THE 34TH ANNUAL INTERNATIONAL CONFERENCE

ON STATISTICS AND MODELLING IN HUMAN

AND SOCIAL SCIENCES

AGENDA

SEPTEMBER 12-13, 2023
CAIRO, EGYPT

FACULTY OF ECONOMICS AND POLITICAL SCIENCE,

CAIRO UNIVERSITY

Conference Organization Committee

General Supervisor

Prof. Hanan Aly, *Acting Dean of Faculty of Economics and Political Science*

Executive Secretariat

Prof. Abdel Nasser Saad, C*hairman of Statistics Department*

Committee Members

Prof. Hussein A. Sayed	Prof. Tarek Emera
Prof. Zeinab Selim	Prof. Abdel Hamed Negm
Prof. Maged Osman	Prof. Sanaa Ismail
Dr. Mahmoud Rashwan	Dr. Nesma Saleh
Dr. Aya Anas	Dr. Rana Nabil
Dr. Heba El Deeb	Dr. Rania Mamdouh
Ms. Reham El-Shaer	Ms. Allaa Hatem
Mr. Moamen Ahmed	Ms. Dina Ashraf

Reviewers (In Alphabetical Order)

Prof. Abdel Hameed Negm	Prof. Abdel Nasser Saad
Prof. Hanan Aly	Prof. Laila El-Zeini
Dr. Mahmoud Rashwan	Dr. Nesma Saleh
Dr. Ola El Sayed	Prof. Ramadan Hamid
Dr. Rania Mamdouh	Dr. Sara Osama
Dr. Wafaa Ibrahim	Dr. Zakaria AbdelSamee

Conference Agenda

Day 1: Tuesday, September 12th

09:00-09:30   Registration

09:30-10:30   Opening Session

Prof. Abdel Nasser Saad, *Chairman of the Statistics Department*

Prof. Hanan Aly, *Acting Dean of Faculty of Economics and Political Science*

Prof. Mahmoud El Said, *Vice President of Cairo University*

Prof. Mohamed Osman El Khosht, *President of Cairo University*

Major General Khairat Barakat, *Chairman of the Central Agency for Public Mobilization and Statistics (CAPMAS)*

10:45-12:30   Seminar (1)

Applied Sampling Methods in Surveys Conducted by the ERF

Chairman: Prof. Hussein A. Sayed (Faculty of Economics and Political Science, Cairo University)

Discussant: Prof. Ramadan Hamid (*Faculty of Economics and Political Science, Cairo University*)

Speakers: Dr. Mai Mokhtar (Faculty of Economics and Political Science, Cairo University)

Co-Presenter: Dr. Nesma Amer (Senior Statistician at the ERF)

12:30-13:00   Coffe Break

13:00-14:30   Session (1): Theoretical Statistics

Chairman: Prof. Abdel Hamed Negm (*Faculty of Economics and Political Science, Cairo University*)

Moderator: Dr. Rana Nabil (*Faculty of Economics and Political Science, Cairo University*)

1- **"Chen-Burr XII Distribution: Properties and Estimation with Applications "**
H. N. Salem, G. R. AL-Dayian and A. A. EL-Helbawy
(*Statistics Department, Faculty of Commerce, AL-Azhar University (Girls' Branch), Cairo, Egypt*)

2- **"Discrete Exponentiated Generalized Family of Distributions "**
Abd EL-Hady, A.E., Hegazy, M. A., EL-Helbawy and G. R. AL-Dayian (*Department of Statistics, Faculty of Commerce, Al-Azhar University (Girls' Branch), Cairo, Egypt*)

3- **" Mixture of Two components of Exponentiated Family"**
F. G., Abd EL-Maksoud, G. R. AL-Dayian (*Department of Statistics, Faculty of Commerce, AL-Azhar University (Girls' Branch), Tafahna Al-Ashraf, Egypt*) and A. A. EL-Helbawy (*Department of Statistics, Faculty of Commerce, AL-Azhar University (Girls' Branch), Cairo, Egypt*)

4- **"Non-Bayesian and Bayesian Prediction for Additive Flexible Weibull Extension-Lomax Distribution"**
H. N. Salem, G. R. AL-Dayian, A. A. EL-Helbawy, and R. E. Abd EL-Kader (*Statistics Department, Faculty of Commerce, Al-Azhar University (Girls' Branch), Cairo, Egypt*)

14:30-16:00   Session (2): Applied Statistics (1)

Chairman: Prof. Abdallah Abdel-Ghaly (*Faculty of Economics and Political Science, Cairo University*)

