Marshall-Olkin Extended Zipf Distribution

Marta Pérez-Casany[a] **and Aina Casellas**[b] [1]
a. Department of Applied Math 2 and Dama-UPC, Technical University of Catalonia
b. Technical University of Catalonia

Abstract: The Zipf distribution also known as scale-free distribution or discrete Pareto distribution, is the particular case of Power Law distribution with support the strictly positive integers. It is a one-parameter distribution with a linear behaviour in the log-log scale. In this paper the Zipfian distribution is generalized by means of the Marshall-Olkin transformation. The new model has more flexibility to adjust the probabilities of the first positive integer numbers while keeping the linearity of the tail probabilities. The main properties of the new model are presented, and several data sets are analyzed in order to show the gain obtained by using the generalized model.

key words and phrases: Scale-free distribution; Zipf distribution; Marshall-Olkin transformation; skew distributions; scale-free network.

1   Introduction

The Zipf distribution [12] appears very often in practice when modelling natural as well as man-made pehnomena. This is because of its simplicity and its suitability to capture the main sample characteristics. Between these characteristics one wants to enhace the following: a) large probability at one in most of its parameter space, b) long right tail, c) linearity in the log-log scale and d) scale-free. Nevertheless, in many cases the proportion of the first few positive integer values observed, differs considerably from the expected probabilities under the Zipf distribution. This is a consequence of the fact that, in those situations, the linear behaviour only holds for large integer values.

The Zipf distribution is a particular case of the discrete Power Law (PL) distribution. In [3] it appears more that twenty situations corresponding to different research areas where the PL distribution has been considered as a candidate distribution. The areas correspond to physics, biology, information sciences, social networking, engineering or social sciences. For example, in real world one observes that a few mega cities contain a population that is orders of magnitude larger than the mean population of cities, and a lot of citites have a much smaller population. In internet one observes that very few sites contain milions of links, but many sites have just one or two links, or that millions of users visit a few select sites, giving little attention to millions of other sites (see [2]).

---

[1]Address for correspondence: Marta Pérez-Casany, Dept. Applied mathematics II and Dama-UPC, Technical University of Catalonia [a], Barcelona, Spain. E-mail: marta.perez@upc.edu

Reserchers from linguistics, ecology, demography, economy, genetics or, more recently, social networking use the Zipf distribution to model data usually presented as rank data or frequencies of frequencies data.

The objective of the paper is to define a two-parameter generalization of the Zipf distribution that is much more flexible in modelling the probability for the first positive integer values while at the same time allowing for both concave and convex representations of the first probabilities in the log-log scale. This is done by applying the transformation defined by Marshall and Olkin in 1997 [9]. In Marshall and Olkin's original paper, the transformation is used to generalize the exponential and the Weibull distributions. Later, the transformation has been used to generalize the Lomax, the Pareto or the Log-normal distributions. Several papers that appear in the last few years apply the generalizations in reliability, in time series and in censored data. See for instance [1], [5, 6] or [7]. In [11] that transformation is presented as a skewing mechanism, and several classes of unimodal and symmetric distributions are extended in that manner.

The paper is organized as follows. Section 2 is devoted to the preliminars. In Section 3 the Marshall-Olkin generalized Zipf distribution is defined and its main properties are presented. In Section 4, three large data sets from very different research areas are analyzed proving the usefulnes of the model presented.

2 Preliminars

2.1 The Zipfian (Zipf) distribution

A random variable (r.v.) X is said to follow a Zipf distribution with *scale* parameter $\alpha > 1$ if, and only if, its probability mass function (pmf) is equal to:

$$P(X = x) = \frac{x^{-\alpha}}{\zeta(\alpha)}, \ \ for \ \ x = 1, 2, 3, \cdots, \tag{1}$$

where $\zeta(\alpha) = \sum_{k=1}^{+\infty} k^{-\alpha}$ is the Riemann zeta function. The Zipf distribution [12] is the particular case of the discrete PL distribution with support the positive integers larger than zero, and it can also be viewed as the discretization of the Pareto (Type I) distribution. The Zipf distribution is often suitable to fit data that correspond to frequencies of frequencies or to ranked data. These type of data show a widespread pattern in their measurements with a very large probability at one and a very small probability at some very large values. Moreover, from (1) one obtains that the Zipf distribution will be appropiate when the data show a linear pattern in a log-log scale, because

$$\log(P(X = k)) = -\alpha \log(x) - \log(\zeta(\alpha)).$$

The continuous PL distribution is the only continuous distribution such that the shape of the distribution curve does not depend on the scale on which the measures are taken [10]. For this reason, it and its discrete version are also konwn as *scale-free* distributions.

