Generalized MLE per Martins and Stedinger

Martins ES and Stedinger JR. (March 2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research* **36**(3):737–744. (henceforth, 2000P)

Martins ES and Stedinger JR. (October 2001). Generalized maximum-likelihood Pareto-Poisson estimators for partial duration series. *Water Resources Research* **37**(10):2551–2557. (henceforth, 2001P)

http://www.ral.ucar.edu/staff/ericg/readinggroup.html

Outline

- Overview of Applications (both papers); predominantly flood frequency analysis (2001P)
- Overview of GEV (2000P)/PDS,AMS (2001P)
  - Review
  - Estimation methods with brief comparison from previous studies
  - Theoretical properties of parameters
  - Lit. Review (2000P)
  - Transformations between GEV/GPD (2001P)
  - Small samples (both papers)
- Small sample simulation
- GMLE
- Results

Overview of GEV/PDS: Review

Extremal Types Theorem: $X_1, \ldots, X_n$ random sample from any distribution.

$$\Pr\{\frac{\max\{X_1, \ldots, X_n\} - b_n}{a_n} \leq z\} \longrightarrow G(z) \text{ as } n \longrightarrow \infty$$

where $G(z)$ is one of three types of distributions.

I. (Gumbel) $G(z) = \exp\{-\exp\left[-\left(\frac{z-b}{a}\right)\right]\}$, $-\infty < z < \infty$.

II. (Fréchet) $G(z) = \exp\{-\left(\frac{z-b}{a}\right)^{-\alpha}\}$, $z > b$ and 0 otherwise.

III. (Weibull) $\exp\{-\left[-\left(\frac{z-b}{a}\right)^{\alpha}\right]\}$, $z < b$ and 1 otherwise.

(where $a > 0$, $\alpha > 0$ and b are parameters).

Overview of GEV/PDS: Review

Extremal Types Theorem

The above three distributions can be combined into a single family of distributions.

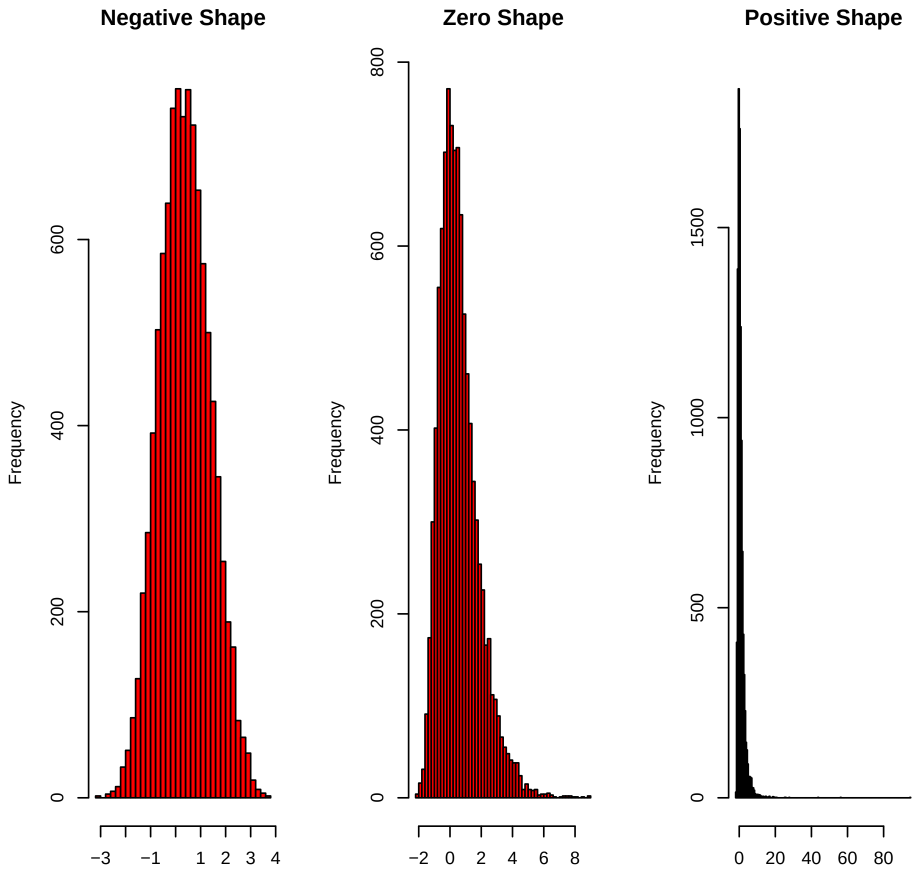
$$G(z) = \exp\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\}$$

G is called the generalized extreme value distribution (GEV).

Three parameters: location (μ), scale (σ) and shape (ξ).

