STATISTICS OF QUOTATIONS REPORTED BY THE INSTITUTE FOR SCIENTIFIC INFORMATION (ISI). A WORKING EXAMPLE OF A CHILEAN INSTITUTION

R. Acevedo[(1)], G. Díaz[(2)] and P. Kittl[(3)]

[(1)] Facultad de Ingeniería. Universidad Mayor
Avenida Manuel Montt 367, Providencia, Santiago, Chile.
E-mail: roberto.acevedo@umayor.cl
[(2)] Departamento de Ciencia de los Materiales,
Facultad de Ciencias Físicas y Matemáticas,Universidad de Chile
P.O.Box 2777, Santiago, Chile. E-mail: gediaz@cec.uchile.cl
[(3)]Departamento de Ingeniería Mecánica,
Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile,
P.O.Box 2777, Correo 21, Santiago, Chile.

SUMMARY

Starting with a list of academicians of a Faculty of a reputed Chilean University, the number of times that their published articles have been quoted in the magazines included in the indexes of the Institute of Scientific Information (ISI) was assessed. The statistical study of the number of quotations of the mentioned academicians shows a J shaped probability distribution that corresponds to a reverse Pareto distribution. Although the fitting of the Pareto distribution parameters shows that the experimental distribution, although following Pareto with a certain degree of approximation, has a singular distribution that indicate the superposition of several statistical populations. The theoretical origin of the Pareto's α parameter and its connection with the origin of the different populations obtained from the sample used is discussed. The explanation for the lack of adaptability of the Pareto distribution - a distribution of second order – may be found in the fact that the obtaining of the quotations by ISI is not entirely based on capability as discussed in the text.

ISI CITATIONS

Table I shows the data on the citations made from the articles written by academicians of a University Faculty in Chile and published in magazines indexed in the registers of the Institute of Scientific Information, ISI [1], up to year 2002.

Table I. Quotations made by ISI of articles written by academicians of a University Faculty in Chile until year 2002.

0	3	10	23	49	93	169	466
0	3	11	24	49	96	172	511
0	4	11	24	52	99	176	604
0	4	12	24	52	101	182	672
0	4	12	25	55	104	185	746
0	4	12	27	55	108	188	772
0	5	12	27	58	113	189	815
0	5	12	28	64	114	215	846
0	5	13	29	64	115	217	1162
0	5	14	31	65	121	229	1284
0	6	14	31	66	124	238	2010
0	6	14	31	66	124	239	
0	6	15	31	69	127	253	
0	7	15	34	71	128	254	
0	7	16	36	71	129	261	
1	7	16	37	71	130	271	
1	7	16	40	75	133	283	
1	8	16	41	76	133	291	
1	8	17	41	77	142	294	
1	8	17	42	78	145	335	
2	8	18	42	78	146	336	
2	9	18	43	79	150	336	
2	9	18	43	81	157	337	
2	9	18	43	82	161	347	
2	9	19	45	83	161	357	
2	9	20	45	85	161	380	
3	9	20	46	86	164	388	
3	10	21	47	86	166	440	
3	10	22	48	88	167	446	
3	10	23	48	92	168	454	

THEORY OF THE PARETO COMPLEMENTARY DISTRIBUTION

The Pareto distribution [2] was originated in the distribution of income that in different cases were approximately equal in respect to parameter α, whose meaning is explained in what follows. In the case of the income, assuming that the specific percentage of its inverse cumulative frequency $\widetilde{F}$ diminishes with the specific increase of income σ, the following equation is obtained:

$$\frac{\frac{d\widetilde{F}}{\widetilde{F}}}{\frac{d\sigma}{\sigma}} = -\alpha \quad , \alpha > 0 \tag{1}$$

that integrated becomes:

$$\widetilde{F} = \frac{{\sigma_L}^{\alpha}}{\sigma^{\alpha}} \tag{2}$$

where σ_L is the lower limit of the income. There is no population under σ_L. The inverse cumulative probability $\widetilde{F}$ is in a real case, obtained as

$$\widetilde{F}(\sigma) = \frac{N - i + 1/2}{N} \qquad \sigma_{i-1} \le \sigma < \sigma_i 0 \tag{3}$$

when the values of σ are ordered it follows:

$$\{\sigma_i\} = \{\sigma_1 \le \sigma_2 \le \ldots \le \sigma_{i-1} \le \sigma_i \le \ldots \le \sigma_{N-1} \le \sigma_N\}$$

when N values of the variable σ have been registered. The function is discontinuous but it is supposed that the theory corresponds to an infinite number of observations, as it is commonly assumed in statistics. The Pareto's complementary distribution is

$$F(\sigma) = 1 - \widetilde{F}(\sigma) = 1 - \left(\frac{\sigma_L}{\sigma}\right)^{\alpha} \tag{4}$$

