UNLV Retrospective Theses & Dissertations

1-1-1996

# Estimating Pareto's constant and Gini coefficient of the Pareto distribution

Brian Christopher Hodge
*University of Nevada, Las Vegas*

Follow this and additional works at: https://digitalscholarship.unlv.edu/rtds

## Repository Citation

Hodge, Brian Christopher, "Estimating Pareto's constant and Gini coefficient of the Pareto distribution" (1996). *UNLV Retrospective Theses & Dissertations*. 3278.
http://dx.doi.org/10.25669/sk4h-ihv5 processed, response: 201

This Thesis is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Thesis has been accepted for inclusion in UNLV Retrospective Theses & Dissertations by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

# INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ESTIMATING PARETO'S CONSTANT

AND GINI COEFFICIENT OF THE

PARETO DISTRIBUTION

by

Brian Christopher Hodge

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science
in
Mathematics

Department of Mathematical Sciences
University of Nevada, Las Vegas
May 1997

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 1385176

UMI Microform 1385176
Copyright 1997, by UMI Company. All rights reserved.

This microform edition is protected against unauthorized
copying under Title 17, United States Code.

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The Thesis of Brian C. Hodge for the degree of Master of Science in Mathematics is approved.

Malwane Ananda

Chairperson, Malwane M. A. Ananda, Ph.D.

Examining Committee Member, Chih-Hsiang Ho, Ph.D.

Ashok K. Singh

Examining Committee Member, Ashok Singh, Ph.D.

T. M. Carroll

Graduate Faculty Representative, Thomas Carroll, Ph.D.

Dean of the Graduate College, Ronald W. Smith, Ph.D.

University of Nevada, Las Vegas
May 1997



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

## ABSTRACT

The Pareto distribution is widely used to describe the distribution of incomes. The distribution can be defined by a location parameter $c$ and a shape parameters $a$. The maximum likelihood estimators (MLE) and the uniform minimum variance unbiased estimators (UMVUE) for parameters $a$ and $c$ are available in current literature. An improved estimator for parameter $a$ will then be found using a frequentist decision theoretic technique. It is then shown that this improved estimator is the generalized bayes estimator for a certain choice of priors.

The Lorenz curve is a tool used to evaluate the share of income of sub-populations or to measure the inequality of individual income distributions. The Lorenz curve for the Pareto distribution will be discussed.

The Gini coefficient is a summary measure of the Lorenz curve. The Gini coefficient for the Pareto distribution will be shown to be a function of shape parameter $a$. The MLE of the Gini coefficient will be discussed. An alternative estimator for the Gini coefficient will then be constructed by using the generalized bayes approach.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS





Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF FIGURES





Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

# ACKNOWLEDGEMENTS

My sincerest gratitude to Dr. Malwane Ananda for his years of guidance and understanding.

I would also like to thank Dr. Singh and Dr. Ho for there devoted time in the classroom.

As a teaching assistant I can appreciate the tedium sometimes felt by instructors.




Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

# CHAPTER 1

# INTRODUCTION TO THE PARETO DISTRIBUTION

## The 'Law' of Income Distribution

Vilfredo Pareto (1848-1923) is a well known leader in early economic theory. Many feel his law of income distribution is his most significant contribution to economic theory. The 'law' still draws attention from economists and econometricians today because they realize the important, early example Pareto set in econometric investigation. Pareto's theory was originally based solely on the income descriptive statistics of Austria. He then expanded his ideas to the income statistics of several other western economies. Born in Paris on July 15, 1848, Vilfredo Pareto was a pioneer in the theory of economics as the western world matured to an advanced stage of industrialization. Pareto studied economic theory up until 1912, when he turned to sociological research. It was in his first major work, the two volume *Cours d'economie Politique* published from 1896 to 1897, that Pareto introduced his 'law' of income distribution and became a major figure in early economic theory.

Pareto believed that above a determined level, the distribution of incomes was static and followed a defined pattern. This was significant in that it argued against the idea that income inequality could be reduced by simply changing the distribution. Pareto's distribution hypothesized that only an increase in production with respect to the population size would decrease income inequality. He initially felt the distribution



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

applied only to the populations that his empirical data were taken from, but he later generalized the distribution to all populations. Yet, however it is applied, the 'law' of income distribution remains a "unique pioneering work of econometric investigation".( Cirillo [4])

## Definition

Pareto's 'Law' of Income Distribution can be generalized by the equation:

$$N_0 = AX^{-a} \qquad (1)$$

where $X$ is annual income, $N_0$ is the number of persons obtaining that or a greater income, and $A$ and $a$ are parameters found through empirical statistics.

The standard form of Pareto's density found in most statistical and econometric texts:

$$f(x;a,c) = ac^a x^{-(a+1)}, \qquad c \le x, \quad 0 < a, \quad 0 < c \qquad (2)$$

where $a$ is known as Pareto's constant and as a shape parameter and $c$ a scale or location parameter.

Equations (1) and (2) can be related in that $N_0 = N \cdot P(X > x_0)$ where N is the population size.

$$N_0 = N \cdot P(X > x_0)$$

$$= N\int_{x_0}^{\infty} f(x)dx = Nac^a \int_{x_0}^{\infty} x^{-a-1}dx$$

$$= Nac^a \left( \frac{x^{-a}}{-a} \right)\Bigg|_{x_0}^{\infty} = Nc^a x_0^{-a}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

so that $A = Nc^a$ and $X = x_0$.

The cumulative distribution function of X, i.e. 'income', can be written:

$$F(x) = \int_c^x f(t)dt \qquad c > 0, \quad a > 0; \quad x \geq c$$

$$= ac^a \int_c^x t^{-a-1} dt = ac^a \left( \frac{t^{-a}}{-a} \right) \Bigg|_c^x \qquad c > 0, \quad a > 0; \quad x \geq c$$

$$= 1 - \left( \frac{c}{x} \right)^a \qquad c > 0, \quad a > 0; \quad x \geq c$$

which is known as the "Pareto distribution of the first kind".(Johnson and Kotz [7])

Several graphs of the pareto distribution follow.