Given that

$$F(X) = P(X \le x) = \frac{1}{\zeta(\alpha)} \sum_{k=1}^{x} k^{-\alpha},$$

the survival or reliability function (SF) of the Zipf distribution is equal to:

$$\overline{F}(X) = P(X > x) = \frac{\zeta(\alpha, x+1)}{\zeta(\alpha)}. \tag{2}$$

where $\zeta(\alpha, x)$ is the Hurwitz zeta function, defined as:

$$\zeta(\alpha, x) = \sum_{k=x}^{+\infty} k^{-\alpha}. \tag{3}$$

The mean of a Zipf distribution is finite for $\alpha > 2$, and the variance is finite only when $\alpha > 3$. Assuming that $\alpha > 3$,

$$E(X) = \frac{\zeta(\alpha - 1)}{\zeta(\alpha)}, \quad \text{and} \quad Var(X) = \frac{\zeta(\alpha - 2)\zeta(\alpha) - (\zeta(\alpha - 1))^2}{(\zeta(\alpha))^2}.$$

Often in practice the Zipf distribution fits well the probabilities for large values of X but not the probabilities for the smaller ones. Plotting the observed values in the log-log scale usually one observes a concave or a convex shape for the smallest values and the linearity holds only for values larger than a given positive value. In order to avoid this problem and increase the goodness of fit of the model, in this paper the Zipf distribution is generalized by means of the Marshall-Olkin transformation.

2.2   The Marshall-Olkin transformation

In 1997, Marshall-Olkin defined a method of generalizing a given probability distribution increasing the number of parameters by one. Assume that X is a r.v. with a given probability distribution with survival function $\overline{F}(x)$, for $-\infty < x < +\infty$. The Marshall-Olkin extension of the initial family is defined to be the family of distributions with survival function equal to:

$$\overline{G}(x) = \frac{\beta \, \overline{F}(x)}{1 - \overline{\beta} \, \overline{F}(x)}, \quad -\infty < x < +\infty, \ \beta > 0, \ \text{and} \ \overline{\beta} = 1 - \beta. \tag{4}$$

The new family contains the initial family as a particular case, obtained when $\beta = 1$. The transformation proposed has an *stability property* in the sense that the result of applying twice the transformation is also in the extended model.

3   The Marshall-Olkin Extended Zipf Distribution

The *Mashall-Olkin Extended Zipf* model (MOEZipf) is defined to be the set of probability distributions with SF:

$$\overline{G}(x; \alpha, \beta) = \frac{\beta \, \overline{F}(X)}{1 - \overline{\beta} \, \overline{F}(X)} = \frac{\beta \, \zeta(\alpha, x + 1)}{\zeta(\alpha) - \overline{\beta} \, \zeta(\alpha + 1)}, \quad \text{for} \quad \beta > 0, \quad \text{and} \quad \alpha > 1, \qquad (5)$$

where $\overline{F}(x)$ is the SF of the Zipf(α) distribution in (2). If Y is a r.v. with a MOEZipf(α, β) distribution, then its probability mass function (pmf) is equal to:

$$\begin{aligned} P(Y = x) & = & \overline{G}(x - 1; \alpha, \beta) - \overline{G}(x; \alpha, \beta) \\ & = & \frac{x^{-\alpha} \, \beta \, \zeta(\alpha)}{[\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x)][\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x + 1)]}, \quad x = 1, 2, 3, \cdots \end{aligned} \qquad (6)$$

In Figure 3 one can see the pmf of the MOEZipf(α, β) distribution for $\alpha = 1.8$ and different values of β. It can be appreciated that the probability at one increases as β tends to zero and decreases as β tends to infinity. This result is proved next, together with some other results that come from comparing the probabilities of the Zipf and the MOEZipf distributions.

Proposition 3.1. *The probability at one of a r.v. Y with a MOEZipf distribution is a decreasing function of β verifying that:*

    a) $P(Y = 1)$ tends to 1 when β tends to 0

    b) $P(Y = 1)$ tends to 0 when β tends to $+\infty$.

Proof: Taking into account that by (3), $\zeta(\alpha, 1) = \zeta(\alpha)$ and that $\zeta(\alpha, 2) = \zeta(\alpha) - 1$, and setting $x = 1$ in (6), one has that:

$$P(Y = 1) = \frac{1}{1 + \beta \, \zeta(\alpha, 2)},$$

which is a decreasing function of β that tends to zero when β tends to infinity and to one when β tends to zero. $\qquad \square$

Proposition 3.2. *For large values of x, parameter β may be interpreted as the ratio between the probabilities of a r.v. Y with a MOEZipf(α, β) distribution and the probabilities of a r.v. X with a Zipf(α) distribution at x.*