Moderator: Dr. Nora Said (*Faculty of Economics and Political Science, Cairo University*)

1- **"Geographically Weighted Zero-Inflated Negative Binomial Regression: An Application to COVID-19 Deaths across the United States"**
Amira Elayouty, Doaa Sakr, Abdelnasser Saad and Mohamed Ismail *(Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt)*

2- **"A spatial statistical downscaling model for combining spatially misaligned maximum temperature data using R-INLA"**
Sylvia A. Shawky(*faculty of Commerce and Business Administration - Future University in Egypt*), Abdelnasser saad, and Amira Elayouty *(Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt)*

3- **" Statistical Data Integration Using Machine Learning Models: Insights from Egyptian Surveys"**
Mohamed Ali Ismail, Mai Sherif Hafez *(Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt)*, and Israa Lewaa (*Faculty of Business Administration, Economics and Political Science, Department of Business Administration, The British University in Egypt*)

16:00-17:00   Lunch

Day 2: Wednesday, September 13th

09:30-11:00   **Session (3): Statistical Modelling and Sampling**

Chairman: *Prof. Ramadan Hamid* (*Faculty of Economics and Political Science, Cairo University*)

Moderator: Dr. Sara Osama (*Faculty of Economics and Political Science, Cairo University*)

1- **"Optimal Stochastic Allocation in Multivariate Stratified Sampling"**
Maha, I., Mahfouz (*National Centre for Social and Criminological Research, Egypt),* Mahmoud, M., Rashwan, and Zeinab, A., Khadr *(Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt)*

2- **" A note on using Kruskal-Wallis test in linear mixed-effects models with heavily skewed error components"**
Yahia S. El-Horbaty (*Department of Mathematics, Insurance, and Applied Statistics, Helwan University, Egypt*)

3- **" A Frailty Model of Generalized Lindley Distribution with the Gompertz Baseline: Estimation and Application"**
Asmaa H.Abd El-Maqsoud, Asmaa S.A.Zeidan, and Salwa A.Mousa
(*Statistics Department, Faculty of Commerce Al-Azhar University (Girls' Branch), Cairo, Egypt*)

 11:00-11:30   **Coffe Break**

11:30-13:00   Session (4): Applied Statistics (2)

Chairman: Prof. Moshira Ismail (*Faculty of Economics and Political Science, Cairo University*)

Moderator: Dr. Maram Salem (*Faculty of Economics and Political Science, Cairo University*)

1- **"Privacy Attitudes in Smart City Context: Political Economy Evidence from a Global Survey Data"**
Abdel Hameed Nawar *(Department of Economics, Faculty of Economics and Political Science, Cairo University, Egypt)* and Siham Rabie *(Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt)*

2- **" Climate change and child malnutrition in Egypt"**
Amira Elayouty *(Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt)*, Hala Abou-Ali *(Department of Economics, Faculty of Economics and Political Science, Cairo University and Institute of National Planning, Egypt)* , Ronia Hawash Rabie (*Lacy School of Business, Butler University, Indianapolis, Indiana, USA)*

3- **"Extreme value statistics in semi-supervised models"**
Hanan Ahmed, John H.J. Einmahl (*Tilburg University, Netherlands*), and Chen Zhou (*Erasmus University, Netherlands*)

13:00-14:30   Session (5): Young Researchers

Chairman: Prof. Fatma El Zanaty (*Faculty of Economics and Political Science, Cairo University)*

1- "What does – being a woman- mean in labor force? "
Rawda Ahmed, Asmaa Khamis , Nouran Mahmoud, and Hany Amr *(Statistics Department, Faculty of Economics and Political Science, Cairo University)*

2- "Quality of academic life for disabled students in Cairo University"

Mayar Mohsen, Roden Raouf, and Mira Taha *(Statistics Department, Faculty of Economics and Political Science, Cairo University)*

3- "Indicators Influencing Poverty in Egypt"

Habiba   Gamal,   Hania   Mohamed   and   Mariam Mahmoud*(Statistics Department, Faculty of Economics and Political Science, Cairo University)*

Abstracts for previous conference (2022) are available online at

http://feps.edu.eg/enn/departments/st/conference.php

CALL FOR PAPERS
Cairo University
Faculty of Economics and Political Science
Department of Statistics

THE 35th ANNUAL INTERNATIONAL CONFERENCE ON STATISTICS AND MODELLING IN HUMAN AND SOCIAL SCIENCES

September 10-11, 2024

Important Dates

Deadline for receiving abstracts: August 1st ,2024

Deadline for receiving papers: August 15th 2024

The conference announces the reward offered by Prof. Nadia Makary that is granted for any research published in the conference proceedings and accepted later for publication in one of the international recognized periodicals.

