arXiv:2602.01144v1 [math.ST] 1 Feb 2026

基于 Sklar定理和经验棋盘近似估计条件分布及其对非参数回归的影响

Kai Sch¨arer*，Wolfgang Trutschnig†

摘要

我们探讨一个自然的问题：能否通过分别估计底层 Copula 的条件分布和边缘分布，利用 Sklar定理来估计条件分布。我们采用所谓的经验棋盘／伯恩斯坦近似，并选择适当的分辨率／次数，首先证明，在非常宽松的正则性假设下，可以实现向真实底层 Copula 的一致弱收敛。在此基础上，代入一元经验边缘分布函数，我们对上述问题给出肯定答案，并证明所得条件分布估计量具有强相合性。此外，我们还表明，对这些估计量进行聚合，可以构造出均值、分位数和期望分位数回归函数等的相合非参数估计量。最后，我们通过一些模拟展示估计量的表现，并给出一个真实数据示例，以补充已建立的理论结果。

Keywords：条件分布、Copula、非参数回归*2020 MSC*：60E05, 62G05, 62G08, 62H05

1 引言

非参数回归技术能够灵活地建模/捕捉响应变量 Y与协变量向量$\mathbf{X}$ 之间的关系，而不施加严格的参数假设，因此在现代统计学及所有以定量研究为导向的领域中发挥着核心作用。经典回归方法通常聚焦于单个点估计，例如条件均值，涵盖可追溯至 Nadaraya (1964) 和 Watson (1964) 的核估计方法，以及局部多项式

*奥地利萨尔茨堡大学人工智能与人类界面系（AIHI），奥地利萨尔茨堡。 kai.schaerer@plus.ac.at

†奥地利萨尔茨堡大学人工智能与人类界面系（AIHI），奥地利萨尔茨堡。 wolfgang@trutschnig.net，通讯作者

建模技术（由 Fan (1996) 提出）。尽管均值回归可谓数据分析中最常用的工具之一，但它只能提供有关条件分布的有限信息，可能会忽略偏度、厚尾和异方差性等重要特征。

一种旨在克服这些局限的自然扩展是非参数分位数回归。该方法直接估计条件分位数函数，因此能够更充分地描述响应变量与协变量之间的关系。非参数分位数回归的关键技术包括 Koenker 和 Gilbert Bassett（1978）提出的线性分位数估计量、Chaudhuri（1991）提出的估计量，以及 Huang 和 Nguyen（2018）提出的基于密度估计的方法。以 Copula 为基础的非参数分位数回归的既有研究聚焦于藤Copula（作为成对Copula构造的一种特例），例如 Tepegjozova 等人（2021）对此进行了讨论。

均值回归的另一种替代方法是期望分位数回归，可将其视为均值回归的非对称版本。它是刻画尾部风险和非对称条件行为的有力工具；Newey 和 Powell（1987）研究的非对称最小二乘估计量，是首个广受认可的期望分位数回归方法。与均值回归和分位数回归的情形一样，人们也开发了非参数方法，例如 Yao 和 Tong（1996）研究的基于核的估计量，以及 Adam 和 Gijbels（2022）分析的局部多项式建模技术。与（非参数）均值回归和分位数回归不同，非参数期望分位数回归的相关文献寥寥无几，更不用说能够同时适用于均值、分位数和期望分位数的统一方法了。

显然，对条件分布函数进行非参数估计，再将其汇总，能够回应上述三个研究问题。Fan（1993）提出的朴素核分布回归面临多种问题，例如非单调性（Hall 等，1999）；而这一问题最近由 Das 和 Politis（2020）加以解决。

聚焦于连续情形，并考虑到根据 Sklar 定理（见 Sklar（1959）），Copula 构成了联合分布函数与其一元边缘分布之间的纽带，因此，通过分别对潜在的 Copula 和边缘分布建模或进行估计来处理（均值、分位数、期望分位数）回归，似乎是自然而然的做法。沿着这一思路，Noh 等人（2013）针对高维预测变量情形，提出了一种用于估计回归函数的新型半参数估计量。然而，正如 Dette 等人（2014）所示并讨论的，即便是在二元情形中，使用参数族时，Copula 类的任何设定错误“都可能导致无效结果”。

本文重新探讨通过 Sklar定理处理回归问题，以及更广义地说，估计条件分布的思路。我们聚焦于二元情形，并提出一种新颖的分布回归方法：通过经验联结函数的所谓棋盘近似和伯恩斯坦近似，对底层 Copula 进行非参数估计（我们分别将其称为经验棋盘近似和经验伯恩斯坦逼近）。利用近期关于（经验）棋盘近似弱条件收敛的研究成果（见 Ansari 等人（2026）），我们证明，在相当宽松的正则性条件下，我们的方法能够得到强相合且计算快速的条件分布估计量。此外，我们还证明，对这些估计量进行聚合，可得到均值回归、分位数回归和期望分位数回归等方法的相合估计量。作为对棋盘方法的补充，我们随后聚焦于经验伯恩斯坦逼近，并推导出类似的一致性结果。据我们所知，

经验棋盘近似和伯恩斯坦近似的收敛速率目前仍然未知，因此我们通过一项小规模模拟研究检验所提出估计量的表现，并将其与其他知名的基于核的非参数回归方法进行比较。

本文其余部分安排如下。第 2 节汇集了后续各节所需的符号和预备知识。第 3 节和第 4 节分别建立了棋盘格方法和伯恩斯坦方法的收敛结果。如第 5 节所示，这些结果为构造前文提到的三类回归函数的相合估计量铺平了道路。第 6 节通过模拟研究展示所建立估计量的性能，第 7 节则以一个基于保险数据的实际案例为本文收尾。

2 符号与预备知识

对于每个度量空间 $(S,\ d)$，其 Borel σ域记作 $\mathcal{B}(S)$，以 $x \in S$ 为中心、半径为 $r > 0$ 的开球记作 $B_r(x)$。闭单位区间记作 $\mathbb{I} := [0,1]$，$\mathcal{B}(\mathbb{I}^2)$ 和 $\mathcal{B}(\mathbb{I})$ 上的勒贝格测度分别记作 λ_2 和 λ。对于每个 $N \geq 2$，我们将使用以下划分

$$I_1^N := \left[0, \tfrac{1}{N}\right], I_2^N := \left(\tfrac{1}{N}, \tfrac{2}{N}\right], \ldots, I_N^N := \left(\tfrac{N-1}{N}, 1\right]$$

的 $\mathbb{I}$，并（为简化本文其余部分的某些公式）令 $I_0^N := \emptyset =: I_{N+1}^N$。我们定义 $Q_{i,j}^N := I_i^N \times I_j^N \subset \mathbb{I}^2$，对所有 $i,j \in \{0,\ldots,N+1\}$，并且给定 $(x,y) \in \mathbb{I}^2$，令 $i_0^N(x)$ $j_0^N(y) \in \{1,\ldots,N\}$ 表示满足 $(x,y) \in Q_{i_0^N(x)\ j_0^N(y)}^N$ 的唯一指标。

$\mathcal{C}$ 表示所有二元 copula 的族，$\mathcal{P}$ 表示 $\mathcal{B}(\mathbb{I}^2)$ 上所有双随机测度的族，Π 表示乘积 copula。众所周知，每个 Copula A 对应于唯一的双随机测度 μ_A，反之亦然，其对应关系由恒等式 $A(x,y) = \mu_A([0,x] \times [0,y])$ 给出，对所有 $x,y \in \mathbb{I}$。若相应的双随机测度 μ_A 是绝对连续的，则称 Copula A 为绝对连续的。照例，d_∞ 表示 $\mathcal{C}$ 上的一致度量，即 $d_\infty(A,B) := \max_{(x,y)\in[0,1]^2} |A(x,y) - B(x,y)|$。众所周知，$(\mathcal{C}, d_\infty)$ 是一个紧致度量空间。对于随机变量 X,Y，若 (X,Y) 具有分布函数 H，则记 $(X,Y) \sim H$；若 A 是 (U,V) 限制在 $\mathbb{I}^2$ 上的分布函数，则记 $(U,V) \sim A \in \mathcal{C}$；在后一种情形下，显然 U 与 V 在 $\mathbb{I}$ 上均匀分布。关于 Copula 的更多背景，请参阅 Durante 和 Sempi（2015）以及 Nelsen（2006）。

假设 $(x_1,\ y_1)$ ……，$(x_n,\ y_n)$ 是 $(X,\ Y)$ 的一个样本，其中 H 是连续的，并假设 A 是其底层 Copula。此外，令 H_n 表示二元经验分布函数，令 $F_n,\ G_n$ 表示一元经验边缘分布函数。Sklar定理表明，存在唯一的子联结函数（参见 Nelsen (2006)）E_n'：Range$(F_n) \times$ Range$(G_n) \to$ Range(H_n)，满足

$$H_n(x,y) = E_n'(F_n(x), G_n(y)),$$

对所有 $(x,\ y) \in \mathbb{R}^2$。由于 H 连续，并列值出现的概率为 0，因此几乎处处有 Range(F_n) = Range$(G_n) = \{0, \frac{1}{n}, \ldots, \frac{n-1}{n}, 1\}$。将 E_n' 扩展为 Copula E_n 的方法有无穷多种；下文中我们只采用双线性插值（见 Junker 等人（2021）；Nelsen（2006））并将这一扩展 E_n 简称为样本 $(x_1,\ y_1)$……$(x_n,\ y_n)$ 的经验联结函数。

若映射 $K\colon \mathbb{R}\times\mathcal{B}(\mathbb{R})\to[0,1]$ 是一个从 $\mathbb{R}$ 到 $\mathcal{B}(\mathbb{R})$ 的马尔可夫核，当且仅当对每个固定的 $B\in\mathcal{B}(\mathbb{R})$，$x\mapsto K(x,B)$ 都是可测的，并且对每个固定的 $x\in\mathbb{R}$，$B\mapsto K(x,B)$ 都是一个概率测度。若马尔可夫核 $K\colon \mathbb{R}\times\mathcal{B}(\mathbb{R})\to[0,1]$ 是随机变量 Y 在给定（另一个随机变量） X 条件下的正则条件分布，则对每个 $B\in\mathcal{B}(\mathbb{R})$

$$K(X(\omega),B)=\mathbb{E}(\mathbf{1}_B\circ Y\mid X)(\omega)$$

对 $\mathbb{P}$-几乎处处的 $\omega\in\Omega$ 成立，其中 $\mathbf{1}_B$ 表示集合 B 的示性函数。众所周知，给定 X 时，Y 的正则条件分布存在，且在 $\mathbb{P}^X$-几乎处处意义下唯一（其中 $\mathbb{P}^X$ 表示 X 的分布，即 $\mathbb{P}$ 在 X 下的推前）。对于每个二元分布函数 H，相应的正则条件分布（即在 $(X,Y)\sim H$ 的情况下，给定 X 时 Y 的正则条件分布）记为 $K_H(\cdot,\cdot)$；相应地，$K_A(\cdot,\cdot)$ 表示该 Copula 的正则条件分布。下文中，我们将 $K_H(\cdot,\cdot)$ 和 $K_A(\cdot,\cdot)$ 简称为 H 和 A 的马尔可夫核（的版本）。

设 (X,Y) 的分布函数为 H。给定 $G\in\mathcal{B}(\mathbb{R}^2)$，并令 $G_x:=\{y\in\mathbb{I}:(x,y)\in G\}$ 表示对每个 $x\in\mathbb{R}$ 的 G 的 x-截断，则解体定理（参见 Kallenberg（2021））意味着

$$\int_{\mathbb{R}}K_H(x,G_x)\,\mathrm{d}\mathbb{P}^X(x)=\mathbb{P}^{(X,Y)}(G)\tag{2.1}$$

成立。特别地，对于 $A\in\mathcal{C}$ 和任意矩形 $G=E_1\times E_2$，满足 $E_1,\ E_2\in\mathcal{B}(\mathbb{I})$，我们有

$$\int_{E_1}K_A(x,E_2)\,\mathrm{d}\lambda(x)=\mu_A(E_1\times E_2).$$

再次假设 $(X,Y)\sim H$，其中 H 连续。令 $F,\ G$ 表示相应的边缘分布函数，A 表示其底层的（唯一）Copula。根据 Mroz 等人（2023），考虑 A 的马尔可夫核，并通过设定以下关系代入边缘分布函数

$$K_H(x,(-\infty,y]):=K_A\left(F(x),[0,G(y)]\right)\tag{2.2}$$

得到一个关于 H 的马尔可夫核 $K_H(\cdot,\cdot)$。我们将利用这种相互关系，把所谓的经验棋盘格／伯恩斯坦近似的收敛性推广到 Y 在给定 X 条件下的条件分布的收敛性。

