Department of Mathematics and Systems Analysis

On Multivariate Extremes

Matias Heikkilä

Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 49/2019

On Multivariate Extremes

Matias Heikkilä

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Science, at a public examination held at the lecture hall M1 of the school on 29 April 2017 at 1pm.

Aalto University
School of Science
Department of Mathematics and Systems Analysis
Mathematics

Supervising professor
Assistant Professor Pauliina Ilmonen, Aalto University School of Science

Preliminary examiners
Assistant Professor Anna Kiriliouk, University of Namur, Belgium
Assistant Professor Christophe Ley, Ghent University, Belgium

Opponent
Assistant Professor Anna Kiriliouk, University of Namur, Belgium

Aalto University publication series
DOCTORAL DISSERTATIONS 49/2019

ISBN 978-952-60-8465-7 (printed)
ISBN 978-952-60-8466-4 (pdf)
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-8466-4

Unigrafia Oy
Helsinki 2019

Finland

Printed matter
4041-0619

Aalto University

Abstract

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Author
Matias Heikkilä

Name of the doctoral dissertation
On Multivariate Extremes

Publisher School of Science

Unit Department of Mathematics and Systems Analysis

Series Aalto University publication series DOCTORAL DISSERTATIONS 49/2019

Field of research Mathematics

Manuscript submitted 12 December 2018 **Date of the defence** 29 April 2019

Permission to publish granted (date) 25 February 2019 **Language** English

☐ **Monograph** ☒ **Article dissertation** ☐ **Essay dissertation**

Abstract

Multivariate extreme value theory has traditionally been developed by studying the componentwise maxima of the observations. Recently, alternative approaches based on measuring the outlyingness of observations, using for example the Mahalanobis distance from the center of the distribution, have been proposed. This thesis features efforts to advance this point of view by, both, developing new results for the previously published methods and proposing new methods.

The existing methods, the separating multivariate Hill estimator in particular, are developed further by settling important open questions concerning their asymptotic properties. As a salient example: We show that, under natural conditions, the separating Hill estimator is both consistent and asymptotically normal not only when the parameters of the underlying distribution are known, but in the practical case when the observations follow an unknown elliptical distribution.

As a somewhat tangential aspect, we study a new way, the Delaunay outlyingness, to measure the outlyingness of multivariate observations that is based on the geometry of the sample. We study Delaunay outlyingness in the case of a compact convex region and a finite number of outliers and show, that at least in this situation, it has favorable asymptotic properties. Delaunay outlyingness is also studied through simulations which suggest that it's applicable beyond compact convex regions.

Keywords extreme value theory, multivariate extremes

ISBN (printed) 978-952-60-8465-7 **ISBN (pdf)** 978-952-60-8466-4

ISSN (printed) 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2019

Pages 114 **urn** http://urn.fi/URN:ISBN:978-952-60-8466-4

Aalto-yliopisto

Tiivistelmä

Aalto-yliopisto, PL 11000, 00076 Aalto www.aalto.fi

Tekijä
Matias Heikkilä

Väitöskirjan nimi
Moniulotteisista ääriarvoista

Julkaisija Perustieteiden korkeakoulu

Yksikkö Matematiikan ja systeemianalyysin laitos

Sarja Aalto University publication series DOCTORAL DISSERTATIONS 49/2019

Tutkimusala Matematiikka

Käsikirjoituksen pvm 12.12.2018 **Väitöspäivä** 29.04.2019

Julkaisuluvan myöntämispäivä 25.02.2019 **Kieli** Englanti

☐ **Monografia** ☒ **Artikkeliväitöskirja** ☐ **Esseeväitöskirja**

Tiivistelmä

Perinteisesti moniulotteinen ääriarvoteoria on perustunut komponenteittain otettujen maksimien tarkasteluun. Hiljattain ilmestyneissä artikkelieissa on tarkasteltu vaihtoehtoista tapaa lähestyä ongelmaa mittaamalla havaintojen äärimmäisyyttä datassa perustuen esimerkiksi havainnon Mahalanobis-etäisyyteen jakauman keskipisteestä. Väitöskirjassa esitellään tätä näkökulmaa edistäviä artikkeleja, joissa sekä edistetään olemassa olevien menetelmien teoriaa, että ehdotetaan uusia tapoja lähestyä moniulotteisia ääriarvoja.

Olemassa olevia menetelmiä, erityisesti erottavaa moniulotteista Hill-estimaattoria, edistetään ratkaisemalla niiden asymptotiikkaan liittyviä merkittäviä avoimia ongelmia. Huomattavana esimerkkinä mainittakoon, että, luonnollisten oletusten vallitessa, erottava Hill-estimaattori on konsistentti ja asymptoottisesti normaali sekä silloin, kun taustalla olevan jakauman parametrit tunnetaan, että käytännön kannalta keskeisessä tilanteessa, kun havainnot syntyvät tuntemattomasta elliptisestä jakaumasta.

Jokseenkin rinnakkaisena aiheena tarkastelemme uutta tapaa, Delaunay-ulkoisuutta, mitata moniulotteisten havaintojen äärimmäisyyttä perustuen havaintojen joukon geometriseen rakenteeseen. Tarkastelemme Delaunay-ulkoisuutta kompaktin konveksin joukon tapauksessa, kun joukon ympärille asetetaan äärellinen määrä ääripisteitä ja näytämme, että vähintään tässä tilanteessa suureella on hyvät asymptoottiset ominaisuudet. Delaunay ulkoisuutta tarkastellaan myös simulaatioiden avulla, jotka antavat näyttöä sen puolesta, että suure soveltuu myös tilanteisiin jossa havainnot eivät ole jakautuneet kompaktin konveksin joukon yli.