The $F(\sigma)$ distribution has the following analytical expression:

$$F(\sigma) = 0 \quad , \ 0 \le \sigma < \sigma_L$$

$$F(\sigma) = 1 - \left(\frac{\sigma_L}{\sigma}\right)^{\alpha} \quad , \ \sigma_L \le \sigma < \infty \ , \ \alpha > 0 \tag{5}$$

in which F is the cumulative probability distribution.
The shape of the distribution may be obtained from the following values:

$$
\begin{aligned}
&F(\sigma_L)=0 \quad , \quad F(\infty)=1 \\
&F'(\sigma)=\frac{\alpha}{\sigma_L}\left(\frac{\sigma_L}{\sigma}\right)^{\alpha-1} \\
&F'(\sigma_L)=\frac{\alpha}{\sigma_L} \quad , \quad F'(\infty)=0 \\
&\alpha>1
\end{aligned}
\tag{6}
$$

The curve then has a J shape (see Figure 1), that has a certain angle with axis σ en σ_L and has as asymptote $F(\sigma)=1$. The mean value is

$$
\begin{aligned}
&\overline{\sigma}=\int_{\sigma_L}^{\infty}\sigma F'(\sigma)d\sigma=\alpha\sigma_L^{\alpha}\int_{\sigma_L}^{\infty}\frac{d\sigma}{\sigma^{\alpha}}=\frac{\alpha}{\alpha-1}\sigma_L \\
&\alpha>1
\end{aligned}
\tag{7}
$$

The quadratic mean value is

$$
\begin{aligned}
&\overline{\sigma^2}=\int_{\sigma_L}^{\infty}\sigma^2F'(\sigma)d\sigma=\alpha\sigma_L^{\alpha}\int_{\sigma_L}^{\infty}\frac{d\sigma}{\sigma^{\alpha-1}}=\frac{\alpha}{\alpha-2}\sigma_L^2 \\
&\alpha>2
\end{aligned}
\tag{8}
$$

The dispersion has a value given by

$$
\Delta\sigma=\sqrt{\left(\overline{\sigma^2}-\overline{\sigma}^2\right)}=\frac{\sigma_L}{\alpha-1}\sqrt{\frac{\alpha}{\alpha-2}}
\tag{9}
$$

From the foregoing the value of the specific dispersions is obtained as

$$
\frac{\Delta\sigma}{\overline{\sigma}}=\frac{1}{\sqrt{\alpha(\alpha-2)}}
\tag{10}
$$

Thus, in the case that $\alpha>2$ the parameters σ_L and α may be obtained from formulas (10), that give α and from (7) or σ_L from (8).

The distribution of formula (5) extends to infinity; this is acceptable when for the last $\sigma_M >> \sigma_L$, see Figure 1. But when only $\sigma_M > \sigma_L$ the upper limit must intervene and then the parameters are α, σ_L y σ_M, see Figure 2.

$$
\begin{aligned}
&F(\sigma) = 0 \quad , \quad 0 \le \sigma < \sigma_L \\
&F(\sigma) = \frac{1-\left(\frac{\sigma_L}{\sigma}\right)^{\alpha}}{1-\left(\frac{\sigma_L}{\sigma_M}\right)^{\alpha}} \quad , \quad \sigma_L \le \sigma < \sigma_M \\
&F(\sigma) = 1 \quad , \quad \sigma_M \le \sigma < \infty
\end{aligned}
\tag{11}
$$

The curve starts with a certain slope $F'(\sigma_L)$ in σ_L, lower limit, and cuts the straight line $F(\sigma)=1$ with the slope $F'(\sigma_M)$. Hence,

$$
\begin{aligned}
&K = 1-\left(\frac{\sigma_L}{\sigma_M}\right)^{\alpha} \\
&F'(\sigma) = \frac{\alpha}{\sigma_L K}\left(\frac{\sigma_L}{\sigma}\right)^{\alpha+1} \\
&F'(\sigma_L) = \frac{\alpha}{\sigma_L K} \\
&F'(\sigma_M) = \frac{\alpha}{\sigma_L K}\left(\frac{\sigma_L}{\sigma_M}\right)^{\alpha+1}
\end{aligned}
\tag{12}
$$

As in the previous case,

$$
\begin{aligned}
&\overline{\sigma} = \int_{\sigma_L}^{\sigma_M} \sigma F'(\sigma) d\sigma = \frac{\alpha}{\alpha-1}\frac{\sigma_L}{K}\left[1-\left(\frac{\sigma_L}{\sigma_M}\right)^{\alpha-1}\right] \quad , \quad \alpha > 1 \\
&\overline{\sigma^2} = \int_{\sigma_L}^{\sigma_M} \sigma^2 F'(\sigma) d\sigma = \frac{\alpha}{\alpha-2}\frac{\sigma_L^{\,2}}{K}\left[1-\left(\frac{\sigma_L}{\sigma_M}\right)^{\alpha-2}\right] \quad , \quad \alpha > 2
\end{aligned}
\tag{13}
$$

