不满足。事实上，我们得到 $F_1$ 在 $n\to\infty$ 时退化。

因此，我们的问题是：能否找到其他边缘分布变换，绕开 $\mathcal{L}\to 0$ 这种情形，并可能找到处理该情形的新方法或解决方案？这个问题之所以重要，在于渐近依赖与 $\mathcal{L}$ 的行为之间的关系通常尚不明确：根据公式（16），我们知道 $\mathcal{L}(n,\ n)$ 的行为以及 $\eta$ 取值与上尾依赖系数 $\lambda_u$ 相关，但尚未探讨 $\mathcal{L}\not\to 0$ 和 $\eta<1$ 的情形。此外，我们还需要考虑 $\mathcal{L}(n,\ \tilde{a}_n)$，而不是 $\mathcal{L}(n,\ n)$。

(b) Ledford 和 Tawn (1998年)：这里仅关注最大值伴随变量 $X_{(n)}$的渐近分布，也就是当 $V_1$ 时（按我们的记号）， $k=1$的渐近分布，即研究 $F_1$。作者指出，在某些情况下，单位 Fréchet 边缘分布并不适用，并建议对变量 $Y$进行标准 Gumbel 变换，从而得到一个新的 BSV 函数 $\mathcal{L}_1$（实际上是 $\mathcal{L}$的一阶近似），以避免 $H_1$ 的极限退化（采用我们的记号）。然而，要将这一方法与两个边缘分布均为单位 Fréchet 的情形进行比较，他们需要假设 $y\to\infty$。

设 $(X_N,\ Y_N)$ 是一个具有标准正态边缘分布且相关系数为 $0<\rho<1$的二元正态随机向量。Joshi 和 Nagaraja (1995) 中的定理 3 指出，在极端情形下（$k$ 固定， $n\to\infty$），有

$$\mathbb{P}\left(\frac{V_1-\tilde{a}_n}{\tilde{b}_n}\le v_1,\frac{V_2-a_n}{b_n}\le v_2\right)\underset{n\to\infty}{\to}\Phi^k(v_1)\mathrm{e}^{-\mathrm{e}^{-v_2}}, \tag{28}$$

其中

$$a_n=(2\log n)^{-1/2}\qquad b_n=\sqrt{2\log n}-\frac{1}{2}\frac{\log(4\pi\log n)}{\sqrt{2\log n}} \tag{29a}$$

$$\tilde{a}_n=\sqrt{1-\rho^2}\qquad \tilde{b}_n=\rho b_n. \tag{29b}$$

公式（28）表示 $V_1$ 与 $V_2$渐近独立，这与如下事实密切相关：

$$F_1(\tilde{a}_n y+\tilde{b}_n\mid a_n x+b_n)=\mathbb{P}\Big(Y_N\le\tilde{a}_n y+\tilde{b}_n\mid X_N>a_n x+b_n\Big)\underset{n\to\infty}{\to}\Phi(y). \tag{30}$$

请注意，定理 1中的模型不适用于此处。事实上，对于任意选择的归一化常数，条件（18）均不满足，因为 $\mathcal{L}(n,\ \tilde{a}_n)\to 0$。因此，这种情形还需要进一步研究。

理解这一问题的一种途径是考察Wadsworth 和 Tawn (2013年。作者研究了二元随机向量的各分量以不同速率增长所产生的影响，并据此提出了一类新的正则变差条件，同时为多元尾部衰减与所考虑的边缘增长率之间的联系提供了新的刻画。

更准确地说，作者将原始向量的边缘分布变换为标准指数分布，记作$(X_E,Y_E)$，并假设对所有$(\zeta_1,\zeta_2)\in\mathbb{R}_+^2\setminus\{\mathbf{0}\}$，

$$\mathbb{P}(X_E>\zeta_1\log n,Y_E>\zeta_2\log n)\underset{n\to\infty}{\sim}L(n;\zeta_1,\zeta_2)n^{-\kappa(\zeta_1,\zeta_2)} \tag{31}$$

其中， $L$ 是 $n$ 中关于 $n\to\infty$的单变量缓变函数，而函数 $\kappa(\zeta_1,\ \zeta_2)$ 将不同的边缘增长率映射为联合尾部衰减率，从而提供相关信息

说明变量在次渐近水平下的依赖程度。假设（31）随后用于描述极端超越的条件概率模型，即

$$\mathbb{P}(X_E > \zeta_1 \log n \mid Y_E > \log n) \underset{n\to\infty}{\sim} L(n;\zeta_1,1)\mathrm{e}^{(1-\kappa(\zeta_1,\zeta_2))\log n}. \tag{32}$$

作者证明，研究32等价于寻找函数 $a^*$， $b^*$ 和$h$，使得

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_E - b^*(\log n)}{a^*(\log n)} > x \,\middle|\, Y_E > \log n\right) = \mathrm{e}^{-h(x)}. \tag{33}$$

使用指数边际分布有助于确定二元高斯分布示例中的归一化常数，从而部分填补了Ledford 和 Tawn(1998年)中留下的空白。在那里，正如前文所述，作者给出了在Fréchet-Gumbel 边际分布情形下处理该问题的方法，尽管该解法仅适用于$Y$的一些较大取值。然而，这番讨论尚未回答我们的核心问题：使用单位 Fréchet 边缘分布时，为什么我们无法找到能使 $\mathcal{L}(nx,\ \tilde{a}_n y + b_n)$产生非退化˜极限的归一化常数? Heffernan andResnick(2007)颇具洞见的论文让我们对这个问题有了更完整的理解。作者在文中探讨了如下假设的含义：存在一个缩放函数 $c_1^*(\cdot)$、一个中心化函数 $c_2^*(\cdot)$，以及一个非零拉东测度 $\mu$，定义在 $[-\infty,\infty]\times(-\infty,\ \infty]$ 的博雷尔子集上，使得对于每个固定的 $y \in E_\xi$: $= \{y \in \mathbb{R}:\ 1+\xi y > 0\}$（其中 $\xi$是相对于 $F_Y$的极值分布的形状参数），都有

(a) $\mu([-\infty,x]\times(y,\ +\infty))$是 $x$中的有限且非退化的分布函数；(b) 给定 $Y$为极端值时，标准化 $X$ 的累积分布函数的极限存在，并且

$$\mathbb{P}\left(\frac{X - c_2^* \circ b^{\leftarrow}(t)}{c_1^* \circ b^{\leftarrow}(t)} \le x \,\middle|\, Y > t\right) \underset{t\to y^+}{\to} \mu([-\infty,x]\times(0,+\infty)), \tag{34}$$

其中， $y^+$ 是 $F_Y$的右端点。值得注意的是，我们可以自由改变 $Y$的边缘分布，而不会影响（34）中的收敛性，但对 $X$变量则并非如此。文中证明，对于二元标准高斯情形，如果将 $X$ 和$Y$ 的边缘分布都变换为帕累托型，则无法在（34）中得到 $\mu$ 的非退化表达式；不过，如果将边缘分布变换为指数分布，仍然可以得到渐近形式。关于这一方面的详细讨论，参见（Heffernan 和 Resnick，2007，第7节）。

4.2. 重新审视高斯情形。 现在，我们对文献中处理高斯情形时所涉及的问题有了更好、更全面的理解，因此我们通过应用模型（32）以及公式（29a）和（29b），给出一种求取合适常数的方法。我们知道，如果$(X_N,\ Y_N)$ 是一个相关系数为 $\rho \in (0,\ 1)$的二元正态随机向量，

$$\mathbb{P}\Big(Y_N > \tilde{a}_n y + \tilde{b}_n \mid X_N > a_n x + b_n\Big) \underset{n\to\infty}{\to} 1 - \Phi(y), \tag{35}$$

其中常数由（29a）和（29b）给出。以下结果成立：

定理2. 设 $(X_E, Y_E)$ 是一个边缘分布为单位指数分布、相关系数为 $\rho \in (0,1)$ 的二元随机向量，且满足 (32)，其中 $\zeta_1 := \zeta_1(n,y)$。那么，我们有

$$\mathbb{P}\Big(Y_E > \tilde{a}_{n,E} y + \tilde{b}_{n,E} \mid X_E > \log n\Big) \to \frac{1}{y\sqrt{2\pi}} e^{-(\rho y)^2/2}, \quad as\ n \to \infty, \tag{36}$$