Avainsanat moniulotteinen ääriarvoteoria

ISBN (painettu) 978-952-60-8465-7 **ISBN (pdf)** 978-952-60-8466-4

ISSN (painettu) 1799-4934 **ISSN (pdf)** 1799-4942

Julkaisupaikka Helsinki **Painopaikka** Helsinki **Vuosi** 2019

Sivumäärä 114 **urn** http://urn.fi/URN:ISBN:978-952-60-8466-4

Preface

Completing this thesis has been an amazing journey and to have been given the opportunity has been a tremendous privilege. Thank you Prof. Pauliina Ilmonen for trusting me with the position of doctoral candidate, and thank you for your earnest guidance and your invaluable advice throughout the process.

I would like to thank the preliminary examinors Prof. Anna Kiriliouk and Prof. Christophe Ley for their careful examination of this thesis and their valuable comments and feedback. I would also like to thank Prof. Kiriliouk for agreeing to act as my opponent.

Thank you Dr. Yves Dominicy and Prof. David Veredas for the delightful opportunity to collaborate throughout these years. I'm also very grateful for Dr. Dominicy and other members of ECARES, Université libre de Bruxelles for providing me the opportunity to work in Brussels.

I acknowledge the financial support from the Magnus Ehrnrooth Foundation that has been crucial for the completion of this thesis. I'm also grateful for all members of Department of Mathematics and Systems Analysis, Aalto University School of Science for providing fantastic circumstances for conducting my research and a friendly, welcoming place to work in. Thank you Niko Liétzen, my long lost second cousin, and Sami Helander for sharing the office with me all these years. I would also like to thank the rest of the entire group of Prof. Ilmonen and I wish them continued success in the future.

Thank you friends and family for your affection and support not only throughout my doctoral studies, but also before them. I would especially like to thank you Mom for working so hard to make sure that a pretty smart, but a bit too hyperactive, kid got through the elementary school

alright. Thank you members of φ for the awesome years at the University of Turku and for years worth of intellectually stimulating discourse in a safe and friendly environment.

Last, but not least: Thank you Sanna for your love and support. Achieving something means a lot more, when I get to do it with you.

Helsinki, March 1, 2019,

Matias Heikkilä

Contents

List of publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I M. Heikkilä, Y. Dominicy, P. Ilmonen. Multivariate moment based extreme value index estimators. *Computational Statistics*, Volume 32, Issue 4, pp 1481–1513, December 2017.

II M. Heikkilä, Y. Dominicy, P. Ilmonen. On multivariate separating Hill estimator under estimated location and scatter. Accepted for publication in *Statistics: A Journal of Theoretical and Applied Statistics*, 18 pp., September 2018.

III M. Heikkilä. Nonparametric Geometric Outlier Detection. Submitted for publication, available at *arXiv:1811.05169*, 22pp., June 2017.

Author's contribution

Publication I: "Multivariate moment based extreme value index estimators"

The idea for the paper came from Ilmonen. Heikkilä derived most of the theoretical results and contributed to implementing the simulations. Dominicy was responsible of the simulations. Heikkilä and Dominicy designed and implemented the real-data examples. Heikkilä wrote the first version of the manuscript. All authors contributed to revising and writing the final version of the manuscript.

Publication II: "On multivariate separating Hill estimator under estimated location and scatter"

The idea for the paper came from Heikkilä and Ilmonen. Heikkilä derived all the theoretical results and contributed to implementing the simulations. Dominicy was responsible of the simulations. Heikkilä and Dominicy designed and implemented the real-data examples. Heikkilä wrote the first version of the manuscript. All authors contributed to revising and writing the final version of the manuscript.

Publication III: "Nonparametric Geometric Outlier Detection"

This is a single-author paper by Heikkilä.

Author's contribution

1. Introduction

Extreme value theory is often introduced through an extrapolation problem: Dikes are built to protect an area from flooding. Based on decades of storm data, how high should they be in order for flooding to only occur with some low probability p? With p sufficiently small, the corresponding quantile of the probability distribution may fall beyond the existing sample and there is no direct way to estimate it. Extreme value theory serves as a rigorous framework for approaching these types of problems in a theoretically justified way.

Extreme value theory includes a probabilistic aspect that's concerned, among other things, with classifying distributions according to the behavior of the extreme observations. The univariate probabilistic extreme value theory started to approach it's contemporary shape during the early 20th century in the works of Fréchet [24], Fisher and Tippett [22], von Mises [39] and, finally yielding a succinct theorem by Gnedenko [27]. Study of the statistical theory, on the other hand, was initiated by Pickands [32] and continued, among many others, by Hill [29].

An immediate obstacle in studying extreme value theory in the multivariate context is the lack of a natural order relation and, consequentially, the lack of a canonical way to decide which observations should be considered extreme. Multivariate extreme value theory is traditionally developed from the point of view of componentwise sample maxima and their limiting distributions. In [15], however, Dominicy, Ilmonen and Veredas chose a different approach to the problem by restricting to the study of elliptical distributions. In this special case the extremity of an observation has a natural interpretation as the value of the generating variate of the distribution, i.e. the Mahalanobis distance of the observation from

the center of the distribution. This point of view is closely related to the notion of statistical depth (the Mahalanobis distance from the center is closely related to the so called Mahalanobis depth). In [15] this approach was used to define multivariate Hill estimators and their properties were studied both analytically and empirically.

The unifying theme in this thesis is an effort to further the study of multivariate extremes by, instead of studying the componentwise maxima, using an intermediary quantity such as the Mahalanobis distance to enable univariate theory to be applied. The articles Publication I and Publication II are closely related to [15]: In Publication I similar generalizations of extreme value index estimators as the ones defined in [15] are studied empirically through a Monte Carlo study. In Publication II, the asymptotic theory of the separating Hill estimator is developed beyond [15].

