Order Statistics for Value at Risk Estimation and Option Pricing

Frederik Herzberg and Christoph Bennemann
d-fine GmbH, Opernplatz 2, D-60313 Frankfurt am Main, Germany
Email: *christoph.bennemann@d-fine.de*

Abstract

We apply order statistics to the setting of VaR estimation. Here techniques like historical and Monte Carlo simulation rely on using the k-th heaviest loss to estimate the quantile of the profit and loss distribution of a portfolio of assets. We show that when the *k*-th heaviest loss is used the expected quantile and its error will be independent of the portfolio composition and the return functions of the assets in the portfolio. This is not the case when a linear combination of simulated losses is used. Furthermore, we briefly demonstrate how order statistics can be applied to pricing options depending on the quantile of a distribution.

Keywords

order statistics, value-at-risk, option pricing, historical simulation, Monte Carlo simulation

1 Introduction

One of the techniques most frequently employed to estimate the Value at Risk from historical data is Historical Simulation. Typically, the basis is data from a 251 day-period from which 250 day-to-day relative changes of the portfolio (denoted by $X_1, \ldots, X_{250}$ in the following) can be computed. Ranking them in increasing order, we obtain a sequence of weakly increasing (logarithmic) relative returns $X_{(1)} \leq \ldots \leq X_{(250)}$. The 1%-quantile, usually used to calcualte regulatory capital, of the underlying common distribution of the relative losses would now be approximated by the 1%-quantile of the histogram of observed data, viz. the "2.5th"-heaviest loss (that is, the "2.5th" value in the weakly increasing sequence; we use the term "loss" here in a rather loose sense, since there is always be the possiblity that the loss is actually a gain), if this was defined. Some practitioners tend to take the 3rd-heaviest loss (that is, the 3rd value, $X_{(3)}$, in the weakly increasing sequence)—while others take the 2nd heaviest loss. Finally, some market participants chose the average between the 2nd- and 3rd-heaviest loss, namely $\frac{X_{(2)}+X_{(3)}}{2}$. Initially is not clear, which one of these three estimators will give the most acurate result, and the objective of this paper is to shed some light on this issue by applying the theory of order statistics to the problem.

At the outset of this short paper, we shall give a brief account of the theory of order statistics and an explanation under which assumptions it can be applied to the setting of VaR estimation. Subsequently, we shall employ basic results from this area to derive explicit integral formulae for the expectation and standard deviation of the set of convex combinations of order statistics (as a class of VaR estimators), as well as for the expected *implied Value-at-Risk levels* (defined as the quantile that is really measured by the estimator) associated to such statistics.

But the use of order statistics in mathematical finance is not limited to finding the optimal VaR estimator in historical simulations. Besides applying the results derived for Historical Simulations to the case where Monte Carlo simulations are used for VaR estimation, order statistics can also be used to price option dependending on the quantile of a distribution. An example for such an option would be an option paying the holder a fixed amoount if at the end of a year the best monthly return exceeds a predefined threshold. Within this

Dr Frederik Herzberg is a researcher at the University of Bonn, while Dr Christoph Bennemann works as Senior Manager at d-fine, the consulting firm. Inquiries should be directed to *christoph.bennemann@d-fine.de*. Both authors would like to thank d-fine for supporting this work.

paper we will shortly demonstrate how to price such options using order statistics.

Though formulæ for most quantities will be derived, not all of them are easily computed. And therefore, in a separate article by Naundorf *et al.* (2006) results of a series of Monte Carlo Simulations will be presented, expanding the theoretical results of this paper.

2 Order Statistics and its Application to VaR-Computation

Several monographs have been written on the subject of order statistics (e.g. the classical work by David and Nagaraja (2003)), however, for our purposes, only the very basic idea needs to be recalled, which can be found for instance as part of web lecture notes by Susan Holmes [2].

Suppose $X_1, \ldots, X_n$, for arbitrary $n \in \mathbb{N}$, are independent identically distributed real-valued random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, X_1 (and hence all other X_i, $i \in \{1, \ldots, n\}$) having a continuous distribution density f and distribution function F.

