Dependent Extremes

Gemai Chen

Department of Mathematics and Statistics, University of Calgary

May 1, 2014 at BIRS

Extreme Distributions

Let $\{Y_t\}$ be a process of interest, where $\{Y_t\}$ may represent, say, river flow rates at a monitoring station. Suppose that we are interested in the block maxima

$$M_n = \max\{Y_{t+1}, \ldots, Y_{t+n}\},$$

say, the annual maximum flow rates.

When $\{Y_t\}$ is an iid sequence, or when $\{Y_t\}$ is stationary with limited dependence, the distribution of M_n can be well approximated by a *generalized extreme value* (GEV) distribution defined for $\{y : 1 + \xi(y - \mu)/\sigma > 0\}$ by

$$G(y) = \begin{cases} \exp\{-\exp[-(y-\mu)/\sigma]\}, & \xi = 0, \\ \exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, & \xi \neq 0, \end{cases}$$

where $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$ are referred to as the location, scale and shape parameters, respectively. (Coles, 2001)

The currently favored methods of parameter estimation include the probability-weighted moments method (for small samples) and the maximum likelihood method. For inference and modeling, the maximum likelihood method is preferred.

However, the GEV family of distributions do not satisfy the regularity conditions for the general asymptotic theory of MLE.

Specifically (Smith, 1985),

- when $\xi < 0.5$, the ML estimators possess the standard asymptotic properties;
- when $0.5 \leq \xi < 1$, the ML estimators are generally obtainable, but do not possess the standard asymptotic properties;
- when $1 \leq \xi$, the ML estimators are unlikely to be obtainable.

In application, a common assumption used in the central models of extreme values is the independence of extremes. Davison and Ramesh (2000) admitted the difficulty to handle the failure of this assumption. Chavez-Demoulin and Davison (2005) argued that dependence between the extremes would not bias estimators of the parameters of a model built upon the independence assumption.

Positive α-Stable Distributions

Let S be a positive α-stable variable in the sense that for $\alpha \in (0,1)$ and $u \geq 0$, the Laplace transform of S satisfies

$$E\left(\exp(-uS)\right) = \exp(-u^{\alpha}). \tag{1}$$

Let X be Gumbel distributed ($\xi = 0$ in the GEV family) with parameters μ and σ. Then $X + \sigma \log S$ is also Gumbel distributed with parameters μ and σ/α.

Let $\{S_t\}$ be iid positive α-stable random variables defined by (1). For $\sigma > 0$, define

$$X_t = \alpha X_{t-1} + \alpha\sigma \log S_t, \;\; t \in \mathbb{Z}. \tag{2}$$

Then Equation (2) has a unique second order stationary solution

$$X_t = \sigma \sum_{j=0}^{\infty} \alpha^{j+1} \log S_{t-j} \tag{3}$$

and X_t is Gumbel distributed with parameters $\mu = 0$ and σ. (Toulemonde, etc. 2010)

There are two ways to construct dependent random sequences $\{Y_t\}$ with the same GEV marginal distributions. In both ways, $Y_t = X_t$ when $\xi = 0$. One way is to define

$$Y_t = \left(Y_{t-1} + \frac{\sigma}{\xi}\right)^{\alpha} S_t^{\alpha\xi} \left(\frac{\sigma}{\xi}\right)^{1-\alpha} - \frac{\sigma}{\xi}, \quad t \in \mathbb{Z}. \quad (4)$$

(Toulemonde, etc. 2010)

Then Equation (4) has a unique second order stationary solution

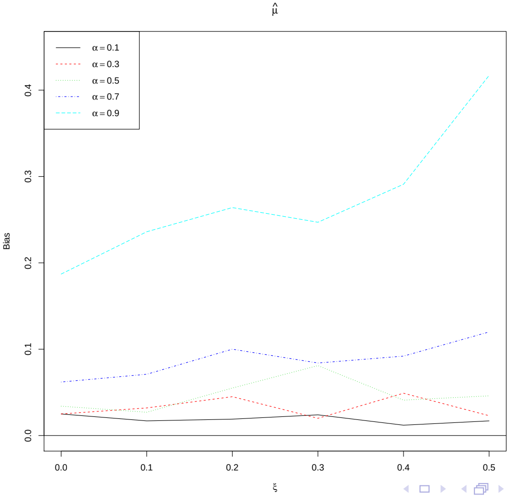
$$Y_t = -\frac{\sigma}{\xi} + \frac{\sigma}{\xi}\prod_{j=0}^{\infty}(S_{t-j})^{\xi\alpha^{j+1}} \tag{5}$$

