Modeling and Estimation of VaR using Extreme Value Theory

By Pouyan Mashayekh Ahangarani

University of Southern California

Economics Department

March 2005

ABSTRACT

A combination of nonparametric approach and extreme value theory will be shown as a method for estimation of VaR. Locally Parametric Nonparametric estimation will be used which is a method that has less bias relative to other standard nonparametric methods. The conditions will be derived in order to estimate the distribution function nonparametrically. The threshold model of extreme values is used in order to circumvent the lack of observation problem at the tail of distribution function. At the end, the results will be reported.

1-Introduction

Recent financial disasters have emphasized the importance of effective risk management for financial institutions. The use of quantitative risk measures has become an essential management tool to be placed in parallel with models of returns. These measures are used for investment decisions, supervisory decisions, risk capital allocation and external regulation.

Value at Risk (VaR) has become the standard measure of risk employed by financial institutions and their regulators. VaR is an estimate of how much a certain portfolio can lose within a given time period and at a given confidence level. More precisely VaR is defined in such a way that the probability of VaR over a particular time horizon is equal to p, a prespecified number. The great popularity that this instrument has achieved among financial practitioners, is essentially due to its conceptual simplicity : VaR reduces the risk associated with any portfolio to just one dollar amount. The summary of many complex bad outcomes in a single number naturally represents a compromise between the needs of different users. This compromise has received the blessing of a wide range of users and regulators.

Despite its conceptual simplicity, the measurement of VaR is a very challenging statistical problem. The existing models for calculating VaR differ in the methodology they use, the assumptions they make and the way they are implemented. The Gaussian i.i.d price increments approach assumes that price increments are normally distributed. It consists of inverting the cumulative normal function that approximates the marginal distribution of returns evaluated at the mean and variance estimated from a sample of past returns. This procedure was suggested by J.P Morgan (1996) and is often used in regulatory purposes. Obviously, the use of a normal approximation of the marginal return distribution results in underestimated tails and disregards excess kurtosis and skewness displayed by the empirical marginal distributions of returns. Nonparametric estimation of distribution function of returns has been another method in order to circumvent this problem. The lack of number of observations especially at extreme values has been the major flaw of this method. The conditional quantile estimation approach was suggested by Engle and Manganelli (1999). These authors introduced a recursive specification for the VaR corresponding to a given level. By recursive substitutions, VaR is expressed as a function of the return history and the parameters. The VaR specifications are not compatible when the parameters related to degree of risk change. Gourieroux and Jasiak (2000) proposed estimation of local extreme risk from Gaussian model and to replace the conditional mean and variance in the Gaussian formula of the VaR by local approximations. Other nonparametric methods are also have been used for finding the distribution of the returns. Application of extreme quantile estimation methods to VaR have been proposed by Danielsson and de Vires (1998) . The intuition here is to exploit results from statistical extreme value theory and to concentrate the attention on the asymptotic form of the tail rather than modeling the whole distribution. Chernozhukov (2001) has used the order statistics in order to estimiate the conditional VaR.

In this paper I try to use Locally Parametric Nonparametric method to find the distribution function of returns. Hjort and Jones (1996) have introduced this method that theoretically has less bias relative to classical nonparametric methods. In order to decrease the bias of estimation I will use the theory of threshold excesses to estimate the tail distribution of returns.

This paper is organized as follows. Section 2 will discuss about the benefits of using LPNE relative to the standard Kernel estimation. Section 3 presents the theory of statistical extreme value used in this paper. Section 4 introduces the data and the empirical results will be presented.

2- Locally Parametric Nonparametric approach

The first step in finding the VaR is to estimate the distribution function of returns. Nonparametric methods can be used in order to find the quantiles. Then the extreme value theory (which will be discussed in section 3) will be used to find the distribution function in tails.

Let $X_1,\ldots,X_n$ be i.i.d with density function f, The traditional Kernel estimator of f is $\tilde{f}(x) = n^{-1}\sum_{i=1}^{n} K_h(x_i - x)$ where $K_h(z)=h^{-1}K(h^{-1}z)$ and K(.) is some chosen unimodel density, symmetric about zero. The basic properties of $\tilde{f}$ are well known and under smoothness assumptions these include

$$E\,\tilde{f}(x) = f(x) + 0.5\,\sigma^2 h^2 f''(x) + O(h^4) \qquad (2.1)$$

$$Var\,\tilde{f}(x) = R(k)\,(nh)^{-1} f(x) - n^{-1}[f(x)]^2 + O(h/n) \qquad (2.2)$$

Where $\sigma^2 = \int z^2 k(z)dz$ and $R(k) = \int [k(z)]^2\,dz$, See scott (1992, chapter6).

A semiparametric approach has been used to obtain the local linear estimator of $\tilde{f}(x)$ by fitting a line, a+b(t-x) locally to x. This amounts to replacing:

$$f(t)= f(x) + f'(x)(t-x)=a+b(t-x)$$

and measuring with respect to a and b. A further extension is to consider local polynomials. Loader(1993) has considered using local exponential functions of the form aExp(b(t-x)). Most of these local polynomial estimators have the advantages, compared to the Kernel estimator, that they perform better in the tails.

Hjort and Jones (1996) have proposed and investigated a new class of semiparametric competitors which have precision comparable to that of $\tilde{f}$ but sometimes better. For any given parametric family, $f(.,\theta) = f(.,\theta_1, \ldots, \theta_p)$ and for each given x, they have presented ways of estimating the locally best approximation to f and then use

$$\hat{f}(x) = f(x, \hat{\theta}_1(x),\ldots, \hat{\theta}_p(x)) \qquad (2.3)$$

Thus the estimated density of x employs a parametric value which depends on x and whose choice is to be tailored to good estimation at x. In other words, the method amounts to a version of nonparametric parameter smoothing within the given parametric class.

