A Closer Look at the Hill Estimator: Edgeworth Expansions and Confidence Intervals

Erich HAEUSLER
University of Giessen
http://www.uni-giessen.de

Johan SEGERS
Tilburg University
http://www.center.nl
INTRODUCTION

- Ordered sample $X_{1:n} \leq \cdots \leq X_{n:n}$ from Pareto-type cdf $F$
- Hill (1975) estimator for positive extreme-value index $\gamma$

$$\hat{H}_n(k) = \frac{1}{k} \sum_{i=1}^{k} \log X_{n-k+i:n} - \log X_{n-k:n}$$

- Simple and popular
- Asymptotic properties well known

$$\sqrt{k_n} \left( \frac{\hat{H}_n(k_n)}{\gamma} - 1 \right) - \mu_n \xrightarrow{d} N(0,1)$$

- intermediate sequence: $k_n \to \infty$, $k_n = o(n)$
- asymptotic bias $\mu_n = O(1)$, depends on $F$ and $k_n$
Confidence intervals and tests

- Confidence intervals and hypothesis tests less studied

- CI of nominal level $1 - \alpha$:
  
  symmetric CI : $\hat{H}_n(k) \left(1 \pm \frac{z}{\sqrt{k}}\right)$
  
  asymmetric CI : $\hat{H}_n(k) \left(1 \mp \frac{z}{\sqrt{k}}\right)$

  with

  $\Phi(z) = 1 - \alpha/2$

- Relevance:
  - Existence of moments
  - CI’s/tests for exceedance probabilities, quantiles, . . . [VANDEWALLE 2004]
Questions

- Which CI to be preferred?
- Yet other CI’s?
- Which $k$ to use for which CI?
- Comparisons between CI’s requires Edgeworth expansions

$$\Pr \left[ \sqrt{k_n} \left( \frac{\hat{H}_n(k)}{\gamma} - 1 \right) \leq x \right] = \Phi(x) + \text{error term}$$
Related literature

  - Useful for one-sided CI's [Cheng & Peng 2001]
  - Insufficiently accurate to analyse two-sided CI's
- Expansions in terms of Gamma distributions [Cheng & de Haan 2001; Guillou & Hall 2001]
  - Insufficiently accurate for two-sided CI's as well
- Note: If $\mu_n \neq o(1)$, then these CI's are inconsistent
  - This is the case for AMSE-minimizing $k_n$
  - Bias-corrected CI's in Ferreira & de Vries (2004)
For proper understanding...

- Won’t talk about:
  - Bias reduction
  - Data-driven methods to choose threshold
  - Comparisons with other estimators
  - Bayesian inference
  - Quantiles, exceedance probabilities
  - Other domains of attraction
  - Temporal dependence, non-stationarity, covariates

- Will talk about:
  - iid variables
  - Positive extreme-value index
  - Performance of various Hill-based CI’s/tests
  - Understanding of impact of intermediate sequence, nominal level, underlying distribution
Outline

- Inference in Pareto model
- CI's and hypothesis tests for extreme-value index
- Edgeworth expansions for normalized Hill estimator
- Main result
- Simulations
- Conclusion
PARETO MODEL

- Cdf and pdf of Pareto(1/γ):

\[ G_\gamma(x) = 1 - x^{-1/\gamma}, \]
\[ p_\gamma(x) = \frac{1}{\gamma} x^{-1-1/\gamma} \]

for \( x > 1 \)

- Inference on \( \gamma > 0 \) from iid \( Y_1, \ldots, Y_k \sim p_\gamma \)?
  - Estimation
  - Testing
  - Confidence intervals
**Likelihood computations**

- Log-likelihood of $\gamma$ given $Y_1, \ldots, Y_k$

$$\ell_k(\gamma) = \sum_{i=1}^{k} \log p_\gamma(Y_i) = -k \left( \frac{\hat{H}_k}{\gamma} + \log(\gamma) \right) + \text{constant}$$

$$\hat{H}_k = \frac{1}{k} \sum_{i=1}^{k} \log(Y_i)$$

- Score

$$\ell_k'(\gamma) = \frac{k}{\gamma} \left( \frac{\hat{H}_k}{\gamma} - 1 \right)$$