Correspondence

Department of Statistics, Faculty of Economics and Political Science, Cairo University

Tel: 35728055 – 37744990   Fax: 35728055 - 37744990

e-mail: ICSM@feps.edu.eg

Website:
http://www.feps.edu.eg/en/departments/statistics/conference/index.ht

Chen-Burr XII Distribution:
Properties and Estimation with Applications

H. N. Salem, G. R. AL-Dayian and A. A. EL-Helbawy
Statistics Department, Faculty of Commerce,
AL-Azhar University (Girls' Branch), Cairo, Egypt

Abstract

In this paper Chen-Burr XII distribution is constructed and graphical description of the probability density function, hazard rate and reversed hazard rate functions of the proposed model is obtained. Also, some statistical properties of Chen-Burr XII distribution are discussed and some new models as special cases from Chen-Burr XII distribution are introduced. Moreover, maximum likelihood estimation of the parameters, reliability, hazard rate and reversed hazard rate functions of Chen-Burr XII distribution is considered. Finally, three applications are considered to demonstrate the applicability of the proposed model and its superiority over some existing distributions.

Keywords: Competing risks, additive model, Chen-Burr XII distribution, maximum likelihood estimation, hazard shape.

Discrete Exponentiated Generalized Family of Distributions

Abd EL-Hady, A.E., Hegazy, M. A., EL-Helbawy and G. R. AL-Dayian

Department of Statistics, Faculty of Commerce,
Al-Azhar University (Girls' Branch), Cairo, Egypt

Abstract

In this paper, exponentiated-G family and the inverted Topp–Leone distribution are used to introduce a new discrete distribution called discrete exponentiated inverted Topp–Leone distribution. This new model possesses only two parameters. The method of maximum likelihood is used under Type-II censored sample for estimating the unknown parameters, survival, hazard rate and alternative hazard rate functions. Also, confidence intervals for the parameters are obtained. A simulation study is performed to investigate the effectiveness of the proposed distribution. Finally, two real data sets are analyzed to illustrate its flexibility for real-life applications.

Keywords: *Survival discretization, Exponentiated-G family, Order statistic, Maximum likelihood.*

Mixture of Two components of Exponentiated Family

F. G., Abd EL-Maksoud[1], G. R. AL-Dayian[2] and A. A. EL-Helbawy[2]

[1]*Department of Statistics, Faculty of Commerce, AL-Azhar University (Girls' Branch), Tafahna Al-Ashraf, Egypt*
[2]*Department of Statistics, Faculty of Commerce, AL-Azhar University (Girls' Branch), Cairo, Egypt*

Abstract

In this paper, the mixture of two components of exponentiated family is introduced as a new family of continuous distributions. Some general properties of the proposed family are discussed, and the maximum likelihood estimation method is used to estimate the unknown parameters. The mixture of two exponentiated inverted Kumaraswamy distribution is studied as a sub-model from the mixture exponentiated family. Some statistical properties are obtained. Also, the maximum likelihood estimators for the unknown parameters of the mixture exponentiated inverted Kumaraswamy distribution are derived. A simulation study is conducted to assess the performance of the maximum likelihood estimators of the parameters of the mixture exponentiated inverted Kumaraswamy distribution. Finally, a real data set is applied to ensure the simulated results.

Keywords: Mixture distribution, Exponentiated family, Inverted Kumaraswamy distribution, Identifiability property, maximum likelihood estimation.

Non-Bayesian and Bayesian Prediction
for Additive Flexible Weibull Extension-Lomax Distribution

H. N. Salem, G. R. AL-Dayian, A. A. EL-Helbawy, and R. E. Abd EL-Kader

Statistics Department, Faculty of Commerce,
Al-Azhar University (Girls' Branch), Cairo, Egypt

Abstract

Prediction of future observations is an important problem in many practical applications. This paper focuses on considering one-sample and two-sample prediction (as a special case of the multi-sample prediction) for a future observation from additive flexible Weibull extension-Lomax distribution. Non-Bayesian and Bayesian prediction based on Type II censoring scheme are studied. The conditional prediction approach is discussed as a non-Bayesian prediction method. Also, Bayesian prediction is obtained under two different loss functions, the squared error and linear-exponential loss functions. Moreover, a simulation study is conducted to evaluate the performance of the derived predictors and three applications of COVID -19 data in some countries are considered.

