Flavio Bonifacio, *An automatic method for detection of extreme values: rules and applications*, 3rd Conference in Actuarial Science and Finance, Samos, Greece, 2-5 September 2004

1. Introduction

Some aspects of power law distributions are analysed. First section investigates the loss of continuity we observe in the experimental counterpart of power law function fit. The observed facts usually analysed with power law models often show a noteworthy change in the measure scale. Furthermore it's quite this measure scale change which is invoked as the testing bench of power law based models efficiency. The most quoted examples are: city sizes, incomes, word frequencies, earthquake magnitudes and others alike. In this section we introduce the idea that the scale change it's indicative of a real change, that is a qualitative change in observed facts. Here we give also some hints to recognize this problem as a measure problem.
Second section is charged to give some examples of the problems that arise when applying power law functions to real facts: we do this using city sizes analysis. The analysis is applied on Italian city sizes data collected in two different periods: 1996 an 2001. In this section we will see that is not problem free describing city size distribution with Zipf law.
Third section presents a method to separate observations that are responsible for the change in measuring scale we introduced in section one. Using the samples showed in section two, a simple analytic method will be presented, based on an elementary linear equations system.
Fourth section applies the findings of section three to a real problem. Applying the formulas given in section three the insurance best performing agencies are singled out by insured risk revenue.

2. The problem

As it is well known Pareto distribution pertains to the wider range of power law distribution functions[1]. Pareto (1848-1923)[2] was interested in picturing the asymmetry of the revenue distribution among people. Reference is often made to Pareto cumulative distribution function as the law of 20%-80%: the minority of population (20%) possesses the majority of the total country revenue (80%). Despite the simplicity of the statement and the obviousness of the interpretation and, I would say,

---

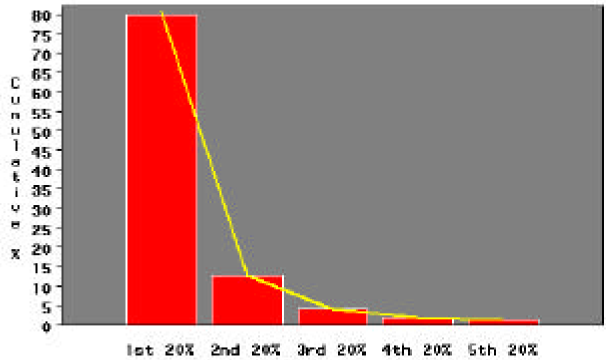
[1] Cfr. Lada A. Adamic, *Zipf, Power-laws and Pareto – a ranking tutorial*, for a short description of this point of view
[2] For a description of Pareto life and work see for instance the following address http://cepa.newschool.edu/het/profiles/pareto.htm

the unpleasant fact, I will report some simple numerical examples showing some particular aspects of Pareto distributions, and more generally of power law functions.
The data of the example are invented and are reported in tab. 1:

Tab. 1 – Input data of Pareto distribution example sorted in descending order

N.	x
	18.400
2	1.747
3	1.154
4	582
5	333
6	233
7	179
8	145
9	122
Total	22.895

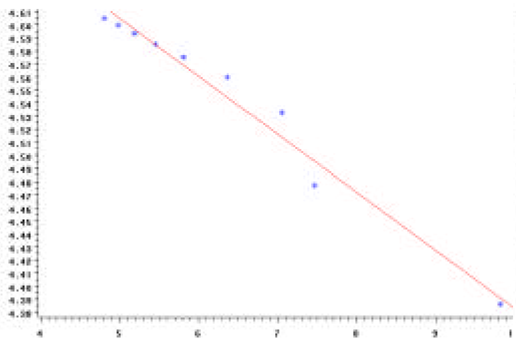
Fig. 1 – Pareto Concentration Curve from data in Table 1



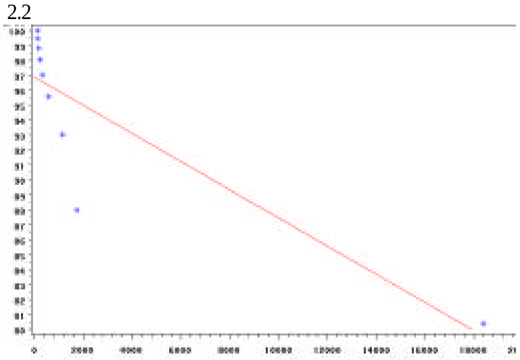
The data in the table are chosen in such a way that they form a Pareto distribution, as is shown in the graph on fig. 1.
The y-axis reports the cumulative proportion of the x_i quantity over the quantity $\sum_i^n x_i$, the x-axis shows the x_i quantities themselves, ranked by dimension in descending order. The distribution shows that the first observation (of nine) counts for the 80% of the total sum of the variable, and the others observations for the remaining 20%.
Drawing on the graph the plot of the power law function, with the parameters estimated by the log-log function, we obtain a better fit than using a simple straight line interpolation[3]. This is shown in Fig. 2.1 by means of the log-log plot, while Fig. 2.2 reports the linear interpolation.

