“Extreme Value Theory as a Risk Management Tool”
By Paul Embrechts, Sidney I. Resnick, and Gennady Samorodnitsky
North American Actuarial Journal, Volume 3, Number 2, April 1999

Copyright 2010 by the Society of Actuaries, Schaumburg, Illinois.
Posted with permission.
EXTREME VALUE THEORY AS A RISK MANAGEMENT TOOL*

Paul Embrechts,† Sidney I. Resnick,‡ and Gennady Samorodnitsky§

ABSTRACT

The financial industry, including banking and insurance, is undergoing major changes. The (re)insurance industry is increasingly exposed to catastrophic losses for which the requested cover is only just available. An increasing complexity of financial instruments calls for sophisticated risk management tools. The securitization of risk and alternative risk transfer highlight the convergence of finance and insurance at the product level. Extreme value theory plays an important methodological role within risk management for insurance, reinsurance, and finance.

1. INTRODUCTION

Consider the time series in Table 1 of loss ratios (yearly data) for earthquake insurance in California from 1971 through 1993. The data are taken from Jaffe and Russell (1996).

On the basis of these data, who would have guessed the 1994 value of 2272.7? Indeed, on the 17th of January of that year the 6.6-Richter-scale Northridge earthquake hit California, causing an insured damage of $10.4 billion and a total damage of $30 billion, making 1994 the year with the third highest loss burden (natural catastrophes and major losses) in the history of insurance. The front-runners are 1992 (the year of hurricane Andrew) and 1990 (the year of the winter storms Daria and Vivian). For details on these, see Sigma (1995, 1997).

The reinsurance industry experienced a rise in both intensity and magnitude of losses due to natural and man-made catastrophes. For the United States alone, Canter, Cole, and Sandor (1996) estimate an approximate $245 billion of capital in the insurance and reinsurance industry to service a country that has $25–30 trillion worth of property. It is no surprise that the finance industry has seized upon this by offering (often in joint ventures with the (re)insurance world) properly securitized products in the realm of catastrophe insurance. New products are being born at an increasing pace. Some of them have only a short life, others are reborn under a different shape, and some do not survive. Examples include:

- Catastrophe (CAT) futures and PCS options (Chicago Board of Trade). In these cases, securitization is achieved through the construction of derivatives written on a newly constructed industry-wide loss-ratio index.
- Convertible CAT bonds. The Winterthur convertible hail-bond is an example. This European-type convertible has an extra coupon payment contingent on the occurrence of a well-defined catastrophic (CAT) event: an excessive number of cars in Winterthur’s...
Swiss portfolio damaged in a hail storm over a specific time period. For details, see Schmock (1997). Further interesting new products are the multiline, multiyear, high-layer (infrequent event) products, credit lines, and the catastrophe risk exchange (CATEX). For a brief review of some of these instruments, see Punter (1997). Excellent overviews stressing the financial engineering of such products are Doherty (1997) and Tilley (1997). Alternative risk transfer and securitization have become major areas of applied research in both the banking and insurance industries. Actuaries are actively taking part in some of the new product development and therefore have to consider the methodological issues underlying these and similar products.

Also, similar methods have recently been introduced into the world of finance through the estimation of value at risk (VaR) and the so-called shortfall; see Bassi, Embrechts, and Kafetzaki (1998) and Embrechts, Samorodnitsky, and Resnick (1998). "Value At Risk for End-Users" (1997) contains a recent summary of some of the more applied issues. More generally, extremes matter eminently within the world of finance. It is no coincidence that Alan Greenspan, chairman of the U.S. Federal Reserve, remarked at a research conference on risk measurement and systemic risk (Washington, D.C., November 1995) that "Work that characterizes the statistical distribution of extreme events would be useful, as well."