其中

$$\tilde{a}_{n,E} = \rho\sqrt{1-\rho^2}\left(\sqrt{2\log n} - \frac{1}{2}\frac{\log(4\pi\log n)}{\sqrt{2\log n}}\right) \tag{37a}$$

$$\textit{and} \quad \tilde{b}_{n,E} = \rho^2\log n - \frac{\rho^2}{2}\log(4\pi\log n) + \frac{\rho^2}{16}\frac{(\log(4\pi\log n))^2}{\log n}. \tag{37b}$$

推论2. 注意，若我们假设 $y$ 足够大（如 *Ledford 和 Tawn* （*1998年*）中所做），则可以使用米尔斯比率得到

$$\mathbb{P}\Big(Y_E > \tilde{a}_{n,E}y + \tilde{b}_{n,E} \mid X_E > \log n\Big) \underset{n\to\infty}{\sim} \rho\,(1-\Phi(\rho y))\,. \tag{38}$$

证明见附录 A 。
（38）中给出的极限分布与Joshi 和 Nagaraja（1995）以及Wadsworth 和 Tawn（2013年）中的结果相似。最终公式之间的差异源于所选泰勒近似阶数不同。

**5. 讨论**。本文深入阐明了二元随机向量（$X$，$Y$）的极值依赖对伴随统计量子集上经过适当重新缩放的最大值之联合渐近行为的影响。研究渐近相关随机变量的框架在实际应用中颇具价值，也丰富了相关文献。

我们的主要结果针对这一问题，给出了两个重新缩放的伴随变量最大值的渐近联合累积分布函数表达式。具体而言，我们利用 Ledford 和 Tawn （1997）提出的联合尾部模型，给出描述 $Y$ 在给定 $X$行为时的所有条件分布的渐近形式；这些分布共同决定了两个伴随变量最大值的渐近联合累积分布函数。我们还详细说明了如何选择适当的归一化常数，以避免此类分布退化（从而防止伴随变量最大值的联合累积分布函数也发生退化）。我们通过风险分析文献中的两个相关实例阐释理论结果：计算条件分布的解析表达式，并通过数值积分得到联合累积分布函数。通过数值模拟，我们展示了渐近近似在所有考察情形下的良好表现。

此外，我们重新审视 Joshi 和 Nagaraja（1995）的定理 2，并说明其中某些假设与刻画母向量的 Copula 的上尾依赖系数之间的联系$(X,\ Y)$，从而为其中的公式提供直观解释，并对所引定理作出新的表述。
更重要的是，我们重新探讨了本文情形下的二元高斯分布示例，并对此进行了全面讨论。我们深入研究了 Ledford 和 Tawn (1997）中的联合尾部模型为何无法给出非退化的渐近分布，并通过展示如何在这种情形下计算合适的归一化常数，弥补了 Wadsworth 和 Tawn（2013）提出的联合尾部模型扩展中的不足。我们的结果与 Joshi 和 Nagaraja（1995）通过直接计算得到的结果一致。

## 附录 A：附录——证明

引理 1的证明 1。使用引理 1中的记号，并分别用 $\mathscr{D}(G_{\gamma_1})$,$\mathscr{D}(G_{\gamma_2})$，$\mathscr{D}(G_0)$ 表示 Fréchet、Weibull 和冈贝尔分布的最大吸引域，对于 $x, y \in \mathbb{R}$，且 $x^+$ 表示该分布的上端点，我们可以写出

$X$的，

$$
\begin{aligned}
&n\mathbb{P}(X > a_n x + b_n, Y > c_n y + d_n) \\
&= \int_{a_n x + b_n}^{x^+} n\mathbb{P}(Y > c_n y + d_n \mid X = t)\, f_X(t)\, \mathrm{d}t \\
&\leq n \int_{a_n x + b_n}^{x^+} f_X(t)\, \mathrm{d}t = n(1 - F_X(a_n x + b_n)) < \infty, \forall n \in \mathbb{N}.
\end{aligned}
$$

正如 Resnick(1987 中命题 2.5 的证明所示，以下结论成立

$$
n a_n f_X(a_n z + b_n) \mathbb{1}\left(x \leq z < a_n^{-1}(x^+ - b_n)\right) \underset{n\to\infty}{\to} \mathbb{1}(z \geq x)\, g(z)
$$

with

$$
g(z) := \begin{cases} \gamma_1 z^{-(\gamma_1+1)},\ z > 0,\ \gamma_1 > 0, & \text{if } F_X \in \mathscr{D}(G_{\gamma_1}), \\ \gamma_2(-z)^{\gamma_2 - 1},\ z < 0,\ \gamma_2 > 0, & \text{if } F_X \in \mathscr{D}(G_{\gamma_2}), \\ e^{-z},\ z \in \mathbb{R}, & \text{if } F_X \in \mathscr{D}(G_0). \end{cases}
$$

注意指示函数 $\mathbb{1}\left(x \leq z < a_n^{-1}(x^+ - b_n)\right)$ 的作用，它确保我们是在 $X$ 的支撑集上积分。为了应用扩展支配收敛定理，我们需要验证

$$
\int_{\mathbb{R}} n a_n f_X(a_n z + x^+) \mathbb{1}(z \geq x)\, \mathrm{d}z = n\left(1 - F_X(a_n x + b_n)\right) \underset{n\to\infty}{\to} \int_{\mathbb{R}} g(z) \mathbb{1}(z \geq x)\, \mathrm{d}z.
$$

对于 Gumbel 情形，我们有

$$
n\left(1 - F_X(a_n x + b_n)\right) \to \mathrm{e}^{-x} \text{ and } \int_{\mathbb{R}} g(z) \mathbb{1}(z \geq x)\, \mathrm{d}z = \mathrm{e}^{-x}.
$$

Fréchet 情形和 Weibull 情形的推理相同。由于这两个量相等，我们可以交换极限与积分，如下所示:

$$
\begin{aligned}
&\lim_{n\to\infty} \int_{\mathbb{R}} n\mathbb{P}(Y > c_n y + d_n \mid X = t)\, f_X(t) \mathbb{1}(t \geq a_n x + b_n)\, \mathrm{d}t \\
&= \int_{\mathbb{R}} \lim_{n\to\infty} \mathbb{P}(Y > c_n y + d_n \mid X = a_n z + b_n) \cdot \lim_{n\to\infty} n a_n f_X(a_n z + b_n) \\
&\qquad \cdot \mathbb{1}\left(z < a_n^{-1}(x^+ - b_n)\right) \cdot \mathbb{1}(z \geq x)\, \mathrm{d}z.
\end{aligned}
$$

由于被积函数为正，积分趋于0 当且仅当后者趋于0，也就是说，当且仅当

$$
\mathbb{P}(Y > c_n y + d_n \mid X = a_n z + b_n) = 1 - F_3\left(c_n y + d_n \mid a_n z + b_n\right) \underset{n\to\infty}{\to} 0
$$

对所有 $z \geq x$ 均成立（注意$\lim_{n\to\infty} n a_n f_X(a_n z + b_n) \neq 0$）。因此结论得证。 □

引理 2的引理 2。为了利用（15），我们用$F$表示所有条件分布，并回顾符号 $C_{\not\perp}$，适用于 $\alpha + \beta = 1$ 且 $\lambda_U > 0$的情况。至于 $F_2(y \mid x)$，我们有

$$
F_2(y \mid x) = \mathbb{P}(Y \leq y \mid X < x) = 1 - \mathbb{P}(Y > y \mid X < x) = 1 - \frac{\bar{F}_Y(y) - \bar{F}(x, y)}{F_X(x)}.
$$

