

## STOCHASTIC ORDERINGS AND PARAMETER ESTIMATION FOR THE HURWITZ-LERCH ZETA DISTRIBUTION

K.W. Liew<sup>1,2</sup> and S. H. Ong<sup>2</sup>

<sup>1</sup>Department of Mathematical and Actuarial Sciences, Universiti Tunku Abdul Rahman  
Jalan Genting Kelang, Setapak, 53300 Kuala Lumpur. Malaysia.

<sup>2</sup>Institute of Mathematical Sciences, Faculty of Science, University of Malaya  
50603 Kuala Lumpur. Malaysia

**ABSTRACT** The Hurwitz-Lerch Zeta (HLZ) family includes many well-known distributions such as the logarithmic distribution, Zipf-Mandelbrot distribution and so on. In this study, the stochastic orderings of the random variables in the HLZ family are established based on the likelihood ratio. These orderings provide a useful tool for comparing the 'magnitudes' of these random variables. The tail probability of the HLZ distribution is shown to have an interesting relation with a generalisation of the logarithmic distribution (GLD) proposed in [1]. To demonstrate the flexibility of the HLZ distribution in empirical modelling, a robust probability generating function (pgf) based estimation method using Hellinger-type divergence is implemented in data-fitting and the results are compared with various other GLD's. An augmented pgf is constructed to overcome the difficulties of this estimation method when some data are grouped.

**ABSTRAK** Keluarga Hurwitz-Lerch Zeta (HLZ) mengandungi beberapa taburan yang ternama seperti taburan logaritma, taburan Zipf-Mandelbrot dan lain-lain. Dalam kajian ini, susunan stokastik pembolehubah-pembolehubah rawak dalam keluarga HLZ telah ditubuhkan berdasar nisbah kebolehjadian. Susunan tersebut boleh digunakan dalam perbandingan 'magnitud' antara pembolehubah-pembolehubah rawak tersebut. Hubungan antara kebarangkalian ekor bagi taburan HLZ dan taburan logaritma teritlak yang dicadangkan dalam [1] terus ditunjukkan. Demi menggambarkan kelenturan taburan HLZ, satu kaedah penganggaran teguh yang menggunakan divergen jenis Hellinger berasaskan fungsi penjana kebarangkalian (fpk) telah dilaksanakan dalam analisis data dan keputusannya dibanding dengan pelbagai generalisasi taburan logaritma yang lain. Fpk ditambah telah dibinakan untuk menyelesaikan masalah pelaksanaan kaedah ini apabila terdapat data yang terkumpul.

**(Keywords:** Generalised logarithmic distribution, Hellinger-type divergence, maximum likelihood, probability generating function based estimation, Zipf-Mandelbrot distribution)

---

### INTRODUCTION

The Zipf's law or the power law that describes decay which is slower than the exponential rate has been applied to various fields where ranking of sizes or frequencies are of particular interest, such as quantitative linguistics [2], human demography [3], economics [4], citations [5] and internet traffic modelling [6]. Discrete distributions that are related to the power law include the Zipf-Mandelbrot distribution, Lotka distribution and so on. A unified representation of these distributions is discussed under the setting of Hurwitz-Lerch Zeta (HLZ) distribution in [7]. The reliability properties of the HLZ distribution were examined in [8] and a set of data was fitted to the model by maximum

likelihood procedure using global optimization technique.