Inductively, the ith order statistic (denoted by $X_{[i]}$), for $i \in \{1, \ldots, n\}$, is defined by

$$X_{[i]} := \min\left(\{X_1, \ldots, X_n\} \setminus \left\{X_{[j]} \ : \ j < i\right\}\right) \wedge \max\{X_1, \ldots, X_n\},$$

that is

$$\begin{aligned} X_{[1]} &:= \min\{X_1, \ldots, X_n\}, \\ X_{[2]} &:= \min\left(\{X_1, \ldots, X_n\} \setminus \left\{X_{[1]}\right\}\right), \\ &\ \ \vdots \\ X_{[n]} &:= \max\{X_1, \ldots, X_n\}. \end{aligned}$$

Note the *a priori* difference between the sequences $\left(X_{[i]}\right)_{i\in\{1,\ldots,n\}}$ and $\left(X_{(i)}\right)_{i\in\{1,\ldots,n\}}$ as defined in Section 1, a difference which will be observable if and only if at least two of the random variables $X_1, \ldots, X_n$ attain the same value: The sequence $\left(X_{(i)}\right)_{i\in\{1,\ldots,n\}}$ is only weakly increasing, whereas the sequence $\left(X_{[i]}\right)_{i\in\{1,\ldots,n\}}$ is strictly increasing until it stabilises in $\max\{X_1, \ldots, X_n\}$.

However, for any two $i, j \in \{1, \ldots, n\}$, the distribution of the random variable $X_i - X_j$ is, due to the independence of X_i and X_j, merely the convolution of the distributions of X_i and $-X_j$, hence also continuous. Thus $\mathbb{P}\left\{X_i - X_j = 0\right\} = 0$, entailing that all $X_1, \ldots, X_n$ are pairwise distinct with probability 1.

As a corollary to this observation, one has for arbitrary $k \in \{1, \ldots, n\}$ and $x \in \mathbb{R}$,

$$\begin{aligned} \mathbb{P}\left\{X_{(k)} \in dx\right\} &= \mathbb{P}\left\{X_{[k]} \in dx\right\} \\ &= \mathbb{P}\left[\bigcup_{\substack{\ell_1,\ldots,\ell_k\in\{1,\ldots,n\} \\ \text{pairwise distinct}}} \left(\left\{X_{\ell_1} \in dx\right\} \cap \bigcap_{i=2}^{k} \left\{X_{\ell_i} < x\right\} \cap \bigcap_{j\notin\{\ell_1,\ldots,\ell_k\}} \left\{X_j \geq x\right\}\right)\right], \end{aligned}$$

hence

$$\begin{aligned} &\mathbb{P}\left\{X_{(k)} \in dx\right\} \\ &\quad = n\mathbb{P}\left[\bigcup_{\substack{1,\ell_2,\ldots,\ell_k\in\{1,\ldots,n\} \\ \text{pairwise distinct}}} \left(\{X_1 \in dx\} \cap \bigcap_{i=2}^{k} \left\{X_{\ell_i} < x\right\} \cap \bigcap_{j\notin\{1,\ell_1,\ldots,\ell_k\}} \left\{X_j \geq x\right\}\right)\right] \\ &\quad = n \cdot \binom{n-1}{k-1}\mathbb{P}\left[\{X_1 \in dx\} \cap \bigcap_{i=2}^{k} \{X_i < x\} \cap \bigcap_{i=k+1}^{n} \{X_i \geq x\}\right] \\ &\quad = n \cdot \binom{n-1}{k-1} f(x)F(x)^{k-1}\left(1 - F(x)\right)^{n-k}. \end{aligned} \tag{1}$$

Where $X \in dx$ is a shorthand for $X \in [x, x + dx]$, and the probability that $X \in dx$ is given by $f(x)$. The first two terms in the final formula represent the number of possible combinations, i.e. there are n possible choices for $X_{(k)}$, and of the remaining $n - 1$ variables $k - 1$ must be smaller than x. The corresponding probability is given by $F(x)$, and conversly the probability that a return is larger (or equal) than x by $1 - F(x)$.

