Cairo University
Institute of Statistical Studies and Research
Department of Mathematical Statistics

On Parameters Estimation of Generalized Pareto Distribution

By

Ekramy Ahmed Hussein

Supervised by

Dr. Abdullah Mohamed Abdul Fattah
Associate professor of Mathematical Statistics
Institute of statistical studies and research

Dr. ELsayed Ahmed Elsherpieny
Associate professor of Mathematical Statistics
Institute of statistical studies and research

A Thesis Submitted To the Institute of Statistical Study and Research, In Partial Fulfillment of the Requirements for the Master Degree of Science in Statistics in the Department of Statistics

2010

APPROVAL SHEET

On Parameters Estimation of Generalized Pareto Distribution

By

Ekramy Ahmed Hussein

This Thesis for the Degree of Master in Mathematical Statistics, Institute of Statistical Studies and Research, Cairo University, has been approved by:

Name	**Signature**
Prof. Samia Abo El-Fetouh Salem	…………………………
Prof. Ahmed Fouad Attia	…………………………
Dr. EL-Sayed Ahmed El-Sherpieny	…………………………

ACKNOWLEDGEMENT

First, I pray and thank my **God** " **ALLAH** " for his uncountable gifts to me, without his assistance I could not be able to do any thing, much more thanks are to him.

I am extremely grateful to my supervisor Dr. Abdullah Mohamed Abdel-Fattah for his sincere help and his effort during his supervision of this thesis, constructive discussions, and especially for his invaluable remarks and encouragement which had been of great help to me.

I am deeply indebted to Dr. EL-Sayed Ahmed El-Sherpieny for his help, stimulating suggestions and encouragement helped me during all the time of research and writing of this thesis.

Moreover, I wish to express my sincere appreciation to my family for their constant encouragement, trying to save a lot of time for me to complete my thesis.

Finally, I would like to express my gratitude to all those who gave me the possibility to complete this thesis.

جـــامعـــة القــــاهــرة
معهــــد الــدراســـات والبحـــوث الاحصــــائيــــة
قســم الاحصـــاء الــريـــاضـى

عن تقـــديـــر معـــالم تـــوزيـــع بـــاريتــــو المعـمم

إعـــداد الطـــالـــب
إكـــرامى احمـــد حســـين خضــــرى

تحـــت إشـــراف

الدكتـــور
عبـــدالله محمـــد عبـــد الفتـــاح
أستاذ مساعـــد بقســـم الاحصاء الرياضى
معهـــد الـــدراسـات والبحـوث الاحصـــائيـة
جــــامعــــة القــــاهــرة

الدكتـــور
السيـــد أحمـــد الشـــربينـــى
أستاذ مساعـــد بقســـم الاحصاء الرياضى
معهـــد الـــدراســـات والبحـــوث الاحصــــائيـــة
جــــامعــــة القــــاهــرة

قـــدمـــت هـــذه الرســـالة إستكمـــالا لمتطلبـــات درجـــة الماجستيـــر فى الإحصــــاء الريــــاضـــى - معهـــد الـــدراســـات والبحـــوث الاحصــــائيـــة جــــامعــــة القـــــاهــرة

2010

الملخص العربى

عن تقدير معالم توزيع باريتو المعمم

ينسب هذا التوزيع الى العالم الاقتصادى الايطالى Vilfredo Pareto (1848-1929) الذى وضع اسس هذا التوزيع.وعلى الرغم من قلة استخدامات هذا التوزيع الا انه لاقى مجالا كبيرا للتطبيق وخصوصا فى موضوع الاقتصاد من خلال دراسة توزيع الدخول Incomes عندما تكون متجاوزة لحد معلوم مثل a أخذ توزيع باريتو أشكال مختلفة ؛ ومن امثلة هذا أرنولد سنة (1983) ليعطى تقريرا واسع المدى فى التاريخ بإستخدامة فى سياق توزيع الدخل ،ودالة الكثافة الاحتمالية لهذا التوزيع تأخذ الشكل التالى :

يقال ان المتغير العشوائى X يتوزع وفق دالة توزيع باريتو بثلاث معالم اذا كانت دالة الكثافة الاحتمالية للمتغير X تأخذ الشكل التالى:

$$f(x;\alpha,\beta,\lambda)= \frac{\alpha}{\beta}\left(1+\left(\frac{x-\lambda}{\beta}\right)\right)^{-(\alpha+1)} \qquad x>\lambda \quad ,\beta>0,\alpha>0$$

- توزيع باريتو العام

فى هذه الرســـالة تم استخدام شكل جديد لدالة توزيع باريتو ويعرف توزيع باريتو العام باربع معالم كما يلى .