Fig. 2 – Log-log plot(2.1) and linear interpolation (2.2) between cumulative percentage and x values of Table 1

2.1



2.2



[3] The power law defines $y = Cx^{-a}$; the derived log-log function is $\log(y) = \log(c) - a\log(x)$

The R^2 values are 0.97 and 0.75 respectively. Furthermore[4] we can easily observe that it's the presence of the outlier point that makes the distribution of the x_i's and their cumulated percentage power law. Erasing the observation n. 1 of table 1 the R^2 values change: the log-log curve R^2 decreases at 0.85 and linear interpolation R^2 increases to 0.98 reversing the goodness of fit hierarchy between the estimated models.
In this example we see also that the magnitude of the outliers is important to suggest the power law form as a better fitting interpolation. That is, it seems that not only outlier observations have to exist in order to make the distribution power law, but also that their dimensions have to be much bigger than the average value of the trimmed distribution, in order to obtain a good fit for a power law function. Here comes the first question: given a distribution of x's , how much big have to be the outliers to make the distribution power law (say Pareto or Zipf) or, in other terms, how much extreme they are expected to be? Second question: established that some of such values exist, how is it possible to identify them? In this paper we suggest a method to answer to the second question and we give some empirical example of his plausibility, letting somebody the more heavy work of eventually proof it in a more general, formal and elegant way. We also let apart the answer to the first question. About this topic we only outline some implications of substantive interest.
There is an important consequence of the above mentioned questions: if a change of measuring scale is implicated by a power law distribution, it may be due to a true qualitative change in the phenomenon. The presence of a significant power law fit, measured in terms of OLS regression, may be the signal of a qualitative change in the subjacent phenomenon or an indicator of discontinuity in the linking function. We will see a possible statistical representation of this by means of variance-covariance analysis in a real application.
As it is well known, one of the most relevant problems in the social sciences, and in sociology more than in psychology and economics, is the measure of objects: sometimes the same object is measured in a different way, sometimes different objects are measured in the same way, sometimes the measure that is supposed to measure one object, measures in fact other objects . Often different stimuli on the same topic yield different results on the topic: for example asking the same question in an open ended form or in a closed form will give totally different results. The observed differences in this case are not to be assigned to a real observed object change, but to the different stimulus that results in different measures. At least the correct attribution of the effect is

[4] In this context we take an empirical point of view: we are thinking about some real applications of the model and not about the model itself.

questionable. All the theory of measuring error concerns this kind of problems. Another, perhaps more pertinent example, concerns different objects measured in the same way: factor analysis produces measures of this type. In this case we have the same set of stimuli that, arranged in some way, rise to a set of measures (usually the principal components) related to different supposed underling latent variables referring to different conceptual objects. Finally, as an example of mismeasurement, let me quote the famous example of storks and children[5]: the number of storks living in a country is a good measure of the number of children born in that country. Or, more directly, storks are carrying children. Where both storks and children are obviously indicators of something else[6]. Lazarsfeld defines this fact as an *hidden relation*.
As in the examples mentioned above, the measure scale change underling a power law distribution fit may be a proxi of some latent phenomenon. That is a phenomenon underlaid to and in some way hidden by the actually measured object.

3. Application fields: city size distributions

What I would say is that in real experiments[7] there is a sort of lack of continuity that make up a power law fit. We have already saw it in the above reported numerical example. We will see it again in a common application field of power law distribution, the City size distribution analysis. Let me briefly introduce the argument by quoting Gabaix and Ioannides[8]: "*The evolution of city size distributions has attracted researchers over a long period of time. The existence of very large cities, the very wide dispersion in city sizes, the remarkable stability of the hierarchy between cities over decades or even centuries, and the role of urbanization in economic development are all particularly*

---

[5] Cfr. Paul F. Lazarsfeld, *Interpretation of statistical relations as a research operation*, in Paul F. Lazarsfeld, M. Rosenberg, *The Language of Social Research*, pp. 115-125, Glencoe, 1955

