管理科学

核密度估计在预测风险价值中的应用

唐林俊[1]，　杨　虎[2]，　张洪阳[1]

(1. 嘉兴学院数学与信息科学学院，浙江　嘉兴　314001)
(2. 重庆大学数理学院，重庆　400044)

摘要：　通过研究核密度估计理论，提出了一种适应估计金融时间序列分布的 Laplace核密度函数．在单变量核密度估计的基础上建立了风险价值 (Valua at Risk，简记为 VaR)预测的预测模型．通过对核密度估计变异系数的加权处理建立了两种加权 VaR 预测模型．最后，通过上证指数收益率对建立的 VaR 预测模型进行了实证分析，结果显示两种加权方法对上证指数收益率的 VaR 预测具有较高的效率．

关键词：　核密度估计；Laplace分布；风险价值 (Valua at Risk)

近年来，新出现的金融风险管理工具—— 风险价值方法 (Value at Risk，简记为 VaR)已得到广泛的重视和应用．在美国，三十小组 (The Group of Thirty)，衍生产品小组 (The Derivatives Product Group)，国际清算银行 (Bank of International Settlement)，以及欧盟 (European)等都在一定程度上将 VaR 作为风险度量的标准．VaR 的预测问题一直没有解决好，传统 VaR 定义不利于预测．本文在文献[1]的基础上提出了基于预测意义下的 VaR 定义，某段时间内的 VaR (置信度为 $1-p$)是这样一个数值，即在这段时间内连续 n 个交易周期中大约 $[np]$ 个交易周期的收益率小于该数值，这个数值可用顺序统计量 $X_{n,[np]+1}$ 刻画，并以 $X_{n,[np]+1}$ 之均值 $EX_{n,[np]+1}$ 作为该值的估计，其数学表达式：

$$VaR(p) = \int_{-\infty}^{+\infty} x f_{X_{n,[np]+1}}(x)\,\mathrm{d}x$$

$X_{n,[np]+1}$ 的分布比单纯的一个总体分位数提供的资产价值潜在的损失的信息更具有时效性．

1　Laplace核的提出

定义　设 $x_1, x_2, \cdots, x_n$ 为单变量 a_x 的一样本，a_x 的概率密度函数 $f(x)$ 的核密度估计定义为

$$\hat{f}(x) = \frac{1}{nh\hat{e}} \sum_{i=1}^{n} K\left(\frac{x - x_i}{\hat{e} h}\right) \tag{1}$$

式中 $K(\cdot)$ 为核函数，是一个给定的概率密度函数，即 $\int_{-\infty}^{+\infty} K(x)\,\mathrm{d}x = 1$；$\hat{e}$ 为变异量，可以选样本方差 (或绝对偏差)，h 为窗宽 (Band-width)系数；n 为样本容量．若研究的对象为多维 (d 维)时，

$$\hat{f}(x) = \frac{1}{nh^d \det(S)^{1/2}} \sum_{i=1}^{n} K\left[\frac{(X - X_i)^T S^{-1} (X - X_i)}{h^2}\right] \tag{2}$$

收稿日期：2003-07-07

式中 $X=(X_1,X_2,\cdots,X_d)^T$, $X_i=(X_{i1},X_{i2},\cdots,X_{id})^T$, $(i=1,2,\cdots,n)$, d为向量 X的维数; S是 X的 $d\times d$维协方差矩阵. (1)、(2)分别为一维和多维核密度估计的定义.

核函数 $K(\cdot)$的确定,在实际工作一般先选定核函数 $K(\cdot)$然后寻求最优窗宽系数 h,核函数的确定一般满足下列条件:① 对称且$\int K(t)\mathrm{d}t=1$;② 一阶矩为零,方差有限;③ 连续.常见的核函数有均匀核 $K(\cdot)$、Epanechnikov核、Bisquare核、高斯核等. 一般多维核密度估计用多维标准正态分布作为其核函数.

本文根据需要,我们采用标准 Laplace密度函数作为核函数,其中标准 Laplace密度函数为

$$f(x)=\frac{1}{2}\mathrm{e}^{-|x|} \tag{3}$$

以下我们称(3)为 Laplace核,它具有许多高斯核的性质. 下面我们进一步考察 Laplace核作为核估计时的一些性质.