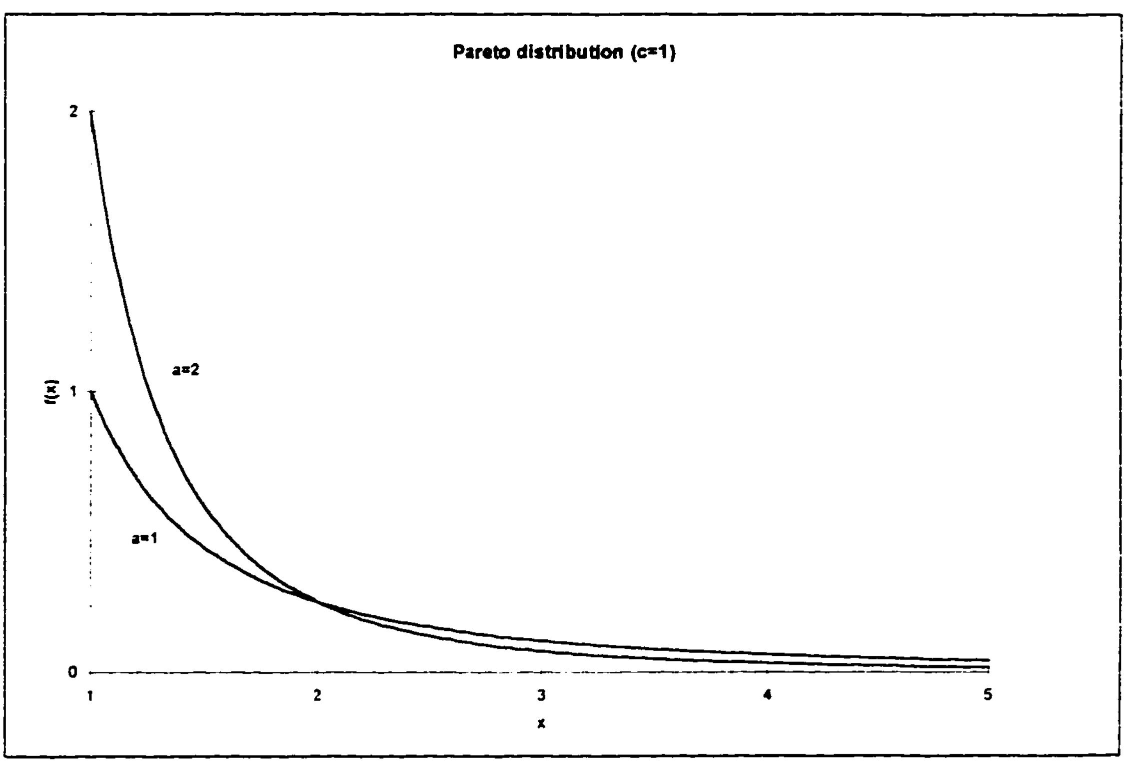


**Figure 1 Graphs of Pareto Density Functions for c=1, a=1 and c=1, a=2.**

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

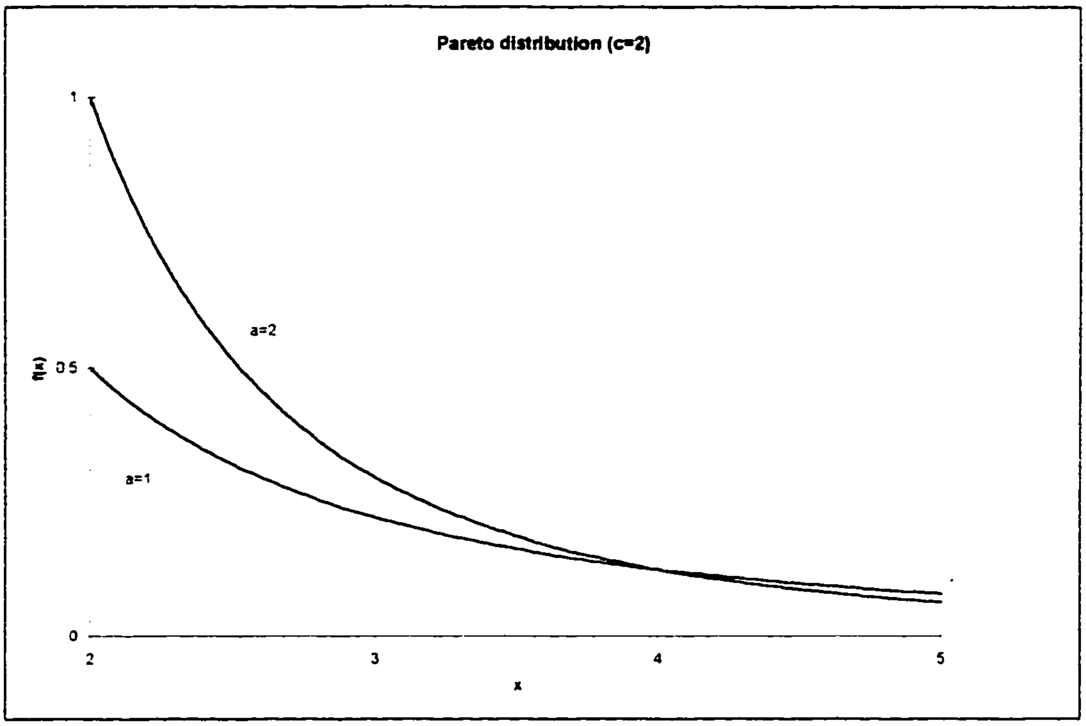


**Figure 2 Graphs of Pareto Density Functions for c=2, a=1 and c=2, a=2.**

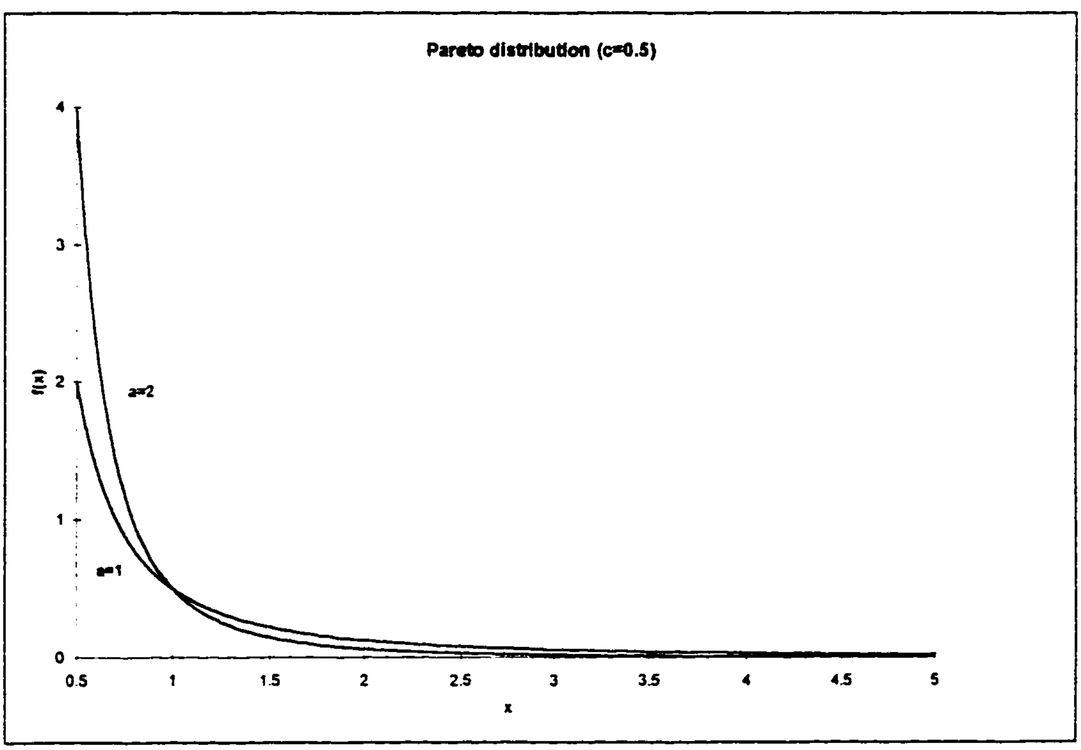


**Figure 3 Graphs of Pareto Density Functions for c=0.5, a=1 and c=0.5, a=2.**

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

# CHAPTER 2

## CLASSICAL ESTIMATION OF PARAMETERS

### Introduction

For the case of unknown shape parameter $a$ and unknown location parameter $c$, it is shown by Malik [10] that a joint sufficient statistic for $a$ and $c$ is

$$\left( \frac{n}{\sum_{i=1}^{n} \ln\left( X_i / X_{(1)} \right)} , \quad X_{(1)} \right). \tag{3}$$

It can also be shown that the statistics $Y = \sum_{i=1}^{n} \ln\left( X_i / X_{(1)} \right)$ and $Z = X_{(1)}$ are stochastically independent. It will be shown that these sufficient statistics are in fact the maximum likelihood estimates of $a$ and $c$. This can be foreseen as the MLE's are simply functions of the sufficient statistics (3). Malik uses the sufficiency and independence of $Y$ and $Z$ to construct the distributions $f_Y(y)$ and $f_Z(z)$. This is equivalent to finding the distributions of the maximum likelihood estimates of $a$ and $c$. The density $f_Y(y)$ will be essential to the development of an improved estimate for $a$ in the following chapter.

The probability density function of the order statistic $Z = X_{(1)}$ is

$$f_Z(z) = n[1 - (1 - (\tfrac{c}{z})^a)]^{n-1} \cdot ac^a z^{-(a+1)} \qquad c < Z$$



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$f_Z(z) = \begin{cases} nac^{na} z^{-(na+1)} & c < Z \\ 0 & otherwise \end{cases}.$$

The probability density function of $Y$ was found by Malik to be

$$f_Y(y) = \begin{cases} \dfrac{a^{n-1} n^{n-1}}{\Gamma(n-1) y^n} \cdot \exp\left(-{}^{na}/_{y}\right) & y \geq 0 \\ 0 & otherwise \end{cases}.$$

## The MLE and UMVUE of *a* and *c*

The Maximum Likelihood Estimators (MLE) of $a$ and $c$ are found using the likelihood function

$$L(a, c) = \prod_{j=1}^{n} \frac{ac^a}{X_j^{a+1}}, \qquad c < x_{(1)}$$

$$= \frac{a^n c^{na}}{\left(\prod_{i=1}^{n} x_i\right)^{a+1}} = \frac{a^n c^{na}}{(x_1 x_2 \cdots x_n)^{a+1}}, \qquad c < x_{(1)} \quad (4)$$

where $L(a, c)$ is the joint probability density function (p.d.f.) of $X_1, X_2, \ldots, X_n$. What then are the values of $a$ and $c$ that would maximize the probability $L(a, c)$ of getting the sample values $x_1, x_2, \ldots, x_n$? These values of $a$ and $c$ are the well known maximum likelihood estimators of $a$ and $c$. For parameter $a$, take the logarithms or alternatively the natural logarithms of both sides of (4). This step is consistent with the fact that $L(a, c)$ and $\ln L(a, c)$ are maximized for the same values of $a$ and $c$. For the known likelihood function, we then have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\ln L = \ln\left(\frac{a^n c^{na}}{(x_1 x_2 \cdots x_n)^{a+1}}\right)$$

$$= n\ln a + na\ln c - (a+1)\ln(x_1 x_2 \cdots x_n).$$

Take the partial derivative of $\ln L$ with respect to $a$, obtaining

$$\frac{\partial \ln L}{\partial a} = \frac{n}{a} + n\ln c - \ln(x_1 x_2 \cdots x_n). \qquad (5)$$

Setting the result (5) equal to zero and solving for $a$, an association between parameters $a$ and $c$ is found to be

$$\hat{a} = \frac{n}{\sum_{i=1}^{n} \ln\left(\frac{x_i}{\hat{c}}\right)}.$$

Simply noting the sample space of a pareto distributed random variable to be $0 < c < x$, the MLE of $c$ can then be found by inspection to be the first order statistic $\hat{c}_{ML} = x_{(1)}$.