In this paper, we consider further properties of the HLZ distribution. We study the stochastic orderings [9] among the members in this family and their tail behaviours. Stochastic orders have found applications in many diverse areas such as reliability theory and survival analysis, life sciences, operations research, insurance, finance and actuarial science, economics, and management science. A probability generating function (pgf) based estimation method proposed in [10] will be applied in data-fitting and a data graduation scheme is proposed to illustrate the feasibility of their method even when some data are grouped. The paper is organized as follows. Section 2 defines the HLZ distribution while

Section 3 discusses stochastic ordering. The tail behavior of the HLZ distribution is examined in Section 4. Section 5 considers parameter estimation with the pgf based method. An augmented pgf is constructed to overcome the difficulties of this estimation method when some data are grouped. A number of data sets have been used to illustrate the parameter estimation and goodness-of-fit: (1) the distribution of number of moth species represented by  $n$  individuals in a sample from the lightly logged rainforest [1]; (2) distribution of 1534 biologists according to the number of research papers to their credit [11]; (3) number of boards that contains at least one sowbug [12]; (4) Corbet's Malayan butterflies data [8]. Section 6 gives the conclusion.

**The Hurwitz-Lerch Zeta distribution**

The HLZ distribution has probability mass function  $p(k) = \frac{1}{T(\theta, s, a)} \frac{\theta^k}{(k+a)^{s+1}}$ ,  $k = 1, 2, \dots$ ;

$$p_X(z) = E[z^X] = \frac{T(\theta z, s, a)}{T(\theta, s, a)} = \frac{z\Phi(\theta z, s+1, a+1)}{\Phi(\theta, s+1, a+1)}, \text{ for } 0 < \theta z \leq 1.$$

A list of the members in the HLZ family and their corresponding parameters can be found in [13]. Some common distributions in the HLZ family are shown in **Table 1**.

Table 1 Some common distributions in the HLZ family

Distribution	$\theta$	$s$	$a$
Lotka	1	1	0
Riemann zeta	1	$s$	0
Zipf-Mandelbrot	1	$s$	$a$
Good	$\theta$	$s$	0
logarithmic	$\theta$	0	0

**Stochastic orderings of HLZ random variables**

Stochastic orders can be defined on random variables in different ways depending on the purpose of research. The usual stochastic ordering, which tells whether a random variable is more likely to take smaller values compared to another, is useful, for example, in comparing risks. Some other orderings may be defined

where  $0 < \theta \leq 1$ ;  $s \in \mathbb{R}$ ;  $a > -1$  and  $s > 0$  when  $\theta = 1$ . Here, we use the notations

$$T(\theta, s, a) = \sum_{k=1}^{\infty} \frac{\theta^k}{(k+a)^{s+1}} = \theta\Phi(\theta, s+1, a+1)$$

and

$$\Phi(z, s, a) = \sum_{k=0}^{\infty} \frac{z^k}{(k+a)^s},$$

is the Lerch transcendent defined for  $a \notin \{-1, -2, \dots\}$  and  $s \in \mathbb{C}$  when  $|z| < 1$  or  $\text{Re } s > 1$  when  $|z| = 1$ .

Let  $X$  be an HLZ random variable. The mean of the distribution is  $\frac{T(\theta, s-1, a)}{T(\theta, s, a)} - a$  and its pgf can be written as

based on the hazard rate, reverse hazard rate or transforms of the random variables, see [9].

Consider two random variables  $X$  and  $Y$  with probability functions  $f(x)$  and  $g(x)$  respectively. Denote by  $F(x)$  and  $G(x)$  the respective cumulative distribution function of  $X$  and  $Y$ , and define their reliability functions to be  $\bar{F}(x) = 1 - F(x)$  and  $\bar{G}(x) = 1 - G(x)$ . We say that  $X$  is smaller than  $Y$  in the usual stochastic order (written  $X \leq_{st} Y$ ) if  $\bar{F}(x) \leq \bar{G}(x)$  for all  $x$ .  $X$  is said to be smaller than  $Y$  in the hazard rate order (written  $X \leq_{hr} Y$ ) if  $\bar{F}(x)\bar{G}(y) \geq \bar{F}(y)\bar{G}(x)$  for all  $x \leq y$ . Similarly,  $X$  is said to be smaller than  $Y$  in the reverse hazard rate order (written  $X \leq_{rh} Y$ ) if  $F(x)G(y) \geq F(y)G(x)$  for all  $x \leq y$ . On the other hand, if  $f(x)g(y) \geq f(y)g(x)$  for

all  $x \leq y$  or equivalently  $\frac{f(x)}{g(x)}$  is monotone

decreasing in  $x$ , then we say  $X$  is smaller than  $Y$  in the likelihood ratio order (written  $X \leq_{lr} Y$ ). The likelihood ratio order is important in the sense that it implies the previous three orderings [9] and is easy to verify. The relations between the stochastic orders are summarized in [14] in a diagram partially shown in **Figure 1**.