性质 **1** 通过 Laplace核得到的核密度估计的尾部是光滑的.

证明

$$\hat{f}(x)=\frac{1}{nh}\sum_{i=1}^{n}K\left(\frac{x-x_i}{h}\right),\quad K(x)=\frac{1}{2}\mathrm{e}^{-|x|} \tag{4}$$

不妨证明上尾部 $x\to+\infty$的情况

当 $x\to+\infty$ $\exists M>0, M>x_i, i=1,2,\cdots,n$

对 $\forall x>M$,有 $x>x_i$, $i=1,2,\cdots,n$

$\therefore \hat{f}(x)=\frac{1}{2nh}\sum_{i=1}^{n}\mathrm{e}^{-\left|\frac{x-x_i}{h}\right|}=\frac{1}{2nh}\sum_{i=1}^{n}\mathrm{e}^{\frac{x_i-x}{h}}$

$\because \forall i=1,2,\cdots,n$,

$\mathrm{e}^{\frac{x_i-x}{h}}$连续可导,即有

$\hat{f}(x)$在 $x>M$时光滑可导,故 $\hat{f}(x)$上尾部光滑,对于下尾部的证明类似.

性质 **2** 式(4)核密度估计的分布如下:

$$\hat{F}(x)=\frac{1}{n}\sum_{i=1}^{n}\Phi_i(x-x_i)$$

其中

$$\Phi_i(x)=\begin{cases}\frac{1}{2}\mathrm{e}^{\frac{x-x_i}{h}}, & x\leq x_i\\ 1-\frac{1}{2}\mathrm{e}^{\frac{x_i-x}{h}}, & x>x_i\end{cases}$$

证明 $\because f_i(x)=\frac{1}{n}K\left(\frac{x-x_i}{h}\right)$

$$\therefore F_i(x)=\int_{-\infty}^{x}f_i(x)\mathrm{d}x=\begin{cases}\frac{1}{2}\mathrm{e}^{\frac{x-x_i}{h}}, & x\leq x_i\\ 1-\frac{1}{2}\mathrm{e}^{\frac{x_i-x}{h}}, & x>x_i\end{cases}$$

$\hat{F}(x)=\frac{1}{n}\sum_{i=1}^{n}\hat{f}_i(x)$,结论成立.

以上两个性质可以看出这个核函数虽然在光滑性质上不如高斯核函数,但是处处可导并且保持了尾部光滑特性,某些点的局部异常正是经济现象的表现;核密度函数估计出来后

概率分布可直接计算出来,不需要通过计算机进行大量的模拟和迭代运算. 在这里我们选取 Laplace核不仅考虑了它的经济意义同时也考虑到它数学计算上的优越性.

对于一个时序序列 $\{x_i\}_n$,将其在某段连续时间内的序列作为某一个总体,如,$\{x_{k-q+1}, X_{k-q} \cdots x_k\}$, $q<k<n$,这里连续时间跨度大小 q称为阶数. 阶数 q的选取是我们常常遇到的问题,比如我们用多少历史数据去预测或者模拟未来某一个交易日或者收益日的收益情况,阶数 q取得太大会导致预测时效性不敏感的问题,如果阶数 q取得太小预测受随机因素影响太大. 通常 q可由 AIC标准确定[2],即

$$AIC(q) = 2q + n\ln \hat{\sigma}_q^2$$

式中 $\hat{\sigma}_t^2$为残差的方差. 当 AIC达到最小时对应 q便为所求,通常 q的确定是根据投资者或决策者的目的而定的,我们在这里暂时不作深入研究.

窗宽 h的确定比较复杂,文献[3]提到了一些方法,如参数参照法,极大似然交叉证实法,最小二乘交叉证实检验法(LSCV)等. 若 h选的过大 $\hat{f}(x)$对 $f(x)$有较大的平滑,使得 $f(x)$的某些特性被掩盖. 如果 h选得过小 $\hat{f}(x)$有较大的波动. 一般取摩根银行研究的结果 $h_q = q^{-0.2}$, h_q的选取与阶数 q的大小有关.