[6] Cfr. Ferenc Moksony, *Small is beautiful. The use and interpretation of R^2 in social research*, "A classic example may serve to illustrate the basic difference between substantive and statistical explanations. The number of births in a given locality can be estimated reasonably well from the number of storks in the same area; if we ran a regression with the number of births as the dependent and the number of storks as the independent variable, we would probably get a fairly large R^2. But does this mean the number of storks explain, in a substantive sense, the level of fertility? Obviously not; the statistical explanatory power of this variable derives entirely from the fact that it is correlated with the real determinant of the number of births - namely, the degree of urbanisation. Rural areas have more storks and they also have a higher birth rate."

[7] Not in mathematical models, where everything is arranged in such a way that all things run well and where the continuity of the function is a precondition for the existence of the function.

[8] X.Gabaix, Y.M.Ioannides, *The Evolution of City Size Distribution*, Department of Economics, Tufts University, 2003, p. 4

interesting qualitative features of urban structure worldwide. Another surprising regularity, Zipf's law for cities, has itself attracted considerable interest by researchers…".

As is noted by Gabaix and Ioannides they selves[9], the first problem in this kind of studies is the city definition: they reported as a relevant definition problem the alternative between city-proper data versus urban agglomeration data (i.e. metropolitan areas). Another problem that arise examining city size distribution is the choice of the population cutting point from which the given definition holds[10].

An experiment conducted by us on Italian cities over 100.000 inhabitants gives results similar to those reported by Gabaix and Ioannides with reference to Brakman, Garretsen and Van Marrewijk[11]. The quoted work reports for cities-proper data a Zipf's exponent mean of 1.13 (over several experiments conducted in 44 countries with a minimum of 0.8 and a maximum of 1.5): we obtained for year 1996 a Zipf coefficient of 1.12 with 41 cities in the sample. The R^2 value is 0.99. The log-log equation is:

$$log(R) = 16.56445 - 1.12068 * log(P) \qquad \text{obs. 41} \qquad (1)$$

Where *log(*R) is the natural logaritm of the Rank of the cities plotted in descending order (1 is the biggest city, 41 the smallest) and –1.12068 is the Zipf coefficient; *log*(P) is the natural logarithm of the city population and 16.56445 the intercept value. Using natural numbers instead of logarithms we will have the usual power law function:

$$R = exp(16.56445) * P^{-1.12068} \qquad (2)$$

For year 2001 we obtained :

$$log(R) = 16.77134 - 1.13975 * log(P) \qquad \text{obs. 42} \qquad (3)$$

still with an R^2 value of 0.99.

The same work conducted by Krugman[12] on USA cities yields this equation:

---

[9] Ibidem, p. 6

[10] "The exponent ζ is sensitive to the choice of the cutoff size above which one selects the cities". Where ζ is taken from the equation: $P(Size > S) = \frac{a}{S^z}$ given that S_i is the size of city *i*. Ibidem, p. 5

[11] Brakman, S., H. Garretsen and C. van Marrewijk, *An Introduction to Geographical Economics*, Cambridge University Press, 2001

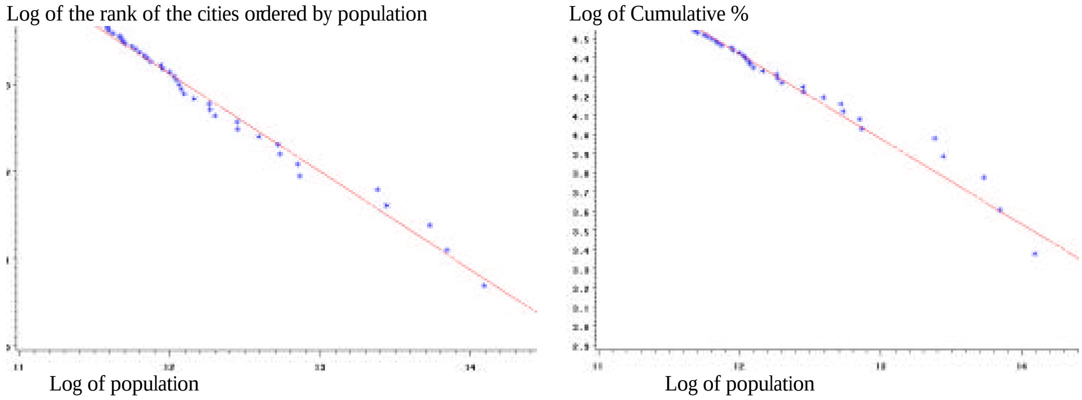
[12] P. Krugman, *The Self-Organizing Economy*, Blakwell Pubblishers, Oxford, 1996, quoted by Gabaix, Joannides, p. 6