尽管 Copula 的马尔可夫核仅在 λ-几乎处处意义下唯一，但很容易验证，对于每个 Copula A，至多存在一个马尔可夫核版本，使得函数 $(x,y)\mapsto K_A(x,[0,y])$ 在 $\mathbb{I}^2$ 上连续。若存在这样的核，我们就称其为 A 的（唯一）连续马尔可夫核，并称 A 具有连续马尔可夫核。若 μ_A 绝对连续，且密度为 k_A，则设定

$$K_A(x,[0,y]):=\int_{[0,y]}k_A(x,s)d\lambda(s),\quad x\in\mathbb{I}\tag{2.3}$$

显然定义了一个 A 的马尔可夫核。有关正则条件分布和解体的更多细节及性质，请参见 Kallenberg（2021）和 Klenke（2020）。

给定 Copula $A,A_1,A_2\ldots$，若对 λ-几乎处处的 $x\in\mathbb{I}$，序列 $(K_A(x,\cdot))_{n\in\mathbb{N}}$

上的概率测度序列 $\mathcal{B}(\mathbb{I})$ 弱收敛至概率测度 $K_A(x,\cdot)$。尽管在若干类 Copula 中，例如二元阿基米德 Copula 族或极值 copula 族，弱条件收敛等价于关于 d_∞ 的收敛（见 Kasper 等人（2021）），但一般而言，这两个概念并不等价；弱条件收敛蕴含关于 d_∞ 的收敛，反之则不然。本文的主要结果（定理 3.1 和定理 4.3 ）建立在以下更强的 Copula 收敛概念之上，我们将其称为一致条件收敛：

定义 2.1 *A sequence $A_n\ _{n\in\mathbb{N}}$ of copulas converges uniformly conditional to a $\cdot(),,,,$ copula A if there are versions $K_A\ K_{A_1}\ K_{A_2}\ \ldots$ of the Markov kernels such that*

$$\lim_{n\to\infty}\sup_{(x,y)\in\mathbb{I}^2}|K_{A_n}(x,[0,y])-K_A(x,[0,y])|=0.$$

接下来我们将使用所谓的棋盘近似和 Bernstein 近似。在介绍下一节的主要结果之前，我们先回顾它们的构造和性质。

2.1 棋盘逼近

设 $2\le N\in\mathbb{N}$。则 N-棋盘逼近 $\mathcal{CB}_N(A)\in\mathcal{C}$ of$A\in\mathcal{C}$ 按定义是一个绝对连续的 Copula，其（某个版本的）密度为 $k_{\mathcal{CB}_N(A)}$，定义如下（见 Li 等人（1998））：

$$\begin{aligned}k_{\mathcal{CB}_N(A)}(x,y)&=N^2\sum_{i,j=1}^{N}\mu_A\Big(Q_{i,j}^N\Big)\mathbf{1}_{Q_{i,j}^N}(x,y)\\&=N^2\sum_{i,j=0}^{N}A\big(\tfrac{i}{N},\tfrac{j}{N}\big)\Big(\mathbf{1}_{I_i^N}(x)-\mathbf{1}_{I_{i+1}^N}(x)\Big)\Big(\mathbf{1}_{I_j^N}(y)-\mathbf{1}_{I_{j+1}^N}(y)\Big)\end{aligned}\tag{2.4}$$

对于所有 $(x,y)\in\mathbb{I}^2$。下文中，遵循 Griessenberger 等人（2022）和 Junker 等人（2021）的做法，我们将 N 称为棋盘逼近的分辨率。根据式（2.3）对式（2.4）积分，可直接得到马尔可夫核 $K_{\mathcal{CB}_N(A)}(\cdot,\cdot)$，它对应于 $\mathcal{CB}_N(A)$，并将在下文中发挥作用：

$$\begin{aligned}K_{\mathcal{CB}_N(A)}(x,[0,y])&=N^2\sum_{i,j=1}^{N}\mu_A\Big(Q_{i,j}^N\Big)\mathbf{1}_{I_i^N}(x)\int_{[0,y]}\mathbf{1}_{I_j^N}\,\mathrm{d}\lambda\\&=N^2\sum_{i,j=0}^{N}A\big(\tfrac{i}{N},\tfrac{j}{N}\big)\Big(\mathbf{1}_{I_i^N}(x)-\mathbf{1}_{I_{i+1}^N}(x)\Big)\int_{[0,y]}\Big(\mathbf{1}_{I_j^N}-\mathbf{1}_{I_{j+1}^N}\Big)\,\mathrm{d}\lambda.\end{aligned}\tag{2.5}$$

使用本节开头提到的索引 $i_0^N(x)$ $j_0^N(x)$，式 (2.5) 中的最后一个表达式可进一步简化为

$$\begin{aligned}K_{\mathcal{CB}_N(A)}(x,[0,y])&=N(j_0^N(y)-yN)\Big[A\Big(\tfrac{i_0^N(x)}{N},\tfrac{j_0^N(y)-1}{N}\Big)-A\Big(\tfrac{i_0^N(x)-1}{N},\tfrac{j_0^N(y)-1}{N}\Big)\Big]\\&\quad+N(yN+1-j_0^N(y))\Big[A\Big(\tfrac{i_0^N(x)}{N},\tfrac{j_0^N(y)}{N}\Big)-A\Big(\tfrac{i_0^N(x)-1}{N},\tfrac{j_0^N(y)}{N}\Big)\Big],\end{aligned}\tag{2.6}$$

这意味着 $\frac{1}{N} K_{\mathcal{CB}_N(A)}(x,[0,y])$ 可以解释为这些差值的凸组合

$$A\Big(\tfrac{i_0^N(x)}{N}, \tfrac{j_0^N(y)-1}{N}\Big) - A\Big(\tfrac{i_0^N(x)-1}{N}, \tfrac{j_0^N(y)-1}{N}\Big), \quad A\Big(\tfrac{i_0^N(x)}{N}, \tfrac{j_0^N(y)}{N}\Big) - A\Big(\tfrac{i_0^N(x)-1}{N}, \tfrac{j_0^N(y)}{N}\Big).$$

注意，直接利用式 (2.6) 与三角不等式，即可对任意 $A_1, A_2 \in \mathcal{C}$ 得到如下不等式：

$$\big|K_{\mathcal{CB}_N(A_1)}(x,[0,y]) - K_{\mathcal{CB}_N(A_2)}(x,[0,y])\big| \le 2N d_\infty(A_1, A_2). \tag{2.7}$$

棋盘逼近的重要性尤其在于如下事实（由 Li 等人（1998）借助马尔可夫算子所确立）：对每一个 Copula $A \in \mathcal{C}$，序列 $(\mathcal{CB}_N(A))_{N\in\mathbb{N}}$ 对 $N \to \infty$ 弱条件收敛于 A（另见 Ansari 等人（2026）第 4 节）。在第 3 节中，我们将改进这一结果，并证明：若 A 允许一个连续马尔可夫核，则序列 $(\mathcal{CB}_N(A))_{N\in\mathbb{N}}$ 甚至一致条件收敛于 A。

2.2 伯恩斯坦逼近

伯恩斯坦 Copula（参见例如 Sancetta 和 Satchell（2004）；Segers 等人（2017）及其参考文献）是另一种基于单位分解的常用 Copula 近似方法，而单位分解在 Li 等人（1998）中于更一般的背景下得到研究。对于棋盘逼近，单位分解由 $\mathbf{1}_{I_1^N}, \ldots, \mathbf{1}_{I_N^N}$ 构成，而伯恩斯坦逼近则建立在伯恩斯坦多项式之上。回顾次数为 $N \in \mathbb{N}$ 的第 s 个伯恩斯坦多项式由下式给出

$$p_{N,s}(u) = \begin{cases} \binom{N}{s} u^s (1-u)^{N-s}, & s \in \{0,\ldots,N\}, \\ 0, & s \notin \{0,\ldots,N\} \end{cases}, \quad u \in \mathbb{I},$$

其中后一种情形仅是为了简化后文计算而加以考虑。对于$s \in \{1,\ldots,N\}$，容易验证其导数由下式给出

$$p'_{N,s}(u) = N(p_{N-1,s-1}(u) - p_{N-1,s}(u))$$

对所有 $u \in \mathbb{I}$。

下面的辅助引理表明，当略去 xN附近的指标时，我们可以对伯恩斯坦多项式之和给出界。

引理2.2 *Suppose that $u \in \mathbb{I}$ and $N \in \mathbb{N}$, consider $t \ge 0$ and set*

$$V := \{0,\ldots,\lfloor uN - t\rfloor, \lceil uN + t\rceil, \ldots, N\},$$

where $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ denote the floor and ceiling functions, respectively. Then the inequal-ity

$$\sum_{i\in V} p_{N,i}(u) \le 2\exp\Big(-\tfrac{2t^2}{N}\Big)$$

holds.

证明 令 X 表示服从二项分布Bin(N,u)的随机变量，我们显然有

$$\sum_{i\in V} p_{N,i}(u) = \mathbb{P}(|X - \mathbb{E}(X)| \ge t).$$

该不等式可直接由切尔诺夫-霍夫丁界（Hoeffding，1963）得出。□

对于每个 Copula $A \in \mathcal{C}$，其 N-伯恩斯坦近似 $\mathcal{B}_N(A)$ 定义为

$$\mathcal{B}_N(A)(x,y) := \sum_{i,j=1}^{N} A\big(\tfrac{i}{N}, \tfrac{j}{N}\big) p_{N,i}(x) p_{N,j}(y)$$

对所有 $x,\ y \in \mathbb{I}$。利用光滑性，并考虑关于 x 的一阶导数，$\mathcal{B}_N(A)$ 的连续马尔可夫核 $K_{B_N(A)}$ 的连续版本如下所示（第一行与第二行相等，显然可验证）

$$\begin{aligned} K_{\mathcal{B}_N(A)}(x,[0,y]) &:= N \sum_{i,j=1}^{N} A\big(\tfrac{i}{N}, \tfrac{j}{N}\big)\big(p_{N-1,i-1}(x) - p_{N-1,i}(x)\big) p_{N,j}(y) \\ &= N \sum_{i,j=1}^{N} \big(A\big(\tfrac{i}{N}, \tfrac{j}{N}\big) - A\big(\tfrac{i-1}{N}, \tfrac{j}{N}\big)\big) p_{N-1,i-1}(x) p_{N,j}(y) \end{aligned} \tag{2.8}$$

对所有 $(x,\ y) \in \mathbb{I}^2$。直接利用公式（2.8）的第一行或第二行以及三角不等式，可得以下不等式（这是对任意 $A_1,\ A_2 \in \mathcal{C}$ 的不等式（2.7）的 Bernstein 版本：

$$\big|K_{\mathcal{B}_N(A_1)}(x,[0,y]) - K_{\mathcal{B}_N(A_2)}(x,[0,y])\big| \le 2N d_\infty(A_1, A_2). \tag{2.9}$$

将 Li 等人（1998）基于马尔可夫算子的结果转用于马尔可夫核，可得与棋盘逼近类似的结论：对于每个 Copula A，序列 $(\mathcal{B}_N(A))_{N\in\mathbb{N}}$ 在 $N \to \infty$ 下条件弱收敛于 A。与第 3 节中的棋盘逼近一样，在第 4 节中我们将进一步完善这一结果，并证明若 A 允许存在连续马尔可夫核，则序列 $(\mathcal{B}_N(A))_{N\in\mathbb{N}}$ 在 A 下甚至条件一致收敛。

3 经验棋盘逼近的一致条件收敛

本节将证明，对于适当选取分辨率 $N = N(n)$ 的经验 Copula E_n，其棋盘逼近 $\mathcal{CB}_{N(n)}(E_n)$ 与经验边缘分布共同构成条件分布的强相合估计量。为使表述简洁，以下将 $\mathcal{CB}_{N(n)}(E_n)$ 称为分辨率为 $N(n)$ 的 *empirical checkerboard approximation*。我们首先陈述本节关于经验棋盘格 Copula 一致条件收敛的主要定理，然后分几个步骤证明该结果。回顾一下，当且仅当存在一个（版本的）马尔可夫核，使函数 $(x,\ y) \mapsto K_A(x,\ [0,y])$ 在 $\mathbb{I}^2$ 上连续时，我们称 A 具有连续马尔可夫核。

定理 3。1 *Suppose that* $(X_1,\ Y_1),\ (X_2,\ Y_2)\ \ldots$ *is a random sample from* $(X,\ Y)$ *and assume that* $(X,\ Y)$ *has continuous distribution function* H *, margins* F *and* G*, and underlying (unique) copula* A*. Furthermore suppose that* A *allows a continuous Markov kernel* K_A *and define* K_H *according to eq.* 2 2 *. Setting* $N\ n\ = \lfloor n^s \rfloor$ *for* $(.)():$

some $s \in \big(0,\ \tfrac{1}{2}\big),\ \ $ *we have that*

$$\lim_{n\to\infty} \sup_{(x,y)\in\mathbb{R}^2} \Big| K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, G_n(y)]) - K_H(x, (-\infty, y])\Big| = 0 \tag{3.1}$$

almost surely.