The MLE's $\hat{a}_{ML}$ and $\hat{c}_{ML}$ of $a$ and $c$ are then

$$\hat{a}_{ML} = \frac{n}{\sum_{i=1}^{n} \ln\left(X_i / X_{(1)}\right)}, \qquad \hat{c}_{ML} = X_{(1)}.$$

It is shown by Lehman [9] that the improved Uniform Minimum Variance Unbiased Estimators (UMVUE) for parameters $a$ and $c$ are

$$\hat{a}_{UMVU} = \frac{n-2}{n} \cdot \hat{a}_{ML}, \qquad \hat{c}_{UMVU} = X_{(1)}\left[1 - \frac{1}{(n-1)\cdot \hat{a}_{ML}}\right].$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

# CHAPTER 3

## INADMISSIBILITY OF THE MLE AND UMVUE OF *a*

### Criteria for Choosing a Decision Rule

In this section, it will be shown that the MLE and UMVUE estimates for parameter $a$ are inadmissible.

Firstly, a decision rule, denoted $\delta(\mathbf{x})$, is a function from our sample space into our action space. This action space is the set of all possible actions.

One criteria for choosing a decision rule d is minimizing the risk function $R(\delta,\theta)$ of a decision rule $\delta(\mathbf{x})$ where $R(\delta,\theta)$ is the expected loss with respect to our standard definition of an average over the random variable X

$$R(\delta,\theta) = E_\theta^X[L(\delta(X),\theta)] = \int_{\forall x} L(\delta(x),\theta)dF^X(x|\theta).$$

Berger [3] defines a decision rule $\delta_1$ R-better than a decision rule $\delta_2$ if $R(\delta_1,\theta) \le R(\delta_2,\theta)$ for all $\theta \in \Theta$, with strict inequality for some $\theta$.

When using the squared error loss function $L(\delta(X),\theta) = (\delta(X)-\theta)^2$, the risk function is defined as:

$$R(\delta,\theta) = E_\theta^X[L(\delta(X),\theta)] = E_\theta^X(\delta(X)-\theta)^2 \qquad (6)$$

Squared error loss is popular for several reasons including its connection with classical least squares estimation and the relative ease of calculation when performing decision investigations.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

**Decision Rule for Parameter *a***

Specifically, we are trying to minimize the risk of parameter $a$. Denote $\delta_\alpha(a) = \alpha\hat{a}$ as the decision rule for parameter $a$. Note that for $\hat{a}_{ML}$ and $\hat{a}_{UMVU}$, a=1 and a=(n-2)/n, respectively. We are trying to find the value of $\alpha$ such that for a decision rule of the form $\delta_\alpha(a) = \alpha\hat{a}$, the risk is minimized.

From (6), the risk function for decision rule $\delta_\alpha(a) = \alpha\hat{a}$ is clearly

$$R(\delta_\alpha, a) = E(\delta_\alpha - a)^2 = E(\alpha\hat{a} - a)^2 = \int_{\forall y} (\alpha\hat{a} - a)^2 f_Y(y)\,dy$$

where the probability density function of $\hat{a}$ was previously given as

$$f_Y(y) = \frac{a^{n-1} n^{n-1}}{\Gamma(n-1) y^n} \cdot \exp\left(-\frac{na}{y}\right) \qquad y \geq 0.$$

Let $\delta_\alpha = \alpha\hat{a} = \alpha y$ so that

$$R(\delta_\alpha, a) = E(\alpha y - a)^2$$

$$= \int_{\forall y} (\alpha y - a)^2 f_Y(y)\,dy$$

$$= \int_0^\infty (\alpha y - a)^2 \cdot \frac{(na)^{n-1} \exp(-na/y)}{\Gamma(n-1) y^n} dy.$$

Expansion of $(\alpha y - a)^2$ then leads to

$$R(\delta_\alpha, a) = \frac{(na)^{n-1}}{(n-2)!} \int_0^\infty \left(\alpha^2 y^2 - 2\alpha a y + a^2\right) e^{-na/y} y^{-n}\,dy$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$=\frac{(na)^{n-1}}{(n-2)!}\left[\alpha^2\int_0^\infty y^{-n+2}e^{-\frac{na}{y}}dy-2\alpha a\int_0^\infty y^{-n+1}e^{-\frac{na}{y}}dy+a^2\int_0^\infty y^{-n}e^{-\frac{na}{y}}dy\right]. \qquad (7)$$

Let $u=\frac{na}{y}$ *so that* $y=\frac{na}{u}$, $du=\frac{-na}{y^2}dy$, *and* $dy=\frac{-y^2}{na}du$. Equation (7) then becomes

$$=\frac{(na)^{n-1}}{(n-2)!}\left[\alpha^2\int_0^\infty y^{-n+2}e^{-u}(\tfrac{y^2}{na})du-2\alpha a\int_0^\infty y^{-n+1}e^{-u}(\tfrac{y^2}{na})du+a^2\int_0^\infty y^{-n}e^{-u}(\tfrac{y^2}{na})du\right]$$

$$=\frac{(na)^{n-1}}{(n-2)!}\left[\frac{\alpha^2}{na}\int_0^\infty y^{-n+4}e^{-u}du-\frac{2\alpha a}{na}\int_0^\infty y^{-n+3}e^{-u}du+\frac{a^2}{na}\int_0^\infty y^{-n+2}e^{-u}du\right]$$

$$=\frac{(na)^{n-1}}{(n-2)!}\left[\frac{\alpha^2}{na}\int_0^\infty (\tfrac{na}{u})^{-n+4}e^{-u}du-\frac{2\alpha a}{na}\int_0^\infty (\tfrac{na}{u})^{-n+3}e^{-u}du+\frac{a^2}{na}\int_0^\infty (\tfrac{na}{u})^{-n+2}e^{-u}du\right]$$

$$=\frac{(na)^{n-1}}{(n-2)!}\left[\alpha^2(na)^{-n+3}\int_0^\infty u^{n-4}e^{-u}du-2\alpha a(na)^{-n+2}\int_0^\infty u^{n-3}e^{-u}du+a^2(na)^{-n+1}\int_0^\infty u^{n-2}e^{-u}du\right]$$

$$=\frac{(na)^{n-1}}{(n-2)!}\left[\alpha^2(na)^{-n+3}\int_0^\infty u^{(n-3)-1}e^{-u}du-2\alpha a(na)^{-n+2}\int_0^\infty u^{(n-2)-1}e^{-u}du+a^2(na)^{-n+1}\int_0^\infty u^{(n-1)-1}e^{-u}du\right]$$

Substitution of the Gamma Integral $\Gamma(\kappa)=\int_0^\infty h^{\kappa-1}e^{-h}dh$ yields

$$=\frac{(na)^{n-1}}{(n-2)!}\left[\alpha^2(na)^{-n+3}\Gamma(n-3)-2\alpha a(na)^{-n+2}\Gamma(n-2)+a^2(na)^{-n+1}\Gamma(n-1)\right]. \qquad (8)$$

Since $\Gamma(\kappa)=(\kappa-1)!$ for any positive integer $\kappa>1$, (8) becomes

$$=\frac{(na)^{n-1}}{(n-2)!}\left[\alpha^2(na)^{-n+3}(n-4)!-2\alpha a(na)^{-n+2}(n-3)!+a^2(na)^{-n+1}(n-2)!\right]$$

$$=\frac{(an)^{n-1}}{(n-2)(n-3)}\left[\alpha^2(na)^{-n+3}-2\alpha a(na)^{-n+2}(n-3)+a^2(na)^{-n+1}(n-2)(n-3)\right]$$

which leads to a quadratic polynomial in terms of $\alpha$. Specifically

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$R(\delta_\alpha, a) = \frac{a^2}{(n-3)(n-2)}\left[n^2\alpha^2 - 2n(n-3)\alpha + (n-2)(n-3)\right]. \tag{9}$$

Now, minimize $R(\delta_\alpha, a)$ by differentiating $R(\delta_\alpha, a)$ with respect to $\alpha$,

$$\frac{\partial R(\delta_\alpha, a)}{\partial \alpha} = \frac{a^2}{(n-3)(n-2)}\left[2n^2\alpha - 2n(n-3)\right],$$

setting the result equal to zero and solving for $\alpha$. This process reveals $\alpha^* = \frac{n-3}{n}$ as the optimal choice for $\alpha$ so as to minimize $R(\delta_\alpha, a)$.