We approximate the unknown distribution f of x by the pseudo family F={f(x,θ),θ varying } on an interval A=[c-h, c+h]. The resulting approximation of the θ parameter is :

$$\tilde{\theta}_{c,h} = \quad (2.4)$$

$$\text{ArgMax } E[(1/h)K((x-c)/h)\log f(x;\theta)]-E[(1/h)K((x-c)/h)]\log\int (1/h)K((x-c)/h)f(x;\theta)dx$$

Which corresponds to the optimization of the Kullback-Leibler criterion. The above definition is valid for any Kernel K and x. The Local Parameter Function (LPF) is defined as the limit of $\tilde{\theta}_{c,h}$ when h tends to zero.

The Locally Parametric Nonparametric estimation shown in (2.3) has these characteristics:

$$E\,\hat{f}\ (x) = f(x) + 0.5\,\sigma^2 h^2 b(x) + O(h^4 + (nh)^{-1})$$

$$\text{Var}\,\hat{f}\ (x) = R(k)\,(nh)^{-1} f(x) - n^{-1}\,[f(x)]^2 + O(h/n)$$

Which is just like (2.2) and (2.3) but with a bias factor function b(x) related to but different from $f''(x)$, with characteristics inherited from the parametric class. To the order of approximation used, the variance is simply the same, regardless of parametric family, The statistical advantage will be that for many "f"s , typically those lying in a broad nonparametric neighborhood of the parametric f(.,θ), b(x) will be smaller in size than $f''(x)$ for most x (Hjort and Jones (1996)).

In this paper the standard normal assumption for Kernel function and distribution function f, has been used, If we plug the normal distribution function in (2.4) we will come up with these conditions for unknown parameters μ_c and σ_c (see Appendix A).

$$\sum_{t=1}^{T} (1/h)K((x_t - c)/h) [\sigma_c^{-2}(x_t - \mu_c) + (\sigma_c^2 + h^2)^{-1} (\mu_c - c)] = 0 \quad (2.5)$$

$$\sum_{t=1}^{T} (1/h)K((x_t - c)/h)[-(\mu_c - c)^2 + (\sigma_c^2 + h^2)^{-1} + \sigma_c^{-4} (x_t - \mu)^2 - \sigma_c^{-2}] = 0 \quad (2.6)$$

3- The theory of statistical extreme value

A disadvantage of the Nonparametric estimation, is the low frequency of observations at tails which leads to the estimations which exhibits a very high frequency. The variance is very high and in some cases even infinite. As a result, it will bring poor estimates of the tails which are very important for VaR estimations. The theory of statistical extreme value mitigates the problem by introducing a parametric distribution functions in tails.

Extreme Value theory has emerged as one of the most important statistical disciplines for the applied sciences over the last decades. Extreme value techniques are also becoming widely used in many other disciplines. Applications of extreme value modeling have been published in the fields of alloy strengthen prediction (Tryon & Cruse, 2000);Ocean wave modeling (Dawson,2000);memory cell failure (McNulty et al., 2000);Wind engineering (Harris, 2001); management strategy (Dahan & Mendelson, 2001);biomedical data processing (Roberts,2000); thermodynamics of earthquakes (Lavenda 2001); non-linear beam vibration (Dunne & Ghanbari, 2001); annd food science (Kawas & Moreira, 2001).

The distinguishing feature of an extreme value analysis is the objective to quantify the stochastic behaviour of a process at unusually large – or small – levels. In particular , extreme value analyses usually require estimation of the probability of events that are more extreme than any that have already been observed.

The cornerstone of the extreme value theory is modeling the maximum. The model focuses on the statistical property of:

$$M_n = Max\{X_1,...,X_n\}$$

Where $X_1,..., X_n$ is a sequence of independent random variables having a common distribution function F. In applications, the X_i usually represent values of a process measured on a regular time-scale- perhaps daily returns of a stock. In theory the distribution of M_n can be derived exactly for all values of n:

$$\Pr\{M_n \leq z\} = \Pr\{X_1 \leq z,..., X_n \leq z\} = \{F(z)\}^n \quad (3.1)$$

However, this is not immediately helpful in practice, since the distribution function F is unknown. One possibility is to use standard statistical techniques to estimate F from observed data, and then to substitute this estimate into (3.1). Unfortunately, very small discrepancies in the estimate of F can lead to substantial discrepancies for the probability distribution defined in (3.1).

An alternative approach is to accept that F is unknown, and to look for approximate families of models defined in (3.1), which can be estimated on the basis of the extreme data only. This is similar to the usual practice of approximating the distribution of sample means by the normal distribution, as justified by the central limit theorem. The following theorem is an extreme value analog of the central limit theory.