The formula can first of all be used to derive a formula for the distribution function of the kth-heaviest loss, $X_{(k)}$, which we will denote by $F_k(x)$:

$$F_k(x) = n \cdot \binom{n-1}{k-1} \int_{-\infty}^{x} f(y)F(y)^{k-1}\left(1 - F(y)\right)^{n-k} dy. \tag{2}$$

To solve this integral, we apply the transformation $z = F(y)$. Using $dz = f(y)dy$ we obtain the following formula:

$$F_k(x) = n \cdot \binom{n-1}{k-1} \int_{0}^{F(x)} z^{k-1}(1 - z)^{n-k} dz. \tag{3}$$

The $F_k(x)$ can be calculated using the following recursion formula:

$$F_{k+1}(x) = F_k(x) - \binom{n}{k}(F(k))^k\left(1 - F(x)\right)^{n-k} \tag{4}$$

Where $F_0(x) = 1$. Using this formula one easily computes:

$$\begin{aligned} F_1(x) &= 1 - (1 - F(x))^n \\ F_2(x) &= 1 - (1 - F(x))^n - nF(x)(1 - F(x))^{n-1} \end{aligned} \tag{5}$$

The expected *implied VaR-level* associated to the estimator $X_{(k)}$, that is, $\mathbb{E}\left[F\left(X_{(k)}\right)\right]$, can also be computed from the above density formula for $X_{(k)}$, yielding

$$\mathbb{E}\left[F\left(X_{(k)}\right)\right] = n \cdot \binom{n-1}{k-1} \int_{-\infty}^{\infty} f(x)F(x)^k\left(1 - F(x)\right)^{n-k} dx$$

We can apply the same transformation, i.e. $z = F(x)$, we applied to calculate the density function to derive an explicit solution:

$$\begin{aligned} \mathbb{E}[F(X_{(k)})] &= n\binom{n-1}{k-1} \int_{0}^{1} z^k(1 - z)^{n-k} dz = n\binom{n-1}{k-1}\frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)} \\ &= \frac{n!}{(k-1)!(n-k)!}\,\frac{k!(n-k)!}{(n+1)!} = \frac{k}{n+1} \end{aligned} \tag{6}$$

Simlilarly, we can calculate the expected variation of the implied VaR-level:

$$\mathbb{E}\left[\Delta F\left(X_{(k)}\right)^2\right] = \frac{2k+1}{(n+1)^2} \tag{7}$$

Note that neither the implied VaR-level nor its variation (and therefore its error) do depend on the specific parametric function. Whether you use Historical Simulation for a portfolio consisting of a single share or a large portfolio with a complex return function, the expected implied VaR-level and its error will only depend on your choice of k and the number of observations. And it will also be independent of the individual return functions of the underlying assets.

Applying these results to our initial problem, one can calculate that in a Historical Simualtion based on 250 portfolio returns the VaR levels implied by the 2nd- and 3rd-heaviest loss will be approximately (0.80± 0.56)% and (1.20± 0.69)% respectively. So loosely speaking if you use the 2nd-heaviest loss you end up about 20% too conservative on average, while if you use the 3rd-heaviest instead you are about 20% not conservative enough. But in both cases the simulation results will vary by a large amount, implying that a large fraction of the results will lay above the desired 1% quantile. Using (4) the corresponding probabilities can be computed, and one finds that for the second-heaviest loss 28.6% of the simulations will have an implied quantile above 1% and 11.0% above 1.5%. For the 3rd-heaviest loss these values rise to 54.3% and 27.5% respectively, i.e. in almost thiry percent of all simulations one will measure a VaR-level above 1.5%. Note that these computations do not tell us how large the error will be in absolute terms, i.e. whether one is off by one or one million Euro, which will depend on the specific parametric function and the portfolio.

In practice this structural problem of Historical Simulations is amplified by the large memory effect caused by the use of historical time series. Therefore, if today's VaR level is off by a large amount, tommorow's implied-VaR will not be much better since it will be based on almost the same time series.

The question is now, is the average of the 2nd- and 3rd-heaviest loss a better estimator? Unfortunately, the situation is slightly more complicated when you opt for using a linear combination of the $X_{(k)}$ as we demonstrate below.