Proof: Taking into account that when x tends to infinity $\zeta(\alpha, x)$ tends to zero, one has that:

$$\lim_{x \to +\infty} \frac{P(Y = x)}{P(X = x)} = \lim_{x \to +\infty} \frac{\beta(\zeta(\alpha))^2}{[\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x)][\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x + 1)]} = \beta.$$
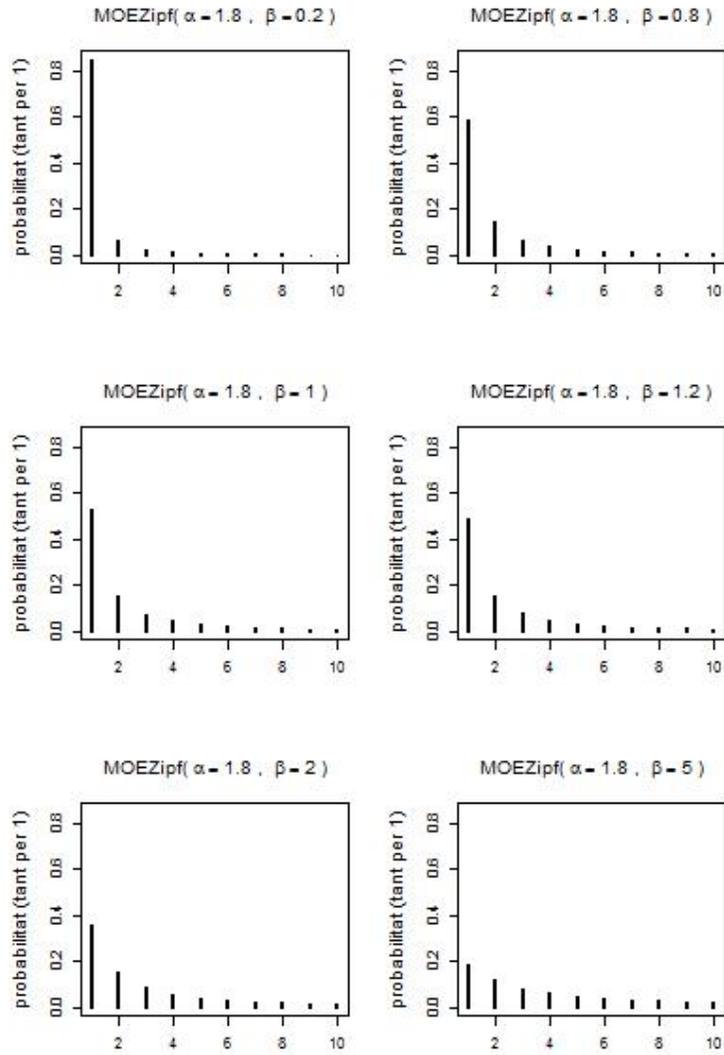
$$\square$$

Figure 1: Pmf's for the MOEZipf(α, β) model for $\alpha = 1.8$ and $\beta = 0.2, 0.8, 1, 1.2, 2$ and 5. For $\beta = 1$, one obtains the Zipf(1.8) distribution.

5

Proposition 3.3. *Let Y be a r.v. with a MOEZipf(α, β) distribution and X be a r.v. with a Zipf(α) distribution. For any $x \geq 1$ one has that: if $\beta = 1$, $P(Y = x) = P(X = x)$, if $\beta > 1$, $\beta P(Y = x) \geq P(X = x)$, and if $\beta < 1$, $P(Y = x) \geq \beta P(X = x)$.*

Proof:

- If $\beta = 1$, the probabilities in (6) correspond to those of the Zipf(α) distribution.

- If $\beta > 1$, $\overline{\beta} < 0$, given that for all $x \geq 1$ $\zeta(\alpha, x) \leq \zeta(\alpha)$, one has that $\overline{\beta}\zeta(\alpha, x) \geq \overline{\beta}\zeta(\alpha)$. Thus,

$$0 \leq \zeta(\alpha) - \overline{\beta}\zeta(\alpha, x) \leq \beta\zeta(\alpha) \quad \Leftrightarrow \quad 0 \leq [\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x)][\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x+1)] \leq (\beta\,\zeta(\alpha))^2$$

$$\Leftrightarrow \quad \frac{x^{-\alpha}\,\beta\,\zeta(\alpha)}{[\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x)][\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x+1)]} \geq \frac{x^{-\alpha}}{\zeta(\alpha)}\frac{1}{\beta}$$

$$\Leftrightarrow \quad P(Y = x) \geq P(X = x)\frac{1}{\beta}.$$

- If $\beta < 1$, $\overline{\beta} > 0$, then $\overline{\beta}\zeta(\alpha, x) \leq \overline{\beta}\zeta(\alpha)$. Thus, for any $x \geq 1$, $0 \leq \zeta(\alpha) - \overline{\beta}\zeta(\alpha, x) \leq \zeta(\alpha)$, which gives that:

$$(\zeta(\alpha))^2 \quad \geq \quad [\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x)][\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x+1)]$$

$$\Leftrightarrow \quad \frac{x^{-\alpha}\beta}{\zeta(\alpha)} \leq \frac{x^{-\alpha}\beta\zeta(\alpha)}{[\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x)][\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x+1)]}$$

$$\Leftrightarrow \quad \beta P(X = x) \leq P(Y = x).$$

□

The MOEZipf distribution is only linear in the log-log scale for large values of x as is proved in the following result:

Proposition 3.4. *Let Y be a r.v. with a MOEZipf(α, β) distribution. For large values of x, $\log(P(Y = x))$ is a linear function of $\log(x)$.*

Proof: The proof is a straightforward consequence of the fact that (3) tends to zero when x tends to infinity. Hence, for large values of x, the denominator of (6) is approximatively equal to $\zeta(\alpha)^2$, and thus one has that:

$$\log(P(Y = x)) \simeq -\alpha \log(x) + \log(\beta) - \log(\zeta(\alpha)). \tag{7}$$

□

REMARK: The result is also true in a larger support of the distribution if α is large. This is because (3) is very small for large values of α if $x \geq 2$.

Figure 3 shows the behaviour of the MOEZipf distribution in the log-log scale, for different parameter values, together with the straight line obtained by changing $\simeq$ by $=$ in (7).

Next proposition compares the ratio of two consecutive probabilities of a MOEZipf and a Zipf distributions with the same α value. As it can be appreciated in Figure 3, the ratio only shows important differences for small values of x. Moreover, for values of β smaller (larger) than one, the ratios corresponding to the MOEZipf distribution are smaller (larger) than the ratios associated to the Zipf distribution. This result is stated in the next proposition.

Proposition 3.5. *If Y is a r.v. with a MOEZipf(α, β) distribution and X is a r.v. with a Zipf(α) distribution, one has that:*

1) If $\beta > (<) 1$, then

$$\frac{P(Y = x + 1)}{P(Y = x)} > (<) \frac{P(X = x + 1)}{P(X = x)}.$$

2) $\forall \beta > 0$,

$$\lim_{x \to +\infty} \frac{P(Y = x + 1)}{P(Y = x)} = \frac{P(X = x + 1)}{P(X = x)}.$$

Proof: By (6) one has that

$$
\begin{aligned}
\frac{P(Y = x + 1)}{P(Y = x)} &= \left(\frac{x}{x+1}\right)^\alpha \frac{\zeta(\alpha) - \overline{\beta}\,\zeta(\alpha, x)}{\zeta(\alpha) - \overline{\beta}\,\zeta(\alpha, x + 2)} \\
&= \frac{P(X = x + 1)}{P(X = x)} \frac{\zeta(\alpha) - \overline{\beta}\,\zeta(\alpha, x)}{\zeta(\alpha) - \overline{\beta}\,\zeta(\alpha, x + 2)}.
\end{aligned}
\tag{8}
$$

Given that for any $\alpha > 1$ and any $x \geq 1$, $\zeta(\alpha, x + 2) < \zeta(\alpha, x)$, it is possible to state that if $0 < \beta < 1$ ($\beta > 1$),

$$1 > (<) \frac{\zeta(\alpha) - (1 - \beta)\zeta(\alpha, x)}{\zeta(\alpha) - (1 - \beta)\zeta(\alpha, x + 2)},$$

which proves point 1). Given that (3) tends to zero when x tends to infinity, point 2) is a straight consequence of (8). $\square$

