DALARNA UNIVERSITY

# Degree Thesis

## Master's level (Second cycle)

# Possible Difficulties in Evaluating University Performance Based on Publications Due to Power Law Distributions: Evidence from Sweden

Authors: Haroon Sadric and Sarah Zia
School: Dalarna University
Supervisors: Arend Hintze
Examiner: Mia Xiaoyun Zhao
Subject/main field of study: Microdata Analysis
Course code: MI4002
Higher education credits: 15 ECTS
Date of examination: 2023.04.03

At Dalarna University it is possible to publish the student thesis in full text in DiVA. The publishing is open access, which means the work will be freely accessible to read and download on the internet. This will significantly increase the dissemination and visibility of the student thesis.

Open access is becoming the standard route for spreading scientific and academic information on the internet. Dalarna University recommends that both researchers, as well as students. publish their work open access.

I give my/we give our consent for full-text publishing (freely accessible on the internet, open access):

Yes ☒ No ☐

Dalarna University – SE-791 88 Falun – Phone +4623-77 80 00

# List of Contents

## List of Tables:



## List of Figures:

## Abstract:

Measuring the research performance of a university is important to the universities themselves, governments, and students alike. Among other metrics, the number of publications is easy to obtain, and due to the large number of publications each university produces during one year, it suggests to be one accurate metric. However, the number of publications depends largely on the size of the institution, suggesting, if not addressed, that larger universities are better. Thus, one might intuitively try to normalize by size and use publications per researcher instead. A better institution would allow individual researchers to have more publications each year. However, publications, like many other things, might follow a power-law distribution, where most researchers have few, and only a few researchers have very many publications. These power-law distributions violate the assumptions the central limit theorem makes, for example, having a well-defined mean or variance. Specifically, one can not normalize or use averages from power-law distributed data, making the comparison of university publications impossible if they indeed follow a power-law distribution. While it has been shown that some scientific domains or universities show this power-law distribution, it is not known if Swedish universities also show this phenomenon. Thus, here we collect publication data for Swedish universities and determine whether or not, they are power-law distributed. Interestingly, if they are, one might use the slope of the power-law distribution as a proxy to determine research output. If the slope is steep, it suggests that the ratio between highly published authors and those with few publications is small. Whereas a flatter slope suggests that a university has more highly published authors than a university with a steeper slope. Thus, the second objective here is to assess if the slope of the distribution can be determined or to which extent this is possible. This study will show that eight of the fifteen Swedish universities considered follow a power-law distribution (Kolmogorov-Smirnov statistic<0.05), while the remaining seven do not. The key determinant is the total number of publications. The difficulty here is that often the total number of publications is so small that one can not reject a power-law distribution, and it is also impossible to determine the slope of the distribution with any accuracy in those cases. While this study suggests that in principle, the slopes of the power-law distributions can be used as a comparative metric, it also showed that for half of Sweden's universities, the data is insufficient for this type of analysis.

**Keywords:** Performance Metric, Swedish Universities, Power Law Distribution, Number of Publications, Slope Value, and Kolmogorov-Smirnov test

## 1. Introduction

The term "performance" is commonly used within the educational industry, and despite an increasing number of studies on university performance, there remains no universally accepted definition within the respective domain. Performance is a crucial consideration for all educational institutions, particularly in terms of research output, reputation, student satisfaction, graduation rates, employability of graduates, research funding, and internationalization as metrics. However, when it comes to evaluating publications, the concept becomes much more complex. This lack of clarity around the definition of performance presents a significant challenge within the educational field. Achieving an understanding of what constitutes a successful university, and identifying the performance indicators that determine success [6, 15].

A significant amount of research has been conducted to evaluate the performance of universities [3, 4, 5, 24]. Furthermore, there have been several studies on the relationship between university performance and economic development, showing that high-performance universities can contribute to the economic growth and development of their regions and countries [11, 12, 15].

Studying the performance of universities has not been a new challenge as different studies have evaluated the performance with the help of traditional parameters such as the size, revenue earned, and results produced by universities. And there have been different approaches to measuring the performance of universities. Furthermore, there is still a debate over which indicators are most appropriate and effective. More research is needed to refine and develop indicators that provide a comprehensive and accurate assessment of university performance [12, 15].

This research would provide confidence to the university administrators, students and parents, policymakers, and researchers in believing publications as performance metrics. As university administrators may want to study university performance to identify areas of strength and weakness, furthermore, prospective students and their parents may be interested in studying university performance to make informed decisions about which university to attend. Moreover, government officials and policymakers may be interested in studying university performance to assess the effectiveness of current policies and make decisions about funding and resource allocation for higher education. In addition, potential researchers may be interested in studying university performance to better understand the factors that contribute to student success and to identify best practices for improving student outcomes.

This study aims to assess Swedish universities' performance. The number of publications as a factor of measurement is easily obtainable and is considered to be accurate due to the high volume of publications produced by each university annually, among other metrics. It is not reasonable to average out the publications depending on the number of authors to assess the performance of universities because some universities have more faculty members with fewer publications and some universities have fewer faculty members with more publications, so their publications are not uniform in numbers. It is not appropriate to compare the averages of different universities to measure their performance or determine which one is better. This is because each university is unique in its number of faculty members.

Now, the challenge here is to figure out a reliable tool that would measure accurately the performance of universities using publications as a performance metric. The slope using the power law is proposed in the methodology section to obtain the output. The study also aims to figure out the possible difficulties that arise during the process and in the output.

The frequency distribution of various phenomena, including the number of publications, is commonly modeled using power law distributions. However, utilizing power law distributions to model publication data poses several challenges. For instance, these distributions are highly sensitive to the quality and size of the data sample used for estimating their parameters. If the data is inaccurate or the sample size is too small, the estimated power law distribution may be biased or not representative of the true distribution. Additionally, with a wide range of power law distributions available, selecting the appropriate one for a specific dataset can be problematic. Choosing an unsuitable distribution can result in inaccurate parameter estimates and poor model fit. Lastly, interpreting power law distributions can be challenging, particularly for non-experts.

### 1.1 Research Questions

RQ 1: Does the number of publications as a factor for performance metrics can be good in Swedish universities?

RQ 2: Can the slope of power law distributions be a potential alternative metric to measure research output?

## 2. Literature Review

### 2.1 Relevance to our Study

This literature review section focused on identifying and evaluating the existing indicators used to measure performance in universities. It would also explore the challenges and limitations of these attributes, particularly in the light of the distribution of the publications per researcher that makes normalization difficult.

The literature review has been categorized into two subsections i.e. citations/publications by authors / traditional metrics, and analytical tools used in the existing literature. The literature review has provided sufficient evidence for understanding the factors influencing the performance of universities. Additionally, the literature review could also explore the potential implication of new metrics on research excellence and collaboration among universities.

Overall, the literature review would aim to provide a comprehensive understanding of the current state of research performance metrics in universities and identify the areas for improvement and innovation in the field.

The above concepts will be discussed in detail in the proceeding sections of the literature review.

### 2.2 Citations / Publications by Authors / Traditional Metrics

The relationship between university publications, citations, and performance has been a topic of interest to many scholars. In this section of the literature review, a comprehensive summary of some key findings and research methods used are mentioned:

An article [15] aimed to study the performance of the universities based in China and the data is extracted from Scopus from 2007 to 2010. The author mentioned that they have presented useful indicators in comparison to US and European universities. The qualitative and quantitative performance of the universities was observed by using these indicators. And then observed the international collaboration, international impact metrics, and international citations. Research performance points were calculated by analyzing all the performances of the universities. There was a gap identified in comparison to the Chinese universities and other regional universities of various field areas. It was observed that the Chinese universities are doing a quite good job in publications but receiving fewer citations on their publications. It is also observed that Chinese universities have a low volume of publications in highly rated journals which could be the cause of not receiving citations.  So the writer proposed a careful selection of publication platforms to receive a high volume of citations.