2 单变量核密度估计预测模型

对于一个序列我们根据其不同特点可以分成两类情况加以考虑,一种是相依时间序列 $\{x_i\}_n$,即 x_t与 $x_{t-1}, x_{t-2}, \cdots, x_{t-q}$序列相关,取 $V_t = (x_{t-1}, x_{t-2}, \cdots x_{t-p})^T$,则 x_t的条件概率密度函数为

$$f(x_t | x_{t-1}, x_{t-2}, \cdots, x_{t-p} = f(x_t | V_t) = \frac{f(x_t, V_t)}{\int f(x_t, V_t)\mathrm{d}x_t} = \frac{f(x_t, V_t)}{f_v(V_t)}$$

式中 $f(x_t, V_t)$为 $q+1$维联合密度函数, $f_V(V_t)$为 q维边缘密度函数. 由多维核密度估计定义知:

$$\hat{f}(x_t, V_t) = \frac{1}{n-q}\sum_{i=q+1}^{n} \frac{1}{(2\pi h^2)^{(q+1)/2}\det(S)1/2}\exp\left[-\begin{pmatrix} x_t - x_i \\ V_t - V_i \end{pmatrix}^T S^{-1}\begin{pmatrix} x_t - x_i \\ V_t - V_i \end{pmatrix} \Big\backslash h\right]$$

$$\hat{f}_v(V_t) = \frac{1}{n-q}\sum_{i=q+1}^{n} \frac{1}{(2\pi h^2)^{q/2}\det(S_v)^{1/2}}\exp\left[-(V_t - V_i)^T S^{-1}(V_t - V_i)/h\right]$$

其中 $S = \begin{pmatrix} S_x & S_{xv} \\ S_{xv}^T & S_v \end{pmatrix}$.

式中 S为 (x_t, V_t)的 $(q+1)\times(q+1)$阶样本协方差矩阵; S_x为 x_t的样本方差; S_{xv}为 x_t与 V_t的 $1\times q$阶样本协方差矩阵; S_{xv}^T为 V_t与 x_t的 $q\times 1$阶样本协方差; S_v为 V_t的 $q\times q$阶样本方差阵, S由实测样本计算得到. $V_i = (x_{i-1}, x_{i-2}, \cdots, x_{i-q})$以及 x_i都来自实测样本 $(i = q+1, q+2, \cdots, n)$.

这个模型理论上非常可行,但是实际上,由于求核密度估计涉及协方差计算,使计算量非常大,其计算复杂度为 nq^2,并且上述模型是假设序列相关的,分布完全由给定的样本数据驱动,估计结果受异常点影响非常大,为了解决这些问题我们定义了下面的模型.

设时间序列 $\{x_i\}_n$随时间是 i. i. d,即 x_t与 $x_{t-1}, x_{t-2}, \cdots, x_{t-q}$序列不相关但同分布. 历史观测值 $x_{t-1}, x_{t-2}, \cdots, x_{t-q}$为对应分布 $Di(x)$的一个样本. 同时由于 $x \sim Di(x)$,所以 x_t的分

布可以通过样本序列 $x_{t-1},x_{t-2},\cdots,x_{t-q}$模拟估计出来. 由于 $x_{t-1},x_{t-2},\cdots,x_{t-q}$是连续历史数据,通过它们估计得到的分布作为 t时刻的分布,这就是典型的分布预测.下面用核密度估计建立分布预测模型:

设 x_i为第 $i(i=t-1,t-2,\cdots t-q)$个序列样本,那么 t时刻分布的核密度估计为:

$$\hat{f}_t(x)=\frac{1}{qh_q}\sum_{i=t-q}^{t-1}K\left(\frac{x-x_i}{\hat{e}_t h_q}\right)$$

其中 $h_q=q^{-0.2}$为所选窗宽,$K(\circ)$选择 Laplace核

$$K(u)=\frac{1}{2}\exp(-|u|)$$

其中 $\hat{e}_t=\frac{1}{q}\sum_{i=t-q}^{t-1}|x_i-\bar{x}_t|$,$\bar{x}_t=\frac{1}{q}\sum_{i=t-q}^{t-1}x_i$.

假定在一定时期内收益率的分布不变在现实中是可行的,这也符合经济“惯性”规律;对于独立性,一般经济序列都不适合,但是可以通过差分来做到,比如收益序列相关非常明显,但是收益率的相关性就不明显,可以认为是独立的序列[4].