A question we can raise is the admissibility of estimators or decision rules. A decision rule $\delta$ is inadmissible if there exists another decision rule $\delta'$ such that $\delta'$ improves $\delta$. From this it is clear that the 'best' estimator within the group of estimators $\alpha y$ is

$$\delta_{\alpha^*}(a) = \left(\frac{n-3}{n}\right)\hat{a} = \left(\frac{n-3}{n}\right)y. \tag{10}$$

For further justification of this fact, substitution of $\alpha = 1$, $\alpha = \frac{n-2}{n}$ and $\alpha^* = \frac{n-3}{n}$ separately into (9) reveals:

$$R(\delta_{\alpha=1}, a) = \frac{a^2}{(n-3)(n-2)}\left[n^2 - 2n(n-3) + (n-2)(n-3)\right] = \frac{a^2(n+6)}{(n-3)(n-2)}$$

$$R(\delta_{\alpha=\frac{n-2}{n}}, a) = \frac{a^2}{(n-3)(n-2)}\left[\frac{n^2(n^2-4n+4)}{n^2} - 2n(n-3)\frac{n-2}{n} + (n-2)(n-3)\right] = \frac{a^2}{n-3}$$

$$R(\delta_{\alpha=\frac{n-3}{n}}, a) = \frac{a^2}{(n-3)(n-2)}\left[\frac{n^2(n^2-6n+9)}{n^2} - 2n(n-3)\frac{(n-3)}{n} + (n-2)(n-3)\right] = \frac{a^2}{n-2}$$

It is then clear that for all $a$,

$$R(\delta_{\alpha=1}, a) = \frac{a^2(n+6)}{(n-3)(n-2)} \quad > \quad R(\delta_{\alpha=\frac{n-2}{n}}, a) = \frac{a^2}{n-3} \quad > \quad R(\delta_{\alpha^*=\frac{n-3}{n}}, a) = \frac{a^2}{n-2}. \tag{11}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It has been determined that the 'best' estimator of the class of estimators $a\hat{a}$ is

$\delta_{a^*}(a) = \left(\frac{n-3}{n}\right)\hat{a} = \left(\frac{n-3}{n}\right)y$. From (11) we have also shown that $\hat{a}_{UMVUE}$ is a 'better' estimate than $\hat{a}_{ML}$. With respect to admissibility, (11) also shows that $\hat{a}_{ML}$ and $\hat{a}_{UMVUE}$ are inadmissible. The admissibility or inadmissibility of the improved estimate given in (10) is unknown.

The corresponding risk functions were graphed versus $a$ for sample sizes $n$=5, $n$=10, $n$=15, and $n$=30.

