MedCrave Step into the World of Research

Biometrics & Biostatistics International Journal

Research Article

Open Access

Simultaneous estimation of adjusted rate of two factors using method of direct standardization

Abstract

This paper presents the use of standardization or adjustment of rates and ratios in comparing two populations using single indices rather than a series of specific rates or ratios. Here the overall adjusted crude rate or the unadjusted crude rate for two populations will have same estimate irrespective of the nature of the standard population distribution. These results are obtained in all cases whenever the two standard distributions are of the total sample. In these cases the overall adjusted crude rates based on the two sets of directly adjusted rates would be equal to each other, although not necessarily always equal to the overall unadjusted crude rate as is found to be the case here. However, if the standard population distribution chosen for a given population is different from that chosen for another, then the two resulting estimated adjusted or standardized crude rates would most likely not be equal to each other.

Keywords: standardization, adjusted specific, unadjusted crude rate, adjusted crude rate, ratios

Volume 2 Issue 7 - 2015

Oyeka ICA,[1] Okeh UM[2]

[1]Department of Statistics, Nnamdi Azikiwe University, Nigeria
[2]Department of Industrial Mathematics and Applied Statistics, Ebonyi State University, Nigeria

Correspondence: Oyeka ICA, Department of Statistics, Nnamdi Azikiwe University, Awka, Nigeria, Email uzomaokey@ymail.com

Received: August 24, 2015 | **Published:** October 27, 2015

Introduction

Standardization or adjustment of rates and ratios is often necessary because it is usually easier in comparing two populations, say, to make the comparison using single summary indices rather than a series of specific rates or ratios. This approach also helps avoid the problem of small imprecise and sometime non-existence of specific rates and ratios.[1-3]

Standardization of rates and ratios may be done for only one factor or several factors of classification of a criterion variable of interest. In particular if a criterion variable or condition is associated with each of two factors of classification which may by themselves also be associated with each other, then standardization of rates or ratios may sometimes be necessary for a clearer analysis and inter-presentation of results to simultaneously standardize or adjust the rates for the two factors of classification, first specific to the levels or categories of one of the factors across the levels of the other factor, and then also specific to the levels of the second factor say holding constant the levels, that is for all levels or categories of the first factor.[4,5] Research interest in this case would be to identify and measure the separate effects of the two factors of classification on the criterion variable or condition. This paper proposes, develops and presents a formatted systematic statistical method for this purpose.

The proposed method

Research interest here is using the direct method of standardization of rates to measure or estimate the separate effects of two factors of classification which may be associated on the variable being studied and to obtain sample estimate of unadjusted and adjusted crude rates specific to the levels of each of the factors holding the levels of the other factor of classification constant.

Now to develop the method of estimation of direct standardized or adjusted rate, suppose A and B are two variables of classification with ‘a’ and ‘b’ groups or levels respectively. Factors A and B may be associated or related. Research interest is to estimate the rates of occurrence of a criterion variable or condition specific to each of the levels of factor A across, that is for all levels of factor B and also the rates of occurrence of the specified condition specific to each of the levels of factor B for all levels of factor A as well as the corresponding marginal rates and overall rate. Suppose a total random sample of size N=N.. of subject are randomly drawn from an antecedent or predisposing population C for all levels of factors A and B of which N_{ij} is the number of subjects at the i[th] level of factor A and j_{th} level of factor B, for i=1,2,…a_j and j=1,2,…..,b. Also suppose there are a total of n=n.. outcomes or cases in condition or set D of cases for all levels of factors A and B of which n_{ij} cases are at the i[th] level of factor A and j[th] level of factor B, for i=1,2,…a_j and j=1,2,…,b where population D is possibly a subset of population C.

Now the rate of occurrence of cases in population D as a function of cases in population C specific to the i[th] level of factor A and j[th] level factor B is

$$r_{ij} = \frac{n_{ij}}{N_{ij}} \quad ....1$$

For i=1,2,…a_j; j=1,2,…b.

Let

$$N_{i.} = \sum_{j=1}^{b} N_{ij} \, ; N_{.j} = \sum_{i=1}^{a} N_{ij} \quad ....2$$

be respectively the total or marginal number of subjects or observations in population C at the i[th] level of factor A and j[th] level of factor B.

