K. Advanced Topics in EVT

1. Efficient Quantile Estimation with POT
2. The POT Method with Dependent Data
3. Dynamic EVT in Time Series Framework
4. An Example with S&P Data
5. VaR Estimation and Backtesting
6. Var for Longer Time Horizons – Scaling Rules
K1. Efficient Quantile Estimation with POT

Estimation of quantiles with POT is a more efficient method than simple empirical quantile estimation. The latter is often used in the historical simulation approach, but gives poor estimates when we are estimating at the edge of the sample.

Recall that we can compare the efficiency of two quantile estimators by comparing their mean squared errors (MSE). If $\hat{x}_q$ is an estimator of $x_q$ then

$$\text{MSE}(\hat{x}_q) = E((\hat{x}_q - x_q)^2)$$

$$= \text{var}(\hat{x}_q) + E(\hat{x}_q - x_q)^2.$$

Good estimators trade variance of against bias to give small MSE.
Comparison of Estimators

Take ordered data $X(1) > \ldots > X(n)$ (no ties) and place threshold $u$ at an order statistic: $u = X_{(k+1)}$.

We emphasize dependence of POT estimator on choice of $k$ by writing

$$\hat{x}_{q,k} = X_{(k+1)} + \frac{\hat{\beta}_k}{\hat{\xi}_k} \left( \left( \frac{n}{k} (1 - q) \right)^{-\hat{\xi}_k} - 1 \right),$$

where $k \in \{ j \in \mathbb{N} : j \geq n(1 - q) \}$.

The empirical quantile estimator is $\hat{x}_q^E = X_{[n(1-q)]+1}$.

**Example.** $n = 1000$ implies $\hat{x}_{0.995}^E = X(6)$.

We compare $\text{MSE}(\hat{x}_{q,k})$ with $\text{MSE}(\hat{x}_q^E)$. 
Simulation Study

For various underlying $F$, various sample sizes $n$ and various quantile probabilities $q$ we compare the MSEs of these estimators. MSEs are estimated by Monte Carlo, i.e. repeated simulation of random samples from $F$.

Examples
Hard: $t$–distribution, $n = 1000$, $q = 0.999$.
Easy: normal distribution $n = 1000$, $q = 0.95$.

We will actually compare

$$\text{RRMSE} (\hat{x}_q) = \frac{\sqrt{\text{MSE}(\hat{x}_q)}}{x_q}$$

to express error relative to original units.
0.999 Quantile of $t$ with $\nu = 2$

Bias (0.999 quantile)

RRMSE (0.999 quantile)
0.95 Quantile of Standard Normal

Bias (0.95 quantile)

RRMSE (0.95 quantile)
K2. Statistical Implications of Dependence

If we believe we have a (strictly) stationary time series with a stationary distribution $F$ in the MDA of an extreme value distribution, then we can still apply the POT method and attempt to approximate the excess distribution $F_u(x)$ by a GPD for some high threshold $u$.

Although the marginal distribution of excesses may be approximately GPD, the joint distribution is unknown. We form the likelihood by making the simplifying assumption of independent excesses.

We can expect our estimation procedure to deliver consistent parameter estimates, but standard errors and confidence intervals may be over-optimistically small. Dependent samples carry less information about extreme events than independent samples of the same size.
Other Possibilities

- Use statistical estimation method for GPD parameters which does not implicitly assume independence of the excesses, such as probability weighted moments. However this method does not deliver standard errors.

- Attempt to make the excesses more independent by the technique of declustering and then use ML estimation. We identify clusters of exceedances and reduce each cluster to a single representative such as the cluster maximum.
K3. EVT in a Time Series Framework

We assume (negative) returns follow stationary time series of the form

\[ X_t = \mu_t + \sigma_t Z_t. \]

Dynamics of conditional mean \( \mu_t \) and conditional volatility \( \sigma_t \) are given by an AR(1)-GARCH(1,1) model:

\[
\begin{align*}
\mu_t &= \phi X_{t-1}, \\
\sigma_t^2 &= \alpha_0 + \alpha_1 (X_{t-1} - \mu_{t-1})^2 + \beta \sigma_{t-1}^2,
\end{align*}
\]

with \( \alpha_0, \alpha_1, \beta > 0, \alpha_1 + \beta < 1 \) and \( |\phi| < 1 \).