3 单变量核密度估计在预测股市 *VaR*中的应用

实证应用中采用股票市场的对数收益率: $X_t=100\ln(P_t/P_{t-1})$,其中 X_t为第 t日的收益率,P_t为第 t日的股票市场价格.下面以股票市场日收益率的 VaR作为对象来研究股票市场日收益率的 VaR的预测,主要计算步骤:

1) 确定核密度函数 $K(\circ)$;估计样本步长(阶数)q;窗宽 h_q; VaR的置信水平 p;

2) 输入作为观测样本的历史数据$\{x_{T-1},X_{T-1}\cdots x_{T-q}\}$;

3) 利用 $x_{T-1},x_{T-2},\cdots,x_{T-q}$进行核密度估计,通过估计得到第 T日收益率分布的密度函数 $\hat{f}_T(x)$,通过上述性质 2得到其分布函数 $\hat{F}_T(x)$;

4) 利用次序分布与总体分布的关系,导出第 $[qp]+1$个顺序统计量 $X^{(T)}_{q,[qp]+1}$的分布密度函数 $f_{X^{(T)}_{q,[qp]+1}}(x)$:

$$f_{X^{(T)}_{q,[qp]+1}}(x)=([qp]+1)C_q^{[qp]+1}[\hat{F}_T(x)]^{[qp]}[1-\hat{F}_T(x)]^{q-[qp]-1}\hat{f}_T(x),\quad x\in R$$

5) 求出 $f_{X^{(T)}_{q,[qp]+1}}(x)$的均值 $EX^{(T)}_{q,[qp]+1}$即为第 T日收益率 VaR的预测值.

下图 1是利用 Laplce核估计得到的上证指数 1997年 8月 2号收益率预测分布图,直观上看下图的分布呈现尖峰厚尾现象,这与股市收益率的尖峰厚尾现象相符合[5].

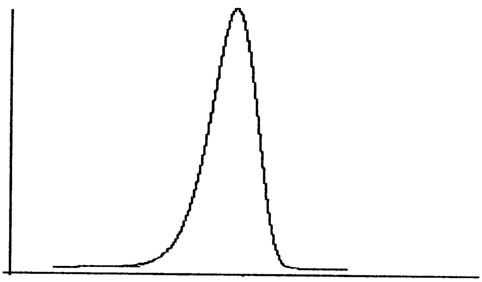

图 1

上面是估计 VaR的方法和步骤,是一种基本的非参数核密度估计.为了及时反映市场波动情况,可以对历史数据样本加权后再做核密度估计,而预测 VaR的方法同上面计算步骤一样.下面分别介绍两种加权方法:

4 预测模型的推广

由于预测分布的历史序列是连续的,为了及时反映股市波动,根据离预测点越近对预测点的信息越重要的原则,我们将信息越重要的历史数据取一个较大的权,反之,取一个较小

的权.下面介绍两种加权方法:

线性加权:记 x_i 为第 i 日的收益率数据,那么对样本 $x_{t-1},\cdots,x_{t-q}$ 赋予权重,$k_0 > k_1 > \cdots > k_{q-1}$,$\sum_{i=0}^{q-1} k_i = 1$ 核密度估计变为:

$$\hat{f}_t(x) = \frac{1}{qh_q}\sum_{i=t-q}^{t-1} K\left(\frac{x - x_i}{\hat{e}_t h_q}\right)$$

$K(u)$取 Laplace核,这里的参数

$$\hat{e}_t = \sum_{i=0}^{q-1} k_i \left| x_{t-q-i} - \bar{x}_t \right|, \quad \bar{x}_t = \sum_{i=0}^{q-1} k_i x_{t-q-i}$$

$q = 40$, $h_q = q^{-0.2}$这两个参数选取都是来自摩根银行研究的结果. k_i 的选取没有一个严格的限制,为了方便可以取 $k_i = \frac{2(i+1)}{q(n+1)}$, $i = 1, 2, \cdots, n$.

这个加权方法其实仅对参数进行加权,其他没有改变,如果我们对核估计也进行加权,并且取相同的权则有:

$$\hat{f}_t(x) = \frac{1}{h_q}\sum_{i=t-q}^{t-1} k_{i-t-q} K\left(\frac{x - x_i}{\hat{e}_t h_q}\right)$$

参数选取与前面加权方法一样.

指数加权 采用指数加权,对历史数据样本 $X_{t-1},\cdots,X_1$,赋予权重 $(1-\lambda)\lambda^i$, $0<\lambda<1$.