These papers use ξ for location, α for scale and κ for shape Also, κ is parametrized differently. Specifically, $\kappa = -\xi$ from the above representation.

Overview of GEV/PDS: Theoretical properties of parameters



Overview of GEV/PDS: Theoretical properties of parameters

1. The GEV is only defined when $1 + \xi\left(\frac{z-\mu}{\sigma}\right) > 0$.
2. Range of data is dependent upon unkown parameters! Hence, regularity conditions for MLE do not necessarily hold.
   (a) For $\xi > 0$, $\mu - \sigma/\xi \leq x$.
   (b) For $\xi < 0$, $x \leq \mu - \sigma/\xi$.
3. For $\xi \geq -0.5$ desirable asymptotic properties of efficiency and normality of MLE's hold.
4. If $\xi < -1$, the density $\longrightarrow \infty$ as $\mu - \sigma/\xi$ approach the largest observation.
5. Even under 2a above, the MLE can perform satisfactorily if the likelihood is modified; but does not help for small samples.

Overview of GEV/PDS: Estimation methods

- Hosking *et al.* (1985) showed L-moments to be superior for GEV to MLE in terms of bias and variance for small sample sizes ($n = 15$ to $n = 100$).
- Madsen *et al.* show MOM quantile estimators have smaller RMSE for $-0.30 < \xi < 0.25$ than both LM and MLE when estimating the 100-year event with $n \in [10, 50]$; with MLE preferable for $\xi < -0.3$ and $n \geq 50$.
- It is straightforward to incorporate censored data (covariates) into MLE; but not with LM/MOM.

Overview of GEV/PDS: Generalized Pareto Distribution (GPD)

Exceedance Over Threshold Model

For X random (with cdf F) and a (large) threshold u

$$\text{Pr}\{X > x | X > u\} = \frac{1 - F(x)}{1 - F(u)}$$

Then for $x > u$ (u *large*), the GPD is given by

$$\frac{1 - F(x)}{1 - F(u)} \approx [1 + \frac{\xi}{\sigma}(x - u)]^{-1/\xi}$$

Overview of GEV/PDS: Transformations between GEV/GPD (2001P)

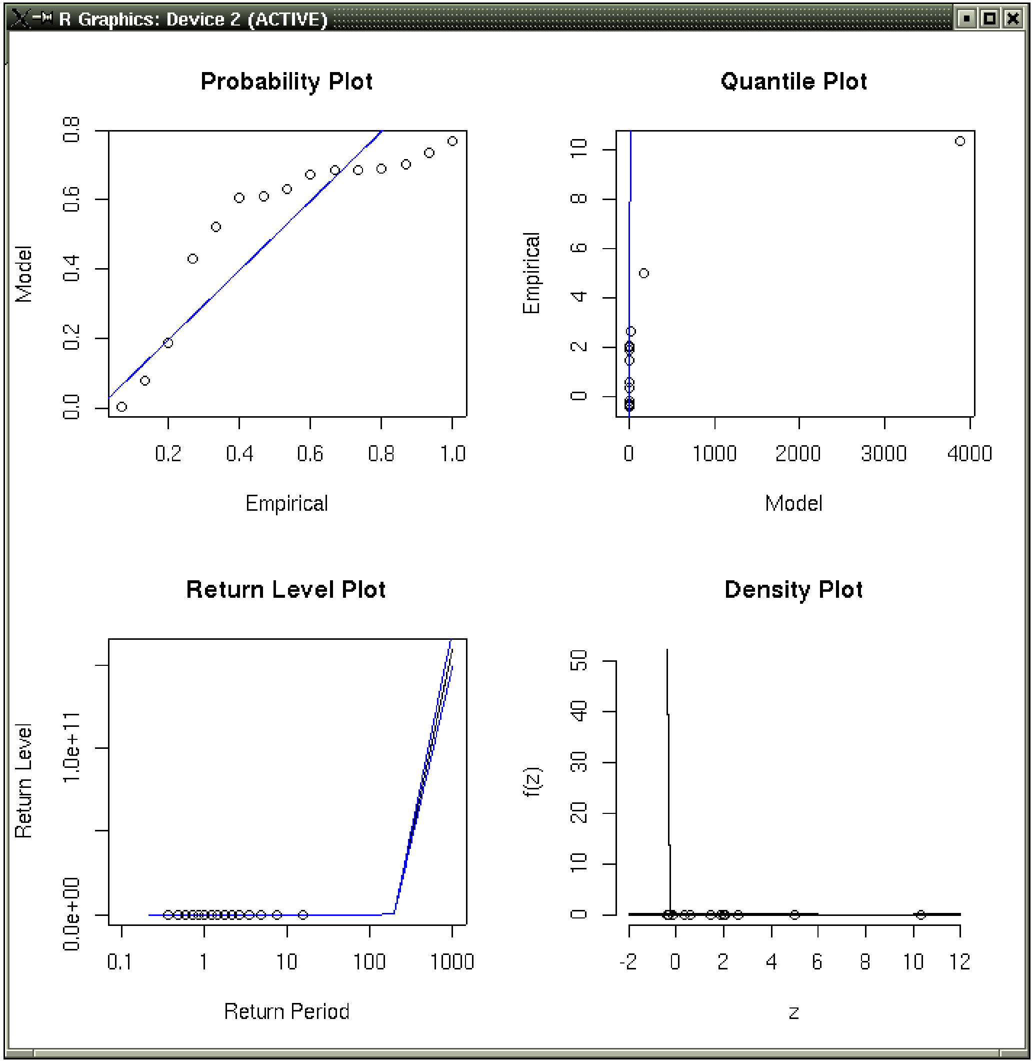
(Here, taken from `extRemes` toolkit tutorial using the $\xi = -\kappa$ parameterization of GEV).