In Publication III, a new quantity for measuring the outlyingness of observations is proposed. This quantity, the Delaunay outlyingness, is based on the geometry of the sample and reflects its structure in a way somewhat similar to k-NN methods (see e.g. [6]), but doesn't involve arbitrary parameters.

2. On univariate extreme value theory

In this section, we review certain results from the univariate extreme value theory. The results have previously been presented in [10] and the references therein.

2.1 Extreme value index

Let F be a distribution function. Assume that there is a distribution function G and sequences $a_n > 0$ and b_n real such that

$$F(a_n x + b_n)^n \to G(x) \tag{2.1}$$

for all continuity points x of G. Such distribution G is called an *extreme value distribution*. Notice that the above condition can be stated in terms of sample maxima $M_n = \max_{1 \leq i \leq n} X_i$ of i.i.d. random variables X_i that follow the distribution F: It's equivalent to

$$\frac{M_n - b_n}{a_n} \to_D G.$$

The Fisher-Tippet-Gnedenko theorem shows the class of distributions G is a simple one-parameter family.

Theorem 2.1 (Fisher-Tippet-Gnedenko)**.** *The class of extreme value distributions is $G_\gamma(ax + b)$ with $a > 0$ and b real, where*

$$G_\gamma(x) = \exp\left(-(1 + \gamma x)^{-1/\gamma}\right), \quad 1 + \gamma x > 0,$$

with γ real and where for $\gamma = 0$ the right hand side is interpreted as $\exp(-e^{-x})$.

The parameter γ that appears in Theorem 2.1 is called the *extreme value index* of F. A distribution with an extreme value index γ is said to be in

the *domain of attraction* of G_γ, denoted $F \in \mathcal{D}(G_\gamma)$. There are results that characterise distributions F with $\mathcal{D}(G_\gamma)$ for different values of γ (see e.g. [9]).

Theorem 2.2. *The distribution function F is in the domain of attraction of the extreme value distribution G_γ if and only if*

1. *for $\gamma > 0$: $x^* = \sup\{x \mid F(x) < 1\}$ is infinite and*

$$\lim_{t\to\infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\gamma}$$

for all $x > 0$.

2. *for $\gamma < 0$: x^* is finite and*

$$\lim_{t\downarrow 0} \frac{1 - F(x^* - tx)}{1 - F(x^* - t)} = x^{-1/\gamma}$$

for all $x > 0$.

3. *for $\gamma = 0$: x^* can be finite or infinite and*

$$\lim_{t\uparrow x^*} \frac{1 - F(t + xf(t))}{1 - F(t)} = e^{-x}$$

for all real x, where f is a suitable positive function. If the above equation holds for some f, then $\int_t^{x^*}(1 - F(s))\,\mathrm{d}s < \infty$ for $t < x^*$ and the above equation holds with

$$f(t) = \frac{\int_t^{x^*}(1 - F(s))\,\mathrm{d}s}{1 - F(t)}.$$

An alternative formulation of the condition $F \in \mathcal{D}(G_\gamma)$ can be expressed in terms of a quantile function U:

$$U = \left(\frac{1}{1 - F}\right)^{\leftarrow}, \tag{2.2}$$

where $f^{\leftarrow}$ denotes the left-continuous inverse of f defined as

$$f^{\leftarrow}(y) = \inf\{x \mid f(x) \geq y\}.$$

It turns out that $F \in \mathcal{D}(G_\gamma)$ is equivalent to

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}, \tag{2.3}$$

for each $x > 0$, where $a(t) = a_{\lfloor t \rfloor}$ is called the scale function. Relation (2.3) hints towards important applications of the extreme value theory in risk analysis as it suggests the following approximation for the quantile function U

$$U(x) \approx U(t) + a(t)\frac{\left(\frac{x}{t}\right)^{\gamma} - 1}{\gamma} \tag{2.4}$$

for $t > 0$ large.

The approximation (2.4) and suggests that the extreme value index γ, the quantile function U and the scale function a could be used to tackle the seemingly intractable problem of estimating probabilities and quantiles beyond the observed data. This turns out to be true in the sense that the approximation (2.4) can be refined to an estimator with desirable asymptotic properties. Empirical problems where extreme value theory is used for extrapolation include estimating the maximum human lifespan [1] [17] [26] [33], sport records [44] [42] [19] [18], earthquake intensity [38][4] and extreme wind speeds [2] [7] [28] [8].

Noticing that the approximation (2.4) depends on the extreme value index γ of the distribution F, the practitioner now faces the problem of estimating γ. One of the earliest such estimators is the Hill estimator

$$\hat{\gamma}_H = \frac{1}{k_n}\sum_{i=0}^{k_n-1} \log X_{n-i,n} - \log X_{k_n,n} \tag{2.5}$$

where k_n is threshold parameter with $k_n/n \to 0$, $k_n \to \infty$ as $n \to \infty$. In [29], where the Hill estimator was proposed, it was derived as a maximum likelihood estimator for the tail-index $\alpha > 0$ of a Pareto distribution $F(x) = 1 - x^{-\alpha}$ and was later shown to be valid for a wider class of distributions as discussed below. However, by definition, the Hill estimator can never be valid in case $\gamma < 0$. Estimators that overcome this limitation include the Pickands estimator

$$\hat{\gamma}_P = \frac{1}{\log 2}\log\left(\frac{X_{n-k_n,n} - X_{n-2k_n,n}}{X_{n-2k_n,n} - X_{n-4k_n,n}}\right), \tag{2.6}$$

introduced in [32], and the moment estimator

$$\hat{\gamma}_M = M_n^{(1)} + 1 - \frac{1}{2}\left(1 - \frac{\left(M_n^{(1)}\right)^2}{M_n^{(2)}}\right), \tag{2.7}$$

introduced in [13], where

$$M_n^{(i)} = \frac{1}{k_n}\sum_{i=0}^{k_n-1} \left(\log X_{n-i,n} - \log X_{k_n,n}\right)^i.$$

In addition to these, several other extreme value index estimators have been proposed [40] [30] [20].