But before we do this, let us briefly discuss how these results apply to the case where Monte Carlo simulations are used for VaR-estimation. In this setting, the expected implied VaR-level will once again be independent of the specific parametric distribution or portfolio. Furthermore, assuming that you choose $k = \alpha n$ (though our results suggest that $k = \alpha(n+1)$ would be a better choice), for large n the error of the quantile measured in the simulations will be given by:

$$\sqrt{\mathbb{E}\left[\Delta F\left(X_{(\alpha n)}\right)^2\right]} \approx \sqrt{\frac{2\alpha}{n}} \tag{8}$$

Hence, improving the accuracy of the confidence level by one order of magnitude requires an increase in the number of simulations by two orders of magnitude. So convergence is slow, using 1000 simulations to estimate the 1%–quantile the error will be about 0.45% (which translates to a relative error of 45%!), and for 10000 simulations it will only drop to 0.14% (still an 14% relative error). Again, this is the error of the quantile the absolute size of the error in Euro terms will depend on the parametric function.

These formulæ for the density of $X_{(k)}$ can also be used to derive explicit integral formulæ for the expectation and variance of $X_{(k)}$:

$$\begin{aligned}
\mathbb{E}\left[X_{(k)}\right] &= n \cdot \binom{n-1}{k-1} \int_{-\infty}^{\infty} x f(x) F(x)^{k-1} \left(1-F(x)\right)^{n-k} dx, \\
\mathbb{V}\left[X_{(k)}\right] &= n \cdot \binom{n-1}{k-1} \int_{-\infty}^{\infty} x^2 f(x) F(x)^{k-1} \left(1-F(x)\right)^{n-k} dx - \left[X_{(k)}\right]^2 .
\end{aligned} \tag{9}$$

From this, one immediately gets an integral formula for the expectation of $\frac{X_{(k)}+X_{(k+1)}}{2}$:

$$\begin{aligned}
\mathbb{E}\left[\frac{X_{(k)}+X_{(k+1)}}{2}\right] &= \frac{\mathbb{E}\left[X_{(k)}\right]+\mathbb{E}\left[X_{(k+1)}\right]}{2} \\
&= \frac{n}{2}\left(\binom{n-1}{k-1}+\binom{n-1}{k}\right) \\
&\quad \cdot \int_{-\infty}^{\infty} x f(x) F(x)^{k-1} \left(1-F(x)\right)^{n-k-1} \left(1-F(x)+F(x)\right) dx \\
&= \frac{n}{2}\binom{n}{k} \int_{-\infty}^{\infty} x f(x) F(x)^{k-1} \left(1-F(x)\right)^{n-k-1} dx.
\end{aligned} \tag{10}$$

In order to derive results on the implied VaR level and higher moments of the estimator $\frac{X_{(k)}+X_{(k+1)}}{2}$, we again need to find a density formula as in (9) first (this time, the joint density of $X_{(k)}$ and $X_{(k+1)}$): For all $x, y \in \mathbb{R}$ and $k < n$,

$$\begin{aligned}
&\mathbb{P}\left[\left\{X_{(k)} \in dx\right\} \cap \left\{X_{(k+1)} \in dy\right\}\right] = \mathbb{P}\left[\left\{X_{(k)} \in dx\right\} \cap \left\{X_{(k+1)} \in dy\right\}\right] \\
&= \mathbb{P}\left[\bigcup_{\substack{\ell_1,\dots,\ell_{k+1}\in\{1,\dots,n\} \\ \text{pairwise distinct}}} \begin{pmatrix} \left\{X_{\ell_1} \in dx\right\} \cap \bigcap_{i=2}^{k} \left\{X_{\ell_i} < x\right\} \cap \{x < y\} \cap \\ \left\{X_{\ell_{k+1}} \in dy\right\} \cap \bigcap_{j\notin\{\ell_1,\dots,\ell_{k+1}\}} \left\{X_j \geq y\right\} \end{pmatrix}\right],
\end{aligned}$$