$$\log \lambda = -\frac{1}{\xi} \log\{1 + \xi \frac{u - \mu}{\sigma}\}$$

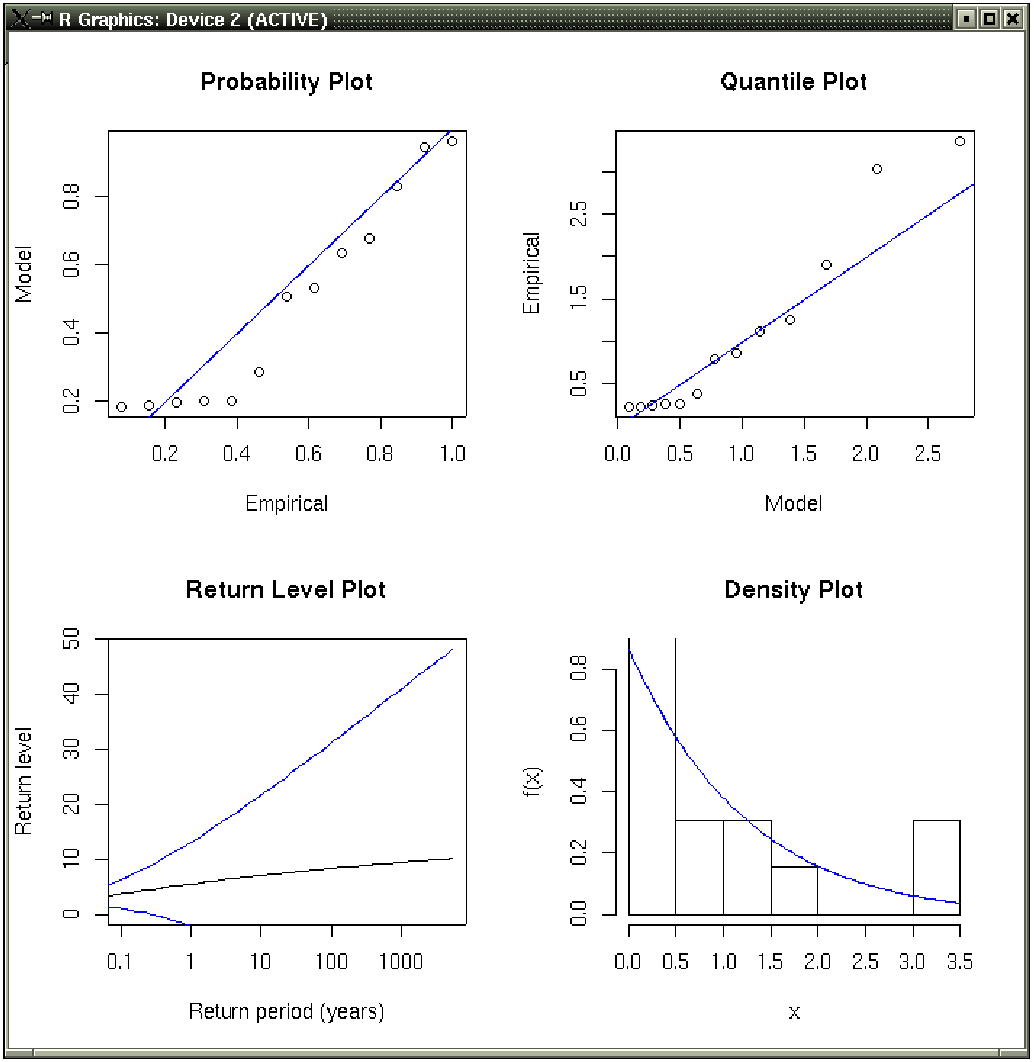
$$\sigma^* = \sigma + \xi(u - \mu)$$

etc...