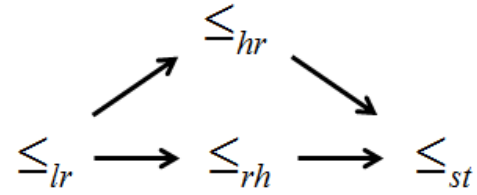


Figure 1. Relationships among the stochastic orders

The likelihood ratio order of the HLZ members will be considered in three different cases depending on the number of different parameters between the models.

Case 1: When one parameter is different.

Let  $\theta_1 < \theta_2$ , then  $\frac{\frac{1}{T(\theta_1, s, a)} \frac{\theta_1^k}{(k+a)^{s+1}}}{\frac{1}{T(\theta_2, s, a)} \frac{\theta_2^k}{(k+a)^{s+1}}} = \frac{T(\theta_2, s, a)}{T(\theta_1, s, a)} \left(\frac{\theta_1}{\theta_2}\right)^k$  is a decreasing function in  $k$ .

Hence  $HLZ(\theta_1, s, a) <_{lr} HLZ(\theta_2, s, a)$

When  $a_1 < a_2$ , the ratio  $\frac{\frac{1}{T(\theta, s, a_1)} \frac{\theta^k}{(k+a_1)^{s+1}}}{\frac{1}{T(\theta, s, a_2)} \frac{\theta^k}{(k+a_2)^{s+1}}} = \frac{T(\theta, s, a_2)}{T(\theta, s, a_1)} \left(1 + \frac{a_2 - a_1}{k + a_1}\right)^{s+1}$  is a decreasing

function in  $k$  when  $s \geq -1$ . Hence,  $HLZ(\theta, s, a_1) <_{lr} HLZ(\theta, s, a_2)$  for  $s \geq -1$ .

Let  $s_2 < s_1$ , then  $\frac{\frac{1}{T(\theta, s_1, a)} \frac{\theta^k}{(k+a)^{s_1+1}}}{\frac{1}{T(\theta, s_2, a)} \frac{\theta^k}{(k+a)^{s_2+1}}} = \frac{T(\theta, s_2, a)}{T(\theta, s_1, a)} (k+a)^{s_2-s_1}$  is a decreasing function in

$k$  implying  $HLZ(\theta, s_1, a) <_{lr} HLZ(\theta, s_2, a)$ .

Case 2: When two parameters are different.

Consider  $\theta$  is common, then  $\frac{\frac{1}{T(\theta, s_1, a_1)} \frac{\theta^k}{(k+a_1)^{s_1+1}}}{\frac{1}{T(\theta, s_2, a_2)} \frac{\theta^k}{(k+a_2)^{s_2+1}}} = \frac{T(\theta, s_2, a_2)}{T(\theta, s_1, a_1)} \frac{(k+a_2)^{s_2+1}}{(k+a_1)^{s_1+1}}$ . Clearly

the ratio is a decreasing function of  $k$  if

$$\frac{(k+a_2)^{s_2+1}}{(k+a_1)^{s_1+1}} > \frac{(k+1+a_2)^{s_2+1}}{(k+1+a_1)^{s_1+1}},$$

that is,

$$\left(\frac{k+a_2}{k+1+a_2}\right)^{s_2+1} > \left(\frac{k+a_1}{k+1+a_1}\right)^{s_1+1}.$$

Hence,  $HLZ(\theta, s_1, a_1) <_{lr} HLZ(\theta, s_2, a_2)$  if, and only if,  $s_2 \leq s_1$  and  $a_2 \geq a_1$ .