如果假设均值 $EX = \bar{X} = 0$,则对于 $\hat{e}_t = \frac{1}{t}\sum_{i=1}^{t} \left| X_i - \bar{X} \right| = \frac{1}{t}\sum_{i=1}^{t} \left| X_i \right|$

如果对历史数据样本赋予权重则有:

$\hat{e} = \sum_{i=1}^{\infty} (1-\lambda)\lambda^i \left| X_i \right|$,其中 $(0 < \lambda < 1)$,当 t 充分大时可以近似用下式表示:

$\hat{e} = \sum_{i=1}^{t} (1-\lambda)\lambda^i \left| X_i \right|$.

上式可以用一个迭代公式表示为

$$\hat{e}_t = (1-\lambda)\left| X_i \right| + \lambda \hat{e}_{t-1}, \quad 0 < \lambda < 1$$

迭代的起点设为 $\hat{e}_{t0} = \frac{1}{q}\sum_{i=0}^{q-1} \left| X_i \right|$, $q < t$. 如果按照 RiskMetrics 中的参数则 $q = 74$, $\lambda = 0.94$,核密度估计可以采用下面的迭代公式估计:

$\hat{f}_{t0} = \frac{1}{qh_q}\sum_{i=0}^{q-1} K\left(\frac{x - X_i}{\hat{e}_i h_q}\right)$,作为迭代起始点,$\hat{f}_t(x) = \lambda \hat{f}_{t-1}(x) + (1-\lambda)\cdot \frac{1}{h_q} K\left(\frac{x - X_t}{\hat{e}_t h_q}\right)$, $t > t_0$,作为迭代公式,其中窗宽 $h_q = q^{-0.2}$, $\lambda = 0.94$, $q = 74$, $K(u)$ 取 Laplace核.

5 实证效果的检验和比较

实证数据取 2000-7-15日到 2002-1-1日上证指数每日对数收益率(数据来源于分析家软件),上证指数的收益率序列如下图 2.序列收益率基本上趋于平稳,局部相关性不显著.

为了检验三种预测方法的实证效果,我们定义了两种判断预测效果的标准.

Ⅰ、如果连续对第 T_1 日到第 T_2 日的 VaR 的进行估计,那么按照本文定义的 VaR,下

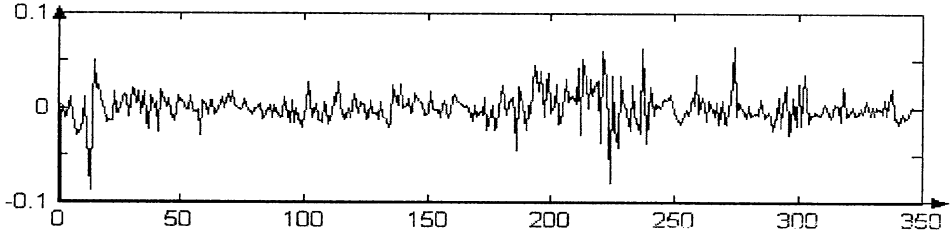

图 2

面比值应该接近尾部概率即 VaR 的置信度水平 $a= 0.05$(针对股票这里只研究下尾部的情况).

$$\frac{\#\ (X_t < VaR_t, T_1 \leq t \leq T_2)}{T_2 - T_1} \approx 0.05$$

其中# (°)表示符合括号内条件的数据个数,X_t 为 t 天实测数据,VaR_t 为 t 天预测 VaR.

II、平均预测绝对偏差,为了更能反映问题这里只考虑负收益率与其预测 VaR 的平均绝对偏差:

$$Error = \frac{1}{T_2 - T_1}\sum_{t=T_1}^{T_2} |X_t - VaR_t|, \quad \text{If} \quad X_t < 0$$

其中,X_t 为 t 天实测数据,VaR_t 为 t 天预测 VaR.