A comprehensive asymptotic theory for the estimators (2.5), (2.6) and (2.7) can be developed on closer examination of relation (2.3). It turns out that the limit (2.3) defines a known class of functions: functions of *extended regular variation*.

2.2 The classes ERV_γ and $2ERV_{\gamma,\rho}$

Definition 2.3. A function $f : \mathbb{R}_+ \to \mathbb{R}$ is said to be of extended regular variation if there is a function $a : \mathbb{R}_+ \to \mathbb{R}_+$ such that for some γ and all $x > 0$

$$\lim_{t\to\infty} \frac{f(tx) - f(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}.$$

In case $\gamma = 0$ the right-hand side is interpreted as $\log x$.

Notation: $f \in ERV_\gamma$. Hence, according to relation (2.3), the conditions $F \in \mathcal{D}(G_\gamma)$ and $U \in ERV_\gamma$ are equivalent. Extended regular variation is a special case of another property known as regular variation. There exists extensive literature on regularly varying functions [3], [25].

Remark 2.4. The function class ERV_γ is natural in the sense that the function $(x^\gamma - 1)/\gamma$ is the only limit ψ appearing in

$$\psi(x) = \lim_{t\to\infty} \frac{f(tx) - f(t)}{a(t)},$$

where f and a are as in Definition 2.3, that is not constant with respect to x.

Hence, the class ERV_γ is useful in characterising the domains of attraction $\mathcal{D}(G_\gamma)$; it turns out that a second-order version of the class ERV_γ is useful in studying the asymptotic properties of several extreme value index estimators, as it provides access to the distribution of the difference between quantiles and order statistics $X_{n-k_n,n} - U\left(\frac{n}{k_n}\right)$ for k_n sufficiently small compared to n.

Theorem 2.5. *Suppose that for some measurable function f and positive functions a and A, with $A(t) \to 0$ as $t \to \infty$, the limit*

$$H(x) = \lim_{t\to\infty} \frac{\frac{f(tx)-f(t)}{a(t)} - c\frac{x^\gamma - 1}{\gamma}}{A(t)}$$

exists for all $x > 0$ *and is not a multiple of* $(x^\gamma - 1)/\gamma$. *There exists real constants* c_1, c_2 *and a parameter* $\rho \leq 0$ *such that for all* $x > 0$,

$$H(x) = c_1 \int_1^x s^{\gamma-1} \int_1^s u^{\rho-1} \mathrm{d}u \, \mathrm{d}s + c_2 \int_1^x s^{\gamma+\rho-1} \mathrm{d}s.$$

Moreover, for $x > 0$

$$\lim_{t\to\infty} \frac{\frac{a(tx)}{a(t)} - x^\gamma}{A(t)} = c_1 x^\gamma \frac{x^\rho - 1}{\rho}$$

and

$$\lim_{t\to\infty} \frac{A(tx)}{A(t)} = x^\rho.$$

Definition 2.6. A measurable function f satisfying the conditions of Theorem 2.5 is said to be of *second-order extended regular variation*, denoted $f \in 2ERV_{\gamma,\rho}$, where $\rho \leq 0$ and γ are as in Theorem 2.5. A distribution function F is said to satisfy the *second-order extreme value condition* (for γ and ρ) if $U \in 2ERV_{\gamma,\rho}$.

Second-order extended regular variation was studied e.g. in [12] and [11]. Second-order extended regular variation enables results that describe the difference between the observed intermediate order statistics of the distribution and the actual quantiles of the distribution. The following result due to [16] can be understood as a description of how the errors $X_{n-i,n} - U\left(\frac{n}{k_n}\right)$ with $1 \leq i \leq k_n$ behave.

Theorem 2.7. *Suppose* $X_1, X_2, \ldots$ *are i.i.d. random variables with* F *satisfying the second-order extreme value condition for some* $\gamma \in \mathbb{R}$ *and* $\rho \leq 0$. *There is a sequence of Brownian motions* $\{W_n(s)\}_{s>0}$ *such that for suitable* a_0, A_0 *and all* $\varepsilon > 0$

$$\sup_{k^{-1}\leq s\leq 1} s^{\gamma+1/2+\varepsilon} \left| \sqrt{k} \left(\frac{X_{(n-\lfloor ks \rfloor,n)} - U\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} - \frac{s^{-\gamma}-1}{\gamma} \right) - s^{-\gamma-1} W_n(s) - \sqrt{k} A_0\left(\frac{n}{k}\right) \Psi_{\gamma,\rho}\left(s^{-1}\right) \right| \to_P 0$$

with $k = k_n$ *such intermediate sequence that* $\sqrt{k}A_0(n/k)$ *is bounded.*

In term $\sqrt{k}A_0\left(\frac{n}{k}\right)\Psi_{\gamma,\rho}\left(s^{-1}\right)$, A_0 is a specific choice of the auxiliary function A appearing in Theorem 2.5 and $\Psi_{\gamma,\rho}$ is limit function H under this choice of A. Hence, this term is a deterministic estimate for the difference

$$\sqrt{k} \left(\frac{X_{(n-\lfloor ks \rfloor,n)} - U\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} - \frac{s^{-\gamma}-1}{\gamma} \right). \tag{2.8}$$

The term $s^{-\gamma-1}W_n(s)$ on the other hand tells us that the error between the difference (2.8) and the deterministic term $\sqrt{k}A_0\left(\frac{n}{k}\right)\Psi_{\gamma,\rho}\left(s^{-1}\right)$ is normally distributed simultaneously for all order statistics under consideration.