Similarly let

$$n_{i.} = \sum_{j=1}^{b} n_{ij} \, ; n_{.j} = \sum_{i=1}^{a} n_{ij} \quad ....3$$

be respectively the total or marginal number of cases or outcomes in population D at the i[th] level of factor A and j[th] level of factor B. Then the estimated unadjusted crude rates of occurrence of cases or outcomes in population D as a function of outcomes in population C specific to the i[th] level of factor A for all levels of factor B for all levels of factor A are respectively the ratios

$$r_{i.;unadj} = \frac{n_{i.}}{N_{i.}} \, ; r_{.j;unadj} = \frac{n_{.j}}{N_{.j}} \quad ....4$$

For i=1,2,….a; j=1,2,…b.

Note that

$$N = N_{..} = \sum_{i=1}^{a} n_{i.} = \sum_{j=1}^{b} N_{.j} = \sum_{j=1}^{b}\sum_{i=1}^{a} N_{ij} \text{...........}(5)$$

And

$$n = n_{..} = \sum_{i=1}^{a} n_{i.} = \sum_{j=1}^{b} n_{.j} = \sum_{j=1}^{b}\sum_{i=1}^{a} n_{ij} \text{..............}(6)$$

Therefore the overall unadjusted crude rate of occurrence of event D as a function of event C for all levels of factors A and B is

$$r_{unadj} = r = \frac{n_{..}}{N_{..}} \text{...........}(7)$$

As noted above research interest is to obtain standardized or adjusted crude rate specific to each level of factor A for all levels of factor B and also specific to each level of factor B for all levels of factor A as well as the overall adjusted or standardized crude rate. To obtain estimates of adjusted or standardized crude rates specific to each level of factor B for all levels of factor A we use the proportionate distribution of total number of observations across the 'a' levels or groups of factor A, namely P_{is} the waiting factor, for i=1,2,..,a.

Thus

$$P_{is} = \frac{N_{i.}}{N_{..}} \text{........}(8)$$

Similarly to obtain estimates of adjusted or standardized crude rate specific to each level of factor A for all levels of factor B we use the proportionate distributions N.. across the 'b' levels or groups of factor B, namely P_{sj} the waiting factor, for j=1,2,…,'b.' Thus

$$P_{sj} = \frac{N_{.j}}{N_{..}} \text{.........}(9)$$

Hence the adjusted or standardized crude rate of condition D as a function of population C specific to the j[th] level of factor B for all levels of factor A is

$$r_{.j;adj} = \sum_{i=1}^{a} p_{is} r_{ij} \text{..........}(10)$$

Similarly the adjusted or standardized crude rate of condition D as a function of population C specific to the i[th] level of factor A for all levels of factor B is

$$r_{i.;adj} = \sum_{j=1}^{b} p_{sj} r_{ij} \text{..........}(11)$$

We then obtain the sample estimate of the overall adjusted crude rate of condition D as a function of population C for all levels of factors A and B as

$$r_{s;adj} = r_{..adj} = \sum_{i=1}^{a} p_{is.} r_{i.} = \sum_{j=1}^{b} p_{sj} . r_{.j} \text{........}12$$

These results are summarized in Table 1.

Table 1 Data format for Estimation of Unadjusted and Adjusted Rates in two Factor Standardization by Direct method

Factor B								
Factor A	**1**	**2**	**……….**	**b**	**Total**	**Proportion**	**Un adjust**	**Adjust**
1	$n_{11}(N_{11})$	$n_{12}(N_{12})$	………	$n_{1b}(N_{1b})$	$(n_{1.}(N_{1.}))$	(p_{sj})	$(r_{.j;adj})$	$(r_{.j;unadj})$
$r_{.1;adj}$								
$\frac{n_{.1}}{N_{.1}}$								
2	$n_{21}(N_{21})$	$n_{22}(N_{22})$	……..	$n_{2b}(N_{2b})$	$n_{2.}(N_{2.})$	p_{2s}	$\frac{n_{2.}}{N_{2.}}$	$r_{2.;adj}$
$r_{.2;adj}$	r_{21}	r_{22}	………	r_{2b}	$r_{2.}$	….	….	…..
$\frac{n_{.2}}{N_{.2}}$								
…								
a	$n_{a1}(N_{a1})$	$n_{a2}(N_{a2})$	……	$n_{ab}(N_{ab})$	$\frac{p_{sa.}}{n_{a.}(N_{a.})}$	p_{as}	$\frac{n_{a.}}{N_{a.}}$	$r_{a.;adj}$