and Y_t is a $GEV(0, \sigma, \xi)$ random variable. The other way is to define

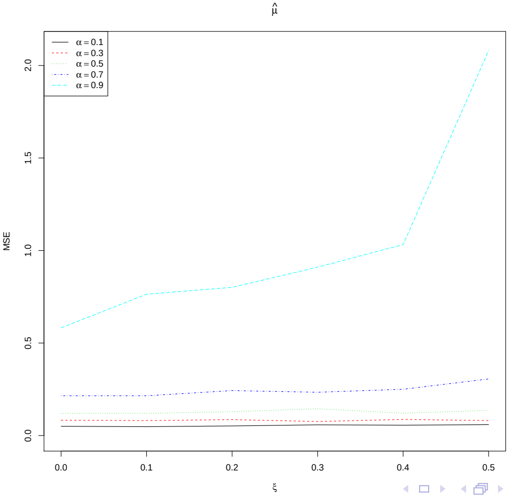
$$Y_t = \left\{\exp\left[\xi\left(\frac{S_t - \mu}{\sigma}\right)\right] - 1\right\}/\xi. \tag{6}$$

Then Y_t is a $GEV(\mu, \sigma, \xi)$ random variable.

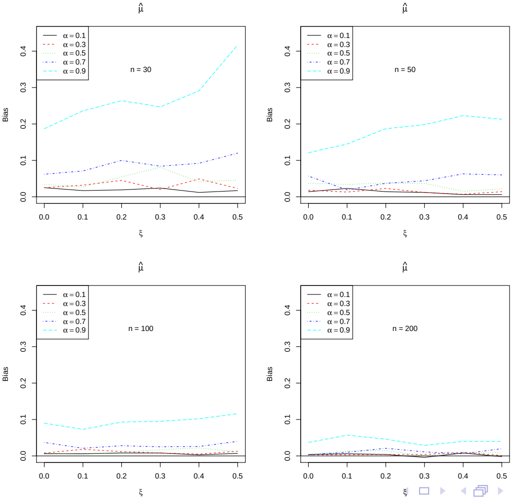
Bias of $\hat{\mu}$



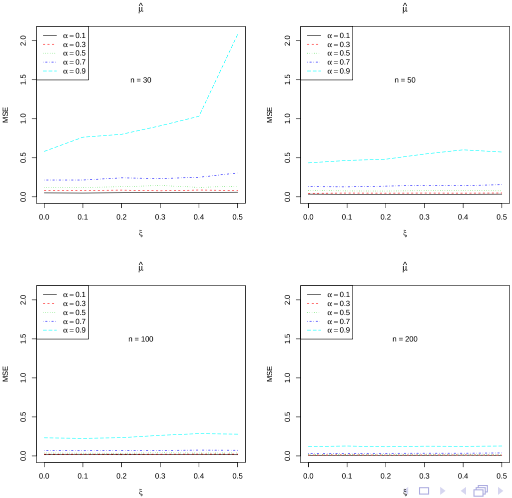
MSE of $\hat{\mu}$



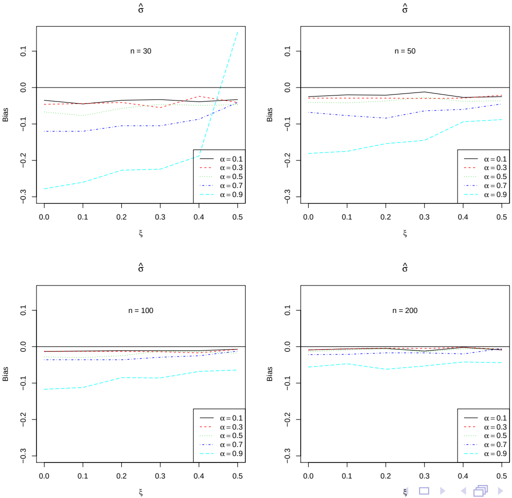
Bias of $\hat{\mu}$



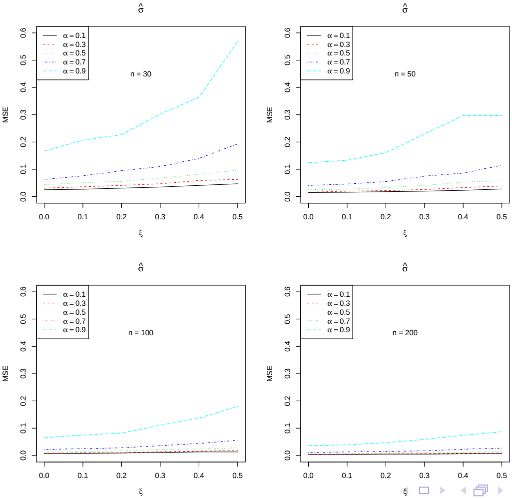
MSE of $\hat{\mu}$