$$log(R) = 10.53 - 1.005*log(P) \quad (4)$$

In that case the cut-off point was set at 250.000 inhabitants and the cities, defined as metropolitan areas, were 131. Due to these differences it's hard to compare the reported results[13]. Anyway working with Italian data and changing the cut-off point it seems that the higher the cutting point is set, the smaller, i.e. closer to 1, the Zipf exponent will be, as we will see later.

For illustrative purpose we report the graphs of the log-log plot between 1) the logarithm of the rank of the cities ordered by population and the logarithm of the city population itself, 2) the logarithm of cumulative percentage of city population and the logarithm of the city population itself. For graphical purpose there are in practice no differences using the log of the cumulative percentage instead of the log of the rank, as the reader may easily see in fig. 3. Logarithm of the rank and logarithm of the cumulative percentage are indeed very similar: the R^2 between them is 0.98.

Fig. 3 – Log-log plot for italian cities > 100.000 inhabitants



If we take the italian population data for years 1996 and 2001, excluding the observation concerning Rome, the exponent of the power law function increase over 1.15-1.20, while the R^2 remains almost constant. The quoted value are indicated by Gabaix and Ioannides as the upper Zipf exponent thresholds of a power law well describing the observed empirical regularities. Equations (5) and (6) show the results of log-log estimates in the case of reduced sample.

1996: $log(R) = 18.47789 - 1.28467*log(P)$ Obs. 40 (5)
2001: $log(R) = 18.73980 - 1.30874*log(P)$ Obs. 41 (6)

[13] Also the data collection years are different: 1996 and 2001 for Italy, 1991 for USA.

To observe changes in Zipf exponent, we may change the cutting point, instead of trimming the upper queue. As noted above, setting the cutting point may be in fact questionable. For example, if we take the Italian administrative definition, it's really impossible to fix it. Simply because there are so many variables used in the city definition, that is impossible to determine common rules. Let me show you some examples.

The mayor of *Monte Porzio* (Provincia of Pesaro Urbino, Center Italy, inhabitants 2.227) so describes the city acknowledgment process:
"The city acknowledgment is given by the Republic President. The guidelines to obtain the city acknowledgment are different. For example, Castel Gandolfo has got it because it has been set to the summer residence of the pope. Another rule may be the inhabitants number. Monte Porzio obtained it for his noteworthy qualities: from Tuscolo to Barco Borghese, from the Information Centre for Young People to the recreation centre..". Where *Tuscolo* is the name of the *Monte Porzio* surrounding area near Rome, recalling a no more existent ancient city receiving the city acknowledgment by no less than Ulisses' and Circes' son Telegonus[14].
In the www site of Carignano (Provincia of Torino, North Italy, inhabitants 8.647) we read: The city acknowledgment comes in 1683, given by the Duke Vittorio Emanuele II.
Monte S'Angelo (Provincia of Foggia, South Italy, inhabitants 13.917) writes: *"The Municipality is entitled by the City title, acknowledged in 1401 (by papal bull "Rerum omnium" of Bonifacio IX).*
And Santa Maria Capua Vetere (Provincia of Caserta, South Italy, inhabitants 30.745) simply notes: *The Municipality of Santa Maria Capua Vetere is entitled by the City title.*
We find in Iesolo www site (Provincia of Venezia, North Italy, Inhabitants 22.698): *Iesolo has got the City Title in 1983 directly by the Republic President Sandro Pertini.*

Furthermore it seems that in the middle age at least in Italy the common rule of the city (*civitas*) acknowledgment was the presence of an Episcopal seat, but also this had to be questionable if Bartolo da Sassoferrato (1313-1357) should say: *"Civitas vero secundum usum nostrum appellatur illa quae habet episcopum: antea tamen quam essent episcopi erant civitates. Et civitati competit potestas eligendi sibi de iure communi defensores, qui habeant iurisdictionem […] et quia secundum canones episcopi debent ordinari in dictis locis ubi sunt officiales […] ideo insurrexit consuetudo quod locum habens*

[14] Telegonus killed his father, who was unknown to him, with a lance pointed with the spine of a trygon, or stingray, which Circe had given him. The roman mythology tells the story that after the death of his father, Telegonus married Ulisses' widow Penelope and had a son named Italus, wich became the king of the Oenotrians or Siculi, the first inhabitants of Italy.