We assume \( (Z_t) \) is strict white noise with \( E(Z_t) = 0 \) and \( \text{var}(Z_t) = 1 \), but leave exact innovation distribution unspecified. Other GARCH-type models could be used if desired.
Dynamic EVT

Given a data sample \(x_{t-n+1}, \ldots, x_t\) from \((X_t)\) we adopt a two-stage estimation procedure. (Typically we take \(n = 1000\).)

- We forecast \(\mu_{t+1}\) and \(\sigma_{t+1}\) by fitting an AR–GARCH model with unspecified innovation distribution by pseudo-maximum-likelihood (PML) and calculating 1–step predictions. (PML yields consistent estimator of GARCH–parameters)

- We consider the model residuals to be iid realisations from the innovation distribution and estimate the tails of this distribution using EVT (GPD-fitting). In particular estimates of quantiles \(z_q\) and expected shortfalls \(E[Z \mid Z > z_q]\) for the distribution of \((Z_t)\) can be determined.
Risk Measures

Recall that we must distinguish between risk measures based on tails of conditional and unconditional distributions of the loss - in this case the negative return.

We are interested in the former and thus calculate risk measures based on the conditional distribution $F[X_{t+1} | F_t]$.

For a one-step time horizon risk measure estimates are easily computed from estimates of $z_q$ and $E [Z | Z > z_q]$ and predictions of $\mu_{t+1}$ and $\sigma_{t+1}$ using

$$\text{VaR}_q(X_{t+1}) = \mu_{t+1} + \sigma_{t+1} z_q,$$

$$\text{ES}_q(X_{t+1}) = \mu_{t+1} + \sigma_{t+1} E [Z | Z > z_q].$$
Dynamic EVT II

Advantages of this approach

We model tails of innovation distribution explicitly, using methods which are supported by statistical theory. Residuals are approximately iid, so use of standard POT procedure is unproblematic.

Alternative Estimation Approaches.

(a) Assume \((X_t)\) is GARCH process with normal innovations and fit by standard ML. In practice high quantiles are often underestimated.

(b) Assume \((X_t)\) is GARCH process with scaled \(t_\nu\)–innovations. Use ML to estimate \(\nu\) and GARCH–parameters at the same time. In practice: this works much better but has some problems with asymmetric return series.
1000 day excerpt from series of negative log returns on Standard & Poors index containing crash of 1987.
“Prewhitening” with GARCH

Series : data

Series : abs(data)

Series : residuals

Series : abs(residuals)
Heavy-Tailedness Remains

QQ-plot of residuals; raw data from S&P

©2005 (Embrechts, Frey, McNeil)
Comparison with Standard Conditional Distributions

Losses

Gains

©2005 (Embrechts, Frey, McNeil)
K5. Backtesting

The ETH Riskometer 🌟

Market Risk Summary for Major Indices on 18/04/00

Dynamic Risk Measures

<table>
<thead>
<tr>
<th>Index</th>
<th>VaR (95%)</th>
<th>ESfall (95%)</th>
<th>VaR (99%)</th>
<th>ESfall (99%)</th>
<th>Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500</td>
<td>3.98</td>
<td>5.99</td>
<td>7.16</td>
<td>9.46</td>
<td>40.1</td>
</tr>
<tr>
<td>Dow Jones</td>
<td>3.66</td>
<td>5.43</td>
<td>6.47</td>
<td>8.47</td>
<td>37.4</td>
</tr>
<tr>
<td>DAX</td>
<td>3.08</td>
<td>4.21</td>
<td>4.89</td>
<td>6.12</td>
<td>29.3</td>
</tr>
</tbody>
</table>

- **VaR and ESfall** prognoses are estimates of potential daily losses expressed as percentages.
- **Volatility** is an annualized estimate expressed as a percentage; click on column heading for recent history.
- **Data** are kindly provided by Olsen & Associates.
- **Developers are** Alexander McNeil and Rüdiger Frey in the group for financial and insurance mathematics in the mathematics department of ETH Zürich.
- **Our methods**, which combine econometric modelling and extreme value theory, are described in our research paper; there are postscript and pdf versions.

**VaR Backtests & Violation Summary**

- DAX backtest table or picture
- Dow Jones backtest table or picture
- S&P backtest table or picture

In all backtest pictures the 95% VaR is marked by a solid red line and the 99% VaR by a dotted red line. Circles and triangles indicate violation respectively.