为了证明定理 3.1，我们首先推导棋盘逼近的一致条件收敛，然后代入经验联结函数，得到经验棋盘的一致条件收敛（$\mathcal{CB}_{N(n)}(A)$）$_{n\in\mathbb{N}}$，最后通过公式 (2.2) 将收敛性推广到 $K_H(\cdot,\ \cdot)$。

引理 3.2 *Suppose that $A \in \mathcal{C}$ allows a continuous Markov kernel K_A. Then $(\mathcal{CB}_N(A))_{N\in\mathbb{N}}$ converges uniformly conditional to A for $N \to \infty$, i.e.,*

$$\lim_{N\to\infty} \sup_{(x,y)\in\mathbb{I}^2} \left|K_{\mathcal{CB}_N(A)}(x,[0,y]) - K_A(x,[0,y])\right| = 0. \tag{3.2}$$

证明 任取并固定 $\varepsilon > 0$。由函数 $(x,y) \mapsto K_A(x,[0,y])$ 在 $\mathbb{I}^2$ 上的一致连续性，存在某个 $\delta > 0$，使得对于每个 $(x_1,y_1),(x_2,y_2) \in \mathbb{I}^2$，以下蕴含关系成立：

$$|x_1 - x_2| + |y_1 - y_2| < \delta \implies |K_A(x_1,[0,y_1]) - K_A(x_2,[0,y_2])| < \tfrac{\varepsilon}{2}$$

取 N_0 足够大，使得对于每个 $N \geq N_0$，都有 $2/N < \delta$。考虑 $N \geq N_0$，并任取且固定 $(x,y) \in \mathbb{I}^2$。

我们首先推导式 (2.6) 的一个替代形式。根据中值定理，可找到 $\xi_1,\xi_2 \in \left(\frac{i_0^N(x)-1}{N}, \frac{i_0^N(x)}{N}\right)$，使得

$$\begin{aligned}
\tfrac{1}{N}K_A\Big(\xi_1, \Big[0, \tfrac{j_0^N(y)-1}{N}\Big]\Big) &= A\Big(\tfrac{i_0^N(x)}{N}, \tfrac{j_0^N(y)-1}{N}\Big) - A\Big(\tfrac{i_0^N(x)-1}{N}, \tfrac{j_0^N(y)-1}{N}\Big)\\
\tfrac{1}{N}K_A\Big(\xi_2, \Big[0, \tfrac{j_0^N(y)}{N}\Big]\Big) &= A\Big(\tfrac{i_0^N(x)}{N}, \tfrac{j_0^N(y)}{N}\Big) - A\Big(\tfrac{i_0^N(x)-1}{N}, \tfrac{j_0^N(y)}{N}\Big)
\end{aligned}$$

成立。直接合并这两个表达式可得

$$\begin{aligned}
K_{\mathcal{CB}_N(A)}(x,[0,y]) = (j_0^N(y) - yN)&K_A\Big(\xi_1, \Big[0, \tfrac{j_0^N(y)-1}{N}\Big]\Big)\\
&+ (yN + 1 - j_0^N(y))K_A\Big(\xi_2, \Big[0, \tfrac{j_0^N(y)}{N}\Big]\Big).
\end{aligned}$$

根据 $\Delta_N := \left|K_{\mathcal{CB}_N(A)}(x,\ [0,y]) - K_A(x,\ [0,y])\right|$ 的三角不等式，我们因此得到

$$\begin{aligned}
\Delta_N \leq (j_0^N(y) - yN)&\Big|K_A\Big(\xi_1, \Big[0, \tfrac{j_0^N(y)-1}{N}\Big]\Big) - K_A(x,[0,y])\Big|\\
&+ (yN + 1 - j_0^N(y))\Big|K_A\Big(\xi_2, \Big[0, \tfrac{j_0^N(y)}{N}\Big]\Big) - K_A(x,[0,y])\Big|.
\end{aligned}$$

考虑 $|\xi_1 - x| + \left|\frac{j_0^N(y)-1}{N} - y\right| < \delta$ 和 $|\xi_2 - x| + \left|\frac{j_0^N(y)}{N} - y\right| < \delta$，这意味着

$$\Delta_N = \left|K_{\mathcal{CB}_N(A)}(x,[0,y]) - K_A(x,[0,y])\right| < \varepsilon.$$

由于 $(x,\ y)$ 是任意的，且 N_0 仅取决于 ε，证明完成。 □

下一个引理表明，若 A 允许连续马尔可夫核，则经验棋盘逼近 $\mathcal{CB}_{N(n)}(E_n)$一致收敛于A。

引理 3。3 *Suppose that* $(X_1,\ Y_1),\ (X_2,\ Y_2)\ \ldots$ *is a random sample from* $(X,\ Y)$ *and assume that* $X\ Y$ *has continuous distribution function* H *marginals* F *and* G *and*$(,\ )\ ,\ \ ,\ (\ ):(\ )$ *underlying (unique) copula* A*. Furthermore suppose that* A *allows a continuous Markov kernel* K_A*. Setting* $N\ n\ \ =\lfloor n^s\rfloor$ *for some* $s\in\ 0,\frac{1}{2}\ \ $ *we have*

$$\lim_{n\to\infty}\sup_{(x,y)\in\mathbb{I}^2}\left|K_{\mathcal{CB}_{N(n)}(E_n)}(x,[0,y])-K_A(x,[0,y])\right|=0 \tag{3.3}$$

almost surely. In other words,

$$\textit{the sequence}\ \ \mathcal{CB}_{N(n)}\ E_n\ \ _{n\in\mathbb{N}}\ \textit{converges uniformly}\ (\ 0)$$

conditional to A *almost surely.*

证明 设定 $\Delta_n(x,\ y):=\left|K_{\mathcal{CB}_{N(n)}(E_n)}(x,\ [0,y])\ -K_A(x,\ [0,y])\right|$，应用三角不等式可得

$$\begin{aligned}\sup_{(x,y)\in\mathbb{I}^2}\Delta_n(x,y)\le&\sup_{(x,y)\in\mathbb{I}^2}\left|K_{\mathcal{CB}_{N(n)}(E_n)}(x,[0,y])-K_{\mathcal{CB}_{N(n)}(A)}(x,[0,y])\right|\\&+\sup_{(x,y)\in\mathbb{I}^2}\left|K_{\mathcal{CB}_{N(n)}(A)}(x,[0,y])-K_A(x,[0,y])\right|\end{aligned}$$

根据引理 3.2，对于 $n\to\infty$，后一个加项收敛于零，因此只需证明前一个加项以概率为一收敛于零。直接使用不等式（2.7）可得

$$\left|K_{\mathcal{CB}_{N(n)}(E_n)}(x,[0,y])-K_{\mathcal{CB}_{N(n)}(A)}(x,[0,y])\right|\le 2N(n)d_\infty(E_n,A)$$

对所有 $(x,\ y)\in\mathbb{I}^2$ 均成立。由于上界不依赖于 $(x,\ y)$，且 $(x,\ y)\in\mathbb{I}^2$ 是任意的，我们已证明

$$\sup_{(x,y)\in\mathbb{I}^2}\left|K_{\mathcal{CB}_{N(n)}(E_n)}(x,[0,y])-K_{\mathcal{CB}_{N(n)}(A)}(x,[0,y])\right|\le 2N(n)d_\infty(E_n,A).$$

根据我们对 s 的选择，以及 Janssen 等人（2012）中的事实，所需结果现即可得出

$$d_\infty(E_n,A)=O\left(\sqrt{\frac{\log(\log(n))}{n}}\right)$$

对 $n\to\infty$，概率为一 . □

现在我们已准备好所有工具，可以证明定理 3.1。

定理 3.1 的证明 设定

$$\Delta_n(x,y):=\sup_{(x,y)\in\mathbb{R}^2}\left|K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x),[0,G_n(y)])-K_H(x,(-\infty,y])\right|$$

并直接利用式 (2.2) 和三角不等式即可得到

$$\begin{aligned}\Delta_n(x,y)=&\sup_{(x,y)\in\mathbb{R}^2}\left|K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x),[0,G_n(y)])-K_A(F(x),[0,G(y)])\right|\\\le&\sup_{(x,y)\in\mathbb{R}^2}\left|K_{\mathcal{CB}_{N(E_n)}}(F_n(x),[0,G_n(y)])-K_A(F_n(x),[0,G_n(y)])\right|\\&+\sup_{(x,y)\in\mathbb{R}^2}\left|K_A(F_n(x),[0,G_n(y)])-K_A(F(x),[0,G(y)])\right|.\end{aligned}$$

根据引理 3.3，前一项几乎处处收敛到零，当 $n \to \infty$ 时。此外，利用函数 $(x,\ y) \mapsto K_A$ $(x,\ [0,y])$ 在 $\mathbb{I}^2$ 上的一致连续性以及 Glivenko-Cantelli 定理，后一项对 $n \to \infty$ 也几乎处处收敛到零，故证明完毕。□

注 3.4 据作者所知，经验棋盘逼近 $\mathcal{CB}_{N(n)}(E_n)$ 的渐近性质仍属未知。预印本 Choudhury 等人（2026）似乎是首篇探讨渐近理论的研究，但其前提假设较强：底层 Copula C 的三阶偏导数有界（这远比要求 C 允许一个连续马尔可夫核更加严格）。

在本节余下部分，我们将证明，对于不允许存在连续马尔可夫核的 Copula，即使是棋盘逼近 $\mathcal{CB}_N(A)$，也无法按照引理 3.2 的意义收敛。在证明这一结果之前，我们先给出一个简单例子，说明可能出现的问题。

例 3.5 对于最小 Copula M，马尔可夫核 $K_M(\cdot,\ \cdot)$ 的一个版本为

$$K_M(x,[0,y]) = \mathbf{1}_{[0,y]}(x), \quad (x,y) \in \mathbb{I}^2.$$

显然，映射 $(x,\ y) \mapsto K_M(x,\ [0,y])$ 在所有形如 $(x,\ x)$ 且满足 $x \in \mathbb{I}$ 的点处不连续。根据式 (2.5)，M 的 N-棋盘逼近 $\mathcal{CB}_{N(n)}(M)$ 的一个马尔可夫核为

$$K_{\mathcal{CB}_{N(n)}(M)}(x,[0,y]) = \begin{cases} 0, & \text{if } y \le \frac{i_0^N(x)-1}{N} \\ yN - i_0^N(x) + 1, & \text{if } y \in \left(\frac{i_0^N(x)-1}{N}, \frac{i_0^N(x)}{N}\right) \\ 1, & \text{if } y \ge \frac{i_0^N(x)}{N} \end{cases}$$

对于每个 $(x,\ y) \in \mathbb{I}^2$。x 考虑到 $K_M(x,\ \{x\}) = 1$ 对每个 $x \in \mathbb{I}$ 都成立，立即可得 $K_M($ $K_{\mathcal{CB}_{N(n)}\ M}\ x\ [0,y]$ y_x $K_{\mathcal{CB}_N\ M}\ x\ [0,y_x] = \frac{1}{2}\cdot,\ [0,y]) \in \{0,1\}$。

另一方面，$_0(,\ )$ 的连续性意味着存在某个，使得 $_0(,\ )$。综上可知，对于每个 $x \in \mathbb{I}$，我们甚至有

$$\sup_{y\in\mathbb{I}} \left| K_M(x,[0,y]) - K_{\mathcal{CB}_N(M)}(x,[0,y]) \right| \ge \tfrac{1}{2}.$$

现在我们转向不允许存在连续马尔可夫核的 Copula 一般情形，并证明一致条件收敛不可能成立。

命题 3.6 *Let $A \in \mathcal{C}$ be a copula and K_A be a version of the Markov kernel, such that the function $x\ y \mapsto K_A\ x\ [0,y]$ the sequence $(,\ )\ (,\ )$ has a discontinuity. Then, $(\Delta_N)_{N\in\mathbb{N}}$, given by*

$$\Delta_N := \sup_{(x,y)\in\mathbb{I}^2} \left| K_{\mathcal{CB}_N(A)}(x,[0,y]) - K_A(x,[0,y]) \right|$$

does not converge to 0 for $N \to \infty$.