يقال ان المتغير العشوائى X يتوزع وفق دالة توزيع باريتو بأربع معالم اذا كانت دالة الكثافة الاحتمالية للمتغير X تأخذ الشكل التالى:

$$f(x;\alpha,\beta,C,\lambda)= \frac{C\alpha}{\beta}\left(\frac{x-\lambda}{\beta}\right)^{C-1}\left[1+\left(\frac{x-\lambda}{\beta}\right)^{C}\right]^{-(\alpha+1)}$$

$\alpha>0\,,C>0$, $\beta>0$, $\lambda<x<\infty$ حيث λ تعتبر Location parameter ، β تعتبر scale parameter ، $(\alpha\,,C)$ يعتبران shape parameter

وتمتاز هذه الدالة لتوزيع باريتو بأنها تشتمل على كثير من الدوال السابقة لها بل ونستطيع من خلال هذه الدالة الحصول على بعض الدوال الاخرى بالتعويض عن المعالم بقيم مختلفة مثل عند التعويض عن قيمة $(C=1)$ نحصل على Three Parameters Pareto Distribution بالإضافة إلى نماذج اخرى

تتناول هذه الدراسة مشكلة تقدير معالم توزيع باريتو العام. وفيما يتعلق بطرق التقدير فقد تم استخدام كلا من طريقة الإمكان الاكبر و طريقة العزوم وذلك لتقدير معالم التوزيع فى حالة كلا من العينات الكاملة والعينات المبتورة.

تتكون هذه الدراسة من ستة ابواب . تم عرض الهدف الرئيسى من هذه الدراسة فى الباب الاول,الباب الثانى تم عرض بعض النعاريف والرموز التى استخدمت .بالنسبة للإستعراض المرجعى فتم عرضه بالباب الثالث. الباب الرابع تم عرض النموذج المقترح ودراسة خصائصة الإحصائية وتقدير معالمة بإستخدام طريقتى الإمكان الأعظم والعزوم ، وذلك فى حالة العينا ت الكاملة من خلال مثال عددى .الباب الخامس تم دراسة الخصائص الإحصائية للنموذج المقترح ، وكذلك تقدير معالمة بإستخدام طريقة العزوم ، وذلك فى حالة العينات المبتورة من خلال أمثلة عددية. اما عن تقدير معالم النوذج المقترح فى حالة العينات المبتورة بإستخدام طريقة الإمكان الأعظم من خلال مجموعة من الأمثلة العددية فتم عرضه بالباب السادس.

Contents

Chapter (I)
Introduction

The Pareto distribution was introduced by Pareto (1897) to develop representation of the income distribution and can be used to represent various other forms of distributions (other than income-data) that arise in human life. It has played a very important role in the investigation of city population sizes, occurrence of natural resources, insurance-risk, business failures etc. Arnald & Press (1983) gives an extensive historical survey of its use in the context of income-distribution as three parameters Pareto distribution. Abdul Fattah, Elsherpieny, and Hussein (2007) generalized the three parameters Pareto distribution by adding another shape parameter and studied the properties of the suggested distribution as well as its relations with other.

The main objective of this thesis is to estimate the unknown parameters of the generalized Pareto distribution using complete and truncated distribution.

It is observed that the four parameters Pareto distribution can be specialized to many distributions and some compound distributions. For the complete and truncated generalized type of Pareto models, some of its properties are investigated. The problem of obtaining the maximum likelihood and moment estimators of the unknown parameters for complete and truncated distribution is considered, numerical examples are presented to illustrate the new results.