Theorem 3.1

If there exists sequences of constants $\{a_n \geq 0\}$ and $\{b_n\}$ such that :

$$\Pr\{(M_n - b_n)/a_n \leq z\} \rightarrow G(z) \quad n \rightarrow \infty$$

for a non-degenerate distribution function G, then G is a member of the Generalized Extreme Value (GEV) family :

$$G(z) = Exp(-(1 + \lambda(\frac{z-\mu}{\sigma}))^{\frac{-1}{\lambda}})$$

Defined on {z : 1+λ(z-μ)/σ>0 }, where -∞<μ<∞ , σ>0 and -∞< λ <∞. □

But it is natural to regard extreme events as those of the X_i that exceed some high threshold u. Denoting an arbitrary term in the X_i sequence by X , it follows that a description of the stochastic behaviour of extreme events is given by he conditional probabilty :

$$\Pr\{X > u + y \mid X > u\} = \frac{1-F(u+y)}{1-F(u)} \quad , y>0$$

If the parent distribution F were known, the distribution of threshold exceedances would be known. Since, in practical applications, this is not the case, approximations that are broadly applicable for high values of the threshold are sought. This parallels the use of the GEV as an approximation to the distribution of maxima long sequences when the parent population is unknown. The following theorem shows how the threshold exceedence are related to the GEV:

Theorem 3.2

Let $X_1,..., X_n$ be a sequence of independent random variables with common distribution function F. Then for large enough u, the distribution function of (X-u) , conditional on X>u is approximately :

$$H(y) = 1-(1+\frac{\theta.y}{\bar{\sigma}})^{\frac{-1}{\theta}} \quad (3.2)$$

Defined on { y: y>0 and $(1+\frac{\theta y}{\bar{\sigma}})$ >0} where $\bar{\sigma} = \sigma + \theta(u-\mu)$

The family of distributions defined by (3.2) is called the generalized Pareto family. So Theorem (3.2) implies that if block maxima have approximating distribution G, then threshold excesses have a corresponding approximate distribution within the generalized Pareto family. Having determined a threshold, the parameters of the generalized Pareto distribution can be estimated by maximum likelihood. Suppose that the values $Y_1,...,Y_n$ are the K excesses of a threshold u. For $\theta \neq 0$ the log-likelihood is derived from (3.2) as:

$$l(\sigma,\theta) = -KLog(\sigma) - (1+1/\theta)\sum_{i=1}^{K}(1+\frac{\theta y_i}{\sigma})^{\frac{-1}{\theta}} \qquad (3.3)$$

provided for i=1,…,K, otherwise $l(\sigma,\theta) = -\infty$.

In the case $\theta = 0$ the log-likelihood is:

$$l(\sigma) = -KLog(\sigma) - \sigma^{-1}\sum_{i=1}^{K} y_i$$

Analytical maximization of the log-likelihood is not possible, so numerical techniques are again required.

4-Empirical results

4-1- Mont Carlo Simulations of two nonparametric methods

The first step is comparing the classical nonparametric method with Locally parametric nonparametric method. Mont Carlo simulation method has been used to compare the two different methods. Sample size of 500 has been chosen for t-distribution with degrees of freedom equal to 5. Then the two methods have been used to find the probability distribution at mean for the sample. 1000 samples have been generated to test the bias of the two methods. For classical nonparametric method the optimal bandwidth has been chosen. The histograms of the values have been shown in Appendix B. The true value is 0.3975. The average bias for LP method is much less than the classical nonparametric approach (0.0005 comparing to -0.0185).

4-2- VaR Estimation

The combination of locally parametric nonparametric method and the extreme value theory has been used to find the VaR of the Vitesse SemiConductor Company (VTSS) and NASDAQ index. The daily returns of VTSS and NASDAQ during April 2003 to April 2004 have been used to find the distribution of returns. The diagram of the returns of VTSS has been plotted in a diagram in Appendix C.

From the Nonparametric method the 10 percent quantile of the distribution function have been found which are as below:

- VTSS = %-5.6 @ %10 VaR
- NASDAQ = %-1.87 @ %10 VaR

These returns have been used as thresholds for finding the parametric functional form (3.2) of distribution function on the left part of tails. The likelihood function defined in (3.3) has been optimized in order to find the parameters. Those functions have been drawn in Appendix D and Appendix E for VTSS and NASDAQ respectively.

4-Conclusion

In this paper, I used the idea of Locally Parametric Nonparametric estimation for finding the VaR. The theory proposed by Hjort and Jones (1996) shows that this method has less bias relative to the standard Kernel estimators. By implementing the normal distribution assumption for functions f and K, the necessary conditions derived from the abstract theory. Mont Carlo simulation method showed that the Locally Parametric Method has less bias relative to the classical nonparametric approach. Then I used the methods to find the VaR of a high tech company.

Much works remain to be done. The difference of two methods should be assessed near the tails. Also the h-effect of kernel method should be considered as a method which gives less bias at tails. The literature in estimation of VaR has been progressing by using factor models to find the distribution at tails. This new approach can give more precise estimates of VaR.

Appendix A – Deriving the locally parameters of Nonparametric estimation

In this section the procedure that the conditions (3.5) and (3.6) are derived from (3.4) will be shown. The assumptions we have used in this papers are:

$f(x;\theta) = (2\pi\sigma^2)^{-1} \exp[-½\sigma^{-2}(x-\mu)^2]$

$K((x-c)/h) = (2\pi)^{-1} \exp[-½h^{-2}(x-c)^2]$

Now if we use the above assumptions, we will have:

$\int (1/h)K((x-c)/h)f(x;\theta)dx =$

$(1/(2\pi\sigma h)) \int \exp[-½(h^{-2}(x-c)^2 + \sigma^{-2}(x-c)^2)]dx$ (A.1)

Now in order to simplify (A.1), the expression in the integral should be simplified. So, we have:

$h^{-2}(x-c)^2 + \sigma^{-2}(x-\mu)^2 = \sigma^{-2}h^{-2}[h^2(x-c)^2 + \sigma^2(x-\mu)^2] =$