Small sample simulation



Small sample simulation



GMLE

- Coles and Dixon (1999)

$$L_{pen}(\mu, \sigma, \xi) = L(\mu, \sigma, \xi) \times P(\xi),$$

where

$$P(\xi) = I_{\xi \leq 0} 1 + I_{0 < \xi < 1} \exp\{-\lambda(\frac{1}{1-\xi} - 1)^{\alpha}\}$$

- Martins and Stedinger (2000, 2001)

$$\text{GL}(\mu, \sigma, \xi | x) = L(\mu, \sigma, \xi) \times \pi(\xi),$$

where $\pi(\xi)$ is a Beta prior.

GMLE

As sample size increases, information in the likelihood *should* dominate the GMLE estimator, so that MLE and GMLE asymptotically have the same desirable properties.

Results

(2000P)

- For $\xi \geq 0$, GMLE does better than MOM and LM at estimating quantiles.
- If $\xi < 0$, then a more appropriate prior should be used with GMLE.
- For $\xi = 0.10$, two-parameter GEV/MLE is better than three-parameter GEV/GMLE (in a narrow region).

(2001P)

- For $\xi \geq 0$, GMLE performs about the same for both PDS and AMS; superior to other quantile estimators.
- MOM is just as good for $\xi = 0$ and better for $\xi \leq 0$.
- Two-parameter PDS/exponential-Poisson MLE is better than three-parameter PDS/GP GMLE in a narrow region.

That's all! Unless you want more.

Estimation methods

- Maximum Likelihood Estimation (MLE)
- Method of L Moments
- Bayesian estimation

MLE

Assuming $Z_1, \ldots, Z_m$ are iid random variables that follow the GEV distribution the log-likelihood is given by the following.

$$\ell(\mu, \sigma, \xi) = -m \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{m} \log \left[1 + \xi \left(\frac{z_i - \mu}{\sigma}\right)\right]$$
$$- \sum_{i=1}^{m} \left[1 + \xi \left(\frac{z_i - \mu}{\sigma}\right)\right]^{-1/\xi}$$

L Moments

Probability Weighted Moments (PWM)

$$M_{p,r,s} = E\left[X^p\{F(X)\}^r\{1 - F(X)\}^s\right]$$

L-moments are based on the special cases $\alpha_r = M_{1,0,r}$ and $\beta_r = M_{1,r,0}$. Specifically, let $x(u)$ be the quantile function for a distribution, then:

$$\alpha_r = \int_0^1 x(u)(1-u)^r du$$

$$\beta_r = \int_0^1 x(u)u^r du$$

Compare to ordinary moments: $E(X^r) = \int_0^1 \{x(u)\}^r du$.

L-moments

Much more to it, but the moments derived in the paper come from:

- $\lambda_1 = \alpha_0 = \beta_0$,
- $\lambda_2 = \alpha_0 - 2\alpha_1 = 2\beta_1 - \beta_0$ and
- $\lambda_3 = \alpha_0 - 6\alpha_1 + 6\alpha_2 = 6\beta_2 - 6\beta_1 + \beta_0$

More generally

$$\lambda_r = \int_0^1 x(u) \sum_{k=0}^{r-1} \frac{(-1)^{r-k-1}(+k-1)!}{(k!)^2(r-k-1)!} du$$

Alternatively

- For $n = 1$, $X_{1:1}$ estimates location. If distribution is shifted to larger values, then $X_{1:1}$ is expected to be larger. (Hence, $\lambda_1 = E(X_{1:1})$)
- For $n = 2$, $X_{2:2} - X_{1:1}$ estimates scale (dispersion). If dist'n is tightly bunched, small value. (Hence, $\lambda_2 = \frac{1}{2}E(X_{2:2} - X_{1:2})$)
- For $n = 3$, $X_{3:3} - 2X_{2:3} + X_{1:3}$ measures skewness. (i.e., $X_{3:3} - X_{2:3} \approx X_{2:3} - X_{1:3}$). (Hence, $\lambda_3 = \frac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3})$)

And in general,

$$\lambda_r = r^{-1} \sum_{j=0}^{r-1} (-1)^j \frac{(r-1)!}{j!(r-j-1)!} E(X_{r-j:r})$$

For more on L-Moments

Hosking JRM and Wallis JR. 1997. *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press.

Some References

Applied Introductory references to extreme value statistical analysis

Coles S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer.

Gilleland E and Katz, RW. 2005. Tutorial to Extremes Toolkit. `http://www.assessment.ucar.edu/toolkit`

Katz RW, Parlange MB, and Naveau P. 2002. Statistics of extremes in hydrology. *Advances in Water Resources*, **25**:1287–1304.

Smith RL. 2002. *Statistics of extremes with applications in environment, insurance and finance*. `http://www.stat.unc.edu/postscript/rs/semstatrls.pdf`