μ̂
MSE
ξ
α = 0.1
α = 0.3
α = 0.5
α = 0.7
α = 0.9
n = 30
n = 50
n = 100
n = 200
0.0
0.5
1.0
1.5
2.0
0.0 0.1 0.2 0.3 0.4 0.5

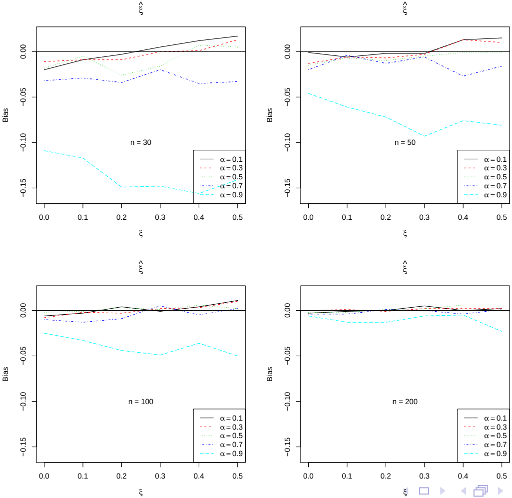
Bias of $\hat{\sigma}$



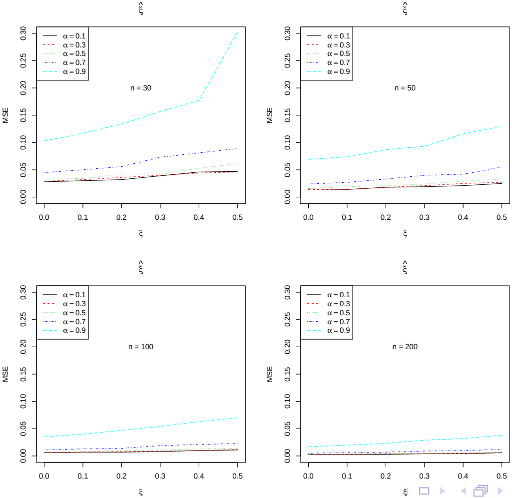
MSE of $\hat{\sigma}$



Bias of $\hat{\xi}$



MSE of $\hat{\xi}$

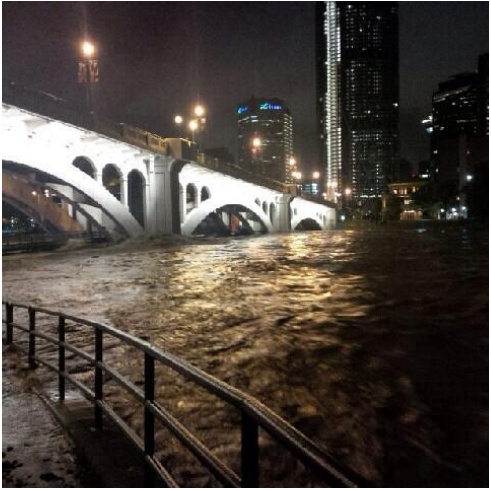


$\hat{\xi}$
MSE
$\alpha = 0.1$
$\alpha = 0.3$
$\alpha = 0.5$
$\alpha = 0.7$
$\alpha = 0.9$
n = 30
n = 50
n = 100
n = 200
0.00 0.05 0.10 0.15 0.20 0.25 0.30
0.0 0.1 0.2 0.3 0.4 0.5
ξ