2.3 Asymptotic theory of certain univariate extreme value estimators

The consistency of the Hill estimator (see [29] and [10]) can be shown as an application of the Rényi representation theorem [43].

Theorem 2.8. *Let $X_1, X_2, \ldots$ be i.i.d. with a distribution function $F \in \mathcal{D}(G_\gamma)$ with $\gamma > 0$ and let γ_H be as in* (2.5)*. If $k_n \to \infty$, $k_n/n \to 0$ as $n \to \infty$,*

$$\gamma_H \to_P \gamma.$$

Asymptotic normality of the estimator can be shown as an application of Theorem 2.7.

Theorem 2.9. *Let $X_1, X_2, \ldots$ be i.i.d. with a distribution function $F \in \mathcal{D}(G_\gamma)$ with $\gamma > 0$. That satisfies the second order condition for $\gamma > 0$ and $\rho \leq 0$. Let γ_H be as in* (2.5) *and assume that the sequence $k_n \to \infty$, $k_n/n \to 0$ as $n \to \infty$, is such that*

$$\lim_{n\to\infty} \sqrt{k_n} A\left(\frac{n}{k_n}\right) = \lambda$$

with λ finite.

Under these conditions

$$\sqrt{k_n}\left(\gamma_H - \gamma\right) \to_D \mathcal{N}\left(\frac{\lambda}{1-\rho}, \gamma^2\right)$$

Theorems similar to Theorem 2.8 and Theorem 2.9 are valid for the Pickands estimator γ_P expressed in (2.6) and the Moment estimator γ_M expressed in (2.7) (see [10] for details).

3. On multivariate extreme value theory

3.1 Approaches to multivariate extremes

Due to the lack of an obvious order relation in the multivariate setting, it's not immediately clear how multivariate extreme value theory should be approached. The most traditional approach is to apply univariate extreme value theory the marginal distributions of multivariate observations $(X_{1,i}, X_{2,i}, \ldots, X_{d,i})$, $d \geq 2$, $1 \leq i \leq n$, and study limits similar to one that appears in Equation (2.1)

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{M_{1,n} - b_{1,n}}{a_{1,n}}, \ldots, \frac{M_{d,n} - b_{d,n}}{a_{d,n}}\right)$$

where $a_{j,n}$ are positive sequences for all $1 \leq j \leq d$ and $M_{j,n} = \max_{1\leq i\leq n} X_{j,i}$. A theory is then developed for approaching estimating probabilities of events of the form $\alpha_1 X_{1,n} + \cdots + \alpha_d X_{d,n}$ with $\alpha_i \geq 0$ for all i. See e.g. [10] and [21] for a thorough review of the topic. There are also several recent developments, whose relationship to the present work could be interesting subjects for future work, including multivariate Hill estimators defined as convex combinations of the componentwise Hill estimators (see e.g. [14] and [34]) and extreme risk region estimation [5].

A potential alternative approach is to assign to each multivariate observation a quantity that measures the outlyingness of that point with respect to the rest of the data. In [15] this was done in case of an elliptical distribution using the Mahalnobis distance to the center of the distribution. In Publication I and Publication II, this direction studied further. In Publication III an alternative quantity for measuring the outlyingness of the observations was developed and its properties were studied both

theoretically and empirically.

3.2 Multivariate extremes under elliptical distributions: The separating Hill estimator

A random variable $X : \Omega \to \mathbb{R}^d$ is said to be elliptically distributed if

$$X \stackrel{D}{=} \mu + \mathcal{R}\Lambda U, \tag{3.1}$$

where $\mu \in \mathbb{R}^d$, the random variables $\mathcal{R} : \Omega \to \mathbb{R}_+$ and U, which is uniformly distributed over the unit sphere $\mathbb{S}^{d-1}$, are independent, and Λ is such that $\Sigma = \Lambda^T \Lambda$ is symmetric, positive definite matrix of rank d. The positive random variable $\mathcal{R}$ is called *the generating variate* of the distribution.

Consider i.i.d. elliptically distributed random variables X_1 and X_2. For $\omega \in \Omega$, it's natural to write $X_1(\omega) \geq X_2(\omega)$, if $\mathcal{R}_1(\omega) \geq \mathcal{R}_2(\omega)$, where $\mathcal{R}_1$ and $\mathcal{R}_2$ are the generating variates of the distributions of X_1 and X_2 respectively. Hence, under ellipticity, it's natural to focus on the behaviour of the generating variate $\mathcal{R}$. To support the intuition, some theoretical justification can be derived from [31], where it was shown that the regular variation of the generating variate of the elliptical distribution is equivalent to the regular variation of the elliptically distributed random variable in the multivariate sense.

Mahalanobis distance is a straightforward way to extract the value of $\mathcal{R}$ when μ and Σ are known:

$$\begin{aligned}
& d_\Sigma(X, \mu) \\
= \; & \left\langle X - \mu, \Sigma^{-1} (X - \mu) \right\rangle^{1/2} \\
= \; & \left\langle \mu + \mathcal{R}\Lambda U - \mu, \left(\Lambda^T \Lambda\right)^{-1} (\mu + \mathcal{R}\Lambda U - \mu) \right\rangle^{1/2} \\
= \; & \left\langle \Lambda^{-1}\mathcal{R}\Lambda U, \Lambda^{-1}\mathcal{R}\Lambda U \right\rangle^{1/2} \\
= \; & \mathcal{R} \left(\|U\|^2 \right)^{1/2} \\
= \; & \mathcal{R}
\end{aligned}$$

Under known location μ and scatter Σ, the tail-index of $\mathcal{R}$ can then be easily estimated using the Mahalanobis distances $d_\Sigma(X, \mu)$ as input. The definition of the *separating Hill estimator* defined in [15] can be formulated this way: The separating Hill estimator is the univariate Hill estimator

(2.5) evaluated with respect to the observations $d_\Sigma(X, \mu)$. Under known location and scatter the asymptotic properties of the separating Hill estimator are inherited directly from those of the univariate Hill estimator.

