arXiv:2301.05354v1 [math.PR] 13 Jan 2023

次线性期望下极大分布的最大似然估计 *

Xinpeng Li† Yue Liu Jiaquan Lu 山东大学数学与交叉科学研究中心，266237，中国青岛

摘要

最大似然估计是根据给定样本估计概率分布参数的一种常用方法。本文旨在介绍次线性期望框架下的最大似然估计。我们通过求解相关的极小极大问题，求得极大分布参数的最大似然估计量，该估计量与 Jin and Peng [8] 给出的最优无偏估计一致。本文还提出了一种适用于具有依赖结构的样本的一般估计方法。该结果为上方和下方方差估计量提供了理论基础，而该估计量广泛应用于金融领域的 G-VaR 预测模型中。

关键词：大数定律；极大分布；最大似然估计；次线性期望

1 引言

Peng [14] 建立的次线性期望理论，是处理许多领域中涉及模型不确定性问题的有力工具，尤其适用于解决金融领域中存在不确定性的动态问题（例如，参见 Epstein 和 Ji [3]），其中底层概率测度的数量可能是无限的。

次线性期望理论中的一种典型分布是极大分布。它通常用于刻画金融领域中的最坏情形风险，尤其是金融资产收益的不确定性（参见 Li 等人 [10] 和 Pei等人 [12]）。在应用中，极大分布随机变量用于建模的主要优势是计算简单。例如，考虑一维情形，在次线性期望 $\mathbb{E}$ 下，极大分布随机变量 X 的分布可由两个参数 $\underline{\mu}$和 $\overline{\mu}$确定，即，

$$\hat{\mathbb{E}}[\varphi(X)] = \max_{\underline{\mu}\le\mu\le\overline{\mu}} \varphi(\mu), \quad \forall\varphi \in C_b(\mathbb{R}).$$

由于 Peng 首创的大数定律（带有不确定性），它描述了许多现实现象，其中 i s 由 Peng [14] 首创（见定理 2.10）。

$-\mathbb{E}[X]$ 以及极大分布随机变量 X 的下均值 $\underline{\mu}$ 的合适估计量 $\hat{\mathbb{E}}[-X]$？最近，Jin and

*本研究得到山东省自然科学基金（编号：ZR2021MA018）、国家自然科学基金（编号：11601281）、国家重点研发计划（编号：2018YFA0703900）以及山东大学青年学者计划的资助。

†通讯作者。电子邮箱：lixinpeng@sdu.edu.cn

Peng [8] 发现，基于独立极大分布样本 $\{X_i\}_{i=1}^n$ ， $\overline{\mu}$ 的最大无偏估计量和 $\underline{\mu}$ 的最小无偏估计量分别可计算为

$$\hat{\overline{\mu}} = \max\{X_1, \cdots, X_n\}, \quad \hat{\underline{\mu}} = \min\{X_1, \cdots, X_n\}. \tag{1}$$

基于这些估计量，Peng 等人 [16] 以及 Peng 和 Yang [15] 在纳斯达克综合指数和标准普尔 500 指数上进行了大量实验，并展示了 G-VaR 预测器的出色性能。该预测器是对经典正态 VaR 模型的非平凡推广。

本文基于经典统计学理论中的最大似然估计（MLE）原理，为这些估计量提供了新的视角（例如，参见 Lehmann 和 Casella [9]）。我们根据经典 MLE 的本质提出了一个极小极大问题。我们最大化具有最小不确定性的样本的"概率"，其中附加的最小化问题旨在降低模型中的不确定性。我们发现，我们针对 $\overline{\mu}$ 和 $\underline{\mu}$ 的 MLE 与 Jin 和 Peng 的最优无偏估计（1）一致。此外，我们的估计量对依赖结构同样有效，并可用于近似不一定服从极大分布的样本。这一新结果为上方和下方方差估计量提供了理论基础，该估计量被广泛应用于 G-VaR 预测器模型中。

本文其余部分安排如下：第 2 节介绍次线性期望理论的一些基本概念和结果，以及极大分布的性质。第 3 节给出极大分布参数的详细最大似然估计。第 4 节研究非极大分布样本的一般估计量。

2 次线性期望理论预备知识

设 Ω 为一个波兰空间，且 $\mathcal{H}$ 是 Ω 上定义的实值函数的线性空间，使得对于每个 $n \in \mathbb{N}$，若 X_1、$\cdots$、$X_n \in \mathcal{H}$，则 $\varphi(X_1, \cdots, X_n) \in \mathcal{H}$、$\forall \varphi \in C_{Lip}(\mathbb{R}^n)$，其中 $C_{Lip}(\mathbb{R}^n)$ 是 Ω 上所有 Lipschitz 函数构成的空间。

定义 2.1 *A sublinear expectation* $\mathbb{E}$ *on* $\mathcal{H}$ *is a functional* $\mathbb{E}$: $\mathcal{H} \to \mathbb{R}$ *satisfying the followingconditions:* $\forall X$、$Y \in \mathcal{H}$、 *we have*

(1) *Monotonicity: if* $X \geq Y$, *then* $\mathbb{E}[X] \geq \mathbb{E}[Y]$;

(2) *Constant preserving:* $\mathbb{E}[c] = c$, $\forall c \in \mathbb{R}$;

(3) *Sub-additivity:* $\mathbb{E}[X+Y] \leq \mathbb{E}[X] + \mathbb{E}[Y]$;

(4) *Positive homogeneity:* $\mathbb{E}[\lambda X] = \lambda \mathbb{E}[X]$, $\forall \lambda \geq 0$.

