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

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

INTRODUCTION

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

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

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

- 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

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- 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 : $\hat{H}_n(k) \Big/ \left(1 \mp \frac{z}{\sqrt{k}}\right)$ with

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

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

Questions

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- 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

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- One-term Edgeworth expansions in CHENG & PAN (1998) and CHENG & PENG (2001)
  - 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. . .

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- 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

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- 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

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- Cdf and pdf of Pareto($1/\gamma$):

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

  for $x > 1$
- Inference on $\gamma > 0$ from iid $Y_1, \ldots, Y_k \sim p_\gamma$?
  - Estimation
  - Testing
  - Confidence intervals

Likelihood computations

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- Log-likelihood of γ 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

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

- Fisher information

$$I(\gamma) = \mathrm{Var}_\gamma\left(\frac{\partial}{\partial\gamma}\log p_\gamma(Y)\right) = \frac{1}{\gamma^2}$$

MLE and deviance statistic

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- $\hat{H}_k$ is sufficient statistic and MLE for γ

$$\sqrt{k}(\hat{H}_k - \gamma) \xrightarrow{d} N(0, \gamma^2), \qquad k \to \infty$$

- Deviance statistic (likelihood ratio) at γ:

$$\begin{aligned} D_k(\gamma) &= 2\left(\ell_k(\hat{H}_k) - \ell_k(\gamma)\right) \\ &= 2k\left(\frac{\hat{H}_k}{\gamma} - 1 - \log\frac{\hat{H}_k}{\gamma}\right) \\ &\xrightarrow{d} \chi_1^2, \qquad k \to \infty \end{aligned}$$

Hypothesis tests (1)

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- 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

Test	*Test statistic* $T_k(\gamma)$
Wald	$k\left(\frac{\hat{H}_k - \gamma}{\hat{H}_k}\right)^2$
Score	$k\left(\frac{\hat{H}_k - \gamma}{\gamma}\right)^2$
Likelihood ratio	$D_k(\gamma) = 2k\left(\frac{\hat{H}_k}{\gamma} - 1 - \log\frac{\hat{H}_k}{\gamma}\right)$
Bartlett-corrected LR	$D_k(\gamma)\Big/\left(1 + \frac{1}{6k}\right)$

Hypothesis tests (2)

● INTRODUCTION
● PARETO MODEL
● CI'S AND TESTS
● EDGEWORTH EXPANSIONS
● MAIN RESULT
● SIMULATIONS
● CONCLUSION

- 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

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

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

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

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

Hill estimator

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- 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)

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

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

Test	*Test statistic* $T_{n,k}(\gamma)$
Wald	$k\left(\frac{\hat{H}_n(k) - \gamma}{\hat{H}_n(k)}\right)^2$
Score	$k\left(\frac{\hat{H}_n(k) - \gamma}{\gamma}\right)^2$
Likelihood ratio	$D_{n,k}(\gamma) = 2k\left(\frac{\hat{H}_n(k)}{\gamma} - 1 - \log\frac{\hat{H}_n(k)}{\gamma}\right)$
Bartlett-corrected LR	$D_{n,k}(\gamma)\Big/\left(1 + \frac{1}{6k}\right)$

Hypothesis tests and CI's (2)

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- 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

EDGEWORTH EXPANSIONS

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- 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

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- 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

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- Tail quantile function

$$V(y) = \inf\{x \in \mathbb{R} : F(x) \geq 1 - 1/y\}, \qquad y > 1$$

- Domain-of-attraction condition equivalent to

$$\log V(ty) - \log V(t) = \gamma \log y + o(1), \qquad t \to \infty$$

for $y > 0$

- Quantify o(1) to capture deviations from Pareto($1/\gamma$) model

Higher-order regular variation

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

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

$$\begin{aligned} &\log V(ty) - \log V(t) \\ &\quad = \gamma \log y + a(t) c h_\rho(y) + o(1), \end{aligned}$$

with $\rho \leq 0$, $a \in \mathrm{RV}_\rho$ with $a(\infty) = 0$, $c \neq 0$, $h_\rho(y) = \int_1^y u^{\rho-1} \mathrm{d}u$
[BINGHAM ET AL. 1987; GELUK & DE HAAN 1987]