由于

$$\frac{\bar{F}_Y(ny)-\bar{F}(nx,ny)}{F_X(nx)}=\mathrm{e}^{\frac{1}{nx}}\left(1-\mathrm{e}^{-\frac{1}{ny}}-\frac{\mathcal{L}(nx,ny)}{n^{\alpha+\beta}x^{\alpha}y^{\beta}}\right)$$

$$\underset{n\to\infty}{\sim}\frac{1}{n}\left(1+\frac{1}{nx}\right)\left(\frac{1}{y}-\frac{\mathcal{L}(nx,ny)}{n^{\alpha+\beta-1}x^{\alpha}y^{\beta}}\right)=\frac{1}{ny}\left(1+\frac{1}{nx}\right)\left(1-\frac{\mathcal{L}(nx,ny)}{n^{\alpha+\beta-1}x^{\alpha}y^{\beta-1}}\right)$$

$$\underset{n\to\infty}{\sim}\frac{1}{y}\left(1+\frac{1}{nx}\right)\left(1-\frac{\mathcal{L}(n,n)}{n^{\alpha+\beta-1}}r(x,ny)x^{-\alpha}y^{1-\beta}\right)\underset{n\to\infty}{\sim}\frac{1}{ny}\left(1-\lambda_U r(x,y)\left(\frac{y}{x}\right)^{\alpha}\mathbb{1}_{C_{\not\perp}}\right),$$

我们得到

$$F_2^n(ny\,|\,nx)=\left[1-\left(\frac{\bar{F}_Y(ny)}{F_X(nx)}-\frac{\bar{F}(nx,ny)}{F_X(nx)}\right)\right]^n\underset{n\to\infty}{\sim}\left[1-\frac{1}{ny}\left(1-\lambda_U r(x,y)\left(\frac{y}{x}\right)^{\alpha}\mathbb{1}_{C_{\not\perp}}\right)\right]^n.$$

Thus

$$F_2^n(ny\,|\,nx)\underset{n\to\infty}{\to}\mathrm{e}^{-\frac{1}{y}\left(1-\lambda_U r(x,y)\left(\frac{y}{x}\right)^{\alpha}\mathbb{1}_{C_{\not\perp}}\right)}.$$

对于 $F_3(y\,|\,x)$，注意到

$$\begin{aligned}F_3(y\,|\,x)&=\frac{1}{f_X(x)}\frac{\partial}{\partial x}F(x,y)=\frac{1}{f_X(x)}\frac{\partial}{\partial x}\left(F_X(x)-\bar{F}(y)+\bar{F}(x,y)\right)\\&=1+\frac{1}{f_X(x)}\frac{\partial}{\partial x}\bar{F}(x,y)=1+x^{2-\alpha}y^{-\beta}\mathrm{e}^{\frac{1}{x}}\left(\frac{\partial}{\partial x}\mathcal{L}(x,y)-\frac{\alpha}{x}\mathcal{L}(x,y)\right).\end{aligned}$$

由此可得

$$\begin{aligned}F_3(ny\,|\,nx)&\underset{n\to\infty}{\sim}1+n^{2-\alpha-\beta}\mathrm{e}^{\frac{1}{nx}}x^{2-\alpha}y^{-\beta}\left(\frac{\partial}{\partial(nx)}\mathcal{L}(nx,ny)-\frac{\alpha}{nx}\mathcal{L}(nx,ny)\right)\\&=1+\mathrm{e}^{\frac{1}{nx}}x^{2-\alpha}y^{-\beta}n^{2-\alpha-\beta}\left(\frac{1}{n}\frac{\partial}{\partial x}\mathcal{L}(nx,ny)-\frac{\alpha}{nx}\mathcal{L}(nx,ny)\right)\\&\underset{n\to\infty}{\sim}1+\mathrm{e}^{\frac{1}{nx}}n^{1-\alpha-\beta}\mathcal{L}(n,n)x^{2-\alpha}y^{-\beta}\left(\frac{\partial}{\partial x}r(x,y)-\frac{\alpha}{x}r(x,y)\right)\\&\underset{n\to\infty}{\to}1+\lambda_U x^{2-\alpha}y^{-\beta}\left(\frac{\partial}{\partial x}r(x,y)-\frac{\alpha}{x}r(x,y)\right)\mathbb{1}_{C_{\not\perp}},\end{aligned}$$

这就完成了证明。 □

定理 1的证明。证明基于引理2，以及Joshi 和 Nagaraja(1995)。首先，我们给出$(V_1,V_2)$的有限样本二元累积分布（ Joshi 和 Nagaraja，1995，公式 (2.1) ），

$$F_{(V_1,V_2)}(v_1,v_2)=\mathbb{E}\left[h\left(v_1,v_2,X_{(n-k)}\right)\right]\tag{39}$$

其中函数 $h$ 定义为

$$h(v_1,v_2,x)=F_1^k(v_1\,|\,x)F_2^{n-k-1}\left(v_2\,|\,x\right)F_3\left(v_2\,|\,x\right).$$

式（39）中的理由见Kaufmann 和 Reiss (1992)。在事件 $\{X_{(n-k)}=x\}$ 的条件下，最大值 $V_1$ 和 $V_2$ 的行为如下： $V_1$ 可视为来自 $F_1(\cdot\,|\,x)$的容量为 $k$ 的随机样本的样本最大值，而 $V_2$ 的行为类似于一个容量为 $n-k$ 的样本的样本最大值，该样本由 $F_2(\cdot\,|\,x)$ 中的 $n-k-1$ 个独立同分布变量和来自 $F_3(\cdot\,|\,x)$的一个随机变量组成。

根据Joshi 和 Nagaraja(1995)，由式（39）可写为

$$F_{(V_1,V_2)}(\tilde{a}_n v_1+\tilde{b}_n,\, c_n v_2+d_n)=\int F_1(\tilde{a}_n v_1+\tilde{b}_n \mid a_n x+b_n) F_2^{n-k-1}\left(c_n v_2+d_n \mid a_n x+b_n\right)$$
$$F_3\left(c_n v_2+d_n \mid a_n x+b_n\right) a_n f_X(a_n x+b_n)\,\mathrm{d}x.$$

由于我们的目标是计算 $V_1$ 和 $V_2$ 在 $n\to\infty$ 时的联合渐近分布，因此关于极限运算与积分之间交换的问题，我们参考 Joshi 和 Nagaraja（1995）以及 Nagaraja 和 David（1994）。由于 $X$ 服从单位 Fréchet 分布，显然边缘分布 $F_X$ 属于 $G_X$（$x$）的吸引域，$=\mathrm{e}^{-1/x}$，且 $g_X$（$x$）$=x^{-2}\mathrm{e}^{-1/x}$。关于当 $n\to\infty$ 时，项 $a_n f_X$（$a_n x+b_n$）收敛于 $g_{(k)}$（$x$）：$=$（$-\log_k G_X\, x^{\,k}$（））！$g_X$（$x$），我们参考 Nagaraja 和 David（1994）中的引理 1 和结果 1。定理 1 随即由条件（18）和引理 2 得出。 □

PROOF OF THEOREM 2。注意，$X_N>a_n x+b_n$ 等价于 $X_N>b_n$，因为当 $a_n\to 0$时 $n\to\infty$。现在，我们用标准正态边缘分布来表述Wadsworth 和 Tawn （2013年）中的模型，即

$$\begin{aligned}&\mathbb{P}(Y_E>\zeta_1(n,y)\log n \mid X_E>\log n)\\ &=\mathbb{P}\left(Y_N>\Phi^{-1}\left(1-n^{-\zeta_1(n,y)}\right)\,\middle|\, X_N>\Phi^{-1}\left(1-n^{-1}\right)\right)\\ &\sim L_1\left(n;\zeta_1(n,y),1\right) n^{1-\kappa(\zeta_1(n,y),1)}\end{aligned} \tag{40}$$

其中

(41)

$$\kappa(\zeta_1(n,y),1)=\frac{\zeta_1(n,y)+1-2\rho\sqrt{\zeta_1(n,y)}}{1-\rho^2},\qquad \text{and}$$

(42)

$$L_1(n;\zeta_1(n,y),1)=(4\pi\log n)^{\frac{2\rho^2-\rho\left(\sqrt{\zeta_1(n,y)}+1/\sqrt{\zeta_1(n,y)}\right)}{2(1-\rho^2)}}\frac{\zeta_1(n,y)^{\frac{1-\rho/\sqrt{\zeta_1(n,y)}}{2(1-\rho^2)}}(1-\rho^2)^{3/2}}{(\sqrt{\zeta_1(n,y)}-\rho)(1-\rho\sqrt{\zeta_1(n,y)})},$$