$x_0\ y_0$ $x\ y \mapsto$

证明 假设 $(K_A\ x\ [0,y]$ $\varepsilon_0 > 0$ $\delta > 0$ $,) \in \mathbb{I}^2$ 是函数 $(,)\ (,)$ 的一个不连续点。那么存在某个值，使得对于任意给定的值，我们都能找到某个 $(x_\delta, y_\delta) \in B_\delta(x_0,y_0)$，满足

$$|K_A(x_\delta,[0,y_\delta]) - K_A(x_0,[0,y_0])| \ge \varepsilon_0.$$

取某个 $N \in \mathbb{N}$，使得 $x_0, y_0 \notin \left\{ \frac{1}{N}, \ldots, \frac{N-1}{N} \right\}$。则 (x_0, y_0) 是 $Q^N_{i_0^N\ x_0\ \ j_0^N\ y_0}$ $\delta_N > 0$ — —— $B_{\delta_N}\ x_0\ y_0\ \subset$ $Q^N_{i_0^N\ x_0\ \ j_0^N\ y_0}$ ()()的内点，因此，取充分小的值，显然有 (,) ()()。

函数 $f\colon (x,y) \mapsto K_{\mathcal{CB}_N(A)}(x,[0,y])$ 在 $Q^N_{i_0^N(x_0)\ j_0^N(y_0)}$ 的内部是利普希茨连续的（事实上，对给定的 y，它关于 x 是常数；对给定的 x，它关于 y 是仿射的）。特别地，f 在 $B_{\delta'_N}(x_0, y_0)$ 上也是利普希茨连续的，因此存在某个 $\delta'_N \in (0, \delta_N)$，使得对于每个满足 $|x - x_0| + |y - y_0| < \delta'_N$ 的 $(x, y) \in \mathbb{I}^2$，不等式

$$\left|K_{\mathcal{CB}_N(A)}(x,[0,y]) - K_{\mathcal{CB}_N(A)}(x_0[0,y_0])\right| < \tfrac{\varepsilon_0}{2}$$

成立。利用反三角不等式（在最后一步），我们得出

$$\begin{aligned}
2\Delta_N &\geq 2 \sup_{(x,y)\in B_{\delta'_N}(x_0,y_0)} \left|K_{\mathcal{CB}_N(A)}(x,[0,y]) - K_A(x,[0,y])\right| \\
&\geq \left|K_{\mathcal{CB}_N(A)}\left(x_{\delta'_N},[0,y_{\delta'_N}]\right) - K_A\left(x_{\delta'_N},[0,y_{\delta'_N}]\right)\right| \\
&\quad + \left|K_{\mathcal{CB}_N(A)}(x_0,[0,y_0]) - K_A(x_0,[0,y_0])\right| \\
&\geq \left|K_A\left(x_{\delta'_N},[0,y_{\delta'_N}]\right) - K_A(x_0,[0,y_0])\right| \\
&\quad - \left|K_{\mathcal{CB}_N(A)}\left(x_{\delta'_N},[0,y_{\delta'_N}]\right) - K_{\mathcal{CB}_N(A)}(x_0,[0,y_0])\right| \\
&> \tfrac{\varepsilon_0}{2}.
\end{aligned}$$

换句话说，我们已经证明

$$\sup_{(x,y)\in\mathbb{I}^2} \left|K_{\mathcal{CB}_N(A)}(x,[0,y]) - K_A(x,[0,y])\right| \geq \varepsilon_0/4$$

对每个满足 $x_0, y_0 \notin \left\{ \frac{1}{N}, \ldots, \frac{N-1}{N} \right\}$ 的 $N \in \mathbb{N}$ 都成立。由于这样的 N 有无穷多个，证毕。□

4 经验伯恩斯坦逼近的一致条件收敛

与经验棋盘逼近类似，下文将经验联结函数

$\mathcal{B}_N\ E_n$ *empirical Bernstein* ( )

approximation 的伯恩斯坦近似称为 。简而言之，本节将证明，为经验棋盘逼近推导出的结果同样适用于经验伯恩斯坦逼近（选择适当的次数 $N = N(s)$）。

引理 4.1 *Suppose $A \in \mathcal{C}$ allows a continuous Markov kernel. Then, we have that*

$$\lim_{N\to\infty} \sup_{(x,y)\in\mathbb{I}^2} \left|K_{\mathcal{B}_N(A)}(x,[0,y]) - K_A(x,[0,y])\right| = 0,$$

i.e., $(\mathcal{B}_N(A))_{N\in\mathbb{N}}$ converges uniformly conditional to A.

证明 任取但固定 $\varepsilon > 0$。由函数 $(x,\ y) \mapsto K_A(x,\ [0,y])$ 在 $\mathbb{I}^2$ 上的一致连续性可知，存在某个 $\delta > 0$，使得对于每个 $(x_1,\ y_1)$，$(x_2,\ y_2) \in \mathbb{I}^2$，以下蕴含关系成立：

$$|x_1 - x_2| + |y_1 - y_2| < \delta \implies |K_A(x_1,[0,y_1]) - K_A(x_2,[0,y_2])| < \tfrac{\varepsilon}{2}$$

选择足够大的 N_0，使得

$$2\exp\left(-2N^{1-2r}\right) < \tfrac{\varepsilon}{2} \quad \text{and} \quad \tfrac{3+2N^{1-r}}{N} < \tfrac{\delta}{2}$$

对每个 $N \geq N_0$均成立，因此 $r \in \left(0, \tfrac{1}{2}\right)$。
根据中值定理，对于每个 $i \in \{1, \ldots, N\}$和 $j \in \{0, \ldots, N\}$，我们可以找到某个 $\xi_{i,j} \in \left(\tfrac{i-1}{N}, \tfrac{i}{N}\right)$，使得

$$\tfrac{1}{N} K_A\left(\xi_{i,j}, \left[0, \tfrac{j}{N}\right]\right) = A\left(\tfrac{i}{N}, \tfrac{j}{N}\right) - A\left(\tfrac{i-1}{N}, \tfrac{j}{N}\right).$$

这表明公式（2.8）中的最后一行可以改写为

$$K_{\mathcal{B}_N(A)}(x, [0, y]) = \sum_{i,j=1}^{N} K_A\left(\xi_{i,j}, \left[0, \tfrac{j}{N}\right]\right) p_{N-1,i-1}(x) p_{N,j}(y).$$

对每个 $(x, y) \in \mathbb{I}^2$，令 $\Delta_N(x, y) := \left|K_{\mathcal{B}_N(A)}(x, [0, y]) - K_A(x, [0, y])\right|$，并利用伯恩斯坦多项式构成单位分解这一事实，我们有

$$\Delta_N(x, y) \quad \leq \quad \sum_{i=1}^{N} \sum_{j=1}^{N} p_{N-1,i-1}(x) p_{N,j}(y) \left|K_A\left(\xi_{i,j}, \left[0, \tfrac{j}{N}\right]\right) - K_A(x, [0, y])\right|. \quad (4.1)$$

定义集合 $U_x, V_y \subset \{0, \ldots, N\}$如下

$$U_x := \{0, \ldots, \lfloor x(N-1) - \Delta_N \rfloor, \lceil x(N-1) + \Delta_N \rceil, \ldots, N-1\}$$
$$V_y := \{0, \ldots, \lfloor yN - \Delta_N \rfloor, \lceil yN + \Delta_N \rceil, \ldots, N\}.$$

公式（4.1）的右端可以表示为

$$\left|K_{\mathcal{B}_N(A)}(x, [0, y]) - K_A(x, [0, y])\right| \leq S_1(x, y) + S_2(x, y),$$

with

$$S_1(x, y) = \sum_{i \in U_x \vee j \in V_y} p_{N-1,i-1}(x) p_{N,j}(y) \left|K_A\left(\xi_{i,j}, \left[0, \tfrac{j}{N}\right]\right) - K_A(x, [0, y])\right|$$

$$S_2(x, y) = \sum_{i \notin U_x, j \notin V_y} p_{N-1,i-1}(x) p_{N,j}(y) \left|K_A\left(\xi_{i,j}, \left[0, \tfrac{j}{N}\right]\right) - K_A(x, [0, y])\right|.$$

由引理 2.2，我们有 $S_1(x, y) \leq 2\exp\left(-2N^{1-2r}\right) < \tfrac{\varepsilon}{2}$，对每一个 $(x, y) \in \mathbb{I}^2$ 以及 $N \geq N_0$ 成立。$S_2(x, y)$ 中的求和仅包含满足 $|\xi_{i,j} - x| + \left|y - \tfrac{j}{N}\right| < \delta$ 的指标，从而蕴含 $S_2(x, y) < \tfrac{\varepsilon}{2}$。综合起来，这表明对每一个 $N \geq N_0$，我们有

$$\left|K_{\mathcal{B}_N(A)}(x, [0, y]) - K_A(x, [0, y]\right| < \varepsilon.$$

由于 (x, y) 是任意的，且 N_0 仅依赖于 ε，证明完成。 □

下一个引理是引理 3.3 的伯恩斯坦版本。我们省略其证明，因为它与引理 3.3 的证明类似，只是用不等式（2.9）代替不等式（2.7）。

引理 4.2 *Let $A \in \mathcal{C}$ be a copula allowing a continuous Markov kernel K_A. Setting $N(n) := \lfloor n^s \rfloor$ for some $s \in \left(0, \tfrac{1}{2}\right)$, we have that*

$$\lim_{n \to \infty} \sup_{(x,y) \in \mathbb{I}^2} \left|K_{\mathcal{B}_{N(n)}(E_n)}(x, [0, y]) - K_A(x, [0, y])\right| = 0$$

almost surely. In other words: the sequence $\left(\mathcal{B}_{N(n)}\ E_n\right)_{n \in \mathbb{N}}$ converges uniformly conditional to A almost surely.

我们以陈述本节的主要结果作为结尾，其证明与定理 3.1 的证明类似。

定理 4。3 *Suppose that* $(X_1, Y_1), (X_2, Y_2) \ldots$ *is a random sample from* (X, Y) *and assume that* (X, Y) *has continuous distribution function* H*, marginals* F *and* G*, and underlying (unique) copula* A*. Furthermore suppose that* A *allows a continuous Markov kernel* K_A *and define* K_H *according to eq.* 2 2 *. Setting* $N\ n\ = \lfloor n^s \rfloor$ *for* $(.)():$

some $s \in \left(0, \frac{1}{2}\right)$*, we have that*

$$\lim_{n\to\infty} \sup_{(x,y)\in\mathbb{R}^2} \left| K_{\mathcal{B}_{N(n)}(E_n)}(F_n(x), [0, G_n(y)]) - K_H(x, (-\infty, y]) \right| = 0 \tag{4.2}$$

almost surely.