Paper [17] explained the performance management system of Finland in academics. They used a mixed approach I.e. survey and qualitative and observed a academics have resulted in a negative view of performance management. The researchers also proposed several ways that could be effective in changing the nature of performance management in academics.

Authors in [1] sought to examine the factors that dictate a paper's "importance" by testing the hypothesis that papers deemed more important are cited more often in published works than those thought to be of lesser importance. Their data, within the experimental design's boundaries, indicated that citation frequency was an accurate measure of a paper's importance as judged by two separate groups. Multiple regression analysis further revealed that a linear combination of citation frequency and impact factor was a better predictor of importance than either one alone.

Villanova University's College of Commerce and Finance wanted to recognize and reward outstanding faculty research by organizing an annual research awards program. However, the process of selecting the best research entries became more challenging than expected, and several attempts to create a fair selection system failed. As a solution, the college decided to use the Analytic Hierarchy Process (AHP) to structure the selection procedure. The AHP method has been used for the past three years, and it has been successful with no major issues. Both the faculty and external advisors agree that the AHP method is better than the previous methods because it accurately reflects the subjective evaluations of the evaluators. In summary, the AHP method helped the College of Commerce and Finance to develop an effective and fair way to choose the best research entries for their annual awards program [2].

Evaluating the quality of published scientific works is a difficult task, and there is no single definitive answer. Ideally, experts in the field should review the research and allocate scores for quality and quantity according to established criteria. However, in reality, peer review is often conducted by committees with a broad range of skills, so they must rely on secondary indicators such as the number of publications, journal prestige, the reputation of the authors and organization, and the perceived importance and relevance of the research. Therefore, other methods such as citation rates and journal impact factors are being used as more objective and directly related measures of published science. Citations data can be obtained from the Institute for Scientific Information (ISI) and is presented as the Science Citation Index (SCI). The annual citation rate of a scientific author or research team can be calculated using the SCI and their publication list. Similarly, the citation rate of a scientific journal, or journal impact factor, can be calculated as the average citation rate of all the articles published in the journal. Journal impact factors are released annually in the SCI Journal Citation Reports [4].

Paper [13] differentiates the six different roles of the researchers. The data on the academic staff was collected from a survey of Norwegian universities over three decades. It was concluded that the main position of the researcher is important to the other small roles. It is also proposed by the academic staff that there is an increase in demand to increase the performance of their roles.

[10] Investigated the impact of publications of international in comparison to local papers. He picked up the top 100 papers, which were cited from three universities in four cities and two countries, taken from four research fields. The results concluded the internationally published papers were not represented in cited papers in the research fields but they were still leading the highly cited papers among all the papers from small cities and countries. The contribution of citations of international papers started to rise from 2001 to 2008 for universities, cities, and countries. It was concluded that the domestic papers from the USA contributed parts of the cited papers from all the papers in the research fields.

Authors in [6] proposed a "systems approach" which utilized citations from international journals to analyze the institutional and cognitive aspects of scientific excellence in national research systems. This methodology concentrates on the most highly cited papers and their distribution across institutions and disciplines. To demonstrate, we present the outcomes of a recent analysis of the research system in the Netherlands in the mid-1990s, with a particular focus on the contribution of the world's most highly cited research papers. The findings suggest that these papers provide an analytical framework that is both transparent and able to differentiate between institution and discipline, while also enabling domestic and international comparison. The average citation scores of these "Centers of Scientific Excellence" may not be an accurate predictor of the production of highly cited papers; thus, further analysis is needed to assess the full potential of this method for use in science policy and evaluation of research performance.

Authors in [9] examine the impact of research universities on the differences in innovation between clusters. Drawing from data on medical device clusters in the U.S. from a 12-year time frame, the authors demonstrate that research universities, which are a source of knowledge spillovers and generate graduates who disperse tacit knowledge within a cluster, are critical for innovative performance in regional technology clusters. Their findings back up the notion that there is spatial variability in cluster innovative performance and that financial, intellectual, and human capital in a cluster also helps its innovative performance.

Authors in [18] conducted interviews with 12 higher education managers in Ireland to explore their experiences, attitudes, and expectations related to the increased use of metrics in their institutions. The results showed that the introduction of metrics led to mixed reactions, with some managers finding them useful for measuring and assessing performance, while others saw negative impacts. The authors analyzed the data using narrative structuring and thematic analysis and identified six themes related to the impact of the transition to a metric-based system.

### 2.3 Analytical Tools Used in Literature

This section of the literature review provides evidence for the analytical tools used in the existing literature:

Previous research [16] explored the appearance of power law distribution on citation data taken from Scopus. The writer utilizes Scopus as a data source for plotting citation distribution, log-normal distribution, and using power law. The evidence that the power law that citation data from Scopus plays the best fit to the power law. The author presented confidence in the selection of the power law approach for representing the distributions explaining the power law on Scopus.

The use of power law based on the Python package helps define the distributions. It presents a set of tools to analyze the distribution based on the fact that the data set follows the power law distribution or not [14]. The main feature includes MLE (maximum likelihood estimation). It also shows the different visualizations explaining the plots and heavy-tailed distribution. The use of power law is suitable for network analysis, economics, finance, and natural science. It is particularly useful for examining the properties of heavy-tailed distributions that are frequently observed in real-world systems and for investigating the potential implications of such distributions for the creation of new models and theories. The power law is user-friendly and highly customizable, and it is available for download under an open-source license.

Authors in [8] outlined the potential problems that may crop up when attempting to fit a power-law distribution, such as the impact of outliers, the choice of model, and the choice of goodness-of-fit measure. It also presented strategies for addressing these issues.

Another research evidence has examined the prevalence of power-law distribution in economics and finance, which are characterized by a small number of extreme values that have a large impact on the overall distribution. The paper reviews empirical evidence for power-law distributions in various economic and financial variables and discusses potential explanations for their prevalence. Gabaix also considers the implications of power-law distributions for economic and financial policy, suggesting that policies to redistribute wealth and income may be necessary to mitigate their effects. Overall, the paper highlights the importance of considering power-law distributions in economic and financial analysis and policymaking [21].

Another research reveals an extensive overview of power-law distributions in natural and social phenomena. The paper defines power-law distributions and reviews empirical evidence for their prevalence in various phenomena, including wealth and income distribution, network structure, and city size distribution. The paper also discusses the mathematical properties of power-law distributions and their potential to inform our understanding of complex systems and to guide interventions to shape their behavior. Overall, the paper emphasizes the importance of power laws in understanding complex systems [21].

"Power laws, Pareto distributions, and Zipf's law" is a comprehensive literature review that explores the relationships between power-law distributions, Pareto distributions, and Zipf's law. The paper defines and compares these distributions and discusses their prevalence in various natural and social phenomena, as well as their potential mechanisms and implications for modeling complex systems. The authors emphasize the need for further research to fully understand these distributions and their properties [22,23].

Based on the literature reviewed, it is evident that performance can be assessed through several key performance indicators such as total publications, total citations, maximum citations, minimum citations, university size, student strength, etc. and these indicators vary depending on the industry. The main research gap observed during the literature review was no such study exists that uses power law distribution on publications of authors to assess performance. The impact of citation frequency and impact factors on a paper's importance may vary across different fields or disciplines. Further research could investigate how these factors differ in fields with different citation practices and publication norms. The cost-function model developed by Glass et al. [3] provides a framework for evaluating the efficiency of universities in terms of research and teaching outputs. However, it is unclear how well this model applies to universities in other countries or regions with different higher education systems or funding structures. Further research could investigate the generalizability of the cost-function model beyond the UK context.