- Fisher information

$$I(\gamma) = \text{Var}_\gamma \left( \frac{\partial}{\partial \gamma} \log p_\gamma(Y) \right) = \frac{1}{\gamma^2}$$
MLE and deviance statistic

- \( \hat{H}_k \) is sufficient statistic and MLE for \( \gamma \)

\[
\sqrt{k}(\hat{H}_k - \gamma) \overset{d}{\rightarrow} N(0, \gamma^2), \quad k \to \infty
\]

- Deviance statistic (likelihood ratio) at \( \gamma \):

\[
D_k(\gamma) = 2 \left( \ell_k(\hat{H}_k) - \ell_k(\gamma) \right)
\]

\[
= 2k \left( \frac{\hat{H}_k}{\gamma} - 1 - \log \frac{\hat{H}_k}{\gamma} \right)
\]

\[
\overset{d}{\rightarrow} \chi^2_1, \quad k \to \infty
\]
### Hypothesis tests (1)

- **Test for** $H_0 : \gamma_0 = \gamma$ **versus** $H_1 : \gamma_0 \neq \gamma$ at nominal level $1 - \alpha$
- $z = z_{1-\alpha/2}$ standard-normal quantile $\Phi(z) = 1 - \alpha/2$
- Reject $H_0 : \gamma_0 = \gamma$ if $T_k(\gamma) > z^2$ where

<table>
<thead>
<tr>
<th>Test</th>
<th>Test statistic $T_k(\gamma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wald</td>
<td>$k \left( \frac{\hat{H}_k - \gamma}{\hat{H}_k} \right)^2$</td>
</tr>
<tr>
<td>Score</td>
<td>$k \left( \frac{\hat{H}_k - \gamma}{\gamma} \right)^2$</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td>$D_k(\gamma) = 2k \left( \frac{\hat{H}_k}{\gamma} - 1 - \log \frac{\hat{H}_k}{\gamma} \right)$</td>
</tr>
<tr>
<td>Bartlett-corrected LR</td>
<td>$D_k(\gamma) / \left( 1 + \frac{1}{6k} \right)$</td>
</tr>
</tbody>
</table>

**Faculty of Economics and Business Administration**
Hypothesis tests (2)

- Wald and score tests also Bartlett correctable
- One-sided tests: similarly
- Corresponding confidence intervals at nominal level $1 - \alpha$:

\[
\{ \text{All } \gamma > 0 \text{ for which } H_0 : \gamma_0 = \gamma \text{ is not rejected at level } 1 - \alpha \}
\]
CI’S AND TESTS FOR EVI

- Pareto domain of attraction
- cdf $F$ has extreme-value index $\gamma > 0$ iff

$$\Pr\left[\frac{X}{u} > x \mid X > u\right] = \frac{1 - F(ux)}{1 - F(u)} \rightarrow x^{-1/\gamma}, \quad u \to \infty$$

- Relative excesses over high thresholds are asymptotically Pareto($1/\gamma$) distributed
Hill estimator

- Heuristic:
  1. Take large threshold \( u = X_{n-k:n} \)
  2. Relative excesses \( Y_{i:k} = X_{n-k+i:n}/X_{n-k:n} \) for \( i = 1, \ldots, k \)
  3. Pretend \( Y_{1:k}, \ldots, Y_{k:k} \) are order statistics from iid Pareto\((1/\gamma)\) sample

- Pseudo-likelihood inference: Hill (1975)

\[
\hat{H}_n(k) = \frac{1}{k} \sum_{i=1}^{k} \log \frac{X_{n-k+i:n}}{X_{n-k:n}}
\]

- Other interpretations [Embrechts et al. 1997; Beirlant et al. 2004]
Hypothesis tests and CI’s (1)

- Fix $k$
- Reject $H_0 : \gamma_0 = \gamma$ at nominal level $1 - \alpha$ if $T_{n,k}(\gamma) > z^2_{1-\alpha/2}$