For the general observer, extremes in the realm of finance manifest themselves most clearly through stock market crashes or industry losses. In Figure 1, we have plotted the events leading up to and including the 1987 crash for equity data (S&P). Extreme value theory (EVT) yields methods for quantifying such events and their consequences in a statistically optimal way. (See McNeil 1998 for an interesting discussion of the 1987 crash example.) For a general equity book, for instance, a risk manager will be interested in estimating the resulting down-side risk, which typically can be reformulated in terms of a quantile for a profit-and-loss function.

EVT is also playing an increasingly important role in credit risk management. The interested reader may browse J.P. Morgan's web site (http://www.jpmorgan.com) for information on CreditMetrics. It is no coincidence that big investment banks are looking at actuarial methods for the sizing of reserves to guard against future credit losses. Swiss Bank Corporation, for instance, introduced actuarial credit risk accounting (ACRA) for credit risk management; see Figure 2. In their risk measurement framework, they use the following definitions:

- Expected loss: the losses that must be assumed to arise on a continuing basis as a consequence of undertaking particular business
- Unexpected loss: the unusual, though predictable, losses that the bank should be able to absorb in the normal course of its business
- Stress loss: the possible—although improbable—extreme scenarios that the bank must be able to survive.

EVT offers an important set of techniques for quantifying the boundaries between these different loss classes. Moreover, EVT provides a scientific language for translating management guidelines on these
boundaries into actual numbers. Finally, EVT helps in the
tooling of diversification factors and the estimation of
diversification factors in the management of bond portfolios. Many more examples can be added.

It is our aim in this paper to review some of the
tools from EVT relevant for industry-wide integrated risk management. Some examples toward the end of this paper will give the reader a better idea of the kind of answers EVT provides. Most of the material covered here (and indeed much more) is found in Embrechts, Klüppelberg, and Mikosch (1997), which also contains an extensive list of further references. For reference to a specific result in this book, we will occasionally identify it as “EKM.”

2. The Basic Theory

The statistical analysis of extremes is key to many of the risk management problems related to insurance, reinsurance, and finance. In order to review some of the basic ideas underlying EVT, we discuss the most important results under the simplifying iid assumption: losses will be assumed to be independent and identically distributed. Most of the results can be extended to much more general models. In Section 4.2 a first indication of such a generalization will be given.

Throughout this paper, losses will always be denoted as positive; consequently we concentrate in our discussion below on one-sided distribution functions (df’s) for positive random variables (rv’s).

Given basic loss data

\[ X_1, X_2, \ldots, X_n \] iid with \( F \),

we are interested in the random variables

\[ X_{\max} = \min(X_1, \ldots, X_n), X_{\min} = \max(X_1, \ldots, X_n). \]

\[ X_{\max} \leq X_{\min} \leq \cdots \leq X_{k,n}. \]

we may be interested in

\[ \sum_{r=1}^{k} h(X_{r,n}), \]

for certain functions \( h \). The basic assumption yields that

\[ X_1, \ldots, X_{100} \] are iid with \( P(X_1 \leq x) \)

\[ = 1 - e^{-x/10}, x \geq 0. \]

Therefore, for \( M_n = \max(X_1, \ldots, X_n) \),

\[ P(M_{100} > x) = 1 - (P(X_1 \leq x))^{100} \]

\[ = 1 - (1 - e^{-x/10})^{100}. \]

From this, we immediately obtain

\[ P(M_{100} \geq 50) = 0.4914, \]

\[ P(M_{100} \geq 100) = 0.00453. \]
However, rather than doing the (easy) exact calculations above, consider the following asymptotic argument. First, for all $n \geq 1$ and $x \in \mathbb{R}$,
\[
P\left(\frac{M_n}{10} - \log n \leq x\right) = P(M_n \leq 10(x + \log n)) = \left(1 - e^{-x}\right)^n,
\]
so that
\[
\lim_{n \to \infty} P\left(\frac{M_n}{10} - \log n \leq x\right) = e^{-x} = \Lambda(x).
\]
Therefore, use the approximation
\[
P(M_n \leq x) = \Lambda\left(\frac{x}{10} - \log n\right)
\]
to obtain
\[
P(M_{100} \leq 50) = 0.4902,
\]
\[
P(M_{100} \geq 100) = 0.00453,
\]
very much in agreement with the exact calculations above.

