Lecture four

❖ Demographic error measures:

It was mentioned previously that some age errors are due to some individuals, preference for ages ending in certain numbers over others, such as zero and five, and also even numbers, which leads to a concentration of the population at these preferred ages.

Consequently, this leads to errors in the age distribution, and some methods have been proposed to measure this bias, including the following:

1. Whipple's Index of Digit Preference

This index is limited to measuring the degree of preference for ages that end in zero or five, in the age range from 23 to 62. The formula of the scale is:

we obtain this number by dividing the total number of people with ages ending in one of the numbers 0 or 5 by (1/ the ratio is the preferred numbers) of the total number of people from the ages of 23 to 62.

$$Weibull\ measurement = \frac{The\ total\ population\ with\ ages\ ending\ at\ 0-5\ in\ the\ range}{1/\text{ratio} * The\ total\ population\ in\ the\ age\ range(23-62)} * 100$$

The value of the Weibull scale ranges from 100 to 500. If the value is 100 as a minimum (then there is unbiased (no focus) on ages ending in 0 or 5, and a maximum of 500 (then there is a complete bias for ages ending in 0 or 5)

In general, the following division reveals the extent of accuracy in:

Provide a statement of age:

Less than 105 very accurate

- 105-109.9 accurate
- 110-124.9 is almost correct
- 125-174.9 is inaccurate
- more than 175 is completely inaccurate

The Weibull number is considered an effective measure of the accuracy of ages, as it enables us to reveal the preference of some numbers and is easily calculated, but

the main flaw in this indicator is that it only measures the preference of numbers, and it only reveals preference for the numbers 0 and 5.

Note: if we want to count the 0 and 5, then the ratio is 1/5, but 0 only or 5 only , then it is 1/10

Example:

Using age distribution data for a West African country (1960), determine the extent of heaping on

(a) Digit '0' and '5'.

(b) Digit '0' only

(c) Digit '5' only.

age	population	age	population
23	38687	45	44654
24	51289	46-49	77102
25	77141	50	52024
26-29	201450	51-54	44751
30	110379	55	14024
31-34	132136	56-59	45283
35	64091	60	38377
36-39	134140	61	4779
40	81515	62	7866
41-44	84422	sum	1304110

Solution:

This table has the numbers of the favorite age 0 or 5 as follows:

Age	Population
25	77141
30	110379
35	64091
40	81515
45	44654
50	52024
55	14024
60	38377
Sum	482205

$$Weibull\ measurement = \frac{The\ total\ population\ with\ ages\ ending\ at\ 0-5\ in\ the\ range}{1/\text{ratio} * The\ total\ population\ in\ the\ age\ range(23-62)} * 100$$

a. Digit ‘0’ and ‘5’.

$$= \frac{482205}{\frac{1}{5} * \ 1304110} * 100 = 185.1$$

The measurement is completely inaccurate

b. Digit ‘0’ only

$$= \frac{282295}{\frac{1}{10} * 1304110} * 100 = \ 216.5$$

The measurement is completely inaccurate

c. Digit ‘5’ only.

$$= \frac{199910}{\frac{1}{10} * 1304110} * 100 = \ 153.3$$

The measurement is inaccurate

2. Myer's Blended Index of Digit Preference

This index is used for evaluating single - year age - sex data. It can give the extent of digit preference for all the digits 0, 1, 2, 3... 9. It can be used to report errors for all ages 10 – 99 years.

The underlying assumption of the method is that in the absence of systematic irregularities in the reporting of age, the blended sum at each terminal digit should be approximately equal to 10% of the total blended population.

If the sum at any given digit exceeds 10% of the total blended population, it indicates over selection of ages ending in that digit (i.e. digit preference).

On the other hand, a negative deviation or sum that is less than 10% of the total blended population indicates an under selection of the ages ending in that digit (i.e. digit avoidances).

If age heaping is non-existent, the index would be approximately zero.

Example

Use the Myer's blended index to assess the quality of age data given below:

Digit	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
0	386	417	307	190	89	80	50	40	10
1	133	93	90	89	41	35	33	10	2
2	341	227	100	40	30	25	15	6	4
3	223	160	90	46	28	22	8	5	1
4	201	138	50	38	25	10	8	5	3
5	298	238	201	154	148	65	25	2	3
6	198	105	70	49	31	18	12	6	4
7	166	86	50	45	25	9	5	4	2
8	255	100	90	40	28	10	5	5	2
9	132	85	40	20	8	6	2	2	2

1. Sum all the populations ending in each digit over the whole range i.e. 10-99
2. Sum figures between ages 20-99.
3. Multiply the sums in (1) by coefficients; 1, 2, 3,4,5,6,7,8,9 and 10.
4. Multiply the sums in (2) by coefficients from 9 descending to 0 i.e. 9 ,8 ,7, 6, 5 ,4 ,3 ,2 ,1 ,0.
5. Add the product of (3) and (4), to obtain the blended sum
6. Add up the blended sum in (5).
7. Find the percent (%) of the total blended sum at different digit ends.
8. Take the deviations of each % in (7) from 10.0. This result indicates the extent of concentration or avoidance of a particular digit.

