Standardization and Decomposition of Rates: A User's Manual

by Prithwis Das Gupta

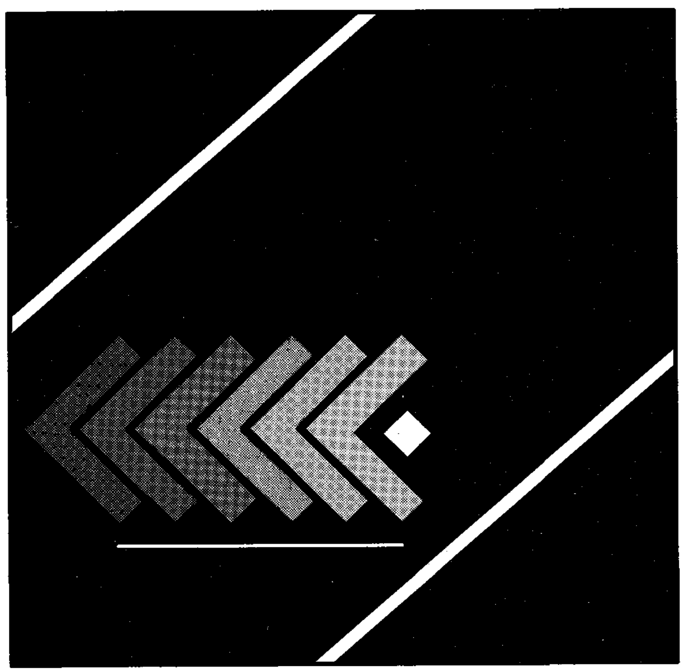
U.S. Department of Commerce
Economics and Statistics Administration
BUREAU OF THE CENSUS

Acknowledgments

This report was prepared in the Population Division, under the general direction of **J. Gregory Robinson**, Chief, Population Analysis and Evaluation Staff. **Donald J. Hernandez** and **Jorge del Pinal** reviewed the draft of the report. **Rheta D. Pemberton** typed the manuscript, and **Tecora B. Jimason** provided statistical assistance. **Michael J. Roebuck** of the Demographic Statistical Methods Division conducted the statistical review.

The staff of the Administrative and Publications Services Division, **Walter C. Odom**, Chief, provided publication planning, design, composition, editorial review, and printing planning and procurement. **Linda H. Ambill** edited and coordinated the publication; **Shirley A. Clark** designed the cover.

Standardization and Decomposition of Rates: A User's Manual



by Prithwis Das Gupta

Issued October 1993

U.S. Department of Commerce
Ronald H. Brown, Secretary

Economics and Statistics Administration
Paul A. London, Acting Under Secretary
for Economic Affairs

BUREAU OF THE CENSUS
Harry A. Scarr, Acting Director

Economics and Statistics Administration
Paul A. London, Acting Under Secretary for Economic Affairs

BUREAU OF THE CENSUS
Harry A. Scarr, Acting Director

William P. Butz, Associate Director for Demographic Programs

POPULATION DIVISION
Arthur J. Norton, Chief

SUGGESTED CITATION

Das Gupta, Prithwis, *Standardization and Decomposition of Rates: A User's Manual*, U.S. Bureau of the Census, Current Population Reports, Series P23-186, U.S. Government Printing Office, Washington, DC, 1993.

For sale by Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402.

Contents

Chapter 5.

Chapter 6.

TABLES

EXAMPLES

PROGRAMS

FIGURES

APPENDIXES

Chapter 1. Introduction

Demographers and other social scientists have traditionally used the technique of direct standardization to eliminate the compositional effects from the overall rates of some phenomenon in two or more populations. Basically, the technique assumes a particular population as standard and recomputes the overall rates in the populations by replacing their compositions by the compositional schedule of the standard population. Numerous authors have dealt with the problem of standardization including Kuczynski (1935, p. 188); Woolsey (1959); Kitagawa (1964); Spiegelman and Marks (1966); Clogg (1978); Little and Pullum (1979); Curtin, Maurer, and Rosenberg (1980); Hoem (1987); and Johansen (1990).

Starting with the classic paper by Kitagawa (1955), another area of research, namely, the decomposition of the difference between the overall rates in two populations, has been fast developing in recent years. The decomposition deals with finding the additive contributions of the effects of the differences in the compositional or rate factors in two populations to the difference in their overall rates. The techniques have been extended to include any number of factors, various functional relationships of the factors with the overall rate including the rate from cross-classified data, and simultaneous considerations of three or more populations. Authors who have contributed to the subject of decomposition include Cho and Retherford (1973); Blake and Das Gupta (1976); Das Gupta (1978, 1988, 1989, 1990, 1991, 1992); Kim and Strobino (1984); Arriaga (1984); Pollard (1988); Nathanson and Kim (1989); and Pullum, Tedrow, and Herting (1989).

The subjects of standardization and decomposition are strictly linked and, logically, one cannot be treated independently of the other. Das Gupta (1992) has recently shown explicitly how these two areas are but parts of the same consistent system. The lack of recognition of a unified system encompassing the two areas has often led to arbitrary selection of standard populations in the past, producing results that are not defensible from the decomposition point of view.

To illustrate this point, let us consider the crude birth rates of 19.435 and 15.899 for the United States for the years 1940 and 1988, respectively, showing a decline of 3.536 points (the so-called "total effect") over the 48-year period. This decline is the combined effects of the changes in the age-sex-specific birth rates and the age-sex structure, and we can compute these two effects separately by controlling for the age-sex structure and the age-sex-specific birth rates, respectively (table 6.12). If we use the 1940 age-sex structure as the standard, then the age-sex-adjusted birth rates for 1940 and 1988 are 19.435 and 16.495, respectively, and, traditionally, we interpret their difference of 2.940 as the effect of the changes in the age-sex-specific birth rates (the so-called "rate effect"). If this interpretation is correct, then, by the same logic, we should be able to use the 1940 age-sex-specific birth rates as the standard to compute the age-sex-specific birth rate-adjusted birth rates of 19.435 and 18.815 for 1940 and 1988, respectively, and interpret their difference of 0.620 as the effect of the changes in the age-sex structure (the so-called "compositional effect"). The sum of these two effects is 3.560, which is, however, different from the total effect of 3.536. (This difference of -0.024 is sometimes called the interaction effect. Section A.2 in appendix A and latter discussions in this chapter explain why there should not be an interaction effect in this case.)

Thus, in this case, use of the 1940 population as the standard produces unacceptable rate and compositional effects and, thereby, unacceptable standardized rates. When there are only two populations and two factors (e.g., age-sex-specific birth rates and age-sex structure), this problem can be easily resolved by using, for each factor, its average over the two populations as the standard (Kitagawa, 1955). However, when more than two populations and/or more than two factors are involved, it is not obvious how to choose standard populations that will not lead to any inconsistencies in the results. The objective of the present report is to provide methodologies for handling the problems of standardization and decomposition corresponding to any number of factors as well as any number of populations, for a variety of relationships of the factors with the overall rate including the rate from cross-classified data.

Chapters 2 through 5 deal with various forms of the overall rate when only two populations are compared. In chapter 2, the rate is expressed as the product of several factors. Bongaarts (1978), for example, expressed the total fertility rate as the product of five factors, namely, proportion married, noncontraception, induced abortion, lactational infecundability, and total fecundity rate.

A more general case is considered in chapter 3, where the rate is expressed as any function of two or more factors. Pullum, Tedrow, and Herting (1989), for example, expressed the mean parity of a cohort of women as a function of the parity progression ratios.

Chapter 4 deals with the rate that is a function of two or more vector-factors, a vector-factor being a factor represented by several numbers, such as the set of six age-specific fertility rates by 5-year age groups in the childbearing period. Smith and Cutright (1988), for example, expressed the illegitimacy ratio as a function of four vector-factors, namely, the age structure of childbearing women, the marital status structure within childbearing age groups, the age-specific nonmarital fertility rates, and the age-specific marital fertility rates.

The most widely used rates for the purpose of standardization and decomposition are those from cross-classified data, and these are discussed in chapter 5. Liao (1989), for example, studied the difference between two crude death rates in terms of the effects of age, race, and age-race-specific death rates. In these examples of cross-classifications, unlike those in the previous chapters, the total number of effects includes the effect of the cell-specific rates and is, therefore, always one higher than the number of variables involved in the cross-classification.

Finally, in chapter 6, the methodologies discussed in chapters 2 through 5 in the context of two populations are extended to include three or more populations. A good example of this topic is the problem of standardization and decomposition for the illegitimacy ratios for five years, considered by Smith and Cutright (1988).

Throughout the report, the applications of the standardization-decomposition techniques are illustrated by numerous examples taken from recently published literature. The report provides a working knowledge of the application of the techniques and interpretation of results without getting the reader lost in the technical mathematical derivations. The users of the techniques are expected to find the extensive supply of computer programs in FORTRAN language extremely helpful for routine applications.

The sources of data used in this report include the censuses of the United States and other countries, the national vital statistics provided by the National Center for Health Statistics, and numerous examples of standardization and decomposition published recently in various professional journals. In three examples (Examples 5.6, 5.7, and 6.8) where the data from the Current Population Survey (CPS) and the Post-Enumeration Survey (PES) are used, the discussions on their errors are available in the references cited. The standard errors used to test the differences in these examples are crude estimates based on standard error parameters from the referenced reports.

The problem of decomposition of the difference between two crude rates into several additive effects is different from the problem of, and cannot be adequately handled by, regression analysis. In other words, "the difference between two crude rates is not the equivalent of a concept like total variance of a dependent variable in regression analysis" (Kitagawa 1955). In the decomposition problem, the rate effect may not always decrease with the addition of each new factor, whereas in the regression analysis, "the addition of each independent variable to the equation increasingly explains the variation in the dependent variable" (Das Gupta 1978). Moreover, a characteristic may play a very important role as an independent variable in a regression equation in explaining the variation in a dependent variable, but the same characteristic may not be an important factor in explaining the difference between two crude rates constructed from the same dependent variable. For example, it is very likely that, in a regression analysis, a person's poverty status would be explained significantly by his (or her) race, but that the difference in the race composition in two years would not be an important factor in explaining the difference in the poverty rates in those years.

In defining the problems of standardization and decomposition, we have adopted a mathematical approach of solving unknowns from algebraic equations rather than a statistical modeling approach involving errors. This is evident from the equations in sections A.1, A.2, and A.3 in appendix A, which do not include error components. The same decomposition problem based on log-linear analysis and the purging method has been studied by Clogg and Eliason (1988); Liao (1989); Santi (1989); and Xie (1989). This interesting statistical modeling approach is handicapped by the fact that it is too complicated to be of any practical use even for data involving only two factors, as Liao's paper and the two-factor example in it amply demonstrate. Also, this approach leads to several widely different sets of results depending on the type of purging used, and it is not clear how to justify choosing one set over all others. On the other hand, the methods of standardization and decomposition provided in this report lead to a single set of solutions, and the computations involved in them are so simple that handling, for example, a six-factor case (Example 5.9) is no more difficult than handling a two-factor case, particularly if one uses the same simple general computer program provided in the report.

Again, unlike the statistical modeling approach, the present method decomposes the difference between two rates into additive main effects and does not involve any interaction effects. This should be a desirable aspect in a decomposition problem because it lends itself to easier and simpler interpretations of the results (for example, even for a four-factor problem, there are as many as 11 interaction terms). This elegance is achieved not by ignoring the parts in the total difference that other models might label interactions, but by fully accounting for the total difference in terms of main effects, and thereby distributing the so-called interactions among the main effects. This distribution does not change our conclusions about the relative importance of the factors, it only simplifies the picture. For example, in the preceding example with the crude birth rates of 1940 and 1988, the compositional effect and the rate effect are 0.620 and 2.940 (with the interaction effect of -0.024); whereas, when the interaction effect is eliminated, the same main effects become 0.608 and 2.928. Thus, the interaction effect in the former case is distributed equally between the two main effects in the latter situation.

As the same example suggests, the interaction term arises because of our using 1940 as the standard population. There is no reason why 1940 should be used as the standard, particularly when the use of the average of the two populations leads to a neat solution without the interaction term. As Kitagawa (1955) has argued, "changes in rates and composition are seldom independent—rather, a change in one is likely to affect the other. It may be argued, therefore, that since both were changing during the period, a logical set of weights for summarizing changes in specific rates, for example, would be the average composition of the population during the period." Finding "average" populations as standards such that the difference between two rates can be expressed as the sum of only the main effects is the crux of the decomposition methodology used in this report.

Expressing the difference between two rates in terms of only the main effects can also be justified by expressing the rate in terms of a linear saturated model with interactions and then solving the unknowns from the same number of equations (see section A.2 in appendix A). It is possible to show that for such models, the **difference** between two rates is always free from two-factor interaction effects, regardless of the number of factors. Since for any set of data, the three-factor and higher order interaction terms are expected to be negligible, it makes sense to find meaningful ways to decompose the difference into the main effects of the factors only by absorbing the interactions into the main effects.

The effects of factors do not necessarily imply any causal relationships. They simply indicate the nature of the association of the factors with the phenomenon being measured. There might be some hidden forces behind the factors that are actually responsible for the numbers we allocate to different factors as effects, but identifying those forces is beyond the scope of the decomposition analysis.

Chapter 2. Rate as the Product of Factors

2.1 INTRODUCTION

The simplest of the decomposition-standardization problems is the situation in which a rate can be expressed as the product of several factors. Some examples are as follows. Bongaarts (1978) expressed the total fertility rate as the product of five factors, namely, the index of proportion married, the index of noncontraception, the index of induced abortion, the index of lactational infecundability, and the total fecundity rate (Example 2.4). For adolescent women, Nathanson and Kim (1989) wrote the proportion of women having a nonmarital live birth as the product of four factors, namely, the proportion of live births among nonmarital pregnancies, the proportion of pregnancies among sexually active single women, the proportion of sexually active women among single women, and the proportion of single women among all women (Example 2.3). Das Gupta (1991) expressed the crude birth rate as the product of the general fertility rate, the proportion of women in the childbearing ages among all women, and the proportion of women in the population (Example 2.2).

In terms of the last example above, if R_1 and R_2 are the crude birth rates in population 1 and population 2, respectively, then questions are addressed separately for the problem of decomposition and for the problem of standardization, but these two areas are tied together by some consistency conditions, as indicated below.

Problem of Standardization

1. What would be the crude birth rates in the two populations if only the general fertility rates in the two populations differed as they did, but if the other two factors, namely, the proportion of women in the childbearing ages among all women and the proportion of women in the population were identical? These conditional crude birth rates are the standardized birth rates controlled (or adjusted) for the latter two factors.
2. As in (1) above, if only the proportions of women in the childbearing ages among all women in the two populations differed as they did, what would be the standardized birth rates controlled for the general fertility rate and the proportion of women in the population?
3. Again, if only the proportions of women in the two populations differed as they did, what would be the standardized birth rates controlled for the general fertility rate and the proportion of women in the childbearing ages among all women?

Problem of Decomposition

4. How much of the difference $R_2 - R_1$ in the crude birth rates in the two populations can be attributed to the difference in their general fertility rates? This amount is the effect of the general fertility rate.
5. As in (4) above, how much of the difference $R_2 - R_1$ is the effect of the proportion of women in the childbearing ages among all women?
6. Again, how much of the difference $R_2 - R_1$ is the effect of the proportion of women in the population?

Consistency Conditions

The decomposition-standardization methodology should be developed in such a way that the results would satisfy the following relationships:

(i) The difference between the standardized rates in question (1) above should give the answer to question (4).

(ii) The difference between the standardized rates in question (2) should give the answer to question (5).

(iii) The difference between the standardized rates in question (3) should give the answer to question (6).

(iv) The answers to questions (4), (5), and (6) should add up to the total difference $R_2 - R_1$ between the crude birth rates in the two populations.

2.2 THE CASE OF TWO FACTORS

Let α and β be the two factors so that the rate R can be expressed as

$$R = \alpha\beta. \tag{2.1}$$

In population 1, α and β take on the values A and B; in population 2, the corresponding values are a and b. The rates R_1 and R_2 in population 1 and population 2 are then

$$R_1 = AB, \quad R_2 = ab . \tag{2.2}$$

Following Das Gupta (1991, formula 6), if the factor α differed in the two populations as it did, and if the factor β remained the same, we have

$$\beta\text{-standardized rate: in population 1} = \frac{b+B}{2}A , \tag{2.3}$$

$$\text{in population 2} = \frac{b+B}{2}a . \tag{2.4}$$

Similarly, if the factor β differed in the two populations while the factor α remained the same, we obtain

$$\alpha\text{-standardized rate: in population 1} = \frac{a+A}{2}B , \tag{2.5}$$

$$\text{in population 2} = \frac{a+A}{2}b . \tag{2.6}$$

Again, we can write the α-effect and β-effect as

$$\alpha\text{-effect} = \frac{b+B}{2}(a-A), \tag{2.7}$$

$$\beta\text{-effect} = \frac{a+A}{2}(b-B). \tag{2.8}$$

We notice that the α-effect in (2.7) is the difference between the β-standardized rates in (2.3) and (2.4), and the β-effect in (2.8) is the difference between the α-standardized rates in (2.5) and (2.6). Again, from (2.2), (2.7), and (2.8), we have the identity

$$R_2 - R_1 = \alpha\text{-effect} + \beta\text{-effect} . \tag{2.9}$$

Therefore, all the consistency conditions in section 2.1 for two factors are satisfied.

Example 2.1

In the data for Black males and White males in table 2.1, equation (2.1) takes on the form

$$\text{Mean earnings based on all persons (R)} = \text{Mean earnings based on those who earned } (\alpha) \times \text{Proportion of persons who earned } (\beta)\,. \quad (2.10)$$

The results shown in table 2.2 can be summarized as follows:

1. The mean earnings (based on all persons) for Black males and White males are $7,846.56 and $13,703.73, respectively. The difference (total effect) is $5,857.17.
2. If the proportions of persons who earned were identical in the two populations, the standardized mean earnings would be $8,437.23 and $12,807.14, respectively. The difference, $4,369.91, gives the effect of the difference in the mean earnings of the earners in the two populations.
3. If the mean earnings of the earners were identical in the two populations, the standardized mean earnings would be $9,878.55 and $11,365.81, respectively. The difference, $1,487.26, gives the effect of the difference in the proportion of earners in the two populations.
4. As expected, the total effect in (1) above is equal to the sum of the effects in (2) and (3). Since both the effects are positive, we can meaningfully express them as percentages of the total effect. Thus, 74.6 percent of the difference between the mean earnings of Black males and White males based on all persons can be attributed to the difference in the mean earnings of the earners. The remaining 25.4 percent can be attributed to the difference in the proportion of earners in the two populations.

Table 2.1. Mean Earnings as the Product of Two Factors for Black Males and White Males, 18 Years and Over: United States, 1980

Measures	Black males (population 1)	White males (population 2)
Mean earnings $= \frac{\text{Total earnings}}{\text{Total population}}$ $(=R)$	$7,846.56 $(=R_1)$	$13,703.73 $(=R_2)$
$\frac{\text{Total earnings}}{\text{Persons who earned}}$ $(=\alpha)$	$10,930 (=A)	$16,591 (=a)
$\frac{\text{Persons who earned}}{\text{Total population}}$ $(=\beta)$	0.717892 (=B)	0.825974 (=b)

Source: U.S. Bureau of the Census (1984a), table 296.

Table 2.2. Standardization and Decomposition of Mean Earnings in Table 2.1

Measures	Standardization		Decomposition	
	White males (population 2)	Black males (population 1)	Difference (effects)	Percent distribution of effects
β-standardized mean earnings [Formulas (2.3) and (2.4)]	$12,807.14	$8,437.23	$4,369.91 ($\alpha$-effect)	74.6
α-standardized mean earnings [Formulas (2.5) and (2.6)]	$11,365.81	$9,878.55	$1,487.26 ($\beta$-effect)	25.4
Mean earnings (R)	$13,703.73	$7,846.56	$5,857.17 (Total effect)	100.0

2.3 THE CASE OF THREE FACTORS

In this case, the rate R can be expressed as

$$R = \alpha\beta\gamma\,, \quad (2.11)$$

where α, β, and γ are the three factors. If these factors assume the values A, B, and C in population 1, and a, b, and c in population 2, then the rates R_1 and R_2 in the two populations are

$$R_1 = ABC \ , \quad R_2 = abc \ . \tag{2.12}$$

From Das Gupta (1991, formula 7), we have

$$\beta\gamma\text{-standardized rate: in population 1} = \left[\frac{bc+BC}{3} + \frac{bC+Bc}{6}\right]A \ , \tag{2.13}$$

$$\text{in population 2} = \left[\frac{bc+BC}{3} + \frac{bC+Bc}{6}\right]a \ , \tag{2.14}$$

$$\alpha\gamma\text{-standardized rate: in population 1} = \left[\frac{ac+AC}{3} + \frac{aC+Ac}{6}\right]B \ , \tag{2.15}$$

$$\text{in population 2} = \left[\frac{ac+AC}{3} + \frac{aC+Ac}{6}\right]b \ , \tag{2.16}$$

$$\alpha\beta\text{-standardized rate: in population 1} = \left[\frac{ab+AB}{3} + \frac{aB+Ab}{6}\right]C \ , \tag{2.17}$$

$$\text{in population 2} = \left[\frac{ab+AB}{3} + \frac{aB+Ab}{6}\right]c \ . \tag{2.18}$$

Also, consistent with the above standardized rates, the factor effects have the following expressions:

$$\alpha\text{-effect} = \left[\frac{bc+BC}{3} + \frac{bC+Bc}{6}\right](a-A) \ , \tag{2.19}$$

$$\beta\text{-effect} = \left[\frac{ac+AC}{3} + \frac{aC+Ac}{6}\right](b-B) \ , \tag{2.20}$$

$$\gamma\text{-effect} = \left[\frac{ab+AB}{3} + \frac{aB+Ab}{6}\right](c-C) \ . \tag{2.21}$$

It is easy to verify from (2.12) and (2.19) through (2.21) that

$$R_2 - R_1 = \alpha\text{-effect} + \beta\text{-effect} + \gamma\text{-effect}. \tag{2.22}$$

Example 2.2

The data in table 2.3 are for Austria and Chile, 1981, in which equation (2.11) assumes the form, as in Das Gupta (1991, equation 11),

Crude birth rate (R) = General fertility rate (α)
x Proportion of women in the childbearing ages among all women (β)
x Proportion of women in the total population (γ).

(2.23)

For convenience, i.e., for making the difference R_2 - R_1 a positive number, we assume Chile, 1981, and Austria, 1981, to be population 2 and population 1, respectively, although the results and the conclusions do not depend on how the two populations are labeled. We will follow this rule of positive R_2 - R_1 in all our examples.

The results in table 2.4 show that the crude birth rates for Chile, 1981, and Austria, 1981, were 32.845 and 12.512 , giving a total difference of 20.333. However, if these rates are standardized with respect to the proportion of women in the childbearing ages among all women and the proportion of women in the population, then the standardized rates become 26.750 and 16.310, producing a difference of 10.440, and this difference is the effect of the difference in the general fertility rates. In other words, the difference between the birth rates for Chile and Austria would have been significantly smaller had the factors other than the general fertility rate been identical in the two populations. Other standardized rates in table 2.4 reveal that the effect of the difference in the proportion of women in the childbearing ages was to make the birth rate for Chile 10.559 points higher than that for Austria. On the other hand, the effect of the difference in the proportion of women in the population was to raise the birth rate for Austria 0.666 point above that for Chile. We have expressed the effects in terms of the percentages of the total effect in the last column of table 2.4, and we will show this percent distribution in all our examples. However, it is easier to interpret these percentages when the factor effects are positive, as in Example 2.1. If an effect is negative, we may ignore the percent of this effect in the last column and interpret the result in terms of the numbers in the preceding three columns.

Table 2.3. Crude Birth Rates as the Product of Three Factors: Austria and Chile, 1981

Measures	Austria, 1981 (population 1)	Chile, 1981 (population 2)
Crude birth rate $= \frac{\text{Births x 1000}}{\text{Total population}}$ $(= R)$	12.512 $(=R_1)$	32.845 $(=R_2)$
General fertility rate $= \frac{\text{Births x 1000}}{\text{Women aged 15-49}}$ $(= \alpha)$	51.76746 (=A)	84.90502 (=a)
$\frac{\text{Women aged 15-49}}{\text{Total women}}$ $(= \beta)$	0.45919 (=B)	0.75756 (=b)
$\frac{\text{Total women}}{\text{Total population}}$ $(= \gamma)$	0.52638 (=C)	0.51065 (=c)

Source: United Nations (1988, table 23; 1989, table 29).