episcopum sit civitas, tamen vere sine episcopo dicitur civitas, eo quod habet officiales praedictos, et iurisdictionem[15]*".*

That is because there were cities before bishops and the reason of cities existence has to be considered strictly a political matter, concerning the territory defence. As it has been noted by Marco Folin[16] this was already said by Brunetto Latini one century before:

«Cittade è uno raunamento di gente fatto per vivere a ragione; onde non sono detti cittadini d'uno medesimo comune perché siano insieme accolti dentro ad uno muro, ma quelli che insieme sono accolti a vivere a una ragione»[17]. Aren't walls that make cities, but the citizens logic goals.

To summarize there are sufficient reasons to conceive the above retained population cut-off points rather arbitrary. These thresholds have to be thought as conventional and liable to be weighted by historical and social conditions. Therefore it seems not so extravagant, at least in the Italian case, to set up a lower cut-off point than the one fixed before, and see what happens with Zipf coefficients. Choosing a cutting-point of 45.000 inhabitants the results of the calculations are the following:

1996: $log(R) = 19.27710 - 1.33561 * log(P)$ obs. 160 (7)

2001: $log(R) = 19.49485 - 1.35660 * log(P)$ obs. 165 (8)

The following table collects the calculated Zipf exponents varying year and cut-off points:

Tab. 2 – Calculated Zipf coefficients by year and cut-off points, Italy

	Cut-off points	
Year	**100.000 *inhab.***	**45.000 *inhab.***
1996	-1.12068	-1.33561
2001	-1.30874	-1.35660

We may conclude that the fit of Zipf curve decreases (the exponent increases in absolute value) if the cut-off point is lower given the year and if the data collection year is more recent given the cutting point.

[15] B. da Sassoferrato, *Tractatus super constitutione Qui sunt rebelles*, quoted in D. Quaglioni, *«Civitas»: appunti per una riflessione sull'idea di città nel pensiero politico dei giuristi medievali,* in *Le Ideologie della città dall'umanesimo al romanticismo*, edited by V. Conti, Firenze, Olschki, 1993, pp. 63-64.

[16] The quoted text and the notes have been collected from M. Folin, *Sui criteri di classificazione degli insediamenti urbani nell'Italia centro-settentrionale (secoli XIV-XVIII)*, to be published in France in *Les mots de la ville*, edited by B. MARIN, Paris, CNRS.

[17] B. Latini, *Le livre du tresor,* edited by F. J. Carmody, Berkeley 1948, p. 391 (cit. in D. Quaglioni, *«Civilis sapientia». Dottrine giuridiche e dottrine politiche fra Medioevo ed Età moderna*, Rimini, Maggioli, 1989, p. 135).

As a final remark on this topic we observe that it seems to be rather reductive to think about the historical conditions influencing the city size distribution just as economic conditions linked to industrial development. In the literature we may in fact find several kind of models[18] but many of them refers to industrial assets. This approach is well documented by Axtell and Florida that gave the following title to a section of their paper: *People Form Firms, Firm Clusters Are Cities*[19]. We suspect that this kind of models may be adequate only in certain period and in well established countries (i.e. USA in the last two centuries), while it would be useful to think also to other social constraints. Perhaps Bartolo da Sassoferrato would say that there have been cities before the firms have been.
From one side we may easily observe that in different countries we have different (big) cities: Sao Paulo is different from New York and New York is different from London, which in turn is different from Paris. They all are different from Peking. From the other side the mentioned cities are similar among them and different from other, smaller cities: Rome, Athens, Madrid, Prague and so on. No matter if a unique Zipf function may describe the distribution. I think that the lack of continuity in the distribution of city sizes, that we noted at the beginning of this paper, is a proxi of qualitative different phenomena. If it's true that any big city doesn't come from nothing, with a few exceptions, (Saint Petersburg in Russia may be one of them), it's also true that, when the difference in size becomes real, it's very hard to find regularities in one size level that explain the phenomena pertaining to the other size level, except adopting an historical and locally focused point of view.