只要 $\rho^2<\min\{\zeta_1(n,y)\ 1\}$。作者解释说，缓变函数 $L$，在标准指数边缘分布的情况下（由公式（32）给出），满足 $L/L_1\sim 1$当 $n\to\infty$。

为简化记号，下文中我们用 $\zeta_1$ 表示 $\zeta_1(n,y)$。我们希望找到一些常数 $\tilde{a}_{n,E}$和 $b_{n,E}$，使得 $\{Y_E>\zeta_1(n,y)\log n\}$ 与 $\{Y_E>\tilde{a}_{n,E}y+b_{n,E}\}$等价，因此我们令

$$\zeta_1=\frac{\tilde{a}_{n,E}y+\tilde{b}_{n,E}}{\log n}. \tag{43}$$

利用（40）中的第一个等式，并注意到当 $Y_N\sim\mathcal{N}(0,1)$时，$-\log\left(1-\Phi(Y_N)\right)\sim$服从 Exp(1)，可得

$$\zeta_1=\frac{-\log\left(1-\Phi(\tilde{a}_n y+\tilde{b}_n\right)}{\log n}. \tag{44}$$

这意味着我们可以通过将边缘分布转换为标准指数分布来改写（35），然后应用模型（40）。结合（43）和（44），可得

$$\begin{aligned}
\tilde{a}_{n,E}y+\tilde{b}_{n,E} &= -\log\left(1-\Phi(\tilde{a}_n y+\tilde{b}_n\right)\\
&\underset{n\to\infty}{\sim} -\log\left(\frac{\phi(\tilde{a}_n y+\tilde{b}_n)}{\tilde{a}_n y+\tilde{b}_n}\right) \text{ through the Mill's ratio, if } \tilde{a}_n\to\infty \text{ or } \tilde{b}_n\to\infty\\
&= -\log\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-(\tilde{a}_n y+\tilde{b}_n)^2/2}\right)+\log\left(\tilde{a}_n y+\tilde{b}_n\right)\\
&= \frac{1}{2}\log(2\pi)+\frac{1}{2}\tilde{a}_n^2 y^2+\tilde{a}_n\tilde{b}_n y+\frac{1}{2}\tilde{b}_n^2+\log\tilde{b}_n+\log\left(1+\frac{\tilde{a}_n}{\tilde{b}_n}y\right)\\
&\underset{n\to\infty}{\sim} \frac{1}{2}\log(2\pi)+\frac{1}{2}\tilde{a}_n^2 y^2+\tilde{a}_n\tilde{b}_n y+\frac{1}{2}\tilde{b}_n^2+\log\tilde{b}_n+\frac{\tilde{a}_n}{\tilde{b}_n}y \quad \text{if } \tilde{a}_n=o\left(\tilde{b}_n\right)\\
&= \tilde{a}_n\left(\tilde{b}_n+\frac{1}{\tilde{b}_n}\right)y+\frac{1}{2}\log(2\pi)+\frac{1}{2}\tilde{a}_n^2 y^2+\frac{1}{2}\tilde{b}_n^2+\log\tilde{b}_n\\
&\underset{n\to\infty}{\sim} \tilde{a}_n\tilde{b}_n y+\frac{1}{2}\tilde{b}_n^2 \quad \text{(the discarded terms are } o\left(\tilde{a}_n\tilde{b}_n\right)\text{).}
\end{aligned}$$

因此，我们可以识别这些项，从而得到

$$\tilde{a}_{n,E}=\tilde{a}_n\tilde{b}_n, \qquad \tilde{b}_{n,E}=\frac{1}{2}\tilde{b}_n^2.$$

注意，到目前为止，我们尚未使用(29b)中的常数，而是得到了关于它们随 $n$ 增大而增长的一般条件。利用(29b)，可得

$$\begin{aligned}
\tilde{a}_{n,E} &= \rho\sqrt{1-\rho^2}\left(\sqrt{2\log n}-\frac{1}{2}\frac{\log(4\pi\log n)}{\sqrt{2\log n}}\right)\\
\tilde{b}_{n,E} &= \rho^2\log n-\frac{\rho^2}{2}\log(4\pi\log n)+\frac{\rho^2}{16}\frac{(\log(4\pi\log n))^2}{\log n}.
\end{aligned}$$

将(43)中的常数代入，得到

$$\zeta_1=\rho^2\left(1+\frac{1}{\rho}\sqrt{\frac{2(1-\rho^2)}{\log n}}y-\frac{1}{2}\frac{\log(4\pi\log n)}{\log n}+o\left((\log n)^{-1}\right)\right). \tag{46}$$

现在注意，由(46)中定义的 $\zeta_1$ 可知，对于较大的 $n$ 以及 $\rho\in(0,1)$，条件 $\rho^2<\min\{\zeta_1(n,y)\ 1\}$ 总是成立，因为它给出 $\rho^2<\rho^2+\mathcal{O}\left((\log n)^{-1/2}\right)$。因此，由于

$$\begin{aligned}
\sqrt{\zeta_1} &= \rho\left(1+\frac{1}{\rho}\sqrt{\frac{2(1-\rho^2)}{\log n}}y-\frac{1}{2}\frac{\log(4\pi\log n)}{\log n}+o\left((\log n)^{-1}\right)\right)^{1/2}\\
&\underset{n\to\infty}{\sim} \rho\left(1+\frac{1}{2\rho}\sqrt{\frac{2(1-\rho^2)}{\log n}}y-\frac{1}{4}\frac{\log(4\pi\log n)}{\log n}-\frac{1-\rho^2}{4\log n}y^2+o\left((\log n)^{-1}\right)\right),
\end{aligned}$$

$\kappa(\zeta_1,\ 1)$ 的表达式变为

$$\kappa(\zeta_1,1)=1+\frac{\rho^2}{2\log n}y^2+o\left((\log n)^{-1}\right).$$

将这些表达式代入方程（40）中，可得

$$\begin{aligned}&\mathbb{P}(Y_E > \zeta_1(n,y)\log n \mid X_E > \log n)\\ &\underset{n\to\infty}{\sim} \frac{\exp\left\{\frac{2\rho^2-\rho\left(\sqrt{\zeta_1}+1/\sqrt{\zeta_1}\right)}{2(1-\rho^2)}\log\left(4\pi\log n\right)+\frac{1-\rho/\sqrt{\zeta_1}}{2(1-\rho^2)}\log\zeta_1+(1-\kappa(\zeta_1,1))\log n\right\}}{(1-\rho^2)^{-3/2}(\sqrt{\zeta_1}-\rho)(1-\rho\sqrt{\zeta_1})}.\end{aligned}$$

由于

$$\begin{aligned}\left(1-\frac{\rho}{\sqrt{\zeta_1}}\right)\log\zeta_1 \underset{n\to\infty}{\sim}& \frac{\log\rho}{\rho}\sqrt{\frac{2(1-\rho^2)}{\log n}}y+\frac{1}{2\rho^2}\frac{2\left(1-\rho^2\right)}{\log n}y^2\\ &-\frac{\log\rho}{2}\frac{\log\left(4\pi\log n\right)}{\log n}-\frac{1-\rho^2}{2}\log\rho\frac{y^2}{\log n}+o\left((\log n)^{-1}\right)\\ =&\frac{\log\rho}{\rho}\sqrt{\frac{2(1-\rho^2)}{\log n}}y-\frac{\log\rho}{2}\frac{\log\left(4\pi\log n\right)}{\log n}\\ &+\left(\log\rho-\frac{2}{\rho^2}\right)\frac{1-\rho^2}{2\log n}y^2+o\left((\log n)^{-1}\right),\end{aligned}$$

$$\begin{aligned}(\sqrt{\zeta_1}-\rho)(1-\rho\sqrt{\zeta_1}) \underset{n\to\infty}{\sim}& \frac{(1-\rho^2)^{3/2}}{\sqrt{2logn}}y-\rho\frac{1-\rho^2}{4}\frac{\log\left(4\pi\log n\right)}{\log n}\\ &-\rho\frac{1-\rho^4}{4\log n}y^2+o\left((\log n)^{-1}\right)\end{aligned}$$

and

$$\begin{aligned}\frac{2\rho^2-\rho\left(\sqrt{\zeta_1}+1/\sqrt{\zeta_1}\right)}{2(1-\rho^2)} \underset{n\to\infty}{\sim}& -\frac{1}{2}+\frac{1}{4\rho}\sqrt{\frac{2(1-\rho^2)}{\log n}}y-\frac{1}{8}\frac{\log\left(4\pi\log n\right)}{\log n}\\ &-\frac{1-\rho^2}{8\log n}y^2+o\left((\log n)^{-1}\right),\end{aligned}$$