Table 2.4. Standardization and Decomposition of Crude Birth Rates in Table 2.3

Measures	Standardization		Decomposition	
	Chile, 1981 (population 2)	Austria, 1981 (population 1)	Difference (effects)	Percent distribution of effects
$\beta\gamma$-standardized birth rates [Formulas (2.13) and (2.14)]	26.750	16.310	10.440 (α-effect)	51.4
$\alpha\gamma$-standardized birth rates [Formulas (2.15) and (2.16)]	26.810	16.251	10.559 (β-effect)	51.9
$\alpha\beta$-standardized birth rates [Formulas (2.17) and (2.18)]	21.651	22.317	−.666 (γ-effect)	−3.3
Crude birth rates (R)	32.845	12.512	20.333 (Total effect)	100.0

2.4 THE CASE OF FOUR FACTORS

When there are four factors α, β, γ, and δ, the rate R is written as

$$R=\alpha\beta\gamma\delta \quad , \tag{2.24}$$

and, using similar notation, we can write the rates in population 1 and population 2 as

$$R_1 = ABCD, \quad R_2 = abcd. \tag{2.25}$$

From Das Gupta (1991, formula 8), we obtain

$$\beta\gamma\delta\text{-standardized rate: in population } 1 = QA, \tag{2.26}$$
$$\text{in population } 2 = Qa, \tag{2.27}$$

so that

$$\alpha\text{-effect} = Q\,(a-A), \tag{2.28}$$

where Q is a function of b,c,d,B,C,D given by

$$Q = Q\,(b, c, d, B, C, D) = \frac{bcd + BCD}{4} + \frac{bcD+bCd+Bcd+BCd+BcD+bCD}{12}. \tag{2.29}$$

Other standardized rates and factor effects can be derived easily by interchanging the letters in equations (2.26) through (2.29). For example, the $\alpha\gamma\delta$-standardized rates and β-effect are obtained by substituting b,a,B,A for a,b,A,B, respectively.

Example 2.3

Table 2.5 provides the data for the example given in Nathanson and Kim (1989). Here, the rates in (2.24) for the White women aged 15 to 19 for 1971 and 1979 are expressed as follows:

Percentage having nonmarital live births (R)
= Percentage having nonmarital live births among nonmarital pregnancies (α)
x Proportion of nonmarital pregnancies among sexually active single women (β)
x Proportion of sexually active single women among total single women (γ)
x Proportion of single women among all women (δ).
(2.30)

The percentages R for 1971 and 1979 are, respectively, 1.434 and 4.423, giving a total difference of 2.989. The eight standardized rates for the two years (standardizing with respect to three factors at a time and allowing the fourth factor to vary) are given in table 2.6. For example, if only the proportions of sexually active single women among total single women (γ) varied as they did in 1971 and 1979, and all the remaining three factors (α, β, and δ) were identical in the two years, then the standardized percentages having nonmarital live births would be 1.989 and 3.372 in 1971 and 1979, respectively, producing a difference of 1.383 as the γ-effect. In other words, as shown in the last column of table 2.6, 46.3 percent of the increase in the percentage having nonmarital live births between 1971 and 1979 can be attributed to the increase in the proportion of sexually active single women among total single women (γ) in the 8-year period. We can make similar comments on other standardized rates and factor effects. The decomposition in table 2.6 agrees with the results shown in table 2 of Nathanson and Kim. The extension of this example to all live births as a 6-factor case is shown in Example 3.6.

Program 2.1

The results in table 2.6 can be easily obtained by using the computer program in FORTRAN (Program 2.1) in which P(1,J)'s are A, B, C, and D and P(2,J)'s are a, b, c, and d from table 2.5, the format of the data input being given in line 3 of the program. The subscripts I, J, and K in R(I,J,K) in line 7 refer to the two populations (1 and 2); the four factors (1, 2, 3, and 4); and the two expressions (1 and 2) on the right-hand side of (2.29). Attaching a value of 1 to the capital letters and a value of 2 to the small letters in (2.29), and adding these values for each three-letter term, we find that the first expression in (2.29) includes terms with 3 and 6 points; the second expression includes terms with 4 and 5 points. M1 and M2 in lines 16 and 17 of the program for M = 1,2 give the above two pairs of points, namely, (3,6) and (4,5). S(I,J)'s in line 24 are the eight standardized rates, and E(J)'s in line 25 are the four factor effects in table 2.6. R2, R1, and T in line 26 are the numbers in the last row of table 2.6 giving R_2 and R_1 in (2.25) and their difference.

Table 2.5. Percentage Having Nonmarital Live Births as the Product of Four Factors for White Women Aged 15 to 19: United States, 1971 and 1979

Measures	1971 (population 1)	1979 (population 2)
$\frac{\text{Nonmarital live births} \times 100}{\text{Total women}}$ $(=R)$	1.434 $(=R_1)$	4.423 $(=R_2)$
$\frac{\text{Nonmarital live births} \times 100}{\text{Nonmarital pregnancies}}$ $(=\alpha)$	25.3 (=A)	32.7 (=a)
$\frac{\text{Nonmarital pregnancies}}{\text{Sexually active single women}}$ $(=\beta)$	.214 (=B)	.290 (=b)
$\frac{\text{Sexually active single women}}{\text{Total single women}}$ $(=\gamma)$	.279 (=C)	.473 (=c)
$\frac{\text{Total single women}}{\text{Total women}}$ $(=\delta)$	.949 (=D)	.986 (=d)

Source: Nathanson and Kim (1989), table 1.

Table 2.6 Standardization and Decomposition of Percentages Having Nonmarital Live Births in Table 2.5

Measures	Standardization		Decomposition	
	1979 (population 2)	1971 (population 1)	Difference (effects)	Percent distribution of effects
$\beta\gamma\delta$-standardized percentages [Formulas (2.26) and (2.27)]	3.044	2.355	0.689 (α-effect)	23.0
$\alpha\gamma\delta$-standardized percentages	3.100	2.288	0.812 (β-effect)	27.2
$\alpha\beta\delta$-standardized percentages	3.372	1.989	1.383 (γ-effect)	46.3
$\alpha\beta\gamma$-standardized percentages	2.792	2.687	0.105 (δ-effect)	3.5
Percentages having nonmarital live births (R)	4.423	1.434	2.989 (Total effect)	100.0

2.5 THE CASE OF FIVE FACTORS

In this case, using analogous notation, we can write the rate as

$$R = \alpha\beta\gamma\delta\epsilon \,, \qquad (2.31)$$

which assumes the values

$$R_1 = ABCDE, \qquad R_2 = abcde, \qquad (2.32)$$

Program 2.1 (Four Factors)

```
      DIMENSION P(2,4),R(2,4,2),E(4),S(2,4)
      READ(5,1) ((P(I,J),J=1,4),I=1,2)
    1 FORMAT(F6.1,3F6.3)
      DO 2 I=1,2
      DO 2 J=1,4
      DO 2 K=1,2
    2 R(I,J,K)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      DO 3 K=1,2
      DO 3 L=1,2
      H=P(I,1)*P(J,2)*P(K,3)*P(L,4)
      IF(I+J+K+L.EQ.4) R1=H
      IF(I+J+K+L.EQ.8) R2=H
      DO 3 M=1,2
      M1=M+2
      M2=9-M1
      IF(J+K+L.EQ.M1.OR.J+K+L.EQ.M2) R(I,1,M)=R(I,1,M)+H
      IF(I+K+L.EQ.M1.OR.I+K+L.EQ.M2) R(J,2,M)=R(J,2,M)+H
      IF(I+J+L.EQ.M1.OR.I+J+L.EQ.M2) R(K,3,M)=R(K,3,M)+H
    3 IF(I+J+K.EQ.M1.OR.I+J+K.EQ.M2) R(L,4,M)=R(L,4,M)+H
      DO 5 J=1,4
      DO 4 I=1,2
    4 S(I,J)=R(I,J,1)/4.+R(I,J,2)/12.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,4),R2,R1,T
    6 FORMAT(40X,3F15.3)
      STOP
      END
****
```

Program 2.2 (Five Factors)

```
      DIMENSION P(2,5),R(2,5,3),E(5),S(2,5)
      READ(5,1) ((P(I,J),J=1,5),I=1,2)
    1 FORMAT(4F8.2,F8.3)
      DO 2 I=1,2
      DO 2 J=1,5
      DO 2 K=1,3
    2 R(I,J,K)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      DO 3 K=1,2
      DO 3 L=1,2
      DO 3 M=1,2
      H=P(I,1)*P(J,2)*P(K,3)*P(L,4)*P(M,5)
      IF(I+J+K+L+M.EQ. 5) R1=H
      IF(I+J+K+L+M.EQ.10) R2=H
      DO 3 N=1,3
      N1=N+3
      N2=12-N1
      IF(J+K+L+M.EQ.N1.OR.J+K+L+M.EQ.N2) R(I,1,N)=R(I,1,N)+H
      IF(I+K+L+M.EQ.N1.OR.I+K+L+M.EQ.N2) R(J,2,N)=R(J,2,N)+H
      IF(I+J+L+M.EQ.N1.OR.I+J+L+M.EQ.N2) R(K,3,N)=R(K,3,N)+H
      IF(I+J+K+M.EQ.N1.OR.I+J+K+M.EQ.N2) R(L,4,N)=R(L,4,N)+H
    3 IF(I+J+K+L.EQ.N1.OR.I+J+K+L.EQ.N2) R(M,5,N)=R(M,5,N)+H
      DO 5 J=1,5
      DO 4 I=1,2
    4 S(I,J)=R(I,J,1)/5.+R(I,J,2)/20.+R(I,J,3)/30.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,5),R2,R1,T
    6 FORMAT(40X,3F15.2)
      STOP
      END
****
```

in population 1 and population 2, respectively.

Using formula 9 in Das Gupta (1991), we have

$$\beta\gamma\delta\epsilon\text{-standardized rate: in population 1} = QA, \tag{2.33}$$

$$\text{in population 2} = Qa, \tag{2.34}$$

so that

$$\alpha\text{-effect} = Q\,(a-A)\ , \tag{2.35}$$

where Q is a function of b,c,d,e,B,C,D,E given by

$$Q = Q(b,c,d,e,B,C,D,E) = \frac{bcde+BCDE}{5} + \frac{bcdE+bcDe+bCde+Bcde+BCDe+BCdE+BcDE+bCDE}{20} + \frac{bcDE+bCdE+bCDe+BCde+BcDe+BcdE}{30}. \tag{2.36}$$

Other standardized rates and factor effects follow directly from those in (2.33) through (2.36).

Example 2.4

Bongaarts (1978) expressed the total fertility rate (TFR) as

$$TFR = C_m \times C_c \times C_a \times C_i \times TF, \tag{2.37}$$

where C_m , C_c , C_a , C_i are, respectively, the indices of proportion married, noncontraception, induced abortion, and lactational infecundability, and TF is the total fecundity rate. We can treat equation (2.37) as equation (2.31) expressing R in terms of five factors α, β, γ, δ, and ϵ. The data corresponding to this equation are given in table 2.7 for South Korea for 1960 and 1970. The results from the application of the standardization and decomposition techniques to these data are shown in table 2.8.

The total fertility rate in South Korea declined 2.08 points during 1960 to 1970, from 6.13 in 1960 to 4.05 in 1970. This decline would have been only 1.23 points (from 5.68 in 1960 to 4.45 in 1970) if only the index of noncontraception (β) declined as it did during 1960 to 1970, and the other four factors were identical. In other words, 59.1 percent of the total decline in the total fertility rate in the decade can be attributed to the increased use of contraception during the same period. Similar conclusions can be drawn from the other standardized rates and factor effects in table 2.8. Again, we should ignore the negative percents in the last column and interpret these results from the corresponding numbers in the other columns. Although Bongaarts provided the data for this example, he did not do any computations for standardization or decomposition similar to those in table 2.8.

Moreno (1991, table 8) used a shorter version of the model in equation. (2.37) given by

$$TFR = C_m \times C_c \times C_i \times \text{Other}, \tag{2.38}$$

for six Latin American countries to decompose the difference between the total fertility rates from the World Fertility Survey and the Demographic and Health Survey, and his results involved interaction terms. The four-factor formulas for standardization and decomposition given in section 2.4 can be easily applied to his data to obtain the results without the interaction terms. The justification for not including the interaction terms separately but absorbing them into the main effects is given in chapter 1.

Table 2.7. Total Fertility Rate as the Product of Five Factors: South Korea, 1960 and 1970

Measures	1970 (population 1)	1960 (population 2)
Total fertility rate (=R)	4.05 ($=R_1$)	6.13 ($=R_2$)
Index of proportion married ($=\alpha$)	0.58 (=A)	0.72 (=a)
Index of noncontraception ($=\beta$)	0.76 (=B)	0.97 (=b)
Index of induced abortion ($=\gamma$)	0.84 (=C)	0.97 (=c)
Index of lactational infecundability ($=\delta$)	0.66 (=D)	0.56 (=d)
Total fecundity rate ($=\epsilon$)	16.573 (=E)	16.158 (=e)

Source: Bongaarts (1978), table 3.

Table 2.8. Standardization and Decomposition of Total Fertility Rates in Table 2.7

Measures	Standardization		Decomposition	
	1960 (population 2)	1970 (population 1)	Difference (effects)	Percent distribution of effects
$\beta\gamma\delta\epsilon$-standardized TFR's [Formulas (2.33) and (2.34)]	5.61	4.52	1.09 (α-effect)	52.4
$\alpha\gamma\delta\epsilon$-standardized TFR's	5.68	4.45	1.23 (β-effect)	59.1
$\alpha\beta\delta\epsilon$-standardized TFR's	5.43	4.70	0.73 (γ-effect)	35.1
$\alpha\beta\gamma\epsilon$-standardized TFR's	4.70	5.54	−0.84 (δ-effect)	−40.4
$\alpha\beta\gamma\delta$-standardized TFR's	5.02	5.15	−0.13 (ϵ-effect)	−6.2
Total fertility rates (R)	6.13	4.05	2.08 (Total effect)	100.0

Program 2.2

The results in table 2.8 can be obtained from Program 2.2, which is almost identical with Program 2.1 except for the minor changes needed for the change in the number of factors from four to five. As before, P(I,J)'s are input data A, B, C, D, E and a, b, c, d, e from table 2.7. The subscripts I, J, K in R(I,J,K) in this program refer to the two populations, the five factors, and the three expressions on the right-hand side of (2.36). Again, attaching a value of 1 to the capital letters and a value of 2 to the small letters in (2.36), and then adding these values for each four-letter term, we find that the first, second, and third expressions in (2.36) include terms with points (4,8), (5,7), and 6, respectively. Accordingly, N1 and N2 in lines 17 and 18 correspond to the three pairs (4,8), (5,7), and (6,6). As in Program 2.1, S(I,J)'s in line 26 are the 10 standardized rates, and E(J)'s in line 27 are the five factor effects in table 2.8. Again, R2, R1, and T in line 28 give the numbers in the last row of table 2.8.

2.6 THE CASE OF SIX FACTORS

When there are six factors so that

$$R = \alpha\beta\gamma\delta\epsilon\eta \, , \qquad (2.39)$$

and in the two populations,

$$R_1 = ABCDE\,F \, , \quad R_2 = abcdef \, , \qquad (2.40)$$

then, formula 10 in Das Gupta (1991) gives

$$\beta\gamma\delta\epsilon\eta\text{-standardized rate: in population } 1 = QA\,, \quad (2.41)$$
$$\text{in population } 2 = Qa\,, \quad (2.42)$$

so that

$$\alpha\text{-effect} = Q\,(a-A)\,, \quad (2.43)$$

where

$$Q = Q(b,c,d,e,f,B,C,D,E,F) = \frac{bcdef+BCDEF}{6}$$
$$+\frac{bcdeF+bcdEf+bcDef+bCdef+Bcdef+BCDEf+BCDeF+BCdEF+BcDEF+bCDEF}{30}$$
$$+\frac{\begin{array}{c}bcdEF+bcDeF+bcDEf+bCdeF+bCdEf+bCDef+BcdeF+BcdEf+BcDef+BCdef\\+BCDef+BCdEf+BCdeF+BcDEf+BcDeF+BcdEF+bCDEf+bCDeF+bCdEF+bcDEF\end{array}}{60}. \quad (2.44)$$

Other rates and effects can be easily obtained from (2.41) through (2.44).

2.7 THE CASE OF P FACTORS

Let us write the rate as the product of P factors as

$$R = \alpha_1\alpha_2....\alpha_p\,. \quad (2.45)$$

In the two populations, this rate assumes the values

$$R_1 = A_1A_2\ldots A_p\;,\;\; R_2 = a_1a_2\ldots a_p\,. \quad (2.46)$$

It follows from formula A6 in Das Gupta (1991) that

$$\alpha_2\alpha_3...\alpha_p\text{-standardized rate: in population } 1 = QA_1\,, \quad (2.47)$$
$$\text{in population } 2 = Qa_1\,, \quad (2.48)$$

so that

$$\alpha_1\text{-effect} = Q\,(a_1-A_1)\,, \quad (2.49)$$

where

$$Q = Q(a_2,a_3, \dots ,a_p, A_2,A_3,\dots ,A_p) = \frac{a_2a_3\dots a_p + A_2A_3\dots A_p}{P}$$

$$+ \frac{\text{sum of all } (P-1)\text{-letter terms with } (P-2) \text{ small letters and } 1 \text{ capital letter or } (P-2) \text{ capital letters and } 1 \text{ small letter}}{P\binom{P-1}{1}}$$

$$+ \frac{\text{sum of all } (P-1)\text{-letter terms with } (P-3) \text{ small letters and } 2 \text{ capital letters or } (P-3) \text{ capital letters and } 2 \text{ small letters}}{P\binom{P-1}{2}}$$

$$+ \dots\dots\dots\dots$$

$$= \sum_{r=1}^{S} \frac{\text{sum of all } (P-1)\text{-letter terms with } (P-r) \text{ small letters and } (r-1) \text{ capital letters or } (P-r) \text{ capital letters and } (r-1) \text{ small letters}}{P\binom{P-1}{r-1}}, \quad (2.50)$$

where

$$S = P/2\text{, when P is even,}$$
$$= (P+1)/2\text{, when P is odd.}$$

2.8 THE GENERAL PROGRAM

From Programs 2.1 and 2.2 corresponding to four and five factors, it is clear how to develop a FORTRAN program for any number of factors higher than five. However, it is not necessary to use different programs for data involving different numbers of factors. A program written for, say, 10-factor data can be used for any number of factors not exceeding 10 by changing the expression for the rate R and the input and output statements and formats in the program. No changes are necessary in the data files previously created to be used with the specific programs.

Assuming that no one is expected to deal with more than 10 multiplicative factors, we provide below a program (Program 2.3) for 10 factors that can be used as a general program for any number of factors up to 10. In order to show how to use this program for a smaller number of factors, we again consider Examples 2.1 through 2.4 involving two to five factors, and indicate the changes in the input and output statements (lines 2,42); the input and output formats (lines 3,43); and in the expression for the rate (lines 18,19) in Program 2.3 that are necessary in these examples to generate the results in tables 2.2, 2.4, 2.6, and 2.8, respectively:

Example 2.1 (two factors)

Lines 2,42: Replace 10 in each line by 2
Lines 3,43: Replace (10F8.4) by (F8.0, F8.6) and 15.3 by 15.2
Lines 18,19: Replace the two lines by H = P(I,1)*P(J,2)

Example 2.2 (three factors)

Lines 2,42: Replace 10 in each line by 3
Lines 3,43: Replace (10F8.4) by (3F10.5) and no change in line 43
Lines 18,19: Replace the two lines by H = P(I,1)*P(J,2)*P(K,3)

Example 2.3 (four factors)

Lines 2,42: Replace 10 in each line by 4
Lines 3,43: Replace (10F8.4) by (F6.1, 3F6.3) and no change in line 43
Lines 18,19: Replace by H = P(I,1)*P(J,2)*P(K,3)*P(L,4)

Example 2.4 (five factors)

Lines 2,42: Replace 10 in each line by 5
Lines 3,43: Replace (10F8.4) by (4F8.2, F8.3) and 15.3 by 15.2
Lines 18,19: Replace by H = P(I,1)*P(J,2)*P(K,3)*P(L,4)*P(M,5).

Program 2.3 (General Program for up to Ten Factors)

```
      DIMENSION P(2,10),R(2,10,5),E(10),S(2,10)
      READ(5,1) ((P(I,J),J=1,10),I=1,2)
    1 FORMAT(10F8.4)
      DO 2 I=1,2
      DO 2 J=1,10
      DO 2 K=1,5
    2 R(I,J,K)=0.0
      DO 3  I=1,2
      DO 3  J=1,2
      DO 3  K=1,2
      DO 3  L=1,2
      DO 3  M=1,2
      DO 3  N=1,2
      DO 3 II=1,2
      DO 3 JJ=1,2
      DO 3 KK=1,2
      DO 3 LL=1,2
      H=P(I,1)*P(J,2)*P(K,3)*P(L,4)*P(M,5)*P(N,6)*
     1  P(II,7)*P(JJ,8)*P(KK,9)*P(LL,10)
      MM=I+J+K+L+M+N+II+JJ+KK+LL
      IF(MM.EQ.10) R1=H
      IF(MM.EQ.20) R2=H
      DO 3 NN=1,5
      N1=NN+8
      N2=27-N1
      IF(MM- I.EQ.N1.OR.MM- I.EQ.N2) R( I, 1,NN)=R( I, 1,NN)+H
      IF(MM- J.EQ.N1.OR.MM- J.EQ.N2) R( J, 2,NN)=R( J, 2,NN)+H
      IF(MM- K.EQ.N1.OR.MM- K.EQ.N2) R( K, 3,NN)=R( K, 3,NN)+H
      IF(MM- L.EQ.N1.OR.MM- L.EQ.N2) R( L, 4,NN)=R( L, 4,NN)+H
      IF(MM- M.EQ.N1.OR.MM- M.EQ.N2) R( M, 5,NN)=R( M, 5,NN)+H
      IF(MM- N.EQ.N1.OR.MM- N.EQ.N2) R( N, 6,NN)=R( N, 6,NN)+H
      IF(MM-II.EQ.N1.OR.MM-II.EQ.N2) R(II, 7,NN)=R(II, 7,NN)+H
      IF(MM-JJ.EQ.N1.OR.MM-JJ.EQ.N2) R(JJ, 8,NN)=R(JJ, 8,NN)+H
      IF(MM-KK.EQ.N1.OR.MM-KK.EQ.N2) R(KK, 9,NN)=R(KK, 9,NN)+H
    3 IF(MM-LL.EQ.N1.OR.MM-LL.EQ.N2) R(LL,10,NN)=R(LL,10,NN)+H
      DO 5 J=1,10
      DO 4 I=1,2
    4 S(I,J)=R(I,J,1)/10.+R(I,J,2)/90.+R(I,J,3)/360.+R(I,J,4)/840.+
     1 R(I,J,5)/1260.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,10),R2,R1,T
    6 FORMAT(40X,3F15.3)
      STOP
      END
```

Chapter 3. Rate as a Function of Factors

3.1 INTRODUCTION

A more general case of standardization and decomposition than that in the preceding chapter is the situation in which the rate can be expressed as any function of two or more factors. Obviously, the rate expressed as the product of factors in chapter 2 is a special case of the present situation. To give an example of a rate that is a function of factors, Pullum, Tedrow, and Herting (1989) expressed the mean parity of a cohort of women as a function of the parity progression ratios (Example 3.7). Again, based on the study by Wojtkiewicz, McLanahan, and Garfinkel (1990), the family headship rate of mothers can be expressed as a function of six factors (Example 3.5). These and other examples of rates expressed as functions of factors are used in this chapter to illustrate the standardization of rates and the corresponding decomposition of rate differences.

3.2 THE CASE OF TWO FACTORS

If there are two factors α and β, the rate R in this case is a function given by

$$R = F(\alpha,\beta)\,. \tag{3.1}$$

If the factors α and β take on the values A and B in population 1 and the values a and b in population 2, then the rates R_1 and R_2 in population 1 and population 2 are

$$R_1 = F(A,B)\,,\quad R_2 = F(a,b)\,. \tag{3.2}$$

If the factor α differed in the two populations as it did, and if the factor β remained the same, then it follows from Das Gupta (1991, formula 1) that

$$\beta\text{-standardized rate: in population 1} = \frac{F(A,b)+F(A,B)}{2}, \tag{3.3}$$

$$\text{in population 2} = \frac{F(a,b)+F(a,B)}{2}. \tag{3.4}$$

Similarly, if the factor β differed in the two populations and the factor α remained the same, we have

$$\alpha\text{-standardized rate: in population 1} = \frac{F(a,B)+F(A,B)}{2}, \tag{3.5}$$

$$\text{in population 2} = \frac{F(a,b)+F(A,b)}{2}. \tag{3.6}$$

The α-effect, as the difference between (3.3) and (3.4), and the β-effect, as the difference between (3.5) and (3.6), are

$$\alpha\text{-effect} = \frac{[F(a,b)-F(A,b)] + [F(a,B)-F(A,B)]}{2}, \tag{3.7}$$

$$\beta\text{-effect} = \frac{[F(a,b)-F(a,B)] + [F(A,b)-F(A,B)]}{2} \tag{3.8}$$

It is easy to verify from (3.2), (3.7), and (3.8) that the sum of the two effects is equal to the difference between the two rates, as in (2.9).