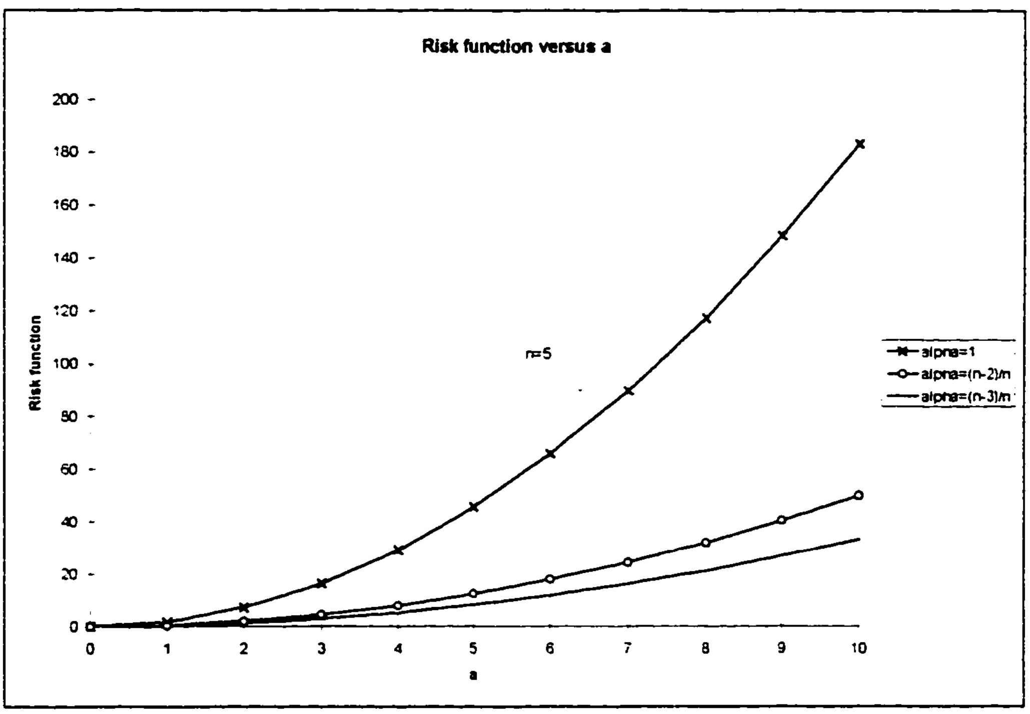


**Figure 4 Risk Functions Graphed Versus a for the MLE, UMVUE, and Improved Estimate with Sample Size n=5**

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

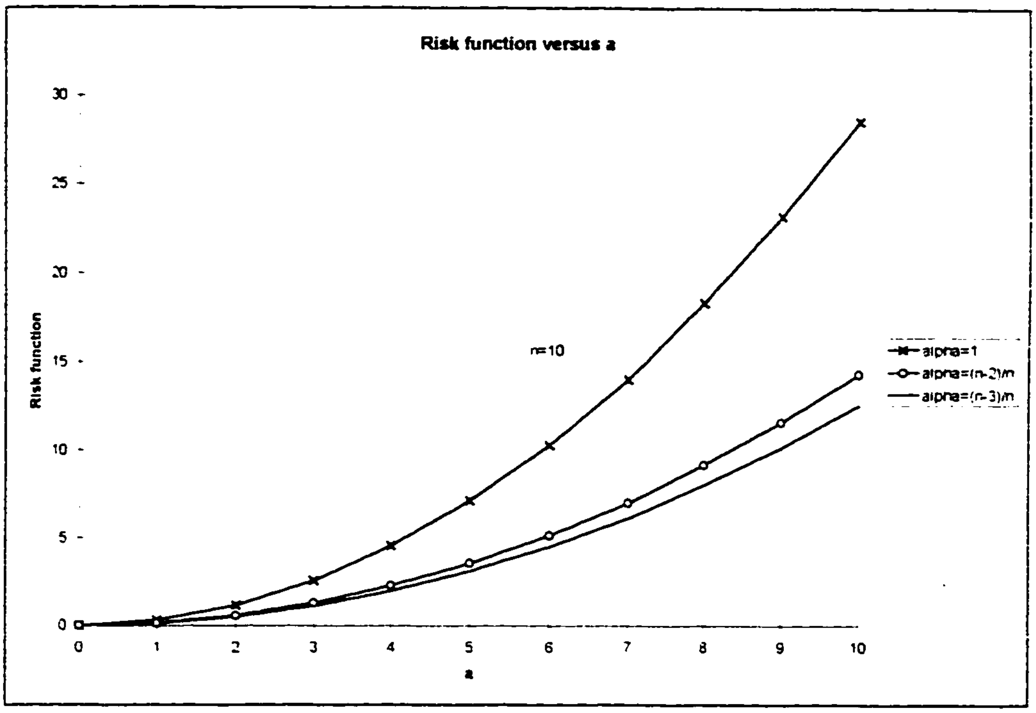


**Figure 5 Risk Functions Graphed Versus a for the MLE, UMVUE, and Improved Estimate with Sample Size =10**

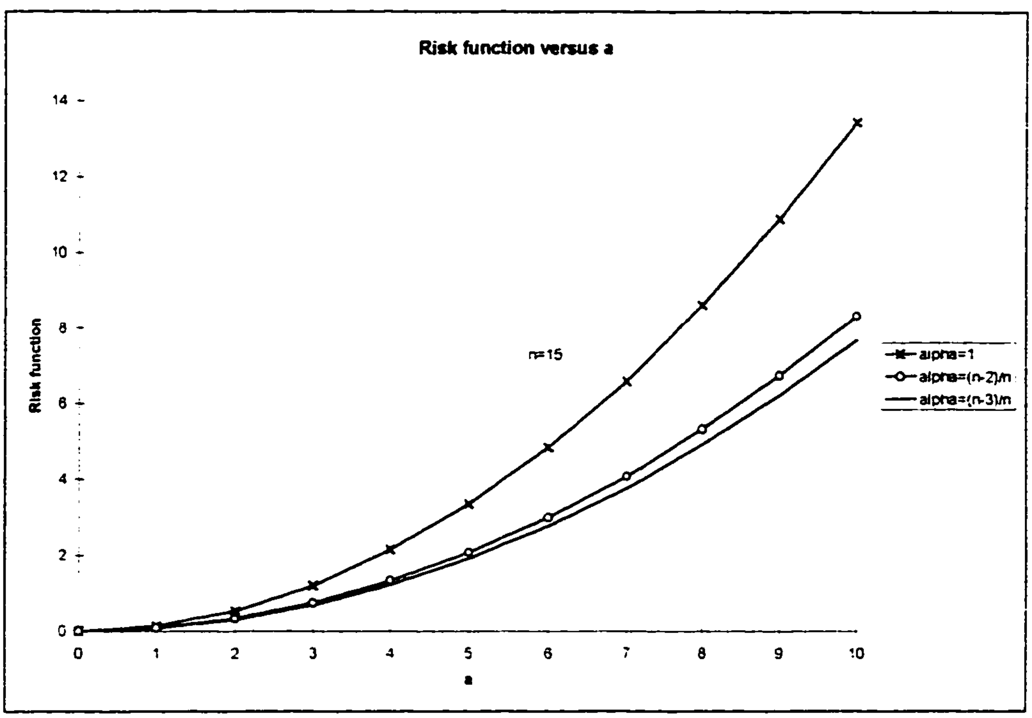


**Figure 6 Risk Functions Graphed Versus a for the MLE, UMVUE, and Improved Estimate with Sample Size n=15**

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

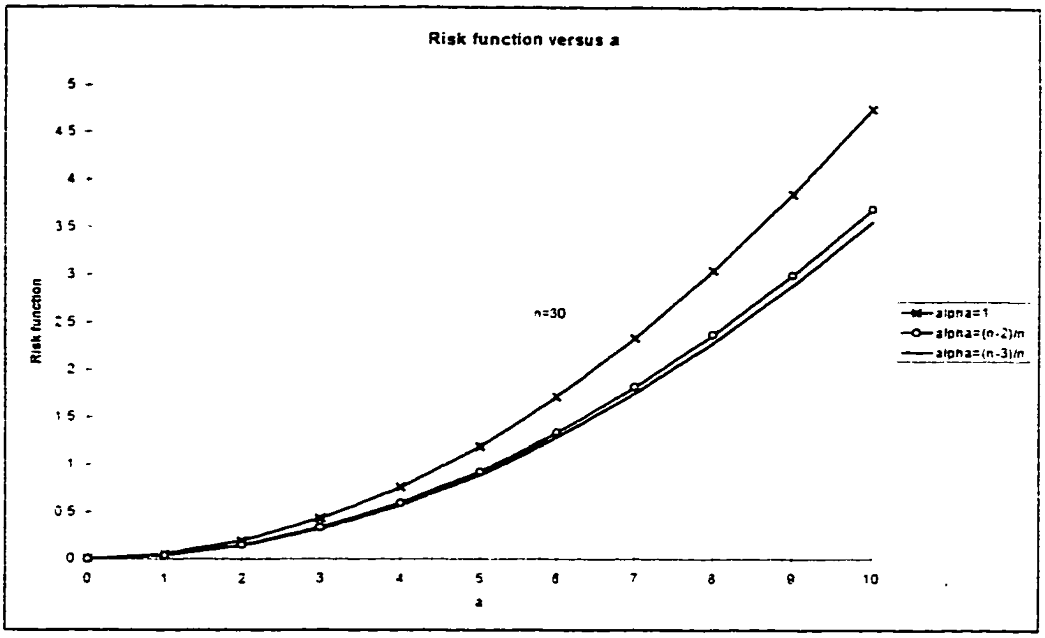


**Figure 7 Risk Functions Graphed Versus a for the MLE, UMVUE, and Improved Estimate with Sample Size n=30**

Now consider the percent risk improvement of the improved estimate $\delta_{\alpha^*}(a) = \left(\frac{n-3}{n}\right)\hat{a}$ over the maximum likelihood estimate $\hat{a}_{ML}$ and uniform minimum variance unbiased estimate $\hat{a}_{UMVU}$. The percent risk improvement $\frac{(R(\delta,\hat{a}) - R(\delta,a))}{R(\delta,a)} \times 100$ is found for varied sample sizes $n$=5, $n$=10, $n$=15, and $n$=30.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE 1 Percent Risk Improvement of the Improved Estimate Over the MLE and UMVUE

| Sample Size | a | percent risk improvement $= \frac{(R(\delta,\hat{a}) - R(\delta,a))}{R(\delta,a)} \times 100$ | |
|---|---|---|---|
| | | For $\delta = \delta_{MLE}$ | For $\delta = \delta_{UMVUE}$ |
| n = 5 | 1 | 81.81% | 33.33% |
| | 5 | 81.81% | 33.33% |
| | 10 | 81.81% | 33.33% |
| | 15 | 81.81% | 33.33% |
| n = 10 | 1 | 56.25% | 12.5% |
| | 5 | 56.25% | 12.5% |
| | 10 | 56.25% | 12.5% |
| | 15 | 56.25% | 12.5% |
| n = 15 | 1 | 42.85% | 7.69% |
| | 5 | 42.85% | 7.69% |
| | 10 | 42.85% | 7.69% |
| | 15 | 42.85% | 7.69% |
| n = 30 | 1 | 25% | 3.57% |
| | 5 | 25% | 3.57% |
| | 10 | 25% | 3.57% |
| | 15 | 25% | 3.57% |

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

**Simulation Experiment**

Simulations were conducted in which random data was taken from pareto distributions with parameters $a$=5 and $a$=10, $c$=1 and $c$=2. For samples of size $n$=5, $n$=10, and $n$=15, the MLE, UMVUE, and improved estimates of $a$ were found.

TABLE 2 Estimation Results for Random Data Taken From the Pareto Distribution

| Parameter Values | | Random Data | MLE of $a$ | UMVUE of $a$ | Improved $a$ |
|---|---|---|---|---|---|
| $a=5$ | $c=1$ | 1.25, 1.12, 1.36, 1.18, 1.15 | 13.2 | 7.92 | 5.28 |
| $a=5$ | $c=2$ | 2.05, 2.26, 2.25, 2.30, 2.90 | 7.60 | 4.56 | 3.04 |
| $a=5$ | $c=1$ | 1.36, 1.03, 1.02, 1.18, 1.16<br>1.11, 1.21, 1.11, 1.01, 1.03 | 10.2 | 8.16 | 7.14 |
| $a=5$ | $c=2$ | 2.03, 2.11, 2.34, 2.02, 2.05<br>2.04, 2.15, 2.10, 2.09, 2.30 | 21.6 | 17.3 | 15.1 |
| $a=5$ | $c=1$ | 1.65, 1.04, 2.92, 1.08, 1.90<br>1.19, 1.90, 1.61, 1.15, 1.04<br>1.12, 1.01, 2.96, 1.18, 1.13 | 2.93 | 2.54 | 2.34 |
| $a=5$ | $c=2$ | 2.31, 2.03, 2.03, 2.07, 2.10<br>2.59, 2.18, 2.59, 2.47, 2.02<br>2.76, 2.18, 2.91, 2.56, 2.04 | 7.48 | 6.48 | 5.98 |
| $a=10$ | $c=1$ | 1.05, 1.07, 1.04, 1.15, 1.30 | 13.6 | 8.16 | 5.44 |

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

| | | | | | |
|---|---|---|---|---|---|
| $a = 10$ | $c = 2$ | 2.06, 2.05, 2.03, 2.39, 2.04 | 26.1 | 15.7 | 10.4 |
| $a = 10$ | $c = 1$ | 1.01, 1.06, 1.05, 1.09, 1.05<br>1.19, 1.00, 1.05, 1.06, 1.07 | 16.7 | 13.4 | 11.7 |
| $a = 10$ | $c = 2$ | 2.23, 2.47, 2.08, 2.02, 2.10<br>2.04, 2.67, 2.02, 2.15, 2.04 | 13.8 | 11.0 | 9.66 |
| $a = 10$ | $c = 1$ | 1.04, 1.03, 1.03, 1.10, 1.01<br>1.01, 1.15, 1.16, 1.05, 1.03<br>1.25, 1.03, 1.00, 1.04, 1.08 | 15.5 | 13.4 | 12.4 |
| $a = 10$ | $c = 2$ | 2.10, 2.25, 2.01, 2.31, 2.15<br>2.16, 2.12, 2.34, 2.40, 2.27<br>2.01, 2.19, 2.01, 2.09, 2.17 | 12.8 | 11.1 | 10.2 |

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

# CHAPTER 4

## A GENERALIZED BAYES ESTIMATE FOR PARAMETER *a*

### Introduction

We will begin the discussion with an introduction of the relevant concepts and definitions. The generalized bayes estimate for $a$ will then be found.

As in the previous case, assume that both the parameters $a$ and $c$ are unknown. Use of the marginal posterior distribution of $a$ given x will be shown to be useful in the inference of the shape parameter $a$.