三元组 $(\Omega, \mathcal{H}, \mathbb{E})$ 称为次线性期望空间，它类似于概率空间 $(\Omega, \mathcal{F}, P)$。

次线性期望的一个典型例子是由以下形式表示的上期望

$$\hat{\mathbb{E}}[X] = \sup_{P \in \mathcal{P}} E_P[X], \quad \forall X \in \mathcal{H}, \tag{2}$$

其中 $\mathcal{P}$ 是 $(\Omega, \mathcal{B}(\Omega))$ 上的某个概率测度集， E_P 是由 P 引入的线性期望。 $\mathcal{P}$ 的大小用于刻画模型的不确定性。在这种情况下，由 $\mathcal{P}$ 引入的相应容量可定义为

$$V(A) := \sup_{P \in \mathcal{P}} P(A), \quad \forall A \in \mathcal{B}(\Omega).$$

同分布和独立性的概念在经典概率论中十分重要，也可以在 Peng [13, 14] 的次线性期望理论框架下进行非平凡的推广。

定义 2.2 *Given an n-dimensional random vector* $X=(X_1,\cdots,X_n)$ *on a sublinear expec* -
ˆ*tation space* $(\Omega\ \mathcal{H},\mathbb{E})$ *where* $X_i\in\mathcal{H},\ 1\le i\le n$, *we define a functional on* $C_{Lip}(\mathbb{R}^n)$ *by*

$$\hat{\mathbb{F}}_X[\varphi]:=\hat{\mathbb{E}}[\varphi(X)],\quad \forall\ \varphi\in C_{Lip}(\mathbb{R}^n).$$

We call ˆ$\mathbb{F}_X[\varphi]$ *the sublinear distribution of* X *under* ˆ$\mathbb{E}$.

容易看出，$(\mathbb{R}^n,C_{Lip}(\mathbb{R}^n)\ \mathbb{F}_X)$构成一个次线性期望空间。

注 2.3 *Given an integrable random variable* X *on the classical probability space* $(\Omega\ \mathcal{F},P)$,
we recall that the distribution function of X *is defined by*

$$F_X(x)=P(X\le x),\quad \forall x\in\mathbb{R}.$$

For each $\varphi\in C_{Lip}(\mathbb{R})$ *we can easily calculate*

$$\hat{\mathbb{F}}_X[\varphi]=E_P[\varphi(X)]=\int_{\mathbb{R}}\varphi(x)dF_X(x).$$

ˆ*Conversely, if we know the value of* $\mathbb{F}_X[\varphi]$ *for every* $\varphi\in C_{Lip}(\mathbb{R})$ *then for each* $x\in\mathbb{R}$, *there exists a sequence of bounded and Lipschitz functions*

$$\varphi_n(y)=\frac{1}{1+n(y-x)^+}$$

such that

Then we obtain $\mathbb{F}_X[\varphi_n\ X\ ]\downarrow F_X\ x$ $\varphi_n(y)\downarrow 1_{(-\infty,x]}(y),\quad \forall y\in\mathbb{R}.$ $\mathbb{F}_X$
in the linear case. But for the sublinear case in particular the sublinear expectation $\mathbb{E}$ *admits* ˆ( )0

. *Thus the distribution function* F_X *is determined by* ˆ, , ˆ*representation (2), we emphasize that the following capacity*

$$V(X\le x):=\sup_{P\in\mathcal{P}}P(X\le x),\quad \forall x\in\mathbb{R},$$

can not always determine the value of ˆ$\mathbb{E}[\varphi(X)]$. *So we directly define* ˆ$\mathbb{E}[\varphi(X)]$ *for each* $\varphi\in C_{Lip}(\mathbb{R})$ *as the distribution of* X.

定义 **2 4** *Let* X *and* Y *be two n-dimensional random vectors defined on sublinear expecta-* 。 $(\Omega,\ \hat{})$
tion spaces $\mathcal{H}\ \mathbb{E}$. *They are called identically distributed if*

$$\hat{\mathbb{E}}[\varphi(X)]=\hat{\mathbb{E}}[\varphi(Y)],\quad \forall\ \varphi\in C_{Lip}(\mathbb{R}^n),$$

denoted by $X\overset{d}{=}Y$.