Example 3.1

In the data for 1940 and 1960 in table 3.1, equation (3.1) takes on the form

$$\underset{\text{increase (R)}}{\text{Crude rate of natural}} = \underset{\text{rate }(\alpha)}{\text{Crude birth}} - \underset{\text{rate }(\beta)}{\text{Crude death}} \qquad (3.9)$$

As shown in table 3.2, the crude rates of natural increase in 1940 and 1960 are 8.60 and 14.20, respectively, their difference being 5.60 (the total effect). If the death rates were identical in the two years, the standardized rates of natural increase would be 9.25 and 13.55, respectively, their difference of 4.30 giving the effect of the difference in the birth rates in the two years. Similarly, the rates of natural increase standardized for birth rate are 10.75 and 12.05 for 1940 and 1960, their difference of 1.30 indicating the effect of the difference in the death rates. As expected, the birth-rate effect and the death-rate effect add up to the total effect. In terms of percentages, 76.8 percent of the change in the rate of natural increase during 1940-1960 can be attributed to the difference in the birth rates and the remaining 23.2 percent, to the difference in the death rates.

Table 3.1 Crude Rate of Natural Increase as a Function of Crude Birth Rate and Crude Death Rate: United States, 1940 and 1960

Measures	1940 (population 1)	1960 (population 2)
Crude rate of natural increase	8.6 $(=R_1)$	14.2 $(=R_2)$
$= \dfrac{\text{(Births - Deaths) x 1000}}{\text{Total population}} = F(\alpha,\beta) = \alpha\text{-}\beta \; (=R)$		
Crude birth rate $= \dfrac{\text{Births x 1000}}{\text{Total population}} \; (=\alpha)$	19.4 (=A)	23.7 (=a)
Crude death rate $= \dfrac{\text{Deaths x 1000}}{\text{Total population}} \; (=\beta)$	10.8 (=B)	9.5 (=b)

Source: National Center for Health Statistics (1990a, table 1-1; 1990b, table 1-2).

Table 3.2. Standardization and Decomposition of Crude Rates of Natural Increase in Table 3.1

Measures	Standardization		Decomposition	
	1960 (population 2)	1940 (population 1)	Difference (effects)	Percent distribution of effects
β-standardized rate of natural increase [Formulas (3.3) and (3.4)]	13.55	9.25	4.30 (α-effect)	76.8
α-standardized rate of natural increase [Formulas (3.5) and (3.6)]	12.05	10.75	1.30 (β-effect)	23.2
Crude rate of natural increase (R)	14.20	8.60	5.60 (Total effect)	100.0

3.3 THE CASE OF THREE FACTORS

In this case, the rate R can be expressed as

$$R = F(\alpha,\beta,\gamma) \; , \qquad (3.10)$$

where α, β, and γ are the three factors. If these factors assume the values A, B, and C in population 1 and a, b, and c in population 2, then the rates in the two populations are

$$R_1 = F(A,B,C) \ , \ R_2 = F(a,b,c) \ . \tag{3.11}$$

It follows from equation (2) in Das Gupta (1991) that

$$\beta\gamma\text{-standardized rate: in population 1} = Q(A) \ , \tag{3.12}$$
$$\text{in population 2} = Q(a) \ , \tag{3.13}$$
$$\alpha\gamma\text{-standardized rate: in population 1} = Q(B) \ , \tag{3.14}$$
$$\text{in population 2} = Q(b) \ , \tag{3.15}$$
$$\alpha\beta\text{-standardized rate: in population 1} = Q(C) \ , \tag{3.16}$$
$$\text{in population 2} = Q(c) \ , \tag{3.17}$$

so that

$$\alpha\text{-effect} = Q(a) - Q(A) \ , \tag{3.18}$$
$$\beta\text{-effect} = Q(b) - Q(B) \ , \tag{3.19}$$
$$\gamma\text{-effect} = Q(c) - Q(C) \ , \tag{3.20}$$

where

$$Q(A) = Q(A;\, b,c,B,C) = \frac{F(A,b,c)+F(A,B,C)}{3} + \frac{F(A,b,C)+F(A,B,c)}{6} \ , \tag{3.21}$$

$$Q(B) = Q(B;\, a,c,A,C) = \frac{F(a,B,c)+F(A,B,C)}{3} + \frac{F(a,B,C)+F(A,B,c)}{6} \ , \tag{3.22}$$

$$Q(C) = Q(C;\, a,b,A,B) = \frac{F(a,b,C)+F(A,B,C)}{3} + \frac{F(a,B,C)+F(A,b,C)}{6} \ , \tag{3.23}$$

and Q(a), Q(b), and Q(c) are, respectively, the same expressions as those in (3.21), (3.22), and (3.23) with A, B, and C replaced by a, b, and c.

We can verify from (3.11) and (3.18) through (3.20) that the three effects add up to the difference between the two rates, as in (2.22). The derivation of effects (3.18) through (3.20) and also their expressions when interactions between the factors are allowed are shown in sections A.1 and A.2 in appendix A.

Example 3.2

The data in table 3.3 for White women in the United States for 1963 and 1983 express the illegitimacy ratio (the ratio of births to unmarried women to total births) as

$$\frac{I}{I+L} = \frac{\frac{U}{W} \cdot \frac{I}{U}}{\frac{U}{W} \cdot \frac{I}{U} + \frac{M}{W} \cdot \frac{L}{M}} \ , \tag{3.24}$$

where U, M, and W are unmarried, married, and total women in the childbearing ages 15 to 44, and I and L are births to unmarried and married women.

Using our notation, equation (3.24) can be written as

$$R = F(\alpha,\beta,\gamma) = \frac{\alpha\beta}{\alpha\beta + (1-\alpha)\,\gamma} \ , \tag{3.25}$$

where α, β, γ represent, respectively, the proportion of unmarried women in the childbearing ages, the nonmarital general fertility rate, and the marital general fertility rate.

Table 3.4 shows that there was an increase of 94.23 in the illegitimacy ratio per 1,000 births in the 20-year period, from 30.95 in 1963 to 125.18 in 1983. If only the nonmarital general fertility rate (β) changed as it did during the two decades but the other two factors were identical, the illegitimacy ratios in 1963 and 1983 would be 50.89 and 87.63, their difference of 36.74 being the effect of the change in β. In other words, only 39.0 percent of the increase in the illegitimacy ratio during 1963-1983 can be attributed to the increase in nonmarital fertility. From the other standardized illegitimacy ratios in table 3.4, it follows that the increase in the proportion of unmarried women and the decrease in marital fertility during the period explain, respectively, 35.4 percent and 25.6 percent of the total increase in the illegitimacy ratio. This example will be discussed again in Example 4.4 with expanded data incorporating age.

Table 3.3. Illegitimacy Ratio for Whites as a Function of Three Factors: United States, 1963 and 1983

Measures	1963 (population 1)	1983 (population 2)
Illegitimacy ratio (=R)	.03095 (=R_1)	.12518 (=R_2)
Proportion unmarried among women aged 15 to 44 years (=α)	.295876 (=A)	.416950 (=a)
Nonmarital general fertility rate (=β)	.010569 (=B)	.019025 (=b)
Marital general fertility rate (=γ)	.139055 (=C)	.095082 (=c)

Source: Smith and Cutright (1988), table 2.

Table 3.4. Standardization and Decomposition of Illegitimacy Ratios in Table 3.3

(For convenience, results obtained from data in table 3.3 are multiplied by 1,000 before presenting them in table 3.4)

Measures	Standardization		Decomposition	
	1983 (population 2)	1963 (population 1)	Difference (effects)	Percent distribution of effects
$\beta\gamma$-standardized illegitimacy ratios	86.04	52.67	33.37 (α-effect)	35.4
$\alpha\gamma$-standardized illegitimacy ratios	87.63	50.89	36.74 (β-effect)	39.0
$\alpha\beta$-standardized illegitimacy ratios	81.80	57.68	24.12 (γ-effect)	25.6
Illegitimacy ratios (R)	125.18	30.95	94.23 (Total effect)	100.0

Program 3.1

The results in table 3.4 can be obtained by using Program 3.1 in which P(1,J)'s are A, B, and C and P(2,J)'s are a, b, and c from table 3.3, the format of the data input being given in line 3 of the program. The subscripts I, J, and K in R(I,J,K) in line 7 refer to the two populations, the three factors, and the two expressions on the right-hand sides of (3.21) through (3.23). Taking any one of these three equations, say, Q(A) in (3.21), we leave the argument A untouched but attach a value of 1 to the other capital letters and a value of 2 to the small letters, and add these two values of the arguments for each F. We find that the first expression in (3.21) includes F's with a total of 2 and 4 points for the arguments. The second expression includes F's with a total of 3 points. L1 and L2 in lines 15 and 16 of the program for L = 1,2

Program 3.1 (Three Factors)

```
      DIMENSION P(2,3),R(2,3,2),E(3),S(2,3)
      READ(5,1) ((P(I,J),J=1,3),I=1,2)
    1 FORMAT(3F10.6)
      DO 2 I=1,2
      DO 2 J=1,3
      DO 2 K=1,2
    2 R(I,J,K)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      DO 3 K=1,2
      H=P(I,1)*P(J,2)/(P(I,1)*P(J,2)+(1.-P(I,1))*P(K,3))
      IF(I+J+K.EQ.3) R1=H
      IF(I+J+K.EQ.6) R2=H
      DO 3 L=1,2
      L1=L+1
      L2=6-L1
      IF(J+K.EQ.L1.OR.J+K.EQ.L2) R(I,1,L)=R(I,1,L)+H
      IF(I+K.EQ.L1.OR.I+K.EQ.L2) R(J,2,L)=R(J,2,L)+H
    3 IF(I+J.EQ.L1.OR.I+J.EQ.L2) R(K,3,L)=R(K,3,L)+H
      DO 5 J=1,3
      DO 4 I=1,2
    4 S(I,J)=R(I,J,1)/3.+R(I,J,2)/6.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,3),R2,R1,T
    6 FORMAT(40X,3F15.5)
      STOP
      END
****
```

Program 3.2 (Four Factors)

```
      DIMENSION P(2,4),R(2,4,2),E(4),S(2,4)
      READ(5,1) ((P(I,J),J=1,4),I=1,2)
    1 FORMAT(4F10.5)
      DO 2 I=1,2
      DO 2 J=1,4
      DO 2 K=1,2
    2 R(I,J,K)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      DO 3 K=1,2
      DO 3 L=1,2
      H=(P(I,1)*P(J,2)+P(L,4)*(1.-P(J,2)))*P(K,3)
      IF(I+J+K+L.EQ.4) R1=H
      IF(I+J+K+L.EQ.8) R2=H
      DO 3 M=1,2
      M1=M+2
      M2=9-M1
      IF(J+K+L.EQ.M1.OR.J+K+L.EQ.M2) R(I,1,M)=R(I,1,M)+H
      IF(I+K+L.EQ.M1.OR.I+K+L.EQ.M2) R(J,2,M)=R(J,2,M)+H
      IF(I+J+L.EQ.M1.OR.I+J+L.EQ.M2) R(K,3,M)=R(K,3,M)+H
    3 IF(I+J+K.EQ.M1.OR.I+J+K.EQ.M2) R(L,4,M)=R(L,4,M)+H
      DO 5 J=1,4
      DO 4 I=1,2
    4 S(I,J)=R(I,J,1)/4.+R(I,J,2)/12.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,4),R2,R1,T
    6 FORMAT(40X,3F15.3)
      STOP
      END
****
```

give the above two pairs of points, namely, (2,4) and (3,3). H in line 11 is the expression for the rate R in (3.25). S(I,J)'s in line 22 are the six standardized rates, and E(J)'s in line 23 are the three factor effects in table 3.4. R2, R1, and T in line 24 are the numbers in the last row of table 3.4 giving R_2 and R_1 in (3.11) and their difference.

3.4 THE CASE OF FOUR FACTORS

When there are four factors α, β, γ, and δ, the rate R is written as

$$R = F(\alpha,\beta,\gamma,\delta)\,, \tag{3.26}$$

and, using similar notation, we can write the rates in population 1 and population 2 as

$$R_1 = F(A,B,C,D)\,,\; R_2 = F(a,b,c,d)\,. \tag{3.27}$$

It follows from equation (3) in Das Gupta (1991) that

$$\beta\gamma\delta\text{-standardized rate: in population 1} = Q(A)\,, \tag{3.28}$$

$$\text{in population 2} = Q(a)\,, \tag{3.29}$$

so that

$$\alpha\text{-effect} = Q(a) - Q(A)\,, \tag{3.30}$$

where

$$Q(A) = Q(A;b,c,d,B,C,D) = \frac{F(A,b,c,d) + F(A,B,C,D)}{4}$$

$$+\frac{F(A,b,c,D) + F(A,b,C,d) + F(A,B,c,d) + F(A,B,C,d) + F(A,B,c,D) + F(A,b,C,D)}{12}\,, \tag{3.31}$$

and Q(a) is the same expression as that in (3.31) with A replaced by a.

Other standardized rates and factor effects can be derived easily by interchanging the letters in equations (3.28) through (3.31).

Example 3.3

This is an extended version of Example 2.2 in which the data on marital and nonmarital births are used for Austria and Chile, 1981, as given in table 3.5. In this case, equation (3.26) assumes the form

$$R = [\alpha\beta + \delta(1-\beta)]\gamma\,, \tag{3.32}$$

where R = Crude birth rate per 1,000 population,
α = Marital general fertility rate
= Marital births per 1,000 married women aged 15 to 49,
β = Proportion of married women among all women aged 15 to 49,
γ = Proportion of women aged 15 to 49 in the total population,
ϵ = Nonmarital general fertility rate
δ = Nonmarital births per 1,000 unmarried women aged 15 to 49.

As shown in table 3.6, the crude birth rates for Chile, 1981, and Austria, 1981, were 32.845 and 12.512, giving a total difference of 20.333. If the proportion of women aged 15 to 49 in the population (γ) differed as it did in the two populations, but all other factors remained identical, then the standardized birth rates for Chile and Austria would be 26.497 and 16.556, their difference of 9.941 being the γ-effect. In other words, 48.9 percent of the excess of the crude birth rate in Chile over Austria is explained by the significantly higher ratio of women in the childbearing ages to the total population in Chile compared with that in Austria. Although the data in Example 2.2 are not exactly the same, this percentage of 48.9 is roughly equal to the combined effect of 48.6 percent for the factors β and γ in Example 2.2, as expected. If only the proportion of married women among all women in the childbearing ages (β) varied as it did, the birth rate in Austria would be 0.994 point higher than that in Chile. As before, the negative percent in the last column should be ignored, and the corresponding numbers in the three preceding columns should be used for interpretation.

Table 3.5. Crude Birth Rate as a Function of Four Factors: Austria and Chile, 1981

Measures	Austria, 1981 (population 1)	Chile, 1981 (population 2)
Crude birth rate $= \dfrac{\text{Births x 1000}}{\text{Total population}}$ (=R)	12.512 (=R_1)	32.845 (=R_2)
Marital general fertility rate per 1,000 (=α)	71.83691 (=A)	115.73732 (=a)
Proportion married among women aged 15 to 49 (=β)	0.58048 (=B)	0.52500 (=b)
Proportion of women aged 15 to 49 in the population (=γ)	0.24171 (=C)	0.38685 (=c)
Nonmarital general fertility rate per 1,000 (=δ)	23.99823 (=D)	50.82674 (=d)

Source: United Nations (1988, tables 23, 33; 1989, table 29).

Table 3.6. Standardization and Decomposition of Crude Birth Rates in Table 3.5

Measures	Standardization		Decomposition	
	Chile, 1981 (population 2)	Austria, 1981 (population 1)	Difference (effects)	Percent distribution of effects
$\beta\gamma\delta$-standardized birth rates	25.496	17.899	7.597 (α-effect)	37.4
$\alpha\gamma\delta$-standardized birth rates	21.493	22.487	−0.994 (β-effect)	−4.9
$\alpha\beta\delta$-standardized birth rates	26.497	16.556	9.941 (γ-effect)	48.9
$\alpha\beta\gamma$-standardized birth rates	23.638	19.849	3.789 (δ-effect)	18.6
Crude birth rates (R)	32.845	12.512	20.333 (Total effect)	100.0

Program 3.2

The results in table 3.6 can be obtained by using Program 3.2. This program is identical to Program 2.1 except for lines 3 and 12. The interpretations of the variables in Program 3.2 are the same as those for Program 2.1, except that the attachment of values of 1 and 2 should be described in a little different way, as indicated in the text for Program 3.1. Line 3 in Program 3.2 is consistent with the data format in table 3.5 (which is different from the data format in table 2.5). Also, H in Line 12 of Program 3.2 gives the expression for R in (3.32), whereas the same line in Program 2.1 gives the expression for R in (2.24).

3.5 THE CASE OF FIVE FACTORS

In this case, using analogous notation, we can write the rate as

$$R = F(\alpha,\beta,\gamma,\delta,\epsilon)\,, \tag{3.33}$$

which assumes the values

$$R_1 = F(A,B,C,D,E)\,,\ R_2 = (a,b,c,d,e)\,, \tag{3.34}$$

in population 1 and population 2, respectively.

Using formula (4) in Das Gupta (1991), we have

$$\beta\gamma\delta\epsilon\text{-standardized rate: in population 1} = Q(A)\,, \tag{3.35}$$

$$\text{in population 2} = Q(a)\,, \tag{3.36}$$

so that

$$\alpha\text{-effect} = Q(a) - Q(A)\,, \tag{3.37}$$

where

$$\begin{aligned} Q(A) = Q(A; b,c,d,e,B,C,D,E) &= \frac{F(A,b,c,d,e) + F(A,B,C,D,E)}{5} \\ &+ \frac{\begin{array}{c} F(A,b,c,d,E) + F(A,b,c,D,e) + F(A,b,C,d,e) + F(A,B,c,d,e) \\ + F(A,B,C,D,e) + F(A,B,C,d,E) + F(A,B,c,D,E) + F(A,b,C,D,E) \end{array}}{20} \\ &+ \frac{\begin{array}{c} F(A,b,c,D,E) + F(A,b,C,d,E) + F(A,b,C,D,e) \\ + F(A,B,C,d,e) + F(A,B,c,D,e) + F(A,B,c,d,E) \end{array}}{30}\,, \end{aligned} \tag{3.38}$$

and Q(a) is the same expression as that in (3.38) with A replaced by a.

Other standardized rates and factor effects follow directly from those in (3.35) through (3.38).

Example 3.4

This is a further extension of Example 3.3 in which the data on total women are used explicitly for Austria and Chile, 1981, as shown in table 3.7. In this case, equation (3.33) assumes the form

$$R = [\alpha\beta + \epsilon\,(1-\beta)]\gamma\delta\,, \tag{3.39}$$

where R = Crude birth rate per 1,000 population,
α = Marital general fertility rate per 1,000,
β = Proportion of married women among all women aged 15 to 49,
γ = Proportion of women aged 15 to 49 among all women,
δ = Proportion of women in the total population ,
ϵ = Nonmarital general fertility rate per 1,000.

The results in table 3.8 are virtually identical with those in table 3.6 except for the fact that the factor γ in Example 3.3 is broken down into two factors γ and δ in Example 3.4. We now see that as high as 52.1 percent of the difference between the crude birth rates of Chile and Austria is explained by the substantially higher proportion of women in the childbearing ages among all women in Chile relative to that in Austria. On the other hand, a smaller proportion of women in the population in Chile had a negative effect on the difference between the birth rates; that is, if all other four factors (except δ) were identical, the birth rate in Chile would be 0.668 point less than that in Austria.

Table 3.7. Crude Birth Rate as a Function of Five Factors: Austria and Chile, 1981

Measures	Austria, 1981 (population 1)	Chile, 1981 (population 2)
Crude birth rate $= \frac{\text{Births} \times 1000}{\text{Total population}}$ $(=R)$	12.512 $(=R_1)$	32.845 $(=R_2)$
Marital general fertility rate per 1,000 $(=\alpha)$	71.83691 (=A)	115.73732 (=a)
Proportion married among women aged 15 to 49 $(=\beta)$	0.58048 (=B)	0.52500 (=b)
Proportion of women aged 15 to 49 among all women $(=\gamma)$	0.45919 (=C)	0.75756 (=c)
Proportion of women in the population $(=\delta)$	0.52638 (=D)	0.51065 (=d)
Nonmarital general fertility rate per 1,000 $(=\epsilon)$	23.99823 (=E)	50.82674 (=e)

Source: See the footnote of table 3.5.

Table 3.8. Standardization and Decomposition of Crude Birth Rates in Table 3.7

Measures	Standardization		Decomposition	
	Chile, 1981 (population 2)	Austria, 1981 (population 1)	Difference (effects)	Percent distribution of effects
$\beta\gamma\delta\epsilon$-standardized birth rates	25.559	17.943	7.616 (α-effect)	37.4
$\alpha\gamma\delta\epsilon$-standardized birth rates	21.545	22.542	−0.997 (β-effect)	−4.9
$\alpha\beta\delta\epsilon$-standardized birth rates	26.872	16.288	10.584 (γ-effect)	52.1
$\alpha\beta\gamma\epsilon$-standardized birth rates	21.700	22.368	−0.668 (δ-effect)	−3.3
$\alpha\beta\gamma\delta$-standardized birth rates	23.696	19.898	3.798 (ϵ-effect)	18.7
Crude birth rates (R)	32.845	12.512	20.333 (Total effect)	100.0

Program 3.3

We can obtain the results in table 3.8 by using Program 3.3. This program is identical with Program 2.2 except for lines 3, 13, and 30. The interpretations of the variables in Program 3.3 are the same as those for Program 2.2 except for the manner in which the values 1 and 2 are attached, as described in the text for Program 3.1. Lines 3 and 30 in Program 3.3 are different because the formats of the input and output data in tables 3.7 and 3.8 are different from the corresponding formats in tables 2.7 and 2.8. Again, H in line 13 of Program 3.3 gives the expression for R in (3.39), whereas the same line in Program 2.2 expresses R in (2.31).