Conclusion: When the dependence among the maximums is on the higher end ($\alpha \geq 0.7$) and when the sample size n is not large ($n \leq 100$), the maximum likelihood estimators of the parameters of the GEV distribution family are noticeably biased.

Calgary 2013 Flood

Calgary 2013 Flood



Calgary 2013 Flood

Calgary 2013 Flood

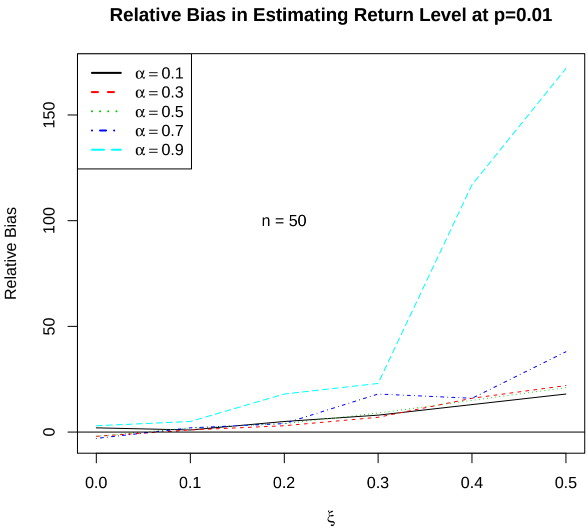
Calgary 2013 Flood

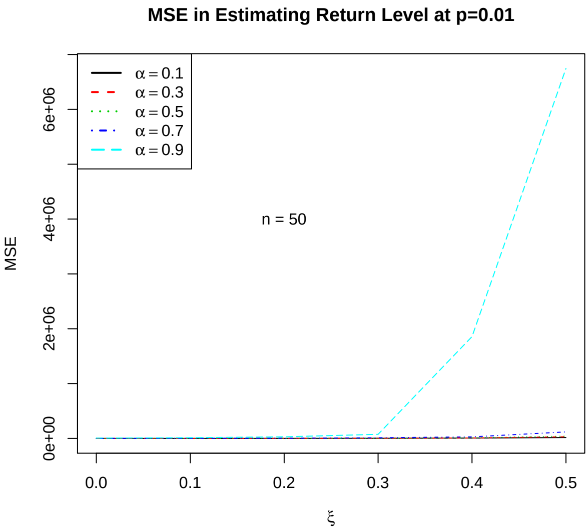
Return Level

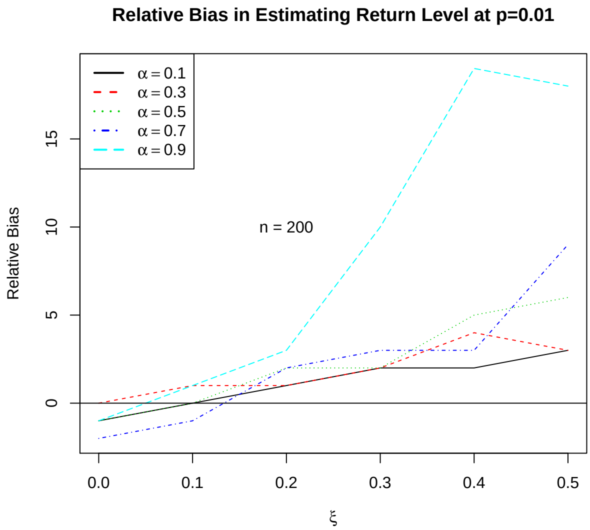
Let p^* be a small upper tail probability and let r^* be the $(1 - p^*)$th quantile of a $GEV(\mu, \sigma, \xi)$ distribution, one has