In the study [10], the analysis was limited to four research fields. Expanding the analysis to include more fields of research could provide a more comprehensive understanding of the impact of international publications on local research. In the study by [9], the focus was on medical device clusters in the U.S. It would be interesting to examine the impact of research universities on innovation in other industries and regions to determine whether the findings of the study can be generalized. This study aims to evaluate the performance of universities, where the "number of publications" will be our key indicator to access the performance of the university to accomplish the task, the study needs to describe the output using the slopes of the power law.

## 3. Methodology

The methodology section explains the procedures to carry out research objectives and perform analyses. Data collection remained an important and sensitive part of this research. We have used Google Scholar as a source for the extraction of data for Swedish universities for seven years, from 2015 to the end of 2021. For this study, we attempted to include data from the most recent years. Specifically, we deemed publications from 2015 as current and took into account data from years up to 2021. We considered the data for the year 2022 incomplete at the time of data collection.

Our methodology section will explain the collection of data and the shaping of the data for analysis. This study has used Python Jupyter Notebook for conducting analyses and Ms. Excel for showing tables.

### 3.1. Data Extraction

Data collection is the process of gathering and measuring information on targeted variables in an established system, which then enables one to answer relevant questions and evaluate outcomes. Data collection is a research component in all study fields, including physical and social sciences, humanities, and business. In this case, no open-source data could be found. Therefore, we need to extract our custom data as per Figure 1. For this, we are using Google Scholar as our main data source.

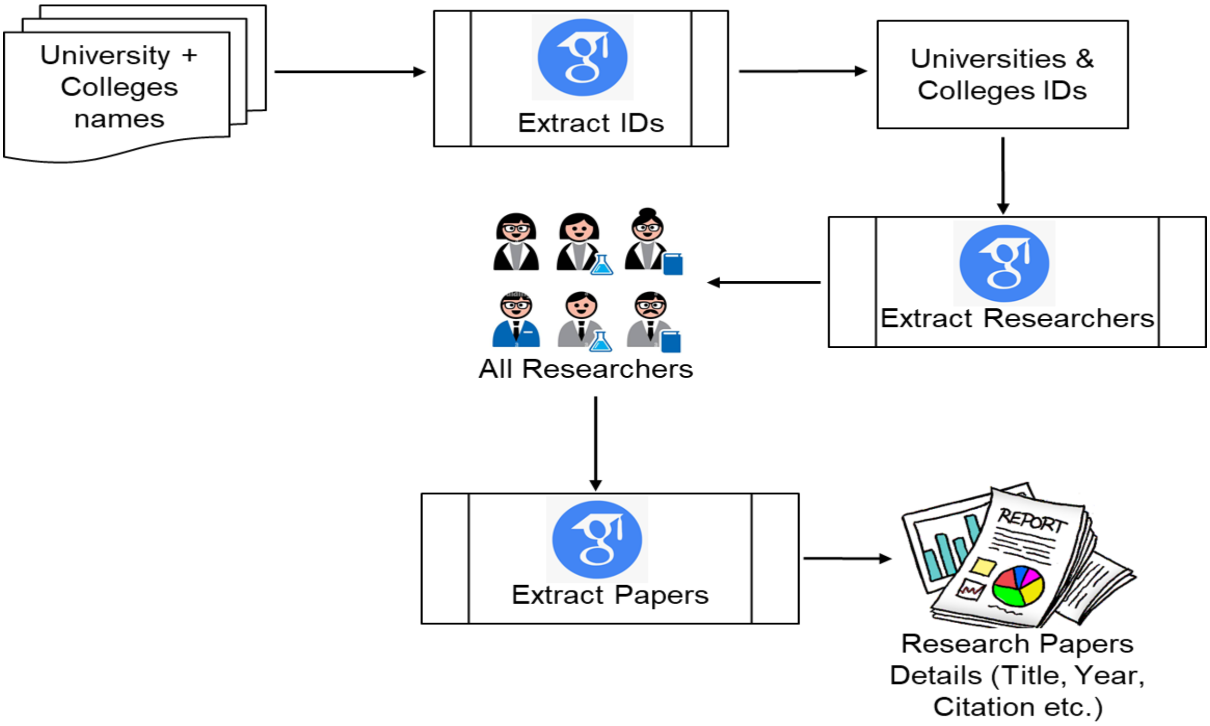


*Figure 1: Data Extraction Pipeline*

#### 3.1.1 Scholarly Data Extraction Tool

We started our data extraction with URLLib Python Library but could not continue with it due to some data extraction limitations, then we tried our data extraction journey with Selenium Framework but some limitations led us to stop the data extraction using Selenium.
After more research, we found Scholarly [20]. Scholarly is a module that allows you to retrieve author and publication information from Google Scholar in a friendly, Pythonic way without having to solve CAPTCHAs.

### 3.2. Data Extraction Steps

We have a pre-defined list of universities (listed in Appendix A). The universities covered areas across Sweden as shown in Figure 2, which shows the physical locations of universities in Sweden. This was only to make sure that our research is not region biased, or location-specific in Sweden.

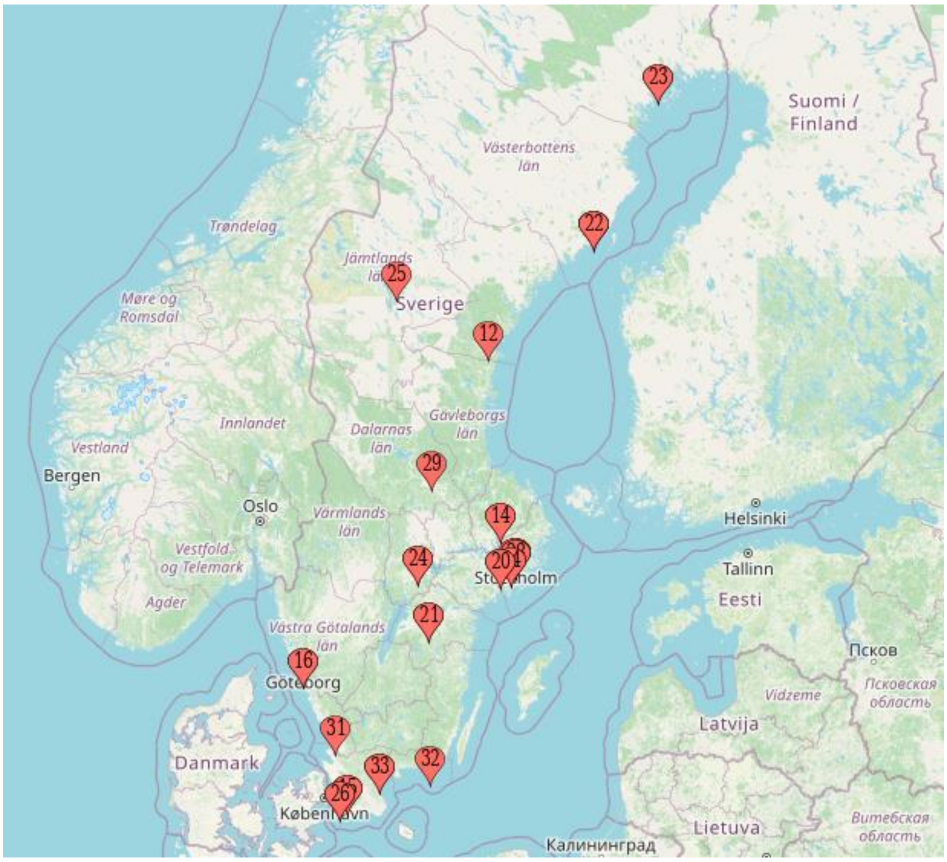


*Figure 2: Universities across Sweden (the focus of this study)*

There are so many options to get data from scholarly, such as you can search organizations by name or by id, searching authors by name or by id or you can search details of the publications. In our case, we have initiated data extraction using universities' names to enable them to access unique universities' IDs. The remaining data extraction process remains the same for extracting the authors' IDs and hence for the publications' IDs. Universities' IDs are unique for being extracted from a Google Scholar. The IDs of universities using names are listed in Table 1 below:

*Table 1: Institutes IDs mapping from Names*

| University Name | Unique ID |
|---|---|
| Uppsala University | 17301068057107671682 |
| Lund University | 11869288113056425402 |
| University of Gothenburg | 16819010087144438317 |
| Stockholm University | 8484286818777024965 |
| Karolinska Institute | 1915422036128507448 |
| Umea University | 14857806903653668437 |
| KTH Royal Institute of Technology | 10294267715845137919 |
| Linkoping University | 17173928468432796250 |
| Swedish University of Agricultural Sciences | 11771272703030442439 |
| Lulea University of Technology | 8354113489406250209 |
| Karlstad University | 14501015063993413853 |
| Orebro University | 13941517108005347753 |
| Mid Sweden University | 5342955148626852968 |
| Linnaeus University | 12745858568459922378 |
| Malmo University | 15973001373153570913 |

### 3.3. Data Description

We have extracted authors and publications of Swedish universities data as below in Table 2:

*Table 2: Data Description*

| Data Description | Value |
|---|---|
| Universities | 15 |
| University Researchers | 15,852 |
| University Research Papers | 1,049,429 |

Table 3 explains the important columns that exist in the extracted data. Each scholar has been assigned a unique id called scholar_id so there is no duplication of the researcher as the unique IDs for researchers are assigned by Google Scholar. Furthermore, publications ids are assigned for his / her publications in the data, likewise. Other important columns are 'name of researcher', 'email domain', 'field of interest', 'paper title', 'publication year', 'citations on paper', and 'cites per year', self-explanatory.

*Table 3: Data Description and Features*

| **Feature** | **Description** |
|---|---|
| scholar_id | Unique ID of the researcher |
| url_picture | The picture of the researcher |
| Name | Name of the researcher |
| Affiliation | Affiliation of researcher with the institute |
| email_domain | The email address being used |
| Interests | Area of interest of a researcher |
| Cited by | Total citation of the researcher so far |
| Hindex | h Index |
| i10index | ith 10 Index |
| Paper Title | Title of the paper |
| Publication Year | Year in which the paper is published |
| Citation on Paper | Citation received on paper |
| cites_per_year | Citations per year |

### 3.4. Power Law Distribution

$$Power\ Law = y = a * x^b \qquad (i)$$

In a power law function, a and b are parameters that determine the shape of the curve. The form of the power law function is referred to in equation 1, where x is the independent variable and y is the dependent variable. The parameter “a” is known as the normalization or scaling factor of the curve. It determines the overall magnitude of the curve and the position of the curve relative to the x-axis. The parameter “b” is known as the power law exponent. It determines the slope of the curve and the rate at which the dependent variable y changes as the independent variable x changes. A larger value of b corresponds to a steeper slope and a faster rate of change, while a smaller value of b corresponds to a flatter slope and a slower rate of change.

Power law distributions are often used in the analysis of citations/publications data because they provide a good fit to the data in many cases [16]. This is because power law distributions have a heavy tail, which means that a few authors will have an exceptionally substantial number of publications, while most authors will have a relatively small number of publications. This reflects the common observation that a small number of authors tends to be heavily published, while most authors only have a few publications [19].

In addition to providing a good fit to the data, power law distributions also have some useful mathematical properties. For example, the power law distribution can be used to model the distribution of many other types of data, such as the sizes of cities or the frequencies of words in a language. This makes it a useful tool for understanding and predicting the behavior of complex systems [7,14].

## 4. Result and Discussion

The result section explains the relationship between the number of publications and the performance of Swedish universities in the following subsections:

i. Exploratory Data Analysis
ii. Power Law Plots
iii. Kolmogorov Smirnov (KS-Test)
iv. Comparison Between Two Universities
v. Power Law Fit on Researchers vs Total Publications

### 4.1. Exploratory Data Analysis

Figure 3 shows the distribution of total papers published from the year 2015 to 2021 for Swedish universities. We can see that some universities have continuous but fewer publications, while some have more publications.

*Figure 3: Total Papers Published per Year for each University*

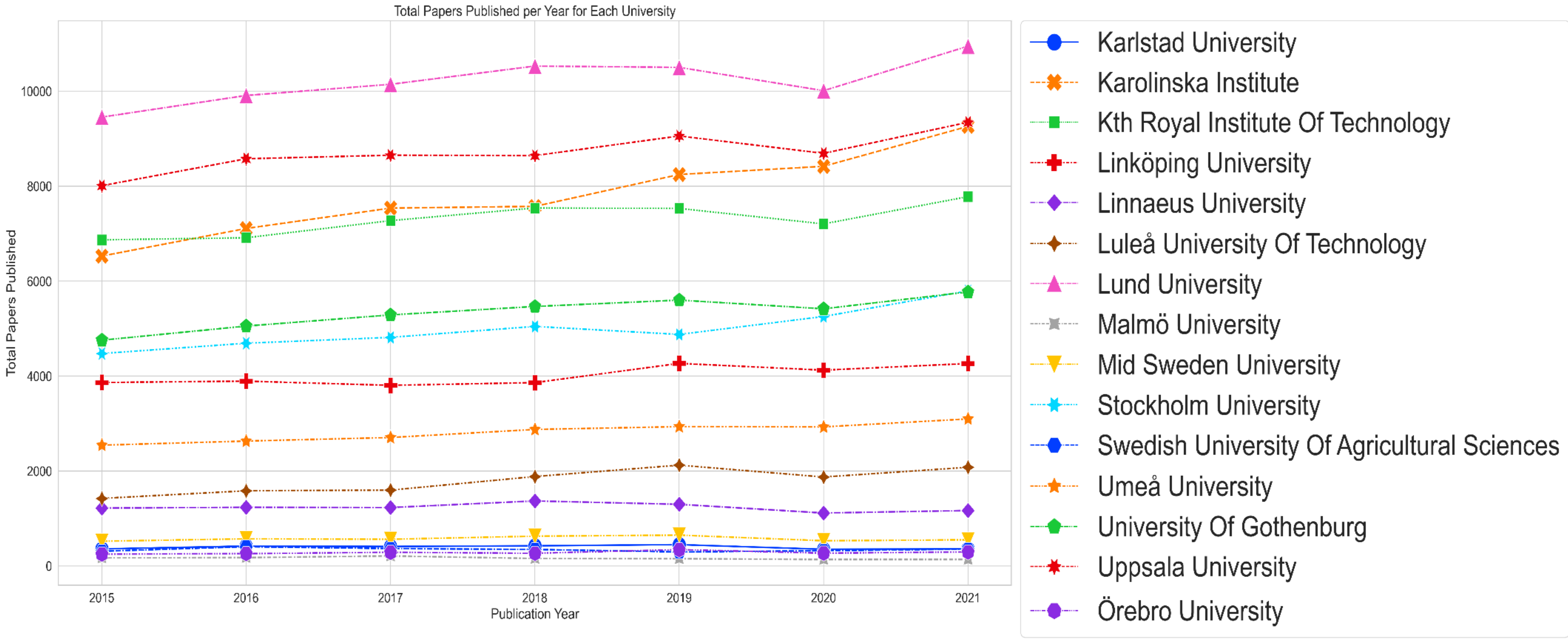

Table 4 shows the overall authors and publications related to a particular university. So, in our data, we have the total number of authors and the total number of research publications for each university.