If  $a$  is common, then 
$$\frac{\frac{1}{T(\theta_1, s_1, a)} \frac{\theta_1^k}{(k+a)^{s_1+1}}}{\frac{1}{T(\theta_2, s_2, a)} \frac{\theta_2^k}{(k+a)^{s_2+1}}} = \frac{T(\theta_2, s_2, a)}{T(\theta_1, s_1, a)} \left(\frac{\theta_1}{\theta_2}\right)^k (k+a)^{s_2-s_1}.$$

Clearly the ratio is a decreasing function of  $k$  if  $\theta_1 \leq \theta_2$  and  $s_2 \leq s_1$ . It is a strictly decreasing function of  $k$  if  $\theta_1 < \theta_2$  and  $s_2 < s_1$ . We have  $HLZ(\theta_1, s_1, a) <_{lr} HLZ(\theta_2, s_2, a)$ .

When  $s$  is common, then 
$$\frac{\frac{1}{T(\theta_1, s, a_1)} \frac{\theta_1^k}{(k+a_1)^{s_1+1}}}{\frac{1}{T(\theta_2, s, a_2)} \frac{\theta_2^k}{(k+a_2)^{s_2+1}}} = \frac{T(\theta_2, s, a_2)}{T(\theta_1, s, a_1)} \left(\frac{\theta_1}{\theta_2}\right)^k \left(\frac{k+a_2}{k+a_1}\right)^{s+1}$$
 and

the ratio is a decreasing function of  $k$  if  $\theta_1 \leq \theta_2$  and  $a_1 \geq a_2$ . Hence,  $HLZ(\theta_1, s, a_1) <_{lr} HLZ(\theta_2, s, a_2)$ .

Case 3: When all parameters are different.

Consider the ratio 
$$\frac{\frac{1}{T(\theta_1, s_1, a_1)} \frac{\theta_1^k}{(k+a_1)^{s_1+1}}}{\frac{1}{T(\theta_2, s_2, a_2)} \frac{\theta_2^k}{(k+a_2)^{s_2+1}}} = \frac{T(\theta_2, s_2, a_2)}{T(\theta_1, s_1, a_1)} \frac{(k+a_2)^{s_2+1}}{(k+a_1)^{s_1+1}} \left(\frac{\theta_1}{\theta_2}\right)^k.$$

$$HLZ(\theta_1, s_1, a_1) <_{lr} HLZ(\theta_2, s_2, a_2)$$

By applying the results in Case 2, this ratio is a decreasing function of  $k$  if  $\theta_1 \leq \theta_2$ ,  $a_1 \geq a_2$  and  $s_2 \leq s_1$ . Thus, based on the discussion above, we can now rank some of the distributions in **Table 1**. For example, by assuming  $0 < s < 1$ ,  $a > 0$  and  $\theta < 1$  to have same values in different models, we have:

Good  $\leq_{lr}$  logarithmic series  $\leq_{lr}$  Lotka (if  $2\theta \leq 1$ )  $\leq_{lr}$  zeta  $\leq_{lr}$  Zipf-Mandelbrot.

**Tail behaviour of HLZ distribution**

The rate of decay of a probability distribution determines whether the distribution has a short or long tail. Usually this can be assessed by finding the limit of the ratio of consecutive probabilities. For the HLZ distribution, we have

$$\lim_{k \rightarrow \infty} \frac{p(k+1)}{p(k)} = \lim_{k \rightarrow \infty} \left(\frac{a+k}{a+k+1}\right)^{s+1} \theta = \theta$$

When  $\theta = 1$ ,  $p(k) \propto (a+k)^{-(s+1)}$  giving the distribution the long-tail property. It shows that the HLZ exhibits either exponential decay or power law decay.