$$
\Delta\sigma = \sqrt{\frac{\alpha}{K(\alpha-2)}}\,\sigma_L \left\{1-\left(\frac{\sigma_L}{\sigma_M}\right)^{\alpha-2} - \frac{\alpha(\alpha-2)}{K(\alpha-1)^2}\left[1-\left(\frac{\sigma_L}{\sigma_M}\right)^{\alpha-1}\right]^2\right\}^{1/2}
$$

$$\frac{\Delta\sigma}{\sigma}=\frac{\alpha-1}{1-\left(\frac{\sigma_L}{\sigma_M}\right)^{\alpha-1}}\sqrt{\frac{K}{\alpha(\alpha-2)}\left\{1-\left(\frac{\sigma_L}{\sigma_M}\right)^{\alpha-2}-\frac{\alpha(\alpha-2)}{K(\alpha-1)^2}\left[1-\left(\frac{\sigma_L}{\sigma_M}\right)^{\alpha-1}\right]^2\right\}^{1/2}}$$

Equations (13) are not linear and must be solved by successive approximations.

If two populations with different parameters were used in proportion $k_1/(k_1+k_2)$ and $k_2/(k_1+k_2)$ the results would be,

$$F_1(\sigma)=\frac{k_1}{k_1+k_2}\left[\frac{1-\left(\frac{\sigma_{L1}}{\sigma}\right)^{\alpha_1}}{1-\left(\frac{\sigma_{L1}}{\sigma_{M1}}\right)^{\alpha_1}}\right] \quad , \qquad F_2(\sigma)=\frac{k_2}{k_1+k_2}\left[\frac{1-\left(\frac{\sigma_{L2}}{\sigma}\right)^{\alpha_2}}{1-\left(\frac{\sigma_{L2}}{\sigma_{M2}}\right)^{\alpha_2}}\right]$$

$$F_1=0 \quad , \quad 0\le\sigma<\sigma_{L_1} \qquad\qquad F_2=0 \quad , \quad 0\le\sigma<\sigma_{L_2} \tag{14}$$

$$F_1=1 \quad , \quad \sigma_{M_1}\le\sigma<\infty \qquad\qquad F_2=1 \quad , \quad \sigma_{M_2}\le\sigma<\infty$$

$$F(\sigma)=F_1(\sigma)+F_2(\sigma)$$

An example may be observed in Figure 3. It would be easy to generalize to a larger number of populations.

$$F(\sigma)=\sum_{i=1}^{n}\frac{k_i}{\sum_{i=1}^{n}k_i}\left[\frac{1-\left(\frac{\sigma_{L_i}}{\sigma}\right)^{\alpha_i}}{1-\left(\frac{\sigma_{L_i}}{\sigma_{M_i}}\right)^{\alpha_i}}\right] \tag{15}$$

In the case that α is not larger than one, or takes any other value, it can be made linear with the following expression;

$$\ln\frac{1}{1-F(\sigma)}=-\alpha\ln\sigma_L+\alpha\ln\sigma \tag{16}$$

Formula (16) may be used advantageously to estimate the α parameter by using the least squares method.

For the given case with N measurements of σ, according to formula (3) and the definition of $F(\sigma)$, formula (4), we have:

$$F(\sigma) = \frac{i - 1/2}{N}$$
$$\sigma_{i-1} \le \sigma < \sigma_i \tag{17}$$
$$\{\sigma_i\} = \{\sigma_1 \le \sigma_2 \le \ldots \le \sigma_{i-1} \le \sigma_i \le \ldots \le \sigma_{N-1} \le \sigma_N\}$$

The mean value and the dispersion for the case are calculated with

$$\overline{\sigma} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i$$
$$\Delta\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\overline{\sigma} - \sigma_i)^2} \tag{18}$$

Differentiating formula (4) we obtain

$$\frac{\frac{dF}{1-F}}{\frac{d\sigma}{\sigma}} = -\alpha \tag{19}$$

This is interpreted as if at the beginning $F \to 0$ and we obtain

$$\lim_{F \to 0, \sigma \to \sigma_L} \frac{\frac{dF}{1-F}}{\frac{d\sigma}{\sigma}} = \frac{dF}{\frac{d\sigma}{\sigma_L}} = +\alpha$$
$$\lim_{F \to 0, \sigma \to \sigma_L} F = \alpha\frac{\sigma}{\sigma_L} \tag{20}$$

The growth is relatively fast; at the end of the curve

$$\lim_{F \to 1, \sigma \to \infty} \frac{dF}{d\sigma} = -\frac{1-F}{\sigma}\alpha = 0 \tag{21}$$

growth is very slow. If we take σ as a measured attribute, it means that for a small attribute the population is large, but when the quality increases the population decreases. In the Pareto distribution, the case of income, many people have a small income and very few people have a large income. A probabilistic model was developed by Kittl [3]. If we throw a coin and play a game in which the first player breaking the sequence head-tails wins, in proportion to the number of heads there are many cases in which someone wins after two throws, and few cases in which many throws are needed to win. For instance, a game needing 15 throws has a frequency of 1/309.