*Table 4: Universities' Researchers vs Research Publications*

| University_Name | Total Authors | Total Publications |
|---|---:|---:|
| Karlstad University | 58 | 3064 |
| Karolinska Institute | 1722 | 62350 |
| Kth Royal Institute Of Technology | 2310 | 57087 |
| Linköping University | 1196 | 31448 |
| Linnaeus University | 371 | 9616 |
| Luleå University Of Technology | 519 | 14206 |
| Lund University | 2639 | 79554 |
| Malmö University | 47 | 1227 |
| Mid Sweden University | 161 | 4456 |
| Stockholm University | 1509 | 38933 |
| Swedish University Of Agricultural Sciences | 30 | 2650 |
| Umeå University | 907 | 22025 |
| University Of Gothenburg | 1527 | 42025 |
| Uppsala University | 2374 | 68478 |
| Örebro University | 30 | 2206 |

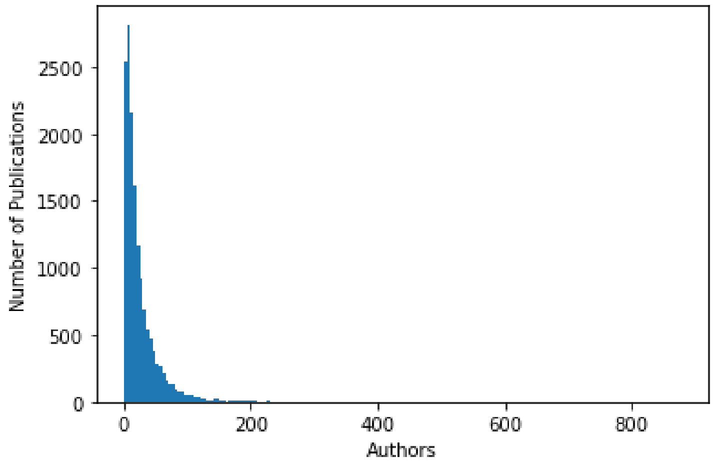


*Figure 4: Histogram of Number of Authors vs Publications (2015-2021)*

Visualizing data with probability density functions. A typical histogram as shown in Figure 4 on linear axis (intercepts) does not help visualize heavy-tailed distributions. On the log-log axis, using logarithmically spaced bins is necessary to accurately represent data as shown in Figure5. Linearly spaced bins obscure the tail of the distribution.

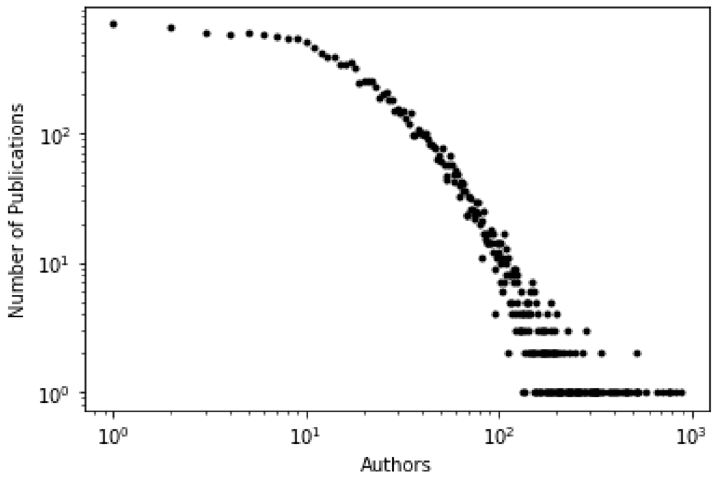


*Figure 5: Loglog plot instead of Histogram*

### 4.2 Power Law Plots

Figures 6 - 9, show the universities' plots on a log-log scale together with fitting power law distribution. The "Alpha (a) in Power Law" in Table 5, shows the value of the alpha parameter in the power law equation for each university (Appendix D). The alpha parameter determines the slope of the power law curve. A negative alpha value indicates a negative power law relationship, meaning that as the independent variable (number of research papers) increases, the dependent variable decreases. A positive alpha value indicates a positive power law relationship, meaning that as the independent variable increases, the dependent variable increases.

The "Beta (b) in Power Law" in Table 5, shows the value of the beta parameter in the power law equation for each university (Appendix D). The beta parameter determines the y-intercept of the power law curve. It represents the expected value of the dependent variable when the independent variable is equal to 1.

A steeper power law curve (a larger absolute value of alpha) indicates a stronger relationship between the two variables, while a shallower power law curve (a smaller absolute value of alpha) indicates a weaker relationship. Similarly, a larger beta value indicates a higher expected value of the dependent variable when the independent variable is equal to 1, while a smaller beta value indicates a lower expected value.

These plots reveal the relationship between authors and their publication in a university. We can see the power law plot (red color) and the actual data (black dots). When the actual data and power law curve follows the same trend (which means they are close to inline or inline), this shows the goodness of fit and vice-versa.

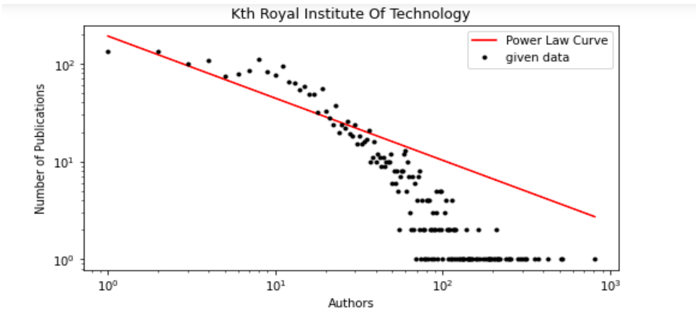


*Figure 6: KTH Institute*

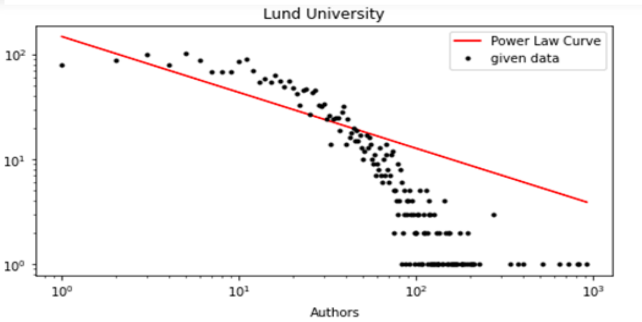


*Figure 7: Lund University*

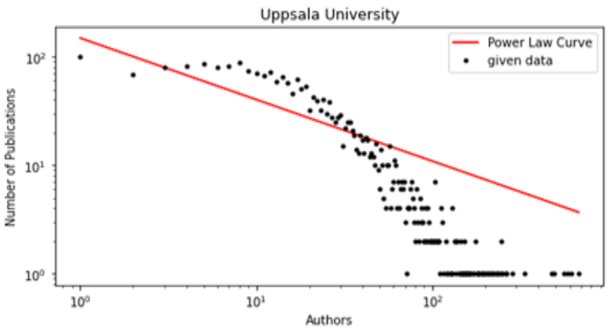


*Figure 8: Uppsala University*

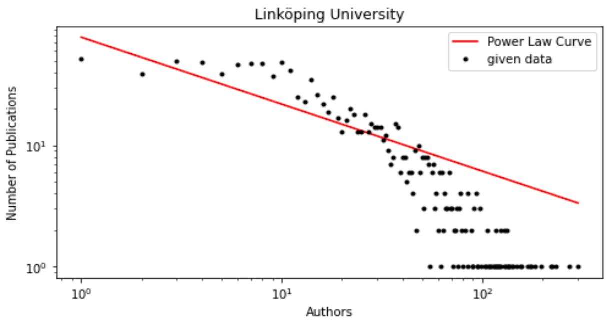


*Figure 9: Linkoping University*

`

### 4.3 Kolmogorov Smirnov (KS-Test)

The Kolmogorov-Smirnov (KS) test is a statistical test used to determine whether a sample of data follows a specific distribution. Following is the information required for the KS test in Appendix C.