The upper tail probability of a random variable is often of interest. Define  $T(x) = \sum_{i=x}^{\infty} p(i)$ , and

for the HLZ case, the tail probability can be expressed in terms of the Lerch transcendent as follows.

$$\begin{aligned}
 T_{HLZ}(x) &= \frac{1}{T(\theta, s, a)} \sum_{i=x}^{\infty} \frac{\theta^i}{(a+i)^{s+1}} \\
 &= \frac{\theta^x}{T(\theta, s, a)} \sum_{i=0}^{\infty} \frac{\theta^i}{(a+x+i)^{s+1}} \\
 &= \frac{\theta^{x-1} \Phi(\theta, s+1, a+x)}{\Phi(\theta, s+1, a+1)}
 \end{aligned}$$

Compare this to the asymptotic upper tail probability of the GLD in [1] which takes the form

$$T_{GLD}(x) \approx -\frac{(q+2\sqrt{pr})^{x+1/2}}{2(\pi\sqrt{pr})^{1/2} \log p} \Phi(q+2\sqrt{pr}, \frac{3}{2}, x)$$

. By taking  $\theta = q + 2\sqrt{pr}$ ,  $s = 0.5$  and  $a = 0$ , it is interesting to see that the asymptotic ratio of the tail probabilities of both distributions can be made independent of  $x$ , that is,

$$\frac{T_{GLD}(x)}{T_{HLZ}(x)} \approx -\frac{\Phi(q+2\sqrt{pr}, 1.5, 1)(q+2\sqrt{pr})^{3/2}}{2(\pi\sqrt{pr})^{1/2} \log p}$$

Figure 2 shows the graph of the above ratio over all possible combinations of  $p$ ,  $q$  and  $r$  ( $p + q + r = 1$ ), we see that the ratio is within the range from 1 to less than 4.5.

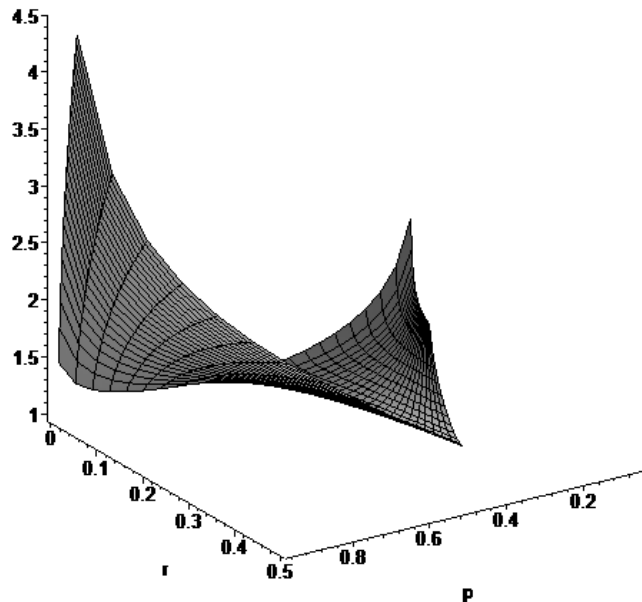


Figure 2. Ratio of tail probabilities of GLD to HLZ for all possible combinations of  $p$  and  $r$

### Estimation of the HLZ parameters

In [8], the maximum likelihood equations for the HLZ parameters are derived and the authors pointed out that the equations result in method of moment equations with different functional forms. However, closed form solution of these equations is unavailable and a global optimization technique is needed to find the maximum likelihood estimates. Following the development of computing technology, more robust estimation methods such as those that minimize the Hellinger-type divergence has become increasingly popular. In [10], an

estimation procedure by minimizing the pgf-based divergence statistic

$$T(\theta, \alpha, n) = \int_0^1 (f_n(t)^\alpha - g(t; \theta)^\alpha)^2 \beta(t) dt, \quad 0 < \alpha \leq 1$$