This thesis consists of fife chapters plus this Introduction. Chapter II consists of some definitions and notation which will be used, Chapter III Pareto distribution a literature review. Chapter IV is concerned with the statistical properties, the maximum likelihood and the moment estimators and confidence interval estimators for the four parameters Pareto distribution. Chapter V deals with the statistical properties and Moment estimators for left, right and doubly truncated of the four parameters Pareto distribution. The maximum likelihood estimators and their approximate asymptotic variance-covariance matrix for left, right and doubly truncated are obtained in Chapter VI.

All computations, in this thesis were performed using computer facilities and MATHCAD PACKAGE (2001). In all chapters IV, V and VI, numerical examples are provided to illustrate the theoretical results in the complete and truncated data.

CHAPTER (II)

Definitions and Notation

This chapter contains the most important definitions and notation used in this thesis.

(2.1) Some Properties of Point Estimation

Various statistical properties of estimators can thus be used to decide which estimator is most appropriate in a given situation, which will expose us to the smallest risk, and give us the most information at the lowest cost, and so forth. Here the properties of point estimation which are unbiasedness efficiency, consistency and sufficiency will be reviewed.

(i) Unbiased Estimator

For any sample size, Larsen (1981) gave the following definition.

Definition (2.1)

Let $(x_1, x_2, \ldots, x_n)$ be a random sample from $f(x;\theta)$, an estimator $W = w(x_1, x_2, \ldots, x_n)$ is said to be unbiased for θ if $E(W) = \theta$, for all θ, otherwise, it will be biased. The bias of W is defined by $bias(W) = E(W) - \theta$, intuitively if W is an unbiased estimator for θ, then the distribution of W is centered around θ, and there is no persistent tendency to under or overestimate θ.

Definition (2.2)

Let $(x_1, x_2, \ldots, x_n)$ be a random sample, whose probability density function depends on an unknown parameter θ, and let W be any statistic then the mean square error of W is

$$\begin{aligned} MSE(W) &= E\left((W-\theta)^2\right) \\ &= Var(W) + (bias(W))^2. \end{aligned}$$

(ii) Efficiency

In general, there may be several unbiased estimators of a parameter θ for samples of any size. Larsen (1981) gave the following definition:

Definition (2.3)

Let W_1 and W_2 be two unbiased estimator for unknown parameter θ with variances $Var(W_1)$ and $Var(W_2)$ respectively, then W_1 will be more efficient than W_2 if $Var(W_1) < Var(W_2)$. Also, the relative efficiency of W_1 with respect to W_2 will be defined as the ratio $Var(W_2)/Var(W_1)$

In the case of unbiased estimator, if there is one estimator from those unbiased estimators have the minimum variance then it is called the most efficient, that is to say, if there was a statistic W has the following properties:

1. W is unbiased estimator of unknown parameter θ
2. W has minimum variance from all the unbiased estimators of unknown parameter θ

Then, W is called the minimum variance unbiased estimator (MVUE) and the variance of W is the same as the following Carmer-Rao Inequality

$$Var(W) \geq \frac{1}{n.E\left(\left(\frac{\partial \ln f(x;\theta)}{\partial \theta}\right)^2\right)} = \frac{1}{-n.E\left(\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2}\right)}$$

Where $f(.;\theta)$ is the value of the population density at x and n is the size of the random sample

(iii) Consistency

Suppose that X is a random variable with probability density function $f(x;\theta)$ and depends on unknown parameter θ. If $(x_1, x_2, \ldots, x_n)$ is a random sample of size n from X, then the statistic W_n is a consistent estimator of the parameter θ, if as the sample increase it is expected that W_n approaches to the parameter. Mendenhall et al. (1981) gave the following definition and theorem:

Definition (2.4)

The estimator W_n is said to be a consistent estimator of the parameter θ if, for any positive number ε,

$$\lim_{n\to\infty} p(|W_n - \theta| \le \varepsilon) = 1,$$

or, equivalently,

$$\lim_{n\to\infty} p(|W_n - \theta| > \varepsilon) = 0,$$

This mean that statistic W_n is a consistent estimator of the parameter θ when expected that the sampling distribution of W_n should become increasingly concentrated around the true parameter value θ.