4. Extreme values[20] detection

Analysing the Italian city sizes we noted that the interception point between the linear interpolation of observed points (population against cumulative percentage distribution) and the linear interpolation of estimated log-log points (population against the cumulative distribution

[18] See for example the ones reviewed by Gabraix and Ioannides in the quoted paper "*The evolution of city distribution*", p. 25
[19] Cfr. R.L. Axtell, R. Florida "*Emergent Cities: a Microeconomic Explanation*", Brookings Institution and Carnegie Mellon University, 2000, Draft Paper, p. 2
[20] As the reader may see our approach refers to the Extreme Value Theory just for the argument covered, but not for the method. Our method is quite empirical instead mathematic and statistics is involved as enquiry tool instead of theoretic instrument.. For a short presentation of the Extreme Value Theory and related references see for example M. Gilli, E.Kellezi, *An Application of Extreme Value Theory for Measuring Risk*, Department of Econometrics, University of Geneva and FAME, Geneva, 2003

estimated by the power law function) performed well as a threshold between normal sized cities and big cities. Given the linear system:

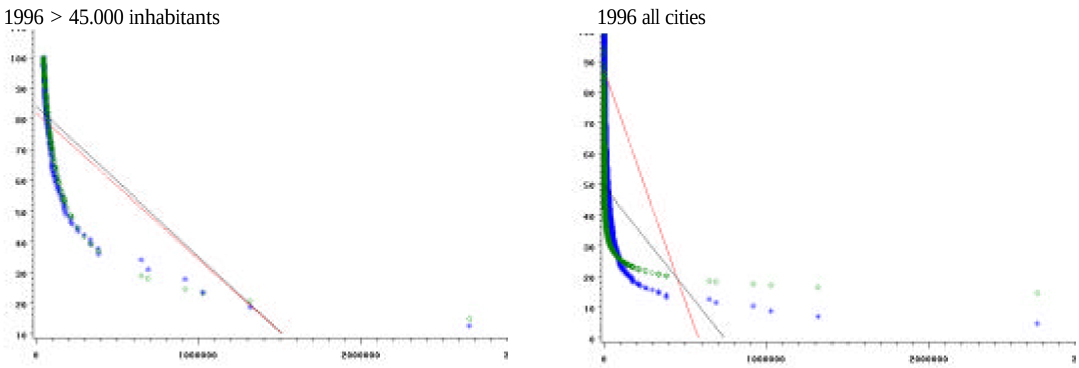
$$\begin{cases} Y = a + bx \\ Y = a' + b'x \end{cases} \tag{9}$$

where the a and b are estimated by the equations described above, the analytic solution for x is:

$$x = \frac{a - a'}{b - b'} \tag{10}$$

The obtained x values are the threshold points. Applying this method we draw the graphs reported in fig.4, showing the interception points for some experiments on Italian city sizes. Only in the case of minimum variance, in which the cities in the experiment for both years have a number of inhabitants comprised between 100.000 and 2.500.000, therefore excluding Rome which was sized 2.546.804 inhabitants in 2001, and comprising Milan which in the same year was 1.256.211. In this case in fact we may say that there are no extreme values and that the distribution doesn't present heavy continuity interruptions.
Whatever is the cut-off point we choose, 100.000 or 45.000 inhabitants, Rome is selected as extreme value. If we consider the entire values distributions, without imposing cut-off points, there are 6 cities considered in both years, as extreme values: Rome, Milan, Napoli, Torino, Palermo and Genova, which are indeed the biggest cities in Italy.

Fig. 4 – Plot of observed and power law function estimated values – 1996 cities > 45.000 inhabitants



5. Using extreme value detection in the insurance field

We used the method defined above to detect the existence of extreme values[21], in the specified sense, for revenue of agencies of an insurance company. We did it by type of insured risk and for all the risks insured in a particular agency. For each risk we calculated the log-log estimates, which are reported in the following table with the resulting R^2 value (first three columns of table 3). By looking to the exponent estimates by the log-log function we first observe that the Zipf coefficients are in almost half of cases enough close to 1 (third, fifth, sixth, seventh and ninth lines), but not in the general case regarding all the risks and in the other cases. Secondly we observe that in each case the R^2 value is less than 0.90 and that in half of cases the simply linear estimation is better than the log-log one: lines 1, 4,7,8,9,12.
If we use the threshold point as a cluster point separating the normal values from the extreme values as defined before and then we estimate two models, one for each cluster of observations (agencies), then the overall models R^2 (column 5) increases in every case by a consistent quantity (column 6) and almost ever his value goes beyond 0.90. This way to estimate the model and the model fit conforms by logic and by numbers with variance and covariance analysis having the thresholds value as a class factor, and the revenue values as a covariate in a model where the cumulated percentage of revenue is the dependent variable.
This means, as we stated before in section 1, a signal of a qualitative change in the subjacent phenomenon or an indicator of discontinuity in the linking function, if the linking function is thought as a single model function (running on all the observations). From a substantive point of view this qualitative change may imply a self adapting model, that is a model able to change his parameters depending on the situation in which it operates.
For illustrative purpose we report in the following figures the plots of the observed and estimated points, with the linear interpolation generating the threshold points for each kind of risk and for the total revenue. For clarity we reports the Pareto concentration curve too. In table 4 we summarize the analysis results, reporting for each combination agency-risk a flag (value 1) indicating if the agency has extreme value for the annual revenue for a particular risk. The marginal totals give us an indication of risks that have at least one agency with high revenues (columns marginal), or give us an indication of the agencies that have at least one risks with high revenue.