<table>
<thead>
<tr>
<th>Test</th>
<th>Test statistic $T_{n,k}(\gamma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wald</td>
<td>$k \left( \frac{\hat{H}_n(k) - \gamma}{\hat{H}_n(k)} \right)^2$</td>
</tr>
<tr>
<td>Score</td>
<td>$k \left( \frac{\hat{H}_n(k) - \gamma}{\gamma} \right)^2$</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td>$D_{n,k}(\gamma) = 2k \left( \frac{\hat{H}_n(k)}{\gamma} - 1 - \log \frac{\hat{H}_n(k)}{\gamma} \right)$</td>
</tr>
<tr>
<td>Bartlett-corrected LR</td>
<td>$D_{n,k}(\gamma) / \left( 1 + \frac{1}{6k} \right)$</td>
</tr>
</tbody>
</table>
Hypothesis tests and CI’s (2)

- Confidence intervals

\[ \{ \text{All } \gamma > 0 \text{ for which } H_0 : \gamma_0 = \gamma \text{ is not rejected} \} \]

- False rejection of \( H_0 : \gamma_0 = \gamma \) (type I error)

\[ \Pr[\text{False rejection}] = \alpha + \text{error term?} \]

- Not considered here but similar: false acceptance of wrong value (type II error)

- Will depend on:
  - type of interval
  - intermediate sequence \( k = k_n \)
  - nominal level
  - underlying distribution
We work under $H_0 : \gamma_0 = \gamma$ for fixed $\gamma > 0$

Intermediate sequence $k_n$

All test statistics can be expressed in terms of

$$H_n = \sqrt{k_n} \left( \frac{\hat{H}_n(k_n)}{\gamma} - 1 \right)$$

We’ll need expansions of the form

$$\Pr[H_n \leq x] = \Phi(x) + \text{error term}$$

*Edgeworth* expansions for the Hill estimator
Two sources of error

- Two reasons why $\Pr[H_n \leq x] \neq \Phi(x)$
  1. Relative excesses only *asymptotically* Pareto$(1/\gamma)$
  2. Even for Pareto$(1/\gamma)$, $H_n$ is standardized Gamma

- These sources of error may work in equal or in opposite directions

- To quantify first effect: higher-order regular variation
Tail quantile function

- Tail quantile function
  \[ V(y) = \inf \{ x \in \mathbb{R} : F(x) \geq 1 - 1/y \}, \quad y > 1 \]

- Domain-of-attraction condition equivalent to
  \[ \log V(ty) - \log V(t) = \gamma \log y + o(1), \quad t \to \infty \]
  for \( y > 0 \)

- Quantify \( o(1) \) to capture deviations from Pareto(1/\( \gamma \)) model
Higher-order regular variation

- Refine domain-of-attraction condition
- **Second-order regular variation:** as $t \to \infty$,

$$
\log V(ty) - \log V(t) = \gamma \log y + a(t)ch_\rho(y) + o(1),
$$

with $\rho \leq 0$, $a \in RV_\rho$ with $a(\infty) = 0$, $c \neq 0$, $h_\rho(y) = \int_1^y u^{\rho-1} du$ [Bingham et al. 1987; Geluk & de Haan 1987]

- **Third-order regular variation:** as $t \to \infty$,

$$
\log V(ty) - \log V(t) = \gamma \log y + a(t)ch_\rho(y) + a(t)b(t)\{B(y) + o(1)\}
$$

with $b \in RV_\tau$ for some $\tau \leq 0$ with $b(\infty) = 0$ and some specified form for $B(y)$ [De Haan & Stadtmüller 1996]
One-term Edgeworth expansion (1)

- Assume
  - Second-order regular variation
  - $\sqrt{k_n}a(n/k_n) = o(1)$

- Expansion of cdf of $H_n = \sqrt{k_n}\left\{ \hat{H}_n(k_n) - \gamma \right\}/\gamma$:

$$
\Pr[H_n \leq x] = \Phi(x) - \varphi(x) \left( \frac{1}{3\sqrt{k_n}} (1 - x^2) + \mu_n \right)
+ o \left( \frac{1}{\sqrt{k_n}} \right) + o(\mu_n)
$$

where

$$
\mu_n = \frac{c}{\gamma(1 - \rho)} \sqrt{k_n}a(n/k_n) \sim E_\infty[H_n]
$$
One-term Edgeworth expansion (2)

- Useful to analyze one-sided tests [Cheng & Peng 2001]
- Insufficient to compare two-sided tests: it only gives

\[
\Pr[\text{False rejection}] = \alpha + o\left(\frac{1}{\sqrt{k_n}}\right) + o(\mu_n)
\]