其中我们忽略了所有相对于$(\log n)^{-1}$衰减更快的项，因此得到

$$\mathbb{P}(Y_E > \zeta_1(n,y)\log n \mid X_E > \log n) \underset{n\to\infty}{\sim} \frac{1}{y\sqrt{2\pi}}\mathrm{e}^{-(\rho y)^2/2}.$$

□

推论2的证明。注意到

$$\mathbb{P}(Y_E > \zeta_1(n,y)\log n \mid X_E > \log n) \underset{n\to\infty}{\sim} \frac{\rho}{\rho y\sqrt{2\pi}}\mathrm{e}^{-(\rho y)^2/2} = \rho\frac{\varphi(\rho y)}{\rho y},$$

其中 $\varphi$ 表示标准正态分布的概率密度函数。假设 $y$ 较大，利用米尔斯比率可以直接证明,

$$\mathbb{P}(Y_E > \zeta_1(n,y)\log n \mid X_E > \log n) \underset{n\to\infty}{\sim} \rho\left(1-\Phi(\rho y)\right),$$

由此完成证明。 □

## 附录 B：附录——关于 $k$ 的讨论

本节展示了第 3 节所讨论示例的一些图形结果。我们回顾一下，每个示例都通过模拟进行 5000 次重复试验，每次试验包含 100000 个双变量样本（$X$，$Y$）。因此，对于每个 $k$，我们得到一个维度为 5000 的（$V_1$，$V_2$）双变量样本（每个（$X$，$Y$）样本的重复试验都会产生一对（$V_1$，$V_2$））。表示极端集合维度的 $k$ 取值固定为 1、50、100、250、500、1000、1500 和 2000。不过，图中并未展示所有取值对应的结果。这样安排是为了便于阅读，因为在同一页上放置许多几乎相同的图可能会显得冗余。快速浏览两个示例的图 4 至图 7，可以得出相同的重要结论：

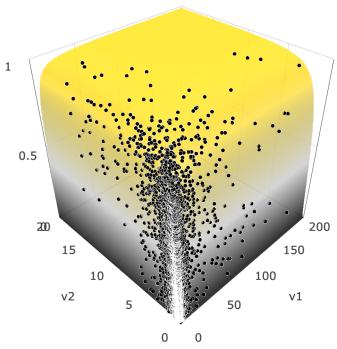
(a) 相对于 $V_1$ 的取值范围远大于对应于 $V_2$ 的取值范围；(b) 图 4 和图 6 中曲面上的黑点云逐渐靠近 $v_1$ 轴，随着 $k$ 增大。

这是 $(X, Y)$ 之间正相关依赖的明显影响。要理解这一点，可以考虑 $(V_1, V_2)$ 的定义：$V_1$ 是最大的 $k$ 个次序统计量所对应伴随变量的最大值，而 $V_2$ 是其余集合中的最大值。两个样本取值范围的明显差异表明，最高的伴随变量对 $V_1$ 的取值有贡献，而最低的伴随变量则导致 $V_2$ 的取值相对较低。此外，随着 $k$ 增大，$V_2$ 取值几乎都低于 1，而 $V_1$ 则可能取大于 100 的值。换言之，伴随变量的高值通常对应于 $X$ 的高次序统计量。

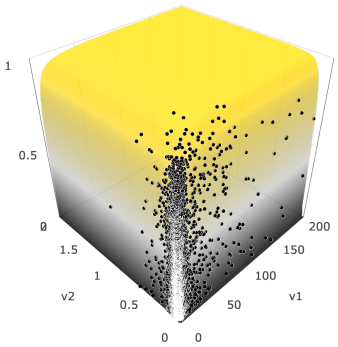
需要指出的是，曲面的绘图范围是为了便于观察而选定的，因为区间过宽不利于观察结果。不过，也可能有一些点落在绘图区域之外：图4e 的中右部包含更多点，而图4f 的右上部包含更多点。这表明，当 $k$ 为1000时，一些数值较高的伴随变量并不属于极端集合；但当 $k$ 增加到2000时，它们所属的划分发生了变化。换言之，随着 $k$ 增加，$V_1$ 会从 $V_2$ 那里“夺走”较高的数值。

边缘分布图支持上述结论：图5a 和图7a所示的对数累积分布函数没有明显差异，这意味着 $V_1$ 中的样本变化不大；而 $V_2$ 分布的上端点似乎向0 收缩（图5b 和图7b）。这再次印证了“夺取”这一说法：随着 $k$ 增加，对 $V_2$ 有贡献的较高伴随变量值进入极端集合，因此只剩下较低的伴随变量值用于计算 $V_2$，使其越来越小。如前所述，这是$(X, Y)$之间正相关依赖的影响。

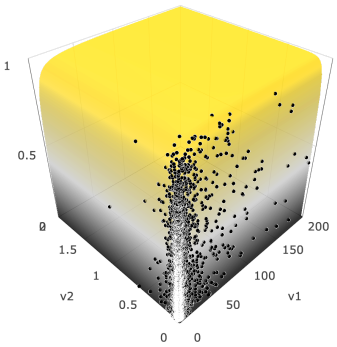
最后，图4g 和图6g 中的箱线图显示误差极低，这确保了渐近近似几乎完美。从图5和图7中展示的边际误差图也可以得出同样的结论。尽管当 $v_1$ 和 $v_2$ 取较低值时误差存在一定波动（而且这种波动本来就很小），但随着 $v_1$ 和 $v_2$增大，误差会迅速趋近于0 。这对实际应用很重要，因为我们通常关注高阈值下的情况。