We may think that a pair head-tail means an attribute; therefore, very few have a great number of attributes that determine that they are outstanding in something: science, art, fraud! …etc.

When $\alpha < 2$ the form of determining it through formula (10) is useless, but we can get around it as follows:

$$\ln(1-F) = -\alpha\ln\sigma + \alpha\ln\sigma_L \tag{22}$$

or what is the same,

$$\ln\left(\frac{1}{1-F(\sigma)}\right) = \alpha\ln\sigma - \alpha\ln\sigma_L$$

In a real case a straight line is adjusted to the experimental points by using the least squares method, calculating the slope α and the constant $-\alpha\ln\sigma_L$, with which the parameters are determined.

Another method would be making α vary between the limits $\alpha_0 \le \alpha_1 \le \alpha_2 \le \ldots \le \alpha_i \le \alpha_{i+1} \le \ldots \le \alpha_N$ and similarly for σ_L, $\sigma_{L_0} \le \sigma_{L_1} \le \sigma_{L_2} \le \ldots \le \sigma_{L_i} \le \sigma_{L_{i+1}} \le \ldots \le \sigma_{L_N}$ and determining the χ^2 for each pair (α_i, σ_{L_i}).

$$\chi^2 = \sum_{j=1}^{M} \frac{\left(k_j - N\Delta F_{ji}\right)^2}{N^2 \Delta F_{ji}^2} \tag{23}$$
$$\Delta F_{ji} = F\left(\sigma_j, \alpha_i, \sigma_{L_i}\right) - F\left(\sigma_{j-1}, \alpha_i, \sigma_{L_i}\right)$$

Where M is the number of classes, and k_i the experimental number of inhabitants. The minimum χ^2 of the α_i and σ_{L_i} is the one representing more closely the data. Finally, when the representation $\ln(1/1-F)$ versus $\ln\sigma$ has not a linear trend but a curve, the following expression may be used:

$$\ln\left(\frac{1}{1-F}\right) = \alpha\ln\sigma + \alpha_1\ln^2\sigma + \alpha_2\ln^3\sigma + \ldots - \alpha\ln\sigma_L - \alpha_1\ln^2\sigma_L - \alpha_2\ln^3\sigma_L - \ldots \tag{24}$$

This would result in the following distribution:

$$F = 1 - \frac{\sigma_L^{\alpha+\alpha_1\ln\sigma_L+\alpha_2\ln^2\sigma_L+\ldots\alpha}}{\sigma^{\alpha+\alpha_1\ln\sigma+\alpha_2\ln^2\sigma+\ldots}} \tag{25}$$

It is quite difficult – if not impossible – to find a model for (25).

ANALYSIS OF THE DATA

The list of quotes by ISI to academicians of a University Faculty in Chile up to year 2000 is shown in title CITAS ISI. The quotations were arranged in an increasing order and $F(\sigma)$ was calculated by using formula (17), thus getting Figure 4 in which σ takes values between 1 and 2010. The parameters obtained are:

$$\begin{aligned}
&\bar{\sigma} = 119 \quad , \quad \Delta\sigma = 225 \\
&\frac{\Delta\sigma}{\sigma} = 1{,}897 = \frac{1}{\sqrt{\alpha(\alpha-2)}} \\
&\alpha = 1 + \sqrt{1 + \left(\frac{\bar{\sigma}}{\Delta\sigma}\right)^2} = 2{,}131 \\
&\sigma_L = \bar{\sigma}\frac{\alpha-1}{\alpha} = 63{,}1
\end{aligned} \tag{26}$$

Figure 4 represents the curve $\alpha = 2{,}131$, $\sigma_L = 63{,}1$ in which a systematic error may be appraised. In the case of linearization the least squares results are $\alpha = 1{,}30$ and $\sigma_L = 3{,}74$. In this case α is very close to the α given by the experience, $\alpha \approx 1{,}2$, when income is studied in different periods and countries. Calculating the values of χ^2 in the intervals $(2 \le \sigma_L \le 65; 0{,}25 \le \alpha \le 2{,}5)$, a minimum is obtained in $\sigma_L = 30$, $\alpha = 1$, matching the afore mentioned values. The outcome of the linearization of ln(1-F) versus lnσ, is seen in Figure 5 where the existence of two populations may be observed.

