Journal of Scientific Research & Reports
3(18): 2459-2465, 2014; Article no. JSRR.2014.18.008

SCIENCEDOMAIN *international*
www.sciencedomain.org

Myers' Blended Method: An Alternative Approach

Jadab Kumar Pal[1], Barun Kumar Mukhopadhyay[1*] and Suchintya Kumar Gupta[2]

[1]Indian Statistical Institute, 203 B T Road, Kolkata – 700108, India.
[2]Saha Institute of Nuclear Physics, Salt Lake, Kolkata - 700064, India.

Authors' contributions

This work was carried out in collaboration between the authors. Author JKP collected the primary data from the field at Ramna Etbar Nagar, within the district of Murshidabad, state of West Bengal, he also contributed major part of this article. Author BKM designed the study, and wrote the first draft of the manuscript. Author SKG contributed literature and funding for its fruition. All authors read and approved the final manuscript.

Original Research Article

Received 3rd April 2014
Accepted 7th June 2014
Published 3rd August 2014

ABSTRACT

Myers' index is widely used in censuses, large scale sample surveys and in many other secondary data to study the digit preference error on single year raw age distribution particularly in developing countries where huge peaks and troughs are observed at the ten digits of age. Historically many modifications are known and theories have been developed on this method by actuaries, demographers and other social scientists (Bachi, [3]; Carrier, [4]; Ramachandran, [5]; Mukhopadhyay, [6] and others). Usually the well known Myers' Blended Method (1940) is done where the figures for ages within a broad band on ten digits from zero to nine are much significant and high. The same method has been tried in the paper to find the quality of age reporting in a small area sample in the district of Murshidabad, West Bengal, India. The index value seems to overestimate may be because of too small values for some of figures corresponding to the ten digits of the entire distribution even within a broad age range of population. An alternative approach replacing the original population by person years is made in the paper to calculate the same index. The modified index value seems to be appropriate keeping in view of the type of high and low values only for a few digits out of the ten digital values.

Keywords: Myers' index; Whipple's index; preference pattern matrix; digit preference.

**Corresponding author: E-mail: barun_mukhopadhyay@yahoo.com;*

1. INTRODUCTION

Historically indices on the digit preference error in single year data on age last birth day have perpetually been constructed by demographers, actuaries and other social scientists. Simultaneous modification on different aspect of those indices had started almost at the same time point. All these have taken place because of methodological drawback or constraint while constructing the same. As a result the process was continued in a fashion from long past, in the recent past and still at present [1,2,3,4,5,6]. Every researcher while doing any socio demographic work or even on the census data of many countries, especially in the developing world at an initial stage use to focus on the data quality particularly on age where large extent of errors in the raw data are observed. Myers' index (1940), to a large extent, has been very useful to measure the extent of digit preference error in the single year age data. It has also been presented in some instance to measure the concentrations particularly at digits, '0' and/or '5' [7]. This index is calculated in a way at another preferable digit in China and Korea, where different kind of age reckoning on 12 different animals is prevalent [8]. Some alternative approach was proposed to justify the usefulness in recent period of a method proposed in 1958, long time ago [9,10]. The present paper tries to calculate the Myers' Index on a very small sample size to know the quality of age reporting in the study area. But at an initial stage the sample values for the ten digits of ages within a broader age range even became very small sometimes of the order of unity and in other digits the figures are too small as well to justify applying the Myers' Blended Method *par se,* while the same is usually done on censuses or large scale survey data where the figures for ages within a broad range on ten digits from zero to nine are much high. Here as the sample size is small an alternative approach in the present paper is made in constructing the index replacing at the outset the original sample figures for ages by person years and then finding the index using the rest of the same calculation.

2. SAMPLE

Using simple random sample a survey was conducted on the small village, Ramna Etbar Nagar in the district of Murshidabad, West Bengal, India, 2008 inhabited mostly by a minority community.

3. METHODOLOGY

Myers' Blended Method (1940) is usually calculated either in censuses or large scale survey data in order to know the quality of age reporting in those areas. In the present paper the sample size is very small so that the population data classified into ten categories of digits from 0 to 9 become further smaller. To make the data more visible a new technique is adopted in the present paper. The population data are replaced by person year of population data.