下面图 3- 5是核密度估计方法、核密度线性和指数加权估计方法得到的结果演示:

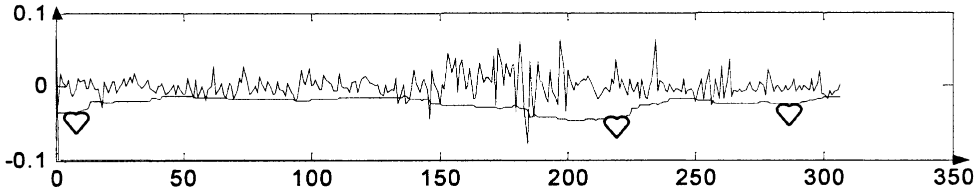

图 3(核密度估计法)

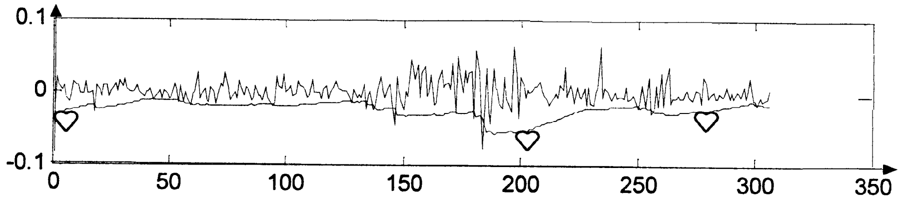

图 4 (线性加权)

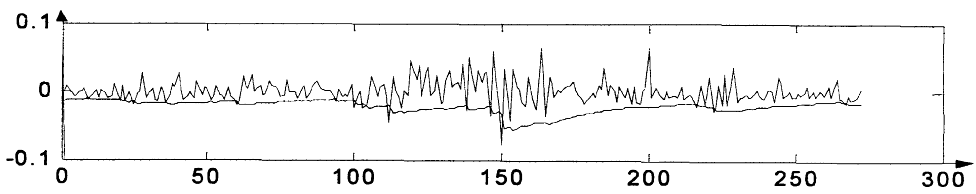

图 5 (指数加权)

通过图 3- 5,可以看出核密度估计方法预测得到的预测值具有比较明显的滞后现象(连续多个观测点的预测 VaR 值明显小于真实值),图 3- 4中标有 标志的地方.线性加权方法虽然还是具有一定的滞后现象,但是大大的减弱了,指数加权在这方面已经做得相当出色.从图表 1可以看出,两种加权方法均优于核密度估计方法.线性加权方法与指数加权方法各有所长,根据不同的标准要看各人的偏好,建议使用指数加权方法.我们在用其他的一些数据计算时得到的结果有所不同,但是加权方法均优于没有加权的情况.

表 1 两种标准下的实证效果检验

	标 准 1	标 准 2
核密度估计	0.0523	0.0142
线性加权	0.0490	0.0140
指数加权	0.0515	0.0135

最后解释一下 VaR 预测结果的意义,比如第 70个交易周期的实测收益率 $R = -0.018$,即损失为 1.8%;而当天使用上述三种方法预测结果分别为: $VaR = -0.019$; -0.0178; -0.0185,其含义是,该股票可以以 95%的可能性(置信度)保证第 70个交易周期的损失不超过 1.9%, 1.78%, 1.85%.

参考文献:

[1] 郑祖康等. 非参数方法在金融风险管理模型中的应用[J]. 系统工程, 1999, 17(5): 25- 32.
[2] 陈希孺等. 非参数统计[M]. 上海科技出版社, 1978.
[3] 张尧庭. 金融市场的统计分析[M]. 广西师范大学出版社, 1998.
[4] Risk-Value Models Rakish Sarin, Martin Weber- UCLA, Anderson Grad. Sch., 405 Hilgard Ave., Los Angeles, CA 90024- 1481.
[5] 唐林俊等. Laplace分布在金融市场中的应用[C]. 2002年中国科协年会学术论文,成都, 2002, 9.

The Application of The Kernel Density Estimates in Predicting *VaR*

TANG Lin-jun[1], YANG Hu[2], ZHANG Hong-yang[1]

(1. The College of Mathematics and Information Science, Jiaxing University, Jiaxing Zhejiang 314001, China)
(2. The College of Mathematics & Physics, Chongqing University, Chongqing 400044, China)

Abstract After studying the theory of kernel density, propose a kind of Laplace Kernel Function, which adapts estimating the distribution of finance time series. Based on single variable kernel density estimator, constructs forecast model of VaR, and by weight variation coefficient get two weighted forecast model of VaR. Finally, Using Shanghai's securities index tests 3 models, shows weighted methods higher efficiency.

Keywords kernel density estimates; laplace-distribution; value at risk