ˆ 以下独立性的概念为联合分布提供了一个简单模型 $\mathbb{E}[\varphi(X,Y)]$，前提是给定边缘分布。

ˆ 定义 2。**5** *Let* $(\Omega\ \mathcal{H},\mathbb{E})$ *be a sublinear expectation space, an n-dimensional random vector* Y *is said to be independent of another m-dimensional random vector* X *under the sublinear expectation* ˆ$\mathbb{E}$, *if* $\forall\ \varphi\in C_{Lip}(\mathbb{R}^{m+n})$,

$$\hat{\mathbb{E}}[\varphi(X,Y)]=\hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(x,Y)]_{x=X}].\tag{3}$$

Moreover, the sequence of random variables $\{X_i\}_{i=1}^{\infty}$ *is said to be independent, if for each* $i\ge 1$,
X_{i+1} *is independent of* $X_1\quad X_i$ $(,\cdots,)$.

注2.6 *In order to explain the equation (3), for simplicity, we only consider two random variables X and Y which are defined on the probability space $(\Omega, \mathcal{F}, P)$ with the joint distribution, function $F(x,y) = P(X \le x, Y \le y)$. If they are independent then $F(x,y) = F_X(x)F_Y(y)$ for all $x, y \in \mathbb{R}$, where F_X and F_Y are distribution functions of X and Y respectively, we further have*

$$\begin{aligned} E_P[\varphi(X,Y)] &= \int_{\mathbb{R}^2} \varphi(x,y)dF(x,y) \\ &= \int_{\mathbb{R}^2} \varphi(x,y)dF_X(x)dF_Y(y) \\ &= \int_{\mathbb{R}} dF_X(x) \int_{\mathbb{R}} \varphi(x,y)dF_Y(y) \\ &= \int_{\mathbb{R}} [E_P[\varphi(x,Y)]dF_X(x) \\ &= E_P[E_P[\varphi(x,Y)]|_{x=X}]. \end{aligned}$$

Thus Definition 2.5 is the natural generalization of classical notion of independence. By Fubini's theorem, we obtain

$$E_P[\varphi(X,Y)] = E_P[E_P[\varphi(x,Y)]|_{x=X}] = E_P[E_P[\varphi(X,y)]|_{y=Y}].$$

But it does not hold for sublinear expectation in general, the notion of independence under sublinear expectation is usually not symmetric, i.e. Y being independent of X can not automatically imply that X is independent of Y. An interesting example can be found in Example 1.3.15 of Peng [14]. More properties of such independence under sublinear expectation and its relations with classical conditional expectations is referred to Guo et al. [6].

注 **2.7** *We note that $f(x) = \mathbb{E}[\varphi(x,Y)]$ may be not continuous (resp. measurable) even if $\varphi(x,y)$ is continuous (resp. measurable). Thus the joint distribution $\mathbb{E}[\varphi(X,Y)] = \mathbb{E}[f(X)]$ is not well-defined by the marginal distributions since $\hat{\mathbb{E}}$ is defined on the domain of continuous (resp. measurable) functions. So we consider the Lipschitz functions in the sublinear expectation theory.*

注 **2.8** *By (3), it is obvious that for independent random variables $\{X_i\}_{i=1}^n$ and bounded Lipschitz functions $\varphi_i \ge 0, 1 \le i \le n$, we have*

$$\hat{\mathbb{E}}[\Pi_{i=1}^n \varphi_i(X_i)] = \Pi_{i=1}^n \hat{\mathbb{E}}[\varphi_i(X_i)], \tag{4}$$

which is equivalent to the classical independence when $\hat{\mathbb{E}}$ is the linear expectation.

We also note that (4) is weaker than (3). Moreover, (4) is an important property to obtain likelihood function in Section 3.

现在我们引入极大分布的概念，它是次线性期望理论中的基本次线性分布之一。

定义 2.9 *Let $(\Omega, \mathcal{H}, \mathbb{E})$ be a sublinear expectation space, an n-dimensional random vector X is said to be maximally distributed if there exists a bounded, closed and convex subset $\Lambda \subset \mathbb{R}^n$ such that*

$$\hat{\mathbb{E}}[\varphi(X)] = \max_{x \in \Lambda} \varphi(x), \quad \forall\, \varphi \in C_{Lip}(\mathbb{R}^n).$$

为简便起见，本文仅考虑一维情形。关于极大分布的更多细节，尤其是相关的极大分布随机场，可参见 Li 和 Peng [11]。

一维极大分布随机变量 X 的次线性分布定义如下

$$\hat{\mathbb{F}}_X[\varphi] = \hat{\mathbb{E}}[\varphi(X)] = \max_{\underline{\mu} \le x \le \overline{\mu}} \varphi(x), \ \forall\, \varphi \in C_{Lip}(\mathbb{R}^n), \tag{5}$$

ˆˆ 其中 $\overline{\mu} := \mathbb{E}[X]$ 且 $\underline{\mu} := -\mathbb{E}[-X]$，称极大分布随机变量 X 为$X \overset{d}{=} M_{[\underline{\mu},\overline{\mu}]}$。区间$[\underline{\mu},\overline{\mu}]$描述了 X的次线性分布的不确定性。由于该区间有界，因此对于所有连续函数 φ，式（5）仍然成立。