Geographically Weighted Zero-Inflated Negative Binomial Regression: An Application to COVID-19 Deaths across the United States

Amira Elayouty,[1] Doaa Sakr,[1,*] Abdelnasser Saad[1] and Mohamed Ismail[1]
[1] Department of Statistics, Faculty of Economic and Political Science, Cairo University, Giza, Egypt.
*Corresponding author. doaasakr@feps.edu.eg

Abstract

In the early stages of any pandemic, the number of cases/deaths related to the contagious infectious disease in a certain area is often skewed, over dispersed and have excessive zeros, as some areas may have no contagions until sometime. For these reasons, this paper proposes a geographically weighted zero inflated negative binomial regression (GWZINBR) that extends the Zero inflated negative binomial (ZINBR) model to the context of Geographically Weighted Regression (GWR). GWR models proved more useful at modelling data recorded across space, where response-predictor relationships are likely to be not stationary across the study region. In this paper, we aim at modelling the number of COVID-19 deaths across the United States counties and the spatial variability in the impacts of the potential socio-economic and environmental risk factors using geographically weighted regression models. The proposed GWZINBR should enable us to analyze the spatial variability in the COVID-19 deaths across the United States and how it relates to a wide range of socio-economic and environmental factors while relaxing the assumptions of stationary relationships and accounting for the specific characteristics of the data such as overdispersion and excess zero counts. The performance of the proposed GWZINBR model is compared to the global negative binomial and ZINBR models. The results indicated that the proposed model has the smallest AIC among all fitted models. The GWZINBR represents an important advancement in the modelling geo-referenced data when the response variable is skewed, over dispersed and has an excessive number of zeros.

Key Words: *Count Data, COVID-19, Spatial regression, Zero Inflated Models, Geographically Weighted Zero Inflated Negative Binomial Regression.*

A spatial statistical downscaling model for combining spatially misaligned maximum temperature data using R-INLA

Sylvia A. Shawky[1a], Abdelnasser saad[b], Amira Elayouty[b]

[a]*faculty of Commerce and Business Administration - Future University in Egypt,*
[b]*dept. of Statistics, Faculty of Economics and political science, Cairo University,*

Abstract

Climate data are essential for analysing and modelling climate variability and trends and their impacts on different health and socio-economic activities. Though Africa is one of the most vulnerable regions to climate change, studying climate change and variability is very scarce in Africa due to the limited availability and access to climate data. The weather stations are often sparse and unevenly distributed across many parts of Africa and suffer from large proportions of missingness over space and time. Therefore, the in-situ geostatistical climate data measured directly from monitoring weather stations are assumed to be accurate within its measurement error. Alternatively, physical climate model output provides another source of climate gridded data that cover large and dense spatial and temporal domains at a certain resolution but not at smaller scales. Physical climate model data do not account for the uncertainty in the data and hence tend to exhibit bias compared to the in-situ observations. To enhance the accuracy of climate model outputs and the spatial coverage of in-situ data, the simulated climate model output data can be calibrated against the real measurements from monitoring weather stations. However, integrating such climate model output data with in-situ observed data for improved interpolation is not trivial as it involves misalignment in space and time which can lead to biases in predictions. In this study, our aim is to present a statistical downscaling framework for combining the daily maximum temperature observations from 52 monitoring weather stations across the Nile Basin countries in 2022 with the 44 x 44 km gridded data simulated from a regional climate model (RCM) across the same study region. To accurately account for uncertainty from the different data sources and propagate it to predictions, a spatially varying coefficient (SVC) coregionalization model that assumes a joint distribution between the covariate (simulated model output) and the response (in-situ observations) is employed. The proposed model is fitted under a hierarchical Bayesian framework using integrated nested Laplace approximation (INLA) coupled with stochastic partial differential equation (SPDE) approach. This spatial model assumes that the true underlying process for both sources of data is a Gaussian process. This allows us to flexibly adjust the spatial latent field of the simulated climate model data and improve the prediction of the daily maximum temperature data along the entire spatial domain at finer resolution.

Keywords: Africa, Coregionalization model, Gaussian process, INLA, Spatially varying coefficient, Statistical downscaling.