Suppose we were asked the same question but had much less specific information on $F(x) = P(X_1 \leq x)$; could we still proceed? This is exactly the point where classical EVT enters. In the above exercise, we have proved the following.

Proposition 1

Suppose $X_1, \ldots, X_n$ are iid with df $F = \exp(\lambda x)$, then for $x \in \mathbb{R}$,
\[
\lim_{n \to \infty} P(\lambda M_n - \log n \leq x) = \Lambda(x).
\]

Here are the key questions:

Q1: What is special about $\Lambda$? Can we get other limits, possibly for other df’s $F$?

Q2: How do we find the norming constants $\lambda$ and $n$ in general—that is, find $a_n$ and $b_n$ so that
\[
\lim_{n \to \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right)
\]
exists?

Q3: Given a limit coming out of Q1, for which df’s $F$ and norming constants from Q2, do we have convergence to that limit? Can one say something about second order behavior, that is, speed of convergence?

The solution to Q1 forms part of the famous Gne- denko, Fisher-Tippett theorem.

Theorem 2 (BKM Theorem 3.2.3)

Suppose $X_1, \ldots, X_n$ are iid with df $F$ and $(a_n), (b_n)$ are constants so that for some nondegenerate limit distribution $G$,
\[
\lim_{n \to \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = G(x), \quad x \in \mathbb{R}.
\]

Then $G$ is of one of the following types:

- Type I (Fréchet):
  \[
  \Phi(x) = \begin{cases} 
  0, & x \leq 0 \\
  \exp\left(\frac{-x}{\alpha}\right), & x > 0 \quad \alpha > 0.
  \end{cases}
  \]

- Type II (Weibull):
  \[
  \Phi(x) = \begin{cases} 
  \exp\left(-\frac{x}{\alpha}\right), & x \leq 0 \\
  1, & x > 0 \quad \alpha > 0.
  \end{cases}
  \]

- Type III (Gumbel):
  \[
  \Lambda(x) = \exp\left(-\frac{e^{-x}}{\xi}\right), \quad x \in \mathbb{R}.
  \]

$G$ is of the type $H$ means that for some $a > 0, b \in \mathbb{R}$, $G(x) = H((x - b)/a), x \in \mathbb{R}$, and the distributions of one of the above three types are called extreme value distributions. Alternatively, any extreme value distribution can be represented as
\[
H_{\xi,\sigma}(x) = \exp\left(-\left(1 + \frac{x}{\sigma}\right)^{-1/\xi}\right), \quad x \in \mathbb{R}.
\]

Here $\xi \in \mathbb{R}, \sigma \in \mathbb{R}$, and $\sigma > 0$. The case $\xi > 0 (\xi < 0)$ corresponds to the Fréchet (Weibull)-type df with $\xi = 1/\alpha (\xi = -1/\alpha)$, whereas by continuity $\xi = 0$.

Some Examples of Extreme Value Distributions

Figure 3

\[
\begin{align*}
\text{Fréchet} & \quad \text{Weibull} \\
\text{Gumbel} & \quad \text{Fréchet} \\
\end{align*}
\]

\[
\begin{align*}
0.6 & \quad 0.5 \\
0.4 & \quad 0.3 \\
0.2 & \quad 0.1 \\
0.0 & \quad 0.0 \\
\end{align*}
\]

\[
\begin{align*}
-5 & \quad -4 \\
-3 & \quad -2 \\
-1 & \quad 0 \\
1 & \quad 2 \\
3 & \quad 4 \\
5 & \quad 6 \\
\end{align*}
\]
corresponds to the Gumbel, or double exponential-type, df.

In Figure 3, some examples of the extreme value distributions are given. Note that the Fréchet case (the Weibull case) corresponds to a model with finite lower (upper) bound; the Gumbel case is two-sided unbounded.