3.1 Analysis of the Data

The age selective migration might have affected the age distribution as is usually observed in a small geographical area, still there are clear indication of preferences and/or dislikes at the even and/or odd digits including the digit '0' the most and '5' the second most preferred digits. Similarly other even digits show concentration, contrary odd digits "disliking". The Table 1 shows the single year age data from age 30 to 60 with per cent figures. It is clear from the figures that ages ending in digits '0' are much higher as compared to surrounding

ages ending in odd digits like ‘1’ and ‘9’. Similarly the ages ending in even digits such as ‘2’, ‘4’ and ‘8’. But the even digit ‘6’ shows heaping only for age 56.

Table 1. Single year age data of population (327) of Ramna Etbar Nagar District of Murshidabad, West Bengal, India, 2008

Age in years	Person	Percent	Age in years	Person	Percent
30	19	5.8	51	0	0.0
31	0	0.0	52	6	1.8
32	14	4.3	53	3	0.9
33	6	1.8	54	3	.9
34	13	4.0	55	13	4.0
35	18	5.5	56	2	0.6
36	9	2.8	57	0	0.0
37	11	3.4	58	4	1.2
38	17	5.2	59	0	0.0
39	3	0.9	60	4	1.2
40	37	11.3			
41	2	0.6			
42	13	4.0			
43	2	0.6			
44	7	2.1			
45	24	7.3			
46	1	0.3			
47	1	0.3			
48	20	6.1			
49	5	1.5			
50	16	4.9			

As far as odd digits are concerned, ‘1’, ‘3’ and ‘9’ show disliking of ages ending in those digits. The following graph on single years of ages from 23 up to 65 shows the huge peaks and troughs indicating thereby digit preference error in the single year age distribution among the sample population of 327 of a small village like Ramna Etbar Nagar in the district of Murshidabad, West Bengal, India. Since the data are small, hence the study is based only on male.

3.2 Original Myers’ Blended Method versus the Alternative Approach

Before applying the new technique of using person years instead of raw population in calculating the Myers’ index based on the blended method, it may be proper first of all to apply the original Myers’ method on the present data to have the idea about whatever the numerical value of the index giving and then comparing that with the value from the proposed new technique. As far as the age range is concerned Myers opted for ages 10 to 69 in his paper (1940). But different arbitrary range is found applied in many countries of the world as per the availability of the data. It is to be kept in mind that the range should be such that there should not be very young population and the very old aged population where there are different kinds of errors existed. Former undercounting is a great hindrance for developing countries to have a good quality of age data in the very young population [11]. For the latter, exaggeration of ages are also problem in population of usually illiterate third world countries. This subject is elaborately discussed in many literatures [12,13,14,15,16].

3.3 Application of Original Myers' Blended Method

The following table gives the calculation of Myers' Index based on the Original Myers' Blended Method.

A perusal of the below table shows the usual pattern of values in the last column with respect to ten digits of age. However the population figures in the 2nd and 3rd columns are in some cases very scarce even of the order of unity which sometimes may act as an out lyre in statistical analysis. However, it cannot be ignored in the present context where each digit has to be studied. In addition to the above low value and some other insufficient figures the final index value of 70.93 which perhaps shows an overestimate of the true value. As a result, the alternative approach is tried here to get a justified index value. This is done by replacing the entire fluctuating values due to small sample in the 2nd and 3rd columns of the billow table by the corresponding values of person years as is done in life table stationary population. The following paragraph hence is presented under a heading "The Alternative Approach".

3.4 Application of Alternative Approach

The idea of revising the above index is nothing new. As already pointed out in the above that there were many modifications and corrections were already done on this particular index in the past. A few may be mentioned for this particular field. For example the limits of Myers' Index (M.I) is $0 \leq M.I \geq 180$. As far as the lower limit is concerned zero indicates there is no digit preference error in the age data of any country's census/survey. But according to Bachi [3] this is not possible, i.e zero is not at all reachable. He then gave an alternative approach by using regression method by which exactly zero is possible. Another attempt was made by Carrier [4] on the subject of Myers' method using number of equation and derivation in his own way. Later Ramachandran [5] again objected Myers' method that did not consider the age structure of the population and varying incidence of digit preferences. He constructed a matrix namely Preference Pattern Matrix and further gave his solution considering the two conditions as mentioned above. For elaborate idea and knowledge the references are already cited. The following table gives the calculation of Myers' Index based on the alternative approach proposed in the paper.