In practice, however, the location μ and the scatter Σ of the distribution are not known. Hence, it's important to study the behaviour of the Hill estimator evaluated with respect to observations of the form $d_{\hat{\Sigma}}(X, \hat{\mu})$, where $\hat{\mu}$ is an estimate of the location μ and $\hat{\Sigma}$ is an estimate of the scatter Σ of the distribution of X. The asymptotic properties of the separating Hill estimator under estimated location and scatter are not obtained in as directly as in the case of known location and scatter.

In Publication II we show that the separating Hill estimator behaves well asymptotically also under estimated location and scatter, when the location and scatter estimators converge. More specifically: Let $X_1, X_2, \ldots$ be i.i.d. elliptically distributed random variables. Denote the location of the distribution by μ and the scatter of the distribution by Σ and let $\hat{\mu}_n$ and $\hat{\Sigma}_n$ be location and scatter estimators, respectively, evaluated with respect to the first n random variables $X_1, X_2, \ldots, X_n$. Let

$$R_i = \left\langle X_i - \mu, \Sigma^{-1}(X_i - \mu) \right\rangle^{1/2}, \tag{3.2}$$

$$E_i^{(n)} = \left\langle X_i - \hat{\mu}_n, \hat{\Sigma}_n^{-1}(X_i - \hat{\mu}_n) \right\rangle^{1/2}. \tag{3.3}$$

The corresponding order statistics are denoted by $R_{(1,n)} \geq R_{(2,n)} \geq \cdots \geq R_{(n,n)}$ and $E_{(1,n)} \geq E_{(2,n)} \geq \cdots \geq E_{(n,n)}$.

The following Lemma, derived in Publication II, essentially states that, for large n, the random variables $E_i^{(n)}$ are good estimates of the random variables R_i, in the sense that the relative error is small.

Lemma 3.1 (Publication II)**.** *Let X be as in* (3.1) *and let $X_1, X_2, \ldots$ be i.i.d. copies of X. Let R_i be as in* (3.2) *and $E_i^{(n)}$ be as in* (3.3)*. If $k_n \to \infty, k_n/n \to 0$ as $n \to \infty$, then, as $n \to \infty$, we have that for all $1 \leq i \leq k_n$,*

$$\left| R_{(i,n)}^2 - E_{(i,n)}^2 \right| \leq K_n R_{(i,n)}^2,$$

where K_n is a sequence of non-negative random variables. If $\hat{\mu}_n \to_P \mu$ and $\hat{\Sigma}_n \to_P \Sigma$, as $n \to \infty$, then $K_n \to_P 0$ as $n \to \infty$. If $\sqrt{n}\,(\hat{\mu}_n - \mu)$ and $\sqrt{n}\left(\hat{\Sigma}_n - \Sigma\right)$ converge in distribution, then $\sqrt{n}K_n$ is bounded in probability.

Using Lemma 3.1, it can be shown that the separating Hill estimator under estimated location and scatter is consistent and asymptotically normal (see Publication II). Only the latter result is displayed here.

Corollary 3.2 (Publication II)**.** *Let X be as in* (3.1) *and let the distribution of $\mathcal{R}$ have an extreme value index $\gamma > 0$. Assume that the distribution of $\mathcal{R}$ is such that its quantile function U, as in Equation* (2.2)*, is in $2ERV_{\gamma,\rho}$. Let $X_1, X_2, \ldots$ be i.i.d. copies of X. Let R_i be as in* (3.2) *and $E_i^{(n)}$ be as in* (3.3)*. Assume that the sequences $\sqrt{n}\left(\hat{\mu}_n - \mu\right)$ and $\sqrt{n}\left(\hat{\Sigma}_n - \Sigma\right)$ converge in distribution. Let A be the auxiliary function of U as in Definition 2.6. If k_n is an intermediate sequence, then as $n \to \infty$,*

$$\sqrt{k_n}\left(\frac{1}{k_n}\sum_{i=1}^{k_n}\log\left(\frac{E_{(i,n)}}{E_{(k_n+1,n)}}\right) - \gamma\right) \to_D \mathcal{N}\left(\frac{\lambda}{1-\rho}, \gamma^2\right),$$

if

$$\lambda = \lim_{n\to\infty}\sqrt{k_n}A\left(\frac{n}{k_n}\right)$$

exists.

3.3 Multivariate moment based extreme value index estimators

The Hill estimator is only valid for a heavy-tailed distribution, i.e. a distribution with a positive extreme value index. This also limits the applicability of the separating Hill estimator to elliptical distributions whose generating variate is heavy tailed. To overcome this limitation, extreme value index estimators such as the Pickands estimator (equation (2.6)) and the moment estimator (equation (2.7)) can be generalized analogously to the way the multivariate separating Hill estimator is defined.

In Publication I, this type of generalizations of the moment estimator and the mixed-moment estimator, introduced and discussed in depth in [23], were studied. By a similar argument as in Section 3.2, under known location and scatter, these estimators enjoy good asymptotic properties. However, under estimated location and scatter, the situation is more complicated. An extensive Monte Carlo study was conducted in Publication I to assess this situation. The simulation study provided additional evidence that the estimators enjoy desirable asymptotic properties also under estimated location and scatter.

3.4 Beyond ellipticity: Geometric outlier detection

A central idea in the separating Hill estimator is that extremity of an observation can be measured in a straightforward way in the case of an elliptically distributed random variable. In the general case, however, this is not a trivial task.