hence

$$\begin{aligned}
&\mathbb{P}\left\{\left(X_{(k)}, X_{(k+1)}\right) \in d(x,y)\right\} \\
&= n(n-1)\mathbb{P}\left[\bigcup_{\substack{1,\ell_2,\dots,\ell_k,n\in\{1,\dots,n\} \\ \text{pairwise distinct}}} \begin{pmatrix} \{X_1 \in dx\} \cap \bigcap_{i=2}^{k} \left\{X_{\ell_i} < x\right\} \cap \{x < y\} \cap \\ \left\{X_{\ell_n} \in dy\right\} \cap \bigcap_{j\notin\{1,\ell_2,\dots,\ell_k,n\}} \left\{X_j \geq y\right\} \end{pmatrix}\right] \\
&= n(n-1) \cdot \binom{n-2}{k-1} \mathbb{P}\begin{bmatrix} \{X_1 \in dx\} \cap \bigcap_{i=2}^{k} \{X_i < x\} \cap \{x < y\} \cap \\ \bigcap_{i=k+1}^{n-1} \{X_i \geq y\} \cap \{X_n \in dy\} \end{bmatrix} \\
&= n(n-1) \cdot \binom{n-2}{k-1} \chi_{\{(u,v)\,:\,u<v\}}(x,y) f(x) f(y) F(x)^{k-1} \\
&\quad \times \left(1-F(y)\right)^{n-k-1} \; dx\, dy.
\end{aligned} \tag{11}$$

The distribution of $X_{(k)} + X_{(k+1)}$ can thus be computed as

$$
\begin{aligned}
&\mathbb{P}\left\{X_{(k)} + X_{(k+1)} < z\right\} \\
&= \int_{\{(u,v)\in\mathbb{R}^2\,:\,u<v,\quad u+v<z\}} (x+y)\mathbb{P}\left\{\left(X_{(k)}, X_{(k+1)}\right) \in d(x,y)\right\} \\
&= n(n-1)\binom{n-2}{k-1}\int_{-\infty}^{z}\int_{x}^{z-x}(x+y)f(x)f(y)F(x)^{k-1} \\
&\quad\times(1-F(y))^{n-k-1}\,dy\,dx \\
&= n(n-1)\binom{n-2}{k-1}\int_{-\infty}^{z}\left(\int_{x}^{z-x}(x+y)f(y)\,(1-F(y))^{n-k-1}\,dy\right) \\
&\quad\times f(x)F(x)^{k-1}\,dx
\end{aligned}
$$

for all $z \in \mathbb{R}$, which implies

$$
\begin{aligned}
&\mathbb{P}\left\{X_{(k)} + X_{(k+1)} \in dz\right\} \\
&= n(n-1)\binom{n-2}{k-1}\begin{pmatrix} z\int_{-\infty}^{z} f(x)F(x)^{k-1}f(z-x)\,(1-F(z-x))^{n-k-1}\,dx \\ +f(z)F(z)^{k-1}\int_{z}^{0}(z+y)f(y)\,(1-F(y))^{n-k-1}\,dy\end{pmatrix} dz
\end{aligned}
\tag{12}
$$

In order to be entitled to apply the results of this Section to the sequence of day-to-day relative changes of the portfolio value, we need to suppose these changes to be independent and identically distributed. This property of the portfolio value process would follow, for instance, if one imposed the assumption of it being a log-Lévy process.

So, by means of (10), we have found a formula to compute the expected arithmetic mean of the second and third-heaviest loss and can compare this to the exact 1%-quantile of the underlying distribution for the relative portfolio change. Thanks to (12), we now obtain formuale for the implied VaR level and the variance of $\frac{X_{(k)}+X_{(k+1)}}{2}$:

$$
\begin{aligned}
&\left[F\left(\frac{X_{(k)}+X_{(k+1)}}{2}\right)\right] \\
&\quad= \frac{n(n-1)}{4}\binom{n-2}{k-1}\cdot \\
&\quad\int_{\mathbb{R}} F(z/2)\begin{pmatrix} z\int_{-\infty}^{z} f(x)F(x)^{k-1}f(z-x)\,(1-F(z-x))^{n-k-1}\,dx \\ +f(z)F(z)^{k-1}\int_{z}^{0}(z+y)f(y)\,(1-F(y))^{n-k-1}\,dy\end{pmatrix} dz
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{V}\left[\frac{X_{(k)}+X_{(k+1)}}{2}\right] \\
&= \frac{n(n-1)}{4}\binom{n-2}{k-1}\int_{\mathbb{R}}\begin{pmatrix} z^3\int_{-\infty}^{z} f(x)F(x)^{k-1}f(z-x)\,(1-F(z-x))^{n-k-1}\,dx \\ +z^2 f(z)F(z)^{k-1}\int_{z}^{0}(z+y)f(y)\,(1-F(y))^{n-k-1}\,dy\end{pmatrix} dz \\
&\quad-\left|\frac{n}{2}\binom{n}{k}\int_{-\infty}^{\infty} xf(x)F(x)^{k-1}\,(1-F(x))^{n-k-1}\,dx\right|^2
\end{aligned}
$$