- Need for higher-order expansions of \(\Pr[H_n \leq x]\)
Two-term Edgeworth expansion

- Assume
  - Third-order regular variation
  - $\sqrt{k_n} a(n/k_n) = o(1)$

- Expansion of cdf of normalized Hill estimator:

$$
\Pr[H_n \leq x] = \Phi(x) - \varphi(x) \left( \frac{1}{3\sqrt{k_n}} (1 - x^2) + \mu_n \right) \\
-x \varphi(x) \left\{ \frac{1}{k_n} P_1(x) + \frac{\mu_n}{\sqrt{k_n}} \left( P_2(x) + \frac{1}{1 - \rho} \right) + \frac{1}{2} \mu_n^2 \right\} \\
+ o\left( \frac{1}{k_n} \right) + o(\mu_n^2) + o(|\mu_n| b(n/k_n))
$$

for known polynomials $P_1$ and $P_2$

- Special case of Cuntz, Haeusler & Segers (2003)
MAIN RESULT

- Assume
  - Third-order regular variation
  - $\sqrt{k_n a(n/k_n)} = o(1)$

- For the four tests considered earlier:

$$\Pr[\text{False rejection}] = \alpha + z \varphi(z) \left\{ \frac{1}{k_n} Q_1(z) + \frac{\mu_n}{\sqrt{k_n}} \left( Q_2(z) + \frac{2\rho}{1 - \rho} \right) + \mu_n^2 \right\}$$

$$+ o \left( \frac{1}{k_n} \right) + o(\mu_n^2) + o(|\mu_n| b(n/k_n))$$

where

- $\Phi(z) = 1 - \alpha/2$
- known polynomials $Q_1$ and $Q_2$, depending on the test
Comments

- Error term may disappear for zero, one or two values of $k_n$
  - Finding such $k_n$ requires estimation of second-order parameters
- LR and Bartlett-corrected LR tests very close
- Small $k_n$: LR tests most accurate
- Larger $k_n$: Cancellation effect may favor Wald or score tests
- $k_n$ too large: too large bias makes tests inconsistent
  - Bias-corrected intervals: see Ferreira & de Vries (2004)
**SIMULATIONS**

- **Burr distribution**, parameters $\rho < 0 < \gamma$:
  
  $$F(x) = 1 - \left( x^{-\rho/\gamma} + 1 \right)^{1/\rho}, \quad x > 0$$

- Third-order regularly varying: $c = \gamma$, $\tau = \rho$, $a(t) = b(t) = t^\rho$

- Predicted and simulated type I errors of
  - Wald test
  - Score test
  - Likelihood ratio test
  - Bartlett-corrected LR test

- Settings:
  - $\gamma = 1$ and $\rho \in \{-1, -0.5\}$
  - nominal type I error $\alpha \in \{0.1, 0.05\}$
  - sample size 500
  - 10,000 samples
Simulations: $\rho = -1, \alpha = 0.1$

![Graph showing Type I error vs. k for different tests](image)

- $\gamma = 1$, $\rho = -1$, $\alpha = 0.1$

**Faculteit van Economie en Bedrijfswetenschappen**
Simulations: \( \rho = -1, \alpha = 0.05 \)

\[ \gamma = 1, \rho = -1, \alpha = 0.05 \]
Simulations: \( \rho = -0.5, \alpha = 0.1 \)

\[ \gamma = 1, \rho = -0.5, \alpha = 0.1 \]
Simulations: $\rho = -0.5, \alpha = 0.05$

$\gamma = 1, \rho = -0.5, \alpha = 0.05$
CONCLUSION

- Inference on positive extreme-value index
- Pseudo-likelihood in Pareto model for relative excesses
  ⇒ Various CI’s/tests:
    - Wald
    - Score
    - Likelihood ratio
    - Bartlett corrections
- Expansions for type I error of two-sided tests
- Tool: two-term Edgeworth expansion for Hill estimator
- Good match between predicted and simulated type I error
- Performance of CI’s and tests depends on
  - nominal level
  - threshold
  - underlying distribution
  - type of CI/test
Thank you!
References (1)

References (2)