[21] It's now clear the way we intend extreme values: they are those values that more contribute to the goodness of fit of a power law function. Therefore nothing to do with Extreme Value Theory.

Tab. 3 – Insurance analysis coefficients summarising Table

Insured risk	(1)Log-Log Intercept	(2)Log-Log (Zipf) Exponent	(3)R^2 Of Log-Log Plot	(4)R^2 Of Linear Plot	(5)R^2 Of Anocova Plot	(6) R^2 Improv Ment (5)-(4)
Car insurance	25.43060	-1.32165	0.80	0.96	0.98	0.02
Legal risk	11.47104	-0.62892	0.76	0.75	0.96	0.21
Civil generic risk	18.61158	-0.97564	0.86	0.76	0.90	0.14
Environmental risk	9.47329	-0.47397	0.74	0.90	0.98	0.08
Social and political risk	14.54892	-0.90807	0.85	0.80	0.94	0.14
Accident risk	19.69320	-1.07360	0.86	0.63	0.89	0.16
Stealing	16.44612	-0.89878	0.82	0.84	0.95	0.11
Cristals	12.08413	-0.72819	0.76	0.83	0.94	0.11
Fire	17.19794	-0.89515	0.79	0.88	0.97	0.09
Health	13.64583	-0.70135	0.59	0.48	0.83	* 0.15
Life	21.64761	-1.20497	0.86	0.77	0.93	0.16
Life one premium	14.15150	-0.65955	0.68	0.73	0.95	0.22
All risks	27.46666	-1.36025	0.88	0.82	0.94	0.12

Fig. 5 – Health Risk Insurance Threshold Point and Concentration Curve

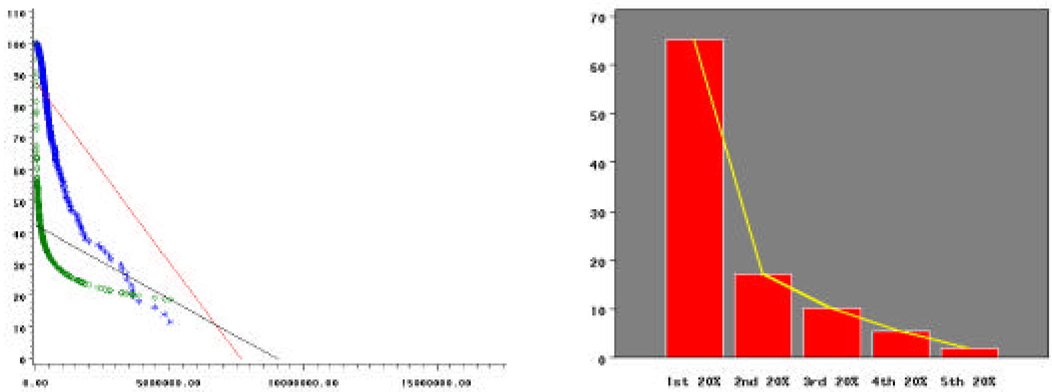
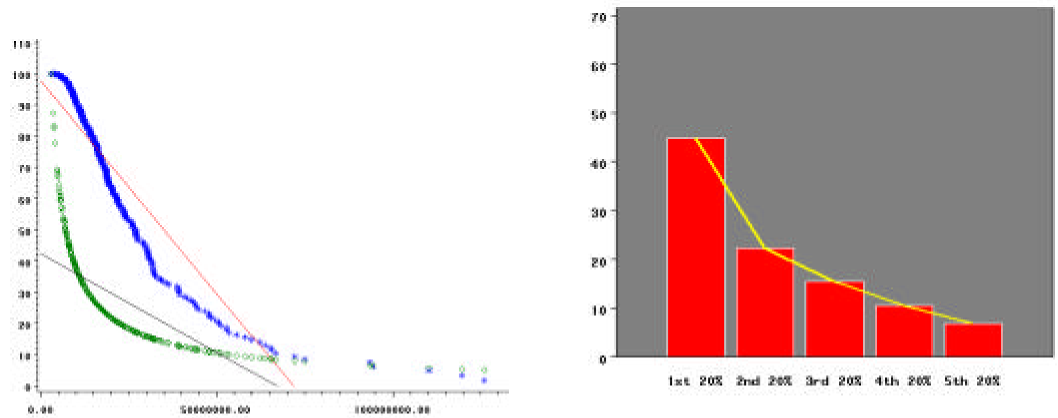