考虑到对于每个 Copula A和每个 $N \in \mathbb{N}$，函数 $(K_{\mathcal{B}_N A}\ x\ [0,y]\ x\ y \mapsto ,)_{()}(,)$ 是连续的，下面的命题可立即得出。

命题 4。4 *Let* $A \in \mathcal{C}$ *be a copula and* K_A *be a version of the Markov kernel, such that the function* $x\ y \mapsto K_A\ x\ [0,y]$ *the sequence* $(,)(,)$ *has a discontinuity. Then,* $(\Delta_N)_{N\in\mathbb{N}}$*, given by*

$$\Delta_N := \sup_{(x,y)\in\mathbb{I}^2} \left| K_{\mathcal{B}_N(A)}(x, [0, y]) - K_A(x, [0, y]) \right|$$

does not converge to 0 *for* $N \to \infty$*.*

5 对非参数回归的影响

在较弱的正则性条件下，即 Copula A 允许马尔可夫核具有连续版本时，定理 3.1 不仅保证估计的条件分布函数弱收敛，甚至保证其一致收敛至真实的条件分布函数。不出所料，这会直接影响条件分布的泛函，尤其包括均值回归、分位数回归和期望分位数回归，以及条件方差等。下文将重点讨论非参数回归；所得的均值回归和分位数回归结果，将 Kaiser 和 Trutschnig（2025）中的二元结果推广到了广义分布情形。为此，我们仅使用经验棋盘逼近；不过，利用定理 4.3，所有结果同样适用于经验伯恩斯坦逼近。

在本节余下部分，我们考虑一对具有连续分布函数 H 的随机变量（X，Y），并令 K_H 表示根据式（2.2）定义的马尔可夫核。

5.1 均值回归

回想一下，条件期望可以通过对条件分布函数进行积分来计算。事实上，假设 Z 服从分布 $K_H(x, \cdot)$ 且可积，则有

$$r_H(x) := \mathbb{E}(Y|X = x) = \int_{(0,\infty)} 1 - K_H(x, (-\infty, t])\, \mathrm{d}\lambda(t) - \int_{(-\infty,0)} K_H(x, (-\infty, t]\, \mathrm{d}\lambda(t).$$

定理 3.1 表明，只要可以交换极限与积分，我们便能直接一致地估计回归函数。以下技术条件保证这种交换有效。

假设 5 1 *For $x \in \mathbb{R}$ let the random variable $Z_n^{(x)}$ have distribution function。 $K_{\mathcal{CB}_{N(n)}\ E_n}\ F_n\ x\ \ [0, G_n\ \cdot\ ]$. We say that the regularity assumption holds in a point $_{()}(\ (\ )\ ())x \in \mathbb{R}$ if the sequence $(Z_n^{(x)})_{n\in\mathbb{N}}$ is uniformly integrable.*

下面的注释表明，一致可积性的条件并不太严格，例如，在一致有界的情况下也成立。

注5.2 假设 Z、 Z_1、 Z_2.....是随机变量，且对每个 $n \in \mathbb{N}$都有 $|Z_n| \le Z$几乎处处成立。那么，如果 Z可积，显然序列$(Z_n)_{n\in\mathbb{N}}$一致可积。回到我们的设定，考虑具有连续分布函数 H的(X, Y)。如果随机变量 Y 有界，那么$(Z_n^{(x)})_{n\in\mathbb{N}}$ 一致可积，证明如下：假设 $|Y| \le a$几乎处处成立，并且(X_1, Y_1)、(X_2, Y_2).....是来自(X, Y)的随机样本。那么显然，对每个$n \in \mathbb{N}$，几乎处处都有 $G_n(a) = G(a) = 1$ 以及 $G_n(-a) = G(-a) = 0$ ，这意味着对每个 $x \in \mathbb{R}$

$$K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, G_n(a)]) = K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, 1]) = 1$$
$$K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, G_n(-a)]) = K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), \{0\}) = 0$$

几乎处处成立。直接推论是，对于每个 $x \in \mathbb{R}$ ，序列 $(Z_n^{(x)})_{n\in\mathbb{N}}$ 由上界 a 和下界 $-a$所限制，因此一致可积。

定理 5.3 *Suppose that $(X_1, Y_1), (X_2, Y_2)\ \ldots$ is a random sample from (X, Y) and assume that (X, Y) has continuous distribution function H marginals F and G and underlying (unique) copula A. Furthermore suppose that A allows a continuous Markov kernel K_A define K_H according to eq. 2 2 and set $N\ n$,(.) $= \lfloor n^s \rfloor$ for some ( ) :$s \in \left(0\ \ \frac{1}{2}\right)$. Then for every $x \in \mathbb{R}$ in which the regularity assumption holds we have ,*

$$r_n(x) := r_n^+(x) - r_n^-(x) \to \mathbb{E}\left(Y \mid X = x\right)$$

almost surely as $n \to \infty$, where

$$r_n^+(x) := \int_{(0,\infty)} 1 - K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, G_n(y)])\, \mathrm{d}\lambda(y),$$
$$r_n^-(x) := \int_{(-\infty,0)} K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, G_n(y)])\, \mathrm{d}\lambda(y).$$

证明 由定理 3.1 以及一致可积性允许交换极限与积分这一事实可直接得出。□

由于经验分布函数 F_n, G_n 是阶梯函数，定理5.3中的两个积分归结为求和，正如下述评注所示。

注5.4 令 $Y_{(k)}$ 表示第 k 个次序统计量，并考虑 $Y_{(k)+} := \max\{0, Y_{(k)}\}$和 $Y_{(k)-} := \max\{0, -Y_{(k)}\}$。那么，定理5.3中的两个积分计算结果为

$$\begin{aligned} r_n^+(x) &= \sum_{i=1}^{n-1}\left(Y_+^{(i+1)} - Y_+^{(i)}\right)\left(1 - K_{\mathcal{CB}_{N(n)}(E_n)}\left(F_n(x), \left[0, G_n\left(Y^{(i)}\right)\right]\right)\right) \\ &= \sum_{i=1}^{n-1}\left(Y_+^{(i+1)} - Y_+^{(i)}\right)\left(1 - K_{\mathcal{CB}_{N(n)}(E_n)}\left(F_n(x), \left[0, \tfrac{i}{n}\right]\right)\right) \end{aligned}$$

以及

$$r_n^-(x) = \sum_{i=1}^{n-1}\left(Y_-^{(i)} - Y_-^{(i+1)}\right)\left(K_{\mathcal{CB}_{N(n)}(E_n)}\left(F_n(x), \left[0, G_n\left(Y^{(i)}\right)\right]\right)\right)$$
$$= \sum_{i=1}^{n-1}\left(Y_-^{(i)} - Y_-^{(i+1)}\right)\left(K_{\mathcal{CB}_{N(n)}(E_n)}\left(F_n(x), \left[0, \tfrac{i}{n}\right]\right)\right)$$

考虑到 r_n 是一个至多取 $N(n)$ 个不同值的阶梯函数，这使我们能够非常高效地计算估计的均值回归。事实上，我们可以以 $\mathcal{O}(N(n)n)$ 的时间计算出 $N(n)$ 个不同的值。因此，计算 $m \in \mathbb{N}$ 个点所需的时间复杂度为 $\mathcal{O}(m + N(n)n)$。其他标准且广为人知的非参数均值回归方法，例如 Nadaraya-Watson 估计量，计算 m 个点需要 $\mathcal{O}(mn)$。这意味着在 $m \geq N(n)$ 的情形下，我们所提出的方法在计算上要快得多。

5.2 分位数回归

接下来，我们关注分位数回归，并利用经验棋盘近似来非参数地估计分位数回归函数。对于 $x \in \mathbb{R}$ 以及分位数水平 $\tau \in (0,1)$，按照 (Jordan 等人, 2022) 将一元分布函数 F 的（区间值） τ-分位数定义为

$$\left[\underline{q}^\tau, \overline{q}^\tau\right] := [\sup\{y \in \mathbb{R} : F(y) < \tau\}, \inf\{y \in \mathbb{R} : F(y) > \tau\}]. \tag{5.1}$$

注意到，$\left[\underline{q}^\tau, \overline{q}^\tau\right]$ 是一个长度为正的区间，当且仅当存在点 $y_1 < y_2$，使得对于每个 $y \in (y_1, y_2)$，$F(y) = \tau$。下面这个简单引理将在后文中用到：

引理 5.5 *Suppose that F, F_1, F_2, ... are univariate distribution functions, such that $(F_n)_{n\in\mathbb{N}}$ converges weakly to F and denote the corresponding τ-quantiles by $\left[\underline{q}^\tau, \overline{q}^\tau\right]$, $\left[\underline{q}_1^\tau, \overline{q}_1^\tau\right]$, $\left[\underline{q}_2^\tau, \overline{q}_2^\tau\right]$ .... Then we have*

$$\lim_{n\to\infty} \max_{y_n \in \left[\underline{q}_n^\tau, \overline{q}_n^\tau\right]} \min_{y \in \left[\underline{q}^\tau, \overline{q}^\tau\right]} |y_n - y| = 0.$$

证明 令 $\varepsilon > 0$ 为任意但固定的值。根据 τ-分位数的定义，显然有 $F\left(\underline{q}^\tau - \frac{\varepsilon}{2}\right) < \tau < F\left(\overline{q}^\tau + \frac{\varepsilon}{2}\right)$。利用弱收敛等价于在 $\tau < F_n\ \overline{q}^\tau + \varepsilon$ $n \geq n_0$ $\mathbb{R}$ $n_0 \in \mathbb{N}$ $F_n\left(\underline{q}^\tau - \varepsilon\right) <$ 的稠密子集上收敛这一事实，我们可以找到某个，使得对于每个，（）成立。这表明

$$\underline{q}^\tau - \varepsilon \leq \underline{q}_n^\tau \leq \overline{q}_n^\tau \leq \overline{q}^\tau + \varepsilon,$$

以及

$$\max_{y_n \in \left[\underline{q}_n^\tau, \overline{q}_n^\tau\right]} \min_{y \in \left[\underline{q}^\tau, \overline{q}^\tau\right]} |y_n - y| \leq \varepsilon$$

对于每个 $n \geq n_0$。由于 ε 是任意的，证明完成。 □

现在我们可以证明，经验棋盘逼近直接导出了 τ 分位数回归函数 $x \mapsto Q_H^\tau(x) = \left[\underline{q}^\tau(x), \overline{q}^\tau(x)\right]$ 的一个相合估计量，其形式为

$$\underline{q}^\tau(x) := \sup\{y \in \mathbb{R} : K_H(x, (-\infty, y]) < \tau\}$$
$$\overline{q}^\tau(x) := \inf\{y \in \mathbb{R} : K_H(x, (-\infty, y]) > \tau\}, \tag{5.2}$$

其中 K_H 仍按式 (2.2) 定义。为简化符号，我们将写作$Q_n^\tau(x) := \left[q_n^\tau(x)\ q_n^\tau(x)\right]$，其中

$$\underline{q}_n^\tau(x) := \sup\Big\{y \in \mathbb{R} : K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, G_n(y)]) < \tau\Big\}$$
$$\overline{q}_n^\tau(x) := \inf\Big\{y \in \mathbb{R} : K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, G_n(y)]) > \tau\Big\}. \tag{5.3}$$

结合定理 3.1 和引理 5.5 可直接得到以下结果:

定理 5.6 *Suppose that* $(X_1, Y_1), (X_2, Y_2) \ldots$ *is a random sample from* (X, Y) *and assume that* (X, Y) *has continuous distribution function* H*, marginals* F *and* G*, and underlying (unique) copula* A*. Furthermore suppose that* A *allows a continuous Markov kernel* K_A *define* K_H *according to eq.* (2.2) *and set* $N(n) = \lfloor n^s \rfloor$ *for some* $s \in \left(0, \frac{1}{2}\right)$*. Then for every* $x \in \mathbb{R}$ *the identity*

$$\lim_{n\to\infty} \max_{y_n \in \left[\underline{q}_n^\tau(x), \overline{q}_n^\tau(x)\right]} \min_{y \in \left[\underline{q}^\tau(x), \overline{q}^\tau(x)\right]} |y_n - y| = 0$$

holds almost surely.