Answering Q2 and Q3 is much more complicated. Below we formulate a complete answer (due to Gne-ndenko) for the Fréchet case. This case is the most important for applications to (re)insurance and fin-
ance being that for

\[ F \] $\left( u \right) = \inf \left\{ x \in \mathbb{R} : x \leq t \right\}, \quad 0 < t < 1. \]

Using this notation, the p-quantile of F is defined as

\[ x_p = F^{-1}(p), \quad 0 < p < 1. \]

Theorem 3 (EKM Theorem 3.3.7)

Suppose \( X_0, \ldots, X_n \) are iid with df F satisfying

\[ \lim_{t \to \infty} \frac{1 - F(t)}{1 - F(t)} = x^-, \quad x > 0, \quad a > 0. \tag{5} \]

Then for \( x > 0, \)

\[ \lim_{t \to \infty} \frac{M_x - b_i}{a_i} = x_- \]

where \( b_i = 0 \) and \( a_i = F^{-1}(1 - 1/n). \) The converse of this result also holds true.

A df F satisfying (5) is called regularly varying with index \( -\alpha, \) denoted by \( F \in \mathbb{R} \). An important consequence of the condition \( F \in \mathbb{R}_+ \) is that for a rv \( X \) with df \( F, \)

\[ \text{EX}^\alpha \left\{ \right. \]

\[ \left. \begin{array}{ll}
\text{for } \beta < \alpha, \\
\text{for } \beta > \alpha.
\end{array} \right\} \tag{6} \]

In insurance applications, one often finds \( \beta \)-values in the range (1, 2), whereas in finance (equity daily log-
returns, say) an interval (2, 5) is common. Theorem 3 is also reformulated thus: The maximal domain of attraction of \( F \) is that of a df \( F \in \mathbb{R}_+ \), which is

\[ \text{MDA}(F) = \mathbb{R}_+. \]

Df’s belonging to \( \mathbb{R}_+ \) are for obvious reasons also called Pareto type. Though we can calculate the norm-
ing constants, the calculation of \( a_0 \) depends on the tail of \( F, \) which in practice is unknown. The construc-
tion of MDA (\( \mathbb{R}_+ \)) is also fairly easy, the main difference being that for \( F \in \text{MDA}(\mathbb{R}_+), \)

\[ x_r = \sup \{ x \in \mathbb{R} : F(x) < 1 \} < \infty. \]

The analysis of MDA(\( \lambda \)) is more involved. It contains such diverse df’s as the exponential, normal, lognor-
mal, and gamma. For details see Embrechts, Klüppel-
berg, and Mikosch (1997, Section 3.3.3).

3. Tail and Quantile Estimation

Theorem 3 is the basis of EVT. In order to show how this theory can be put into practice, consider, for in-
stance, the pricing of an XL treaty. Typically, the prior-
ity (or attachment point) \( u \) is determined as a t-year event corresponding to a specific claim event with claim size df \( F, \) for example. This means that

\[ u = u_t = F^{-1} \left( \frac{1 - \frac{1}{t}}{t} \right). \tag{7} \]

In our notation used before, \( u_t = x_{1-1/t} \). Whenever \( t \)

is large—typically the case in the catastrophic, that is, rare, event situation—the following result due to

Balkema, de Haan, Gnedenko, and Pickands (see EKM, Theorem 3.4.13(b)) is very useful.

Theorem 4

Suppose \( X_1, \ldots, X_n \) are iid with df F. Equivalent are:

i) \( F \in \text{MDA}(\mathbb{R}_+), \) \( \xi \in \mathbb{R}_+ \),

ii) for some function \( \beta : \mathbb{R}_+ \to \mathbb{R}_+ ; \)

\[ \lim_{u \to \infty} \sup_{0 < x < \xi < \infty} |F(x) - G_{\xi,\beta}(x)| = 0. \tag{8} \]

where \( F(x) = P(X - u \leq x | X > u), \) and the gen-
eralized Pareto df is given by

\[ G_{\xi,\beta}(x) = 1 - \left( 1 + \frac{x}{\xi} \right)^{-1/\beta}, \tag{9} \]

for \( \beta > 0. \)

It is exactly the so-called excess df \( F_x \) that risk man-
gers as well as reinsurers should be interested in.