is proposed, where  $f_n(t)$  is the empirical pgf,  $g(t; \theta)$  is the pgf with parameter vector  $\theta$  and  $\beta(t)$  is a weight function. For simplicity and shorter execution time, the statistics

$$T_1(\theta, n) = \int_0^1 (f_n(t) - g(t; \theta))^2 dt$$

$$\text{and } T_2(\boldsymbol{\theta}, n) = \int_0^1 \left( \sqrt{f_n(t)} - \sqrt{g(t; \boldsymbol{\theta})} \right)^2 dt$$

will be used in this section to estimate the parameters of the HLZ distribution. A global optimization method that searches over the parameter space for optimal solution has been implemented and the integrals are evaluated by a 6-point Gaussian quadrature. Datasets that have been fitted to various generalizations of the logarithmic distribution will be used for comparison purpose. The performances of the estimates can be assessed based on the  $p$ -value.

In [1], the distribution of number of moth species represented by  $n$  individuals in a sample from the lightly logged rainforest was fitted to a GLD arising from the inverse trinomial distribution and giving a good fit with  $p$ -value of 0.77. We may expect the HLZ to give a good fit to the data following the discussion in Section 4. Using the same grouping, our result, obtained by minimizing  $T_1(\boldsymbol{\theta}, n)$ , shows adequate fit with a slightly smaller  $p$ -value of 0.66 as shown in **Error! Reference source not found..**

Table 2 Distribution of number of moth species represented by  $n$  individuals in a sample from the lightly logged rainforest fitted to the HLZ distribution by minimizing  $T_1(\boldsymbol{\theta}, n)$ .

$N$	Number of species	HLZ
1	140	138.75
2	36	39.21
3	17	17.65
4	13	9.60
5	6	5.78
6	2	3.72
7	4	
8	2	1.74
9	2	
10	0	0.90
11	1	0.67
12	0	0.50
13	0	
14	0	0.29
15	1	0.22
>15	0	0.84
Total	224	224
$df$		3
$\chi^2$		1.61
$p$ -value		0.6561

$$\hat{\theta} = 0.8600; \hat{s} = -0.0173 \text{ and } \hat{a} = 0.5856$$

the same set of data by using another generalization of logarithmic distribution and get a  $p$ -value of 0.323. Our fits, obtained by minimizing  $T_1(\boldsymbol{\theta}, n)$  and  $T_2(\boldsymbol{\theta}, n)$  are marginal with  $p$ -values of 0.1359 and 0.1117 respectively, see **Table 3**.

**Table 4** shows the data of number of boards that contains at least one sawbug which is fitted to the Good distribution in [12] by using maximum likelihood estimation and a quadratic distance estimator. Their reported  $p$ -values for the two estimates are 0.5351 and 0.117 respectively

whereas minimizing  $T_1(\theta, n)$  and  $T_2(\theta, n)$  respectively.  
 gives  $p$ -values of 0.4819 and 0.4543

Table 3 Distribution of 1534 biologists according to the number of research papers to their credit fitted to the HLZ distribution by minimizing  $T_1(\theta, n)$  and  $T_2(\theta, n)$

Number of papers per author	Observed Frequencies	$T_1(\theta, n)$	$T_2(\theta, n)$
1	1062	1056.52	1057.44
2	263	283.80	282.75
3	120	104.47	103.84
4	50	44.90	44.76
5	22	21.17	21.25
6	7	10.61	10.77
7	6	5.56	5.71
8	2	3.01	3.14
9	0	1.67	1.77
10	1	0.95	1.02
11	1	0.55	0.60
>11	0	0.80	0.93
Total	1534	1534.00	1534.00
$df$		4	4
$\chi^2$		7.00	7.50
$p$ -value		0.1359	0.1117
$\theta$		0.6814	0.7050
$s$		0.8443	1.0184
$a$		0.5232	0.6217

Table 4 Number of boards that contains at least one sowbug fitted to Good's distribution by minimizing  $T_1(\theta, n)$  and  $T_2(\theta, n)$