Peng [14] 中的以下大数定律在次线性期望理论中发挥着重要作用。

定理 **2.10** *Let $\{X_i\}_{i=1}^{\infty}$ be an independent and identically distributed (i.i.d.) sequence of random variables defined on $\mathcal{H}$ $\mathbb{E}$ and we further assume that X_1 is uniformly integrable under $\mathbb{E}$ i.e.*

$$\lim_{\lambda\to\infty} \hat{\mathbb{E}}[(|X_1| - \lambda)^+] = 0. \tag{6}$$

Then for all $\varphi \in C_{Lip}(\mathbb{R})$ we have

$$\lim_{n\to\infty} \hat{\mathbb{E}}\left[\varphi\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)\right] = \max_{\mu\in[\underline{\mu},\overline{\mu}]} \varphi(\mu). \tag{7}$$

注释 **2.11** *Recently Fang et al. [4] Song [17] and Hu et al. [7] obtained the convergence rate of (7) under higher moment conditions. If we further assume that $\hat{\mathbb{E}}[X_1^2] < \infty$ in Theorem 2.10 then we have*

$$\left|\hat{\mathbb{E}}\left[d^2_{[\underline{\mu},\overline{\mu}]}\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right]\right| \le \frac{\hat{\mathbb{E}}[X_1^2]}{n}, \tag{8}$$

where $d_{[\underline{\mu},\overline{\mu}]}(x) = $信息娱乐网格$_{y\in[\underline{\mu},\overline{\mu}]} |x - y|$.

In addition, Chen [1] and Zhang [18] established the corresponding strong law of large numbers.

命题 **2.12** *Let X be a random variable defined on sublinear expectation spaces $(\Omega, \mathcal{H}, \mathbb{E})$ we further assume that X is uniformly integrable under $\mathbb{E}$. Then X is maximally distributed if and only if*

$$aX + b\bar{X} \overset{d}{=} (a+b)X, \quad \forall a, b \ge 0. \tag{9}$$

where $\bar{X}$ is an independent copy of X, i.e., $\bar{X}$ is independent of X and $\bar{X} \overset{d}{=} X$.

证明. 设 $X \overset{d}{=} M_{[\underline{\mu},\overline{\mu}]}$ 和 $\Lambda = [\underline{\mu},\overline{\mu}]$，则有，$\forall \varphi \in C_{Lip}(\mathbb{R})$,

$$\begin{aligned}
\hat{\mathbb{E}}[\varphi(aX + b\bar{X})] &= \hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(ax + b\bar{X})]_{x=X}] \\
&= \max_{x\in\Lambda} \max_{\bar{x}\in\Lambda} \varphi(ax + b\bar{x}) \\
&= \max_{\mu\in\Lambda} \varphi[(a+b)\mu] \\
&= \hat{\mathbb{E}}[\varphi((a+b)X)],
\end{aligned}$$

因此（9）成立。

反过来，我们构造一个独立同分布序列 $\{X_i\}_{i=1}^{\infty}$，其满足 $X_1 \overset{d}{=} X$，并定义

$$\eta_n := \frac{1}{2^n}(X_1 + X_2 + \cdots + X_{2^n}), \quad \forall n \in \mathbb{N}$$

特别地，将 $a = b = 1$ 代入式（9），可得

$$\begin{aligned}
X_1 + X_2 &\overset{d}{=} 2X_1, \\
&\cdots \\
X_{2^n-1} + X_{2^n} &\overset{d}{=} 2X_{2^n-1}.
\end{aligned}$$

通过归纳法，对每个 $n \in \mathbb{N}$，可得

$$\begin{aligned}\eta_n &\overset{d}{=} \frac{1}{2^n}(2X_1 + 2X_3 + \cdots + X_{2^n-1}) \\ &\overset{d}{=} \frac{1}{2^{n-1}}(X_1 + X_3 + \cdots + X_{2^n-1}) \\ &\overset{d}{=} \cdots \overset{d}{=} X_1 \overset{d}{=} X.\end{aligned}$$

由定理 2.10 可得

$$\lim_{n\to\infty} \hat{\mathbb{E}}[\varphi(\eta_n)] = \max_{\mu\in\Lambda} \varphi(\mu),$$

这意味着

$$\hat{\mathbb{E}}[\varphi(X)] = \max_{\mu\in\Lambda} \varphi(\mu).$$

因此 X 是极大分布的。 ■

ˆ 推论 2.13 *An uniformly integrable random variable X on $(\Omega\ \mathcal{H}, \mathbb{E})$ is maximally distributed if and only if*

$$X + \bar{X} \overset{d}{=} 2X,$$

where $\bar{X}$ is the independent copy of X.