Bayesian analysis, named after the English philosopher Thomas Bayes who in 1763 discovered the discrete result in probability theory, combines prior information and sample information into the construction of a posterior distribution. This posterior distribution is the tool from which we develop all inference, in this case a point estimate of the parameter $a$. The posterior distribution of $\theta$ given x, commonly denoted as $\pi(\theta|\mathbf{x})$, is known as the distribution of $\theta$ given the sample observations $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Note that at times '$\theta$' will represent a vector of parameters, while other times it may represent a single parameter.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

**Definitions**

The conditional density of random variable X, given the continuous random variable $\theta$ with value $\theta$, is defined as $f(x|\theta) = \dfrac{f(x,\theta)}{f(\theta)}$. The joint density of $\theta$ and X is given by

$$f(x,\theta) = f(\theta) f(\mathbf{x}|\theta). \qquad (12)$$

Denote the density $f(\theta)$ as $\pi(\theta)$ so that $f(x,\theta) = \pi(\theta) f(\mathbf{x}|\theta)$. Here $\pi(\theta)$ represents the 'prior' distribution while $f(\mathbf{x}|\theta)$ is the likelihood function. A large value of $f(\mathbf{x}|\theta)$ for a value $\theta$ is considered to be the basis in reasoning that the particular $\theta$ is more 'likely' than any other $\theta$. For this reason $f(\mathbf{x}|\theta)$ is naturally called the likelihood function. Consider the marginal density of X:

$$m(x) = \int_{\forall\theta} f(\mathbf{x}|\theta) dF^{\pi}(\theta) = \int_{\forall\theta} f(\mathbf{x}|\theta)\pi(\theta) d\theta$$

in the continuous case. The construction of the posterior distribution, using our definitions previously developed, is then

$$\pi(\theta|\mathbf{x}) = \frac{h(x,\theta)}{m(x)}. \qquad (13)$$

This posterior then incorporates and combines the prior beliefs about $\theta$ with the current information about $\theta$ obtained from the sample **x**.

From (12) and (13) we can then state

$$\pi(\theta|\mathbf{x}) = \frac{h(x,\theta)}{m(x)} = \frac{\pi(\theta) f(\mathbf{x}|\theta)}{m(x)}. \qquad (14)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

All inference of $\theta$ will be based on this posterior density $\pi(\theta|\mathbf{x})$. It is noted that the posterior density $\pi(\theta|\mathbf{x})$ can be shown, in certain cases or groups of priors, to be proportional to the numerator of (14).

### Construction of a Posterior Density of *a*

Consider the joint prior $\pi(a, c)$ as the product of independent priors $\pi(a)$ and $\pi(c)$. Let independent priors $\pi(a)$ and $\pi(c)$ for $a$ and $c$ be of the form

$$\pi(a) = \frac{1}{a^h} \quad a > 0$$

and

$$\pi(c) = \frac{1}{c^r} \quad c > 0$$

so that the joint prior is of the form

$$\pi(a, c) = \pi(a) \cdot \pi(c) = \frac{1}{a^h c^r} \quad a > 0, c > 0 .$$

The posterior for this given prior can now be found using the known likelihood function and the concept of proportionality.

The posterior joint density $\pi(a, c|\mathbf{x})$, as defined by (14), is proportional to the product of the joint prior distribution of $a$ and $c$ and the likelihood function for the pareto distribution.

$$\pi(a, c|\mathbf{x}) \propto \pi(a, c) \cdot f(\mathbf{x}|a, c) = \pi(a, c) \cdot L(a, c)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\pi(a, c|\mathbf{x}) \propto \frac{1}{a^h c^r} \cdot \frac{a^n c^{na}}{\left(\prod_{i=1}^{n} x_i\right)^{a+1}}$$

$$\pi(a, c|\mathbf{x}) \propto a^{n-h} c^{na-r} \exp[-a\sum_{i=1}^{n} \ln x_i] \quad a > 0, 0 < c < x_{(1)}$$

The joint posterior can now be used to find the marginal posterior density for $a$.

$$\pi(a|\mathbf{x}) = \int_{\forall c} \pi(a, c|\mathbf{x}) dc$$

$$\pi(a|\mathbf{x}) \propto \int_0^{x_{(1)}} a^{n-h} c^{na-r} \exp[-a\sum_{i=1}^{n} \ln x_i] dc$$

$$\pi(a|\mathbf{x}) \propto a^{n-h} \exp[-a\sum_{i=1}^{n} \ln x_i] \cdot \int_0^{x_{(1)}} c^{na-r} dc$$

$$\pi(a|\mathbf{x}) \propto a^{n-h} \exp[-a\sum_{i=1}^{n} \ln x_i] \cdot \left. \frac{c^{na-r+1}}{na-r+1} \right|_0^{x_{(1)}}$$

$$\pi(a|\mathbf{x}) \propto a^{n-h} \exp[-a\sum_{i=1}^{n} \ln x_i] \cdot \frac{x_{(1)}^{na-r+1}}{na-r+1} \qquad a > 0$$

Now, consider the case where $r = 1$ so that the marginal density of $a$ is

$$\pi(a|\mathbf{x}) \propto a^{n-h} \exp[-a\sum_{i=1}^{n} \ln x_i] \cdot \frac{x_{(1)}^{na}}{na}$$

$$\pi(a|\mathbf{x}) \propto a^{n-h-1} \exp[-a(\sum_{i=1}^{n} \ln x_i - n \ln x_{(1)})] \tag{15}$$

Now consider the exact posterior density of $a$ so that the posterior mean can be considered. It can be shown that the constant

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{(\sum_{i=1}^{n} \ln x_i - n \ln x_{(1)})^{n-h}}{\Gamma(n-h)}$$

can be applied to the proportional posterior density of $a$ (15), so that the density is proper. The posterior density of $a$ is then

$$\pi(a|\mathbf{x}) = \frac{a^{n-h-1} \exp[-a(\sum_{i=1}^{n} \ln x_i - n \ln x_{(1)})]}{\Gamma(n-h)(\sum_{i=1}^{n} \ln x_i - n \ln x_{(1)})^{-(n-h)}} \quad a > 0.$$

**Estimation of Parameter *a***

Given the posterior density $\pi(\theta|\mathbf{x}) = \pi(a|\mathbf{x})$, there are several approaches one may use in problems of inference. Several involve the use of classical estimation techniques, such as finding the M.L.E. of parameter $a$ with respect to the posterior density. Other Bayesian point estimates include the posterior mean $\mu^{\pi}(x) = E^{\pi(\theta|\mathbf{x})}[\theta]$ or posterior median. For both of the previous methods, it is common to include a measure of accuracy. A typical measure of accuracy is the posterior variance $V^{\pi}(x) = E^{\pi(\theta|\mathbf{x})}[(\theta - \mu^{\pi}(x))^2]$ of the estimate. Another approach is to construct a "credible set" for $\theta$. This credible set is the Bayesian equivalent of the frequentist confidence set or confidence interval. Hypothesis testing is another option. In order to maintain consistency with our previous analysis in presenting an improved estimate for parameter $a$ given that $c$ is known, we will use the Bayesian approach to decision theory.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The Bayes risk of a decision rule $\delta$, with respect to a prior distribution $\pi$ on $\theta$, is defined as

$$r(\delta, \pi) = E^{\pi}[R(\delta, \theta)].$$

The Bayes rule $\delta^{\pi}$ is then a rule which minimizes the Bayes risk $r(\delta, \pi)$. It is also known that a Bayes rule $\delta^{\pi}$ is an action which ,under certain conditions, minimizes the posterior expected loss. The posterior expected loss of an action a is

$$\rho(a, \pi(\theta|\mathbf{x})) = \int_{\forall\theta} L(a, \theta) dF^{\pi(\theta|\mathbf{x})}(\theta)$$

where $L(a, \theta)$ is the loss function and $\pi(\theta|\mathbf{x})$ is the posterior density of $\theta$. A posterior Bayes action, further denoted $\delta^{\pi}(\mathbf{x})$, is an action a which minimizes $\rho(a, \pi(\theta|\mathbf{x}))$. To be consistent with our earlier frequentist decision theoretic analysis, "squared error" (ie quadratic) loss $L(a, \theta) = (a - \theta)^2$ will be used and the resulting posterior expected loss is

$$\rho(a, \pi(\theta|\mathbf{x})) = \int_{\forall\theta} (a - \theta)^2 dF^{\pi(\theta|\mathbf{x})}(\theta).$$

It is a trivial result then to show that the value of $a$ that minimizes the posterior expected loss, also known as the Bayes rule $\delta^{\pi}(\mathbf{x})$, is the mean of the posterior distribution of $\theta$ given x, $\delta^{\pi}(\mathbf{x}) = E^{\pi(\theta|\mathbf{x})}[\theta]$.

Since the posterior density of $a$ is in the form of a Gamma distribution, the posterior mean is easily found.

$$E^{\pi(a|\mathbf{x})}(A) = \frac{n-h}{\left(\sum_{i=1}^{n} \ln x_i - n \ln x_{(1)}\right)}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$= \frac{n-h}{\sum_{i=1}^{n} (\ln x_i - \ln x_{(1)})}$$

$$= \frac{n-h}{\sum_{i=1}^{n} \ln\left( \frac{x_i}{x_{(1)}} \right)}$$

$$= \frac{n-h}{\frac{n}{y}} = \frac{n-h}{n} \cdot y$$

where $y = \frac{n}{\sum_{i=1}^{n} \ln\left( \frac{x_i}{x_{(1)}} \right)}$. Note that setting $h = 0$ and $h = 2$, the generalized bayes estimate for $a$ is the MLE and UMVUE of $a$, respectively. Similarly, note that when $h = 3$, the generalized bayes estimate for $a$ becomes the improved estimate for $a$, $\frac{n-3}{n} y$.