Number of sowbug	Observed	$T_1(\theta, n)$	$T_2(\theta, n)$
1	28	26.91	27.24
2	14	16.29	16.12
3	11	11.32	11.12
4	8	8.33	8.16
5	11	6.32	6.20
6	2	4.89	4.81
7	3	{ 3.84	{ 3.79
8	3	{ 3.04	{ 3.02
9	3	{ 2.43	{ 2.43
10	3	{ 1.95	{ 1.96
11	2	{ 1.58	{ 1.59
12	0	{ 1.28	{ 1.30
13	1	{ 1.04	{ 1.07
14	2	{ 0.85	{ 0.88
15	1	{ 0.69	{ 0.72
16	0	{ 0.57	{ 0.60
17	2	{ 0.47	{ 0.49
>18	0	{ 2.21	{ 2.49
Total	94	94.00	94.00
$df$		7	7
$\chi^2$		6.51	6.76
$p$ -value		0.4819	0.4543
$\theta$		0.8450	0.8551
$s$		-0.5186	-0.4691

**Table 5** consists of the Corbet's Malayan butterflies data that was fitted to the HLZ distribution in [8] by using maximum likelihood estimates. An immediate problem in implementing the pgf based estimation method is that classes larger than 24 has been grouped into one class. If we ignore the last class completely and minimizing the divergence between the 'empirical pgf constructed only from the first 24 classes' and the 'pgf that only consists of 24 terms', the deviance is very large. If we assume the 119 grouped observations to be uniformly distributed with one observation in each class after 24, the fit is still not satisfactory. The arbitrariness in handling the grouped data is a

disadvantage in the pgf based estimation. This drawback can be removed by using a more objective data graduation scheme as shown below.

When  $s$  is small, the ratios between the expected frequencies from consecutive classes in the HLZ distribution are approximately  $\theta$ . In finding the optimal solution, for each new random search of  $\theta$ , we make use of the observation from class 24 and construct the sequence  $\{[3\theta], [3\theta^2], \dots\}$ , where  $[x]$  refer to  $x$  rounded to the nearest positive integer. The process continues until the sum of the sequence



Table 5 Corbet's Malayan butterflies data fitted to the HLZ distribution based on three different ways to construct the empirical pgf

Individuals/Species	Number of Species	Truncating the tail at 24	Assuming 1 observation in each class after 24	Graduation of data using HLZ ratio
1	118	119.14	117.75	118.93
2	74	67.25	68.89	65.64
3	44	45.87	47.33	45.67
4	24	34.36	35.42	35.03
5	29	27.22	27.96	28.37
6	22	22.39	22.88	23.78
7	20	18.91	19.24	20.42
8	19	16.30	16.50	17.84
9	20	14.27	14.38	15.80
10	15	12.65	12.69	14.14
11	12	11.33	11.31	12.77
12	14	10.23	10.18	11.61
13	6	9.31	9.23	10.62
14	12	8.52	8.42	9.76
15	6	7.85	7.72	9.02
16	9	7.26	7.12	8.36
17	9	6.74	6.59	7.78
18	6	6.28	6.13	7.25
19	10	5.87	5.72	6.79
20	10	5.51	5.35	6.36
21	11	5.18	5.02	5.98
22	5	4.88	4.72	5.63
23	3	4.61	4.45	5.31
24	3	4.37	4.21	5.02
>24	119	143.70	140.79	122.13
Total	620	620.00	620.00	620.00
<i>df</i>		21	21	21
$\chi^2$		31.26	31.88	20.32
<i>p</i> -value		0.0695	0.0602	0.5011
$\theta$		0.9924	0.9942	0.9803
<i>s</i>		0.1340	0.2339	-0.1311
<i>a</i>		0.5508	0.8630	0.0674

is 119. This sequence will be used to represent the observations in classes larger than 24 when constructing the empirical pgf. Using the augmented empirical pgf, we manage to get estimates that greatly improved the fit as shown in the last column in **Table 5**. The smaller chi-square value compared to that in [8] could also be due to our extended parameter space to include negative values for  $s$ .