注释

2 14 *A counterexample in Guo and Li [5] shows that the law of large numbers fails 。without the uniformly integrable condition (6). It is still open that whether Proposition 2.12 still holds without such integrable condition.*

3 独立样本的最大似然估计

最大似然估计的思路是，找到合适的参数，使具有给定概率测度的总体中观测样本的概率 $P\ X_1 = x_1 \quad X_n = ($, $x_n \quad \{x_i\}_{i=1}^n \quad X \quad P$ $\cdots,)$ 最大。类似地，对于极大分布，我们希望使以下容量最大，并称其为似然函数，

$$V(x_1, \cdots, x_n) := \hat{\mathbb{E}}[I_{\{X_1=x_1,\cdots,X_n=x_n\}}].$$

值得指出的是， $I_{\{X_1=x_1,\cdots,X_n=x_n\}} \notin \mathcal{H}$ 因为示性函数不连续。但由于此类示性函数可以用 Lipschitz 函数逼近（见定理 3.3），因此可以对其进行良好定义。

此外，如果 $\hat{\mathbb{E}}$ 可以表示为

$$\hat{\mathbb{E}}[\cdot] = \sup_{P\in\mathcal{P}} E_P[\cdot],$$

那么自然可以定义

$$V(x_1, \cdots, x_n) := \sup_{P\in\mathcal{P}} P(X_1 = x_1, \cdots, X_n = x_n).$$

特别地，对于参数为 $\underline{\mu}$ 和 $\overline{\mu}$ 的极大分布总体 X，我们将相应的似然函数记为

$$V(x_1, \cdots, x_n; \underline{\mu}, \overline{\mu}).$$

显然，当区间 $[\underline{\mu}, \overline{\mu}]$ 扩大时，似然函数的值会增大。

令 $\Delta = \overline{\mu} - \underline{\mu}$ 为极大分布的不确定度，我们也希望在最大化似然函数时推导出不确定性，因此参数$\underline{\mu}$和 $\overline{\mu}$ 的最大似然估计是求解以下极小极大问题：

$$\min_{\Delta} \max_{\underline{\mu},\overline{\mu}} V(x_1, \cdots, x_n; \underline{\mu}, \overline{\mu}). \tag{10}$$

为了求解这类极小极大问题，我们首先建立一个以狄拉克测度表示极大分布的表示定理，其中点 $\mu \in \mathbb{R}$ 上的狄拉克测度记为 δ_μ，满足

$$\delta_\mu(A) = \begin{cases} 1, & \mu \in A, \\ 0, & \mu \notin A, \end{cases} \qquad \forall A \in \mathcal{B}(\mathbb{R}).$$

定理 3.1 *Let* $(\Omega\ \mathcal{H}, \hat{\mathbb{E}})$ *be a sublinear expectation space and* $X \overset{d}{=} M[\underline{\mu}, \overline{\mu}]$, *then for each* $\varphi \in C_{Lip}(\mathbb{R})$,

$$\begin{aligned} \hat{\mathbb{E}}[\varphi(X)] &= \max_{\underline{\mu} \le x \le \overline{\mu}} \varphi(x) \\ &= \max_{\underline{\mu} \le \mu \le \overline{\mu}} \int_{\underline{\mu}}^{\overline{\mu}} \varphi(y) \delta_\mu(dy). \end{aligned}$$

证明。另一方面，显然

$$\begin{aligned} \max_{\underline{\mu} \le \mu \le \overline{\mu}} \int_{\underline{\mu}}^{\overline{\mu}} \varphi(y) \delta_\mu(dy) &\le \max_{\underline{\mu} \le \mu \le \overline{\mu}} \max_{\underline{\mu} \le y \le \overline{\mu}} \varphi(y) \int_{\underline{\mu}}^{\overline{\mu}} \delta_\mu(dy) \\ &= \max_{\underline{\mu} \le y \le \overline{\mu}} \varphi(y). \end{aligned}$$

另一方面，存在 $\mu^* \in [\underline{\mu}, \overline{\mu}]$ 使得 $\varphi(\mu^*) = \max_{\underline{\mu} \le x \le \overline{\mu}} \varphi(x)$。于是我们有

$$\begin{aligned} \max_{\underline{\mu} \le \mu \le \overline{\mu}} \int_{\underline{\mu}}^{\overline{\mu}} \varphi(y) \delta_\mu(dy) &\ge \int_{\underline{\mu}}^{\overline{\mu}} \varphi(y) \delta_{\mu^*}(dy) \\ &= \varphi(\mu^*) = \max_{\underline{\mu} \le x \le \overline{\mu}} \varphi(x). \end{aligned}$$

■

注 3.2 *Let* $\{\delta_\mu, \mu \in [\underline{\mu}, \overline{\mu}]\}$ *be a sequence of Dirac measures with* $\delta_{\mu_n} \Rightarrow \delta_{\mu^*}$ *in* $[\underline{\mu}, \overline{\mu}]$. *Then there exists a subsequence* $\{\delta_{\mu_{n_i}}\}_{i=1}^{\infty}$ *such that* $\delta_{\mu_{n_i}} \Rightarrow \delta_{\mu^*}$ *provided* $\mu_{n_i} \to \mu^*$. *Thus the set of Dirac measures* $\mathcal{P} = \{\delta_\mu, \mu \in [\underline{\mu}, \overline{\mu}]\}$ *is weakly compact.*

其次，根据样本独立性，我们可以直接计算似然函数。

定理 3.3 *Let* $\{X_i\}_{i=1}^n$ *be a sequence of independent random variables on sublinear expectation space* $(\Omega\ \mathcal{H}\ \hat{\mathbb{E}})$. *We further assume that* $\hat{\mathbb{E}}$ *can be represented by*

$$\hat{\mathbb{E}}[X] = \max_{P \in \mathcal{P}} E_P[X], \quad \forall X \in \mathcal{H},$$

where $\mathcal{P}$ *is a weakly compact set of probability measures on* $(\Omega\ \mathcal{B}(\Omega))$。