- Sample size: This is the number of observations in the sample being tested.
- Significance level (p-value): This is the probability of observing a test statistic as extreme as the one calculated, assuming that the null hypothesis is true. It is often reported as a decimal number between 0 and 1, with values less than 0.05 indicating significant differences between the sample and the theoretical distribution.

As there are many possible combinations, we have written generic formula for sample size and p-value in Table 5 which shows the detailed values of power law parameters (alpha and beta) and the KS test parameters (Calculated KS, Actual KS, P-value, and sample size) to conclude the distribution. Based on the results, we can see that the values of alpha (a) and beta (b) vary among the universities, indicating that the relationship between the two quantities is not the same across all universities.

The goodness of fit is evaluated at Karolinska Institute, Kth Royal Institute Of Technology, Linköping University, Lund University, Stockholm University, Umeå University, University of Gothenburg, and Uppsala University. These universities have a strong power law relationship between their publication data, indicating that a few authors have more publications, while the majority have fewer publications count. The KS test results for these universities indicate that the p-values are less than 0.05 for almost all universities, so we will reject the null hypothesis that the data came from the normal distribution and calculated KS statistics test values are less than 0.05 for the above mentioned eight universities in Table 5. This means that their publication data distribution is a goodness of fit by a power law model which is strong enough to address research question 2 that Power Law slopes can be used as an alternate metric to measure the performance output.

Whereas, Luleå University of Technology, Karlstad University, Linnaeus University, Malmö University, Mid Sweden University, Swedish University of AgriSciences, and Örebro University, have a weak power law relationship among their publication data, indicating that their publication data is more evenly distributed. The KS test results for these universities indicate that their data distributions are not well fit by a power law model, with calculated KS statistics test values larger than 0.05 for these seven universities.

In summary, the power law analysis of the publications for the universities in Sweden shows that eight universities have a strong power law relationship with their publication data, while the rest of the seven universities have a weak power law relationship. This information can be used by stakeholders to understand the distribution of research impact across universities in Sweden and inform strategic decision-making which is addressing research question 1.

*Table 5: Power Law Distribution Parameters for Each University*

| University Name | No. of Authors | No. of Publications | Alpha (a) in Power Law | Beta (b) in Power Law | KS Statistics (D) | PValue | Sample Size (n) | Calculated KS Statistics (D') |
|---|---|---|---|---|---|---|---|---|
| Karlstad University | 58 | 3064 | -1.435935e-01 | 2.114115 | 0.174304 | 1.110223e-16 | 58 | 0.178577 |
| Karolinska Institute | 1722 | 62350 | -4.911104e-01 | 79.412359 | 0.048343 | 1.887379e-15 | 1722 | 0.032773 |
| Kth Royal Institute Of Technology | 2310 | 57087 | -6.363194e-01 | 192.165718 | 0.033970 | 5.551115e-16 | 2310 | 0.028297 |
| Linköping University | 1196 | 31448 | -5.527070e-01 | 78.063181 | 0.057241 | 2.220446e-16 | 1196 | 0.039325 |
| Linnaeus University | 371 | 9616 | -4.988371e-01 | 23.668862 | 0.092946 | 9.992007e-16 | 371 | 0.070608 |
| Luleå University Of Technology | 519 | 14206 | -5.169483e-01 | 34.280812 | 0.064065 | 0.000000e+00 | 519 | 0.059697 |
| Lund University | 2639 | 79554 | -5.332336e-01 | 147.977794 | 0.028190 | 1.110223e-16 | 2639 | 0.026474 |
| Malmö University | 47 | 1227 | -3.617794e-01 | 4.183546 | 0.137148 | 8.783863e-12 | 47 | 0.198376 |
| Mid Sweden University | 161 | 4456 | -2.643850e-01 | 5.805695 | 0.082662 | 0.000000e+00 | 161 | 0.107183 |
| Stockholm University | 1509 | 38933 | -5.244865e-01 | 90.048050 | 0.037589 | 0.000000e+00 | 1509 | 0.035010 |
| Swedish University Of Agri Sciences | 30 | 2650 | -4.054731e-10 | 1.000000 | 0.158084 | 1.942602e-07 | 30 | 0.248301 |
| Umeå University | 907 | 22025 | -4.892301e-01 | 51.332918 | 0.064468 | 4.440892e-16 | 907 | 0.045158 |
| University Of Gothenburg | 1527 | 42025 | -5.132771e-01 | 86.067338 | 0.053263 | 3.330669e-16 | 1527 | 0.034803 |
| Uppsala University | 2374 | 68478 | -5.678224e-01 | 149.141229 | 0.038703 | 1.998401e-15 | 2374 | 0.027912 |
| Örebro University | 30 | 2206 | -1.214757e-02 | 1.085206 | 0.184165 | 1.060263e-13 | 30 | 0.248301 |

### 4.4. Comparison of Two Universities

The comparison of the two universities in the result section is explaining a clear understanding of the two groups of universities. There is a goodness of fit for Uppsala University in the result of power law distribution and Malmo University is not a well fit in explaining the data distribution following the power law.

Let's compare the data of two universities and then the power law parameters and corresponding KS statistics. We have selected Uppsala University having values from Table 5, (-0.56 alpha and 149 beta) & (0.0279 calculated KS statistics and 0.00e-00 p-value). The data shows that the alpha and beta values follow the power law distribution. On the other hand, Malmo University with small data of publications and authors, having statistical values from Table 5, (-0.36 alpha and 4.18 beta) & (0.198 calculated KS statistics and 8.783863e-12 p-value). This sample shows that Malmo University is weak-fitted power law distribution as a comparison to Uppsala

University's power law plot. Figure 10 shows the power law plots comparison of both universities.

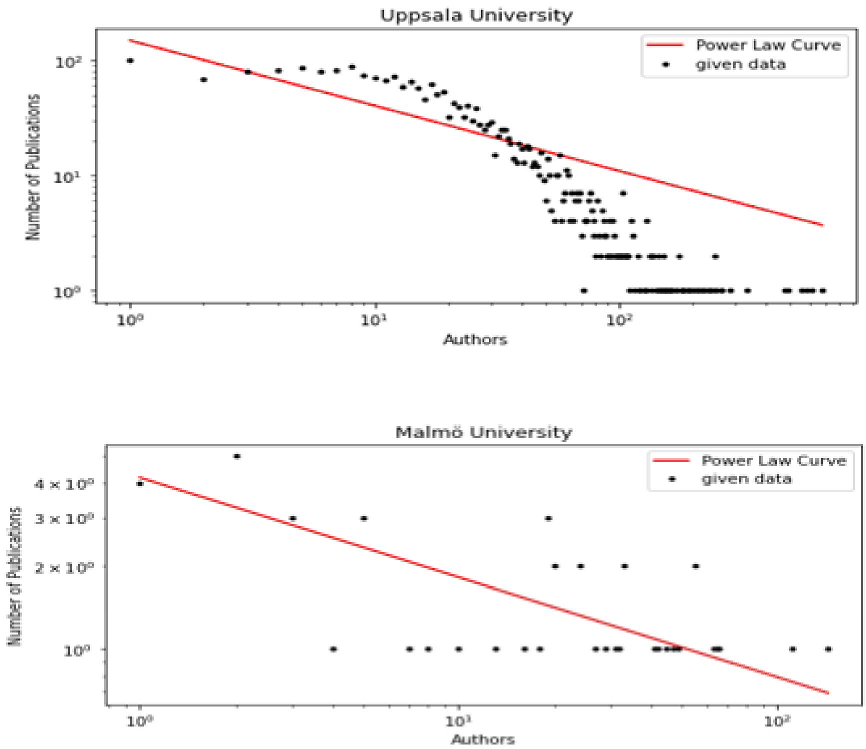


*Figure 10: Comparison of two universities*

### 4.5. Power Law Fit on Researchers vs Total Publications

The fit result "Fitted power law curve:

$$y = 2.323e^{-30} * x^{7.61} \qquad ii$$

refers to the parameters of a power law curve that has fitted to a set of data points.