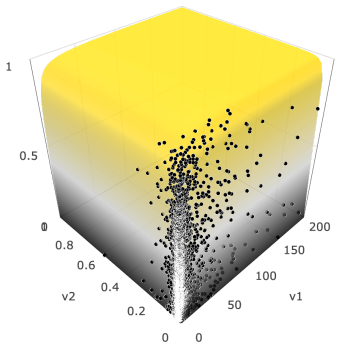

(a) $k = 1$。



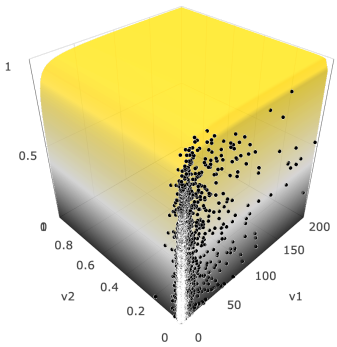

(b) $k = 50$。



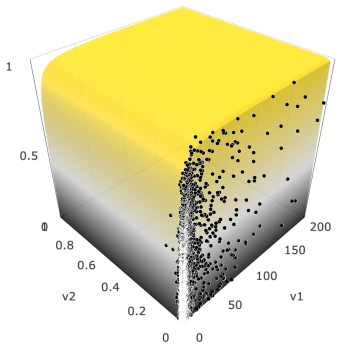

(c) $k = 100$。



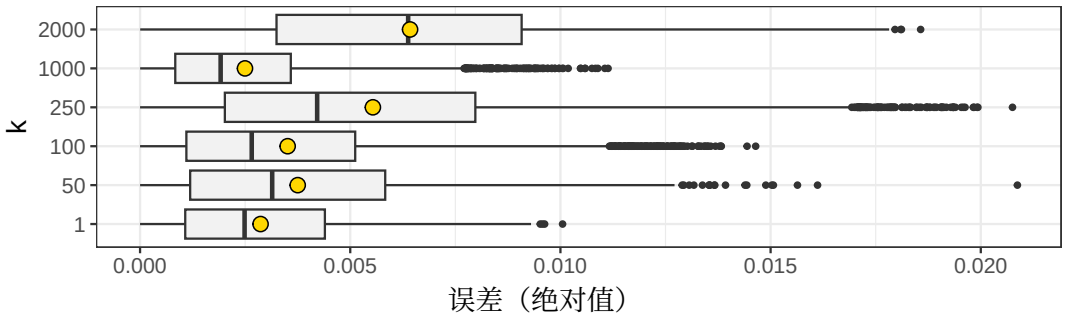

(d) $k = 250$。




(e) $k = 1000$。

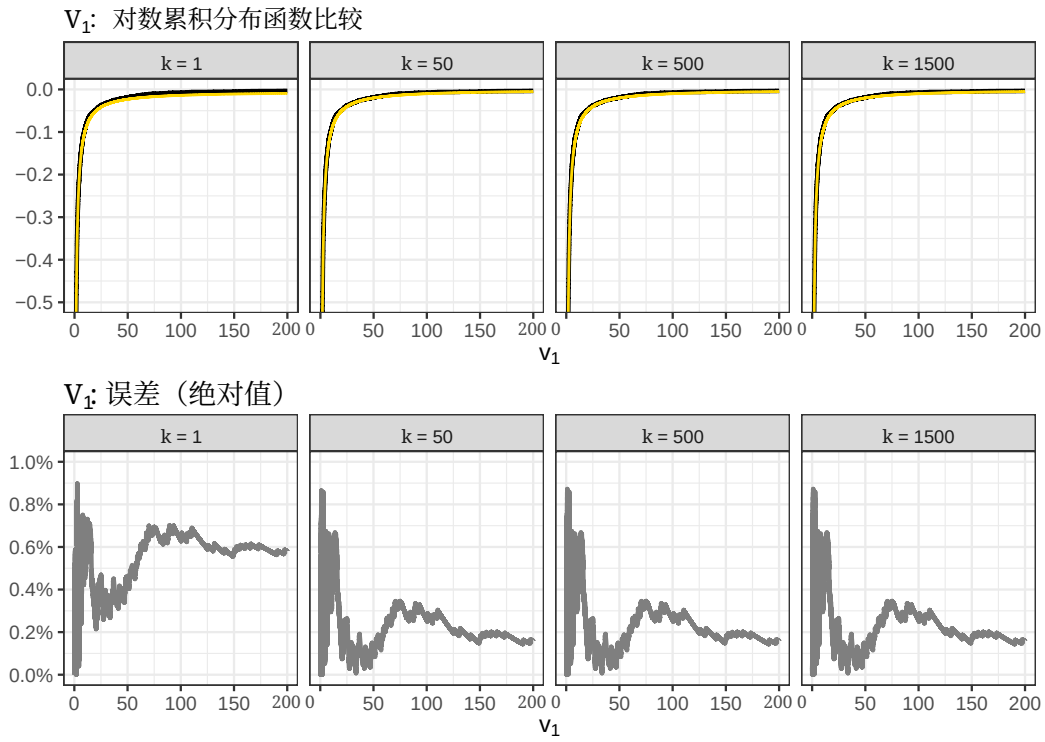



(f) $k = 2000$。

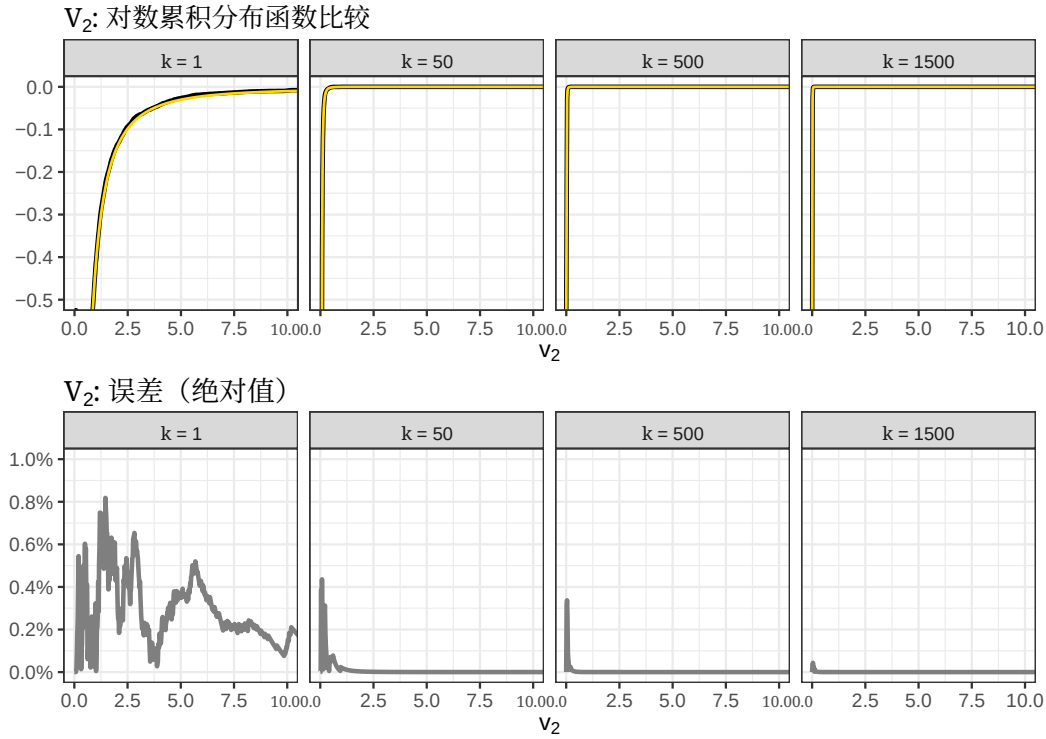



(g)误差箱线图。

图4：生存 Clayton$(\theta = 2)$。通过对方程（19）进行数值积分计算得到的二元累积分布函数，针对不同的 $k$ 取值。黑色圆点对应模拟样本的二元经验累积分布函数。面板 g) 显示了在 $(V_{1,i}, V_{2,i})$ 中计算得到的二元渐近累积分布函数与对应的二元经验累积分布函数之间绝对值误差的箱线图，$i = 1, \ldots, 5000$。金色圆点表示误差均值。

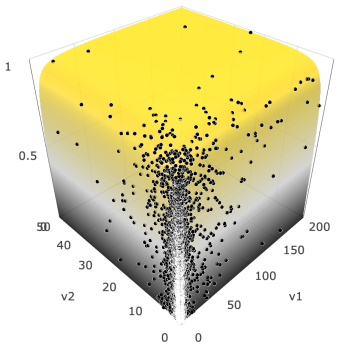



(a) $V_1$：关于 $v_1$的对数边缘累积分布函数比较和边缘误差行为。

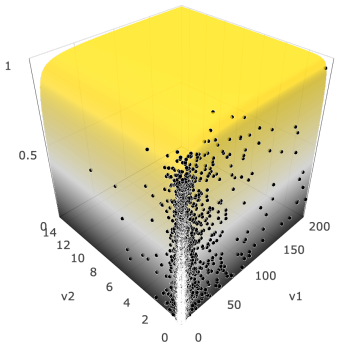



(b) $V_2$：关于 $v_2$的对数边缘累积分布函数比较和边缘误差行为。

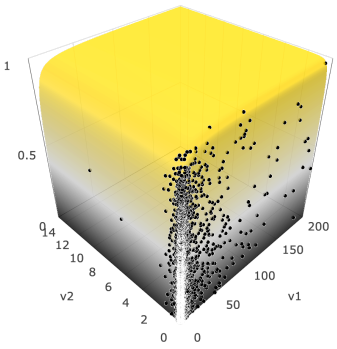
图5：生存 Clayton($\theta = 2$)。和 $V_2$的经验单变量累积分布函数，以及通过对渐近联合累积分布函数（19）进行数值积分得到的渐近理论值。