Unfortunately, the independence of the X_i, $i \in \{1,\ldots,n\}$, does not imply the independence of $X_{(k)}, X_{(k+1)}$—otherwise, Bienaymé's identity could have been applied to derive a simpler expression for $\mathbb{V}\left[\frac{X_{(k)}+X_{(k+1)}}{2}\right]$ in terms of f and F.

In contrast to the VaR estimators based on the kth-heaviest loss, this time the implied VaR level does depend on the parametric function. And calculating it requires solving the above integral. Only for simple cases this integral can be computed analytically, but even then the calculatons are cumbersome. If the X_i are distributed uniformly between 0 and 1 (the simplest case) one finds:

$$
\mathbb{E}\left[F\left(\frac{X_{(k)}+X_{(k+1)}}{2}\right)\right] = \frac{1}{2}\frac{2k+1}{n+1} \tag{13}
$$

Hence, in this case, the implied VaR level is given by the average of the VaR levels implied by the kth and and $(k+1)$st-heaviest loss. That this is not a general result can be seen if one assumes the X_i to be exponential distributed. In this case one obtains:

$$
\begin{aligned}
\mathbb{E}\left[F\left(\frac{X_{(k)}+X_{(k+1)}}{2}\right)\right] &= 1 - \frac{(n-k)(n-k+1)}{(n+1)(n-k+\frac{1}{2})} \\
&= \mathbb{E}\left[F\left(X_{(k)}\right)\right] + \frac{n-k+1}{2(n+1)(n-k+\frac{1}{2})}
\end{aligned}
\tag{14}
$$

The only general statement that can be made about the implied VaR level is that it will be bounded by the VaR levels implied by the kth and $(k+1)$st-heaviest loss, because:

$$
\begin{aligned}
&X_{(k)} \le \frac{X_{(k)}+X_{(k+1)}}{2} \le X_{(k+1)} \\
&\implies F\left(X_{(k)}\right) \le F\left(\frac{X_{(k)}+X_{(k+1)}}{2}\right) \le F\left(X_{(k+1))}\right) \\
&\implies \mathbb{E}\left[F\left(X_{(k)}\right)\right] \le \left[F\left(\frac{X_k+X_{(k+1)}}{2}\right)\right] \le \mathbb{E}\left[F\left(X_{(k+1)}\right)\right]
\end{aligned}
$$

Returning to the problem which spurred our interest in order statistics, namely whether the average of the 2nd- and 3rd-heaviest will in general be a better VaR estimator, we can only conclude that it will be more conservative than the 3rd, but it might still not be conservative enough.

Note that it was only for the sake of simplicity that we have confined ourselves to studying the estimator $\frac{X_{(k)}+X_{(k+1)}}{2}$ as the most "sophisticated" object here—in a similar vein, arbitrary convex (and even general linear) combinations of order statistics can be investigated as well.

3 Option Pricing with Order Statistics

Let us assume that you observe the performance of a share (or any other financial asset) over N periods of equal length, e.g. 12 months, and X_i is the logarithmic performance of the share in the ith period, i.e.:

$$
X_i = \ln\left(\frac{S_{i+1}}{S_i}\right),
$$

where S_i is the price of the share at the beginning of the ith period. Assuming that the X_i are independent identically normal distributed, we know from standard option pricing theory that we have to use the risk neutral measure if wish to price any option dependent on the X_i. Hence,