A quantity such as a statistical depth is a way to approach the issue (see [47] for a general notion of statistical depth). A well-known statistical depth is the half-space depth introduced in [45] is defined as follows:

$$HD(x, P) = \inf \{P(H) \mid H \text{ a closed halfspace}, x \in H\}$$

Notice that this quantity is not limited to any particular type of distribution. Other examples of depth functions include the simplicial depth proposed in [36], the majority depth discussed in [37], the zonoid depth introduced in [35] and the spatial depth introduced in [46].

In Publication III, we study a quantity called the Delaunay outlyingness that is based on the Delaunay triangulation of the sample (see [41] for an introduction to the Delaunay triangulation and related combinatorial topics). The Delaunay outlyingness assigns each observation a quantity whose inverse is analogous to the notion of depth. As Delaunay outlyingness is based on the geometry of the sample, the resulting outlyingness also reflects the spatial structure of the sample.

4. Summaries of the articles

Publication I Multivariate moment-based extreme value index estimators are introduced for elliptical distributions with a general extreme value index (as opposed to the multivariate separating Hill estimator that's only valid for elliptical distributions whose generating variate has a positive extreme value index). We show that their asymptotic properties are directly inherited to their univariate counterparts under known location and scatter and conduct a Monte Carlo study to assess their asymptotic properties under estimated location and scatter. Based on the simulations it seems that the estimators enjoy desirable asymptotic properties also under estimated location and scatter.

We also study empirical examples. The first one concerns financial data where the observations are the residuals of a AR(2)-GARCH(1,1) process fitted to the multivariate observations consisting of the stock prices of General Electric Company and Nokia Corporation from January 3rd 2007 to June 17th 2016. The residuals can be considered elliptically distributed and the multivariate moment-based extreme value index estimators proposed in the paper can be used to study the extreme value index of the distribution.

The second empirical example concerns temperature data collected from Barcelona, Spain from June 1st 2005 to June 1st 2015. Multivariate observations were formed by using the consecutive differences in the daily maximum temperatures as the first component and the daily minimum temperatures as the second component.

Publication II Multivariate separating Hill estimator introduced in [15] is studied further. In particular, it is proven that the estimator is con-

sistent and asymptotically normal under estimated location and scatter, a result of high practical importance as, in practice, the location and the scatter of the distribution are not known. The proof is based on an elementary argument which yields Lemma 3.1 and allows for control on the relative error between the true and estimated values of the generating variate.

A Monte Carlo study is also conducted to empirically assess the asymptotic behavior of the multivariate separating Hill estimator. Under different sample sizes, 2000 simulated values of the separating Hill estimator are collected under different distributions and the values of the median and 1st and 3rd quartile are plotted to illustrate their behavior as the sample size increases. As an empirical example, the separating Hill estimator is applied to several financial indexes. And the relationship between the resulting extreme value indices and the threshold value of the Hill estimator is studied in detail.

Publication III The notion of Delaunay outlyingness is introduced. Its theoretical properties are studied in the case of a convex compact region and a fixed, finite number of outliers. More specifically, a finite number of observations are scattered outside a compact convex region K and an increasing number of observations distributed over K is collected. As the sample size increases, the Delaunay outlyingness of the observations in K decreases, while that of the observations scattered outside of K remains at least δ for some $\delta > 0$.

Delaunay outlyingness is also studied through simulations, which suggest that the method also correctly recognizes outliers when K is not convex. Finally, an empirical example is given to illustrate the manner in which the outlyingness measure behaves with respect to the structure of the point cloud formed by the observations.

References

[1] Karin Aarssen and Laurens de Haan. On the maximal life span of humans. *Mathematical Population Studies*, 4(4):259–281, 1994. PMID: 12318734.

[2] Ying An and M.D. Pandey. A comparison of methods of extreme wind speed estimation. *Journal of Wind Engineering and Industrial Aerodynamics*, 93(7):535 – 545, 2005.

[3] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1987.

[4] Paul W. Burton. Seismic risk in southern europe through to india examined using gumbel's third distribution of extreme values. *Geophysical Journal of the Royal Astronomical Society*, 59(2):249–280.

[5] Juan-Juan Cai, John H. J. Einmahl, and Laurens de Haan. Estimation of extreme risk regions under multivariate regular variation. *Ann. Statist.*, 39(3):1803–1826, 06 2011.

[6] Yumin Chen, Duoqian Miao, and Hongyun Zhang. Neighborhood outlier detection. *Expert Systems with Applications*, 37(12):8745–8749, 2010.

[7] Nicholas J. Cook, R. Ian Harris, and Richard Whiting. Extreme wind speeds in mixed climates revisited. *Journal of Wind Engineering and Industrial Aerodynamics*, 91(3):403 – 422, 2003.

[8] N.J. Cook. Towards better estimation of extreme winds. *Journal of Wind Engineering and Industrial Aerodynamics*, 9(3):295 – 323, 1982.

[9] Laurens de Haan. *Slow Variation and Characterization of Domains of Attraction*, pages 31–48. Springer Netherlands, Dordrecht, 1984.

[10] Laurens de Haan and Ana Ferreira. *Extreme value theory: an introduction*. Springer-Verlag New York, 2006.

[11] Laurens de Haan and Sidney Resnick. Second-order regular variation and rates of convergence in extreme-value theory. *The Annals of Probability*, 24(1):97–124, 1996.

[12] Laurens de Haan and Ulrich Stadtmüller. Generalized regular variation of second order. *Journal of the Australian Mathematical Society. Series A. Pure Mathematics and Statistics*, 61(3):381–395, 1996.

[13] A. L. M. Dekkers, J. H. J. Einmahl, and L. De Haan. A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, 17(4):1833–1855, 12 1989.