$= (h^{-2}+\sigma^{-2})x^2 - 2(ch^{-2} + \mu\sigma^{-2})x + (c^2h^{-2} + \mu^{-2}\sigma^{-2}) =$

$= (h^{-2}+\sigma^{-2})[x^2 - 2(ch^{-2} + \mu\sigma^{-2})(h^{-2}+\sigma^{-2})^{-1} + (c^2h^{-2} + \mu^{-2}\sigma^{-2})(h^{-2}+\sigma^{-2})^{-1}] =$

$= (h^{-2}+\sigma^{-2})[x^2 - 2(c\sigma^2 + \mu h^2)(h^2+\sigma^2)^{-1} + (c^2\sigma^2 + \mu^2h^2)(h^2+\sigma^2)^{-1}] =$

$= (h^{-2}+\sigma^{-2})[(x-(c\sigma^2 + \mu h^2)(h^2+\sigma^2)^{-1})^2 + (c^2\sigma^2 + \mu^2h^2)(h^2+\sigma^2)^{-1} - (c\sigma^2 + \mu h^2)^2(h^2+\sigma^2)^{-2}]$

After some algebraic simplification we will have:

$= \sigma^{-2}h^{-2}(h^{-2}+\sigma^{-2})(x-(c\sigma^2 + \mu h^2)(h^2+\sigma^2)^{-1})^2 + (\mu-c)^2(h^2+\sigma^2)^{-1}$ (A.2)

$\int (1/h)K((x-c)/h)f(x;\theta)dx =$

$(2\pi(h^2+\sigma^2))^{-½} \exp(-½(\mu-c)^2(h^2+\sigma^2)^{-1}) \left\{ (h^2+\sigma^2)^{½}(2\pi h^2\sigma^2)^{-½} \int \exp[-½(h^2+\sigma^2)h^{-2}\sigma^{-2}(x-(c\sigma^2 + \mu h^2)(h^2+\sigma^2)^{-1})^2 dx \right\}$

Since the expression in the bracket is equal to 1, we will have:

$\log \int (1/h)K((x-c)/h)f(x;\theta)dx = -.5(h^2+\sigma^2)^{-1}(\mu-c)^2 - .5\log[2\pi(h^2+\sigma^2)]$ (A.3)

Also we have :

$\log f(x;\theta) = -.5\sigma^{-2}(x-\mu)^2 - .5\log[2\pi\sigma^2]$ (A.4)

Now if we plug (A.3) and (A.4) in (3.4) we will have :

$$(\theta_c\ ,\ \mu_c) = \text{ArgMax}\ \ -.5 \left\{ \sum_{t=1}^{T} (1/h)K((x_t - c)/h)\ [\sigma^{-2} (x-\mu)^2 + \log(2\pi\sigma^2)] - \sum_{t=1}^{T} (1/h)K((x_t - c)/h)\ [(h^2+\sigma^2)^{-1} (\mu - c)^2 + \log(2\pi) + \log(h^2+\sigma^2)\ ] \right\} =$$

$$= \text{ArgMax} \sum_{t=1}^{T} (1/h)K((x_t - c)/h)\ [(h^2+\sigma^2)^{-1} (\mu - c)^2 - \sigma^{-2} (x-\mu)^2 + \log(1+h^2\sigma^{-2})]$$

So if we take the derivative relative to μ we will have:

$$\frac{\partial}{\partial \mu}(.) = \sum_{t=1}^{T} (1/h)K((x_t - c)/h)\ [\sigma_c^{-2}(x_t - \mu_c) + (\sigma_c^{2} + h^2)^{-1} (\mu_c - c)] = 0$$

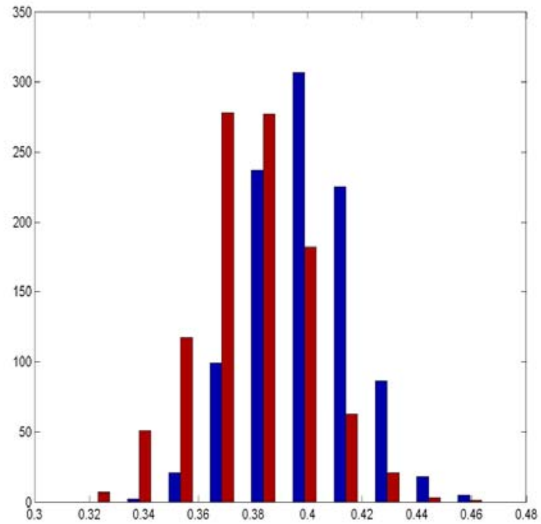
Which is conditioned defined as (3.5) and if we take the derivative relative to σ^2 we will have:

$$\frac{\partial}{\partial \sigma^2}(.) = \sum_{t=1}^{T} (1/h)K((y_t - c)/h)[-(\mu_c - c)^2 + (\sigma_c^{2} + h^2)^{-1} + \sigma_c^{-4} (x_t - \mu)^2 - \sigma_c^{-2}] = 0$$