The above table shows the estimated figures in columns (2) and (3) as person years calculated from the original population figures which incidentally show some kind of dearth of values in some cases due to small sample about which already a mention has been made. The pattern of the values in the last column is quite justified in the sense of what has been observed in the graph of the single year data (Fig. 1) of the population in the study area where illiteracy is high and economic condition is very poor in general. However the final estimated index value of 66.4 after modification of Myers' original method is quite consistent.

Table 2. Calculation of original Myers' index

Digit	Person years		Weight for		Blended sum = Col 2*Col 4 +Col 3*Col 5	Percentages	10% deviation
	30-49	40-59	Col 2	Col 3			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0	56	53	1	9	533	26.74360261	16.74360261
1	2	2	2	8	20	1.003512293	-8.996487707
2	27	19	3	7	214	10.73758154	0.737581535
3	8	5	4	6	62	3.110888108	-6.889111892
4	20	10	5	5	150	7.526342198	-2.473657802
5	42	37	6	4	400	20.07024586	10.07024586
6	10	3	7	3	79	3.963873557	-6.036126443
7	12	1	8	2	98	4.917210236	-5.082789764
8	37	24	9	1	357	17.91269443	7.912694431
9	8	5	10	0	80	4.014049172	-5.985950828

Index=70.93

Table 3. Calculation of Myers' index on the basis of person years

Digit	Person years		Weight for		Blended sum = Col 2*Col 4 +Col 3*Col 5	Percentage	10% deviation
	30-49	40-59	Col 2	Col 3			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0	2050	2280	1	9	22570	22.30545728	12.30545728
1	82	82	2	8	820	0.810388789	-9.189611211
2	994	858	3	7	8988	8.88265175	-1.11734825
3	284	245	4	6	2606	2.575455102	-7.424544898
4	750	470	5	5	6100	6.028501967	-3.971498033
5	1710	1795	6	4	17440	17.23558595	7.235585951
6	370	158	7	3	3064	3.028086889	-6.971913111
7	454	47	8	2	3726	3.682327595	-6.317672405
8	1606	1192	9	1	15646	15.46261341	5.462613405
9	362	245	10	0	3620	3.57757002	-6.42242998