3.6. THE CASE OF SIX FACTORS

When there are six factors so that

$$R = F(\alpha,\beta,\gamma,\delta,\epsilon,\eta)\,, \tag{3.40}$$

and in the two populations,

Program 3.3 (Five Factors)

```
      DIMENSION P(2,5),R(2,5,3),E(5),S(2,5)
      READ(5,1) ((P(I,J),J=1,5),I=1,2)
    1 FORMAT(5F10.5)
      DO 2 I=1,2
      DO 2 J=1,5
      DO 2 K=1,3
    2 R(I,J,K)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      DO 3 K=1,2
      DO 3 L=1,2
      DO 3 M=1,2
      H=(P(I,1)*P(J,2)+P(M,5)*(1.-P(J,2)))*P(K,3)*P(L,4)
      IF(I+J+K+L+M.EQ. 5) R1=H
      IF(I+J+K+L+M.EQ.10) R2=H
      DO 3 N=1,3
      N1=N+3
      N2=12-N1
      IF(J+K+L+M.EQ.N1.OR.J+K+L+M.EQ.N2) R(I,1,N)=R(I,1,N)+H
      IF(I+K+L+M.EQ.N1.OR.I+K+L+M.EQ.N2) R(J,2,N)=R(J,2,N)+H
      IF(I+J+L+M.EQ.N1.OR.I+J+L+M.EQ.N2) R(K,3,N)=R(K,3,N)+H
      IF(I+J+K+M.EQ.N1.OR.I+J+K+M.EQ.N2) R(L,4,N)=R(L,4,N)+H
    3 IF(I+J+K+L.EQ.N1.OR.I+J+K+L.EQ.N2) R(M,5,N)=R(M,5,N)+H
      DO 5 J=1,5
      DO 4 I=1,2
    4 S(I,J)=R(I,J,1)/5.+R(I,J,2)/20.+R(I,J,3)/30.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,5),R2,R1,T
    6 FORMAT(40X,3F15.3)
      STOP
      END
```

Program 3.4 (Six Factors)

```
      DIMENSION P(2,6),R(2,6,3),E(6),S(2,6)
      READ(5,1) ((P(I,J),J=1,6),I=1,2)
    1 FORMAT(F5.0,3F5.3,F5.0,F5.3)
      DO 2 I=1,2
      DO 2 J=1,6
      DO 2 K=1,3
    2 R(I,J,K)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      DO 3 K=1,2
      DO 3 L=1,2
      DO 3 M=1,2
      DO 3 N=1,2
      H=P(I,1)*P(J,2)*P(K,3)*P(L,4)+P(M,5)*P(N,6)*(1.-P(L,4))
      IF(I+J+K+L+M+N.EQ. 6) R1=H
      IF(I+J+K+L+M+N.EQ.12) R2=H
      DO 3 KK=1,3
      K1=KK+4
      K2=15-K1
      IF(J+K+L+M+N.EQ.K1.OR.J+K+L+M+N.EQ.K2) R(I,1,KK)=R(I,1,KK)+H
      IF(I+K+L+M+N.EQ.K1.OR.I+K+L+M+N.EQ.K2) R(J,2,KK)=R(J,2,KK)+H
      IF(I+J+L+M+N.EQ.K1.OR.I+J+L+M+N.EQ.K2) R(K,3,KK)=R(K,3,KK)+H
      IF(I+J+K+M+N.EQ.K1.OR.I+J+K+M+N.EQ.K2) R(L,4,KK)=R(L,4,KK)+H
      IF(I+J+K+L+N.EQ.K1.OR.I+J+K+L+N.EQ.K2) R(M,5,KK)=R(M,5,KK)+H
    3 IF(I+J+K+L+M.EQ.K1.OR.I+J+K+L+M.EQ.K2) R(N,6,KK)=R(N,6,KK)+H
      DO 5 J=1,6
      DO 4 I=1,2
    4 S(I,J)=R(I,J,1)/6.+R(I,J,2)/30.+R(I,J,3)/60.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,6),R2,R1,T
    6 FORMAT(40X,3F15.2)
      STOP
      END
```

$$R_1 = F(A,B,C,D,E,F)\,,\ R_2 = (a,b,c,d,e,f)\,, \tag{3.41}$$

then, formula (5) in Das Gupta (1991) gives

$$\beta\gamma\delta\epsilon\eta\text{-standardized rate: in population 1} = Q(A)\,, \tag{3.42}$$

$$\text{in population 2} = Q(a)\,, \tag{3.43}$$

so that

$$\alpha\text{-effect} = Q(a) - Q(A)\,, \tag{3.44}$$

where

$$Q(A) = Q(A; b,c,d,e,f,B,C,D,E,F) = \frac{F(A,b,c,d,e,f) + F(A,B,C,D,E,F)}{6}$$

$$+ \frac{\begin{gathered}F(A,b,c,d,e,F) + F(A,b,c,d,E,f) + F(A,b,c,D,e,f) + F(A,b,C,d,e,f) + F(A,B,c,d,e,f)\\ + F(A,B,C,D,E,f) + F(A,B,C,D,e,F) + F(A,B,C,d,E,F) + F(A,B,c,D,E,F) + F(A,b,C,D,E,F)\end{gathered}}{30}$$

$$+ \frac{\begin{gathered}F(A,b,c,d,E,F) + F(A,b,c,D,e,F) + F(A,b,c,D,E,f) + F(A,b,C,d,e,F) + F(A,b,C,d,E,f)\\ + F(A,b,C,D,e,f) + F(A,B,c,d,e,F) + F(A,B,c,d,E,f) + F(A,B,c,D,e,f) + F(A,B,C,d,e,f)\\ + F(A,B,C,D,e,f) + F(A,B,C,d,E,f) + F(A,B,C,d,e,F) + F(A,B,c,D,E,f) + F(A,B,c,D,e,F)\\ + F(A,B,c,d,E,F) + F(A,b,C,D,E,f) + F(A,b,C,D,e,F) + F(A,b,C,d,E,F) + F(A,b,c,D,E,F)\end{gathered}}{60}, \tag{3.45}$$

and Q(a) is the same expression as that in (3.45) with A replaced by a.

Other standardized rates and factor effects follow directly from those in (3.42) through (3.45).

Example 3.5

The data in table 3.9 are taken from Wojtkiewicz, McLanahan, and Garfinkel (1990) where the family headship rates per 1,000 for White mothers, 18 to 59 years, for 1950 and 1980 are expressed as follows:

$$\frac{\text{Mothers who are family heads x 1000}}{\text{Total women}}$$

$$= \frac{\text{Formerly married mothers who are family heads x 1000}}{\text{Formerly married mothers}}$$

$$\times \frac{\text{Formerly married mothers}}{\text{Ever-married mothers}} \times \frac{\text{Ever-married mothers}}{\text{Ever-married women}} \times \frac{\text{Ever-married women}}{\text{Total women}} \tag{3.46}$$

$$+ \frac{\text{Never-married mothers who are family heads x 1000}}{\text{Never-married mothers}}$$

$$\times \frac{\text{Never-married mothers}}{\text{Never-married women}} \times \frac{\text{Never-married women}}{\text{Total women}}\,,$$

which, in our notation, reduces to

$$R = F(\alpha,\beta,\gamma,\delta,\epsilon,\eta) = \alpha\beta\gamma\delta + \epsilon\eta(1-\delta)\ . \tag{3.47}$$

The family headship rates increased from 22.70 to 55.02 during 1950 to 1980, producing a total increase of 32.32 points. The standardized rates and the effects of the six factors are shown in table 3.10. For example, if only the proportions of formerly married mothers among ever-married mothers (β) varied as they did in 1950 and 1980, and all the remaining five factors were identical in the two years, then the standardized headship rates would be 26.36 and 49.14 in 1950 and 1980, respectively, producing a difference of 22.78 as the β-effect. In other words, 70.5 percent of the increase in the headship rate between 1950 and 1980 can be attributed to the increase in the proportion of the formerly married mothers among ever-married mothers (β) in the three decades. Similar observations can be made about other numbers in table 3.10. Wojtkiewicz, McLanahan, and Garfinkel decomposed the difference between the

numbers of female family heads rather than between the female family headship **rates** and also considered interaction between the factors. Their results are, therefore, not directly comparable with those presented here. This example will be extended to four populations for the years 1950, 1960, 1970, and 1980 in Example 6.4 (tables 6.7 and 6.8).

Table 3.9. Family Headship Rate for Mothers, 18 to 59 Years, as a Function of Six Factors: United States, White, 1950 and 1980

Measures	1950 (population 1)	1980 (population 2)
Family headship rate of mothers per 1,000 total women (=R)	22.70 ($=R_1$)	55.02 ($=R_2$)
Formerly married mothers who are family heads per 1,000 formerly married mothers ($=\alpha$)	688 (=A)	878 (=a)
Proportion of formerly married mothers among ever-married mothers ($=\beta$)	0.067 (=B)	0.129 (=b)
Proportion of ever-married mothers among ever-married women ($=\gamma$)	0.571 (=C)	0.562 (=c)
Proportion of ever-married women among total women ($=\delta$)	0.851 (=D)	0.808 (=d)
Never-married mothers who are family heads per 1,000 never-married mothers ($=\epsilon$)	509 (=E)	623 (=e)
Proportion of never-married mothers among never-married women ($=\eta$)	0.004 (=F)	0.030 (=f)

Source: Wojtkiewicz, McLanahan, and Garfinkel (1990), table 2.

Table 3.10. Standardization and Decomposition of Family Headship Rates in Table 3.9

Family headship rates	Standardization		Decomposition	
	1980 (population 2)	1950 (population 1)	Difference (effects)	Percent distribution of effects
$\beta\gamma\delta\epsilon\eta$-standardized rates	42.03	33.31	8.72 (α)	27.0
$\alpha\gamma\delta\epsilon\eta$-standardized rates	49.14	26.36	22.78 (β)	70.5
$\alpha\beta\delta\epsilon\eta$-standardized rates	37.84	38.42	−0.58 (γ)	−1.8
$\alpha\beta\gamma\epsilon\eta$-standardized rates	37.43	38.89	−1.46 (δ)	−4.5
$\alpha\beta\gamma\delta\eta$-standardized rates	38.21	37.87	0.34 (ϵ)	1.0
$\alpha\beta\gamma\delta\epsilon$-standardized rates	39.25	36.73	2.52 (η)	7.8
Crude headship rates (R)	55.02	22.70	32.32 (Total effect)	100.0

Program 3.4

The results in table 3.10 can be obtained by using Program 3.4. The format of this program and the interpretations of the variables are the same as those for Program 3.3 except for the changes that are needed to go from five to six factors. Also, the input and output formats in lines 3 and 32 in Program 3.4 are made consistent with the numbers in tables 3.9 and 3.10. The equation in (3.47) is expressed in line 14. The subscripts I, J, and K in R(I,J,K) in line 7 of Program 3.4 refer to the two populations, the six factors, and the three expressions on the right-hand side of (3.45). In (3.45), we leave the argument A untouched but attach a value of 1 to the other capital letters and a value of 2 to the small letters, and add these two values of the arguments for each F. We find that the first expression in (3.45) includes F's with a total of 5 and 10 points for the arguments, the second expression includes F's with a total of 6 and 9 points, and the third expression includes F's with a total of 7 and 8 points. K1 and K2 in lines 18 and 19 of Program 3.4 for KK = 1,3 give the above three pairs of points, namely, (5,10), (6,9), and (7,8).

Example 3.6

Exactly the same six-factor model in equation (3.47) can also be used to extend the four-factor model by Nathanson and Kim (1989), given in equation (2.30), to all live births (nonmarital and marital) by defining R as the percentage having live births, and adding to equation (2.30) the term $\epsilon\eta(1\text{-}\delta)$ where

$$\begin{aligned} \epsilon &= \text{Percentage having marital live births among marital pregnancies,} \\ \eta &= \text{Proportion of marital pregnancies among total married women,} \\ 1\text{-}\delta &= \text{Proportion of married women among all women.} \end{aligned} \tag{3.48}$$

The data for this example are provided in table 3.11, and the corresponding standardized rates and the factor effects, in table 3.12. The percentages R for 1971 and 1979 are, respectively, 3.592 and 4.846, giving a total difference of 1.254. Although this difference is much smaller than the difference of 2.989 in table 2.6 based on nonmarital live births, the absolute values of α, β, and γ effects are identical in tables 2.6 and 3.12. These results follow from a comparison of equations (2.24) and (3.47) since, for given values of ϵ, η, and δ, the additional term in (3.47) does not have any effect on the difference. It is interesting to note that the increase in the proportion of single women among all women (δ) during 1971-1979 tended to increase the percentage having nonmarital live births (0.105 in table 2.6) and decrease the percentage having live births (-1.237 in table 3.12) during the same period. A significant decline in the proportion of marital pregnancies among married women (η) during the 8-year period also had a negative effect on the difference between the percentages having live births in table 3.12.

The results in table 3.12 can be obtained by using the data in table 3.11 and Program 3.4. The only changes needed in the program are the input and output format statements in lines 3 and 32 as follows:

Line 3: 1 FORMAT (F6.1, 3F6.3, F6.1, F6.3)

Line 32: 6 FORMAT (40X, 3F15.3)

Table 3.11. Percentage Having Live Births as a Function of Six Factors, for White Women Aged 15 to 19: United States, 1971 and 1979

Measures	1971 (population 1)	1979 (population 2)
Percentage having live births (=R)	3.592 (=R_1)	4.846 (=R_2)
Percentage having nonmarital live births among nonmarital pregnancies (=α)	25.3 (=A)	32.7 (=a)
Proportion of nonmarital pregnancies among sexually active single women (=β)	.214 (=B)	.290 (=b)
Proportion of sexually active single women among total single women (=γ)	.279 (=C)	.473 (=c)
Proportion of single women among all women (=δ)	.949 (=D)	.986 (=d)
Percentage having marital live births among marital pregnancies (=ϵ)	92.0 (=E)	91.4 (=e)
Proportion of marital pregnancies among total married women (=η)	.460 (=F)	.331 (=f)

Source: Nathanson and Kim (1989), tables 1 and 4; table 2.5 in this report.

Table 3.12. Standardization and Decomposition of Percentages Having Live Births in Table 3.11

Percentages having live births	Standardization		Decomposition	
	1979 (population 2)	1971 (population 1)	Difference (effects)	Percent distribution of effects
$\beta\gamma\delta\epsilon\eta$-standardized percentages	4.260	3.572	0.688 (α)	54.9
$\alpha\gamma\delta\epsilon\eta$-standardized percentages	4.317	3.504	0.813 (β)	64.8
$\alpha\beta\delta\epsilon\eta$-standardized percentages	4.588	3.205	1.383 (γ)	110.3
$\alpha\beta\gamma\epsilon\eta$-standardized percentages	3.299	4.536	−1.237 (δ)	−98.7
$\alpha\beta\gamma\delta\eta$-standardized percentages	3.960	3.968	−0.008 (ϵ)	−0.6
$\alpha\beta\gamma\delta\epsilon$-standardized percentages	3.735	4.120	−0.385 (η)	−30.7
Percentages having live births (R)	4.846	3.592	1.254 (Total effect)	100.0

3.7 THE CASE OF P FACTORS

Let us write the rate as a function of P factors as

$$R = F(\alpha_1, \alpha_2, \ldots, \alpha_p) \ , \tag{3.49}$$

and, in the two populations, this rate assumes the values

$$R_1 = F(A_1, A_2, \ldots, A_p) \ , \quad R_2 = F(a_1, a_2, \ldots, a_p) \ . \tag{3.50}$$

It follows from formula A5 in Das Gupta (1991) that

$$\alpha_2\alpha_3 \ldots \alpha_p\text{-standardized rate: in population 1} = Q(A_1) \ , \tag{3.51}$$

$$\text{in population 2} = Q(a_1) \ , \tag{3.52}$$

so that

$$\alpha_1\text{-effect} = Q(a_1) - Q(A_1) \ , \tag{3.53}$$

where

$$Q(A_1) = Q(A_1; a_2, a_3, \ldots, a_p, A_2, A_3, \ldots, a_p) = \frac{F(A_1, a_2, a_3, \ldots, a_p) + F(A_1, A_2, A_3, \ldots, A_p)}{P}$$

$$+ \frac{\text{sum of all F's with } A_1, (P-2) \text{ small letters and 1 capital letter or } A_1, (P-2) \text{ capital letters and 1 small letter}}{P\binom{P-1}{1}}$$

$$+ \frac{\text{sum of all F's with } A_1, (P-3) \text{ small letters and 2 capital letters or } A_1, (P-3) \text{ capital letters and 2 small letters}}{P\binom{P-1}{2}} \tag{3.54}$$

$$+ \ldots\ldots\ldots\ldots$$

$$= \sum_{r=1}^{S} \frac{\text{sum of all F's with } A_1, (P-r) \text{ small letters and } (r-1) \text{ capital letters or } A_1, (P-r) \text{ capital letters and } (r-1) \text{ small letters}}{P\binom{P-1}{r-1}}$$

where

$$S = P/2 \text{, when P is even ,}$$
$$= (P+1)/2 \text{, when P is odd.}$$

3.8 THE GENERAL PROGRAM

From Programs 3.1 through 3.4 corresponding to three to six factors, a FORTRAN program can be developed for any number of factors higher than six. However, it is not necessary to use different programs for data involving different numbers of factors. A program written for, say, 10-factor data can be used for any number of factors not exceeding 10 by changing the expression for the rate R and the input and output statements and formats in the program, as suggested in section 2.8. Again, no changes are needed in the data files previously created to be used with the specific programs.

As a matter of fact, the general program for up to 10 factors (Program 2.3) given in section 2.8 can also be used for any number of factors up to 10 for the standardization and decomposition problems in chapter 3, i.e., when the rate is a function of the factors. As before, the only changes needed in Program 2.3 are in the input and output statements in lines 2 and 42, the input and output formats in lines 3 and 43, and in the expression for the rate in lines 18 and 19. We show below the specific changes in Program 2.3 that will be needed to generate the results in tables 3.2, 3.4, 3.6, 3.8, 3.10, and 3.12 corresponding to Examples 3.1 through 3.6 in this chapter:

Example 3.1 (two factors)

Lines 2,42: Replace 10 in each line by 2
Lines 3,43: Replace (10F8.4) by (2F8.1) and 15.3 by 15.2
Lines 18,19: Replace the two lines by H = P(I,1)-P(J,2)

Example 3.2 (three factors)

Lines 2,42: Replace 10 in each line by 3
Lines 3,43: Replace (10F8.4) by (3F10.6) and 15.3 by 15.5
Lines 18,19: Replace the two lines by line 11 in Program 3.1

Example 3.3 (four factors)

Lines 2,42: Replace 10 in each line by 4
Lines 3,43: Replace (10F8.4) by (4F10.5) and no change in line 43
Lines 18,19: Replace the two lines by line 12 in Program 3.2

Example 3.4 (five factors)

Lines 2,42: Replace 10 in each line by 5
Lines 3,43: Replace (10F8.4) by (5F10.5) and no change in line 43
Lines 18,19: Replace the two lines by line 13 in Program 3.3

Example 3.5 (six factors)

Lines 2,42: Replace 10 in each line by 6
Lines 3,43: Replace (10F8.4) by (F5.0, 3F5.3, F5.0, F5.3) and 15.3 by 15.2
Lines 18,19: Replace the two lines by line 14 in Program 3.4

Example 3.6 (six factors)

Lines 2,42: Replace 10 in each line by 6
Lines 3,43: Replace (10F8.4) by (F6.1, 3F6.3, F6.1, F6.3) and no change in line 43
Lines 18,19: Replace the two lines by line 14 in Program 3.4

3.9 EXAMPLE 3.7 (TEN FACTORS)

Pullum, Tedrow, and Herting (1989) expressed the mean parity M of a cohort of women by

$$M = P_0 + P_0P_1 + P_0P_1P_2 + \ldots\ldots + P_0P_1P_2 \ldots P_9 , \tag{3.55}$$

where P_i is the parity progression ratio for transition from parity i to parity i+1 (we assume here that the highest possible parity is 10).

In terms of our notation, equation (3.55) can be written as

$$\begin{aligned} R &= F(\alpha,\beta,\gamma,\delta,\epsilon,\eta,\theta,\lambda,\mu,\upsilon) \\ &= \alpha + \alpha\beta + \alpha\beta\gamma + \ldots\ldots + \alpha\beta\gamma\delta\epsilon\eta\theta\lambda\mu\upsilon. \end{aligned} \tag{3.56}$$

The values of the two rates and the 10 factors for White women for 1908 and 1933 cohorts are shown in table 3.13.

In the 25-year period from 1908 to 1933, the mean parity of a cohort increased by .854, from 2.247 in 1908 to 3.101 in 1933. As shown in table 3.14, the mean parities in 1908 and 1933 would have been 2.454 and 2.854 if only the parity progression ratio from parity 0 to parity 1 (α) changed as it did between 1908 and 1933, and all other parity progression ratios were equal in the two years. Therefore, .400 (46.8 percent) of the increase in the mean parity in the 25-year period was contributed by the increase in the parity progression ratio from parity 0 to parity 1. It is interesting to note that the first four parity progression ratios made positive contributions to the total increase in the mean parity, and the remaining ratios contributed negatively. The decomposition in table 3 of Pullum, Tedrow, and Herting by and large agrees with that presented in the last two columns of table 3.14.

Table 3.13. Mean Parity of a Cohort as a Function of Ten Factors (Parity Progression Ratios), for White Women: United States, 1908 and 1933 Cohorts

Mean parity and parity progression ratios (PPR's)	1908 cohort (population 1)	1933 cohort (population 2)
Mean Parity (=R)	2.247 (=R_1)	3.101 (=R_2)
PPR for transition from parity 0 to 1 (=α)	0.7921 (=A)	0.9215 (=a)
PPR for transition from parity 1 to 2 (=β)	0.7247 (=B)	0.8950 (=b)
PPR for transition from parity 2 to 3 (=γ)	0.5937 (=C)	0.7198 (=c)
PPR for transition from parity 3 to 4 (=δ)	0.5924 (=D)	0.6016 (=d)
PPR for transition from parity 4 to 5 (=ϵ)	0.6057 (=E)	0.5354 (=e)
PPR for transition from parity 5 to 6 (=η)	0.6353 (=F)	0.5267 (=f)
PPR for transition from parity 6 to 7 (=θ)	0.6396 (=G)	0.5214 (=g)
PPR for transition from parity 7 to 8 (=λ)	0.7948 (=H)	0.6381 (=h)
PPR for transition from parity 8 to 9 (=μ)	0.7468 (=I)	0.5522 (=i)
PPR for transition from parity 9 to 10 (=υ)	0.6746 (=J)	0.4162 (=j)

Source: Pullum, Tedrow, and Herting (1989), table 1 (data extended for higher parities).

Table 3.14. Standardization and Decomposition of Mean Parities in Table 3.13

Mean parities standardized for—	Standardization		Decomposition	
	1933 cohort (population 2)	1908 cohort (population 1)	Difference (effects)	Percent distribution of effects
All PPR's except α	2.854	2.454	.400 (α)	46.8
All PPR's except β	2.842	2.464	.378 (β)	44.3
All PPR's except γ	2.761	2.549	.212 (γ)	24.8
All PPR's except δ	2.664	2.654	.010 (δ)	1.2
All PPR's except ϵ	2.637	2.683	−.046 (ϵ)	−5.4
All PPR's except η	2.639	2.680	−.041 (η)	−4.8
All PPR's except θ	2.646	2.672	−.026 (θ)	−3.0
All PPR's except λ	2.651	2.667	−.016 (λ)	−1.9
All PPR's except μ	2.653	2.664	−.011 (μ)	−1.3
All PPR's except υ	2.656	2.662	−.006 (υ)	−0.7
Mean parities (R)	3.101	2.247	0.854 (Total effect)	100.0

The data in table 3.13 and the general program (Program 2.3) in chapter 2 can be used to obtain the results in table 3.14 if the following changes are made in Program 2.3:

Lines 2,42: No changes
Lines 3,43: No changes
Lines 18,19: Replace the two lines by equation (3.56), i.e., by

```
      H=P(I,1)*(1.+P(J,2)*(1.+P(K,3)*(1.+P(L,4)*(1.+P(M,5)*(1.+P(N,6)*
     1 (1.+P(II,7)*(1.+P(JJ,8)*(1.+P(KK,9)*(1.+P(LL,10))))))))))
```

Chapter 4. Rate as a Function of Vector-Factors

4.1 INTRODUCTION

In many situations, a factor may be represented by several numbers. For example, six age-specific fertility rates together may be considered one factor. Such factors may be called vector-factors (as opposed to scalar-factors). Cho and Retherford (1973), for example, expressed the crude birth rate as a function of three vector-factors, namely, the age-specific marital fertility rates (assuming that no births occur to unmarried women), the proportions of married women among total women in the age groups, and total women in the age groups as proportions of the total population (Example 4.3). Again, Smith and Cutright (1988) expressed the illegitimacy ratio as a function of four vector-factors, namely, the proportional age distribution of women in the childbearing period, the proportions of unmarried women to total women in the childbearing age groups, the age-specific nonmarital fertility rates, and the age-specific marital fertility rates (Example 4.4). The expressions for standardization and decomposition for both scalar- and vector-factors are identical except that we should use different symbols to distinguish between them, as shown in the following sections.

4.2 THE CASE OF TWO VECTOR-FACTORS

We express the two vector-factors as

$$\overline{\alpha} = (\alpha_1,\alpha_2, ...,\alpha_{n_1}) \ , \ \overline{\beta} = (\beta_1,\beta_2, ...,\beta_{n_2}) \ , \tag{4.1}$$

n_1 and n_2 being the numbers of elements in the two vectors. In many situations, as in the two examples in section 4.1, the numbers n_1 and n_2 are equal.

We write the rate R as

$$R = F(\alpha_1,\alpha_2, ...,\alpha_{n_1}, \beta_1,\beta_2, ...,\beta_{n_2}) = F(\overline{\alpha},\overline{\beta}) \ , \tag{4.2}$$

and equations (3.2) for population 1 and population 2 change to

$$R_1 = F(\overline{A},\overline{B}) \ , \ R_2 = F(\overline{a},\overline{b}) \ , \tag{4.3}$$

where

$$\overline{A} = (A_1,A_2,... ,A_{n_1}), \quad \overline{B} = (B_1,B_2,... ,B_{n_2}) \ , \tag{4.4}$$

$$\overline{a} = (a_1,a_2,... ,a_{n_1}), \quad \overline{b} = (b_1,b_2,... ,b_{n_2}) \ .$$

In spite of the fact that R in (4.2) depends on (n_1+n_2) scalar numbers, we do not treat this as a (n_1+n_2)-factor case because we do not allow all these factors to take on values from population 1 and population 2 independently of each other. We impose here the condition that the n_1 scalars $(\alpha_1, \alpha_2, ..., \alpha_{n_1})$ must take on either the values $(A_1, A_2, ..., A_{n_1})$ or the values $(a_1, a_2, ..., a_{n_1})$. Had we treated this as a (n_1+n_2)-factor case, it would have been possible to have a set of values such as $(A_1, a_2, a_3, ..., A_{n_1})$ for $\overline{\alpha}$. Similar restrictions apply to the elements of $\overline{\beta}$.