Theorem 4 states that for large \( u, F_u \) has a generalized Pareto df (9). Now, to estimate the tail \( F(u + x) \) for a fixed large value of \( u \) and all \( x \) \( \equiv 0, \) consider the trivial identity

\[ F(u + x) = F(u) F_x(x), \quad u, x \geq 0. \tag{10} \]

In order to estimate \( F(u + x), \) one first estimates \( F(u) \) by the empirical estimator

\[ \left( F^\ast(u) \right)^n = \frac{N_u}{n}, \]

where \( N_u = \# \left( 1 \leq i \leq n : X_i > u \right). \) In order to have a “good” estimator for \( F(u), \) we need \( u \) not too large:
the level \( u \) has to be well within the data. Given such a \( u \)-value, we approximate \( F_{\xi}(x) \) via (8) by

\[
(F_{\xi}(x)) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{(x > u)}(X_i)
\]

for some estimators \( \xi \) and \( \beta(u) \), depending on \( u \). For this to work well, we need \( u \) large (indeed, in Theorem 4(ii), \( u \mid x \), the latter being \( \rightarrow \) in the Frechet case). A "good" estimator is obtained via a trade-off between these two conflicting requirements on \( u \).

The statistical theory developed to work out the above program runs under the name Peaks over Thresholds Method and is discussed in detail in Embrechts, Klüppelberg, and Mikosch (1997, Section 6.5), McNeil and Saladin (1997), and references therein. Software (S-plus) implementation can be found at http://www.math.ethz.ch/~mcneil/software.

This maximum-likelihood-based approach also allows for modeling of the excess intensity \( N_u \), as well as the modeling of time (or other co-variable) dependence in the relevant model parameters. As such, a highly versatile modeling methodology for external events is available. Related approaches with application to insurance are to be found in Beirlant, Teugels, and Vynckier (1996), Reiss and Thomas (1997), and the references therein. Interesting case studies using up-to-date EVT methodology are McNeil (1997), Resnick (1997), and Rootzen and Tajvidi (1997). The various steps needed to perform a quantile estimation within the above EVT context are nicely reviewed in McNeil and Saladin (1997), where a simulation study is also found. In the next section, we illustrate the methodology on real and simulated data relevant for insurance and finance.

4. Examples

4.1 Industrial Fire Insurance Data

In order to highlight the methodology briefly discussed in the previous sections, we first apply it to 8043 industrial fire insurance claims. We show how a tail-fit and the resulting quantile estimates can be obtained. Clearly, a full analysis (as found, for instance, in Rootzen and Tajvidi 1997 for windstorm data) would require much more work.

Figure 4 contains the log histogram of the data. The right-skewness stresses the long-tailed behavior of the underlying data. A useful plot for specifying the long-tailed nature of data is the mean-excess plot given in Figure 5. In it, the mean-excess function

\[
e(u) = E(X - u \mid X > u)
\]

is estimated by its empirical counterpart

\[
e_n(u) = \frac{1}{\# \{1 \leq i \leq n : X_i > u \}} \sum_{i=1}^{\# \{1 \leq i \leq n : X_i > u \}} (X_i - u)^+
\]

The Pareto df can be characterized by linearity (positive slope) of \( e(u) \). In general, long-tailed df's exhibit an upward sloping behavior, exponential-type df's have roughly a constant mean-excess plot, whereas short-tailed data yield a plot decreasing to 0. In our case, the upward trend clearly stresses the long-tailed behavior. The increase in variability toward the upper end of the plot is characteristic of the technique, since toward the largest observation \( X_{\max} \), only a few data points go into the calculation of \( e_n(u) \). The main aim of our EVT analysis is to find a fit of the underlying df \( F(x) \) (or of its tail \( F(x) \)) by a generalized
Pareto df, especially for the larger values of $x$. The empirical df of $F$ is given in Figure 6 on a doubly logarithmic scale. This scale is used to highlight the tail region. Here an exact Pareto df corresponds to a linear plot.