Index=66.40

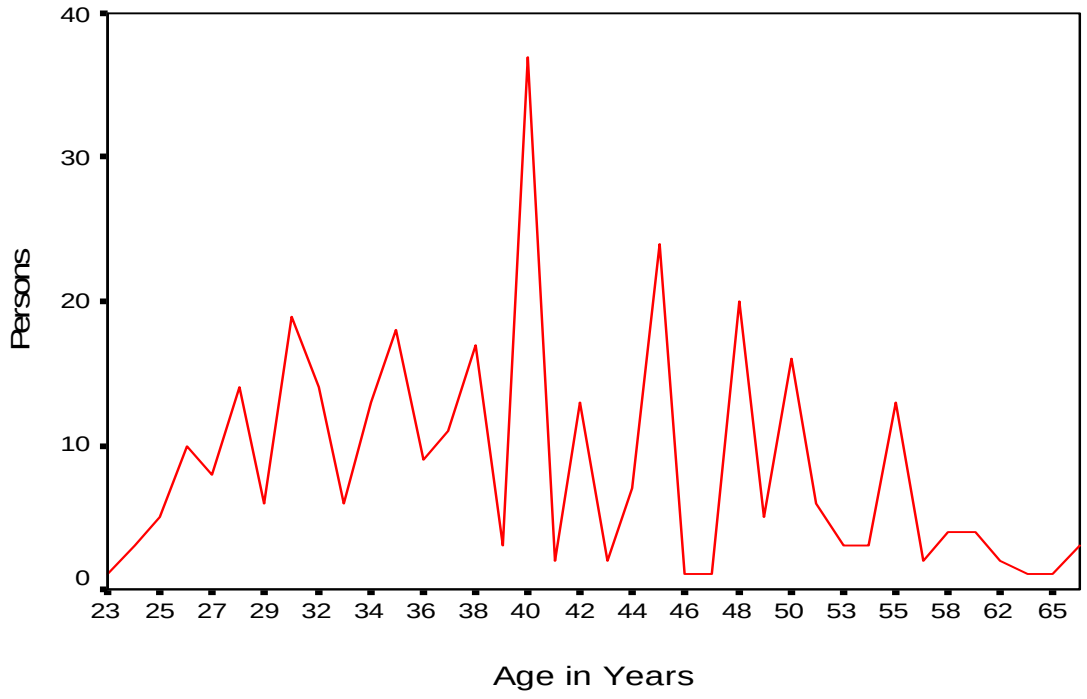

Fig. 1. Single year age data of population (327) of Ramna Etbar Nagar, 2008

4. CONCLUSION

The present study is a kind of analysis which applies on a very small sample data obtained from a particular locality of rural West Bengal, India. The area is dominated by a particular section of Indian community. As far as their reporting of ages as usual was found no diversion from overall reporting pattern in India. There are preferences for particular digits especially '0' and '5' and even digits too. Contrary there are disliking for odd digits as well. The total sample population itself being very small and further diverting it into ten categories,- here digits from '0' to '9', the single cell figures are found sometimes very small even of the order of only unity. While applying Myers' blended method where population figures are the main ingredients which must have some legitimate values in each of the ten digits. A new kind of approach is proposed in the article to have been replaced the entire original sample population figures as per the ten digits by some estimate in terms of person years in life table terminology. As a result the entire distribution of the figures turned to a legitimate one. The final result based on the new methodology shows more appropriateness of the index value of 66.4 having a parity of high figures only for a few ages like 40, 45 and 48 whereas the original Myer's index of 70.9 seems to be over estimating because of random fluctuation from small sample size.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. King G. The new national life tables. Journal of the Institute of Actuaries. 1915;49. Oxford.
2. Myers RJ. Errors and bias in the reporting of ages in census data. Transaction of the Actuarial Society of America. 1940;XLI(2):104.
3. Bachi R. The tendency to round off age returns, measurement and correction. Bulletin of the International Statistical Institute. 1951;33(4):195-222.
4. Carrier NH. A note on the measurement of digital preference in age recordings. Journal of the Institute of Actuaries; 1959. Oxford.
5. Ramachandran KV. An index to measure of digit preference error in age data, proceedings of the UN's world population conference (Summary). Belgrade, Yugloslavia; 1965.
6. Mukhopadhyay BK. Shortcoming in some Index: Improvisation through mathematical modeling, accepted for oral presentation at SMTDA international conference. Lisbon, Portugal; 2014.
7. Whipple GC. Vital statistics- an introduction to the science of demography, Chapman and Hall; 1919.
8. Poston DL. The age and sex composition of the republic of Korea and its provinces, Genus, Italy. 2003;LIX.
9. Mukhopadhyay BK. Non-heaping in age DATA: A new technique under review. American Statistician; 2013.
10. Turner SH. Patterns of heaping in the reporting of numerical data. Proceedings of the Social Statistics Section (American Statistical Association); 1958.
11. Mukhopadhyay BK. An age adjustment of very young children of India, 1981 and reappraisal of fertility and mortality rates. A model approach, Genus, Roma, Italy. 1986;XLII-n:3-4.
12. Bogue DJ, Arriaga EE, Anderton DL. (eds). Readings in population research methodology, Basic Tools. Chicago. Illinois: Social Development Center and United Nations Population Fund. 1993;1.
13. Ewbank DC. Age misreporting and age-selective under-enumeration: Sources, patterns, and consequences for demographic analysis. Washington: National Academy Press, Committee on Population and Demography, Report No. 4; 1981.
14. Srinivasan K. Basic demographic techniques and applications. New Delhi: Sage Publications; 1998.
15. Shryock HH, Seigel JS. Associates. The methods and materials of demography, US Bureau of the Census, fourth printing (rev.), U.S. Government Printing Office, Washington D.C. 1980;1.
16. United Nations. Methods of estimating basic demographic measures from incomplete data Manual 4, Population Studies 42, New York; 1967.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=611&id=22&aid=5624