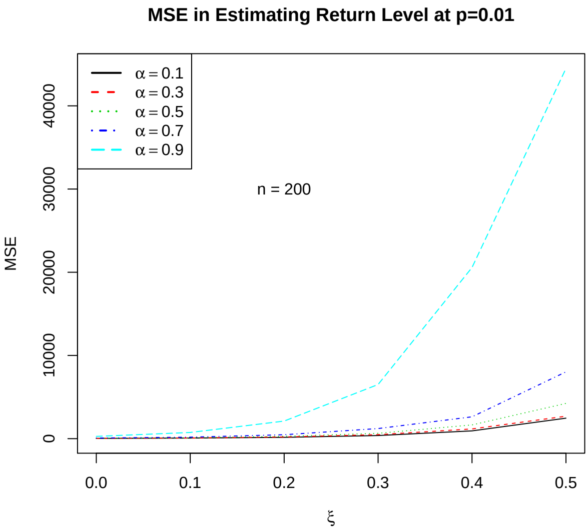
$$r^* = \begin{cases} \mu - \frac{\sigma}{\xi}\left\{1 - [-\ln(1-p^*)]^{-\xi}\right\} & \text{if } \xi \neq 0, \\ \mu - \sigma \ln[-\ln(1-p^*)] & \text{if } \xi = 0. \end{cases}$$

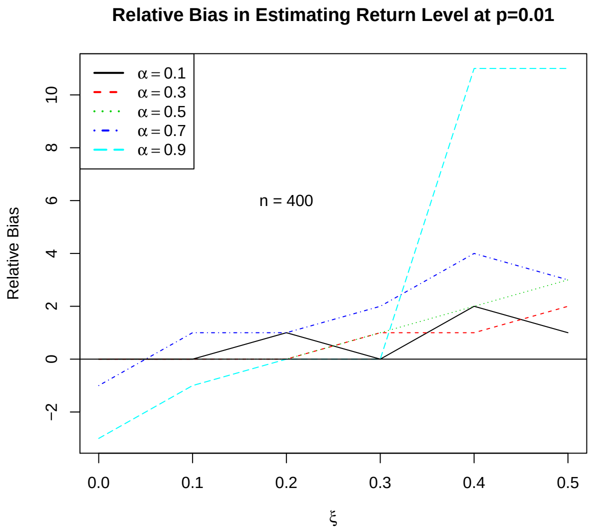
In hydrology, this is the $(1/p^*)$-year return level.

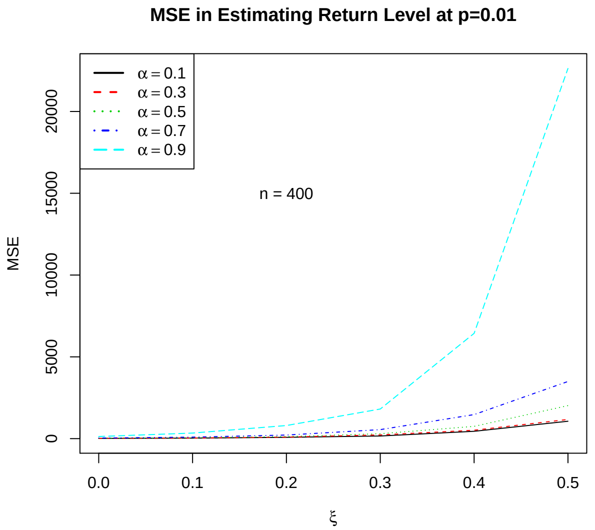

Relative Bias in Estimating Return Level at p=0.01
α = 0.1
α = 0.3
α = 0.5
α = 0.7
α = 0.9
n = 50
Relative Bias
0
50
100
150
0.0
0.1
0.2
0.3
0.4
0.5
ξ


MSE in Estimating Return Level at p=0.01
α = 0.1
α = 0.3
α = 0.5
α = 0.7
α = 0.9
n = 50
MSE
0e+00
2e+06
4e+06
6e+06
0.0
0.1
0.2
0.3
0.4
0.5
ξ