Citation: Oyeka ICA, Okeh UM. Simultaneous estimation of adjusted rate of two factors using method of direct standardization. *Biom Biostat Int J.* 2015;2(7):210–214. DOI: 10.15406/bbij.2015.02.00051

Table Continued

Factor B								
Factor A	**1**	**2**	**………..**	**b**	**Total**	**Proportion**	**Un adjust**	**Adjust**
$r_{.j;adj}$	r_{a1}	r_{a2}	…….	r_{ab}	$r_{a.}$			
Total	$n_{.1}(N_{.1})$	$n_{.2}(N_{.2})$		$n_{.b}(N_{.b})$				
$\left(n_{.j}(N_{.j})\right)$	$r_{.1}$	$r_{.2}$		$r_{.b}$				
proportion $\left(p_{sj}\right)$	p_{s1}	p_{s2}	……..	p_{sb}	….	…..	….	
Un adjust $\left(r_{.j;adj}\right)$	$\frac{n_{.1}}{N_{.1}}$	$\frac{n_{.1}}{N_{.1}}$	…..	$\frac{n_{.b}}{N_{.b}}$	……	…	$r_{.unadj} = \frac{n_{..}}{N_{..}}$	
Adjust $\left(r_{.j;unadj}\right)$	$r_{.1;adj.}$	$r_{.2;adj}$	……	$r_{.b;adj}$				$r_{..;adj}$

Illustrative Example

We now illustrate the proposed method with the sample data of Table 2 on premature and live births by birth order and age of mother in a certain population.

Table 2 Sample Data on Premature and Live births by Birth order and Maternal age in a population

Birth Order							
Maternal Age	**1**	**2**	**3**	**4**	**5+**	**Total** $\left(n_{i.}(N_{1.})\right)$	**Proportion of total births** (p_{is})
Under 20	11(23)	3(72)	3(32)	1(43)	0(33)	18(203)	
	0.478	0.042	0.094	0.023	0.000	0.089	0.066
20-24	14(329)	15(327)	7(176)	3(69)	8(67)	47(968)	
	0.043	0.046	0.040	0.043	0.119	0.049	0.012
25-29	6(115)	11(209)	11(207)	6(132)	6(123)	40(786)	
	0.052	0.053	0.053	0.045	0.049	0.051	0.254
30-34	4(78)	8(83)	10(117)	9(98)	12(150)	43(526)	
	0.051	0.096	0.085	0.092	0.080	0.082	0.170
35-39	4(42)	8(56)	11(90)	14(56)	3(104)	40(348)	
	0.095	0.143	0.122	0.050	0.029	0.115	0.112
40 and above	3(34)	4(457)	8(72)	10(48)	4(68)	29(267)	
	0.088	0.089	0.111	0.208	0.059	0.109	0.086
Total $\left(n_{.j}(N_{.j})\right)$	42(621)	49(792)	47(694)	45(446)	33(545)	217(3098)	
	0.068	0.010	0.068	0.096	0.060	0.070	0.070
Proportion of total births (p_{sj})	0.200	0.256	0.256	0.224	0.144	0.176	

Citation: Oyeka ICA, Okeh UM. Simultaneous estimation of adjusted rate of two factors using method of direct standardization. *Biom Biostat Int J.* 2015;2(7):210–214. DOI: 10.15406/bbij.2015.02.00051

The data of Table 2 is used to obtain estimates of the unadjusted and adjusted crude rate specific to each of the levels or groups of the two factors of classification.