### CONCLUSION

As shown in the numerical examples, the HLZ family is a very flexible distribution comparable to other GLD's and is able to accommodate variations that are covered in other models as well. The likelihood ratio orders between the members in the HLZ family have been established to allow easy deduction of other orders when conditions are satisfied. The pgf based estimation method proposed by [10] (see [16] for the implementation in the multivariate case) performed well with the HLZ distribution. The fits obtained are comparable to that by the maximum likelihood method as shown in **Table 4**. A simple data graduation scheme is proposed to avoid the arbitrariness in handling grouped data when constructing the empirical pgf and the result obtained is encouraging.

### REFERENCES

1. Khang, T.F. and S.H. Ong (2007). A new generalization of the logarithmic distribution arising from the inverse trinomial distribution. *Communications in Statistics-Theory and Methods*. **36**(1):3-21.
2. Calderon, F., S. Curilef, and M.L. Ladron de Guevara (2009). Probability distribution in a quantitative linguistic problem. *Brazilian Journal of Physics*. **39**(2A):500-502.
3. Gan, L., D. Li, and S. Song (2006). Is the Zipf law spurious in explaining city-size distributions? *Economics Letters*. **92**(2):256-262.
4. Zhang, J., Q. Chen, and Y. Wang (2009). Zipf distribution in top Chinese firms and an economic explanation. *Physica A: Statistical Mechanics and its Applications*. **388**(10):2020-2024.

5. Perc, M. (2010). Zipf's law and log-normal distributions in measures of scientific output across fields and institutions: 40 years of Slovenia's research as an example. *Journal of Informetrics*. **4**(3):358-364.
6. Clegg, R.G., C. Di Cairano-Gilfedder, and S. Zhou (2010). A critical look at power law modelling of the Internet. *Computer Communications*. **33**(3):259-268.
7. Zörnig, P. and G. Altmann (1995). Unified representation of Zipf distributions. *Computational Statistics & Data Analysis*. **19**(4):461-473.
8. Gupta, P.L., R.C. Gupta, S.H. Ong, and H.M. Srivastava (2008). A class of Hurwitz-Lerch Zeta distributions and their applications in reliability. *Applied Mathematics and Computation*. **196**(2):521-531.
9. Shaked, M. and J.G. Shanthikumar (2007). *Stochastic orders*. Springer series in statistics. Springer, New York.
10. Sim, S.Z. and S.H. Ong (2010). Parameter Estimation for Discrete Distributions by Generalized Hellinger-Type Divergence Based on Probability Generating Function. *Communications in Statistics-Simulation and Computation*. **39**(2):305-314.
11. Jain, G.C. and R.P. Gupta (1973). A logarithmic series type distribution. *Trabajos de Estadística*. **24**:99-105.
12. Doray, L.G. and A. Luong (1997). Efficient estimators for the Good family. *Communications in Statistics-Simulation and Computation*. **26**(3):1075-1088.
13. Kemp, A.W. (2010). Families of power series distributions, with particular reference to the Lerch family. *Journal of Statistical Planning and Inference*. **140**(8):2255-2259.
14. Shaked, M. and T. Wong (1997). *Stochastic orders based on ratios of*

- Laplace transforms. *Journal of Applied Probability*. **34**(2):404-419.
15. Tripathi, R.C. and R.C. Gupta (1988). Another generalization of the logarithmic series and the geometric distributions. *Communications in Statistics-Theory and Methods*. **17**(5):1541-1547.
  16. Ng, C. M., S. H. Ong and H. M. Srivastava (2012). Parameter estimation by Hellinger type distance for multivariate distributions based upon probability generating function. *Applied Mathematical Modelling* (submitted).