3.1 Parameter estimation

In this subsection two ways of estimating the parameters of the MOEZipf(α, β) distribution are considered. The maximum likelihood estimator (m.l.e) denoted by $(\hat{\alpha}, \hat{\beta})$,
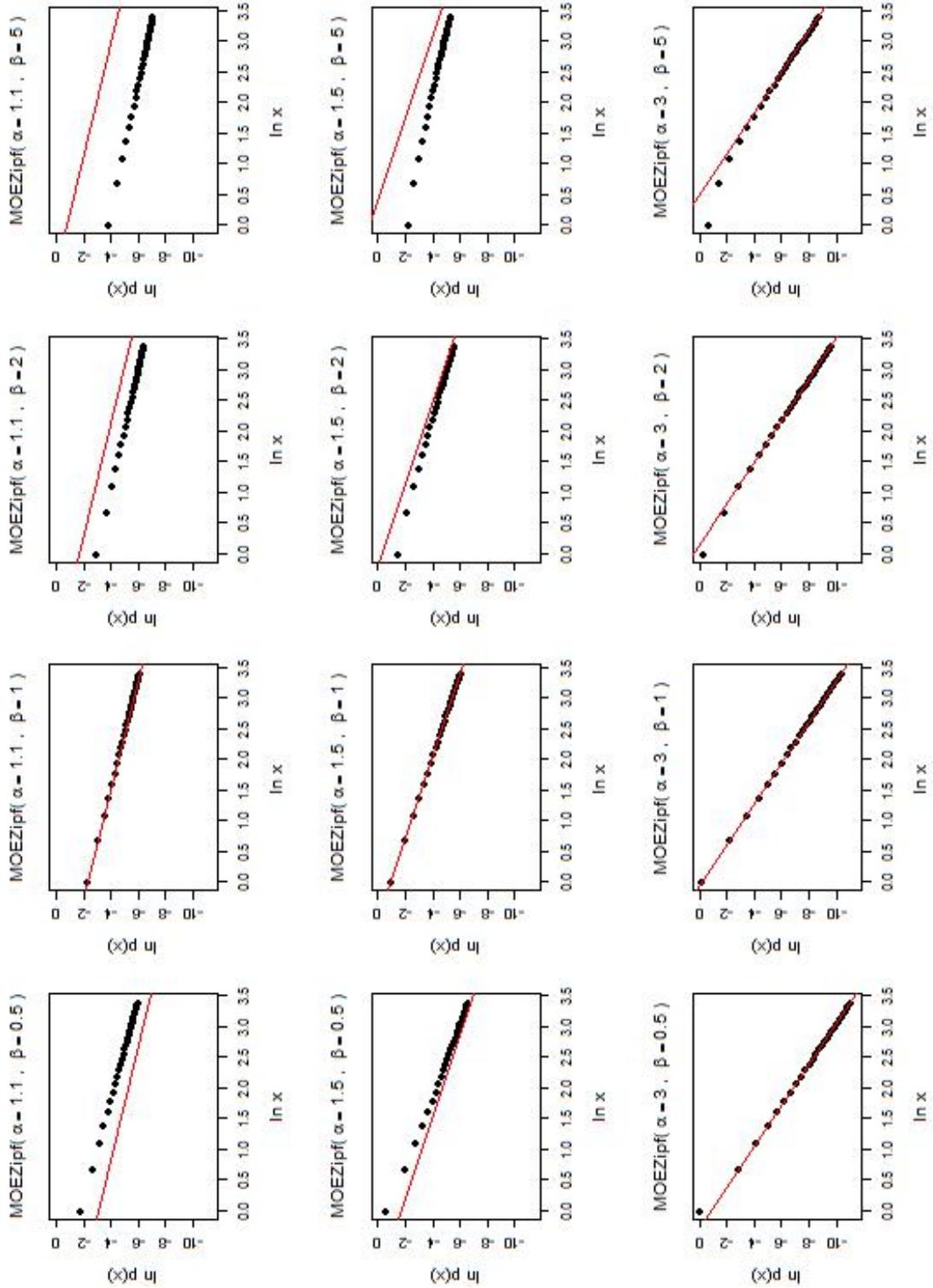
Figure 2: Probabilities of the MOEZipf(α, β) distribution in the log-log scale. Each row corresponds to a different value of α and each column to a different value of β. The parameter values considered are: $\alpha = 1.1, 1.5$ and 3 and $\beta = 0.5, 1, 2$ and 5.
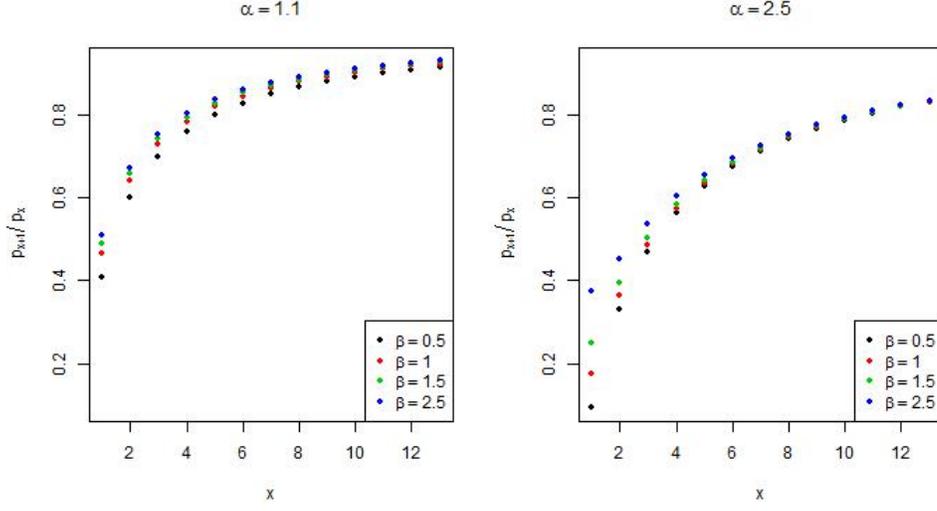
8