[14] A. Dematteo and S. Clémençon. On tail index estimation based on multivariate data. *Journal of Nonparametric Statistics*, 28(1):152–176, 2016.

[15] Yves Dominicy, Pauliina Ilmonen, and David Veredas. Multivariate Hill estimators. *International Statistical Review*, 85(1):108–142.

[16] Holger Drees. On smooth statistical tail functionals. *Scandinavian Journal of Statistics*, 25(1):187–210, 1998.

[17] Jesson Einmahl, John Einmahl, and L.F.M. de Haan. Limits to human life span through extreme value theory. Discussion Paper 2017-051, Tilburg University, Center for Economic Research, 2017.

[18] John H. J. Einmahl and Jan R. Magnus. Records in athletics through extreme-value theory. *Journal of the American Statistical Association*, 103(484):1382–1391, 2008.

[19] John H. J. Einmahl and Sander G. W. R. Smeets. Ultimate 100-m world records through extreme-value theory. *Statistica Neerlandica*, 65(1):32–42.

[20] Michael Falk. Some best parameter estimates for distributions with finite endpoint. *Statistics: A Journal of Theoretical and Applied Statistics*, 27:115–125, 01 1995.

[21] Michael Falk. It was 30 years ago today when laurens de haan went the multivariate way. *Extremes*, 11(1):55–80, Mar 2008.

[22] R. A. Fisher and L. H. C. Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24:180, 1928.

[23] M. Isabel Fraga Alves, M. Ivette Gomes, Laurens de Haan, and Cláudia Neves. Mixed moment estimator and location invariant alternatives. *Extremes*, 12(2):149–185, Jun 2009.

[24] Maurice (1878-1973) Frechét. Sur la loi de probabilité de l'écart maximum, 1928.

[25] J. Galambos and E. Seneta. Regularly varying sequences. *Proceedings of the American Mathematical Society*, 41(1):110–116, 1973.

[26] Samuel Gbari, Michel Poulain, Luc Dal, and Michel Denuit. Extreme value analysis of mortality at the oldest ages: A case study based on individual ages at death. *North American Actuarial Journal*, 21(3):397–416, 2017.

[27] B. Gnedenko. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics*, 44(3):423–453, 1943.

[28] R.I. Harris. Gumbel re-visited - a new look at extreme value statistics applied to wind speeds. *Journal of Wind Engineering and Industrial Aerodynamics*, 59(1):1 – 22, 1996.

[29] Bruce M. Hill. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174, 09 1975.

[30] J. R. M. Hosking and J. R. Wallis. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29(3):339–349, 1987.

[31] Henrik Hult and Filip Lindskog. Multivariate extremes, aggregation and dependence in elliptical distributions. *Advances in Applied Probability*, 34(3):587–608, 2002.

[32] James Pickands III. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1):119–131, 01 1975.

[33] ASA Kathryn A. Watts MSc, Debbie J. Dupuis PhD, and FCIA Bruce L. Jones PhD, FSA. An extreme value analysis of advanced age mortality data. *North American Actuarial Journal*, 10(4):162–178, 2006.

[34] Moosup Kim and Sangyeol Lee. Estimation of the tail exponent of multivariate regular variation. *Annals of the Institute of Statistical Mathematics*, 69(5):945–968, Oct 2017.

[35] Gleb Koshevoy and Karl Mosler. Zonoid trimming for multivariate distributions. *Ann. Statist.*, 25(5):1998–2017, 10 1997.

[36] Regina Y. Liu. On a notion of data depth based on random simplices. *Ann. Statist.*, 18(1):405–414, 03 1990.

[37] Regina Y. Liu, Jesse M. Parelius, and Kesar Singh. Multivariate analysis by data depth: descriptive statistics, graphics and inference, (with discussion and a rejoinder by liu and singh). *Ann. Statist.*, 27(3):783–858, 06 1999.

[38] C Kostas Makropoulos and W Paul Burton. Hazan: A fortran program to evaluate seismic-hazard parameters using gumbel's theory of extreme value statistics. *Comput. Geosci.*, 12(1):29–46, April 1984.

[39] R von Mises. La distribution de la plus grande de n valeurs. *Rev. math. Union interbalcanique*, 1:141–160, 1936.

[40] Liang Peng and Yongcheng Qi. Maximum likelihood estimation of extreme value index for irregular cases. *Journal of Statistical Planning and Inference*, 139(9):3361 – 3376, 2009.

[41] Franco P. Preparata and Michael I. Shamos. *Computational Geometry: An Introduction*. Springer-Verlag, Berlin, Heidelberg, 1985.

[42] Michael E. Robinson and Jonathan A. Tawn. Statistics for exceptional athletics records. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 44(4):499–511, 1995.

[43] Alfréd Rényi. On the theory of order statistics. *Acta Mathematica Hungarica*, 4(3-4):191–231, 1953.

[44] Richard L. Smith. Forecasting records by maximum likelihood. *Journal of the American Statistical Association*, 83(402):331–338, 1988.

[45] John W. Tukey. Mathematics and the Picturing of Data. In Ralph D. James, editor, *International Congress of Mathematicians 1974*, volume 2, pages 523–532, 1974.

[46] Yehuda Vardi and Cun-Hui Zhang. The multivariate l1-median and associated data depth. *Proceedings of the National Academy of Sciences*, 97(4):1423–1426, 2000.

[47] Yijun Zuo and Robert Serfling. General notions of statistical depth function. *Ann. Statist.*, 28(2):461–482, 04 2000.

Aalto-DD 49/2019

ISBN 978-952-60-8465-7 (printed)
ISBN 978-952-60-8466-4 (pdf)
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University
School of Science
Department of Mathematics and Systems Analysis
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

**DOCTORAL
DISSERTATIONS**