Theorem (2.1)

If W_n is an unbiased estimator of the parameter θ and $Var(W_n) \to 0$ as $n \to \infty$, then W_n is a consistent estimator of θ.

(iv) Sufficient Estimator

A statistic W is sufficient estimator for the unknown parameter θ if W contains all of the information about the unknown parameter θ that is available in the entire data variable X. This idea is clear in the following definition.

Definition (2.5)

Let $(x_1, x_2, \ldots, x_n)$ denote a random sample from a probability distribution with unknown parameter θ. Then the statistic $W = w(x_1, x_2, \ldots, x_n)$. is said to be sufficient for θ if the conditional distribution of $(x_1, x_2, \ldots, x_n)$ given W does not depend on θ [Mendenhall et al. (1981)].

From definition (2.5), sufficient statistic is not very workable for the following reasons. First, it does not tell which statistic is likely to be sufficient and second it requires deriving a condition which may not be easy especially for continuous random variable. So it is usually easier to base it instead on the following factorization theorem to find sufficient statistic which was mentioned by [Mendenhall et al. (1981)].

Theorem (2.2)

Let $(x_1, x_2, \ldots, x_n)$ be a random sample of size n from the density $f(.;\theta)$ where the parameter θ may be a vector. A statistic $W = w(x_1, x_2, \ldots, x_n)$is sufficient if and only if the joint density of $(x_1, x_2, \ldots, x_n)$ which is $\prod_{i=1}^{n} f(x_i, \theta)$, a factor as

$$f(x_1, x_2, \ldots, x_n) = g(W, \theta)h(x_1, x_2, \ldots, x_n)$$

where the function $h(x_1, x_2, \ldots, x_n)$ does not involve the parameter θ and the function $g(W, \theta)$ depends on W and θ.

(2.2) Some Methods of Point Estimation:

The theory of estimation consists of these methods by which we make inference or generalization about a population parameter. Let Ω denoted the set of all possible values of the unknown parameter θ and is called the parameter space. Suppose that we wish to estimate the population parameter θ, there are two distinct ways. First, to obtain a single number that should be close to the unknown population parameter θ , this type of estimation is called point estimation. Second, it might say that θ will fall between two numbers which are intended to enclose the parameter of interest, this type of estimation is called interval estimation. A point estimation procedure utilizes the information in the sample to arrive at a single number or point that estimates the target parameter.

Definition (2.6) Point Estimation

An approximate value of a population parameter on a basic sample statistic is called an estimator of the population parameter and it depends on the random sample.

Definition (2.7) Statistic

A function of random sample $T = u(x_1, x_2, \ldots, x_n)$ that does not depend on any unknown parameters is called a statistic.

Definition (2.8) Estimator

A statistic $T = u(x_1, x_2, \ldots, x_n)$ that is used to estimate the value of θ and observed value $u(x_1, x_2, \ldots, x_n)$ is called an estimator of θ. There are different procedures to obtain point estimators, some of possible estimators better than the others; different statistical properties can be used to decide which estimator is most appropriate in a certain situation. The four criteria that optimum or best estimators should satisfy are unbiasedness, efficiency, consistency and sufficiency.

Some measures of statistical properties of estimator, method of maximum likelihood and moment of estimation and Bayesian estimation will be discussed in the following section.

(2.2.1) Method of Moments

Let $(X_1, X_2, \ldots, X_n)$ be independent identically distributed random variables with pdf $f(X_i, \theta_1, \theta_2, \ldots, \theta_s)$ having S parameters $\theta_1, \theta_2, \ldots, \theta_s$ let $\mu_r^{\backslash}(\theta_1, \theta_2, \ldots, \theta_s) = \mu_r^{\backslash} = E(X^r)$ be the r^{th} population moment (if it exists) and let $m_r^{\backslash} = \sum_{i=1}^{n} \frac{x_i^r}{n}$ be the corresponding r^{th} sample moment. The method of moments consists of equating a suitable number of population moments $\mu_r^{\backslash}$ to the corresponding sample moments $m_r^{\backslash}$ and solving the resulting equating equations for $(\theta_1, \theta_2, \ldots, \theta_s)$, i.e. we solve