*Alexander McNeil (mcneil@math.ethz.ch)*

©2005 (Embrechts, Frey, McNeil)
Dynamic EVT: 95% and 99% VaR Predictions

DAX Returns: losses (+ve) and profits (-ve)
## Backtesting II – numbers of violations

<table>
<thead>
<tr>
<th></th>
<th>S&amp;P</th>
<th>DAX</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Length of Test</strong></td>
<td>7414</td>
<td>5146</td>
</tr>
<tr>
<td><strong>0.95 Quantile</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>371</td>
<td>257</td>
</tr>
<tr>
<td>Conditional EVT</td>
<td>366 (0.41)</td>
<td>258 (0.49)</td>
</tr>
<tr>
<td>Conditional Normal</td>
<td>384 (0.25)</td>
<td>238 (0.11)</td>
</tr>
<tr>
<td>Conditional t</td>
<td>404 (0.04)</td>
<td>253 (0.41)</td>
</tr>
<tr>
<td>Unconditional EVT</td>
<td>402 (0.05)</td>
<td>266 (0.30)</td>
</tr>
<tr>
<td><strong>0.99 Quantile</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>74</td>
<td>51</td>
</tr>
<tr>
<td>Conditional EVT</td>
<td>73 (0.48)</td>
<td>55 (0.33)</td>
</tr>
<tr>
<td>Conditional Normal</td>
<td>104 (0.00)</td>
<td>74 (0.00)</td>
</tr>
<tr>
<td>Conditional t</td>
<td>78 (0.34)</td>
<td>61 (0.11)</td>
</tr>
<tr>
<td>Unconditional EVT</td>
<td>86 (0.10)</td>
<td>59 (0.16)</td>
</tr>
</tbody>
</table>

**Remark:** Performance of ES estimates even more sensitive to suitability of model in the tail region.
K6. Multi-day returns: Simulation of P&L

We adopt a Monte Carlo procedure and simulate from our dynamic model. We simulate iid noise from composite distribution made up of empirical middle and GPD tails.

\[ F_Z(z) = \begin{cases} 
\frac{k}{n} \left( 1 + \xi_k^{(2)} \frac{|z-z_{(n-k)}|}{\beta_k^{(2)}} \right)^{-1/\xi_k^{(2)}} & \text{if } z < z_{(n-k)}, \\
\frac{1}{n} \sum_{i=1}^{n} 1\{z_i \leq z\} & \text{if } z_{(n-k)} \leq z \leq z_{(k+1)}, \\
1 - \frac{k}{n} \left( 1 + \xi_k^{(1)} \frac{z-z_{(k+1)}}{\beta_k^{(1)}} \right)^{-1/\xi_k^{(1)}} & \text{if } z > z_{(k+1)}. 
\end{cases} \]

For an $h$-day calculation we simulate 1000 (say) conditionally independent future paths $x_{t+1}, \ldots, x_{t+h}$ and compute simulated iid observations $x_{t+1} + \ldots + x_{t+h}$. Risk measures are estimated from simulated data.
Empirical Multi-day Results

**Goal:** assess performance and compare with “square root of time rule” (valid for iid normally distributed returns).

<table>
<thead>
<tr>
<th></th>
<th>S&amp;P</th>
<th>DAX</th>
<th>BMW</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = 10$; length of test</td>
<td>7405</td>
<td>5136</td>
<td>5136</td>
</tr>
<tr>
<td>0.95 Quantile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>370</td>
<td>257</td>
<td>257</td>
</tr>
<tr>
<td>Conditional EVT ($h$-day)</td>
<td>403</td>
<td>249</td>
<td>231</td>
</tr>
<tr>
<td>Square-root-of-time</td>
<td>623</td>
<td>318</td>
<td>315</td>
</tr>
<tr>
<td>0.99 Quantile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>74</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>Conditional EVT ($h$-day)</td>
<td>85</td>
<td>48</td>
<td>53</td>
</tr>
<tr>
<td>Square-root-of-time</td>
<td>206</td>
<td>83</td>
<td>70</td>
</tr>
</tbody>
</table>

Square root of time scaling does not seem sophisticated enough!
Note that formal statistical testing difficult because of overlapping returns.
References

On EVT and Market Risk Management

- [McNeil and Frey, 2000]
- [McNeil, 1999]
- Papers in the Part “Applications to Finance” of [Embrechts, 2000]