Using the theory presented in Theorems 2 and 4, a maximum-likelihood-based approach yields estimates for the parameters of the extreme value df $H$, and the generalized Pareto df $G$. In order to start this procedure, a threshold value $u$ has to be chosen, as estimates depend on the excesses over this threshold. The estimates of the key shape parameter $\xi$ as a function of $u$ (alternatively as a function of the number of order statistics used) is given in Figure 7. Approximate 95% confidence intervals are given. The picture shows a rather stable behavior for values of $u$ below 300. An estimate in the range (0.7, 0.9) results, which should be remarked that the “optimal” value of the threshold $u$ to be used is difficult (if not impossible) to obtain. See Embrechts, Klüppelberg, and Mikosch (1997, p. 351) and Beirlant, Teugels, and Vynckier (1996) for some discussion. We also would like to stress that in order to produce Figure 7, a multitude of models (one for each $u$ chosen) has to be estimated.

For each given $u$, a tail fit for $F_u$ and $F$ (as in (10)) can be obtained. For the former, in the case of $u = 100$ an estimate $\xi = 0.747$ results. A graphical representation of $F_{100}$ is given in Figure 8. Using the parameter estimates corresponding to $u = 100$ in (10), the tail fit of $F$ on a doubly logarithmic scale is given in Figure 9.

Though we have extended the generalized Pareto fit to the left of $u = 100$, clearly only the range above this $u$-value is relevant. The fitting method is designed only for the tail. Below $u$ (where typically data are abundant) one could use a smooth version of the empirical df. From the latter plot, quantile estimates can be deduced.

Figure 10 contains as an example the estimate for the 99.9% quantile $x_{0.999}$, together with the profile likelihood. The latter can be used to find confidence intervals for $x_{0.999}$. The 95% and 99% intervals are given. Figure 11 contains the same picture, but the (symmetric) confidence intervals are calculated using the Wald statistic. Finally, the 99.9% quantile estimates across a whole range of models (depending on the threshold value, or number of exceedances used) are given in Figure 12. Though the estimate of $x_{0.999}$ settles between 1400 and 1500, the 95% Wald intervals are rather wide, ranging from 500 to about 2200.

The above analysis yields a summary about the high quantiles of the fire insurance data based on the information on extremes available in the data. The analysis can be used as a tool in the final pricing of risks corresponding to high layers (catastrophic, rare events). All the methods used are based on extremes and are fairly standard.

### 4.2 An ARCH Example

To further illustrate some of the available techniques, we simulated an ARCH(1) time series of length 99,000. The time series, called `testarch`, has the form...
From known results of Kesten (1973) (see also EKM, Theorem 8.4.12; Goldie 1991; Vervaat 1979)

\[ P(\xi_n > x) \sim c(x^{-\alpha}), \quad x \to \infty, \quad (12) \]

and we get from Table 3.2 of de Haan et al. (1989) that \( \alpha = 2.365 \) (see also Hooghiemstra and Meester 1995).

There are several reasons why we choose to simulate an ARCH process:
• Despite the fact that the ARCH process is dependent, much of the classical extreme value analysis applies with suitable modifications.
• The ARCH process has heavy tails, which matches what is observed in data sets emerging from finance.
• Although it is often plausible to model large insurance claims as iid, data from the finance industry such as exchange rate or equity data are demonstrably not iid. For some of these examples, the data look remarkably uncorrelated, but squares or absolute values of the data appear to have high correlations. It is this property that the ARCH process and its cousins were designed to model. See, for instance, Taylor (1986) for more details.