If the value of σ is enlarged between 0 and 100, as seen in Figure 6, the presence of three families is detected in this zone. Covering all the available data spectrum, i.e. adding to the three families mentioned the data between 100 and 2010, four families are originated that may be characterized as follows:

$$\begin{cases} 0 \le \sigma \le 40 \\ \alpha = 2{,}522 \end{cases}, \quad \begin{matrix} \bar{\sigma} = 11{,}5 \\ \sigma_L = 6{,}93 \end{matrix}, \quad \begin{matrix} \Delta\sigma = 10 \\ k_1 = 0{,}5 \end{matrix}, \quad \Delta\sigma/\bar{\sigma} = 0{,}872$$

$$\begin{cases} 40 < \sigma \le 70 \\ \alpha = 6{,}73 \end{cases}, \quad \begin{matrix} \bar{\sigma} = 51{,}5 \\ \sigma_L = 40 \end{matrix}, \quad \begin{matrix} \Delta\sigma = 9{,}12 \\ k_1 = 0{,}1 \end{matrix}, \quad \Delta\sigma/\bar{\sigma} = 0{,}177$$

$$\begin{cases} 70 < \sigma \le 100 \\ \alpha = 10{,}3 \end{cases}, \quad \begin{matrix} \bar{\sigma} = 83{,}6 \\ \sigma_L = 70 \end{matrix}, \quad \begin{matrix} \Delta\sigma = 9{,}01 \\ k_3 = 0{,}1 \end{matrix}, \quad \Delta\sigma/\bar{\sigma} = 0{,}108$$

$$\begin{cases} 100 < \sigma \le 2010 \\ \alpha = 2{,}43 \end{cases}, \quad \begin{matrix} \bar{\sigma} = 329 \\ \sigma_L = 194 \end{matrix}, \quad \begin{matrix} \Delta\sigma = 320 \\ k_4 = 0{,}3 \end{matrix}, \quad \Delta\sigma/\bar{\sigma} = 1{,}028$$

$$\begin{cases} F(\sigma) = \sum k_i \left[1 - \left(\frac{\sigma_{L_i}}{\sigma}\right)^{\alpha_i}\right] \\ \sigma_{L_{i-1}} < \sigma \le \sigma_{L_i}, \sum k_i = 1 \end{cases} \tag{27}$$

Depicting these data in Figure 6 it may be observed that there are systematic errors. The values σ of the interval $\sigma_{Li-1} \leq \sigma \leq \sigma_{Li,}$ are used to estimate the coefficients α_i, but in order to represent the data in the corresponding graph, all the values such that $\sigma \geq \sigma_{Li}$ must be employed. It is possible to improve the fitting by proceeding with a finer analysis, as shown in Figure 7, although the systemic deviations persist.

CONCLUSIONS

The analytic expression of the Pareto complementary curve, Formula 19, tells us the specific increment of the population having an attribute larger or equal to a σ which is $dF/1-F$, has in respect to the specific growth of the attribute σ , a constant negative value α. This means that the bigger is the exigency of an attribute σ, the smaller the increment of F for one of σ. When the system saturates $F \to 1$ the growth of F is almost null. We would be then in front of a typical growth based on the "struggle for life". However, the lack of adaptability to this distribution denotes the existence of factors that alter the system that probably are those mentioned in what follows:

- Dissertations of persons who opt for a Doctor's degree in a foreign country, and whose thesis advisors are eminent persons that ensure a good number of citations in magazines ISI; they correspond to papers published early in their career.
- Works in collaboration with foreign scientists that produce many citations in ISI. They can be detected by studying who the authors are.
- Friendly relations with editors of ISI magazines.
- Second order works serving the purpose of solving some loose problems in theories elaborated in the scientific world of the developed countries.
- The method of the self-citation.
- The system of mutual praise.

Although these practices allow for a sizeable increase in the number of citations ISI, they at least imply a considerable effort and a good dose of talent. Inversely, if a work is too innovative, or introduces doubt, or what is worst, it demonstrates that the results obtained by an eminent figure from the milieu of editors of ISI magazines is wrong, it will be blocked by an obsequious service of referees. It is known how a combination of delays and probable plagiarism of a work by Giambiagi and Bollini [4] produced a Nobel prize for t'Hooft and Voltman [5,6]. All this could be corrected with the implementation of Latin American scientific magazines. The structures above mentioned to multiply the citations could indeed be reproduced. Anyway we guess that this might not be so serious.

REFERENCES

1. Data on the ISI.
2. Pareto, W., Cours d'Economie Politique, Lausanne, 1896, Vol. II, book III.
3. Kittl, P., Diaz, G., Gibert, J., El desarrollo científico y tecnológico, particularmente en Chile. *(The scientific and technological development particularly in Chile)* May be consulted in libraries like the National of Chile and the Congress Library of the U.S.A)
4. Bollini, C. Gy Giamblliagei J.J., Lowest order "divergent" graphs in v-dimensional space, Physics Letters, 40B (1972) 566-568.
5. t'Hooft, G. and Voltman, M., Regularization and renormalization of gauge fields, Nuclear Physics, B44 (1972) 189-213.
6. Nobel Prize, Ciencia Hoy *(Science Today).* http://www.ciencia-hoy.retina.ar/hoy55/nobel/.html