5.3 期望回归

作为定理 3.1 的第三个也是最后一个推论，我们推导出一种新颖的期望分位数回归方法。据我们所知，目前已有的非参数期望回归方法并不多，因此，我们的方法或许能够填补这一空白，成为一项有价值的贡献。

考虑 $x \in \mathbb{R}$，令 $\alpha \in (0,1)$ 为期望水平。回顾一下，对于随机变量 Y， α-期望值定义为满足下式的唯一 $e_\alpha \in \mathbb{R}$

$$\alpha \mathbb{E}\left((Y - e_\alpha)_+\right) = (1-\alpha)\mathbb{E}\left((e_\alpha - Y)_+\right).$$

回到二元情形，考虑一对随机变量 (X, Y)，其连续分布函数为 H。令 K_H 按照式 (2.2) 定义。根据定义，给定 $X = x$ 时 Y 的条件 α-期望值 $e_\alpha(x)$ 是下式的唯一解

$$\alpha \int_{e_\alpha(x)}^{\infty} 1 - K_H(x, (-\infty, y])\, \mathrm{d}\lambda(y) = (1-\alpha) \int_{-\infty}^{e_\alpha(x)} K_H(x, (-\infty, y])\, \mathrm{d}\lambda(y).$$

定理 5.7 *Suppose that* $(X_1, Y_1), (X_2, Y_2) \ldots$ *is a random sample from* (X, Y) *and assume that* (X, Y) *has continuous distribution function* H*, marginals* F *and* G*, and underlying (unique) copula* A*. Furthermore suppose that* A *allows a continuous Markov kernel* K_A *define* K_H *according to eq.* (2.2) *and set* $N(n) = \lfloor n^s \rfloor$ *for some* $s \in \left(0, \frac{1}{2}\right)$*. For* $x \in \mathbb{R}$ *fulfilling the regularity assumption and every* $n \in \mathbb{N}$ *define* $t_n \in \mathbb{R}$ *as the solution of*

$$\alpha \int_{t_n}^{\infty} 1 - K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, G_n(y)])\, \mathrm{d}\lambda(y) = (1-\alpha) \int_{-\infty}^{t_n} K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, G_n(y)]\, \mathrm{d}\lambda(y).$$

Then we have that $(t_n)_{n\in\mathbb{N}}$ *converges to* $e_\alpha(x)$ *almost surely.*

证明 利用三角不等式以及我们可以交换积分与极限这一假设，我们有

$$\begin{aligned}
&\left|\alpha \int_{t_n}^{\infty} 1 - K_H(x, (-\infty, y])\, \mathrm{d}\lambda(y) - (1-\alpha) \int_{-\infty}^{t_n} K_H(x, (-\infty, y])\, \mathrm{d}\lambda(y)\right| \\
&\quad \leq \alpha \left| \int_{t_n}^{\infty} K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, G_n(y)]) - K_H(x, (-\infty, y])\, \mathrm{d}\lambda(y)\right| \\
&\qquad + (1-\alpha) \left| \int_{-\infty}^{t_n} K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x), [0, G_n(y)]) - K_H(x, (-\infty, y])\, \mathrm{d}\lambda(y)\right|
\end{aligned}$$

且后两项和几乎处处收敛到 0，当 $n \to \infty$ 时。这意味着$t_n \to t$ 当 $n \to \infty$ 时，因为函数

$$c \mapsto \left|\alpha \int_{c}^{\infty} 1 - K_H(x, (-\infty, y])\, \mathrm{d}\lambda(y) - (1-\alpha) \int_{-\infty}^{c} K_H(x, (-\infty, y])\, \mathrm{d}\lambda(y)\right|$$

是凸函数，因此具有唯一最小值。 □

6 模拟研究

本节的目的有二：首先，展示前几节推导的结果；其次，在回归场景下将我们提出的新估计量与现有估计量进行比较。与前文一样，我们首先聚焦于 Copula 框架，并通过模拟展示引理 3.3。随后，我们开展关于定理 5.3 和定理 5.6 的模拟，最后通过一项模拟强调，经验棋盘逼近也可用于估计其他感兴趣的条件量，例如条件方差。

6.1 Copula 框架

作为第一个例子，我们考虑 Ali-Mikhail-Haq（AMH）族，它由允许连续马尔可夫核的 Copula 构成。

例6.1 参数为 $\theta \in [-1, 1)$ 的 AMH Copula（参见 Durante 和 Sempi (2015, 例6.5.19)）由下式给出

$$C_\theta(x, y) = \frac{xy}{1 - \theta(1-x)(1-y)}.$$

C_θ 的马尔可夫核的连续版本由下式给出

$$K_{C_\theta}(x, [0, y]) = \frac{\theta(y-1)y + y}{(1 - \theta(1-x)(1-y))^2}$$

对于 $x, y \in \mathbb{I}$。考虑 $\theta = 0.75$，从 $(X, Y) \sim C_\theta$ 中抽取规模不断增大的样本 $n \in \{100, 250, 2500, 10000, 25000\}$，并计算

$$\left|K_{C_\theta}(x, [0, y]) - K_{\mathcal{CB}_{N(n)}(E_n)}(x, [0, y])\right|$$

在来自（U， V）的样本量为 $m = 2 \cdot \lfloor n^{0.45} \rfloor^2$ 的随机样本上 $\sim \Pi$，并计算这些点上的最大值，对每个样本量 n 得到一个值 $v \in \mathbb{I}$。将该值 v 解释为对

$$\sup_{(x,y) \in \mathbb{I}^2} \left|K_{C_\theta}(x, [0, y]) - K_{\mathcal{CB}_{N(n)}(E_n)}(x, [0, y])\right|,$$

并将模拟重复 $R = 2.000$ 次，得到汇集于图 1a 左图中的结果。随着 n 增大，误差似乎趋于零，从而证实了引理 3.2。

作为第二个例子，我们考察 Clayton copula。它不允许存在连续版本的马尔可夫核，并表明在这种情况下，经验棋盘逼近在条件意义下不一致收敛。

例 6.2 参数为 $\theta \in [-1, \infty) \setminus \{0\}$ 的 Clayton copula（见 Durante 和 Sempi（2015，例 6.5.17））定义如下：

$$C_\theta(x, y) = \max\left\{\left(x^{-\theta} + y^{-\theta} - 1\right)^{-1/\theta}, 0\right\}.$$

在这种情况下，对于 $x,\ y \in (0, 1]$，马尔可夫核的一个版本定义如下：

$$K_{C_\theta}(x, [0, y]) = \max\left\{x^{-\theta-1}\left(x^{-\theta} + y^{-\theta} - 1\right)^{-\frac{\theta+1}{\theta}}, 0\right\},$$

它在每个 $(x,\ y) \in (0, 1]^2$ 处连续。容易证明，C_θ 不允许存在连续马尔可夫核。事实上，考虑任意 Copula$C$，解体意味着

$$0 = \mu_C(\mathbb{I} \times \{0\}) = \int_{\mathbb{I}} K_C(x, \{0\})\, \mathrm{d}\lambda(x),$$

因此，对于 λ-几乎处处的 $x \in \mathbb{I}$，每个核都必须满足 $K_C(x,\ \{0\}) = 0$。直接推论是，每个连续马尔可夫核 K_C 都必须满足 $K_C(0,\ \{0\}) = 0$。回到 Clayton copula C_θ，假设 $(x,\ y) \mapsto K_{C_\theta}(x,\ [0, y])$ 在 $\mathbb{I}^2$ 上连续，并利用上述性质，可得

$$K_{C_\theta}(0, \{0\}) = K_{C_\theta}(0, [0, 0]) = 0 \neq 2^{-\frac{\theta+1}{\theta}} = \lim_{\varepsilon \to 0} K_{C_\theta}(\varepsilon, [0, \varepsilon]),$$

这就产生了矛盾。因此，C_θ 不允许存在连续马尔可夫核。考虑$\theta = 2$ 以及与 AHM 族相应的模拟设定，我们得到图 1b 右图所示的结果。箱线图表明，与 AHM 模拟相反，当 $n \to \infty$ 时，最大距离并不趋于零。

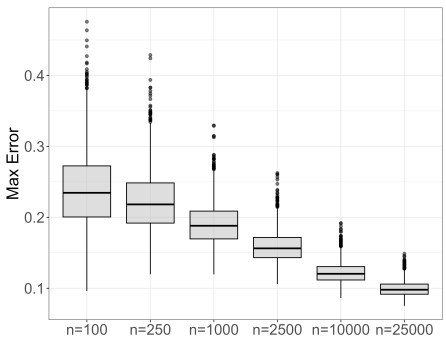
6.2 一般回归设定

我们考虑如下模拟设定。令 $s, \theta :\ [0, 10] \to (0, \infty)$ 为光滑的正函数，分别充当伽马分布的形状参数和尺度参数。为了确保正则性假设在全局成立，我们对该伽马分布函数作如下近似：令 c 为一个常数，使得 $c \gg s \cdot \theta$，并假设

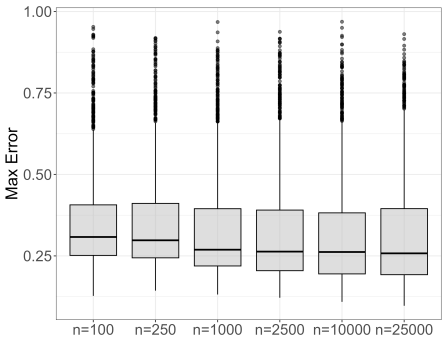
$$\begin{aligned} X &\sim 10 \cdot \mathrm{Beta}(\alpha, \beta), \\ Y \mid X &\sim c \cdot \mathrm{Beta}(\alpha'(X), \beta'(X)), \end{aligned}$$

其中 $\alpha' > 0$ 和 $\beta' > 0$ 是连续函数，且

$$\mathbb{E}\left(Y \mid X\right) = s(X)\theta(X), \quad \mathbb{V}(Y \mid X) = s(X)\theta(X)^2$$



(a) Ali-Mikhail-Haq copula, $\theta = 0.75$



(b) Clayton copula, $\theta = 2$

图1：AMH Copula（a）和 Clayton copula（b）的棋盘逼近收敛情况。

由于 $|Y|$ 被 c所界定，注5.2表明，定理5.3和定理 5.6中使用的正则性假设均得到满足。

现在我们考虑均值回归，并将第 5.1 节中分析的经验棋盘估计量（CBE）与常用的 Nadaraya-Watson 估计量（NWE）进行比较。根据注5.4，在需要进行大量预测时，CBE 的计算速度快于 NWE。值得注意的是，CBE 的另一项优势是无需指定带宽超参数。随后，我们使用 Silverman 经验法则（Silverman，1998）根据 $h_n = \widehat{\sigma}_n（X）\, n^{-\frac{1}{5}}$ 选择带宽 h_n。

例6.3 假设 $X \sim \mathcal{U}$(0,10)，并考虑

$$s(x) = \max\{0.5, \sqrt{x}\}, \quad \text{and} \quad \theta(x) = \min\{\max\{1, x\}, 6\}.$$

我们选择这个例子，是因为它有助于直观展示 CBE 和 NWE 这两种方法在异方差性下的表现。图2展示了来自该分布的 $n = 10.000$ 个样本点，以及 CBE（蓝色）、NWE（橙色）和真实回归函数（红色）。与连续的 NWE 相比，我们可以清楚地看到 CBE 的分段常数形式。在低方差区域，CBE 和 NWE 的表现相近。在高方差区域，CBE 的平滑度较低。计算最大误差和平均绝对误差（在由 $\mathbb{P}^X$ 中随机生成的 $m = 2.000$ 个点组成的网格上），重复该过程 $R = 2.000$ 次，并将所得数值汇总为箱线图，得到的结果如图3a和图3b所示。图中显示，在这个非常规则的设置中，NWE 比我们的 CBE 收敛更快，因为 NWE 按照构造是加权平均，因此能够提供平滑估计。

在下一个示例中，我们考虑这样一种情形：NWE 对均值回归函数的平滑程度过高，表现也不如 CBE。

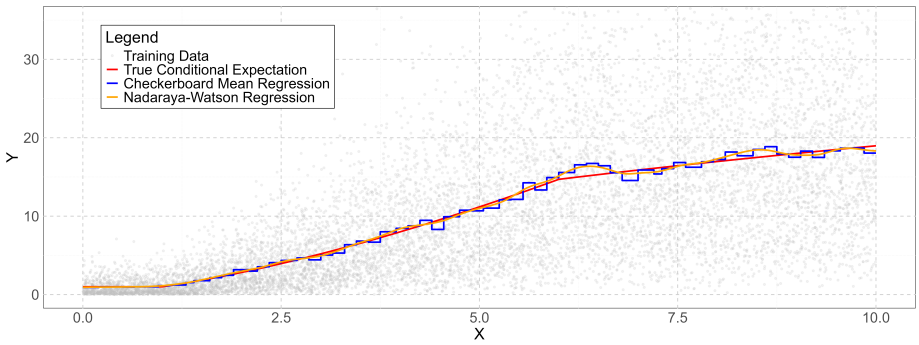
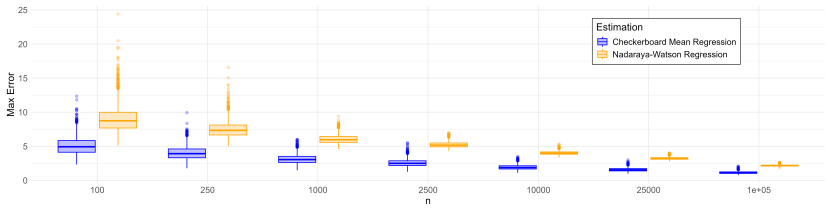
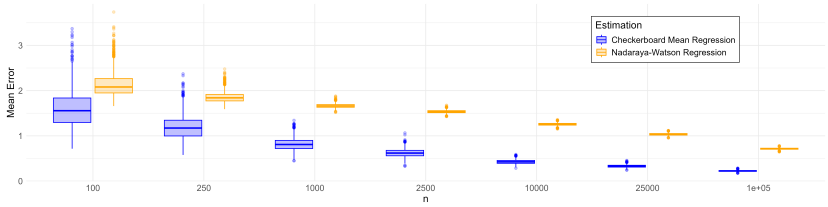