Digit	Sum (10 – 99)	Coef-ficient	Product	Sum (20-99)	Coef-ficient	Product	Blended sum	Percent dist	Dev from 10%	Remark
0	1569	1	1569	1183	9	10647	12216	21.7	11.7	Preference
1	526	2	1052	393	8	3144	4196	7.5	-2.5	Avoidance
2	788	3	2364	447	7	3129	5493	9.8	-0.2	Avoidance
3	583	4	2332	360	6	2160	4492	8.0	-2.0	Avoidance
4	478	5	2390	277	5	1385	3775	6.7	-3.3	Avoidance
5	1134	6	6804	836	4	3344	10148	18.0	8.0	Preference
6	493	7	3451	295	3	885	4336	7.7	-2.3	Avoidance
7	392	8	3136	226	2	452	3588	6.4	-3.6	Avoidance
8	535	9	4815	280	1	280	5095	9.0	-1.0	Avoidance
9	297	10	2970	165	0	0	2970	5.3	-4.7	Avoidance

3. United Nations Age-gender Accuracy Index

This index which was proposed by the United Nation is used for evaluation of five-year age- gender data. The index is also referred to as **Joint Score**. It has three components;

a. **Average gender ratio score (S)**: This score is obtained by fist calculating the gender ratio at each age group. Successive differences irrespective of sign are added and averaged.

$$Age - specific\ gender\ ratio(S) = \frac{P_m}{P_f} * 100$$

b. **Average male age ratio score (M):** For each age group for males, calculate the age ratios computed as:

$$age\ ratio(M) = \frac{P_{(m)x}}{{}^1/_2\,(P_{(m)x-5} + P_{(m)x+5})} * 100$$

The deviations from unity irrespective (بغض النظر) of sign are added and averaged (M).

c. **Average female age ratio score (F)**: For each age group for females, the age ratios are calculated using the same formulae as for males. The deviations from unity irrespective of sign are added and averaged (F).

$$age\ ratio(M) = \frac{P_{(f)x}}{{}^1/_2\,(P_{(f)x-5} + P_{(f)x+5})} * 100$$

d. The index is then computed as: UNAI = 3(S) + M + F

The reported age-gender data for a given population is presumed to be accurate if:

- the age-gender accuracy index is between 0 and 19.9,
- inaccurate if the index is between 20 and 39.9,
- and highly inaccurate if the index is above 40.

Example

Use the United Nation age-sex accuracy index to assess the age – gender reporting of the data shown in the table below:

Age Group	Male Population	Female Population	Gender Ratio	First Difference	Male Age Ratio	Dev from 100	Female Age Ratio	Dev from 100
0-4	2376	2350	101.1	5	-	-	-	-
5-9	1983	1972	100.6	-0.3	99.1	-0.9	99.5	-0.5
10-14	1628	1614	100.9	-0.1	99.9	-0.1	99.7	-0.3
15-19	1277	1265	101.0	-0.1	97.3	-2.7	97.3	-2.7
20-24	997	986	101.1	-2.5	95.7	-4.3	96.5	-3.5
25-29	807	779	103.6	1	97.4	-2.6	95.6	-4.4
30-34	661	644	102.6	-0.8	97.4	-2.6	98.2	-1.8
35-39	551	533	103.4	-0.8	95.4	-4.6	104.3	4.3
40-44	394	378	104.2	8.2	96.3	-3.7	93.2	-6.8
45-49	267	278	96.0	10.9	90.8	-9.2	91.8	-8.2
50-54	194	228	85.1	4.6	90.7	-9.3	95.4	-4.6
55-59	161	200	80.5	-2.9	97.6	-2.4	102.3	2.3
60-64	136	163	83.4	-0.3	103.0	3	100.9	0.9
65-69	103	123	83.7	-11.2	97.6	-2.4	101.7	1.7
70-74	75	79	94.9	-10.2	90.9	-9.1	86.8	13.2
75-79	59	59	105.1	-	-	-	-	-
				54.4		56.9		55.2
				S=3.63		M=4.06		F=3.94

UNAI = 3(S) + M + F

=3 (3.63) + 4.06 + 3.94

= 10.89 + 4.06 + 3. 94 = 18.89

Comment: Age – gender reporting is accurate.