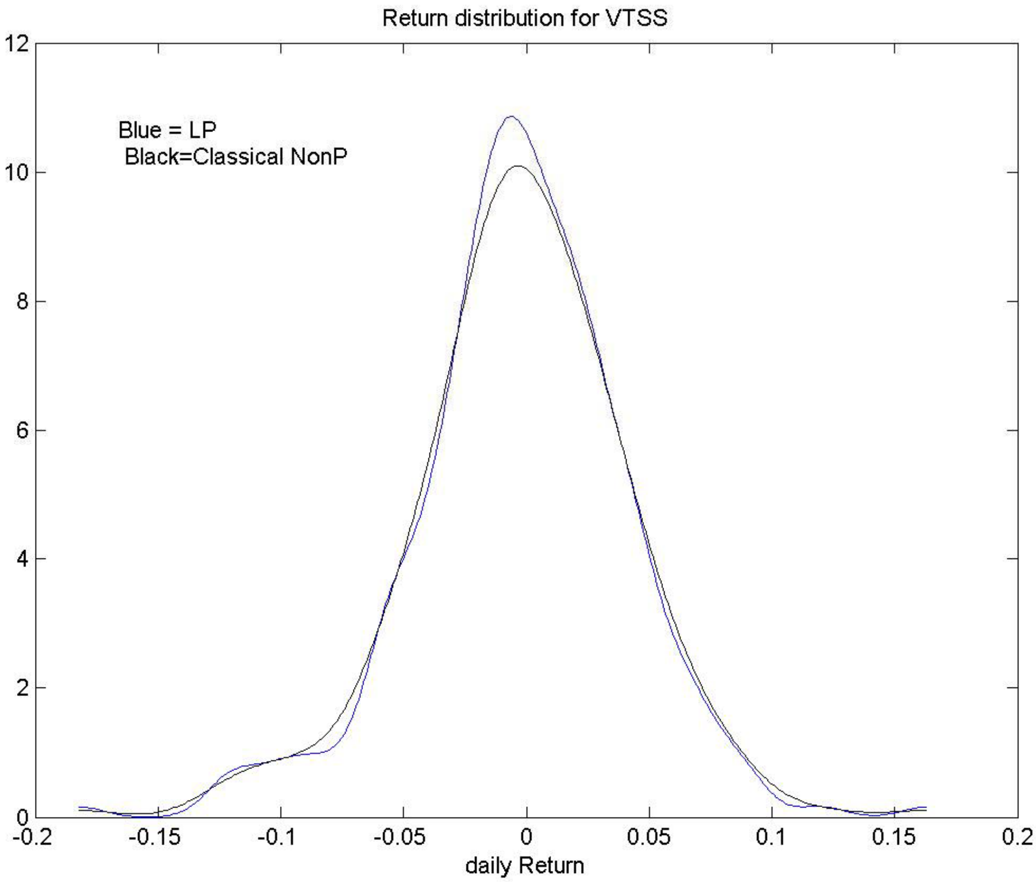
Which is the expression shown as (3.6).

Appendix B

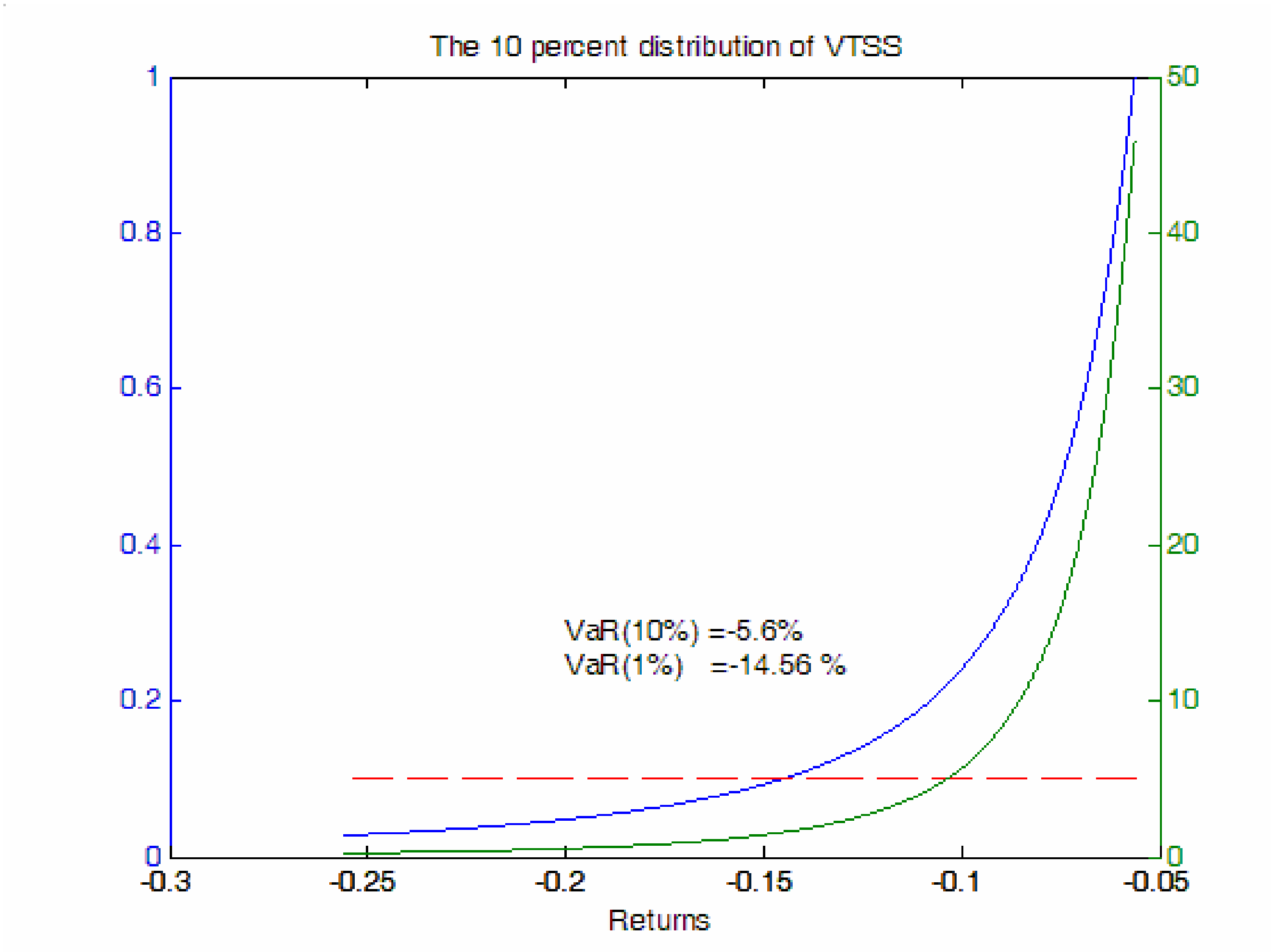
Figure 1. Estimated pdf for two methods
Average bias:
LP=0.0005
Classical=-0.0185