It has been shown that use of the priors $\pi(c) = \frac{1}{c}$ and $\pi(a) = \frac{1}{a^3}$ will lead to the generalized bayes estimator $\frac{n-3}{n} y$. This is equivalent to the improved estimator of $a$ found in chapter 3. It may also be possible to use the above priors in establishing the admissibility of the improved estimator.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

# CHAPTER 5

# THE LORENZ CURVE

## History and Definition

The Gini coefficient is known as the most notable summary statistic of the Lorenz Curve of the income distribution. The primary distribution of interest will be the Pareto distribution. The Lorenz Curve and the corresponding Gini coefficients will be found for the Pareto distribution.

The century began with an increased interest in the inequality, and for some, the eventual equality of income and the wealth of individuals. M. O. Lorenz is generally known to be the first to propose a method in which to evaluate one income distribution versus one or more other distributions. Lorenz [10] first introduced the "Lorenz Curve" in 1905. The Lorenz Curve is used to measure the income shares of groups in a population. It is useful in determining a time trend for a certain geographic location or comparing income distributions between different regions. The Lorenz Curve, further denoted as L(p), is found by first ordering the income units by magnitude, from the smallest to the largest. L(p) is then the function that determines the proportion of total income that the lowest 100p% members of the population obtain. It is then clear that $0 \leq p \leq 1$ and $0 \leq L(p) \leq 1$. In other words, the cumulative proportion of the population, as determined by the particular income distribution, is plotted against the cumulative proportion of total income obtained by the ordered income units. If all incomes were



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

equal, the resulting Lorenz Curve would be the perfect 45 degree line running from zero to one. Actual Lorenz Curves typically are convex and lie under the above described "line of perfect equality", increasing in slope as the proportion of the total population increases. Lorenz states "With an unequal distribution, the curves always begin and end in the same points as with an equal distribution, but they will be bent in the middle; and the rule of interpretation will be, as the bow is bent, concentration increases." Here, an equal distribution is obviously the equivalent of a uniform distribution. This makes sense in that after arranging the population in order of income from smallest to largest, the individuals earning the upper ended incomes will earn more of the total income than the individuals in the lower end of the ordering. Lorenz's idea can be represented in the form of an inequality as stated by Marshall and Olkin [12]. "Letting $x_1, \ldots, x_n$ represent the wealth of individuals for the distribution of total wealth T that leads to (the curve or line of perfect equality). Similarly let $y_1, .., y_n$ lead to (a curve not equal to, under the line of perfect equality). Then $(x_1, \ldots, x_n)$ represents a more even distribution of wealth than does $(y_1, \ldots, y_n)$ if and only if

$$\sum_1^k x_{(i)} \geq \sum_1^k y_{(i)} \qquad k = 1, \ldots, n-1$$

where $\sum_1^n x_{(i)} = \sum_1^n y_{(i)} = T$."

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

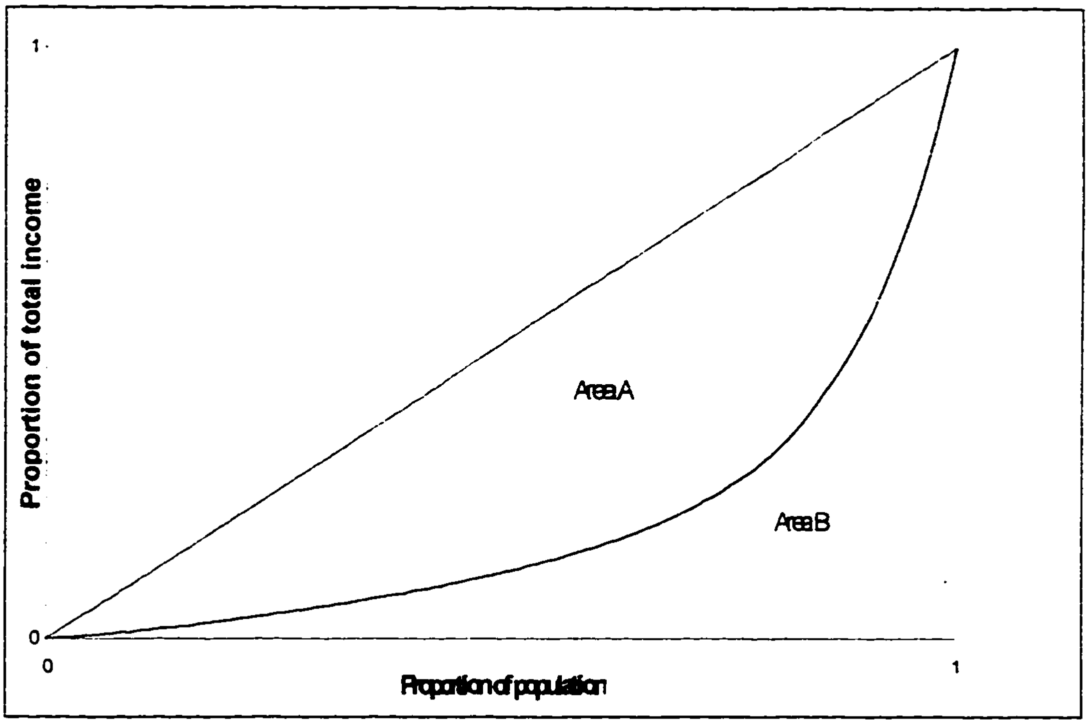


**Figure 8 Lorenz Curve**

Note that income is traditionally lognormally distributed with the upper tail being pareto distributed. Determining the most accurate distribution of income is itself an important topic. A great deal of material has been written on the subject; reference Harrison [5] and Harrison [6]. As the primary purpose of this paper is not to argue the most accurate income\earnings distribution, the subject is left to the reader for further analysis.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

## The Lorenz Curve for Pareto Populations

Given a specific distribution, the resulting Lorenz curve can be found. Mathematically, Lambert [8] defines the Lorenz curve L(p) by

$$p = F(y) \Rightarrow L(p) = \int_0^y \frac{xf(x)dx}{\mu} \quad 0 < p < 1 \qquad (16)$$

Equation (16) is a result of the facts that for each p within the interval (0,1) there is a single level of income y given a rank p. This satisfies p=F(y).

Let us look at the derivation of the Lorenz curve for the Pareto distribution. The density function of the Pareto distribution with location parameter $c$ and shape parameter $a$ $(c > 0 \, and \, a > 0)$ is given by

$$f(x) = \begin{cases} 0 & x < c \\ ac^a x^{-a-1} & x \geq c \end{cases}$$

with the minimum income level $c$. The resulting distribution function is given by

$$F(x) = \int_c^x f(t)dt = \begin{cases} 0 & x < c \\ 1 - (c/x)^a & x \geq c \end{cases} \qquad (17)$$

with mean $\mu = \dfrac{ac}{a-1}$.

Using equation (16):

$$p = F(y) = 1 - c^a y^{-a}$$

and thus

$$y = \frac{c}{(1-p)^{-a}}. \qquad (18)$$

Using equation (16), (17) and result (18):

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$L(p)=\int_c^y x\cdot ac^a x^{-a-1}\cdot\frac{(a-1)}{ac}dx=\frac{(a-1)c^a}{c}\cdot\frac{x^{-a+1}}{-a+1}\Bigg|_c^y$$

$$=c^{a-1}c^{-a+1}-c^{a-1}y^{-a+1}=1-c^{a-1}\left(\frac{c}{(1-p)^{-a}}\right)^{-a+1}$$

$$=1-(1-p)^{\frac{(a-1)}{a}} \qquad (19)$$

The Lorenz curve for the Pareto distribution is typically given in the form of result (19).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

# CHAPTER 6

# THE GINI COEFFICIENT

## History and Definition

The Gini coefficient, first proposed by C. Gini in 1912, is an area measure of the magnitude in which a particular Lorenz curve deviates from the 45 degree "line of perfect equality". Specifically, the Gini coefficient is the area between the Lorenz curve and the 45 degree line divided by the total area under the 45 degree line. This is equivalent to twice the area between the Lorenz curve and the 45 degree line.

$$G = \frac{A}{A+B} = 2A = 2\left(\frac{1}{2} - B\right) = 1 - 2B$$

From this definition we know that G can run anywhere from zero to one-half, i.e. the total area under the 45 degree line. One important use of the Gini coefficient is the ability to rank different income distributions in order to describe their inequality.

## The Gini Coefficient for Pareto Populations

Using result (19) and our knowledge of the Gini coefficient, the Gini coefficient for the pareto distribution can be found:

$$G = 1 - 2\int_0^1 L(p)dp = 1 - 2\int_0^1 1 - (1-p)^{(a-1)/a} dp.$$



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Let $u = 1 - p$ and $du = -dp$ so that

$$G = 1 + 2\int_{1}^{0} 1 - u^{(a-1)/a}\, du = 1 + 2\int_{1}^{0} 1 - u^{1-1/a}\, du$$

$$= 1 + 2\left(u - \frac{u^{2-1/a}}{2 - 1/a}\right)\Bigg|_{1}^{0} = 1 + 2\left[0 - \left(1 - \frac{1}{2 - 1/a}\right)\right]$$

$$= 1 + 2\left(-1 + \frac{a}{2a-1}\right)$$

$$= \frac{1}{2a-1}$$

## Estimation of the Gini Coefficient

It has been shown that the Gini coefficient $G$ for the pareto distribution is $G = \frac{1}{2a-1}$. Given the earlier development of the maximum likelihood and improved estimators for parameter $a$, the MLE for $G$ and the generalized bayes estimator for $G$ will be found.