Specifically to estimate adjusted or standardized crude rates specific to birth order, we apply the proportionate distribution of the total life births across maternal age as the standard population, namely p_{is} in the last column of Table 2 to each of the columns of rates, $p_{:}$ of the Table, for j=1,2,3,4,5.Similarly to estimate adjusted or standardize crude rate specific to Maternal age we apply the proportionate distribution of total life births across birth order as the standard population, namely p_{sj} in the last row of Table 2 to each of the rows of rates, $p_{:}$ of the Table, for i=1,2,3,4,5,6.The results are presented in Table 3.

Table 3 Simultaneous Estimates of Unadjusted Adjusted Premature Birth rates by maternal age and Birth order: Direct Standardization

Birth order													
Maternal Age	**Proportion of total birth** (p_{is})	r_{i1}	**1**	r_{i2}	**2**	r_{i3}	**3**	r_{i4}	**4**	r_{i5}	**5+**	**Unadjusted crude rate** $(r_{1.;unadj})$	**Adjusted crude rate** $(r_{1.;adj})$
less than 20	0.066	0.478		0.042		0.094		0.023		0.000		0.089	0.131
20-24	0.312	0.042		0.046		0.040		0.043		0.119		0.049	0.057
25-29	0.254	0.052		0.052		0.053		0.045		0.049		0.050	0.051
30-34	0.170	0.051		0.096		0.085		0.092		0.080		0.082	0.081
35-39	0.112	0.095		0.143		0.122		0.250		0.029		0.115	0.157
40 and over	0.086	0.088		0.008		0.111		0.208		0.059		0.109	0.594
Proportion of total birth $(p_{.j})$													
Unadjusted cruderate $(r_{.j;unadj})$		0.068		0.062		0.068		0.096		0.061		0.070	
Adjusted crude rate $(r_{.j;adj})$			0.086		0.070		0.069		0.088		0.068		0.070

Summary and conclusion

The adjusted crude rate of premature births specific to birth order for all age groups shown in the last row of Table 3 are estimated using equation 10,while the corresponding adjusted crude rate specific to maternal age for all birth orders shown in the last column of Table 3 are estimated using equation 11. Thus the last two rows of Table 3 show rates specific for birth order and directly adjusted for maternal age, with the standard maternal age distribution of births being that of the total sample of births. The last two columns of the Table show rates specific for maternal age and directly adjusted for birth order, with the standard birth order distribution of birth being that of total sample of births.

The estimated adjusted specific premature birth rate of Table 3 seem to indicate that incidence of premature births may not be strongly associated with birth order, but may probably be some how associated with increasing maternal age, especially from age 25 years. The overall adjusted crude premature birth rate is estimated to be severally 70 per 1000 live births whether the standard population distribution is either the proportionate distribution of total birth by birth order or by maternal age. The unadjusted crude rate is also here estimated to be 70 per 1000 live births. These results are usually the case whenever the two standard distributions are those of the total sample. In these cases the overall adjusted crude rates based on the two sets of directly adjusted rates would be equal to each other, although not necessarily always equal to the overall unadjusted crude rate as is found to be the case here. However, if the standard population distribution chosen for population A (here maternal age)is different from that chosen for factor B(here birth order),then the two resulting estimated adjusted or standardized crude rates would most lively not be equal to each other.

Acknowledgement

None.

Conflict of interest

None.

References

1. Fleiss JL. Statistical Method for Rates and proportions. 2nd edn. New York. John Wiley & Sons; 1981.
2. Pepe MS. The Statistical Evaluation of Medical Tests for Classification and Prediction. Oxford statistical series 28, Oxford. UK. University Press; 2003.
3. Greenberg RS, Daniels SR, Flanders WD, et al. Medical Epidemiology. London. Lange-McGraw-Hill; 2001.
4. Cochran WG. The comparison of percentages in matched samples. Biometrika. 1950;37:256-266.
5. Gibbon JD. Non Parametric Statistical Inference. McGraw Hill Inc; 1971.

Citation: Oyeka ICA, Okeh UM. Simultaneous estimation of adjusted rate of two factors using method of direct standardization. *Biom Biostat Int J.* 2015;2(7):210–214. DOI: 10.15406/bbij.2015.02.00051