>

we assume:

$$X_i \sim N\left(\left(r - \frac{\sigma^2}{2}\right)\Delta t, \sigma\,\Delta t^{\frac{1}{2}}\right),$$

where σ is the annualized volatility of the share, r the short rate, and Δt the length of each period. The distribution function $F(X_i)$ of the X_i is given by:

$$\begin{aligned} F(X_i) &= \frac{1}{2\pi}\int_{-\infty}^{Z_i} e^{-\frac{z^2}{2}}dz \equiv N(Z_i) \\ Z_i &= \frac{X_i - (r - \frac{1}{2}\sigma^2)\Delta t}{\sigma\,\Delta t^{\frac{1}{2}}} \end{aligned} \tag{15}$$

The results of the last section can be used to price various options dependent on the kth-heaviest loss, $X_{(k)}$, of the time series. And for binary options it is even possible to derive analytical solutions. The value PV_k of a binary option paying one currency unit if the kth-heaviest loss is above the strike level X (i.e. the option will pay one currency unit if the kth worst performance is better than X, which itself will be a logarithmic-performance) is given by:

$$PV_k = e^{-rT}\mathbb{E}(H(X_{(k)} - X)) = e^{-rT}\left(1 - F_k(X)\right), \tag{16}$$

where H is the Heavyside function, i.e. $H(x)$ is 1 if $x > 0$ and zero otherwise, and F_k is given by (3).

As an example, let us take an option paying one currency unit if the worst monthly (logarithmic) performance in the next 12 months lies above the strike level X. Using (5) we compute the value of the option to be:

$$PV_1 = e^{-rT}\left(1 - F(X)\right)^{12} \tag{17}$$

Note that in practice one would probably use a relative rather than a logarithmic strike level, but since it is easy to transform one into the other this is not really a challenge. More challenging is to calculate the option value during its lifetime, where the performance in past months and in the current month (month to date performance) have to be taken into account. Some simple cases are easily solvable, e.g. for a binary call option on the best monthly performance you only have to work out the probability to exceed the strike level in the current month as well as the probability of exceeding it in one of the remaining months, if you happen to miss it in the current. A full discussion of this topic is beyond the scope of this paper.

4 Discussion and Conclusion

The theory of order statistics has various applications in mathematical finance. In this paper we used it too calculate various properties of the quantile measured in Historical Simulations (and Monte Carlo simulations). We found that the VaR level implied by the kth-heaviest loss (the expected quantile) does not depend on the actual parametric function, which in particular means that it is independent of the portfolio and the return functions of the underlying assets. In a Historical Simulation based on 250 relative portfolio returns the 2nd-heaviest loss implies a VaR level of approximately 0.8% and the 3rd-heaviest loss of 1.2%. So neither is a perfect estimator, because one is (looseley speaking) 20% too conservative, while the other is about 20% not conservative enough. Worse, both values vary widely, and when the 2nd-heaviest loss is used the fraction of results effectively yielding a quantile above the desired 1% level is 28.6%, which rises to 54.3 % if the 3rd-heaviest loss is used.

Unfortunately, order statistics does not allow us to check whether the VaR level implied by the average of the 2nd- and 3rd-heaviest loss is in general a better estimator, since it will depend on the parametric distribution function. However, since it is bounded by the VaR levels implied by the 2nd- and 3rd-heaviest loss, we know that at least it will be a more conservative estimator than the 3rd-heaviest loss.

Note that there are plenty of other applications of order statistics in mathematical finance. For example, Pritsker (1997) uses order statistics to calculate errors on VaR. And order statistics can also be applied to price options, as we have briefly demonstrated above (again there are other examples, e.g. Akahori (1995)).

REFERENCES

- H A David, H N Nagaraja, *Order statistics*, 3rd ed, John Wiley & Sons, Chichester 2003.
- S Holmes, *Order Statistics*, http://www-stat.stanford.edu/~susan/courses/s116/node79.html, http://www-stat.stanford.edu/~susan/courses/s116/node80.html.
- B Naundorf, J Raufeisen, C Bennemann, F S Herzberg, *You cannot always get what you want—estimating the Value at Risk from historical data with limited statistics*, Wilmott 26 (2006), Frankfurt 2006.
- M Pritsker *Evaluating Value at Risk Methodologies: Accuracy versus Computational Time*, Board of Governors of the Federal Reserve System, July 1997.
- J Akahori, *On Dassios Formula of Brownian Quantiles, Working Paper (1995).*