## Maximum Likelihood Estimate of Gini Coefficient

From the development of the maximum likelihood estimates for parameters $a$ and $c$,

$\hat{a}_{ML} = \dfrac{n}{\sum_{i=1}^{n} \ln\left(X_i / X_{(1)}\right)}$ and $\hat{c}_{ML} = X_{(1)}$, the exact maximum likelihood estimate of the Gini coefficient $G$ can be found. Bain and Engelhardt [2] define the Invariance Property with respect to more than one unknown parameter:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

If $\hat{\Theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k)$ denotes the MLE of $\Theta = (\theta_1, \ldots, \theta_k)$, then the MLE of

$T = (\tau_1(\Theta), \ldots, \tau_r(\Theta))$ is $\hat{T} = (\hat{\tau}_1, \ldots, \hat{\tau}_r) = (\tau_1(\hat{\Theta}), \ldots, \tau_r(\hat{\Theta}))$ for $1 \le r \le k$.

For the case of the one-to-one transformation Gini coefficient $G = \frac{1}{2a-1}$, we can construct the MLE of $G$ simply by substituting the MLE of $a$ for $a$ in G:

$$\hat{G}_{ML}(a) = G(\hat{a}_{ML}) = \frac{1}{2\hat{a}_{ML} - 1}.$$

## The Generalized Bayes Estimator for the Gini Coefficient

In chapter 4, the improved estimator for $a$ was found using the generalized bayes approach. The generalized bayes estimator for the Gini coefficient $G$ will now be constructed by use of priors densities $\pi(c) = \frac{1}{c}$ and $\pi(a) = \frac{1}{a^3}$. These priors will lead to the identical posterior $\pi(a|\mathbf{x})$ found in chapter 4.

The posterior density of $a$ was found, given the priors $\pi(a) = \frac{1}{a^h}$ and $\pi(c) = \frac{1}{c}$, to be

$$\pi(a|\mathbf{x}) = \frac{a^{n-h-1} \exp[-a(\sum_{i=1}^{n} \ln x_i - n \ln x_{(1)})]}{\Gamma(n-h)(\sum_{i=1}^{n} \ln x_i - n \ln x_{(1)})^{-(n-h)}} \qquad a > 0.$$

The posterior mean $E^{\pi(\theta|\mathbf{x})}[A]$ was then found to be the improved estimate when $h = 3$.

The posterior density of $a$ would then become

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\pi(a|\mathbf{x}) = \frac{a^{(n-3)-1}\exp[-a(\sum_{i=1}^{n}\ln x_i - n\ln x_{(1)})]}{\Gamma(n-3)(\sum_{i=1}^{n}\ln x_i - n\ln x_{(1)})^{-(n-3)}} \quad a > 0.$$

This is in the form of a gamma density where the first shape parameter is equal to $n$-$3$ and the second shape parameter is equal to $(\sum_{i=1}^{n}\ln x_i - n\ln x_{(1)})^{-1}$. It should be noted that we must limit the use of this posterior in inferences based on samples of size greater than or equal to 4.

The generalized bayes estimator for $G$ will now be found by finding the posterior mean of $G$ $E^{\pi(a|\mathbf{x})}[G]$ with respect to the posterior density $\pi(a|\mathbf{x})$. $\pi(a|\mathbf{x})$ can be used because the Gini coefficient for the pareto distribution is a function of $a$. We are then evaluating

$$E^{\pi(a|\mathbf{x})}[G] = E\left(\frac{1}{2a-1}\right) = \int_{\forall a}\frac{1}{2a-1}\cdot\pi(a|\mathbf{x})da.$$

This integral will be evaluated using a numerical technique in which twenty-five thousand random data values from a gamma distribution will be transformed using the known relationship $G = \frac{1}{2a-1}$. The mean of these transformed values will then be found to be an estimate of $E^{\pi(a|\mathbf{x})}[G]$.

Using the pareto random data from table 2, the MLE and generalized bayes estimates of $G$ were found and are presented in table 3.

Again note that the estimation of the Gini coefficient in this case is not true bayesian estimation. This is clear in that particular values of parameter $a$ are given in order to find estimates of $G$. This is not consistent with the fact that we are basing our estimation on

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the posterior density of parameter $a$. By definition this density is based on the random parameter $a$.

TABLE 3 Estimation Results for the Gini Coefficient for Random Data Taken From the Pareto Distribution

| Parameter Values | | Random Data | MLE of $G$ | Bayes $G$ | Actual $G$ |
|---|---|---|---|---|---|
| $a = 5$ | $c = 1$ | 1.25, 1.12, 1.36, 1.18, 1.15 | .039 | .253 | .111 |
| $a = 5$ | $c = 2$ | 2.05, 2.26, 2.25, 2.30, 2.90 | .070 | .691 | .111 |
| $a = 5$ | $c = 1$ | 1.36, 1.03, 1.02, 1.18, 1.16<br>1.11, 1.21, 1.11, 1.01, 1.03 | .052 | .091 | .111 |
| $a = 5$ | $c = 2$ | 2.03, 2.11, 2.34, 2.02, 2.05<br>2.04, 2.15, 2.10, 2.09, 2.30 | .024 | .040 | .111 |
| $a = 5$ | $c = 1$ | 1.65, 1.04, 2.92, 1.08, 1.90<br>1.19, 1.90, 1.61, 1.15, 1.04<br>1.12, 1.01, 2.96, 1.18, 1.13 | .205 | .318 | .111 |
| $a = 5$ | $c = 2$ | 2.31, 2.03, 2.03, 2.07, 2.10<br>2.59, 2.18, 2.59, 2.47, 2.02<br>2.76, 2.18, 2.91, 2.56, 2.04 | .072 | .102 | .111 |
| $a = 10$ | $c = 1$ | 1.05, 1.07, 1.04, 1.15, 1.30 | .038 | .303 | .053 |

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

| | | | | | |
|---|---|---|---|---|---|
| $a = 10$ | $c = 2$ | 2.06, 2.05, 2.03, 2.39, 2.04 | .020 | .082 | .053 |
| $a = 10$ | $c = 1$ | 1.01, 1.06, 1.05, 1.09, 1.05<br>1.19, 1.00, 1.05, 1.06, 1.07 | .031 | .053 | .053 |
| $a = 10$ | $c = 2$ | 2.23, 2.47, 2.08, 2.02, 2.10<br>2.04, 2.67, 2.02, 2.15, 2.04 | .038 | .065 | .053 |
| $a = 10$ | $c = 1$ | 1.04, 1.03, 1.03, 1.10, 1.01<br>1.01, 1.15, 1.16, 1.05, 1.03<br>1.25, 1.03, 1.00, 1.04, 1.08 | .033 | .043 | .053 |
| $a = 10$ | $c = 2$ | 2.10, 2.25, 2.01, 2.31, 2.15<br>2.16, 2.12, 2.34, 2.40, 2.27<br>2.01, 2.19, 2.01, 2.09, 2.17 | .041 | .056 | .053 |

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

# REFERENCES


[1] Arnold, B. C. and Press, S. J. (1983), Bayesian Inference for Pareto Populations, Journal of Econometrics, 21, 287-306.

[2] Bain, L. J. and Engelhardt, M. (1992), Introduction To Probability and Mathematical Statistics, Duxbury Press, Belmont, California.

[3] Berger, J. O. (1985), Statistical Decision Theory and Bayesian Analysis, Springer-Verlag, New York.

[4] Cirillo, R. (1979), The Economics of Vilfredo Pareto, Frank Cass and Company, London.

[5] Harrison, A. (1979), The Upper Tail of the Earnings Distribution: Pareto or Lognormal, Economics Letters 2, 191-195.

[6] Harrison, A. (1981), Earnings by Size: A Tale of Two Distributions, Review of Economic Studies, 621-631.

[7] Johnson, N. L. and Kotz, S. (1970), Continuous Univariate Distributions, Houghton Mifflin, Boston.

[8] Lambert, P. J. (1993), The Distribution and Redistribution of Income: A Mathematical Analysis, Manchester University Press, London.

[9] Lehmann, E. L. (1991), Theory of Point Estimation, Wadsworth and Brooks, Pacific Groove, California.

[10] Lorenz, M. O. (1905), Methods of Measuring the Concentration of Wealth, Publications of the American Statistical Association, New Series, 70, 209-219.

[11] Malik, H. J. (1970), Estimation of the Parameters of the Pareto Distribution, Metrika, 16, 126-132.

[12] Marshall, A. W. and Olkin, I. (1979), Inequalities: Theory of Majorization and Its Applications, Academic Press, New York.

[13] Zellner, A. (1971), An Introduction to Bayesian Inference in Econometrics, Wiley, New York.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.