Where "x" is the number of authors and "y" is the number of publications. In other words, "a" is the normalization factor and b is the power law exponent. In this case, the normalization factor is 2.3203e-30 and the power law exponent is 7.61.

The power law curve is a mathematical model that describes the relationship between the independent variable x and the dependent variable y in the data. The parameters “a” and “b” are chosen to minimize the difference between the model curve and the data points.

The goodness-of-fit measure, the R^2 value (also known as the coefficient of determination) is a measure of the percentage of the variance in the dependent variable that is explained by the model. R^2 value of 0.657 means that the power law curve explains approximately 65% of the variance in the data. However, individual university's performance using the power law equation in context with the goodness of fit is also referred to in Appendix D.

## 5. Conclusion

This study utilized a dataset spanning from 2015 to 2021 to analyze the performance of Swedish universities based on the number of publications (author-wise) as a performance metric. The study also aims to determine the possible difficulties arising during the process and obtain the output.

The power law method was employed to fit a curve and compare publication patterns and distributions across universities. Results revealed that eight universities, including Karolinska Institute, KTH Royal Institute Of Technology, Linköping University, Lund University, Stockholm University, Umeå University, University of Gothenburg, and Uppsala University, have strong goodness of fit to power law slopes, while the remaining seven universities, including the Luleå University of Technology, Karlstad University, Linnaeus University, Malmö University, Mid Sweden University, Swedish University of AgriSciences, and Örebro University, exhibit a weak power law distribution. This was determined through calculated KS statistics test values, with those below 0.05 indicating a good fit to the power law model. Therefore, a KS Statistics (D') < 0.05 for eight universities supports Power Law as a performance parameter for data, answering research question two effectively.

The study examined the goodness of fit of power law on the publications data output of fifteen universities. However, it was discovered that for seven of these universities, the data did not fit well with the power law curve. Additionally, the study encountered the following difficulties in formulating the publications data output using power law:

i. Power law distributions are sensitive to the quality and size of the data sample used to estimate their parameters. If the data is not accurate or the sample size is too small, the estimated power law exponent may be biased or not representative of the true distribution.
ii. There are many different types of power law distributions, and selecting the right one for a given dataset can be difficult. Choosing the wrong distribution can result in inaccurate parameter estimates and poor model fit.
iii. Power law distributions can be difficult to interpret, especially for non-experts. The parameters of the distribution may not have intuitive meaning and may require careful explanation.

This study is contributing to analyzing the performance of fifteen Swedish universities based on the number of publications using power law as a performance metric. Eight universities exhibited a strong fit to the power law distribution while the remaining seven showed a weak fit. The study faced difficulties such as data quality, model selection, and interpretation in formulating the publications' data output using power law. Overall, the study suggests that power law can be a useful parameter, but careful consideration must be given to these challenges when applying it to publications data.

## 6. Future Work and Limitations

After analyzing publication data from various Swedish universities, we concluded that some follow a power-law distribution. Therefore, comparing them based on the mean number of papers per author is not possible. For other universities, we lacked sufficient data to determine whether a power-law distribution exists, and if so, what slope it might have. Although the results are promising for comparing some universities using the slope of the power law, more work is needed to apply this method to the remaining universities. This could involve pooling data over multiple years, using a sliding window, or analyzing larger periods. Additionally, citation data might provide sufficient information and solve the problem of repetition in publication as we can generate a unique Id for an author but generating a unique Id for publications might not be possible in some cases as there could be multiple authors for a single publication. It's important to note that other metrics could also be appropriate and should be explored. We also didn't investigate the impact of the Covid pandemic on publications, which presents an interesting avenue for further research.

## 7. Acknowledgement

Our gratitude goes to the Lord Almighty for blessing us with the wisdom to explore the intriguing topic of "possible difficulties in evaluating university performance based on publications due to power law evidence from Sweden". Our journey was made smooth with the valuable guidance of our esteemed supervisor, Arend Hintze, and the support of the program director, Ilias Thomas, and examiner, Mia Xiaoyun. Additionally, we would like to say thanks to our teachers at Dalarna University and our family members for their unwavering support throughout the completion of our thesis project.

## 8. References

[1] Virgo, J. A. (1977). A statistical procedure for evaluating the importance of scientific papers. The Library Quarterly, 47(4), 415-430.

[2] Liberatore, M. J., Nydick, R. L., & Sanchez, P. M. (1992). The evaluation of research papers (or how to get an academic committee to agree on something). Interfaces, 22(2), 92-100.

[3] Glass, J. C., McKillop, D. G., & Hyndman, N. (1995). Efficiency in the provision of university teaching and research: An empirical analysis of UK universities. Journal of Applied Econometrics, 10(1), 61-72.

[4] Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. BMJ, 314(7079), 497.

[5] Miyata, Y. (2000). An empirical analysis of the innovative activity of universities in the United States. Technovation, 20(8), 413-425.

[6] Tijssen, R., Visser, M., & van Leeuwen, T. (2002). Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? Scientometrics, 54(3), 381-397.

[7] Goldstein, M. L., Morris, S. A., & Yen, G. G. (2004). Problems with fitting to the power-law distribution. The European Physical Journal B-Condensed Matter and Complex Systems, 41(2), 255-258.

[8] Drewes, T., & Michael, C. (2006). How do students choose a university?: an analysis of applications to universities in Ontario, Canada. Research in Higher Education, 47(7), 781-800.

[9] Rothaermel, F. T., & Ku, D. N. (2008). Intercluster innovation differentials: the role of research universities. IEEE Transactions on Engineering Management, 55(1), 9-22.

[10] Persson, O. (2010). Are highly cited papers more international? Scientometrics, 83(2), 397-401.

[11] Azma, F. (2010). Qualitative Indicators for the Evaluation of universities performance. Procedia-Social and Behavioral Sciences, 2(2), 5408-5411.

[12] Perkmann, M., Neely, A., & Walsh, K. (2011). How should firms evaluate success in university–industry alliances? A performance measurement system. R&D Management, 41(2), 202-216.

[13] Kyvik, S. (2013). The academic researcher role: Enhancing expectations and improved performance. Higher Education, 65(4), 525-538.

[14] Power law: a Python package for analysis of heavy-tailed distributions Jeff Alstott1,2014 no.14

[15] Zhu, J., Hassan, S. U., Mirza, H. T., & Xie, Q. (2014). Measuring recent research performance for Chinese universities using bibliometric methods. Scientometrics, 101(1), 429-443.

[16] Brzezinski, M. (2015). Power laws in citation distributions: evidence from Scopus. Scientometrics, 103(1), 213-228.