---

[1]Corresponding author, *email: silvia.shawky2012@feps.edu.eg*

Preprint submitted to the 34th Annual International Conference on Statistics and Modelling in Human and Social Science   *August 28, 2023*

14

Statistical Data Integration Using Machine Learning Models: Insights from Egyptian Surveys

Mohamed Ali Ismail[1], Mai Sherif Hafez[2], Israa Lewaaelhamd[3,*]

[1]Professor of Statistics, Faculty of Economics and Political Science, Department of Statistics, Cairo University (mismail@feps.edu.eg)

[2] Assistant Professor of Statistics, Faculty of Economics and Political Science, Department of Statistics, Cairo University (mai.sherif@feps.edu.eg)

[3]Assisstant Professor of Statistics, Faculty of Business Administration, Economics and Political Science, Department of Business Administration, The British University in Egypt

[*]Corresponding author: Tel.: +201029499446; E-mail: israalewaa@feps.edu.eg or israa.lewaa@bue.edu.eg

Abstract

This study presents an innovative approach to the amalgamation of social survey data by means of a combination of machine learning techniques (ML) and compares them with the existing methods of statistical matching (SM). The objective of this integration is to investigate the connection between variables that are not simultaneously found in a single survey. The outcomes of the novel matching procedure appear to be highly encouraging as they surpass the effectiveness of conventional matching methods.

Key Words: Machine Learning Models; Statistical Data Integration; Real dataset

Optimal Stochastic Allocation in Multivariate Stratified Sampling

Maha, I., Mahfouz[a*], Mahmoud, M., Rashwan[b], Zeinab, A., Khadr[c]

a National Centre for Social and Criminological Research, Egypt, e-mail: mimismail@msa.edu.eg

b Department of Statistics, Faculty of Economics and Political Sciences, Cairo University, Egypt

c Department of Statistics, Faculty of Economics and Political Sciences, Cairo University, Egypt

Abstract

In this paper, a multivariate optimal stochastic compromise allocation is proposed using multi-objective mathematical programming model that minimizes the total survey cost as well as the individual variances of the overall stratified mean of each of the characteristics of interest, simultaneously. The proportional increase in the variance of the estimator under optimum variance of the estimators from the variance of the estimators under optimum cost is set as a constraint and is upper-bounded by a pre-determined quantity. The Weighted Goal Programming is adopted as a solution technique. However, in practice, because the population data are usually unknown and are estimated under uncertainty, thus, their sample estimates should be better treated as random variables. Therefore, the allocation model is stated as a stochastic mathematical programming problem, that is to be solved using the Chance-Constrained Programming technique. Simulation-based comparative study is conducted to assess the performance of the proposed allocation versus other optimal allocation techniques chosen from the literature. Based on the criteria used for comparison, it is shown that the suggested models have the highest efficiency in obtaining the estimators with the highest precision.

A note on using Kruskal-Wallis test in linear mixed-effects models with heavily skewed error components

Yahia S. El-Horbaty[1]

Department of Mathematics, Insurance, and Applied Statistics, Helwan University, Egypt

Abstract

Testing the need for random effects is a common practice in linear models for one-way independent group designs. Though many existing tests can be used to achieve this task, we shed light on the outstanding performance of the Kruskal-Wallis (KW) test compared to the commonly used likelihood ratio (LR) and F tests when the distributions of the error components are heavily skewed. A permutation-based test procedure using the KW statistic is proposed where the necessary exchangeability condition for valid permutation of the residual ranks is verified. Using simulation experiments, the new permutation-based test possesses correct empirical Type-I error rates as well as a dominant power.

Keywords. Exchangeability, Residuals Ranks, ANOVA, Robust Tests, Non-normality.

---

[1] Correspondence: yahia_mohamed@commerce.helwan.edu.eg

A Frailty Model of Generalized Lindley Distribution with the Gompertz Baseline: Estimation and Application

Asmaa H. Abd El − Maqsoud*, Asmaa S. A. Zeidan, and Salwa A. Mousa

Statistics Department, Faculty of Commerce

Al-Azhar University (Girls' Branch), Cairo, Egypt.

*Corresponding author: Asmaa.Hamed@azhar.edu.eg

Abstract

The hazard function plays a very important role in survival analysis. According to the homogeneous assumption, the hazard function distribution is the same for each individual in the study. However, this is practically impossible to be satisfied. This means that there is unobserved heterogeneity; the values of the covariates for everyone may have different distributions. The frailty model considers the unobserved heterogeneity by introducing a random effect in the model where it has a multiplicative effect on the baseline hazard function. The main goal of this paper is to suggest the generalized Lindley frailty model with Gomprtez baseline hazard function and assume the model in a univariate case using the ML method to estimate the parameters. In addition to the modernity of the data used to evaluate the performance of the suggested model. Several properties of the model are discussed such Laplace transform function which is useful to obtain marginal survival and marginal hazard functions. The maximum likelihood estimation method is used in estimating the parameters of the proposed model. Simulation studies are conducted to evaluate the performance of the estimates of parameters. Finally, an application of the proposed model is illustrated to a real lifetime dataset to analyze times of the effect of epinephrine doses on the prevention of spinal hypotension during Cesarean. In addition, a comparison study for the same data is presented by using model selection criteria and the appropriate model has been suggested.