FIGURES

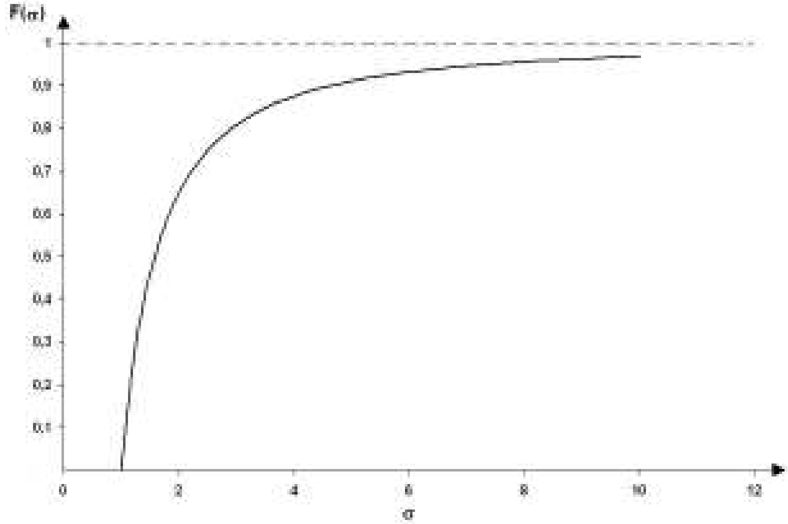

Figure 1: Pareto's Curve in $\sigma_M \rightarrow \infty$. In this case the curve is described by $F(\sigma)=1-\left(\frac{\sigma_L}{\sigma}\right)^{\alpha}$, α = 1,5 y σ_L = 1.

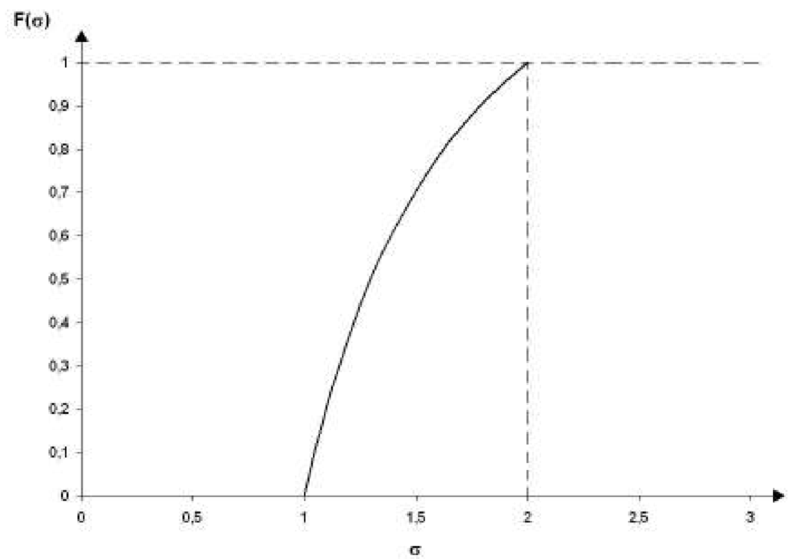

Figure 2: Pareto's curve with finite upper limit σ_M. In this case the curve is described by

$$F(\sigma)=\frac{1-\left(\frac{\sigma_L}{\sigma}\right)^{\alpha}}{1-\left(\frac{\sigma_L}{\sigma_M}\right)^{\alpha}}$$

, α = 1,5 ,σ_L = 1 y σ_M = 2.

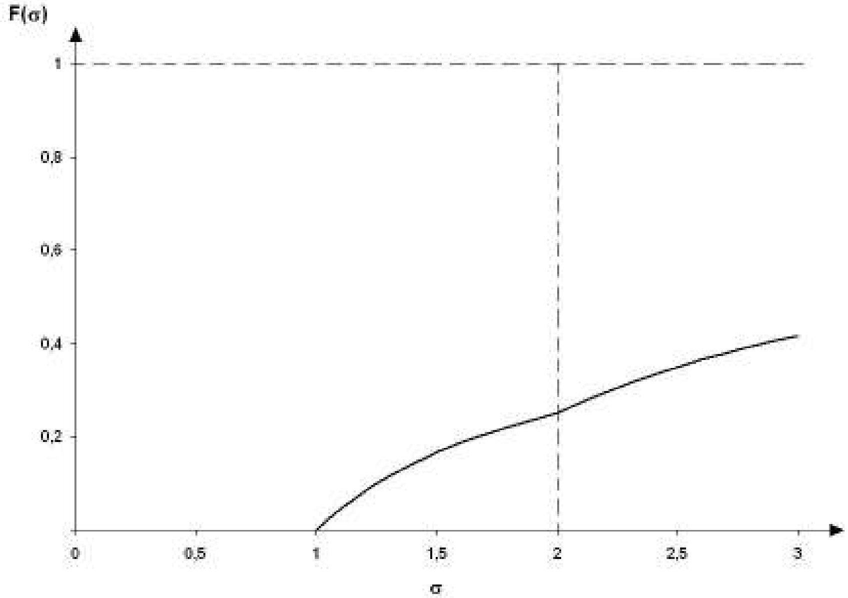

Figure 3: Sum of two populations with different parameters. In this case the curve is described by $F(\sigma)=\sum_{i=1}^{2} k_i\left[1-\left(\frac{\sigma_{L_i}}{\sigma}\right)^{\alpha_i}\right]$, and the parameters of each population are respectively: $\sigma_{L_1}=1; \alpha_1=1; k_1=1/2$ y $\sigma_{L_2}=2; \alpha_2=1; k_2=1/2$.