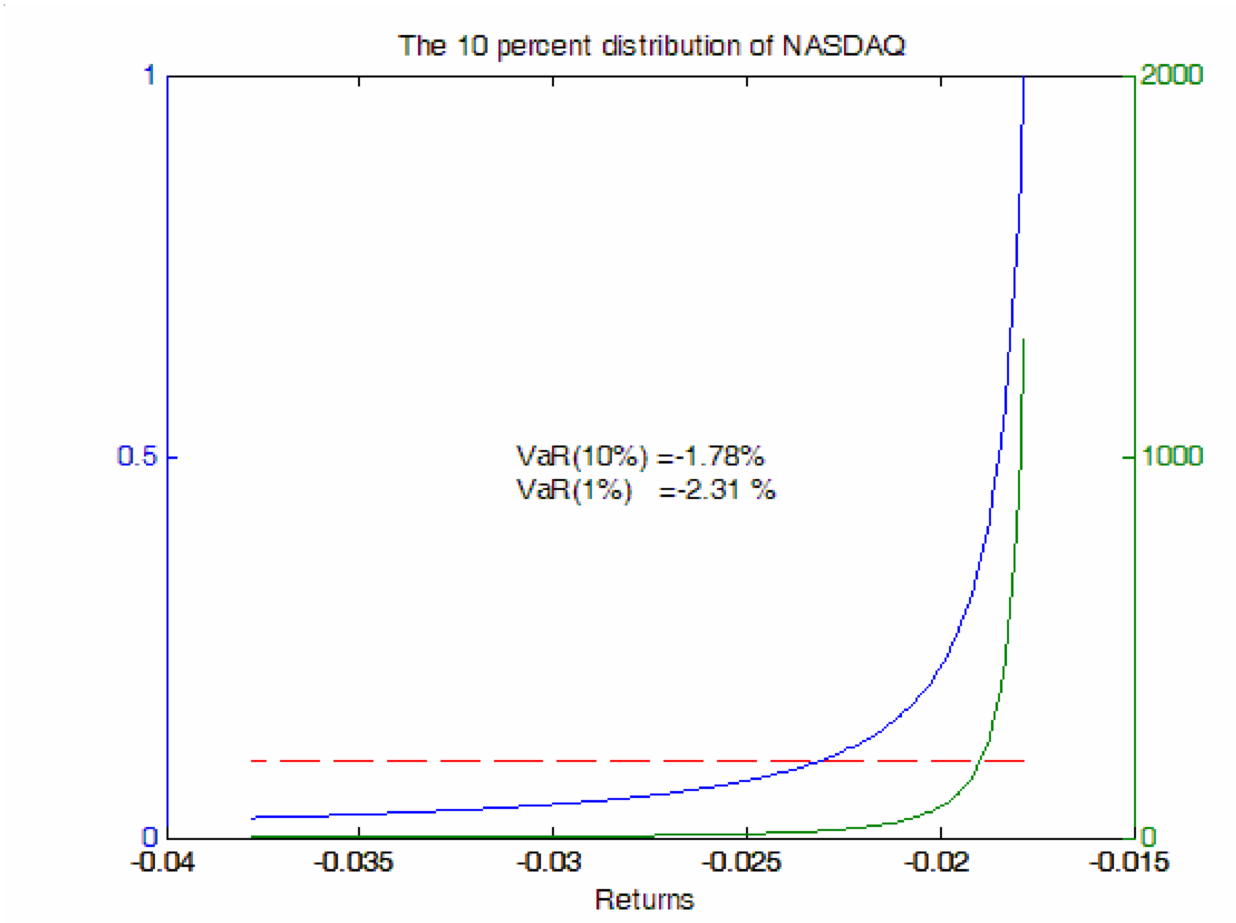
Appendix C- The distribution of VTSS



Appendix D – The Tail Distribution of VTSS



Appendix E – The Tail Distribution of NASDAQ



Appendix F – Programs and their outputs

Nonparam.m

```
tic
clear
clc

unasdaq=0.1
uvtss=0.1

d=wk1read('c:\user\vtssnasdaq.wk1');
vtss=d(:,1);
nasdaq=d(:,2);
ss=size(vtss);
n=ss(1,1);

logvtss=log(vtss);
lognasdaq=log(nasdaq);

for i=1:n-1
   rvtss(i)=logvtss(i+1)-logvtss(i);
   rnasdaq(i)=lognasdaq(i+1)-lognasdaq(i);
end

maxrvtss=max(rvtss);
minrvtss=min(rvtss);
meanrvtss=mean(rvtss);
stdrvtss=std(rvtss);
hrule=1.06*stdrvtss*n^(-0.2);

i=0;
for l=minrvtss:0.001:maxrvtss
   i=i+1;

   for j=1:n-1
      g=(l-rvtss(1,j))/hrule;
      k(j)=pdf('norm',g,0,1);
   end
```

```
    fhrulevtss(i)=mean(k)/(hrule);
    x(i)=l;

end

figure
hist(rvtss)
hold on
plot(x,fhrulevtss,'r-')

r=rnasdaq;
maxrnasdaq=max(rnasdaq);
minrnasdaq=min(rnasdaq);
meanrnasdaq=mean(rnasdaq);
stdrnasdaq=std(rnasdaq);
hrule=1.06*stdrnasdaq*n^(-0.2);

i=0;
for l=minrnasdaq:0.0002:maxrnasdaq
    i=i+1;

    for j=1:n-1
        g=(l-rnasdaq(1,j))/hrule;
        k(j)=pdf('norm',g,0,1);
    end

    fhrulenasdaq(i)=mean(k)/(hrule);
    xx(i)=l;

end

figure
hist(rnasdaq)
hold on
plot(xx,fhrulenasdaq,'r-')
```

```
% In this loop the integral will be calculated
intsum=0;
k=0;
i=1;
while intsum<uvtss

    k=k+1;
    intsum=intsum+0.5*0.001*(fhrulevtss(i)+fhrulevtss(i+1));
    i=i+1;
end
varvtss=minrvtss+k*0.001
kvtss=k;

j=1;
for i=1:n-1
    if rvtss(i)<varvtss
        dvtss(j)=varvtss-rvtss(i);
        j=j+1;
    else
    end
end
pvtss=varvtss-rvtss;
svtss=diag((pvtss>0))*(pvtss');

wk1write('c:\user\datatresh',dvtss);
yt=[.48,.77];
    [v,FVAL,EXITFLAG]=fminsearch(@threshparam,yt,optimset('MaxFunEvals',1220));
    sigmavtss=v(1,1)
    landavtss=v(1,2)


    datvtss=[uvtss;sigmavtss;landavtss];
    wk1write('c:\user\dat',datvtss);
    l=0;
    for i=0:0.0001:0.20
        l=l+1;
        gvtss(l)=fextrem(i);
```

```
        fvtss(l)=(fextrem(i)-fextrem(i+.0001))/0.0001;
        xvtss(l)=varvtss-i;
    end
    sizevtss=l;

    bord10=ones(1,l)*0.1;
    figure;
    plotyy(xvtss,gvtss,xvtss,fvtss)
    hold on