To experiment with these ARCH data we took the first 10,000 observations to form a data set shortarch, which will be used for estimation. Based on the estimation, some model-based predictions can then be made and compared with actual data in testarch shorten.

Figure 13 shows a time series plot of shortarch. The plot exhibits the characteristic heavy tail appearance. The Hill estimator is a popular way of detecting heavy tails and estimating the Pareto index 1/α. In our case, the Hill estimator is trying to estimate 2.365, when the number of upper order statistics was 300.) Applying this technique to the full testarch data produced estimates of 2α = 3.861008 and c = 1.319316, when the number of upper order statistics was 300.

Based on these estimates, we experiment with some predictions and compare them with what is observed from that part of the data set testarch called playarch, obtained by removing the 10,000 shortarch observations. Thus the length of playarch is 99,000 — 10,000 = 89,000. In Table 2 we give estimated marginal probabilities that the ARCH variable exceeds x for x = 5, 10, 15, 20. Note that we are predicting values that are beyond the range of the data and have not been observed. The second row gives the estimate (12) based on the fitted values for c and 2α. In the third row we compute the empirical frequency that elements of playarch exceed x. The last row gives the corresponding probabilities 1 — H(x, µ, σ²) based on a normal distribution whose mean and variance are the sample mean and variance computed from shortarch. One can see from Table 2 the penalty paid for ignoring extreme value analysis and relying on more conventional normal-distribution-based analysis.

The extreme value theory for the ARCH process is somewhat complicated by the ARCH dependence structure not present for an iid sequence. A quantity called the extremal index must be accounted for; see Embrechts, Klüppelberg, and Mikosch (1997, Section 8.1). From (11) and de Haan et al. (1989, Table 3.2), we have
EXTREME VALUE THEORY AS A RISK MANAGEMENT TOOL

Figure 14
Hill Plots of Shortarch

Table 2
Exceedance Probabilities
for the ARCH Example

Table 3 gives a few representative values.

4.3 Value at Risk: A Word of Warning

We have already pointed out the similarity in estimating attachment points or retentions in reinsurance andVaR calculations in finance. Both are statistically based methods, where the basic underlying risk measure corresponds to a quantile estimate \( \xi_p \) of an unknown df. Through the work of Artzner et al. (1996, 1998) we know that a quantile-based risk measure for general (nonnormal) data fails to be coherent—such a measure is not subadditive, creating inconsistencies in the construction of risk capital based upon it. This
Table 3
Quantile Estimates for the ARCH Example

<table>
<thead>
<tr>
<th>y</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(\text{max})</td>
<td>0.28229</td>
<td>0.65862</td>
<td>0.83794</td>
<td>0.91613</td>
</tr>
<tr>
<td>p</td>
<td>0.05</td>
<td>0.01</td>
<td>0.005</td>
<td>0.0005</td>
</tr>
<tr>
<td>x_p</td>
<td>34.47153</td>
<td>52.63015</td>
<td>63.04769</td>
<td>114.692</td>
</tr>
</tbody>
</table>

Figure 16
Time Series and Mean-Excess Plots of BMW Return Data

More work is needed to combine the ideas presented in this paper with detailed statistical information on financial time series before risk measures such as conditional VaR (15) can be precisely formulated and reliably estimated. Once more, the interplay between statisticians, finance experts, and actuaries should prove to be fruitful toward achieving this goal.

ACKNOWLEDGMENTS
Sidney Resick and Gennady Samorodnitsky were partially supported by NSF Grant DMS-97-04892 at Cornell University. Paul Embrechts gratefully acknowledges the hospitality of the School of Operations Research and Industrial Engineering, Cornell University, during Fall 1996.

REFERENCES


Discussions on this paper can be submitted until October 1, 1997. The authors reserve the right to reply to any discussion. See the Submission Guidelines for Authors for detailed instructions on the submission of discussions.