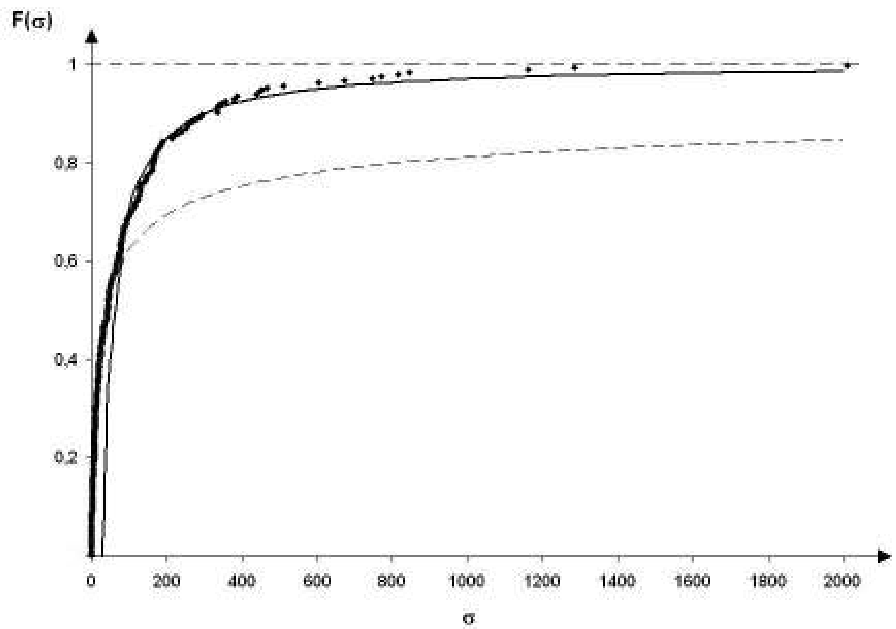

Figure 4: Representation of the effective data and with the parameters $\alpha = 2{,}131$, $\sigma_L = 63{,}1$ (···). In the real case $F_i = \frac{i-1/2}{N}$, N is the total number of observations, in this case, $N = 221$. The Pareto's complementary curve is $F(\sigma) = 1 - \left(\frac{\sigma_L}{\sigma}\right)^{\alpha}$. The dotted curve (-·-·) with $\alpha = 1$, $\sigma_L = 30$ corresponds to the parameters that make χ^2 a minimum for $\alpha \geq 1$. If $\alpha \geq 0$ is accepted a minimum is obtained $\alpha = 0{,}3$, $\sigma_L = 3{,}8$ that adapts well for $3{,}8 \leq \sigma \leq 50$ (- - - -).

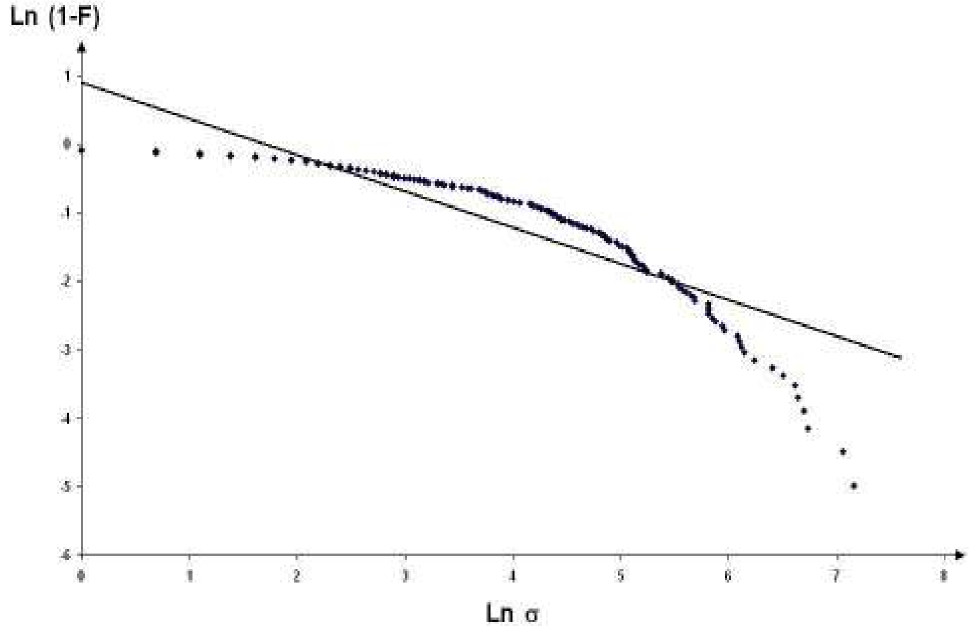

Figure 5: Linearization of the of the Pareto's function using the real data $\ln(1-F) = -\alpha \ln(\sigma) + \alpha \ln(\sigma_L)$, with α = 0,464 and σ_L = 3, As it may be appraised, the Pareto's complementary curve represents a systematic difference with the real data.