Keywords: Generalized Lindley distribution; Frailty models; Maximum likelihood method; Unobserved heterogeneity; Survival analysis.

Privacy Attitudes in Smart City Context: Political Economy Evidence from a Global Survey Data

Dr. Abdel Hameed Nawar   Dr. Siham Rabie

Abstract

The application of smart technology and Artificial intelligence in many facets of life have undergone a technological revolution in recent years. With the global urban megatrend to transform cities to smart ecosystems, numerous smart tools and innovative ways have emerged and deployed to manage assets, resources and services efficiently and to improve the quality of life. Today, although the majority of people in many cities worldwide have *access* to these smart applications, there are still differences in the proportion of people who *use* them.

Using a unique dataset, this study offers a new perspective to investigate the impact on privacy attitudes of a host of individual and political economy factors ecosystems including human development levels, rates of non- cash transactions, e-trust in authorities due to online dissemination of public information in general and the existence of AI policy are investigated.

Evidence from the study findings shows that there are obvious differences of people responses towards the privacy aspects according to different levels of human development and prevalence of cashless economy among the world's cities, while e-trust in authorities and the existence of a national AI policy do not have sufficient significant impact. A concerted approach is needed to deal with the global challenge of building more trust in smart technologies, especially in the emerging era of "generative AI."

Keywords: Human Development Index, Smart City Index, Artificial Intelligence (AI) Applications, economics of privacy, Political Economy

Climate change and child malnutrition in Egypt

Amira Elayouty[1], Hala Abou-Ali[2], Ronia Hawash[3]

Abstract

Children's nutritional status is expected to be negatively impacted by global climate change given their relative vulnerability to food insecurity shocks. This study investigates the impact of climate change on the geographical variation of the prevalence of stunting and wasting among children under the age of five in Egypt using the latest available Demographic and Health Survey. Data are spatially and temporally merged with high resolution climate change datasets to investigate whether and how the change in temperatures and precipitation has an influence on children's malnutrition. The prevalence of stunting and wasting among children under five years of age and its socioeconomic determinants are modelled using a Bayesian geospatial regression model at the individual level and a multi-scale geographic weighted regression at the kism/shyakha level. The prevalence of stunting and wasting as well as the association between them and socio-economic and climatic factors varied from one area to another across Egypt. The result of this paper highlights the fact that social policies and public health interventions targeted to reduce the burden of childhood stunting should consider geographical heterogeneity and adaptable risk factors.

Keywords: Egypt, Climate Change, Bayesian geo-spatial model, Multiscale Geographically Weighted Regression, Stunting, Wasting

---

[1] Assistant Professor, Department of Statistics, FEPS, Cairo University, Egypt. a.ayouti@feps.edu.eg

[2] Professor, Department of Economics, FEPS, Cairo University, Egypt and Institute of National Planning, Cairo, Egypt.

[3] Assistant Professor, Lacy School of Business, Butler University, Indianapolis, Indiana, USA. rhawash@butler.edu.

Extreme value statistics in semi-supervised models

Hanan Ahmed
Tilburg University

John H.J. Einmahl
Tilburg University

Chen Zhou
Erasmus University

August 29, 2023

Abstract. We consider extreme value analysis in a semi-supervised setting, where we observe, next to the n data on the target variable, $n + m$ data on one or more covariates. This is called the semi-supervised model with n labeled and m unlabeled data. By exploiting the tail dependence between the target variable and the covariates, we derive an estimator for the extreme value index of the target variable in this setting and establish its asymptotic behavior. Our estimator substantially improves the univariate estimator, based on only the n target variable data, in terms of asymptotic variance whereas the asymptotic bias remains unchanged. We present a simulation study in which the asymptotic results are confirmed and also an extreme quantile estimator is derived and its improved performance is shown. Finally the estimation method is applied to rainfall data in France.

Key words. Asymptotic normality, extreme value index, semi-supervised inference, tail dependence, variance reduction.