Changing the notation from scalar to vector in (3.3) through (3.8), we obtain

$$\overline{\beta}\text{-standardized rate: in population 1} = \frac{F(\overline{A},\overline{b}) + F(\overline{A},\overline{B})}{2} \ , \tag{4.5}$$

$$\text{in population 2} = \frac{F(\overline{a},\overline{b}) + F(\overline{a},\overline{B})}{2} \ , \tag{4.6}$$

$$\bar{\alpha}\text{-standardized rate: in population 1} = \frac{F(\bar{a},\bar{B})+F(\bar{A},\bar{B})}{2}, \quad (4.7)$$

$$\text{in population 2} = \frac{F(\bar{a},\bar{b}) + F(\bar{A},\bar{b})}{2}, \quad (4.8)$$

$$\bar{\alpha}\text{-effect} = \frac{[F(\bar{a},\bar{b}) - F(\bar{A},\bar{b})] + [F(\bar{a},\bar{B}) - F(\bar{A},\bar{B})]}{2}, \quad (4.9)$$

$$\bar{\beta}\text{-effect} = \frac{[F(\bar{a},\bar{b}) - F(\bar{a},\bar{B})] + [F(\bar{A},\bar{b}) - F(\bar{A},\bar{B})]}{2}. \quad (4.10)$$

Example 4.1

Keyfitz (1968, p. 189) considered the decomposition of the difference between two intrinsic growth rates into the effects of changes in the age-specific fertility and mortality rates. Table 4.1 gives the stationary populations ${}_5L_x$ from the abridged life tables for females and the fertility rates ${}_5m_x$ for females (based on the female births only) by 5-year age groups for 1960 and 1965. These two series of data for a year serve as the vector-factors $\bar{\alpha}$ and $\bar{\beta}$ for that year.

For a given set of $\bar{\alpha},\bar{\beta}$, the female intrinsic growth rate $R = F(\bar{\alpha},\bar{\beta})$ can be obtained iteratively by the Newton-Raphson Method (Scarborough, 1962, p. 199) as follows:

We compute

$$\mu_o = \sum_{i=1}^{9} \alpha_i\beta_i/100000, \quad (4.11)$$

$$\mu_1 = \sum_{i=1}^{9} (5i+7.5)\alpha_i\beta_i/100000. \quad (4.12)$$

The first approximation r_1 is given by

$$r_1 = (\log_e\mu_0) \cdot \mu_0/\mu_1. \quad (4.13)$$

With the above value of r_1, we compute

$$N(r_1) = \sum_{i=1}^{9} \exp[-r_1(5i+7.5)]\alpha_i\beta_i/100000, \quad (4.14)$$

$$D(r_1) = \sum_{i=1}^{9} (5i+7.5)\exp[-r_1(5i+7.5)]\alpha_i\beta_i/100000. \quad (4.15)$$

The second approximation r_2 is

$$r_2 = r_1 - \frac{N(r_1)-1}{D(r_1)}. \quad (4.16)$$

This process is continued until

$$|r_n - r_{n-1}| \leq .0000001, \quad (4.17)$$

and at this point, r_n is taken as the intrinsic growth rate R.

The intrinsic growth rates $R_1 = F(\bar{A},\bar{B})$ and $R_2 = F(\bar{a},\bar{b})$ for 1965 and 1960 are, respectively, 12.14 and 20.77 per 1,000, their difference being 8.63. Table 4.2 gives the four standardized rates and the two factor effects. For example, the mortality-standardized intrinsic growth rates in 1960 and 1965 are 20.81 and 12.10; i.e., if only the fertility varied as it did in 1960 and 1965, and the mortality were the same in the two years, then the intrinsic growth rate would decline from 20.81 to 12.10 in the 5-year period. This decline of 8.71 is even higher than the actual decline of 8.63. Therefore, the change (decline) in mortality during 1960-1965 had a slight dampening effect on the total decline in the intrinsic growth rate. Keyfitz used the Australian data and, therefore, his decomposition is not directly comparable with our decomposition on the U.S. data.

Table 4.1. Female Intrinsic Growth Rate per Person as a Function of Two Vector-Factors: United States, 1960 and 1965

Age groups x to x+5	i	${}_5L_x$ (α_i)		${}_5m_x$ (β_i)	
		1965 (population 1) A_i	1960 (population 2) a_i	1965 (population 1) B_i	1960 (population 2) b_i
10 to 15	1	486446	485434	.00041	.00040
15 to 20	2	485454	484410	.03416	.04335
20 to 25	3	483929	492905	.09584	.12581
25 to 30	4	482046	481001	.07915	.09641
30 to 35	5	479522	478485	.04651	.05504
35 to 40	6	475844	474911	.02283	.02760
40 to 45	7	470419	469528	.00631	.00758
45 to 50	8	462351	461368	.00038	.00045
50 to 55	9	450468	449349	.00000	.00001
Intrinsic growth rate (R)		R_1 (1965) = .01214,		R_2 (1960) = .02077	

Source: National Center for Health Statistics (1962, tables 2-13, 5-3; 1963, table 2-1; 1967a, tables 1-48, 4-2; 1967b, table 5-1).

Table 4.2. Standardization and Decomposition of Female Intrinsic Growth Rates per Person in Table 4.1

(For convenience, results obtained from data in table 4.1 are multiplied by 1,000 before presenting them in table 4.2)

Female intrinsic growth rate	Standardization		Decomposition	
	1960 (population 2)	1965 (population 1)	Difference (effects)	Percent distribution of effects
$\bar{\beta}$ (fertility)-standardized growth rates	16.41	16.49	−.08 ($\bar{\alpha}$)	−0.9
$\bar{\alpha}$ (mortality)-standardized growth rates	20.81	12.10	8.71 ($\bar{\beta}$)	100.9
Overall intrinsic rates (R)	20.77	12.14	8.63 (Total effect)	100.0

Program 4.1

The results in table 4.2 can be obtained by using Program 4.1 in which V(I,J,K)'s in line 2 are the data from table 4.1 corresponding to I = 1,2 (1965 and 1960); J = 1,2 (mortality and fertility); and K = 1,9 (nine age groups). In other words, the data file consists of four lines with the four vectors $\bar{A}$,$\bar{B}$, $\bar{a}$, and $\bar{b}$ in table 4.1, each line having nine elements. Equations (4.11) through (4.17) are given in lines 12 through 14 and 18 through 21 of the program. As in Program 3.1, S(I,J)'s in line 28 are the four standardized rates and E(J)'s in line 29 are the two vector-factor effects in table 4.2.

Program 4.1 (Two Factors)

```
      DIMENSION V(2,2,9),R(2,2),E(2),S(2,2)
      READ(5,1) (((V(I,J,K),K=1,9),J=1,2),I=1,2)
    1 FORMAT(9F8.0/9F8.5)
      DO 2 I=1,2
      DO 2 J=1,2
    2 R(I,J)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      H1=0.0
      H2=0.0
      DO 7 K1=1,9
      H1=H1+V(I,1,K1)*V(J,2,K1)/100000.
    7 H2=H2+(5.*K1+7.5)*V(I,1,K1)*V(J,2,K1)/100000.
      H=ALOG(H1)*H1/H2
    8 SI=0.0
      SD=0.0
      DO 9 K1=1,9
      SI=SI+EXP(-H*(5.*K1+7.5))*V(I,1,K1)*V(J,2,K1)/100000.
    9 SD=SD-(5.*K1+7.5)*EXP(-H*(5.*K1+7.5))*V(I,1,K1)*V(J,2,K1)/100000.
      H=H-(SI-1.)/SD
      IF(ABS(SI-1.).GT..0000001) GO TO 8
      IF(I+J.EQ.2) R1=H
      IF(I+J.EQ.4) R2=H
      R(I,1)=R(I,1)+H
    3 R(J,2)=R(J,2)+H
      DO 5 J=1,2
      DO 4 I=1,2
    4 S(I,J)=R(I,J)/2.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,2),R2,R1,T
    6 FORMAT(40X,3F15.5)
      STOP
      END
****
```

Program 4.2 (Two Factors)

```
      DIMENSION V(2,2,480),R(2,2),E(2),S(2,2)
      READ(5,1) (((V(I,J,K),K=1,480),J=1,2),I=1,2)
    1 FORMAT(8F10.7)
      DO 2 I=1,2
      DO 2 J=1,2
    2 R(I,J)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      H1=0.0
      H2=0.0
      DO 7 K1=1,480
      H1=H1+V(I,1,K1)*V(J,2,K1)
    7 H2=H2+(1.-V(I,1,K1))*V(J,2,K1)
      H=0.0
      DO 8 K1=1,480
    8 H=H+50.*ABS(V(I,1,K1)*V(J,2,K1)/H1-(1.-V(I,1,K1))*V(J,2,K1)/H2)
      IF(I+J.EQ.2) R1=H
      IF(I+J.EQ.4) R2=H
      R(I,1)=R(I,1)+H
    3 R(J,2)=R(J,2)+H
      DO 5 J=1,2
      DO 4 I=1,2
    4 S(I,J)=R(I,J)/2.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,2),R2,R1,T
    6 FORMAT(40X,3F15.3)
      STOP
      END
****
```

Example 4.2

Bianchi and Rytina (1986) decomposed the difference between the indices of male-female occupational dissimilarity for 1970 and 1980 in order to eliminate from this difference the effect of the change in the occupational structure during the decade. The index of dissimilarity may be written as

$$\text{Index} = \frac{1}{2} \sum_i \left| (M_i/M) \times 100 - (F_i/F) \times 100 \right| , \tag{4.18}$$

where M_i and F_i are the numbers of males and females in occupation i, and M and F are the total males and the total females.

Equation (4.18) can also be written in terms of our notation as

$$R = F(\overline{\alpha},\overline{\beta}) = 50 \sum_i \left| \frac{\alpha_i\beta_i}{\sum_j \alpha_j\beta_j} - \frac{(1-\alpha_i)\beta_i}{\sum_j (1-\alpha_j)\beta_j} \right| , \tag{4.19}$$

where

$$\alpha_i = M_i/T_i \ , \ \beta_i = T_i/T \ , \ T_i = M_i + F_i \ , \ T = M + F \ . \tag{4.20}$$

Table 4.3 gives a sample of the 480 elements of the vectors $\overline{\alpha}$ and $\overline{\beta}$ for 1970 and 1980. The indices of male-female occupational dissimilarity based on these data are 59.285 and 67.683 for 1980 and 1970, respectively. The standardization of these indices and the decomposition of their difference of 8.398 are shown in table 4.4. It shows, for example, that if the occupational structures in 1970 and 1980 were identical, then the indices of dissimilarity in 1970 and 1980 would be 67.017 and 60.271, producing a difference of 6.746. This difference is, obviously, the effect of the change in the occupational sex segregation during the decade. In other words, 80.3 percent of the decline in the index of male-female occupational dissimilarity during 1970-1980 is contributed by the decline in the occupational sex segregation during the decade. The decomposition by Bianchi and Rytina is in agreement with these results except that it included an interaction term. (With a slightly different set of data producing a total effect of 8.5, their results were 6.4 and 1.4 for the $\overline{\alpha}$ and $\overline{\beta}$ effects and 0.7 for the interaction effect.) The approximate method by Das Gupta (1987) applied to the same set of data produced a slightly different result. Again, arguments in favor of using only the main effects that absorb the interactions are given in chapter 1.

Program 4.2

The results in table 4.4 can be obtained by using Program 4.2 in which V(I,J,K)'s in line 2 are the data from table 4.3 corresponding to I = 1,2 (1980 and 1970); J = 1,2 (M_i/T_i's and T_i/T's); and K = 1,480 (480 occupations). The data file consists of 240 lines, each of the four vectors $\overline{A}$,$\overline{B}$, $\overline{a}$, and $\overline{b}$ occupying 60 lines in the same order with eight numbers in each line. Equation (4.19) is expressed in the program in line 16. Program 4.2 is basically the same as Program 4.1 except for the fact that in Program 4.1, there are nine elements in a vector-factor, and it uses lines 9 through 21 to compute the rate R (i.e., H in the program) whereas in Program 4.2, there are 480 elements in a vector-factor, and it uses lines 9 through 16 to compute H. Consequently, Program 4.2 is five lines shorter than Program 4.1.

Table 4.3. Index of Male-Female Occupational Dissimilarity as a Function of Two Vector-Factors: United States, 1970 and 1980 (Partial Data)

i	Occupation	$(M_i/T_i) = \alpha_i$ 1980 (population 1) A_i	1970 (population 2) a_i	$(T_i/T) = \beta_i$ 1980 (population 1) B_i	1970 (population 2) b_i
1	Legislators, etc., public administration..........	.7443052	1.0000000	.0004481	.0001251
2	Administrators, public administration...........	.6637712	.7826656	.0028553	.0031344
3	Administrators, protective services.............	.9059344	1.0000000	.0002531	.0003277
4	Financial Managers..........................	.6861586	.8058082	.0039536	.0027667
--	--....................................	--	--	--	--
--	--....................................	--	--	--	--
479	Wholesale and retail trade....................	.8197265	.8040852	.0032985	.0032290
480	All other industries...........................	.8203055	.8557602	.0024285	.0022611
Index of dissimilarity (R)...........................		R_1 (1980) = 59.285,		R_2 (1970) = 67.683	

Source: U.S. Bureau of the Census (1984b), pp. 7-15. Total males (M) and total females (F=T-M) in 1980 are 59,592,657 and 44,069,629, and those in 1970 are 49,405,944 and 30,285,210, respectively (excluding the experienced unemployed not classified by occupation, and the seven occupations with no persons in 1970).

Table 4.4. Standardization and Decomposition of Indices of Male-Female Occupational Dissimilarity in Table 4.3

Index of male-female occupational dissimilarity	Standardization 1970 (population 2)	1980 (population 1)	Decomposition Difference (effects)	Percent distribution of effects
$\bar{\beta}$ (occupational structure)-standarized index of dissimilarity	67.017	60.271	6.746 ($\bar{\alpha}$)	80.3
$\bar{\alpha}$ (occupational sex segregation)-standardized index of dissimilarity	64.470	62.818	1.652 ($\bar{\beta}$)	19.7
Overall index of dissimilarity (R)	67.683	59.285	8.398 (Total effect)	100.0

4.3 THE CASE OF THREE VECTOR-FACTORS

We express the three vector-factors as

$$\bar{\alpha} = (\alpha_1,\alpha_2,\ldots,\alpha_{n_1}) \ , \ \bar{\beta} = (\beta_1,\beta_2,\ldots,\beta_{n_2}) \ , \ \bar{\gamma} = (\gamma_1,\gamma_2,\ldots,\gamma_{n_3}) \ , \tag{4.21}$$

and write the rate R as

$$\begin{aligned} R &= F(\alpha_1,\alpha_2,\ldots,\alpha_{n_1}, \beta_1,\beta_2,\ldots,\beta_{n_2},\gamma_1,\gamma_2,\ldots,\gamma_{n_3}) \ , \\ &= F(\bar{\alpha},\bar{\beta},\bar{\gamma}) \ . \end{aligned} \tag{4.22}$$

Equations (3.11) for population 1 and population 2 in this case change to

$$R_1 = F(\bar{A},\bar{B},\bar{C}) \ , \ R_2 = F(\bar{a},\bar{b},\bar{c}) \ . \tag{4.23}$$

Equations (3.12) through (3.23) remain unchanged except that the scalars α,β,γ, A, B, C, a, b, and c in these equations should be replaced by the corresponding vectors $\bar{\alpha},\bar{\beta},\bar{\gamma}$, $\bar{A}$, $\bar{B}$, $\bar{C}$, $\bar{a}$, $\bar{b}$, and $\bar{c}$.

Example 4.3

For East Asian countries, Cho and Retherford (1973) expressed the crude birth rate per 1,000 population as

$$\frac{1000B}{P} = \sum_i \frac{1000B_i}{M_i} \cdot \frac{M_i}{W_i} \cdot \frac{W_i}{P} \ , \tag{4.24}$$

where B_i, M_i, and W_i are, respectively, the number of births, the number of married women, and the number of total women in age group i, and B and P are the total number of births and the total population. In terms of our notation, we can write equation (4.24) as

$$R = F(\overline{\alpha},\overline{\beta},\overline{\gamma}) = \sum_i \alpha_i\beta_i\gamma_i \,, \tag{4.25}$$

where the vector-factors $\overline{\alpha}$,$\overline{\beta}$, and $\overline{\gamma}$ represent, respectively, the age-specific marital fertility rates per 1,000 women (it is assumed that all births occur to married women), the proportions of married women among total women in the age groups, and total women in the age groups as proportions of the total population.

Table 4.5 gives the three vector-factors for Taiwan for the years 1960 and 1970. The crude birth rates for 1960 and 1970 based on these data are, respectively, 38.77 and 27.20, the total difference being 11.57. The results in table 4.6 show that if, for example, neither the within-age group marital status structure ($\overline{\beta}$) nor the age-sex structure ($\overline{\gamma}$) was different in 1960 and 1970, but the age-specific marital fertility rates ($\overline{\alpha}$) varied as they did in the two years, then the crude birth rates in 1960 and 1970 would be 36.73 and 29.44, giving a difference of 7.29. The percent contributions of the vector-factors $\overline{\alpha}$, $\overline{\beta}$, and $\overline{\gamma}$ to the total difference of the two crude birth rates are, respectively, 63.0, 23.5, and 13.5. The decomposition in table 1 of Cho and Retherford agrees closely with these percentages.

It should be noted here that we can also express equation (4.24) as

$$\frac{1000B}{P} = \sum_i \frac{1000B_i}{M_i} \cdot \frac{M_i}{M} \cdot \frac{M}{P} \,, \tag{4.26}$$

where M is the total number of married women in the childbearing ages. $\overline{\beta}$ and $\overline{\gamma}$ in (4.26) represent, respectively, the age-structure of the married women, and the marital status-sex structure. Equations (4.24) and (4.26) are two different "hierarchical" models (Kim and Strobino, 1984; Das Gupta, 1989) and generate two different sets of results. By contrast, chapter 5 (Rate from Cross-Classified Data) deals with "symmetrical" models in which the results do not depend on the order in which the factors are considered.

Table 4.5. Crude Birth Rate per 1,000 as a Function of Three Vector-Factors: Taiwan, 1960 and 1970

Age groups	i	1970 (population 1)			1960 (population 2)		
		$1000B_i/M_i = A_i$	$M_i/W_i = B_i$	$W_i/P = C_i$	$1000B_i/M_i = a_i$	$M_i/W_i = b_i$	$W_i/P = c_i$
15 to 19	1	488	.082	.058	393	.122	.043
20 to 24	2	452	.527	.038	407	.622	.041
25 to 29	3	338	.866	.032	369	.903	.036
30 to 34	4	156	.941	.030	274	.930	.032
35 to 39	5	63	.942	.026	184	.916	.026
40 to 44	6	22	.923	.023	90	.873	.020
45 to 49	7	3	.876	.019	16	.800	.018
Crude birth rate (R) = 1000B/P		R_1 = 27.20			R_2 = 38.77		

Source: Cho and Retherford (1973), tables 2, 3, 4.

Table 4.6. Standardization and Decomposition of Crude Birth Rates in Table 4.5

Birth rates	Standardization		Decomposition	
	1960 (population 2)	1970 (population 1)	Difference (effects)	Percent distribution of effects
$\overline{\beta}\overline{\gamma}$-standardized rates	36.73	29.44	7.29 ($\overline{\alpha}$)	63.0
$\overline{\alpha}\overline{\gamma}$-standardized rates	34.47	31.75	2.72 ($\overline{\beta}$)	23.5
$\overline{\alpha}\overline{\beta}$-standardized rates	33.83	32.27	1.56 ($\overline{\gamma}$)	13.5
Crude birth rates (R)	38.77	27.20	11.57 (Total effect)	100.0

Program 4.3

The results in table 4.6 can be obtained by using Program 4.3 in which V(I,J,K)'s in line 2 are the data from table 4.5 corresponding to I = 1,2 (1970 and 1960); J = 1,3 ($\overline{\alpha}$, $\overline{\beta}$, and $\overline{\gamma}$); and K = 1,7 (seven age groups). The data file consists of six lines with the six vectors $\overline{A}$, $\overline{B}$, $\overline{C}$, $\overline{a}$, $\overline{b}$, and $\overline{c}$ in table 4.5, each line having seven elements. Program 4.3 is basically the same as Program 3.1 for three factors in chapter 3 except that it takes three lines (lines 11 through 13) to compute the rate R (i.e., H in the program) instead of a single line (line 11) used in Program 3.1. Program 4.3 is, therefore, two lines longer than Program 3.1.

4.4 THE CASE OF FOUR VECTOR-FACTORS

In this case, the rate R is written as

$$R = F(\overline{\alpha},\overline{\beta},\overline{\gamma},\overline{\delta}) \ , \tag{4.27}$$

and, therefore, in population 1 and population 2, the rates are

$$R_1 = F(\overline{A},\overline{B},\overline{C},\overline{D}) \ , \ R_2 = F(\overline{a},\overline{b},\overline{c},\overline{d}) \ . \tag{4.28}$$

The expressions for the standardized rates and the factor effects are the same as those in equations (3.28) through (3.31) except that the scalars α, β, γ, δ, A, B, C, D, a, b, c, and d should be replaced by the corresponding vectors $\overline{\alpha}$, $\overline{\beta}$, $\overline{\gamma}$, $\overline{\delta}$, $\overline{A}$, $\overline{B}$, $\overline{C}$, $\overline{D}$, $\overline{a}$, $\overline{b}$, $\overline{c}$, and $\overline{d}$.

Example 4.4

Smith and Cutright (1988) expressed the illegitimacy ratio (the ratio of births to unmarried women to total births) as

$$\frac{I}{I+L} = \frac{\sum_i \frac{W_i}{W}\cdot\frac{U_i}{W_i}\cdot\frac{I_i}{U_i}}{\sum_i \frac{W_i}{W}\cdot\frac{U_i}{W_i}\cdot\frac{I_i}{U_i}+\sum_i \frac{W_i}{W}\cdot\frac{M_i}{W_i}\cdot\frac{L_i}{M_i}}, \tag{4.29}$$

where U_i, M_i, and W_i are unmarried, married, and total women in age group i, and I_i and L_i are births to unmarried and married women in age group i. W, I, and L are the corresponding totals in the childbearing ages 15 to 44.

Using our notation, equation (4.29) can be written as

$$R = F(\overline{\alpha},\overline{\beta},\overline{\gamma},\overline{\delta}) = \frac{\sum_i \alpha_i\beta_i\gamma_i}{\sum_i \alpha_i\beta_i\gamma_i + \sum_i \alpha_i\ (1-\beta_i)\delta_i}, \tag{4.30}$$

where the vector-factors $\overline{\alpha}$, $\overline{\beta}$, $\overline{\gamma}$, and $\overline{\delta}$ represent, respectively, the age-structure of the women in the childbearing ages, the marital status structure within childbearing age groups, the age-specific nonmarital fertility rates, and the age-specific marital fertility rates.