图2：从例6.3所考虑的分布中抽取的大小为 $n = 10.000$ 的样本。回归函数的CBE（蓝色）、NWE（橙色）和真实回归函数（红色）。



（a）最大误差



（b）平均误差

图3：例6.3所考虑的分布下，CBE（蓝色）和 NWE（橙色）的平均绝对误差（近似）；通过在 $m = 2.000$ 个随机生成的点上求值，计算平均/最大绝对误差。

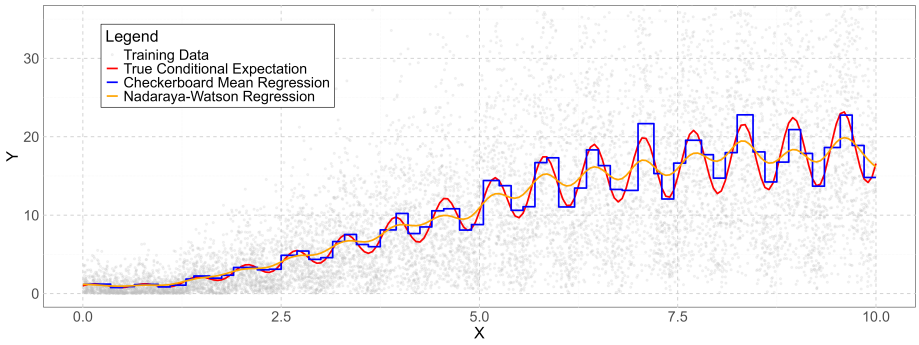


图4：来自例 6.4 所考虑分布的大小为 $n = 10.000$ 的样本。回归函数的 CBE（蓝色）、NWE（橙色）和真实回归函数（红色）。

例 6.4 仍考虑 $X \sim \mathcal{U}(0,10)$，但这次令

$$s(x) := \max\{1, \sqrt{x}\}\left(1 + \frac{\sin(10x)}{2}\right), \quad \theta(x) := \min\{\max\{1, x\}, 6\}.$$

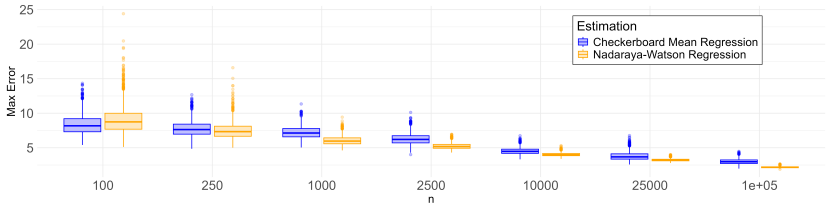
从该分布中抽样、进行计算并绘制 CBE、NWE 和真实回归函数，得到图4所示的结果。显然，NWE 在局部极值附近表现较差，而 CBE 的表现更好，在局部极值附近仅略有偏差（这是 CBE 作为真实回归函数估计量会产生阶梯函数的自然结果）。

对该分布重复例6.3中提到的模拟，得到图 5 汇总的结果。箱线图表明，CBE 和 NWE 估计量都一致收敛（且在 L^1-意义下收敛）。图 5 上图显示，CBE 和 NWE 的最大绝对误差相当接近。然而，下图显示，就平均绝对误差而言，CBE 的表现优于 NWE，这证实了一个自然的想法：在不那么规则／平滑的情形下，CBE 可能比 NWE 产生更好的预测。

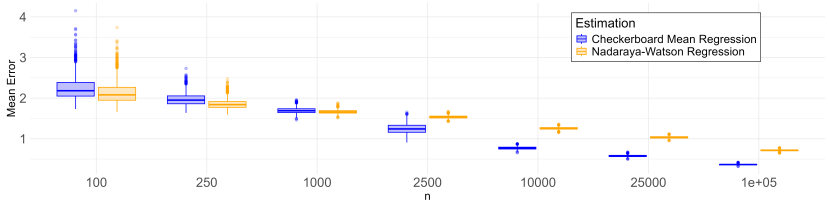
在最后一个关于均值回归的例子中，为了比较两个估计量在数据稀疏区域的表现，我们不再假设 X 服从均匀分布。

例6.5 设 $X \sim$ 服从Beta(2,4)分布，并令 $Y \mid X$ 与例6.3中的分布相同。这样一来，高协变量区域的数据就很稀疏。图 6显示，对于较大的 x，CBE比NWE能给出更好的预测，这表明在数据充足的区域，CBE可能优于NWE。值得注意的是，鉴于我们的方法基于对（秩）经验联结函数进行聚合，因此它比NWE表现出更强的稳健性也就不足为奇了。图7显示，CBE和NWE似乎在最大值意义和均值意义下都收敛。此外，就最大绝对误差而言，CBE的表现要好得多（见上图）；而就平均绝对误差而言，NWE的误差则略小。

接下来，我们通过一项模拟来展示 CBE 在分位数回归中的表现。在此背景下，我们比较我们的经验棋盘分位数



(a) 最大误差



(b) 平均误差

图 5：针对例 6.4 中考虑的分布，CBE（蓝色）和 NWE（橙色）的平均/最大绝对误差（近似）；平均/最大绝对误差通过在 $m = 2.000$ 个随机生成的点上进行评估计算得出。第二个面板清楚地显示，CBE 在平均绝对误差方面收敛更快。

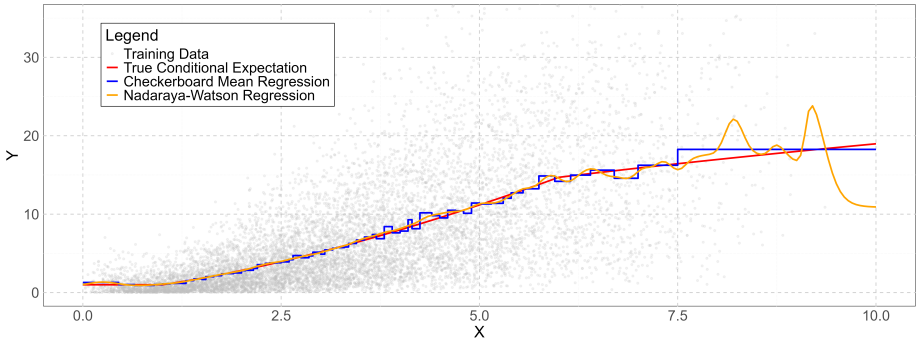
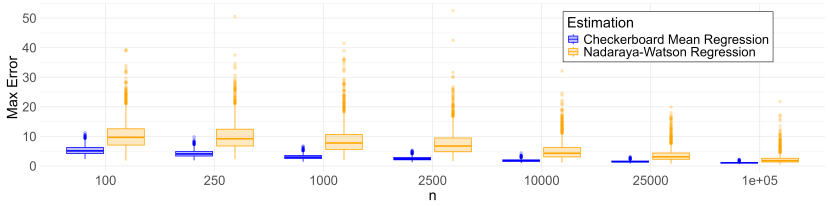
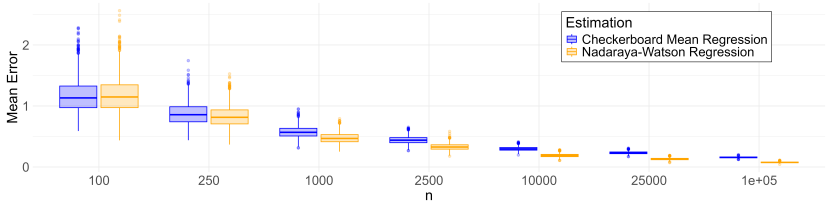


图 6：从例 6.5 中考虑的分布中抽取的大小为 $n = 10.000$ 的样本。回归函数的 CBE （蓝色）、 NWE （橙色）和真实回归函数（红色）。



(a) 最大误差



(b) 平均误差

图7：对于例6.5中考虑的分布，CBE（蓝色）和 NWE（橙色）的平均绝对误差（近似值）；平均/最大绝对误差通过在 $m = 2.000$ 个随机生成的点上进行评估计算得出。第一个面板清楚地表明，CBE 在最大绝对误差方面收敛更快。

第5.2节研究的估计量（CBQE）与用于分位数回归的 Nadaraya-Watson 估计量（Hall等，1999）（简称 NWQE）。请注意，对于预测 m 个点，我们提出的 CBQE 的时间复杂度为 $\mathcal{O}(m + N(n)n)$，而 NWQE 的时间复杂度为 $\mathcal{O}(mn\log(n))$。NWQE 中额外的 $\log(n)$ 项源于计算时所需的额外排序步骤。在 CBQE 中，只有预备步骤需要排序，此时 $\log(n)$ 被 $N(n)$ 所支配。

例6.6 我们再次考虑例6.3中讨论的分布，并聚焦于中位数（即 $\tau = \frac{1}{2}$）。计算并绘制 CBQE、NWQE 和真实分位数函数，得到图8所示的结果。与例6.3类似，我们可以清楚地观察到，CBQE 呈分段常数形式，而 NWQE 则生成光滑函数。再次比较高方差区域和低方差区域，可以观察到准确度存在差异。重复例6.3中的模拟，我们得到图9，其结果可用与图3类似的方式概括。

注6.7 需要特别注意的是，定理 3.1 中所述的经验棋盘逼近的一致条件收敛，其影响并不局限于回归。一个简单的例子是条件方差的非参数估计。考虑例6.3中的设置（调整样本量为 $n = 100.000$ 和 $s = 0.35$），并计算具有分布函数 $K_{\mathcal{CB}_{N(n)}(E_n)}(F_n(x)\ [0, G_n(\cdot)])$ 的随机变量 Z^x 的方差，在一组 x 值构成的网格上进行计算，得到图10中以蓝色显示的阶梯函数，它是条件方差的估计量；真实条件方差函数以红色显示。可以看到，估计量正确地捕捉到了真实条件方差函数的总体趋势

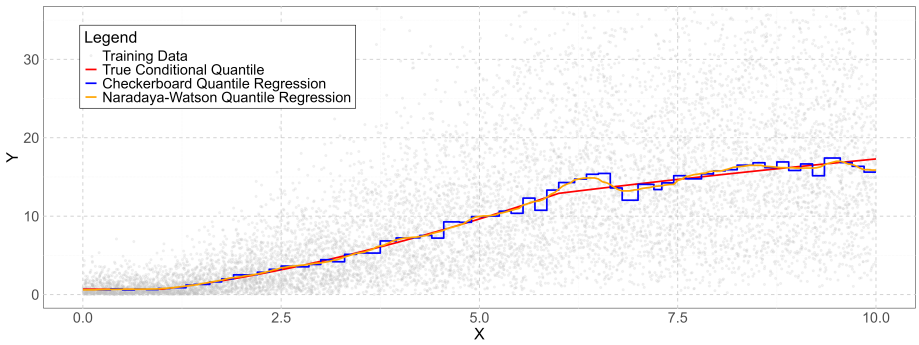
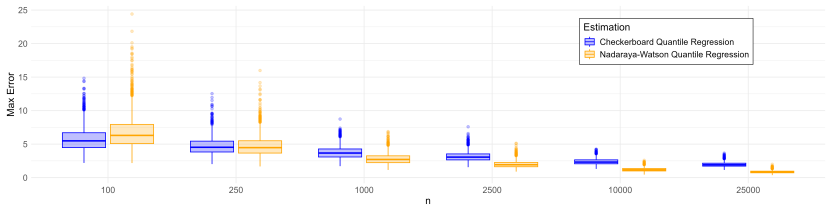
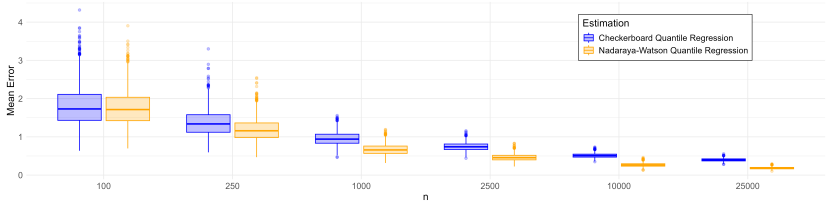


图8：从例6.6中考虑的分布中抽取的大小为 $n = 10.000$ 的样本。回归函数的CBQE（蓝色）、NWQE（橙色）和真实分位数函数（红色）。



（a）最大误差



（b）平均误差

图9：例6.6中考虑的分布下 CBQE（蓝色）和 NWQE（橙色）的平均绝对误差（近似值）；通过在 $m = 2.000$ 个随机生成的点处求值，计算平均/最大绝对误差。