Then the terms $\hat{\mathbb{E}}[I_{\{X_1 = x_1, \cdots, X_n = x_n\}}]$ *and* $\hat{\mathbb{E}}[I_{\{X_i = x_i\}}]$, $1 \le i \le n$ *are well-defined.*

Furthermore, $\forall x_i \in \mathbb{R}$, $1 \le i \le n$,

$$\hat{\mathbb{E}}[I_{\{X_1 = x_1, \cdots, X_n = x_n\}}] = \Pi_{i=1}^n \hat{\mathbb{E}}[I_{\{X_i = x_i\}}]. \tag{11}$$

证明。示性函数 $I_{x^*}(\cdot)$， $x^* \in \mathbb{R}$ 不连续，因此 $I_{\{X_i=x_i\}} \notin \mathcal{H}$。但它 ˆ 可以用 Lipschitz 函数逼近，因此 $\mathbb{E}[I_{\{X_i=x_i\}}]$ 可以被良好定义，ˆ$\mathbb{E}[I_{\{X_1=x_1,\cdots,X_n=x_n\}}]$亦然。

事实上，对于固定的 $x^* \in \mathbb{R}$，考虑函数

$$\varphi_k^{x^*}(x) = \frac{1}{1+k|x-x^*|}.$$

容易看出，$\varphi_k^{x^*}(\cdot)$ 是一个非负有界连续函数，并且 $\varphi_k^{x^*}(\cdot) \downarrow I_{x^*}(\cdot)$。对于每个 $k \in \mathbb{N}$，我们还有

$$\begin{aligned}
|\varphi_k^{x^*}(x_1) - \varphi_k^{x^*}(x_2)| &= \left| \frac{1+k|x_2-x^*| - 1 - k|x_1-x^*|}{(1+k|x_1-x^*|)(1+k|x_2-x^*|)} \right| \\
&\leq \frac{k|x_2-x_1|}{(1+k|x_1-x^*|)(1+k|x_2-x^*|)} \\
&\leq k|x_2-x_1|.
\end{aligned}$$

因此，$\varphi_k^{x^*}$ 是一个 Lipschitz 函数。

于是，由 Denis等人 的定理31， [2]，我们得到

$$\hat{\mathbb{E}}[\varphi_k^{x^*}(X_i)] \downarrow \hat{\mathbb{E}}[I_{\{X_i=x^*\}}].$$

ˆ 同样，对于每个 $n \in \mathbb{N}$，项 $\mathbb{E}[I_{\{X_1=x_1,\cdots,X_n=x_n\}}]$ 都是良定义的，可以近似为

$$\hat{\mathbb{E}}[\Pi_{i=1}^n \varphi_k^{x_i}(X_i)] \downarrow \hat{\mathbb{E}}[I_{\{X_1=x_1,\cdots,X_n=x_n\}}].$$

根据 $\{X_i\}_{i=1}^n$ 的独立性，我们得到， $\forall x_i \in \mathbb{R}$， $1 \leq i \leq n$，

$$\begin{aligned}
\hat{\mathbb{E}}[I_{\{X_1=x_1,\cdots,X_n=x_n\}}] &= \lim_{k\to\infty} \hat{\mathbb{E}}[\Pi_{i=1}^n \varphi_k^{x_i}(X_i)] \\
&= \lim_{k\to\infty} \Pi_{i=1}^n \hat{\mathbb{E}}[\varphi_k^{x_i}(X_i)] \\
&= \Pi_{i=1}^n \lim_{k\to\infty} \hat{\mathbb{E}}[\varphi_k^{x_i}(X_i)] \\
&= \Pi_{i=1}^n \hat{\mathbb{E}}[I_{\{X_i=x_i\}}].
\end{aligned}$$

■

由定理 3.3 和 3.1 可知，如果 $X \overset{d}{=} M[\underline{\mu}, \overline{\mu}]$，那么

$$\hat{\mathbb{E}}[I_{\{X=x^*\}}] = \max_{\underline{\mu} \leq \mu \leq \overline{\mu}} \delta_\mu(x^*).$$

现在我们可以求解极小极大问题 (10)。

定理 3.4 *Let $\{X_i\}_{i=1}^n$ be an i.i.d. sequence with $X_1 \overset{d}{=} M_{[\underline{\mu},\overline{\mu}]}$. Then the MLE of parameters $\underline{\mu}$ and $\overline{\mu}$ is given by*

$$\hat{\underline{\mu}} = \min\{X_1, \cdots, X_n\}; \quad \hat{\overline{\mu}} = \max\{X_1, \cdots, X_n\}. \tag{12}$$