Table 4.7 gives the values of the elements of the four vector-factors for White women for 1963 and 1983. The illegitimacy ratios based on these data and their standardization and decomposition are shown in table 4.8. There was an increase of 94.23 in the illegitimacy ratio in the 20-year period, from 30.95 in 1963 to 125.18 in 1983. This increase would have been only 27.06 (28.7 percent of the total increase) if only the age-specific nonmarital fertility rates changed as they did during the two decades but the other three factors were identical. On the other hand, the increase in the illegitimacy ratio would have been as high as 48.66 (51.7 percent of the total increase) if only the within-age group marital status structure changed as it did but the other three factors were identical. Thus, although the illegitimacy rates (i.e., the nonmarital fertility rates) by definition do not depend on the marital-status structure of the women in the

Program 4.3 (Three Factors)

```
      DIMENSION V(2,3,7),R(2,3,2),E(3),S(2,3)
      READ(5,1) (((V(I,J,K),K=1,7),J=1,3),I=1,2)
    1 FORMAT(7F6.0/7F6.3/7F6.3)
      DO 2 I=1,2
      DO 2 J=1,3
      DO 2 K=1,2
    2 R(I,J,K)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      DO 3 K=1,2
      H=0.0
      DO 7 L1=1,7
    7 H=H+V(I,1,L1)*V(J,2,L1)*V(K,3,L1)
      IF(I+J+K.EQ.3) R1=H
      IF(I+J+K.EQ.6) R2=H
      DO 3 L=1,2
      L1=L+1
      L2=6-L1
      IF(J+K.EQ.L1.OR.J+K.EQ.L2) R(I,1,L)=R(I,1,L)+H
      IF(I+K.EQ.L1.OR.I+K.EQ.L2) R(J,2,L)=R(J,2,L)+H
    3 IF(I+J.EQ.L1.OR.I+J.EQ.L2) R(K,3,L)=R(K,3,L)+H
      DO 5 J=1,3
      DO 4 I=1,2
    4 S(I,J)=R(I,J,1)/3.+R(I,J,2)/6.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,3),R2,R1,T
    6 FORMAT(40X,3F15.2)
      STOP
      END
****
```

Program 4.4 (Four Factors)

```
      DIMENSION V(2,4,6),R(2,4,2),E(4),S(2,4)
      READ(5,1) (((V(I,J,K),K=1,6),J=1,4),I=1,2)
    1 FORMAT(6F6.3)
      DO 2 I=1,2
      DO 2 J=1,4
      DO 2 K=1,2
    2 R(I,J,K)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      DO 3 K=1,2
      DO 3 L=1,2
      H1=0.0
      H2=0.0
      DO 7 M1=1,6
      H1=H1+V(I,1,M1)*V(J,2,M1)*V(K,3,M1)
    7 H2=H2+V(I,1,M1)*(1.-V(J,2,M1))*V(L,4,M1)
      H=H1/(H1+H2)
      IF(I+J+K+L.EQ.4) R1=H
      IF(I+J+K+L.EQ.8) R2=H
      DO 3 M=1,2
      M1=M+2
      M2=9-M1
      IF(J+K+L.EQ.M1.OR.J+K+L.EQ.M2) R(I,1,M)=R(I,1,M)+H
      IF(I+K+L.EQ.M1.OR.I+K+L.EQ.M2) R(J,2,M)=R(J,2,M)+H
      IF(I+J+L.EQ.M1.OR.I+J+L.EQ.M2) R(K,3,M)=R(K,3,M)+H
    3 IF(I+J+K.EQ.M1.OR.I+J+K.EQ.M2) R(L,4,M)=R(L,4,M)+H
      DO 5 J=1,4
      DO 4 I=1,2
    4 S(I,J)=R(I,J,1)/4.+R(I,J,2)/12.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,4),R2,R1,T
    6 FORMAT(40X,3F15.5)
      STOP
      END
****
```

childbearing ages, the significant shift in this latter structure during 1963-1983 in favor of nonmarriage had a tremendous boosting effect on the illegitimacy ratio. In table 4 of Smith and Cutright, the standardization was performed by holding one factor constant at a time, whereas our standardization holds three factors constant simultaneously allowing the fourth factor to vary. The two sets of standardizations are, therefore, not directly comparable. This example will be discussed further with five populations for five years in Example 6.5 (tables 6.9 and 6.10).

Table 4.7. Illegitimacy Ratio as a Function of Four Vector-Factors: United States, Whites, 1963 and 1983

Age groups	i	1963 (population 1)				1983 (population 2)			
		$W_i/W = A_i$	$U_i/W_i = B_i$	$I_i/U_i = C_i$	$L_i/M_i = D_i$	$W_i/W = a_i$	$U_i/W_i = b_i$	$I_i/U_i = c_i$	$L_i/M_i = d_i$
15 to 19	1	.200	.866	.007	.454	.169	.931	.018	.380
20 to 24	2	.163	.325	.021	.326	.195	.563	.026	.201
25 to 29	3	.146	.119	.023	.195	.190	.311	.023	.149
30 to 34	4	.154	.099	.015	.107	.174	.216	.016	.079
35 to 39	5	.168	.099	.008	.051	.150	.199	.008	.025
40 to 44	6	.169	.121	.002	.015	.122	.191	.002	.006
Illegitimacy ratio (R) = I/(I+L)		$R_1 = .03095$				$R_2 = .12518$			

Source: Smith and Cutright (1988), tables 2 and 3.

Table 4.8. Standardization and Decomposition of Illegitimacy Ratios in Table 4.7

(For convenience, results obtained from data in table 4.7 are multiplied by 1,000 before presenting them in table 4.8)

Illegitimacy ratios	Standardization		Decomposition	
	1983 (population 2)	1963 (population 1)	Difference (effects)	Percent distribution of effects
$\overline{\beta}\overline{\gamma}\overline{\delta}$ -standardized ratios	71.51	77.71	−6.20 ($\overline{\alpha}$)	−6.6
$\overline{\alpha}\overline{\gamma}\overline{\delta}$ -standardized ratios	96.08	47.42	48.66 ($\overline{\beta}$)	51.7
$\overline{\alpha}\overline{\beta}\overline{\delta}$ -standardized ratios	86.30	59.24	27.06 ($\overline{\gamma}$)	28.7
$\overline{\alpha}\overline{\beta}\overline{\gamma}$ -standardized ratios	84.34	59.63	24.71 ($\overline{\delta}$)	26.2
Overall illegitimacy ratios (R)	125.18	30.95	94.23 (Total effect)	100.0

Program 4.4

The results in table 4.8 can be obtained by using Program 4.4 in which V(I,J,K)'s in line 2 are the data from table 4.7 corresponding to I = 1,2 (1963 and 1983); J = 1,4 ($\overline{\alpha}, \overline{\beta}, \overline{\gamma}$, and $\overline{\delta}$); and K = 1,6 (six age groups). The data file consists of eight lines with the eight vectors $\overline{A}, \overline{B}, \overline{C}, \overline{D}, \overline{a}, \overline{b}, \overline{c}$, and $\overline{d}$ in table 4.7, each line having six elements. Program 4.4 is basically the same as Program 3.2 for four factors in chapter 3 except that it takes six lines (lines 12 through 17) to compute the rate R (i.e., H in the program) instead of a single line (line 12) used in Program 3.2. Program 4.4 is, therefore, five lines longer than Program 3.2.

4.5 THE CASE OF FIVE VECTOR-FACTORS

In this case, we can write the rate as

$$R = F(\overline{\alpha},\overline{\beta},\overline{\gamma},\overline{\delta},\overline{\epsilon}) \, , \tag{4.31}$$

which assumes the values

$$R_1 = F(\overline{A},\overline{B},\overline{C},\overline{D},\overline{E}) \, , \quad R_2 = F(\overline{a},\overline{b},\overline{c},\overline{d},\overline{e}) \tag{4.32}$$

in population 1 and population 2, respectively.

The standardized rates and the factor effects have the same expressions as those in (3.35) through (3.38) with the scalars α, β, γ, δ, ϵ, A, B, C, D, E, a, b, c, d, and e in them replaced by their corresponding vectors.

Example 4.5

Arriaga (1984) studied changes in life expectations as a result of changes in mortality rates in different age groups. In terms of a complete life table extending to age 109, we can express the expectation of life at birth $\overset{o}{e}_o$ as

$$\overset{o}{e}_o = \frac{L_o}{100000} + \frac{1-q_o}{2} + \sum_{y=1}^{109} \prod_{x=o}^{y} (1-q_x) \;, \tag{4.43}$$

where L_o is the stationary population in the age interval 0-1, and q_x is the probability that a person of exact age x will die before reaching the exact age x+1.

In table 4.9, $\overset{o}{e}_o$'s for White males for 1940 and 1980 are shown as a function of five vector-factors as follows:

$$\overset{o}{e}_o = R = F(\overline{\alpha},\overline{\beta},\overline{\gamma},\overline{\delta},\overline{\epsilon}) \;, \tag{4.44}$$

where, from the values of L_o and q_o in the two life tables, L_o in (4.33) is expressed as

$$L_o = 100054 - 86065\, q_o \;, \tag{4.35}$$

and

$$\overline{\alpha} = (q_o,q_1,...,q_{19}) \;, \overline{\beta} = (q_{20},q_{21},...,q_{39}) \;, \overline{\gamma} = (q_{40},q_{41},...,q_{59}) \;, \tag{4.36}$$

$$\overline{\delta} = (q_{60},q_{61},...,q_{79}) \;, \overline{\epsilon} = (q_{80},q_{81},...,q_{109}) \;.$$

There was an increase of 8.005 in the expectation of life at birth for White males in the four decades 1940-1980, from 62.812 in 1940 to 70.817 in 1980. The standardization and decomposition in table 4.10 show how this increase in $\overset{o}{e}_o$ can be attributed to the decrease in the mortality rates in the age groups 0 to 20, 20 to 40, 40 to 60, 60 to 80, and 80 and over. From the last column, we find that, in terms of percentages, the contributions made by these age groups towards the overall increase in $\overset{o}{e}_o$ are, respectively, 44.3, 10.3, 22.0, 18.9, and 4.5. Arriaga's decompositions do not include one that corresponds to the data in table 4.9; therefore, we cannot compare our results with his.

Suchindran and Koo (1992) used this formulation to decompose the difference between two mean ages at last birth into the effects of the differences in five factors (which include one scalar factor and four vector-factors), namely, age at first birth, earlier parity progression ratios, later parity progression ratios, earlier birth intervals, and later birth intervals.

Table 4.9. Expectation of Life at Birth ($\mathring{e}_o$) as a Function of Five Vector-Factors: United States, White Males, 1940 and 1980

1940 (population 1)						1980 (population 2)					
Age x	q_x		Age x	q_x		Age x	q_x		Age x	q_x	
0	.04812		60	.02548		0	.01231		60	.01762	
1	.00487		61	.02743		1	.00092		61	.01933	
2	.00265		62	.02952		2	.00066		62	.02119	
3	.00190		63	.03177		3	.00053		63	.02316	
4	.00153		64	.03420		4	.00043		64	.02523	
5	.00138		65	.03685		5	.00039		65	.02738	
6	.00124		66	.03975		6	.00037		66	.02968	
7	.00114		67	.04293		7	.00034		67	.03218	
8	.00106		68	.04643		8	.00030		68	.03495	
9	.00102	$\bar{A}$	69	.05028	$\bar{D}$	9	.00024	$\bar{a}$	69	.03805	$\bar{d}$
10	.00100		70	.05454		10	.00019		70	.04148	
11	.00101		71	.05924		11	.00019		71	.04516	
12	.00106		72	.06443		12	.00028		72	.04901	
13	.00114		73	.07014		13	.00046		73	.05295	
14	.00127		74	.07637		14	.00071		74	.05703	
15	.00143		75	.08313		15	.00096		75	.06146	
16	.00158		76	.09040		16	.00118		76	.06642	
17	.00172		77	.09818		17	.00137		77	.07180	
18	.00186		78	.10647		18	.00151		78	.07762	
19	.00199		79	.11530		19	.00163		79	.08394	
20	.00212		80	.12471		20	.00175		80	.09099	
21	.00223		81	.13472		21	.00186		81	.09886	
22	.00232		82	.14537		22	.00193		82	.10733	
23	.00238		83	.15668		23	.00193		83	.11613	
24	.00241		84	.16859		24	.00189		84	.12523	
25	.00243		85	.18104		25	.00183		85	.13507	
26	.00245		86	.19395		26	.00177		86	.14592	
27	.00251		87	.20727		27	.00172		87	.15691	
28	.00259		88	.22091		28	.00168		88	.16774	
29	.00268	$\bar{B}$	89	.23482	$\bar{E}$	29	.00167	$\bar{b}$	89	.17875	$\bar{e}$
30	.00279		90	.24894		30	.00166		90	.19058	
31	.00291		91	.26322		31	.00165		91	.20389	
32	.00306		92	.27760		32	.00166		92	.21864	
33	.00323		93	.29202		33	.00169		93	.23453	
34	.00342		94	.30642		34	.00175		94	.25061	
35	.00363		95	.32076		35	.00184		95	.26617	
36	.00387		96	.33496		36	.00196		96	.28001	
37	.00414		97	.34898		37	.00209		97	.29311	
38	.00443		98	.36275		38	.00224		98	.30545	
39	.00476		99	.37623		39	.00240		99	.31703	
			100	.38935					100	.32784	
40	.00513		101	.40205		40	.00261		101	.33791	
41	.00554		102	.41429		41	.00287		102	.34724	
42	.00600		103	.42599		42	.00316		103	.35588	
43	.00650		104	.43712		43	.00348		104	.36384	
44	.00706		105	.44760		44	.00382		105	.37117	
45	.00766		106	.45738		45	.00420		106	.37790	
46	.00833		107	.46640		46	.00463		107	.38407	
47	.00904		108	.47462		47	.00514		108	.38971	
48	.00981	$\bar{C}$	109	.48000		48	.00573	$\bar{c}$	109	.39486	
49	.01064					49	.00639				
50	.01155					50	.00706				
51	.01253					51	.00775				
52	.01360		$\mathring{e}_o$ (1940)			52	.00850		$\mathring{e}_o$ (1980)		
53	.01476					53	.00934				
54	.01602					54	.01027				
55	.01737		$= R_1$			55	.01125		$= R_2$		
56	.01881					56	.01227				
57	.02034		= 62.812			57	.01338		= 70.817		
58	.02195					58	.01464				
59	.02366					59	.01605				

Source: U.S. Bureau of the Census (1946), table 5; National Center for Health Statistics (1985), table 5.

Table 4.10. Standardization and Decomposition of Expectations of Life at Birth in Table 4.9

Expectation of life at birth	Standardization		Decomposition	
	1980 (population 2)	1940 (population 1)	Difference (effects)	Percent distribution of effects
$\overline{\beta}\overline{\gamma}\overline{\delta}\overline{\epsilon}$-standardized expectations	68.463	64.917	3.546 ($\overline{\alpha}$)	44.3
$\overline{\alpha}\overline{\gamma}\overline{\delta}\overline{\epsilon}$-standardized expectations	67.112	66.286	.826 ($\overline{\beta}$)	10.3
$\overline{\alpha}\overline{\beta}\overline{\delta}\overline{\epsilon}$-standardized expectations	67.566	65.802	1.764 ($\overline{\gamma}$)	22.0
$\overline{\alpha}\overline{\beta}\overline{\gamma}\overline{\epsilon}$-standardized expectations	67.433	65.925	1.508 ($\overline{\delta}$)	18.9
$\overline{\alpha}\overline{\beta}\overline{\gamma}\overline{\delta}$-standardized expectations	66.876	66.515	.361 ($\overline{\epsilon}$)	4.5
Overall expectation of life at birth (R)	70.817	62.812	8.005 (Total effect)	100.0

Program 4.5

The results in table 4.10 can be obtained by using Program 4.5 in which V(I,J,K)'s in line 2 are the data from table 4.9 giving 220 q_x's corresponding to I = 1,2 (1940 and 1980); J = 1,5 ($\overline{\alpha}$, $\overline{\beta}$, $\overline{\gamma}$, $\overline{\delta}$, and $\overline{\epsilon}$); and K = 1,20 (20 single-year age groups for the first four vector-factors) or K = 1,30 (30 single-year age groups for the fifth vector-factor). The data file, therefore, consists of 22 lines: lines 1 through 11 are for 110 q_x values for 1940 and lines 12 through 22 are for 110 q_x values for 1980 (each of lines 1 through 22 having 10 values), the format being as shown in line 3 of the program. Program 4.5 is basically the same as Program 3.3 for five factors in chapter 3 except that it has 11 additional lines (lines 13 through 23) for the computation of the rate R (i.e., H in the program). Program 4.5 is, therefore, 11 lines longer than Program 3.3.

4.6 THE CASE OF SIX VECTOR-FACTORS

When there are six vector-factors so that

$$R = F(\overline{\alpha},\overline{\beta},\overline{\gamma},\overline{\delta},\overline{\epsilon},\overline{\eta}), \quad (4.37)$$

and in the two populations,

$$R_1 = F(\overline{A},\overline{B},\overline{C},\overline{D},\overline{E},\overline{F}), \quad R_2 = F(\overline{a},\overline{b},\overline{c},\overline{d},\overline{e},\overline{f}), \quad (4.38)$$

then the standardized rates and the factor effects have the same expressions as those in (3.42) through (3.45) except that the scalars have to be replaced by their corresponding vectors.

Example 4.6

As in Example 4.5, the changes in life expectations can also be decomposed into the effects of changes in mortality by different causes of death (Pollard, 1988; Myers, 1991). Table 4.11 gives the data from the U.S. total abridged life tables for 1962 and 1987 expressing the expectation of life at birth $\overset{o}{e}_o$ (=R) as

$$R = F(\overline{\alpha},\overline{\beta},\overline{\gamma},\overline{\delta},\overline{\epsilon},\overline{\eta}), \quad (4.39)$$

where

$$\overline{\alpha} = ({}_1q_o{}^{(1)}, {}_4q_1{}^{(1)}, \ldots, {}_5q_{80}{}^{(1)}), \overline{\beta} = ({}_1q_o{}^{(2)}, {}_4q_1{}^{(2)}, \ldots, {}_5q_{80}{}^{(2)}),$$
$$\overline{\gamma} = ({}_1q_o{}^{(3)}, {}_4q_1{}^{(3)}, \ldots, {}_5q_{80}{}^{(3)}), \overline{\delta} = ({}_1q_o{}^{(4)}, {}_4q_1{}^{(4)}, \ldots, {}_5q_{80}{}^{(4)}), \quad (4.40)$$
$$\overline{\epsilon} = ({}_1q_o{}^{(5)}, {}_4q_1{}^{(5)}, \ldots, {}_5q_{80}{}^{(5)}), \overline{\eta} = ({}_1q_o{}^{(6)}, {}_4q_1{}^{(6)}, \ldots, {}_5q_{80}{}^{(6)}),$$

Program 4.5 (Five Factors)

```
      DIMENSION V(2,5,30),R(2,5,3),E(5),S(2,5),Q(110)
      READ(5,1) (((V(I,J,K),K=1,20),J=1,4),(V(I,5,K),K=1,30),I=1,2)
    1 FORMAT(10F8.5)
      DO 2 I=1,2
      DO 2 J=1,5
      DO 2 K=1,3
    2 R(I,J,K)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      DO 3 K=1,2
      DO 3 L=1,2
      DO 3 M=1,2
      DO 7 N1=1,30
      IF(N1.GT.20) GO TO 7
      Q(N1    )=V(I,1,N1)
      Q(N1+20)=V(J,2,N1)
      Q(N1+40)=V(K,3,N1)
      Q(N1+60)=V(L,4,N1)
    7 Q(N1+80)=V(M,5,N1)
      H=(100054.-86065.*Q(1))/100000.+(1.-Q(1))/2.
      QQ=1.-Q(1)
      DO 8 N1=2,110
      QQ=QQ*(1.-Q(N1))
    8 H=H+QQ
      IF(I+J+K+L+M.EQ. 5) R1=H
      IF(I+J+K+L+M.EQ.10) R2=H
      DO 3 N=1,3
      N1=N+3
      N2=12-N1
      IF(J+K+L+M.EQ.N1.OR.J+K+L+M.EQ.N2) R(I,1,N)=R(I,1,N)+H
      IF(I+K+L+M.EQ.N1.OR.I+K+L+M.EQ.N2) R(J,2,N)=R(J,2,N)+H
      IF(I+J+L+M.EQ.N1.OR.I+J+L+M.EQ.N2) R(K,3,N)=R(K,3,N)+H
      IF(I+J+K+M.EQ.N1.OR.I+J+K+M.EQ.N2) R(L,4,N)=R(L,4,N)+H
    3 IF(I+J+K+L.EQ.N1.OR.I+J+K+L.EQ.N2) R(M,5,N)=R(M,5,N)+H
      DO 5 J=1,5
      DO 4 I=1,2
    4 S(I,J)=R(I,J,1)/5.+R(I,J,2)/20.+R(I,J,3)/30.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,5),R2,R1,T
    6 FORMAT(40X,3F15.3)
      STOP
      END
****
```

Program 4.6 (Six Factors)

```
      DIMENSION V(2,6,18),A(19),B(19),R(2,6,3),E(6),S(2,6)
      READ(5,1) (((V(I,J,K),J=1,6),K=1,18),I=1,2)
    1 FORMAT(6F10.6)
      READ(5,8) (A(K),B(K),K=1,19)
    8 FORMAT(2F10.4)
      DO 2 I=1,2
      DO 2 J=1,6
      DO 2 K=1,3
    2 R(I,J,K)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      DO 3 K=1,2
      DO 3 L=1,2
      DO 3 M=1,2
      DO 3 N=1,2
      EL=1.0
      H=0.0
      DO 9 N1=1,18
      Q=V(I,1,N1)+V(J,2,N1)+V(K,3,N1)+V(L,4,N1)+V(M,5,N1)+V(N,6,N1)
      H=H+EL*(A(N1)+B(N1)*Q)
    9 EL=EL*(1.-Q)
      H=H+EL*(A(19)+B(19)*(1.-EL))
      IF(I+J+K+L+M+N.EQ. 6) R1=H
      IF(I+J+K+L+M+N.EQ.12) R2=H
      DO 3 KK=1,3
      K1=KK+4
      K2=15-K1
      IF(J+K+L+M+N.EQ.K1.OR.J+K+L+M+N.EQ.K2) R(I,1,KK)=R(I,1,KK)+H
      IF(I+K+L+M+N.EQ.K1.OR.I+K+L+M+N.EQ.K2) R(J,2,KK)=R(J,2,KK)+H
      IF(I+J+L+M+N.EQ.K1.OR.I+J+L+M+N.EQ.K2) R(K,3,KK)=R(K,3,KK)+H
      IF(I+J+K+M+N.EQ.K1.OR.I+J+K+M+N.EQ.K2) R(L,4,KK)=R(L,4,KK)+H
      IF(I+J+K+L+N.EQ.K1.OR.I+J+K+L+N.EQ.K2) R(M,5,KK)=R(M,5,KK)+H
    3 IF(I+J+K+L+M.EQ.K1.OR.I+J+K+L+M.EQ.K2) R(N,6,KK)=R(N,6,KK)+H
      DO 5 J=1,6
      DO 4 I=1,2
    4 S(I,J)=R(I,J,1)/6.+R(I,J,2)/30.+R(I,J,3)/60.
    5 E(J)=S(2,J)-S(1,J)
      T=R2-R1
      WRITE(6,6) (S(2,J),S(1,J),E(J),J=1,6),R2,R1,T
    6 FORMAT(40X,3F15.3)
      STOP
      END
****
```

and

$$ {}_nq_x = \sum_{i=1}^{6} {}_nq_x^{(i)} , \qquad (4.41) $$

i (= 1,2,...,6) being the six categories of causes of death as shown in table 4.11.

The ${}_nq_x^{(i)}$-values are obtained from the corresponding ${}_nq_x$ of the abridged life table and from the death statistics of the six causes. For example, in the age group 5 to 10 for 1987, ${}_5q_5$ is .001225 and the deaths in the six causes are, respectively, 149, 46, 681, 151, 2, 231, and 1,043, the total number of deaths being 4,301. We compute

$$ {}_5q_5^{(1)} = .001225x(149/4301) = .000043 , \qquad (4.42) $$

$$ {}_5q_5^{(2)} = .001225x(46/4301) = .000013 , $$

and so on. These values are shown in table 4.11.