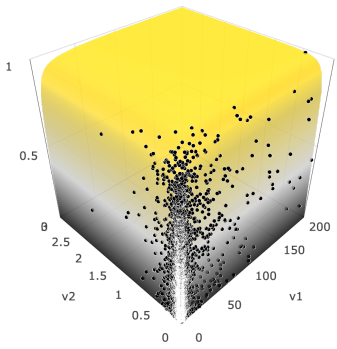

(a) $k = 1$。



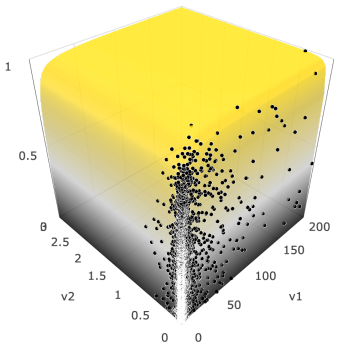

(b) $k = 50$。



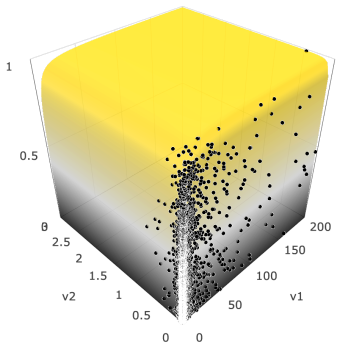

(c) $k = 100$。



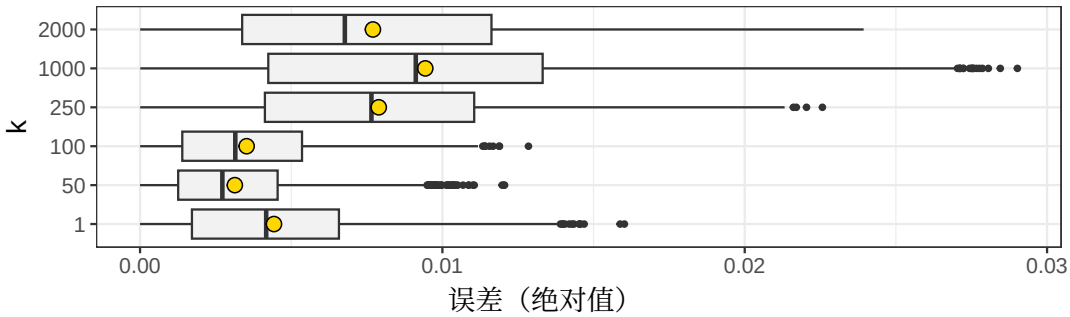

(d) $k = 250$。




(e) $k = 1000$。

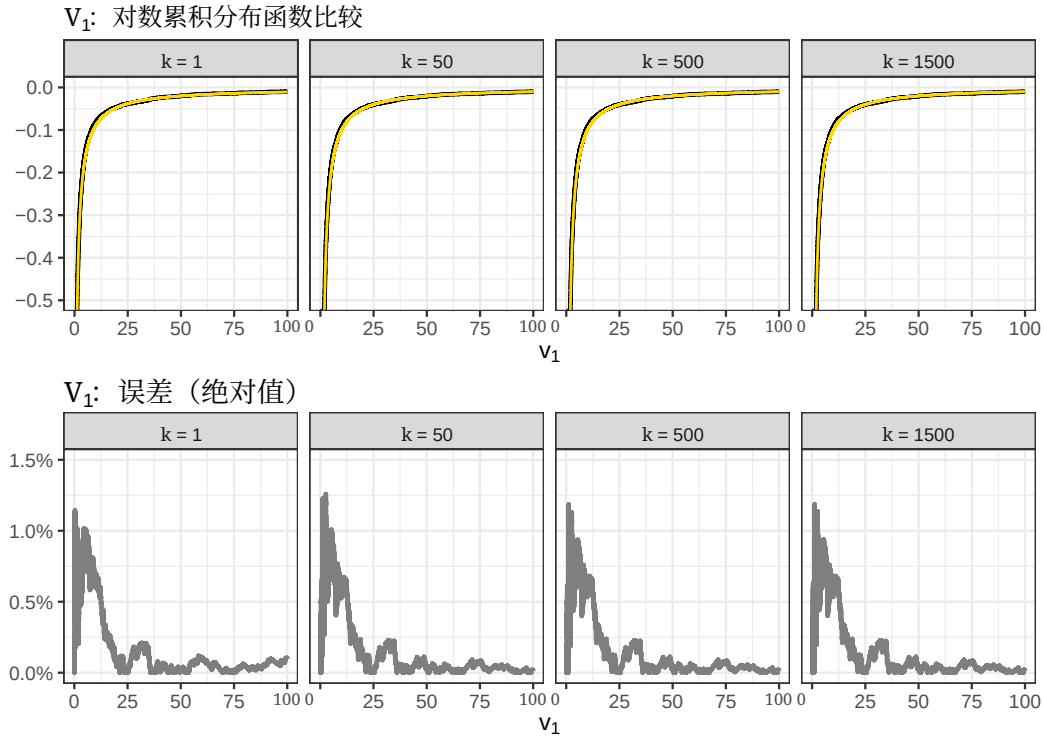



(f) $k = 2000$。

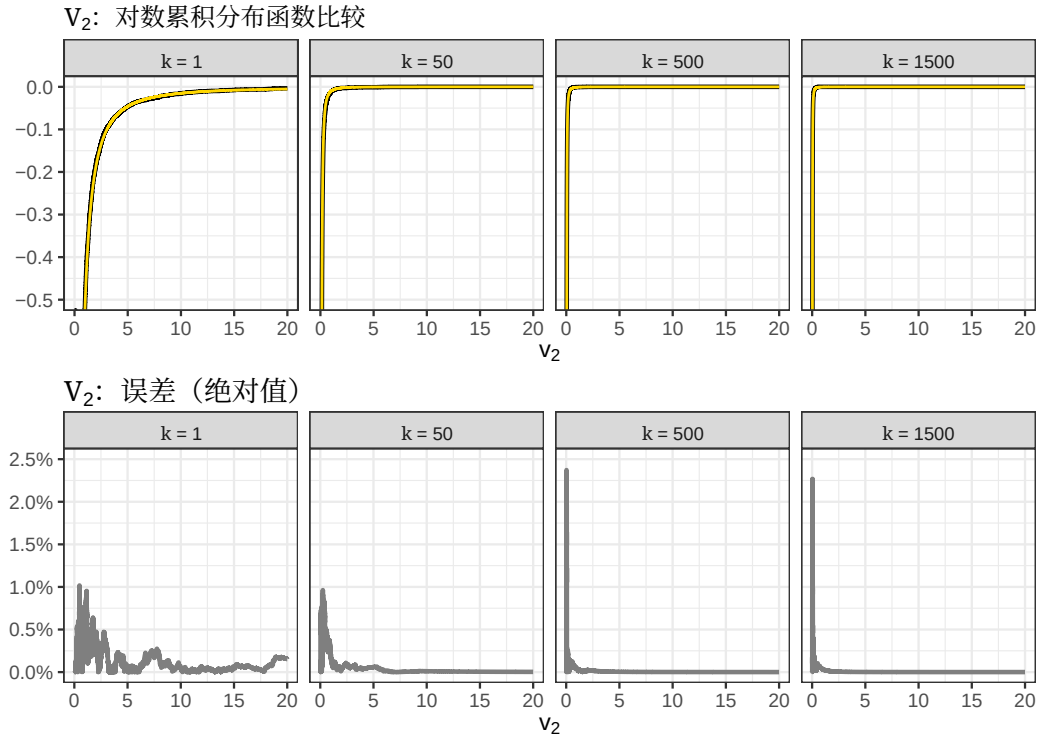



(g)误差箱线图。

图 6：逻辑分布($\gamma = 1/2$)。通过对方程（19）进行数值积分得到的二元累积分布函数，取 $k$ 的不同值。黑色圆点对应模拟样本的二元经验累积分布函数。面板 g）显示了在 $(V_{1,i}, V_{2,i})$ 中计算的二元渐近累积分布函数与相应二元经验累积分布函数之间绝对误差的箱线图，其中 $i = 1, \ldots\ldots, 5000$。金色圆点表示误差均值。




(a) $V_1$：对数边缘累积分布函数比较和边缘误差行为，相对于 $v_1$。




(b) $V_2$：对数边缘累积分布函数比较和边缘误差行为，相对于 $v_2$。

图7：逻辑分布($\gamma = 1/2$) 示例。 $V_1$ 和 $V_2$的经验单变量累积分布函数，以及通过对渐近联合累积分布函数（19）进行数值积分得到的渐近理论值。