证明。设 $\{x_i\}_{i=1}^n$ 为总体 $\{X_i\}_{i=1}^n$ 的样本。

由定理 3.1 和定理 3.3 可得

$$\begin{aligned}
V(x_1, \cdots, x_n; \underline{\mu}, \overline{\mu}) &= \hat{\mathbb{E}}[I_{\{X_1=x_1,\cdots,X_n=x_n\}}] \\
&= \Pi_{i=1}^n \hat{\mathbb{E}}[I_{\{X_i=x_i\}}] \\
&= \Pi_{i=1}^n \max_{\underline{\mu} \leq \mu \leq \overline{\mu}} \delta_\mu(x_i).
\end{aligned}$$

显然

$$V(x_1,\cdots,x_n;\underline{\mu},\overline{\mu})=\begin{cases}1, & \underline{\mu}\le \underline{x}\le \overline{x}\le \overline{\mu},\\ 0, & \text{otherwise},\end{cases} \tag{13}$$

其中 $\underline{x}=$最小值$\{x_1,\cdots,x_n\}$，且 $\overline{x}=$最大值$\{x_1,\cdots,x_n\}$。

因此，极小极大问题 (10) 的解为

$$\hat{\underline{\mu}}=\min\{X_1,\cdots,X_n\},\quad \hat{\overline{\mu}}=\max\{X_1,\cdots,X_n\}.$$

■

注 **3 5** *The MLE in (12) is the same as the largest unbiased estimator for $\overline{\mu}$ and smallest unbiased estimator for $\underline{\mu}$ of a parameter θ for which $\theta=\theta$ is the $f_n=$ is continuous on $\mathbb{R}^n$.*

4 一般最大似然估计量

本节首先证明，在不要求独立性的情况下，定理 3.4 仍然成立，这表明我们的结果也适用于具有依赖结构的样本。

定理 **4.1** *Let $\{X_i\}_{i=1}^n$ be identically distributed sequence with $X_1\overset{d}{=}M_{[\underline{\mu},\overline{\mu}]}$. Then the MLE of parameters $\underline{\mu}$ and $\overline{\mu}$ is given by*

$$\hat{\underline{\mu}}=\min\{X_1,\cdots,X_n\};\quad \hat{\overline{\mu}}=\max\{X_1,\cdots,X_n\}. \tag{14}$$

证明。设 $\{x_i\}_{i=1}^n$ 为总体 $\{X_i\}_{i=1}^n$的样本。显然，$\forall\ 0\le\varphi_i\in C_{Lip}(\mathbb{R})\ \ 1\le i\le n$，我们有

$$\hat{\mathbb{E}}[\Pi_{i=1}^n\varphi_i(X_i)]\le\Pi_{i=1}^n\max_{\underline{\mu}\le\mu_i\le\overline{\mu}}\varphi_i(\mu_i).$$

By 采用前面证明中的类似论证，我们得 到

$$\hat{\mathbb{E}}[\Pi_{i=1}^n I_{\{X_i=x_i\}}]\le\Pi_{i=1}^n\hat{\mathbb{E}}[I_{\{X_i=x_i\}}].$$

于是我们有

$$\begin{aligned}\bar{V}(x_1,\cdots,x_n;\underline{\mu},\overline{\mu}):&=\hat{\mathbb{E}}[I_{\{X_1=x_1,\cdots,X_n=x_n\}}]\\ &=\hat{\mathbb{E}}[\Pi_{i=1}^n I_{\{X_i=x_i\}}]\\ &\le\Pi_{i=1}^n\hat{\mathbb{E}}[I_{\{X_i=x_i\}}]\\ &=V(x_1,\cdots,x_n;\underline{\mu},\overline{\mu}),\end{aligned}$$

其中 $V(x_1,\cdots,x_n;\underline{\mu},\overline{\mu})$ 的定义与式 (13) 中相同。

由于样本 $\{x_i\}_{i=1}^n$ 已实现，我们有

$$0<\bar{V}(x_1,\cdots,x_n;\underline{\mu},\overline{\mu})\le 1,$$

thus

$$V(x_1,\cdots,x_n;\underline{\mu},\overline{\mu})=1,$$

这意味着

$$\hat{\underline{\mu}}=\min\{x_1,\cdots,x_n\},\quad \hat{\overline{\mu}}=\max\{x_1,\cdots,x_n\}.$$

■

在许多实际情形中，来自非极大分布总体 X 的样本 $\{X_i\}_{i=1}^n$ 往往满足独立同分布条件，而我们希望基于独立同分布样本 $\{X_i\}_{i=1}^n$ 估计某个 φ 的值 $\hat{\mathbb{E}}[\varphi(X)]$。得益于大数定律，我们的结果也可用于近似非极大分布总体。

一种典型的应用是基于历史时间序列 $\{Z_{t-i}\}_{i\geq 1}$ 估计具有模型不确定性的总体 Z 的上、下方差（参见 Li 等人 [10], 、Peng 等人 [16] 以及 Peng 和 Yang [15]）。

我们假设 Z 是零均值的，即

$$\hat{\mathbb{E}}[Z] = \hat{\mathbb{E}}[-Z] = 0.$$

Z的上、下方差分别定义为 $\overline{\sigma}^2 = \hat{\mathbb{E}}[Z^2]$ 和 $\underline{\sigma}^2 = -\hat{\mathbb{E}}[-Z^2]$，可按如下方式估计：