Fig. 6 – All Risks Insurance Threshold Point and Concentration Curve



Tab. 4 – Insurance Analysis - Extreme values summarised by Agency and Risk Type

	FLAG VALORE ESTREMO													
	TOTAL EXT. VAL.	RISK O_01	RISK O_02	RISK O_03	RISK O_05	RISK O_07	RISK O_09	RISK O_10	RISK O_15	RISK O_20	RISK O_21	RISK O_4A	RISK O_4B	ALL RISK
CODICE AGENZIA														
119	1	1	0	0	0	0	0	0	0	0	0	0	0	0
132	1	0	0	0	0	0	0	0	0	0	1	0	0	0
137	1	0	0	0	0	0	0	1	0	0	0	0	0	0
155	1	0	0	0	0	0	0	0	0	0	1	0	0	0
159	1	0	1	0	0	0	0	0	0	0	0	0	0	0
180	5	0	0	1	0	1	1	0	0	1	0	0	0	1
188	1	0	0	0	0	0	1	0	0	0	0	0	0	0
192	1	0	1	0	0	0	0	0	0	0	0	0	0	0
207	1	1	0	0	0	0	0	0	0	0	0	0	0	0
210	9	0	0	1	1	1	0	1	1	1	0	1	1	1
232	1	0	0	0	0	0	0	0	0	0	0	0	1	0
235	10	0	1	1	1	1	0	1	0	1	1	1	1	1
251	3	0	1	0	1	0	0	0	0	0	0	1	0	0
253	7	0	0	0	1	1	1	1	0	0	0	1	1	1
254	1	0	0	0	0	0	1	0	0	0	0	0	0	0
260	2	0	0	0	0	1	1	0	0	0	0	0	0	0
279	4	1	1	0	0	0	0	0	0	0	0	1	1	0
465	8	1	0	1	1	1	0	1	0	1	0	1	0	1
470	1	0	0	0	0	0	0	1	0	0	0	0	0	0
510	1	0	0	1	0	0	0	0	0	0	0	0	0	0
613	1	0	0	0	0	1	0	0	0	0	0	0	0	0
742	1	1	0	0	0	0	0	0	0	0	0	0	0	0
745	1	1	0	0	0	0	0	0	0	0	0	0	0	0
750	1	0	0	0	0	0	0	0	0	1	0	0	0	0
TOTAL EXTREM VAL.	64	6	5	5	5	7	5	6	1	5	3	6	5	5

6. Final remarks

In first section we showed the loss of continuity we observe in the experimental counterpart of power law function fit. Few observations may have a great influence on the model fit. We concluded that power law functions must be weighted with historical and local facts and their general validity (outside the specified condition of existence) may be questionable. Furthermore it seems that the existence of a power law fit may be used as an indicator of a qualitative change in the measure scale. There are some facts indeed that indicate the existence of measure problems.

Second section shows, among other, that the cut-off point from which the cities are considered to be cities is questionable and it is responsible of the efficiency of the Zipf power law simulation. Once again it seems that local and historical variables must be taken in consideration.

Third section outlines that there is an elementary method to separate observations that are responsible to the change in measuring scale: it is sufficient to draw straight lines on observed and on log-log estimated points, converted in power law form, to separate extreme values from "normal" values. The extreme values are defined as the values that most contribute to make a power law function the best fit in the model.

In fourth section we simply apply the findings of third section for identifying the best performing agencies of an insurance company. We obtained a table in which are listed the agencies with one extreme revenue value in at least one kind of risk. Obviously inverting the values it would be possible to identify the worse performing agencies.

We didn't do any formal proof of the general validity of the above presented threshold finding method. We will try to do it in the future. For now we shall only say that the presented method seems to work well: just like the Zipf distribution. Nobody knows why, but it works.

Index