Figure 3: Ratio of two consecutive probabilities for the MOEZipf(α, β) and Zipf(α) distributions with α equal to 1.1 and 2.5 and β equal to $0.5, 1, 1.5$ and 2.5.

is obtained by maximizing the corresponding log-likelihood function, that in that case takes the form:

$$
\begin{aligned}
l(\alpha, \beta; y_i) &= n \log(\beta) + n \log(\zeta(\alpha)) - \alpha \sum_{i=1}^{n} \log(y_i) - \sum_{i=1}^{n} log(\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x)) \\
&\quad - \sum_{i=1}^{n} \log(\zeta(\alpha) - \overline{\beta}\zeta(\alpha, x+1)).
\end{aligned}
$$

The second method of estimation considered consists of solving numerically the system of equations that comes from equating the observed and expected probabilities at one, and the sample mean to the expected value of the distribution. If one denotes by f_1 the observed frequency at one, and by $\overline{y}$ the sample mean, that leads to solving

$$
\begin{cases}
(\zeta(\alpha, 2)\,\beta + 1)^{-1} = \frac{f_1}{n} \\
E(Y) = \overline{y},
\end{cases}
\tag{9}
$$

which can be done by first solving the following equation in α:

$$
\frac{n - f_1}{f_1 \cdot \zeta(\alpha, 2)} \zeta(\alpha) \sum_{x=1}^{\infty} \frac{x^{-\alpha+1}}{[\zeta(\alpha) - \frac{f_1 \cdot \zeta(\alpha) - n}{f_1 \cdot \zeta(\alpha, 2)}\zeta(\alpha, x)][\zeta(\alpha) - \frac{f_1 \cdot \zeta(\alpha) - n}{f_1 \cdot \zeta(\alpha, 2)}\zeta(\alpha, x+1)]} = \overline{y},
$$

and then estimating β by substituing the estimation of α in the first equation of (9). The solution obtained by this method is denoted by $(\tilde{\alpha}, \tilde{\beta})$.

4 Data analysis

In this section three sets of data are fitted by means of the Zipf and the MOEZipf models, and the results obtained are compared. All the data sets have a very large sample size and correspond to real data obtained from different areas.

4.1 Example from Linguistics

The data of this example corresponds to the frequency of occurrence of words in the novel *Moby Dick* by Herman Melville and can be found in:

http://tuvalu.santafe.edu/~aaronc/powerlaws/data.htm.

This set of data was first analyzed in [12] which is the reference where the Zipf distribution was defined. More recently [4] and [3] have also considered this set of data. The first ones use the data set to compare real with random texts, and the second ones fit the data by means of a general PL distribution. The set contains the frequencies of a total of 18855 words and nearly 75% of the observations correspond to the first three positive integer values. The three most frequent words appear 14086, 6414 and 6260 times. The observations larger than or equal to 53 have been grouped, in order to be able to compare the two models by means of the χ^2 goodness-of-fit statistic.

The m.l.e estimations of α obtained assuming the two different models do not differ considerably. Nevertheless, the m.l.e of parameter β under the MOEZipf model is 50% larger than under the Zipf model ($\beta = 1$). The reduction obtained in the χ^2 Pearson statistic using the proposed model instead of the Zipf model is equal to 79.64%. The AIC as well as the log-likelihood show that the generalized model estimating the parameters by maximum likelihood gives the best fit. In Figure 4.1 one can see the data together with the two fitted models in a log-log scale.

Distrib.	Param.	Estimat.	log-like.	X^2	p-val.	AIC
Zipf	$\hat{\alpha}$	1.775	-40196.00	272.38	0	80394.00
MOEZipf	$\tilde{\alpha}$	1.908	-40086.28	62.96	0.097	80176.56
(2nd method)	$\tilde{\beta}$	1.429				
MOEZipf	$\hat{\alpha}$	1.944	**-40082.42**	**55.45**	0.293	**80168.83**
(m.l.e)	$\hat{\beta}$	1.523				

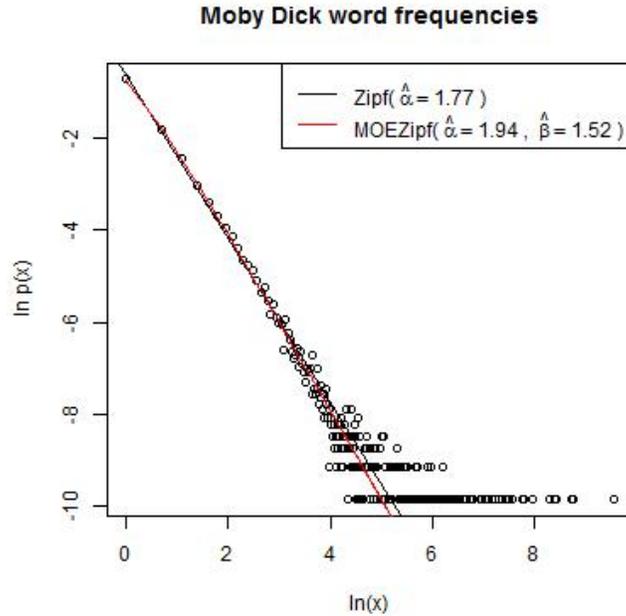Table 1: Reults of fitting the variable: *Frequency of occurence of words*, in *Moby Dick*.