[17] Kallio, K. M., Kallio, T. J., Tienari, J., & Hyvönen, T. (2016). Ethos at stake: Performance management and academic work in universities. Human Relations, 69(3), 685-709.

[18] Stephens, S., & Gallagher, P. (2021). Metrics, metrics, metrics: the emergence of technological universities in Ireland. Quality Assurance in Education.

[19] Sandstro¨m U, van den Besselaar P (2016) Quantity and/or Quality? The Importance of Publishing Many Papers

[20] data source scholarly: https://pypi.org/project/scholarly/

[21] Gabaix, X. (2009). Power-laws in economics and finance. New York University, 1-39.

[22] Newman, M. (2005). Power laws, Pareto distributions, and Zipf's law, Cornell University, 46(1), 323-351.

[23] Shalizi, C. R., & Gelman, A., (2003), Power laws, Pareto distributions, and Zipf's law.

[24] Cotelnic, A. (2022). University performance: how we define it and how we measure IT. Eastern European Journal for Regional Studies (EEJRS), 8(1), 21-29.

## 9. Appendixes

### Appendix A: List of Swedish Universities Used in this Research

| University | Established as a university | First establishment | Students Population | Research Grants |
|---|---|---|---|---|
| Uppsala University | 1477 | 1477 | 1478 | 4.112 |
| Lund University | 1666 | 1425 | 1426 | 4.874 |
| University of Gothenburg | 1954 | 1891 | 1892 | 3.446 |
| Stockholm University | 1960 | 1878 | 1879 | 2.633 |
| Karolinska Institutet | 1965 | 1810 | 1811 | 4.805 |
| Umeå University | 1965 | 1965 | 1966 | 2.336 |
| KTH Royal Institute of Technology | 1970 | 1827 | 1828 | 2.836 |
| Linköping University | 1975 | 1969 | 1970 | 1.892 |
| Swedish University of Agricultural Sciences | 1977 | 1775 | 1776 | 2.083 |
| Luleå University of Technology | 1997 | 1971 | 1972 | 0.857 |
| Karlstad University | 1999 | 1977 | 1978 | 0.341 |
| Örebro University | 1999 | 1977 | 1978 | 0.364 |
| Mid Sweden University | 2005 | 1993 | 1994 | 0.371 |
| Linnaeus University | 2010 | 1967 | 1967 | 0.435 |
| Malmö University | 2018 | 1998 | 1999 | 1.382 |

Source: https://en.wikipedia.org/wiki/List_of_universities_and_colleges_in_Sweden

### Appendix B: Division of Project Work

| Research Work | Course of Conduct | Sarah Zia | Haroon Sadric |
|---|---|---|---|
| Introduction | The introduction was carried out by both co-partners as it presented the main idea of the research. | 50% | 50% |
| Literature Review | Literature was studied individually by both students however the entire literature work was reviewed by both students. | 40% | 60% |
| Methodology | Data extraction and analysis were carried out by both students. | 50% | 50% |
| Results, Findings, and Conclusion | Results, findings, and conclusions were also extracted and discussed between both students. Both students put a lot of effort into results, findings, and conclusions. | 50% | 50% |

## Appendix C: KS Test Critical Values

| SAMPLE SIZE (N) | LEVEL OF SIGNIFICANCE FOR D = MAXIMUM [ $F_0(X)$ - $S_n(X)$ ] | | | | |
|---|---|---|---|---|---|
| | .20 | .15 | .10 | .05 | .01 |
| 1 | .900 | .925 | .950 | .975 | .995 |
| 2 | .684 | .726 | .776 | .842 | .929 |
| 3 | .565 | .597 | .642 | .708 | .828 |
| 4 | .494 | .525 | .564 | .624 | .733 |
| 5 | .446 | .474 | .510 | .565 | .669 |
| 6 | .410 | .436 | .470 | .521 | .618 |
| 7 | .381 | .405 | .438 | .486 | .577 |
| 8 | .358 | .381 | .411 | .457 | .543 |
| 9 | .339 | .360 | .388 | .432 | .514 |
| 10 | .322 | .342 | .368 | .410 | .490 |
| 11 | .307 | .326 | .352 | .391 | .468 |
| 12 | .295 | .313 | .338 | .375 | .450 |
| 13 | .284 | .302 | .325 | .361 | .433 |
| 14 | .274 | .292 | .314 | .349 | .418 |
| 15 | .266 | .283 | .304 | .338 | .404 |
| 16 | .258 | .274 | .295 | .328 | .392 |
| 17 | .250 | .266 | .286 | .318 | .381 |
| 18 | .244 | .259 | .278 | .309 | .371 |
| 19 | .237 | .252 | .272 | .301 | .363 |
| 20 | .231 | .246 | .264 | .294 | .356 |
| 25 | .210 | .220 | .240 | .270 | .320 |
| 30 | .190 | .200 | .220 | .240 | .290 |
| 35 | .180 | .190 | .210 | .230 | .270 |
| OVER 35 | $\frac{1.07}{\sqrt{N}}$ | $\frac{1.14}{\sqrt{N}}$ | $\frac{1.22}{\sqrt{N}}$ | $\frac{1.36}{\sqrt{N}}$ | $\frac{1.63}{\sqrt{N}}$ |

## Appendix D: Results of Each University Using Power Law Equation

### Power Law=$y=a*x^b$

| | |
|---|---|
| University name = Karlstad University<br>Fitted power law curve: y = 392.88 * x ^ 0.01<br>Goodness-of-fit: R^2 = 0.002 | University name = Mid Sweden University<br>Fitted power law curve: y = 546.68 * x ^ 0.04<br>Goodness-of-fit: R^2 = 0.101 |
| University name = Karolinska Institute<br>Fitted power law curve: y = 6314.27 * x ^ 0.17<br>Goodness-of-fit: R^2 = 0.905 | University name = Stockholm University<br>Fitted power law curve: y = 4339.35 * x ^ 0.11<br>Goodness-of-fit: R^2 = 0.740 |
| University name = Kth Royal Institute Of Technolog y<br>Fitted power law curve: y = 6805.15 * x ^ 0.06<br>Goodness-of-fit: R^2 = 0.694 | University name = Swedish University Of Agricultur al Sciences<br>Fitted power law curve: y = 349.78 * x ^ -0.02<br>Goodness-of-fit: R^2 = 0.011 |
| University name = Linköping University<br>Fitted power law curve: y = 3753.28 * x ^ 0.05<br>Goodness-of-fit: R^2 = 0.519 | University name = Umeå University<br>Fitted power law curve: y = 2487.09 * x ^ 0.10<br>Goodness-of-fit: R^2 = 0.923 |
| University name = Linnaeus University<br>Fitted power law curve: y = 1255.10 * x ^ -0.01<br>Goodness-of-fit: R^2 = 0.022 | University name = University Of Gothenburg<br>Fitted power law curve: y = 4763.14 * x ^ 0.09<br>Goodness-of-fit: R^2 = 0.921 |
| University name = Luleå University Of Technology<br>Fitted power law curve: y = 1383.54 * x ^ 0.21<br>Goodness-of-fit: R^2 = 0.818 | University name = Uppsala University<br>Fitted power law curve: y = 8066.82 * x ^ 0.06<br>Goodness-of-fit: R^2 = 0.775 |
| University name = Lund University<br>Fitted power law curve: y = 9498.26 * x ^ 0.06<br>Goodness-of-fit: R^2 = 0.694 | University name = Örebro University<br>Fitted power law curve: y = 249.55 * x ^ 0.10<br>Goodness-of-fit: R^2 = 0.335 |
| University name = Malmö University<br>Fitted power law curve: y = 186.70 * x ^ -0.11<br>Goodness-of-fit: R^2 = 0.287 | |