## 附录 C：附录——冯·米塞斯条件

本节介绍冯·米塞斯条件，依据Resnick(1987)中的表述，并沿用其符号。

命题 *1.15 (Resnick (1987)):* 假设 $F$ 在$f$ 的某邻域内绝对连续且具有正密度 $\infty$。

(a) 若存在某个 $\alpha > 0$

$$\lim_{x\to\infty} xf(x)/(1-F(x)) = \alpha \tag{47}$$

则 $F \in \mathscr{D}(\Phi_\alpha)$。我们可以选取 $a_n$ 使其满足 $a_n f(a_n) \sim \alpha/n$。
(b) 若 $f$ 是非递增的，且 $F \in \mathscr{D}(\Phi_\alpha)$ 则(47) 成立。
(c) 等式 (47) 成立，当且仅当存在 $z_0$ 且对所有 $x > z_0$，均有

$$1 - F(x) = c\exp\left\{-\int_{z_0}^{x} t^{-1}\alpha(t)\,dt\right\}$$

其中 极限$_{t\to\infty}\alpha(t) = \alpha$，且 $c$ 是正常数。

命题 *1.16*（*Resnick (1987)*）：假设 $F$具有有限右端点 $x_0$，且在 $x_0$ 的左邻域内绝对连续，并具有正密度 $f$。

(a) 若存在某个 $\alpha > 0$

$$\lim_{x\uparrow x_0} (x_0 - x)\, f(x)/(1-F(x)) = \alpha \tag{48}$$

则 $F \in \mathscr{D}(\Psi_\alpha)$。
(b) 若 $f$ 是非递增的，且 $F \in \mathscr{D}(\Psi_\alpha)$ 则 (48) 成立。
(c) 等式 (48) 成立，当且仅当 $c(x)$ 在表示中可取为在 $x_0$ 的某个左邻域内为常数

$$1 - F(x) = c(x)\exp\left\{-\int_{x_0-1}^{x} \delta(t)/(x_0 - t)\,dt\right\} \quad \textit{and for } x < x_0$$

其中 $\lim_{t\uparrow x_0}\delta(t) = \alpha$，$\lim_{t\uparrow x_0} c(t) = c_0 > 0$。

命题 *1.17 (Resnick(1987)):* 设 $F$ 在$x_0$ 的某个左邻域内绝对连续，且密度为 $f$。

(a) *If*

$$\lim_{t\uparrow x_0} f(x)\int_{x}^{x_0} (1-F(t))\,dt/(1-F(x))^2 = 1 \tag{49}$$

那么 $F \in D(\Lambda)$。在这种情况下，我们可以取

$$f(t) = \int_{t}^{x_0} (1-F(s))\,ds/(1-F(x))$$

$$b_n = (1/(1-F))^{\leftarrow}(n), \quad a_n = f(b_n).$$

(b) 若 $f$ 非递增且 $F \in \mathscr{D}(\Lambda)$，则成立。

(c) 方程 (49)当且仅当成立

$$1 - F(x) = c \exp\left\{ - \int_{z_0}^{x} (g(t)/f(t))\, dt \right\}, \quad z_0 < x < x_0, \tag{50}$$

其中，当 $x \uparrow x_0$ 时，$\lim_{t \uparrow_0} g(x) = 1$，且 $f$ 绝对连续，其密度为 $f(x) \to 0$。*(d)* 方程（*49*）或（*50*）等价于 $tf((1/(1-F))^{\leftarrow}(t)) \in RV_0$（*Sweeting*（*1985*））。

**致谢**。本文主要撰写于第一作者在法国巴黎的 ESSEC商学院访学期间，以及第二作者在意大利都灵的都灵大学 ESOMAS系访学期间。两位作者均对接待机构表示感谢。

**经费支持**。作者感谢 Labex MME DII (ANR-11-LABX-0023- 01) 和 ESOMAS系（都灵大学）提供的经费支持。

## 参考文献

BALCI, S., AKKAYA, A. D. 和 ULGEN, B. E.（2013）。使用秩集抽样的修正最大似然估计量。《计算与应用数学杂志》238 171-179。https://doi.org/10.1016/j.cam.2012.08.030 COLES, S., HEFFERNAN, J. 和 TAWN, J.（1999）。《极值分析的依赖度量》。《极值》2 339–365。DACOROGNA, M. M. 和 KRATZ, M.（2015）。《生活在随机世界中并管理复杂风险》。可于 SSRN 获取：https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2668468。https://doi.org/10.2139/ssrn.2668468 DE HAAN, L. 和 ZHOU, C.（2011）。随机向量和过程的极端残差依赖。《应用概率进展》43 217 – 242。https://doi.org/10.1239/aap/1300198520 DELL, T. R. 和 CLUTTER, J. L.（1972）。《秩集抽样理论：次序统计量基础》。《生物计量学》28 545–555。FISHER, R. A. 和 TIPPETT, L. H. C.（1928）。样本中最大或最小观测值的频数分布的极限形式。《剑桥哲学学会数学论文集》24 180-190。https://doi.org/10.1017/S0305004100015681 GNEDENKO, B.（1943）。关于随机序列最大项的极限分布。《数学年刊》44 423–453。HEFFERNAN, J. E.（2000）。《尾部依赖系数目录》。《极值》3 279–290。HEFFERNAN, J. E. 和 RESNICK, S. I.（2007）。具有极端分量的随机向量的极限定律。《应用概率年刊》17 537 – 571。https://doi.org/10.1214/105051606000000835 JOSHI, S. N. 和 NAGARAJA, H. N.（1995）。次序统计量子集的伴随变量最大值的联合分布。《伯努利》1 245 – 255。https://doi.org/10.3150/bj/1193667817 KAUFMANN, E. 和 REISS, R. D.（1992）。最近邻的条件分布。《多元分析杂志》42 67-76。https://doi.org/10.1016/0047-259X(92)90079-U LEDFORD, A. W. 和 TAWN, J. A.（1997）。《联合尾部区域内的依赖建模》。《英国皇家统计学会杂志》。B辑（方法论）59 475–499。LEDFORD, A. W. 和 TAWN, J. A.（1998）。《极值的伴随尾部行为》。《应用概率进展》30 197–215。https://doi.org/10.1239/aap/1035228000 LEONETTI, P. 和 KHORRAMI CHOKAMI, A.（2022）。多元极值分布的最大吸引域很小。《概率电子通讯》27 1 – 8。https://doi.org/10.1214/22-ECP501 NAGARAJA, H. N. 和 DAVID, H. A.（1994）。《选定次序统计量伴随变量最大值的分布》。《统计年刊》22 478–494。https://doi.org/10.1214/aos/1176325380 OZTURK, O.（2019）。采用秩集抽样设计的两阶段整群样本。《统计数学研究所年刊》71 63–91。RESNICK, S. I.（1987）。《极值、正则变差与点过程》。应用概率4。施普林格，纽约。初版。https://doi.org/10.1007/978-0-387-75953-1 SHAO, X.（2010）。《依赖野生自助法》。《美国统计学会杂志》105 218-235。https://doi.org/10.1198/jasa.2009.tm08744

STOKES, S. L. and SAGER, T. W. (1988). 对等级集样本的刻画及其在估计分布函数中的应用。《美国统计学会杂志》**83** 374–381. SWEETING, T. J.（1985）。极值理论中一致局部吸引域。 概率年刊**13** 196 – 205. https://doi.org/10.1214/aop/1176993075 WADSWORTH, J. L. 和 TAWN, J. A.（2013）。多元尾概率的一种新表示。《伯努利》**19** 2689 – 2714。https://doi.org/10.3150/12-BEJ471 WANG, X., AHN, S. 和 LIM, J.（2017）。整群随机研究中的不平衡等级集抽样。统计规划与推断杂志**187** 1-16。https://doi.org/10.1016/j.jspi.2017.02.005 WANG, X., LIM, J. 和 STOKES, L.（2016）。在整群随机设计中使用秩集抽样以改进处理效应推断。《美国统计学会杂志》**111** 1576-1590。https://doi.org/10.1080/01621459.2015.1093946 WANG, X., STOKES, L., LIM, J. 和 CHEN, M.（2006）。多元次序统计量的伴随变量及其在判断后分层中的应用。《美国统计学会杂志》**101** 1693–1704。ZAMANZADE, E. 和 WANG, X.（2018）。存在并列信息时秩集抽样中的比例估计。计算统计学**33** 1349–1366。