Figure 4: Observed and expected data in the log-log scale, for the word frequency data.

4.2 Example from electronic mail

Given a database containing different electronic mail addresses, one can count how many connections one address has had in a given period of time. The table of frequencies of such a r.v. tends to have large probability at one (most of the addresses only have one contact), and a very small probability at some large values (just few addresses have lots of contacts). The data set analyzed in this example corresponds to the number of connections of a total of 225409 electronic addresses, and may be found in http://snap.stanford.edu/data/email-EuAll.html. They were collected between october 2003 and may 2005. In [8] this data set is analyzed by fitting a PL distribution in the tail of the distribution.

Here 85% of the observations are equal to one. The observed probabilities of the first five values decrease very quickly and after these values, they decrease more slowly. The three addresses with the largest number of contacts have exactly 930, 871 and 854 contacts. After grouping the data larger than 65, the χ^2 statistic is reduced in a 93.74% by using the generalized model instead of the original by m.l. The AIC criterium concludes that the MOEZipf model is the best one.

Figure 4.2 shows the observed and fitted probabilities in the log-log scale. It can be appreciated that the convex behaviour of the MOEZipf model gives place to a better fit, not only for the first values of the distribution but also for the values in the tail.

Distrib.	Param.	Estimat.	log-like.	X^2	p-val.	AIC
Zipf	$\hat{\alpha}$	2.968	-156765.21	13714.84	0	313532.42
MOEZipf	$\tilde{\alpha}$	2.126	-154526.75	968.34	0	309057.51
(2nd method)	$\tilde{\beta}$	0.321				
MOEZipf	$\hat{\alpha}$	2.284	**-154399.82**	**858.27**	0	**308803.64**
(m.l.e.)	$\hat{\beta}$	0.390				

Table 2: Results of fitting the variable: *Number of relations by electronic mail.*
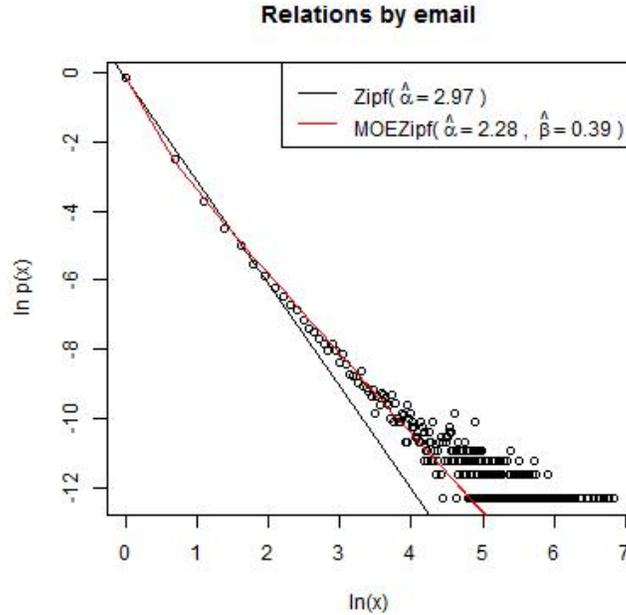
Figure 5: Observed and expected data in the log-log scale, for the e-mail example.

4.3   Example from citations

The last example considered corresponds to the number of times that a given paper is cited in a given database. This is an important variable because it allows one to calculate the impact factor of the scientific journals. The database analyzed has a total of 32158 papers in the area of *High-energy physics*, published in arXiv.org between January 1993 and April 2003, and may be found in http://snap.stanford.edu/data/citHepPh.html. This data set has also been analyzed in [8].

From that data one observes that the 26% of the probability corresponds to the first two values, meaning that more than a quarter of the papers are cited at most twice. For model fitting, we have grouped the values larger than 119. As in the previous examples, Table 3 indicates that the MOEZipf model estimating by m. l. provides the

best fit. The inclusion of the β parameter implies a reduction of a 87.73% in the χ^2 statistic. It is important to note that $\hat{\beta}$ is equal to 13.1 which is a very large value, compared with the value of 1 that corresponds to the Zipf distribution. In Figure 4.3 it is possible to appreciate that the MOEZipf model shows a concave behaviour that improves considerably the fit of the first values as well as the one in the tail of the distribution.