    plot(xvtss,bord10,'r--')
    title(' The 10 percent distribution of VTSS')
    xlabel('Returns')

intsum=0;
k=0;
i=1;
while intsum<unasdaq

    k=k+1;
    intsum=intsum+0.5*0.0002*(fhrulenasdaq(i)+fhrulenasdaq(i+1));
    i=i+1;
end
varnasdaq=minrnasdaq+k*0.0002
knasdaq=k

j=1;
for i=1:n-1
    if rnasdaq(i)<varnasdaq
        dnasdaq(j)=varnasdaq-rnasdaq(i);
        j=j+1;
    else
    end
end

pnasdaq=varnasdaq-rnasdaq;
snasdaq=diag((pnasdaq>0))*(pnasdaq');
```

```
wk1write('c:\user\datatresh',dnasdaq);
yt=[.48,.77];
   [v,FVAL,EXITFLAG]=fminsearch(@threshparam,yt,optimset('MaxFunEvals',1220));
   sigmanasdaq=v(1,1)
   landanasdaq=v(1,2)

   datnasdaq=[unasdaq;sigmanasdaq;landanasdaq];
   wk1write('c:\user\dat',datnasdaq);
   l=0;
   for i=0:0.0001:0.02
      l=l+1;
      gnasdaq(l)=fextrem(i);
      fnasdaq(l)=(fextrem(i)-fextrem(i+0.0001))/0.0001;
      xnasdaq(l)=varnasdaq-i;
   end
   sizenasdaq=l;

   bord10=ones(1,l)*0.1;
   figure;
   plotyy(xnasdaq,gnasdaq,xnasdaq,fnasdaq)
   hold on
   plot(xnasdaq,bord10,'r--')
   title(' The 10 percent distribution of NASDAQ')
   xlabel('Returns')


for i=1:n-1
   if svtss(i)~=0
      vtssbar(i)=-inv(log(1-uvtss*(1+landavtss*svtss(i)/sigmavtss)^(-1/landavtss)));
   else
   vtssbar(i)=0;
   end
end

for i=1:n-1
   if snasdaq(i)~=0
      nasdaqbar(i)=-inv(log(1-unasdaq*(1+landanasdaq*snasdaq(i)/sigmanasdaq)^(-1/landanasdaq)));
```

```
    else
    nasdaqbar(i)=0;
    end
end

datbar=[vtssbar' nasdaqbar'];
wk1write('c:\user\datbar',datbar);
vars=[varvtss varnasdaq];
wk1write('c:\user\vars',vars);

ainit=0.5
[alfa,FVAL,EXITFLAG]=fminsearch(@fexroot,ainit);

figure
plot(rnasdaq,rvtss,'r.')

k=0;
for i=1:200
    xk(i)=varvtss-k;
    x3(i)=-inv(log(1-uvtss*(1+landavtss*k/sigmavtss)^(-1/landavtss)));

    k=k+0.01/5;
end

k=0;
for j=1:200
    yk(j)=varnasdaq-k;
    y3(j)=-inv(log(1-unasdaq*(1+landanasdaq*k/sigmanasdaq)^(-1/landanasdaq)));
    k=k+0.0005;

end
for i=1:200
    for j=1:200
        z3(i,j)=1-exp(-(x3(i)^(-1/alfa)+y3(j)^(-1/alfa))^alfa);
    end
end
```

```
d1=[xk' z3];
d2=[0 yk];
z=[d2 ;d1];
wk1write('c:\user\z',z);

%figure
%surf(yk,xk,z3)
%axis([varvtss varvtss+0.2 varnasdaq varnasdaq+0.025 0 1])

toc
```

