STA 590, Spring 05. Some Aspects on Extreme Value Analysis

- If $X_1, X_2, X_3, X_4 \ldots$ forms a sequence of independent random variables, consider

  \[ M_n = \max\{X_1, \ldots, X_n\} \]

- Develop statistical models and study the statistical behavior of $M_n$.

- The $X_i's$ could be ozone levels, rainfall, sea levels, temperature or a financial index.

- Model tail behavior of the distribution.
• If the $X_i$'s have a common distribution function $F(x)$,
\[
Pr[M_n \leq z] = Pr[X_1 \leq z, X_2 \leq z, \ldots, X_n \leq z] = F^n(z)
\]

• Result not very useful, since $F$ is unknown.

• If $z_+$ is the smallest value for which $F(z) = 1$, if $z < z_+$ then $F(z) < F(z_+) = 1$ and $\lim_{n \to \infty} F^n(z) = 0$.

• Consider a new random variable $M_n^* = (M_n - a_n)/b_n$ where $\{a_n > 0\}$ and $\{b_n\}$ are sequences of numbers

• Focus: Limit distribution of $M_n^*$.

• Coles, S. (2001). *An Introduction to Statistical
Modeling of Extreme Values. Springer Verlag: New York, discusses the following theorem.

- **Extremal Theorem:** If \( \{a_n > 0\} \) and \( \{b_n\} \) are such that

\[
Pr\{(M_n - b_n)/a_n \leq z\} \rightarrow G(z); \ n \rightarrow \infty
\]

then \( G \) must be a member of the so called Generalized Extreme Value (GEV) family of distributions.

- The Generalized Extreme Value (GEV) distribution function is:

\[
G(z) = \exp \left\{ -\left[1 + \xi(z - \mu)/\sigma\right]_+^{1/\xi} \right\}
\]
\( -\infty < \mu < \infty \) is a location parameter; \( \sigma > 0 \) is a scale parameter; \( -\infty < \xi < \infty \) is a shape parameter.

- + denotes the positive part of the argument.
- \( \xi > 0 \) gives the Fréchet (type I) family;
- \( \xi < 0 \) defines the Weibull (type II) family;
- \( \xi \to 0 \) leads to the Gumbel (type III) family.

\[
G(z) = \exp \left\{ -\exp \left( -\left( z - \mu \right) / \sigma \right) \right\} ; -\infty < z < \infty
\]

**Example:** If \( X_1, X_2, \ldots, X_n \) are \( U(0, 1) \). For \( z < 0; \ n > -z \); let \( a_n = 1/n \) and \( b_n = 1 \)

\[
Pr\left\{ \left( M_n - b_n \right) / a_n \leq z \right\} = Pr\left\{ M_n \leq z / n + 1 \right\} = (1 + n^{-1} z)^n \to e^z
\]
which is a *Weibull* type distribution $\xi = -1$.

- The difficulty with the normalizing constants is "easily" resolved. Equivalently,

$$Pr\{M_n \leq z\} \approx G\{(z - b_n)/a_n\} = G^*(z)$$

for large enough $n$. $G^*(\cdot)$ also belongs to the GEV family.

- *Estimation of the GEV:* Consider

$$M_{n,k} = \max\{X_{k,1}, X_{k,2}, \ldots, X_{k,n}\}$$

- $n$ block size; $k = 1, \ldots, m$ number of blocks.

- To simplify the notation,

$$z_1 = M_{n,1}, z_2 = M_{n,2}, \ldots, z_m = M_{n,m}.$$
• Let $\theta = (\mu, \sigma, \xi)$, if the $z_i$'s are independent $z_i \sim GEV(\theta); i = 1, \ldots, m$, the log-likelihood is
\[
\ell(\theta) = -m \log \sigma - \left(1 + 1/\xi\right) \sum_{i=1}^{m} \log \{1 + \xi(z_i - \mu)/\sigma\} \\
- \sum_{i=1}^{m} \{1 + \xi(z_i - \mu)/\sigma\}^{-1/\xi}
\]
provided that $1 + \xi(z_i - \mu)/\sigma > 0; i = 1, 2, \ldots, m$.

• This log-likelihood function cannot be maximized analytically.

• To obtain the MLE, we require some kind of computational method (Newton-Raphson, EM?).

• S. Coles created a S-plus/R package ismev to find the
MLE for the parameters of the GEV distribution.

- The Splus version can be downloaded from http://www.maths.bris.ac.uk/ masgc/ismev/summary.html.
- The R-version (Alec Stephenson) is available at: http://cran.r-project.org/
- Pages 185-187 of the book by Coles give a description of the functions.
- The main functions are `gev.fit` and `gpd.fit`.
- Bayesian inference for $\theta$ can be performed using MCMC.
- A trivariate normal prior on $\theta' = (\mu, \log\sigma, \xi)$ leads to
the prior density.

\[
\pi(\theta) \propto \frac{1}{\sigma} \exp \left\{ -\frac{1}{2} (\theta' - \nu)^T \Sigma^{-1} (\theta' - \nu) \right\}
\]

- Includes the case of independent priors on \( \mu, \sigma, \xi \).
- If \( \Sigma \) is a diagonal matrix, then

\[
\pi(\theta) \propto \pi(\mu) \pi(\log(\sigma)) \pi(\xi)
\]

- Other priors: Beta Distributions for Probability Ratios and Gamma Distribution for Quantile Differences.
- Set \( G(q_p) = 1 - p \) so \( q_p \) is the \( 1 - p \) quantile of the
GEV distribution, then
\[
\exp \left\{ -\left[1 + \xi (q_p - \mu)/\sigma \right]^{1/\xi} \right\} = 1 - p
\]

- The solution for \( q_p \) is:
\[
q_p = \mu + \sigma (x_p^{\xi} - 1)/\xi
\]
with \( x_p = -\log(1 - p) \)

- A prior can be constructed in terms of quantiles \( q_{p_1}, q_{p_2}, q_{p_3} \) for probabilities \( p_1 > p_2 > p_3 \).