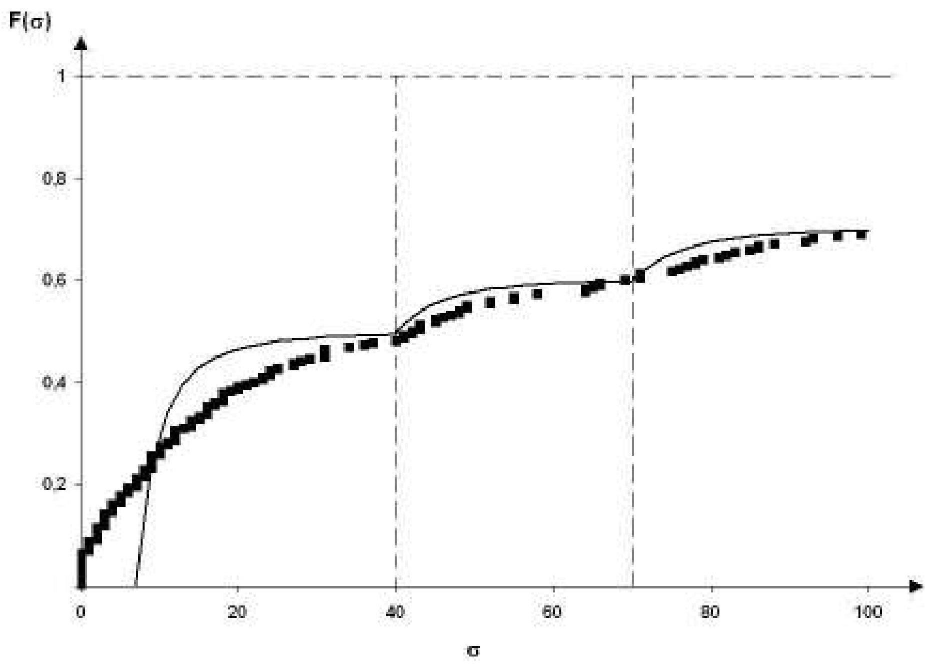

Figure 6: Enlargement of the zone $0 \leq \sigma \leq 100$. Three families may be observed; they present systematic deviations when analyzed with the method of the moments.

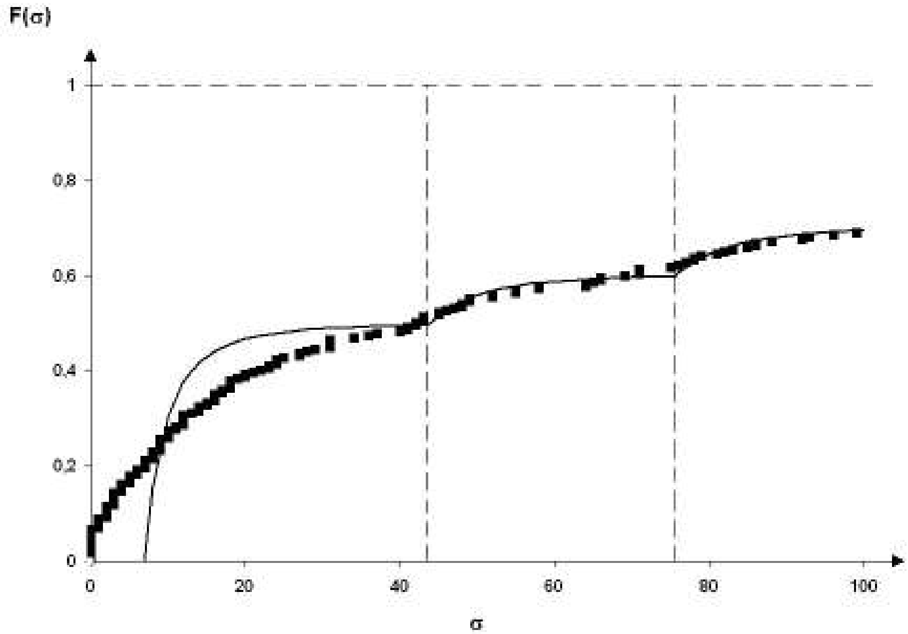

Figure 7: The same data used in Figure 6 but with best fit. The systematic deviations persist.