Table 4.11 also shows the values of ${}_nG_x$ and ${}_nH_x$ where

$$ ({}_nL_x/l_x) = {}_nG_x + {}_nH_x \cdot {}_nq_x , \; x = 0,1,5...,80 , $$

$$ ({}_\infty L_{85}/l_{85}) = {}_\infty G_{85} + {}_\infty H_{85} \cdot {}_{85}q_0 , \qquad (4.43) $$

and where each of the 19 straight lines is fitted from the two points corresponding to the abridged life tables for 1962 and 1987. For example, for age group 5 to 10, ${}_5L_5$, l_5, and ${}_5q_5$ are 484,912, 97100, and .00225541 for 1962 and 493,611, 98788, and .00122485 for 1987. Therefore,

$$ {}_5H_5 = [(484912/97100) - (493611/98788)]/(.00225541 - .00122485) $$
$$ = -2.6444 , \qquad (4.44) $$

and

$$ {}_5G_5 = (484912/97100) + 2.6444x.00225541 = 4.9999 . $$

Again, for solving the last equation in (4.43), we have ${}_\infty L_{85}$, l_{85}, and ${}_{85}q_o$ equal to 88,325, 19101, and .80899 for 1962 and 183,453, 30220, and .69780 for 1987. Therefore,

$$ {}_\infty H_{85} = [(88325/19101) - (183453/30220)]/(.80899 - .69780) \qquad (4.45) $$
$$ = -13.0091 , $$

and

$$ {}_\infty G_{85} = (88325/19101) + 13.0091x.80899 = 15.1483 . $$

It is evident from (4.39) and from the formulas for six vector-factors similar to (3.45) that we need to compute the expectation of life at birth for 2^6 combinations of the vector-factors for the two years. These computations for a particular combination may proceed as follows:

$$ l_o = 1.0 , $$

$$ {}_nq_x = \sum_{i=1}^{6} {}_nq_x^{(i)} , \; {}_nL_x = l_x({}_nG_x + {}_nH_x \cdot {}_nq_x) , $$

$$ l_{x+n} = l_x(1 - {}_nq_x) , \; x = 0,1,5,...,80 , $$

$$ {}_\infty L_{85} = l_{85}[{}_\infty G_{85} + {}_\infty H_{85}(1 - l_{85})] , \qquad (4.46) $$

$$ \overset{o}{e}_o = \sum_{x=0,1,5,...,85} {}_nL_x . $$

Table 4.11. Expectation of Life at Birth ($\overset{o}{e}_o$) as a Function of Six Vector-Factors: United States, Total, 1962 and 1987

Age interval (x to x+n)	Diseases of heart[1] $_nq_x^{(1)}$	Other diseases of circulatory system $_nq_x^{(2)}$	Neoplasms $_nq_x^{(3)}$	Diseases of respiratory system $_nq_x^{(4)}$	Accidents (E800-E999) $_nq_x^{(5)}$	Residual $_nq_x^{(6)}$	Straight line fits
	1962 (population 1)						$_nG_x$
	$\bar{A}$	$\bar{B}$	$\bar{C}$	$\bar{D}$	$\bar{E}$	$\bar{F}$	
0 to 1	.000067	.000008	.000092	.002767	.000926	.021391	1.0004
1 to 5	.000042	.000006	.000404	.000688	.001231	.001476	4.0003
5 to 10	.000043	.000005	.000414	.000178	.000923	.000693	4.9999
10 to 15	.000061	.000011	.000340	.000129	.000968	.000617	5.0013
15 to 20	.000122	.000032	.000415	.000155	.002789	.000831	5.0068
20 to 25	.000248	.000070	.000488	.000173	.003960	.001190	4.9994
25 to 30	.000454	.000109	.000720	.000221	.003312	.001602	4.9946
30 to 35	.001010	.000173	.001256	.000265	.003163	.002192	4.9956
35 to 40	.002386	.000286	.002222	.000449	.003198	.003253	4.9993
40 to 45	.004994	.000407	.003905	.000658	.003350	.004654	5.0003
45 to 50	.009531	.000631	.006811	.000998	.003625	.006940	5.0016
50 to 55	.017209	.001011	.011176	.001615	.004022	.010226	5.0010
55 to 60	.027175	.001568	.016085	.002609	.004120	.014418	5.0023
60 to 65	.043154	.002850	.023030	.004191	.004429	.021757	5.0095
65 to 70	.066734	.005040	.031189	.006720	.004839	.033799	5.0049
70 to 75	.095732	.008388	.037083	.009641	.005880	.050108	5.0102
75 to 80	.140593	.014886	.043462	.014269	.008150	.077066	5.0483
80 to 85	.205785	.028057	.049336	.022368	.012517	.117353	5.1383
85+	–	–	–	–	–	–	15.1483
	1987 (population 2)						$_nH_x$
	$\bar{a}$	$\bar{b}$	$\bar{c}$	$\bar{d}$	$\bar{e}$	$\bar{f}$	
0 to 1	.000250	.000049	.000042	.000324	.000337	.009107	−0.8989
1 to 5	.000087	.000019	.000164	.000117	.000905	.000738	−2.4973
5 to 10	.000043	.000013	.000194	.000043	.000635	.000297	−2.6444
10 to 15	.000052	.000017	.000161	.000052	.000803	.000242	−2.9417
15 to 20	.000105	.000027	.000231	.000074	.003341	.000423	−3.8883
20 to 25	.000169	.000066	.000300	.000096	.004284	.000741	−2.3457
25 to 30	.000288	.000114	.000465	.000129	.003728	.001293	−1.6032
30 to 35	.000557	.000189	.000820	.000204	.003336	.002185	−1.8383
35 to 40	.001169	.000352	.001584	.000301	.003128	.002814	−2.2804
40 to 45	.002517	.000603	.002968	.000413	.002781	.003070	−2.3455
45 to 50	.004949	.000987	.005901	.000684	.002724	.003843	−2.3983
50 to 55	.009167	.001703	.010753	.001331	.002683	.005110	−2.3768
55 to 60	.015050	.002655	.017474	.002504	.002780	.006882	−2.4042
60 to 65	.024774	.004722	.026434	.004935	.002931	.009766	−2.4935
65 to 70	.037764	.007857	.035297	.008276	.003129	.013547	−2.4388
70 to 75	.059333	.014468	.046072	.014341	.003904	.020403	−2.4771
75 to 80	.090378	.025200	.054446	.022443	.005236	.031492	−2.6499
80 to 85	.143461	.045047	.061850	.033924	.007245	.049208	−2.8923
85+	–	–	–	–	–	–	−13.0091
$\overset{o}{e}_o = R$	R_1 (1962) = 70.035, R_2 (1987) = 74.963						

[1]Codes 390-398, 402, 404-429.

Source: National Center for Health Statistics (1964, tables 1-23, 5-1; 1990b, tables 1-26, 6-1).

Table 4.12. Standardization and Decomposition of Expectations of Life at Birth in Table 4.11

Expectation of life at birth	Standardization		Decomposition	
	1987 (population 2)	1962 (population 1)	Difference (effects)	Percent distribution of effects
$\overline{\beta\gamma\delta\epsilon\eta}$ -standardized expectations	73.587	71.315	2.272 ($\overline{\alpha}$)	46.1
$\overline{\alpha\gamma\delta\epsilon\eta}$ -standardized expectations	72.321	72.649	–.328 ($\overline{\beta}$)	–6.6
$\overline{\alpha\beta\delta\epsilon\eta}$ -standardized expectations	72.405	72.562	–.157 ($\overline{\gamma}$)	–3.2
$\overline{\alpha\beta\gamma\epsilon\eta}$ -standardized expectations	72.530	72.427	.103 ($\overline{\delta}$)	2.1
$\overline{\alpha\beta\gamma\delta\eta}$ -standardized expectations	72.585	72.353	.232 ($\overline{\epsilon}$)	4.7
$\overline{\alpha\beta\gamma\delta\epsilon}$ -standardized expectations	73.853	71.047	2.806 ($\overline{\eta}$)	56.9
Overall expectation of life at birth (R)	74.963	70.035	4.928 (Total effect)	100.0

The results of the standardization and decomposition of the expectations of life at birth in table 4.11 are shown in table 4.12. The expectations of life at birth were 70.035 in 1962 and 74.963 in 1987, the total increase in $\overset{o}{e}_o$ during the 25-year period being 4.928. If the mortality rates from the diseases of the heart differed as they did in 1962 and 1987, and those from all other causes of death were identical in the two years, then the $\overset{o}{e}_o$'s in 1962 and 1987 would be 71.315 and 73.587, respectively, showing an increase of 2.272. In other words, 46.1 percent of the increase in the $\overset{o}{e}_o$ during the 25-year period can be attributed to the decline in the mortality rates from the diseases of the heart. On the other hand, other diseases of the circulatory system and neoplasms had negative effects on the increase in the $\overset{o}{e}_o$; i.e., without changes in the other four cause-of-death categories, the $\overset{o}{e}_o$ in 1987 would have been lower than that in 1962.

The techniques in Examples 4.5 and 4.6 can be easily combined to handle both age groups and causes of death simultaneously, as Pollard did. His results on the Australian data are not directly comparable with ours.

Program 4.6

The results in table 4.12 can be obtained by using Program 4.6 in which V(I,J,K)'s in line 2 are the data from table 4.11 corresponding to I = 1,2 (1962 and 1987); J = 1,6 ($\overline{\alpha}, \overline{\beta}, \overline{\gamma}, \overline{\delta}, \overline{\epsilon}$, and $\overline{\eta}$); and K = 1,18 (18 age groups 0 to 1, 1 to 5,..., 80 to 85). A(K)'s and B(K)'s in line 4 of the program are 19 pairs of straight line parameters $_nG_x$'s and $_nH_x$'s given in table 4.11. The data file, therefore, consists of 55 lines: lines 1 through 18 are for $_nq_x^{(i)}$ values for 1962 corresponding to 18 age groups, each line having six such values for six cause-of-death categories; lines 19 through 36 give the same data for 1987; and lines 37 through 55 are for 19 straight line parameters for the 19 age groups, each line having two values. The formats for these data inputs are given in lines 3 and 5 of the program. Program 4.6 is basically the same as Program 3.4 for six factors in chapter 3 except that it has two additional lines (lines 4,5) for data input and six additional lines (lines 16 through 21) for the computation of the rate R (i.e., H in the program), as shown in the equations in (4.46). Program 4.6 is, therefore, eight lines longer than Program 3.4.

4.7 P VECTOR-FACTORS AND THE GENERAL PROGRAM

When there are P vector-factors so that

$$R = F(\overline{\alpha}_1, \overline{\alpha}_2, ..., \overline{\alpha}_p) , \tag{4.47}$$

and in the two populations,

$$R_1 = F(\overline{A}_1, \overline{A}_2, ..., \overline{A}_p), \quad R_2 = F(\overline{a}_1, \overline{a}_2, ..., \overline{a}_p) , \tag{4.48}$$

then the standardized rates and the factor effects have the same expressions as those in (3.51) through (3.54) except that the scalars have to be replaced by their corresponding vectors.

The general program for up to 10 factors (Program 2.3) given in section 2.8 can also be used for any number of factors up to 10 for the standardization and decomposition problems in chapter 4 (i.e., when the rate is a function of vector-factors). The only changes needed in Program 2.3 are in the dimension statement in line 1, the input statement and format in lines 2 and 3, the output statement and format in lines 42 and 43, and in the expression for the rate in lines 18 and 19. In particular, the computation of the rate may take several lines in the program because of more involved data. We show below the specific changes in Program 2.3 that will be needed to generate the results in tables 4.2, 4.4, 4.6, 4.8, 4.10, and 4.12 corresponding to Examples 4.1 through 4.6 in this chapter. As before, no changes are needed in the data files previously created to be used with the specific programs.

Example 4.1 (two factors)

Line 1:	Replace P(2,10) by V(2,10,500)
Lines 2,3:	Replace by lines 2,3 in Program 4.1
Lines 18,19:	Replace by lines 9-21 in Program 4.1
Lines 42,43:	Replace 10 by 2 and 15.3 by 15.5

Example 4.2 (two factors)

Line 1:	Replace P(2,10) by V(2,10,500)
Lines 2,3:	Replace by lines 2,3 in Program 4.2
Lines 18,19:	Replace by lines 9-16 in Program 4.2
Lines 42,43:	Replace 10 by 2

Example 4.3 (three factors)

Line 1:	Replace P(2,10) by V(2,10,500)
Lines 2,3:	Replace by lines 2,3 in Program 4.3
Lines 18,19:	Replace by lines 11-13 in Program 4.3
Lines 42,43:	Replace 10 by 3 and 15.3 by 15.2

Example 4.4 (four factors)

Line 1:	Replace P(2,10) by V(2,10,500)
Lines 2,3:	Replace by lines 2,3 in Program 4.4
Lines 18,19:	Replace by lines 12-17 in Program 4.4
Lines 42,43:	Replace 10 by 4 and 15.3 by 15.5

Example 4.5 (five factors)

Line 1:	Replace P(2,10) by V(2,10,500) and add Q(110)
Lines 2,3:	Replace by lines 2,3 in Program 4.5
Lines 18,19:	Replace by lines 13-24 in Program 4.5
Lines 42,43:	Replace 10 by 5

Example 4.6 (six factors)

Line 1:	Replace P(2,10) by V(2,10,500) and add A(19), B(19)
Lines 2,3:	Replace by lines 2-5 in Program 4.6
Lines 18,19:	Replace by lines 16-22 in Program 4.6
Lines 42,43:	Replace 10 by 6

Chapter 5. Rate From Cross-Classified Data

5.1 INTRODUCTION

Most of the papers on standardization and decomposition published so far deal with the case in which the techniques are performed on cross-classified data involving one or more factors. For example, Liao (1989) decomposed the difference between two crude death rates into the effects of age and race (Example 5.3). Sweet (1984) studied the growth of households as a result of the changes in age and marital status composition (Example 5.6). Again, Wilson (1988) decomposed the difference in the mobility rates in terms of age and education (Example 5.7).

Unlike the situations in the preceding chapters, the decomposition in the case of cross-classified data involves an additional effect, namely, the effect of the differences in the cell-specific rates, called the rate-effect. In other words, if the cross-classification involves, say, three factors, namely, age (I), sex (J), and marital status (K), then the decomposition generates four additive effects: the age (I)-effect, the sex (J)-effect, the marital status (K)-effect, and the rate (R)-effect. The most crucial part in the development of decomposition technique in this case is expressing the proportion of population in a cell in the cross-classification in terms of the product of a number of symmetrical expressions (equal to the number of factors) that represent the factors involved, as in equation (5.7) for two factors and in equation (5.15) for three factors.

5.2 THE CASE OF ONE FACTOR

When there is only one factor I, N_i and T_i are the number of persons and the rate for the ith category of I in population 1, N. and T. being the corresponding total number of persons and the crude rate. For population 2, analogous symbols are used with lower-case letters n and t.

The crude rates can be expressed as

$$T. = \sum_i \frac{T_i N_i}{N.}, \quad t. = \sum_i \frac{t_i n_i}{n.}. \tag{5.1}$$

Writing

$$\frac{N_i}{N.} = A_i, \quad \frac{n_i}{n.} = a_i, \tag{5.2}$$

it follows from Das Gupta (1991, formula 18) that

$$\begin{aligned} t.-T. &= \text{R (rate)-effect} + \text{I-effect} \\ &= [R(\bar{t}) - R(\bar{T})] + [I(\bar{a}) - I(\bar{A})], \end{aligned}$$

where

$$\begin{aligned} R(\bar{T}) &= \text{I-standardized rate in population 1} \\ &= \sum_i \frac{\frac{n_i}{n.} + \frac{N_i}{N.}}{2} T_i, \end{aligned} \tag{5.3}$$

$$\begin{aligned} I(\bar{A}) &= \text{R-standardized rate in population 1} \\ &= \sum_i \frac{t_i + T_i}{2} A_i, \end{aligned} \tag{5.4}$$

and $R(\bar{t})$ and $I(\bar{a})$ have the same expressions as those in (5.3) and (5.4), respectively, with T_i in (5.3) replaced by t_i and A_i in (5.4) replaced by a_i.

It is clear from the discussions in section 4.2 that the present case can also be treated as a case of two vector-factors, the vectors in (4.4) being

$$\bar{A} = \left(\frac{N_1}{N.}, \frac{N_2}{N.}, \ldots\right), \quad \bar{B} = (T_1, T_2, \ldots), \tag{5.5}$$
$$\bar{a} = \left(\frac{n_1}{n.}, \frac{n_2}{n.}, \ldots\right), \quad \bar{b} = (t_1, t_2, \ldots),$$

so that the rates in (4.3) are

$$R_1 = T. = F(\bar{A}, \bar{B}), \quad R_2 = t. = F(\bar{a}, \bar{b}) .$$

Example 5.1

The data in table 5.1 are taken from Santi (1989) where the percentage distribution of population and household headship rates by age groups are given for 1970 and 1985 for the United States. The headship rates are 44.727 and 47.694 for 1970 and 1985, respectively, the difference between them being 2.967. Table 5.2 shows that if the age-specific headship rates varied as they did in 1970 and 1985, but the age structures of the populations were identical in the two years, then the headship rates in 1970 and 1985 would be, respectively, 45.331 and 47.071, giving a difference of 1.740. In other words, 41.4 percent of the total difference between the headship rates in 1970 and 1985 is due to the difference in the age structures of the populations in the two years. The remaining 58.6 percent of the difference is the so-called "real" difference (i.e., the effect of the difference in the age-specific headship rates). We will discuss this problem again in Example 6.2 (tables 6.3 and 6.4) to compare Santi's results with ours when four populations for the four years 1970, 1975, 1980, and 1985 are considered simultaneously.

Table 5.1. Population Sizes (Percents) and Household Headship Rates per 100 by Age Groups: United States, 1970 and 1985

Age groups	i	1970 (population 1)		1985 (population 2)	
		Size N_i	Rate T_i	Size n_i	Rate t_i
15 to 19	1	12.9	1.9	10.1	2.2
20 to 24	2	10.9	25.8	11.2	24.3
25 to 29	3	9.5	45.7	11.6	45.8
30 to 34	4	8.0	49.6	10.9	52.5
35 to 39	5	7.8	51.2	9.4	56.1
40 to 44	6	8.4	51.6	7.7	55.6
45 to 49	7	8.6	51.8	6.3	56.0
50 to 54	8	7.8	54.9	6.0	57.4
55 to 59	9	7.0	58.7	6.3	57.2
60 to 64	10	5.9	60.4	5.9	61.2
65 to 69	11	4.7	62.8	5.1	63.9
70 to 74	12	3.6	66.6	4.0	68.6
75+	13	4.9	66.8	5.5	72.2
All ages	i = .	100.0	44.727	100.0	47.694

Source: Santi (1989), table 1.

Program 5.1

The results in table 5.2 can be obtained by using Program 5.1 in which P(I,J)'s are N_i's and n_i's, and T(I,J)'s are T_i's and t_i's in table 5.1. In other words, the data file consists of four lines corresponding to the data in the last four columns in table 5.1 in the same order, each line having 13 numbers with the format specified in line 4 of the program. The four standardized rates in table 5.2 are given by ER(J)'s and S(J)'s in lines 17 and 18 of the program. The two effects in table 5.2 are denoted by ERR and U in lines 20 and 21 of the program.

Table 5.2. Standardization and Decomposition of Household Headship Rates in Table 5.1

Household headship rates	Standardization		Decomposition	
	1985 (population 2)	1970 (population 1)	Difference (effects)	Percent distribution of effects
R (rate)-standardized headship rates	46.815	45.588	1.227 (I=age)	41.4
I (age)-standardized headship rates	47.071	45.331	1.740 (R=rate)	58.6
Overall headship rates	47.694	44.727	2.967 (Total effect)	100.0

Alternatively, in view of (5.5), we can also use Program 4.2 and the same data file to obtain the results in table 5.1. The only changes needed in Program 4.2 are as follows:

Lines 1,2: Replace 480 by 13

Line 3: Replace 8F10.7 by 13F5.1

Lines 9-16: Replace by the following three lines:

```
   H = 0.0
   DO 7 K1 = 1,13
7  H = H+V(I,1,K1)*V(J,2,K1)/100.
```

Example 5.2

We consider another one-factor data from Clogg and Eliason (1988) in table 5.3 where the percent desiring more children is compared for two groups of women: parity 1 and parity 4+. The women and the percentage of them desiring more children are given by age groups. The issue here is how to eliminate the effect of the difference in the age structures in the two parity groups from the overall difference in the percents desiring more children. Of the women with parity 1, 72.093 percent desire more children, whereas the corresponding percentage for women with parity 4+ is only 11.489, producing a difference of 60.604 in these percentages. Table 5.4 shows that if the age structures of the women in the two groups were held constant and the age-specific percents desiring more children were allowed to vary as they did in the two parity groups, the overall percents desiring more children would be 55.849 and 18.317, giving a difference of 37.532 as the rate effect. In other words, 38.1 percent of the difference in the desires in the two parity groups is explained by the difference in their age structures. This problem will be taken up again in Example 6.3 (tables 6.5 and 6.6) to compare our results with those of Clogg and Eliason when four parity groups are treated simultaneously.

We can use Program 5.1 to obtain the results in table 5.4 if the following changes are made in the program:

1. Replace the number of age groups 13 by 5 throughout the program.
2. For the same reason, replace 14 by 6 throughout the program.
3. Replace 13F5.1 in line 4 by 5F8.0/5F8.3 .

The data file, again, should be made in four lines corresponding to the last four columns in table 5.3 in the same order, each line having five numbers with the format for each of the two pairs of lines as specified in (3) above.

Program 5.1 (One Factor +Rate)

```
      DIMENSION P(14,2),T(13,2),S(2),ET(2),ER(2)
      DO 1 J=1,2
    1 READ(5,2) (P(I,J),I=1,13),(T(I,J),I=1,13)
    2 FORMAT(13F5.1)
      DO 3 J=1,2
      P(14,J)=0.0
      DO 3 I=1,13
    3 P(14,J)=P(14,J)+P(I,J)
      DO 5 J=1,2
      ET(J)=0.0
      ER(J)=0.0
      S(J)=0.0
      DO 5 I=1,13
      ET(J)=ET(J)+P(I,J)*T(I,J)/P(14,J)
      Q=(P(I,1)/P(14,1)+P(I,2)/P(14,2))*.5
      R=(T(I,1)+T(I,2))*.5
      ER(J)=ER(J)+Q*T(I,J)
    5 S(J)=S(J)+R*P(I,J)/P(14,J)
      ETT=ET(2)-ET(1)
      ERR=ER(2)-ER(1)
      U=S(2)-S(1)
      WRITE(6,4) S(2),S(1),U,ER(2),ER(1),ERR,ET(2),ET(1),ETT
    4 FORMAT(40X,3F15.3)
      STOP
      END
****
```

Program 5.2 (Two Factors +Rate)

```
      DIMENSION P(12,3,2),T(11,2,2),R(2,2),U(2),S(2,2),ET(2),ER(2)
      DO 1 K=1,2
    1 READ(5,2) ((P(I,J,K),I=1,11),J=1,2),((T(I,J,K),I=1,11),J=1,2)
    2 FORMAT(11F7.0/11F7.0/11F7.3/11F7.3)
      DO 6 K=1,2
      DO 3 J=1,2
      P(12,J,K)=0.0
      DO 3 I=1,11
    3 P(12,J,K)=P(12,J,K)+P(I,J,K)
      DO 4 I=1,12
      P(I,3,K)=0.0
      DO 4 J=1,2
    4 P(I,3,K)=P(I,3,K)+P(I,J,K)
    6 CONTINUE
      DO 5 K=1,2
      ET(K)=0.0
      ER(K)=0.0
      DO 5 I=1,11
      DO 5 J=1,2
      ET(K)=ET(K)+P(I,J,K)*T(I,J,K)/P(12,3,K)
      Q=(P(I,J,1)/P(12,3,1)+P(I,J,2)/P(12,3,2))*.5
    5 ER(K)=ER(K)+Q*T(I,J,K)
      ETT=ET(2)-ET(1)
      ERR=ER(2)-ER(1)
      DO 7 I=1,2
      DO 7 J=1,2
    7 R(I,J)=0.0
      DO 11 II=1,2
      DO 11 JJ=1,2
      H=0.0
      DO 10 IS=1,11
      DO 10 JS=1,2
      W=1.0
      DO 15 I1=1,2
      IF(I1.EQ.1) J=JS
      IF(I1.EQ.2) J=3
      DO 8 NL=1,2
      GO TO (13,14),NL
   13 A=P(IS,J,II)/P(12,J,II)
      GO TO 8
   14 IF(I1.EQ.1) I=IS
      IF(I1.EQ.2) I=12
      A=P(I,JS,JJ)/P(I,3,JJ)
    8 W=W*A
   15 CONTINUE
   10 H=H+(T(IS,JS,1)+T(IS,JS,2))*.5*W**(1./2.)
      R(II,1)=R(II,1)+H
   11 R(JJ,2)=R(JJ,2)+H
      DO 9 J=1,2
      DO 12 I=1,2
   12 S(I,J)=R(I,J)/2.
    9 U(J)=S(2,J)-S(1,J)
      WRITE(6,16) (S(2,J),S(1,J),U(J),J=1,2),ER(2),ER(1),ERR,
     1 ET(2),ET(1),ETT
   16 FORMAT(40X,3F15.3)
      STOP
      END
****
```

Table 5.3. Population Size and Percent Desiring More Children (Rate) by Age Groups for Parity 1 and Parity 4+ Women: 1970 National Fertility Survey

Age groups	i	Parity 4+ (population 1)		Parity 1 (population 2)	
		Size N_i	Rate T_i	Size n_i	Rate t_i
20 to 24	1	27	37.037	363	90.083
25 to 29	2	152	19.079	208	76.923
30 to 34	3	224	15.179	96	56.250
35 to 39	4	239	5.021	59	20.339
40 to 44	5	211	6.161	48	10.417
All ages	i = .	853	11.489	774	72.093

Source: Clogg and Eliason (1988), table 1.