取足够大的 $L \in \mathbb{N}$，并使用窗口 L按如下方式计算局部样本方差

$$\sigma_j^2 = \frac{\sum_{i=1}^{L}(Z_{t-L+i-j} - \mu_j)^2}{L-1},\ 1 \leq j \leq K,$$

其中，$\mu_j = \frac{\sum_{i=1}^{L} Z_{t-L+i-j}}{L}$是相同窗口 L 下的样本均值，数值 $K \in \mathbb{N}$ 根据先验知识固定。K 的值可用于刻画模型的不确定性。K 越大，意味着我们倾向于认为模型具有更大的不确定性。

我们注意到，当 L 足够大时，项 μ_j 接近 0，因为 $\hat{\mathbb{E}}[Z] = \hat{}$

$$\sigma_j^2 \approx \frac{\sum_{i=1}^{L} Z_{t-L+i-j}^2}{L-1},\quad 1 \leq j \leq K.$$

对于每个 $j \in \{1, \cdots, K\}$，根据大数定律（定理 2.10），当 L 足够大时，σ_j^2 可视为 $[\underline{\sigma}^2, \overline{\sigma}^2]$ 上的极大分布随机变量。事实上，此类近似的误差项可由式（8）估计。

根据定理 4.1，我们得到 $\overline{\sigma}^2$ 和 $\underline{\sigma}^2$ 的最大似然估计：

$$\hat{\overline{\sigma}}^2 = \max_{1\leq j\leq K} \sigma_j^2;\ \ \hat{\underline{\sigma}}^2 = \min_{1\leq j\leq K} \sigma_j^2.$$

基于对方差上界和下界的此类估计，对纳斯达克综合指数和标准普尔 500 指数进行的大量实验表明，G-VaR 模型表现优异。该模型是一种基于所谓 G-正态分布的全新风险价值基准预测器，优于大多数现有的风险价值基准预测器（参见 [16] 和 [15]）。

注 **4 2** *Theorem 3.4 or Jin and Peng's optimal unbiased estimator can not be applied to* 。, *this situation since the sequence* $\{\sigma_j^2\}_{j=1}^K$ *is not independent. In fact*

it is L-dependent in sublinear case. Our results provide the theoretical foundation for the widely used estimator of upper and lower variances in finance.

参考文献

[1] Chen, Z. (2016)，次线性期望下的强大数定律。《中国科学：数学》，59(5)：945–954。[2] Denis, L., Hu, M. 和 Peng, S. (2011)，与次线性期望相关的函数空间与容量：在 G-布朗运动路径中的应用。Potential Analysis，34(2)：139–161。[3] Epstein, L. 和 Ji, S. (2013)，连续时间下的模糊波动率与资产定价。Review of Financial Studies，26(7)：1740C-1786。

[4] Fang, X., Peng, S., Shao, Q. 和 Yong, S. (2019)，次线性期望下具有收敛速度的极限定理。伯努利，25(4A)：2564–2596。[5] Guo, X. 和 Li, X. (2021)，次线性期望下伪独立随机变量的大数定律。《统计与概率快报》，172：第109042号。[6] Guo, X., Li, S. 和 Li, X. (2023)，次线性期望理论中的Peng独立性札记。《统计与概率快报》，193：第109719号。[7] Hu, M., Li, X. 和 Li, X. (2021)，次线性期望下Peng大数定律的收敛速度。《概率、不确定性与定量风险》，6(3)：261–266。[8] Jin, H. 和 Peng, S. (2021)，极大分布的最优无偏估计。《概率、不确定性与定量风险》，6(3)：189–198。[9] Lehmann, E. L. 和 Casella, G. (1998)，点估计理论。第二版。Springer统计学丛书。施普林格出版社，纽约。[10] Li, S., Li, X. 和 Yang, X. (2022)，模型不确定性下的上方和下方方差及其在金融中的应用。《国际金融工程杂志》，9(1)：第2250007号。[11] Li, X. 和 Peng, S. (2022)，次线性期望下的极大分布随机场。载于《随机分析、滤波与随机优化》，339–356，施普林格。[12] Pei, Z., Wang, X., Xu, Y. 和 Yue, X. (2021)，基于 G-VaR的最坏情形风险度量。《应用数学学报（英文版）》，37(2)：421–440。[13] Peng, S. (2019)，非线性期望下的大数定律与中心极限定理。《概率、不确定性与定量风险》，4，第4号。[14] Peng, S. (2019)，不确定性下的非线性期望与随机微积分：包含稳健中心极限定理与G-布朗运动。施普林格。[15] Peng, S. 和 Yang, S. (2022)，由 G-期望度量的金融时间序列分布不确定性。《概率论及其应用》，66(4)：729–741。[16] Peng, S., Yang, S. 和 Yao, J. (2020)，在模型不确定性下改进风险价值预测。《金融计量经济学杂志》，nbaa022（在线）。[17] Song, Y. (2021)，次线性期望下大数定律的Stein方法。《概率、不确定性与定量风险》，6(3)：199–212。[18] Zhang, L. (2021)，次线性期望下强大数定律的充分必要条件，arXiv:2104.08471。