Threshparam.m

```
function x=threshparam(p)
sigma=p(1,1);
landa=p(1,2);
data=wk1read('c:\user\datatresh');
nn=size(data);
n=nn(1,1);
for i=1:n
   l(i)=log(1+landa*data(i)/sigma);
end

x=+n*log(sigma)+(1+1/landa)*sum(l);
```

Fexroot.m

```
function xl=fexroot(a)

dat=wk1read('c:\user\datbar');
xx=dat(:,1);%vtssbar
yy=dat(:,2);%nasdaqbar
x=xx';
y=yy';

var=wk1read('c:\user\vars');
ux=var(1,1);
```

```
uy=var(1,2);

nn=size(x);
n=nn(1,2);

for i=1:n
   if (x(i)==0 & y(i)==0)
      l(i)=exp(-(ux^(-1/a)+uy^(-1/a))^a);
   end
   if (x(i)==0 & y(i)~=0)
      l(i)=-((1+a)/a)*log(y(i))+(a-1)*log(ux^(-1/a)+y(i)^(-1/a))-(ux^(-1/a)+y(i)^(-1/a))^a;
   end
   if (x(i)~=0 & y(i)==0)
      l(i)=-((1+a)/a)*log(x(i))+(a-1)*log(x(i)^(-1/a)+uy^(-1/a))-(x(i)^(-1/a)+uy^(-1/a))^a;
   end
   if (x(i)~=0 & y(i)~=0)
      l(i)=-((1+a)/a)*(log(x(i))+log(y(i)))-(x(i)^(-1/a)+y(i)^(-1/a))^a+(a-2)*log(x(i)^(-1/a)+y(i)^(-
1/a))+log(1/a-1+(x(i)^(-1/a)+y(i)^(-1/a))^a);
   end
end

xl=-sum(l);
```

Fextrem.m

```
function x=fextrem(p)

dat=wk1read('c:\user\dat');
u=dat(1,1);
sigma=dat(2,1);
landa=dat(3,1);

x=(1+landa*p/sigma)^(-1/landa);
```

References

Chernozhukov, J. , Len Umantsev (2001), Conditional value-at-risk: Aspects of modeling and estimation, *Empirical Economics.*

Dahan, E. & Mendelson, H.(2001). An extreme value model of concept testing. *Management Science* 47, 102-116.

Danielsson, J and C.G. de Vries (1998), Beyond the sample: extreme quantile and probability estimation, London Scholl of Economics, *Discussion paper 298.*

Dawson, T. H.(2000). Maximum wave crests in heavy seas. *Journal of Offshore Mechanics and Arctic Engineering_Transactions of ASME 122, 222-224.*

Dunne, J. F. & Ghanbari, M. (2001). Effiecient extreme value prediction for nonlinear beam vibrations using measured random response histories. *Nonlinear Dynamics* 24, 143-171.

Engle, R., and S. Mangenelli. (1999). CaViar: Conditional Autoregressive Value at Risk by Regression Quantiles, Discussion paper, University of California, San Diego.

Gourieroux, C., Joan Jasiak , 2000. Local Likelihood Density Estimation and value at Risk, Discussion paper, CREST, Paris.

Gourieroux, C., Joan Jasiak , 2001, Truncated maximum likelihood, goodness of fit tests and tail analysis , Unpublished manuscript.

Gourieroux, C., Joan Jasiak , 2001, Financial Econometrics. *Princeton University Press*.

Harris, R. I. (2001). The accuracy of design values predicted from extreme value analysis. *Journal of Wind Engineering and Industrial Aerodynamics* 89,153-167.

Hjort, N., M. C. Jones (1996), Locally Parametric Nonparametric Density Estimation, *The annals of Statistics*, Vol 24, No 4, 1619-1647.

Kawas, M. L. & Moreira, R. G. (2001). Characterization of prosuct quality attributes of tortilla chips during the frying process. *Journal of Food Engineering* 47, 97-107.

Lavenda, B. H. & Cipollone, E. (2000). Extreme value statistics and thermodynamics of earthquackes:aftershock sequences. Annali di geofisca 43, 967-982.

Loader, C. R. (1993), "Local Likelihood Density Estimation," manuscript, AT&T laboratories, Murray Hill.

McBulty , P. J., Scheick, L. Z., Roth, D. R. , Davis, M. G. ,&Tortora, M. R. S. (2000). First failure predictions for EPROMs of the type flown on the MPTB satellite. *IEEE Transactions on Nuclear Science* 47, 2237-2243.

Roberts,S. I. (2000) Extreme Value statistics for novelty detection in biomedical data processing. *IEE Proceedings-Science Measurement and Technology* 147,363-367.

Scott, D.W. (1992). *Multivariate Density Estimation: Theory, Practice and Visualization*. Wiley, NewYork.

Tryon, R. G. & Cruse, T.A (2000). Probabilistic mseomechanics for high cycle fatigue life prediction. *Journal of Engineering Materials and Technology-Transactions of the AMSE* 122, 209-214.