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

$$\begin{aligned} &\log V(ty) - \log V(t) \\ &\quad = \gamma \log y + a(t) c h_\rho(y) + a(t) b(t) \{B(y) + o(1)\} \end{aligned}$$

with $b \in \mathrm{RV}_\tau$ 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)

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- Assume
  - Second-order regular variation
  - $\sqrt{k_n} a(n/k_n) = o(1)$
- Expansion of cdf of $H_n = \sqrt{k_n}\{\hat{H}_n(k_n) - \gamma\}/\gamma$:

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

where

$$\mu_n = \frac{c}{\gamma(1-\rho)}\sqrt{k_n} a(n/k_n) \sim \mathrm{E}_\infty[H_n]$$

One-term Edgeworth expansion (2)

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- Slight generalization of Cheng & Pan (1998) and Cheng & Peng (2001)
- 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

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- Assume
  - Third-order regular variation
  - $\sqrt{k_n} a(n/k_n) = o(1)$
- Expansion of cdf of normalized Hill estimator:

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

  for known polynomials P_1 and P_2
- Special case of CUNTZ, HAEUSLER & SEGERS (2003)

MAIN RESULT

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- Assume
  - Third-order regular variation
  - $\sqrt{k_n}a(n/k_n) = o(1)$
- For the four tests considered earlier:

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

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

Comments

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- 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

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

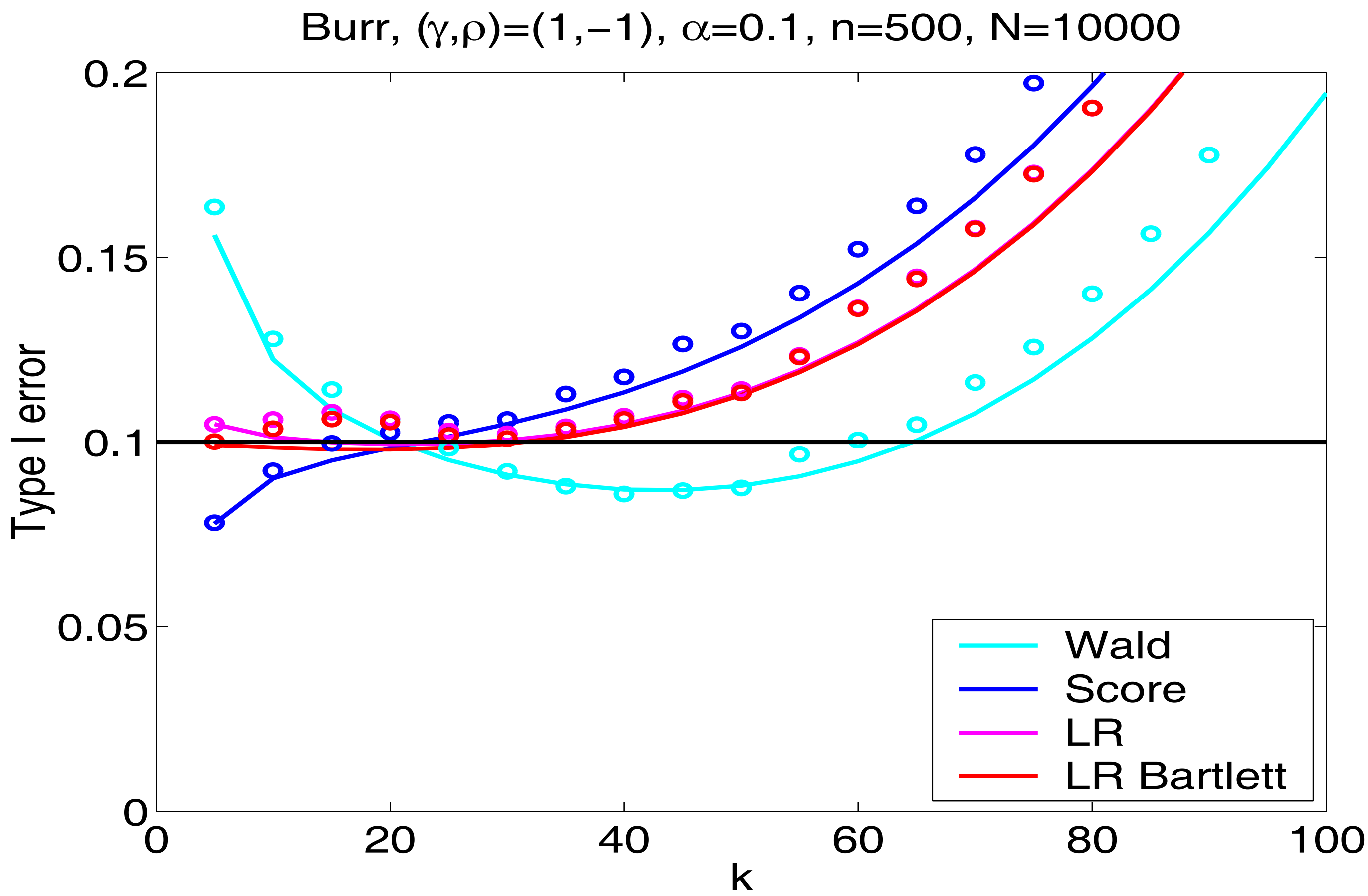
- *Burr* distribution, parameters $\rho < 0 < \gamma$:

$$F(x) = 1 - \left(x^{-\rho/\gamma} + 1\right)^{1/\rho}, \qquad x > 0$$

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

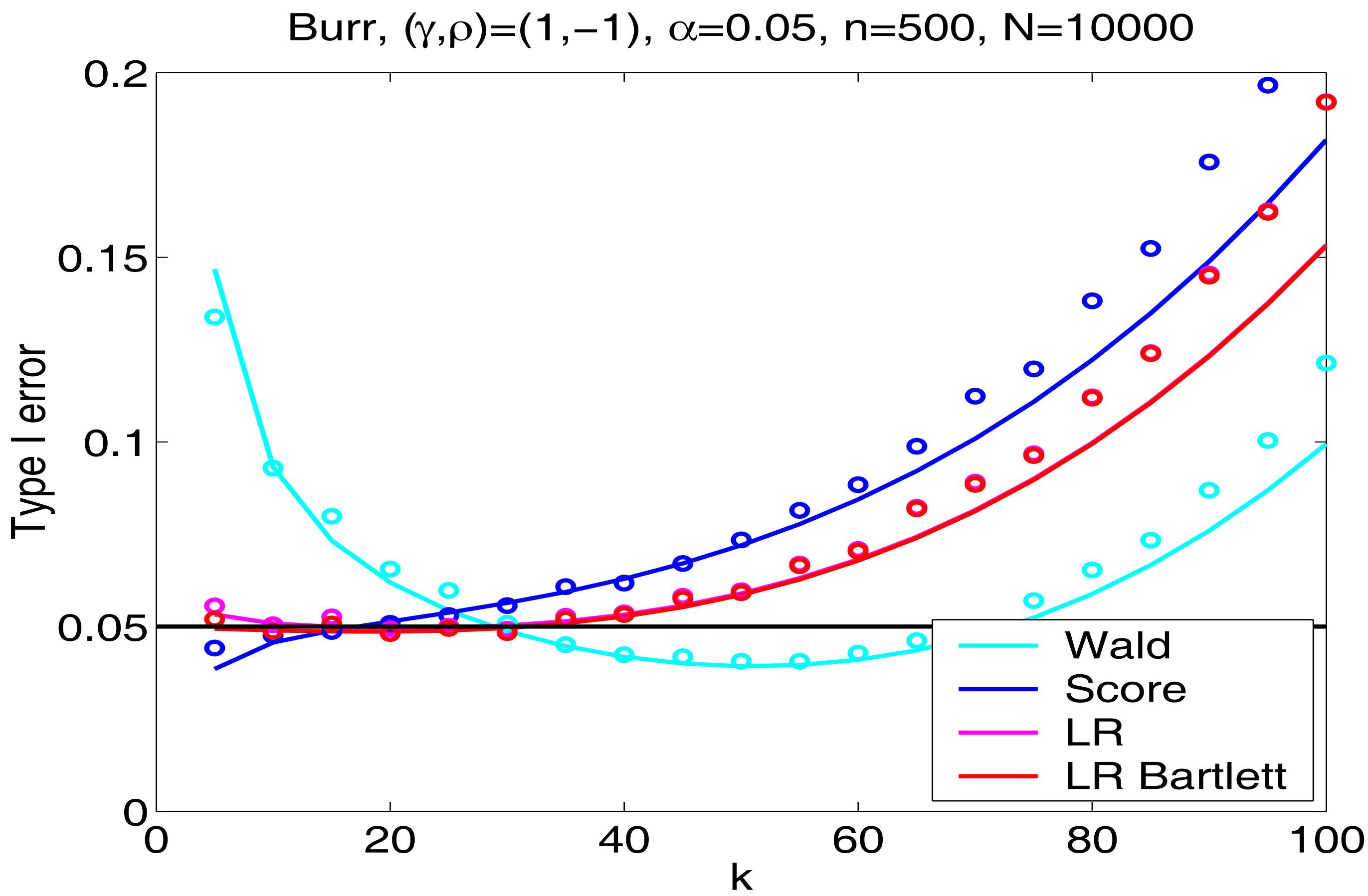
- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION



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

Simulations: $\rho = -1$, $\alpha = 0.05$

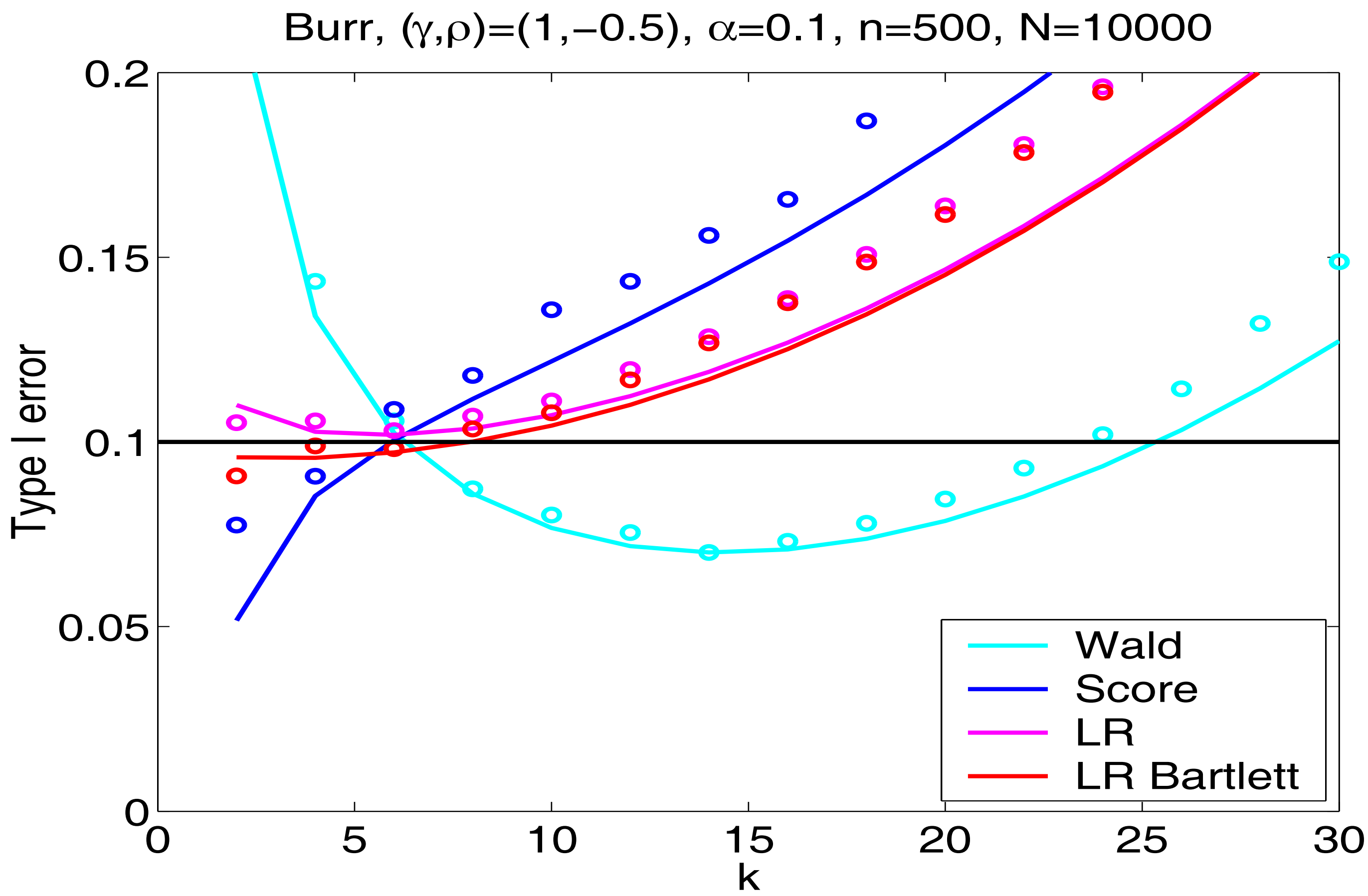
- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION



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

Simulations: $\rho = -0.5$, $\alpha = 0.1$

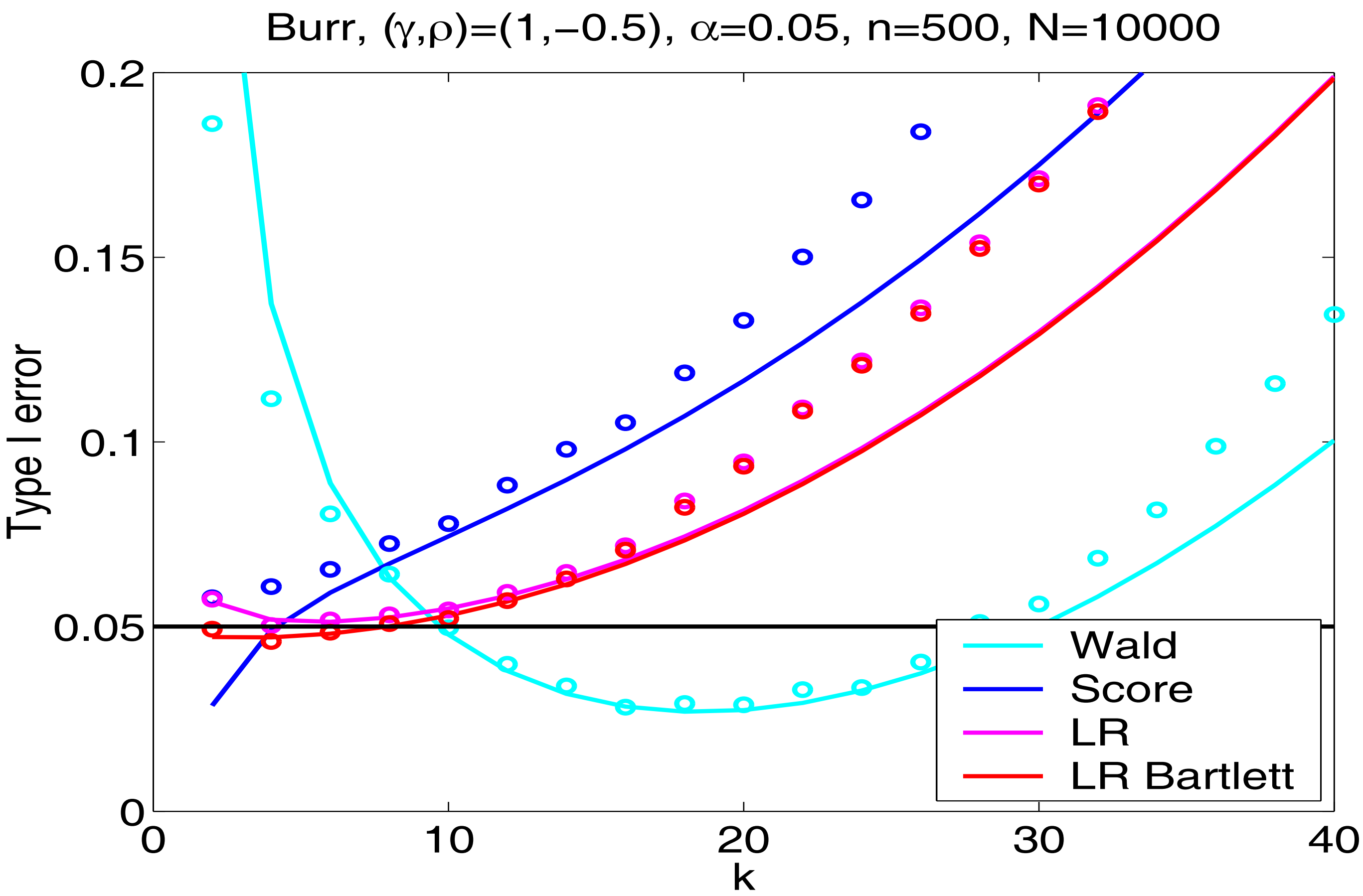
- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION



$\gamma = 1$, $\rho = -0.5$, $\alpha = 0.1$

Simulations: $\rho = -0.5$, $\alpha = 0.05$

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION



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

CONCLUSION

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- Inference on positive extreme-value index
- Pseudo-likelihood in Pareto model for relative excesses
  $\Rightarrow$ 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

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

Thank you!

References (1)

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- Beirlant, J, Goegebeur, Y, Segers, J & Teugels, J (2004) *Statistics of Extremes: Theory and Applications.* Wiley
- Bingham, NH, Goldie, CM & Teugels, JL (1987) *Regular Variation.* Cambridge University Press
- Cheng, S & Pan, J (1998) *Scand. J. Statist.* 25, 717–728
- Cheng, S & Peng, L (2001) *Bernoulli* 7, 751–760
- Cuntz, A, Haeusler, E & Segers, J (2003) CentER discussion paper 2003-08, `center.uvt.nl/pub/dp`
- Embrechts, P, Klüppelberg, C & Mikosch, T (1997) *Modelling Extremal Events for Insurance and Finance.* Springer-Verlag
- Ferreira, A & de Vries, CG (2004), `www.tinbergen.nl`
- Geluk, J & de Haan, L (1987) *Regular variation, extensions and Tauberian theorems.* CWI Tract 40

References (2)

- INTRODUCTION
- PARETO MODEL
- CI'S AND TESTS
- EDGEWORTH EXPANSIONS
- MAIN RESULT
- SIMULATIONS
- CONCLUSION

- Guillou, A & Hall, P (2001) *J. R. Statist. Soc.* B 63, 293–305
- de Haan, L & Stadtmüller, U (1996) *J. Austral. Math. Soc., Ser. A* 61, 381–395
- Hall, P (1992) *The Bootstrap and Edgeworth Expansion.* Springer-Verlag
- Hill, BM (1975) *Ann. Statist.* 3, 1163–1174
- Vandewalle, B (2004) PhD thesis, Catholic University Leuven