Relative Bias in Estimating Return Level at p=0.01
α = 0.1
α = 0.3
α = 0.5
α = 0.7
α = 0.9
n = 200
Relative Bias
0
5
10
15
0.0
0.1
0.2
0.3
0.4
0.5
ξ


MSE in Estimating Return Level at p=0.01
α = 0.1
α = 0.3
α = 0.5
α = 0.7
α = 0.9
n = 200
MSE
0
10000
20000
30000
40000
0.0
0.1
0.2
0.3
0.4
0.5
ξ


Relative Bias in Estimating Return Level at p=0.01
α = 0.1
α = 0.3
α = 0.5
α = 0.7
α = 0.9
n = 400
Relative Bias
−2
0
2
4
6
8
10
0.0
0.1
0.2
0.3
0.4
0.5
ξ


MSE in Estimating Return Level at p=0.01
α = 0.1
α = 0.3
α = 0.5
α = 0.7
α = 0.9
n = 400
MSE
0
5000
10000
15000
20000
0.0
0.1
0.2
0.3
0.4
0.5
ξ

1. Chavez-Demoulin, V. and Davison, A. C. (2005), Generalized additive modelling of sample extremes. *Applied Statistics*, C, 54, 207-222.

2. Coles, S. G. (2001), *An Introduction to Statistical Modelling of Extreme Values*. London: Springer.

3. Davison, A. C. and Ramesh, N. I. (2000), Local likelihood smoothing of sample extremes. *Journal of Royal Statistical Society*, B, 42, 191-208.

4. Smith, R. L. (1985), Maximum likelihood estimation in a class of non-regular cases. *Biometrika*, 72 67-90.

5. Toulemonde, G., Guillou, A., Naveau, P., Vrac, M. and Chevallier, F. (2010), Autoregressive models for maxima and their applications to CH_4 and N_2O. *Environmetrics*, 21, 189-207.

Thank You!

The extreme value method.

Divide the (log) returns $r_1, \ldots, r_T$ into

$$\{r_1, \ldots, r_n | r_{n+1}, \ldots, r_{2n} | \cdots, | r_{(g-1)n+1}, \ldots, r_{gn}\}$$

and let $r_{n,i}$ be the maximum of the ith subsample. When n is large and the dependence among $\{r_t\}$ is weak, $\{r_{n,i}\}_{i=1}^{g}$ can be regarded as a sample from the generalized extreme value (GEV) distribution defined by

$$F(r) = \begin{cases} \exp\left\{-\left[1 + \frac{\xi_n(r-\beta_n)}{\alpha_n}\right]^{-1/\xi_n}\right\} & \text{if } \xi_n \neq 0, \\ \exp\left[-\exp\left(-\frac{r-\beta_n}{\alpha_n}\right)\right] & \text{if } \xi_n = 0, \end{cases}$$

where ξ_n, α_n and β_n are the shape, scale and location parameters, and $1 + \xi_n(r - \beta_n)/\alpha_n > 0$ for $\xi_n \neq 0$.

Let p^* be a small upper tail probability and let r_n^* be the $(1 - p^*)$th quantile of the above GEV distribution, one has

$$r_n^* = \begin{cases} \beta_n - \frac{\alpha_n}{\xi_n}\left\{1 - [-\ln(1 - p^*)]^{-\xi_n}\right\} & \text{if } \xi_n \neq 0, \\ \beta_n - \alpha_n \ln[-\ln(1 - p^*)] & \text{if } \xi_n = 0. \end{cases}$$

This is the (approximate) expression for the VaR of the subsample maximums $\{r_{n,i}\}_{i=1}^{g}$.

To get the VaR for the original returns $\{r_t\}$, the following relationship is assumed in the literature:

$$1 - p^* = P(r_{n,i} \leq r_n^*) = P(r_t \leq r_n^*)^n,$$

from which one has

$$\text{VaR} = \begin{cases} \beta_n - \frac{\alpha_n}{\xi_n}\left\{1 - [-n\ln(1-p)]^{-\xi_n}\right\} & \text{if } \xi_n \neq 0, \\ \beta_n - \alpha_n \ln[-n\ln(1-p)] & \text{if } \xi_n = 0. \end{cases}$$