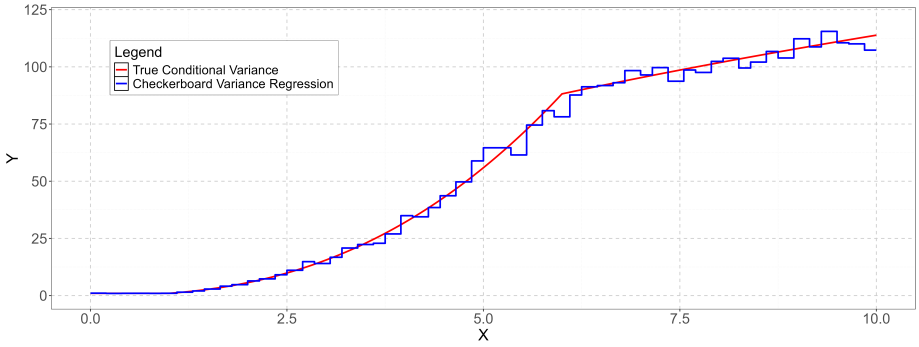


图10：从例 6.3 中考虑的分布抽取的样本，样本量为 $n = 100.000$。条件方差函数的 CBE（蓝色）和真实条件方差函数（红色）。

敛。尽管收敛速度似乎相当缓慢，但这凸显了定理 3.1 的灵活性及其广泛的应用范围。

注 6.8 请注意，如果不再局限于阶梯函数，而是希望得到更平滑的估计量，可以考虑将基于经验棋盘的估计量结合起来，并采用选定的平滑方法对得到的阶梯函数进行平滑。本文没有这样做，因为我们的主要关注点是经验棋盘逼近及其主要性质。

7 真实数据示例：保险损失与理赔处理费用

将我们的估计量应用于一个真实数据示例时，我们考虑了保险服务办公室提供的数据；Frees 和 Valdez（1998）对该数据进行了分析，其中包含 $n = 1.466$ 个未删失观测值。该数据集可通过 R 包 copula（Hofert 等，2025）公开获取，名称为 loss。我们关注协变量“对数赔偿金”（loss）与响应变量“对数分摊损失理赔款”（已分摊损失理赔费用，ALAE）之间的关系。图 11 展示了 $n = 1.466$ 个数据点，以及 CBE（蓝色）和 NWE（橙色）。可以看出，CBE 在尾部稳定得多，而众所周知，NWE 对单个点/离群值很敏感。为了比较 CBE 和 NWE 的表现，我们进行了以下小型实验，共重复 $R = 10.000$ 次：首先，将数据随机划分为训练集（80%）和测试集（20%），计算训练集的 CBE 和 NWE，并将其与真实观测值进行比较。图 12a 展示了所得结果，表明 CBE 和 NWE 的表现相近。

我们观察到，采用最大绝对误差进行评估时，我们提出的 CBE 表现更好，而在平均绝对误差方面，两者的精度相近。这表明，与广受欢迎的 NWE 相比，我们的方法并未牺牲准确度，同时计算效率更高，在尾部也更加稳定。

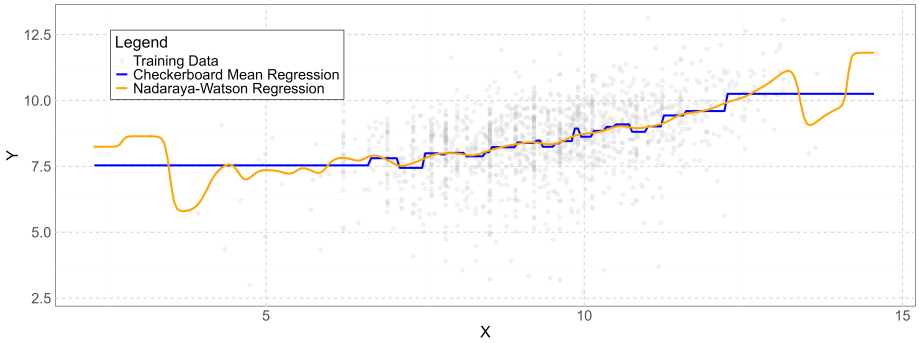
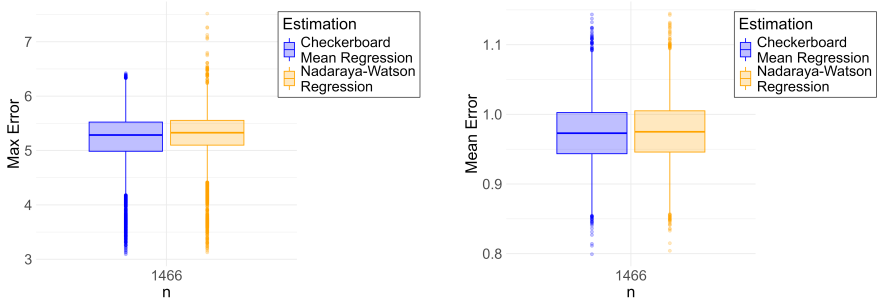


图 11：保险数据的 CBE（蓝色）和 NWE（橙色）。



（a）最大误差　　（b）平均误差

图 12：通过随机划分数据集，在训练集上训练模型并计算其在测试集上的误差，得到的 CBE 和 NWE 的最大/平均绝对误差（$R = 10.000$ 次运行）。

致谢。 两位作者谨此感谢 WISS 2025 项目“IDA-lab Salzburg”（ 20204-WISS/225/348/3-2023 和 20102/F2300464- KZP）提供的支持。

参考文献

Adam, C., Gijbels, I., 2022. 局部多项式期望分位数回归。《统计数学研究所年刊》74, 341–378. https://doi.org/10.1007/s10463-021-00799-y. Ansari, J., Langthaler, P.B., Fuchs, S., Trutschnig, W., 2026. 基于条件分布敏感性量化并估计依赖关系。《伯努利》32, 179–204. https://doi.org/10.3150/25-BEJ1854. Chaudhuri, P., 1991. 回归分位数的非参数估计及其局部 Bahadur 表示。《统计年刊》19, 760–777. https://doi.org/10.1214/ aos/1176348119. Choudhury, M., Das, D., Ghosh, S., 2026. 棋盘格 Copula 尾部依赖度量的渐近理论及乘子自助法的有效性。https: // arxiv.org/abs/2601.10252, arXiv:2601.10252. Das, S., Politis, D.N., 2020. 回归边界点处条件分布的非参数估计。《美国统计学家》74, 233–242. https://doi.org/10. 1080/00031305.2018.1558109. Dette, H., Hecke, R.V., Volgushev, S., 2014. 关于基于 Copula 的回归的一些评论。《美国统计学会杂志》109, 1319–1324. https://doi.org/10.1080/ 01621459.2014.916577. Durante, F., Sempi, C., 2015. Copula 理论原理。第 1 版，Chapman and Hall/CRC. https://doi.org/10.1201/b18674. Fan, J., 1993. 局部线性回归平滑器及其极小极大效率。《统计年刊》21, 196–216. http://www.jstor.org/stable/3035587. Fan, J., 1996. 局部多项式建模及其应用：统计学与应用概率专著第 66 卷（第 1 版）。Routledge. https://doi.org/10. 1201/9780203748725. Frees, E.W., Valdez, E.A., 1998. 使用 Copula 理解关系。《北美精算杂志》2, 1–25. https://doi.org/10.1080/10920277.1998.10595667. Griessenberger, F., Junker, R.R., Trutschnig, W., 2022. 关于一种基于多元 Copula 的依赖度量及其估计。《电子统计学杂志》16, 2206–2251. https://doi.org/10.1214/22-EJS2005. Hall, P., Wolff, R.C.L., Yao, Q., 1999. 条件分布函数的估计方法。《美国统计学会杂志》94, 154–163. https://doi.org/10.2307/ 2669691. Hoeffding, W., 1963. 有界随机变量之和的概率不等式。《美国统计学会杂志》58, 13–30. https://doi.org/10.2307/2282952.

Hofert, M., Kojadinovic, I., Maechler, M., Yan, J., 2025. copula：基于 Copula 的多元依赖。https://CRAN.R-project.org/package=copula。R 包版本 1.1-6. Huang, M.L., Nguyen, C., 2018. 分位数回归的非参数方法。《统计分布与应用杂志》5. https://doi.org/10.1186/s40488-018-0084-9. Janssen, P., Swanepoel, J., Veraverbeke, N., 2012. 伯恩斯坦 Copula 估计量的大样本行为。《统计规划与推断杂志》142, 1189–1197. https: //doi.org/10.1016/j.jspi.2011.11.020. Jordan, A.I., M¨ uhlemann, A., Ziegel, J.F., 2022. 可识别泛函保序回归问题最优解的刻画。《统计数学研究所年刊》74, 489–514. https://doi.org/10.1007/s10463-021-00808-0. Junker, R.R., Griessenberger, F., Trutschnig, W., 2021. 二元分布尺度不变的有向依赖之估计。《计算统计与数据分析》153, 107058 10 1016 2020 107058 . https://doi.org/ . /j.csda. . . Kaiser, H., Trutschnig, W., 2025. Copula 框架下的均值回归与分位数回归：性质、紧确界及关于估计的说明。https://arxiv.org/abs/2510. 03804, arXiv:2510.03804. Kallenberg, O., 2021. 现代概率论基础。概率论与随机建模。第三版，Springer Cham. https://doi.org/10. 1007/978-3-030-61871-1. Kasper, T.M., Fuchs, S., Trutschnig, W., 2021. 关于二元阿基米德 Copula 和极值 Copula 的弱条件收敛及其对非参数估计的影响。《伯努利》27, 2217–2240. https://doi.org/10.3150/20-BEJ1306. Klenke, A., 2020. 概率论。Universitext。第三版，Springer Cham. https://doi.org/10.1007/978-3-030-56402-5. Koenker, R., Gilbert Bassett, J., 1978. 回归分位数。《计量经济学》46, 33–50. https://doi.org/10.2307/1913643. Li, X., Mikusi ´nski, P., Taylor, M.D., 1998. Copula 的强逼近。《数学分析与应用杂志》225, 608–623. https://doi.org/10.1006/jmaa.1998.6056. Mroz, T., S ´anchez, J.F., Fuchs, S., Trutschnig, W., 2023. 关于具有固定边缘分布且使联合违约概率或先验违约概率最大化的分布、估计及相关结果。《统计规划与推断杂志》223, 33–52. https://doi.org/10.1016/ j.jspi.2022.07.005. Nadaraya, E.A., 1964. 关于回归估计。《概率论及其应用》9, 141–142. https://doi.org/10.1137/1109020. Nelsen, R.B., 2006. Copula 导论。Springer 统计学丛书。第二版，Springer 纽约，NY。https://doi.org/10.1007/0-387-28678-0. Newey, W.K., Powell, J.L., 1987. 非对称最小二乘估计与检验。《计量经济学》55, 819–847. https://www.jstor.org/stable/1911031.

Noh, H., Ghouch, A.E., Bouezmarni, T., 2013. 基于 Copula 的回归估计与推断。《美国统计学会杂志》108, 676–688. https://doi.org/10.1080/ 01621459.2013.783842. Sancetta, A., Satchell, S., 2004. 伯恩斯坦 Copula 及其在多元分布建模与近似中的应用。经济理论 20, 535–562. https://doi.org/10.1017/S026646660420305X. Segers, J., Sibuya, M., Tsukahara, H., 2017. 经验贝塔 Copula. 多元分析期刊 155, 35–51. https://doi.org/10.1016/j.jmva.2016.11.010. Silverman, B.W., 1998. 统计与数据分析中的密度估计。第 1 版，Routledge. https://doi.org/10.1201/9781315140919. Sklar, M., 1959. n 维分布函数及其边缘分布。Ann. I.S.U.P. VIII, 229–231. https://hal.science/hal-04094463v1. Tepegjozova, M., Zhou, J., Claeskens, G., Czado, C., 2021. 基于非参数 c-vine 和 d-vine 的分位数回归。依赖模型 10, 1–21. https://doi.org/10. 1515/demo-2022-0100. Watson, G.S., 1964. 平滑回归分析。Sankhya A. 26, 359–372. https: //www.jstor.org/stable/25049340. Yao, Q., Tong, H., 1996. 非对称最小二乘回归估计：一种非参数方法。非参数统计期刊 6, 273–292. https://doi.org/10.1080/ 10485259608832675.