Distrib.	Param.	Estim.	log-like	X^2	p-val.	AIC
Zipf	$\hat{\alpha}$	1.421	-105839.81	13172.05	0	211681.61
MOEZipf	$\tilde{\alpha}$	2.214	-99490.01	1615.25	0	198984.03
(2nd met.)	$\tilde{\beta}$	11.677				
MOEZipf	$\hat{\alpha}$	2.161	**-99197.93**	**816.62**	0	**198399.87**
(m.l.e)	$\hat{\beta}$	13.058				

Table 3: Results of fitting the r.v.: *Number of citations of a given paper.*
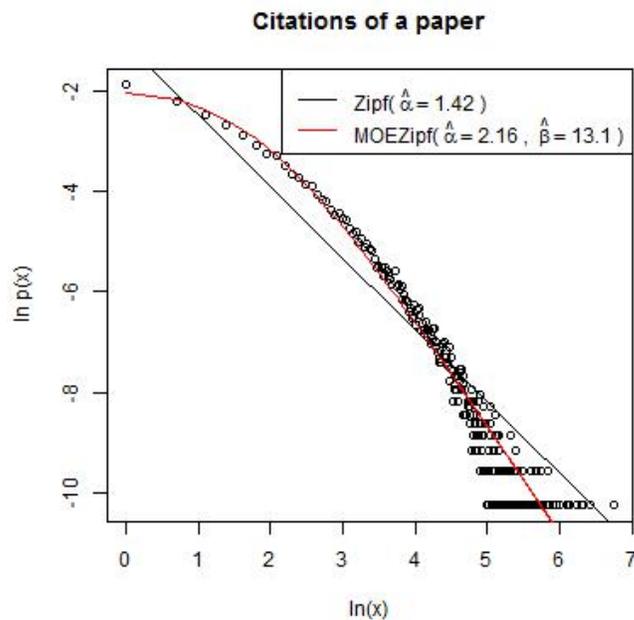

Figure 6: Observed and expected data in the log-log scale, for the citations example.

5 Conclusions

The Marshall-Olkin transformation has proved to be useful for generalizing the Zipf distribution both in terms of providing good properties as well as in terms of improving the goodness of fit obtained in the data sets analyzed. The extended model can show

concavity or convexity in the first part of the domain, as it is shown by means of the examples. The linear behaviour is always observed in the tail of the distribution. The extra-parameter also allows for ratios between two consecutive probabilities larger or smaller than the corresponding ratio of a Zipf distribution. In the three data sets considered, the fittings obtained for the first values are considerably better than the ones corresponding to the Zipf distribution, but they are also better in the tail. The reduction in the χ^2 goodness-of-fit statistic has always been larger or equal than 80%. The AIC points out the MOEZipf model as the better model in all the examples considered.

References

[1] T. Alice and K. Jose (2005). Marshall-Olkin semi-weibull minification processes. *Recent Advances in Statistical Theory and Applications*, I pp. 6-17.

[2] L.A. Adamic and B.A. Huberman (2002) Zipf's law in internet, *Glottometrics*, vol 3, pp 143-150.

[3] A. Clausset, C.R. Shalizi, and M.E. Newman (2009). Power-law distributions in empirical data, *SIAM Review*, vol. 51 pp. 661-703.

[4] R. Ferrer and R.V. Solé (2002). Zipf's law and random texts. *Advances in Complex Systems*, vol 5, pp. 1-6.

[5] M.E. Ghitany, E.K. Al-Hussaini, and R.A. Al-Jarallah, (2005). Marshall-Olkin extended Weibull distribution and its application to censored data, *Journal of Applied Statistics*, vol. 32, pp. 1025-1034.

[6] M.E. Ghitany, F.A. Al-Awadhi, and L.A. Alkhalfan, (2007). Marshall-Olkin Extended Lomax distribution and its application to censored data. *Communications in Statistics-Theory and Methods*, vol. 36, pp 1855-1866.

[7] W. Gui (2013). A Marshall-Olkin Power Log-normal distribution and its application to survival data. *International Journal of Statistics and Probability*, vol. 2. (No.1) pp. 63-71.

[8] J. Leskovec, (2008) *Dynamics of Large Metworks*. PhD thesis, School of Computer Science, Carnegie Mellon University.

[9]   A.W. Marshall and I. Olkin (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families, *Biometrika*, vol 84, pp 641-652.

[10] M.E.J. Newman (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, vol. 46 pp. 323-351.

[11] F.J. Rubio and F.J. Mark (2012). On the Marshall-Olkin transformation as a skewing mechanism. *Computational Statistics and Data Analysis*, vol. 56 (No. 7) pp. 2251-2257.

[12] G.K. Zipf, (1949). *Human behaviour and the principle of least effort*, Addison-Wesley Press.