- Since \( q_{p_1} < q_{p_2} < q_{p_3} \) it is simpler to deal with the differences \( \tilde{q}_{p_1}, \tilde{q}_{p_2}, \tilde{q}_{p_3} \) where
\[
\tilde{q}_{p_i} = q_{p_i} - q_{p_{i-1}}; \quad i = 1, 2, 3
\]
• Fix $q_{p_0}$ ($= 0$) as a lower end point.
• A proposed prior on the quantile differences is:
  $$\tilde{q}_{p_i} \sim Gamma(\alpha_i, \beta_i); \alpha_i > 0; \beta_i > 0; i = 1, 2, 3$$
• The prior for $\theta$ is then
  $$\pi(\theta) \propto J \prod_{i=1}^{3} [\tilde{q}_{p_i}^{-1} \exp(-\beta_i \tilde{q}_{p_i})]$$
where $(\alpha_1, \alpha_2, \alpha_3), (\beta_1, \beta_2, \beta_3), p_1, p_2, p_3$ must all be specified.
• For posterior inference a Hybrid MCMC method is used.
• The full conditional distribution of each parameter is
simulated with a Metropolis-Hastings step.

- A simple choice is to specify random walks in the 3 model parameters:

\[
\begin{align*}
\mu^* &= \mu + \epsilon_\mu \\
\phi^* &= \phi + \epsilon_\phi \\
\xi^* &= \xi + \epsilon_\xi
\end{align*}
\]

where \( \epsilon_\mu, \epsilon_\phi, \epsilon_\xi \) are normal RVs with zero mean and variances \( v_\mu, v_\phi, v_\xi \) respectively. \( \phi = \log(\sigma) \)

- After some tuning, it is possible to obtain decent MCMC simulations.

- The output of `ismev` can be used to tune the
The R-library \textit{evdbayes} available at http://cran.r-project.org/ provides function for the Bayesian Analysis of the GEV distribution.

This library was written by Alec Stephenson and can also be downloaded from http://www.maths.lancs.ac.uk/~stephena/.

The package includes a \textit{user guide}.

An alternative to extreme value analysis is to consider \textit{threshold models}.

\( X_1, X_2, \ldots \) iid observations. Extreme event: \( X_i > u \).

Model exceedances: \( P(X - u > y|X > u) \)
For any $y > 0$, $P(X > y + u | X > u) = \frac{1 - F(u+y)}{1 - F(u)}$

where $F$ is the distribution of $X$.

Under similar conditions for the Extremal theorem, it can be shown that for large enough $u$, the distribution of $Y = X - u$ conditional on $X > u$ is defined by the Generalized Pareto Distribution (GPD).

The GPD family is given by the expression:

$$H(y) = 1 - \left( 1 + \frac{\xi y}{\tilde{\sigma}} \right)^{-1/\xi}$$

defined for $y > 0$ and $1 + \frac{\xi y}{\tilde{\sigma}} > 0$
• If $\xi \to 0$ then

$$H(y) = 1 - \exp \left( \frac{-y}{\tilde{\sigma}} \right)$$

• For specific applications, given $u$, the parameters $\xi$ and $\tilde{\sigma}$ can be estimated by maximum likelihood (gpd.fit) or with Bayes approaches based on MCMC.

• In general, is difficult to determine a reasonable value or to estimate $u$.

• Extensions: Model changes across time. Trend or seasonality.

• Traditional approach: $z_1, z_2, \ldots, z_m$;

$$z_t \sim GEV(\mu_t, \sigma, \xi)$$
- Deterministic functions: $\mu_t = \beta_0 + \beta_1 t$; 
  $\mu_t = \beta_0 + \beta_1 + \beta_2 t + \beta_3 t^2$ or $\mu_t = \beta_0 + \beta_1 X_t$.

- Non-stationarity can also be included for the shape and/or scale parameters: $\sigma_t = \exp(\beta_0 + \beta_1 t)$; 
  $\xi_t = \beta_0 + \beta_1 t$ or $\xi_t = \beta_0 + \beta_1 t + \beta_2 t^2$.

- Alternatively, we propose the use of Dynamic Linear Models (DLM) as in West and Harrison (1997) to model the parameter changes in time. (see paper *Time-Varying Models for Extreme Values* on my personal web page).

- **Model Checking**: Consider $z_p$ such that
\( G(z_p) = 1 - p \). Then,
\[
z_p = \mu - \frac{\sigma}{\xi} \left(1 - y_p^{-\xi}\right)
\]
where \( y_p = -\log(1 - p) \)

- A return level plot is given by the points
  \[\{(\log y_p, z_p); 0 < p < 1\}\]

- If we have a point estimate of the parameters, we can obtain a point estimate of the return level plot.

- For a Bayesian approach, applying this transformation to samples of \((\mu, \sigma, \xi)\) leads to samples of the return level plot.

- This curve can be compared to the Empirical Return
level given by

\[(\log(-\log(i/m)), z_{(i)}); i = 1, \ldots m\]

where \(z_{(i)}\) denotes the ordered data.

- If empirical and theoretical return levels match, then we have a good fit of the GEV distribution.
- For the Gumbel case (\(\xi = 0\)), \(z_p = \mu - \sigma \log(y_p)\)
- For the non-stationary case a return level can be obtained for every value of \(t\).