Table 5.4. Standardization and Decomposition of Percents Desiring More Children in Table 5.3

Percents desiring more children	Standardization		Decomposition	
	Parity 1 (population 2)	Parity 4+ (population 1)	Difference (effects)	Percent distribution of effects
R (rate)-standardized percents	48.619	25.547	23.072 (I=age)	38.1
I (age)-standardized percents	55.849	18.317	37.532 (R=rate)	61.9
Overall percents desiring more children	72.093	11.489	60.604 (Total effect)	100.0

Alternatively, as in the case of Example 5.1, we can also use Program 4.2 to obtain the results in table 5.4 by making the following changes in Program 4.2:

Line 1: Replace 480 by 5 and add VV(2)

Line 2: Replace 480 by 5

Line 3: Replace 8F10.7 by 5F8.0/5F8.3

Line 6: Add the following two lines for the two totals after line 6:

```
        VV(1) = 853.
        VV(2) = 774.
```

Lines 9-16: Replace by the following three lines:

```
  H = 0.0
  DO 7 K1 = 1,5
7 H = H+V(I,1,K1)*V(J,2,K1)/VV(I)
```

5.3 THE CASE OF TWO FACTORS

When there are two factors I and J, N_{ij} and T_{ij} are the number of persons and the rate for the (i,j)-category in population 1; $N_{i.}$ and $T_{i.}$ are the number of persons and the rate for the ith category of I, and $N_{.j}$ and $T_{.j}$ are the corresponding number of persons and the rate for the jth category of J. As before, $N_{..}$ and $T_{..}$ are the total number of persons and the crude rate. Analogous symbols are used for population 2 with lower-case letters n and t.

The crude rates can be expressed as

$$T_{..} = \sum_{i,j} \frac{T_{ij}N_{ij}}{N_{..}}, \quad t_{..} = \sum_{i,j} \frac{t_{ij}n_{ij}}{n_{..}}. \qquad (5.6)$$

Writing

$$\frac{N_{ij}}{N_{..}} = \left(\frac{N_{ij}}{N_{.j}} \cdot \frac{N_{i.}}{N_{..}}\right)^{\frac{1}{2}} \cdot \left(\frac{N_{ij}}{N_{i.}} \cdot \frac{N_{.j}}{N_{..}}\right)^{\frac{1}{2}} = A_{ij}B_{ij}, \tag{5.7}$$

$$\frac{n_{ij}}{n_{..}} = \left(\frac{n_{ij}}{n_{.j}} \cdot \frac{n_{i.}}{n_{..}}\right)^{\frac{1}{2}} \cdot \left(\frac{n_{ij}}{n_{i.}} \cdot \frac{n_{.j}}{n_{..}}\right)^{\frac{1}{2}} = a_{ij}b_{ij},$$

we notice that the two ratios in A_{ij} and a_{ij} represent only the I-effect, and the two ratios in B_{ij} and b_{ij} represent only the J-effect.

It follows from equations (19) and (21) in Das Gupta (1991) that

$$\begin{aligned} t_{..} - T_{..} &= \text{R-effect} + \text{I-effect} + \text{J-effect} \\ &= [R(\bar{t}) - R(\bar{T})] + [I(\bar{a}) - I(\bar{A})] + [J(\bar{b}) - J(\bar{B})], \end{aligned} \tag{5.8}$$

where

$R(\bar{T})$ = (I,J)-standardized rate in population 1

$$= \sum_{i,j} \frac{\frac{n_{ij}}{n_{..}} + \frac{N_{ij}}{N_{..}}}{2} T_{ij}, \tag{5.9}$$

$I(\bar{A})$ = (J,R)-standardized rate in population 1

$$= \sum_{i,j} \frac{t_{ij} + T_{ij}}{2} \frac{b_{ij}+B_{ij}}{2} A_{ij}, \tag{5.10}$$

$J(\bar{B})$ = (I,R)-standardized rate in population 1

$$= \sum_{i,j} \frac{t_{ij} + T_{ij}}{2} \frac{a_{ij}+A_{ij}}{2} B_{ij}, \tag{5.11}$$

and $R(\bar{t})$, $I(\bar{a})$, and $J(\bar{b})$ for population 2 have the same expressions as those in (5.9) through (5.11), respectively, with T_{ij} in (5.9) replaced by t_{ij}, A_{ij} in (5.10) replaced by a_{ij}, and B_{ij} in (5.11) replaced by b_{ij}.

We note here that $I(\bar{A})$ and $J(\bar{B})$ can also be written as

$$I(\bar{A}) = \sum_{i,j} \frac{t_{ij} + T_{ij}}{2} \text{ [Expression (2.3) with subscripts ij in each letter]}, \tag{5.12}$$

$$J(\bar{B}) = \sum_{i,j} \frac{t_{ij} + T_{ij}}{2} \text{ [Expression (2.5) with subscripts ij in each letter]}. \tag{5.13}$$

Unlike the **hierarchical** approaches by Cho and Retherford (1973) and Kim and Strobino (1984), the effects of the factors in the decomposition (5.8) remain unchanged irrespective of which one of the factors is regarded as I and which one as J. In other words, the treatment of the factors I and J is **symmetrical** in the present approach.

Example 5.3

Table 5.5 is from Liao (1989), which shows the cross-classification of the population and the death rates by age and race for the United States for the years 1970 and 1985. The standardization and decomposition of the crude death rates from these data are shown in table 5.6. The crude death rate for 1970 was .686 point higher than that for 1985. However, if only the age structures of the populations differed as they did in the two years but the race structures and the age-race-specific death rates were identical in 1970 and 1985, then the overall death rate in 1985 would be 1.522 points higher than that for 1970. The differences in the age and race structures in 1970 and 1985 dampened the difference between the crude death rates in these two years. If the rates were standardized with respect to both age (I) and race (J), the difference between the standardized rates would be as high as 2.228. Table 2 in Liao's paper showed four sets of widely different decompositions for these data using the modeling approach, each set involving an interaction term. The results from only the marginal CG method (namely, -1.57, -0.06, and 2.23 for the I, J, and R effects and 0.08 for the interaction effect) are comparable to our decomposition in table 5.6. There is a discussion in chapter 1 that it is unnecessary to complicate the model by including the interaction effects.

Table 5.5. Population (in thousands) and Death Rates (per 1,000 Population) by Age and Race: United States, 1970 and 1985

Race j	Age i	1985 (population 1) Size N_{ij}	1985 (population 1) Rate T_{ij}	1970 (population 2) Size n_{ij}	1970 (population 2) Rate t_{ij}
1	1	3,041	9.163	2,968	18.489
1	2	11,577	0.462	11,484	0.751
1	3	27,450	0.248	34,614	0.391
1	4	32,711	0.929	30,992	1.146
1	5	35,480	1.084	21,983	1.287
1	6	27,411	1.810	20,314	2.672
1	7	19,555	4.715	20,928	6.636
1	8	19,795	12.187	16,897	15.691
1	9	15,254	27.728	11,339	34.723
1	10	8,022	64.068	5,720	79.763
1	11	2,472	157.570	1,315	176.837
2	1	707	17.208	535	36.993
2	2	2,692	0.738	2,162	1.352
2	3	6,473	0.328	6,120	0.541
2	4	6,841	1.103	4,781	2.040
2	5	6,547	2.045	3,096	3.523
2	6	4,352	3.724	2,718	6.746
2	7	3,034	8.052	2,363	12.967
2	8	2,540	17.812	1,767	24.471
2	9	1,749	34.128	1,149	45.091
2	10	804	68.276	448	74.902
2	11	236	125.161	117	123.205
j = .	i = .	238,743	8.736	203,810	9.422

Source: Liao (1989), table 1. Age i = 1, 2,..., 11 correspond to less than 1, 1-4, 5-14, 15-24, ..., 75-84, 85+. Race j = 1, 2 correspond to White and non-White.

Table 5.6. Standardization and Decomposition of Crude Death Rates in Table 5.5

Death rates per 1,000 population	Standardization: 1970 (population 2)	Standardization: 1985 (population 1)	Decomposition: Difference (effects)	Decomposition: Percent distribution of effects
(J,R)-standardized rates	8.385	9.907	−1.522 (I)	−221.9
(I,R)-standardized rates	9.136	9.156	−0.020 (J)	−2.9
(I,J)-standardized rates	10.258	8.030	2.228 (R)	324.8
Crude death rates	9.422	8.736	0.686 (Total effect)	100.0

Program 5.2

The results in table 5.6 can be obtained by using Program 5.2 in which P(I,J,K)'s are N_{ij}'s and n_{ij}'s, and T(I,J,K)'s are T_{ij}'s and t_{ij}'s in table 5.5. The data file consists of eight lines corresponding to the data in the last four columns in table 5.5 in the same order—two lines of 11 numbers for each column—with the format specified in line 4 of the program. The six standardized rates in table 5.6 are given by ER(K)'s and S(I,J)'s in lines 22 and 51 of the program. The three effects in table 5.6 are denoted by ERR and U(J)'s in lines 24 and 52 of the program.

Example 5.4

Kitagawa (1955) used the data in table 5.7 to decompose the difference between the job mobility rates (i.e., mean number of jobs held) in Los Angeles and Philadelphia in terms of the effects of time spent in the labor force and migrant status. The overall job mobility rates in Los Angeles and Philadelphia were 3.145 and 2.379, respectively, producing a difference of .766. Table 5.8 decomposes this total difference into .024 as the I (time spent in the labor force)-effect, .330 as the J (migrant status)-effect, and .412 as the R (rate)-effect. Thus, the factors I and J together explain 46.2 percent of the difference between the job mobility rates in Los Angeles and Philadelphia. These results are not very different from the decomposition in table 1 in Kitagawa's paper except that she attributed 7 percent of the total difference to the interaction between I and J (which she called Joint IJ). This 7 percent is distributed equally between the I and J effects in table 5.8.

We can use Program 5.2 to obtain the results in table 5.8 if the following changes are made in the program:

1. Replace the number of age groups 11 by the number of categories 3 in the time spent in the labor force, throughout the program.
2. *For the same reason, replace* 12 *by* 4 *throughout the program.*
3. Replace the format in line 4 by 6F5.0/6F5.2 .

The data file should be made in four lines corresponding to the last four columns in table 5.7 in the same order, each line having six numbers with the format for each of the two pairs of lines as specified in (3) above.

Table 5.7. Population Size (Percents) and Job Mobility Rates (Mean Number of Jobs Held) by Migrant Status and Time Spent in the Labor Force: Philadelphia and Los Angeles, Men, 1940 to 1949

Migrant status	Time in labor force	Philadelphia (population 1)		Los Angeles (population 2)	
		Size	Rate	Size	Rate
j	i	N_{ij}	T_{ij}	n_{ij}	t_{ij}
1	1	1	2.29	6	2.89
1	2	4	3.43	17	4.07
1	3	8	3.15	24	3.79
2	1	6	2.45	5	2.92
2	2	22	3.23	13	3.49
2	3	59	1.88	35	2.20
j = .	i = .	100	2.379	100	3.145

Source: Kitagawa (1955), table 1. Time in labor force i = 1,2,3 correspond to less than 5 years, 5 but less than 9.5 years, 9.5 to 10 years. Migrant status j = 1,2 correspond to migrants, nonmigrants.

Example 5.5

Another two-factor case is presented in tables 5.9 and 5.10 to study the effects of birth weights (I) and age of mother (J) on the difference between the neonatal mortality rates for White and non-White live births in 1960. Kim and Strobino (1984) used a different set of data to study the same problem. However, as mentioned earlier in this section, they used a hierarchical approach (as opposed to the symmetrical approach presented here) in the treatment of the two factors. They decomposed the combined effect of the two factors into the effect of age of mother and the effect of birth weight **within** age of mother. The same hierarchical approach can also be used to decompose the combined effect of the two factors into the effect of birth weight and the effect of age of mother **within** birth weight. In general, these two

Table 5.8. Standardization and Decomposition of Job Mobility Rates in Table 5.7

Job mobility rates	Standardization		Decomposition	
	Los Angeles (population 2)	Philadelphia (population 1)	Difference (effects)	Percent distribution of effects
(J,R)-standardized rates	2.749	2.725	.024 (I)	3.1
(I,R)-standardized rates	2.902	2.572	.330 (J)	43.1
(I,J)-standardized rates	2.940	2.528	.412 (R)	53.8
Overall job mobility rates	3.145	2.379	.766 (Total effect)	100.0

alternative ordering of the two factors will lead to two different sets of results, whereas the symmetrical approach produces a unique set of results. We apply this approach to the data in table 5.9, and the results are shown in table 5.10. The crude neonatal mortality rate for non-Whites is 8.91 points higher than that for Whites. It is interesting to note that when the rates are standardized with respect to both age of mother and birth weight, the White rate becomes higher than the non-White rate by 1.50 points. The effects of birth weight (I) and age of mother (J) are, respectively, 10.19 and 0.22, suggesting that unfavorable distribution of birth weight for non-Whites is primarily responsible for the significant difference in the neonatal mortality rates between Whites and non-Whites, and that age of mother is only marginally important in explaining this difference.

We can use Program 5.2 to obtain the results in table 5.10 if the following changes are made in the program:

1. Replace the number of age groups 11 by the number of birth weight categories 10, throughout the program.
2. For the same reason, replace 12 by 11 throughout the program.
3. Replace the number of race categories 2 by the number of age groups of mother 7, throughout the program.
4. For the same reason, replace 3 by 8 throughout the program.
5. Replace lines 3 and 4 by the following four lines:

```
      READ (5,2) ((P(I,J,K), I=1,10),J=1,7)
    1 READ (5,17) ((T(I,J,K),I=1,10),J=1,7)
    2 FORMAT (10F8.0)
   17 FORMAT (10F8.2)
```

6. Replace 15.3 in line 55 by 15.2 .

The data file should be made in 28 lines with the data in the last four columns in table 5.9 in the same order, each column occupying seven lines of 10 numbers. The formats of the numbers should be according to the specifications in (5) above.

Two more examples of two-factor decomposition from cross-classified data are the study by Gibson (1976) of the contributions of changes in marital status and marital fertility to the decline in the U.S. fertility during 1961-1975, and the research by Hernandez (1984) on the relationship between the decline in the birth rates in the developing countries and the corresponding changes in age-sex composition and marital status composition.

5.4 THE CASE OF THREE FACTORS

Using symbols analogous to those in the preceding sections, we can write the crude rates in population 1 and population 2 as

$$T_{...} = \sum_{i,j,k} \frac{T_{ijk}N_{ijk}}{N_{...}}, \quad t_{...} = \sum_{i,j,k} \frac{t_{ijk}n_{ijk}}{n_{...}}. \tag{5.14}$$

Table 5.9. Single Live Births and Neonatal Mortality Rates (per 1,000 Live Births), by Age of Mother and Birth Weight: White and Non-White, 1960

Age of mother	Birth weight	White (population 1)		Non-White (population 2)	
		Live births	Rate	Live births	Rate
j	i	N_{ij}	T_{ij}	n_{ij}	t_{ij}
1	1	2258	899.47	1494	852.07
1	2	3065	607.50	1851	463.53
1	3	6626	232.87	3666	131.48
1	4	22769	44.36	13043	27.22
1	5	86436	8.73	39304	8.98
1	6	184474	3.74	49181	6.28
1	7	119071	3.43	18679	7.71
1	8	26892	3.53	3225	12.71
1	9	3351	6.86	637	6.28
1	10	244	20.49	62	64.52
2	1	4461	899.57	1732	886.84
2	2	5383	576.63	1948	448.67
2	3	11793	228.19	4096	149.90
2	4	47905	46.30	15477	29.20
2	5	212061	9.15	53818	7.66
2	6	481385	3.93	79255	5.61
2	7	337526	2.88	36723	4.96
2	8	85994	3.37	7701	7.66
2	9	12802	5.62	1397	15.03
2	10	1180	13.56	157	31.85
3	1	3500	944.86	1215	889.71
3	2	3674	553.35	1302	413.21
3	3	8033	217.23	2551	132.50
3	4	34133	47.46	9778	32.62
3	5	152928	10.09	34454	8.65
3	6	355446	4.16	56245	6.26
3	7	271301	2.99	31039	5.12
3	8	78027	3.11	7739	9.05
3	9	14134	5.80	1648	10.92
3	10	1728	15.05	223	35.87
4	1	2493	911.75	825	876.36
4	2	2444	545.42	826	406.78
4	3	5586	200.32	1795	132.59
4	4	22080	49.68	6431	35.45
4	5	91004	11.65	20650	11.23
4	6	209931	4.91	35030	6.88
4	7	171323	3.49	21873	8.14
4	8	55454	3.50	6333	10.26
4	9	11603	6.98	1633	10.41
4	10	1606	21.79	312	41.67
5	1	1293	936.58	368	855.98
5	2	1469	492.85	410	431.71
5	3	3360	192.26	952	138.66
5	4	12309	55.89	3327	38.17
5	5	45476	12.69	10399	14.23
5	6	104558	5.83	17520	8.22
5	7	90093	4.14	12045	9.88
5	8	31815	4.71	3849	13.51
5	9	7295	6.72	1242	16.91
5	10	1194	22.61	222	40.54
6	1	316	936.71	71	915.49
6	2	423	468.09	100	400.00
6	3	959	212.72	252	146.83
6	4	3539	64.42	878	56.95
6	5	11570	17.46	2656	19.20
6	6	25515	9.05	4294	12.34
6	7	22477	6.32	3253	10.45
6	8	8829	6.68	1164	13.75
6	9	2183	9.16	419	16.71
6	10	419	16.71	87	45.98

Table 5.9. Single Live Births and Neonatal Mortality Rates (per 1,000 Live Births), by Age of Mother and Birth Weight: White and Non-White, 1960—Continued

Age of mother	Birth weight	White (population 1)		Non-White (population 2)	
j	i	Live births N_{ij}	Rate T_{ij}	Live births n_{ij}	Rate t_{ij}
7	1	24	708.33	4	1000.00
7	2	30	533.33	12	333.33
7	3	69	246.38	20	.00
7	4	221	117.65	81	74.07
7	5	590	32.20	181	27.62
7	6	1414	12.02	293	10.24
7	7	1204	9.14	226	13.27
7	8	477	6.29	90	22.22
7	9	113	.00	35	28.57
7	10	24	.00	6	.00
$j = .$	$i = .$	3,531,362	15.32	639,804	24.23

Source: National Center for Health Statistics (1972), tables 5 and 6. Birth weight $i = 1, 2, 3, \ldots, 9, 10$ correspond to (in grams) 1000 and less, 1001-1500, 1501-2000, ..., 4501-5000, 5001 and above. Age of mother $j = 1, 2, \ldots, 6, 7$ correspond to under 20, 20-24, ..., 40-44, 45 and over.

Table 5.10. Standardization and Decomposition of Neonatal Mortality Rates in Table 5.9

Neonatal mortality rates	Standardization		Decomposition	
	Non-White (population 2)	White (population 1)	Difference (effects)	Percent distribution of effects
(J,R)-standardized rates	25.56	15.37	10.19 (I)	114.4
(I,R)-standardized rates	20.57	20.35	0.22 (J)	2.5
(I,J)-standardized rates	19.76	21.26	−1.50 (R)	−16.9
Overall neonatal mortality rates	24.23	15.32	8.91 (Total effect)	100.0

As shown in equation (22) in Das Gupta (1991), we express the cell proportions as

$$\frac{N_{ijk}}{N_{...}} = A_{ijk} B_{ijk} C_{ijk}, \tag{5.15}$$

where

$$A_{ijk} = \left(\frac{N_{ijk}}{N_{.jk}}\right)^{\frac{1}{3}} \cdot \left(\frac{N_{ij.}}{N_{.j.}} \cdot \frac{N_{i.k}}{N_{..k}}\right)^{\frac{1}{6}} \cdot \left(\frac{N_{i..}}{N_{...}}\right)^{\frac{1}{3}},$$

$$B_{ijk} = \left(\frac{N_{ijk}}{N_{i.k}}\right)^{\frac{1}{3}} \cdot \left(\frac{N_{ij.}}{N_{i..}} \cdot \frac{N_{.jk}}{N_{..k}}\right)^{\frac{1}{6}} \cdot \left(\frac{N_{.j.}}{N_{...}}\right)^{\frac{1}{3}}, \tag{5.16}$$

$$C_{ijk} = \left(\frac{N_{ijk}}{N_{ij.}}\right)^{\frac{1}{3}} \cdot \left(\frac{N_{i.k}}{N_{i..}} \cdot \frac{N_{.jk}}{N_{.j.}}\right)^{\frac{1}{6}} \cdot \left(\frac{N_{..k}}{N_{...}}\right)^{\frac{1}{3}}.$$

Equations (5.16) are derived in section A.3 in appendix A. $n_{ijk}/n_{...}$ is similarly expressed in terms of lower-case letters a, b, c, and n.

As in (5.8) through (5.13), we can write

$$\begin{aligned} t_{...} - T_{...} &= \text{R-effect} + \text{I-effect} + \text{J-effect} + \text{K-effect} \\ &= [R(\bar{t}) - R(\bar{T})] + [I(\bar{a}) - I(\bar{A})] + [J(\bar{b}) - J(\bar{B})] + [K(\bar{c}) - K(\bar{C})], \end{aligned} \tag{5.17}$$

where

$R(\bar{T}) = (I,J,K)$-standardized rate in population 1

$$= \sum_{i,j,k} \frac{\frac{n_{ijk}}{n_{...}} + \frac{N_{ijk}}{N_{...}}}{2} T_{ijk}, \tag{5.18}$$

$I(\bar{A}) = (J,K,R)$-standardized rate in population 1

$$= \sum_{i,j,k} \frac{t_{ijk}+T_{ijk}}{2} \left[\frac{b_{ijk}c_{ijk}+B_{ijk}C_{ijk}}{3} + \frac{b_{ijk}C_{ijk}+B_{ijk}c_{ijk}}{6}\right] A_{ijk}$$

$$= \sum_{i,j,k} \frac{t_{ijk}+T_{ijk}}{2} \text{ [Expression (2.13) with subscripts ijk in each letter]}, \tag{5.19}$$

$J(\bar{B}) = (I,K,R)$-standardized rate in population 1

$$= \sum_{i,j,k} \frac{t_{ijk}+T_{ijk}}{2} \left[\frac{a_{ijk}c_{ijk}+A_{ijk}C_{ijk}}{3} + \frac{a_{ijk}C_{ijk}+A_{ijk}c_{ijk}}{6}\right] B_{ijk}$$

$$= \sum_{i,j,k} \frac{t_{ijk}+T_{ijk}}{2} \text{ [Expression (2.15) with subscripts ijk in each letter]}, \tag{5.20}$$

$K(\bar{C}) = (I,J,R)$-standardized rate in population 1

$$= \sum_{i,j,k} \frac{t_{ijk}+T_{ijk}}{2} \left[\frac{a_{ijk}b_{ijk}+A_{ijk}B_{ijk}}{3} + \frac{a_{ijk}B_{ijk}+A_{ijk}b_{ijk}}{6}\right] C_{ijk}$$

$$= \sum_{i,j,k} \frac{t_{ijk}+T_{ijk}}{2} \text{ [Expression (2.17) with subscripts ijk in each letter]}. \tag{5.21}$$

$R(\bar{t})$, $I(\bar{a})$, $J(\bar{b})$, and $K(\bar{c})$ for population 2 have the same expressions as those in (5.18) through (5.21), respectively, with T_{ijk} in (5.18) replaced by t_{ijk}, A_{ijk} in (5.19) replaced by a_{ijk}, B_{ijk} in (5.20) replaced by b_{ijk}, and C_{ijk} in (5.21) replaced by c_{ijk}.

Example 5.6

Table 5.11 shows a three-factor cross-classification of the population and the household headship rates by age (I), marital status (J), and sex (K) for the United States, 1970 and 1980. Sweet (1984) considered similar data to study the components of change in the **number** of households during the decade. Since our present example deals with the change in the household **headship rate**,[1] the two sets of results are not comparable. The overall headship rate increased by 4.39 points during 1970-1980. However, as shown in table 5.12, if the age-marital status-sex distributions were identical in the two years, this increase would have been 3.81. The headship rate in 1980 would be only .49 point[2] higher than that in 1970 if the age structures differed as they did in 1970 and 1980 but if everything else (namely, marital status, sex, and the cell-specific headship rates) were identical in the two years. The differences in age, marital status, and sex structures explain only 13.2 percent of the difference in the household headship rates in 1970 and 1980.

[1]Beginning with the 1980 CPS, the Bureau of the Census discontinued the use of the term "head of household" and started using the term "householder," instead.
[2]Not significant at 90-percent level.

Table 5.11. Population and Household Headship Rates per 100 Persons, by Age, Sex, and Marital Status: United States, 1970 and 1980

(Population in thousands)

Sex k	Marital status j	Age i	1970 (population 1) Size N_{ijk}	1970 (population 1) Rate T_{ijk}	1980 (population 2) Size n_{ijk}	1980 (population 2) Rate t_{ijk}
1	1	1	3239	92.50	2950	90.07
1	1	2	9710	98.47	11374	95.23
1	1	3	9661	99.31	9943	96.40
1	1	4	9501	99.32	8979	96.46
1	1	5	7225	99.39	8130	95.76
1	1	6	3979	99.04	5200	95.79
1	1	7	1742	97.70	2190	93.52
1	2	1	206	18.45	241	32.37
1	2	2	351	30.48	586	59.56
1	2	3	377	45.09	415	70.84
1	2	4	294	45.58	368	65.49
1	2	5	233	68.24	284	63.03
1	2	6	159	50.94	146	76.03
1	2	7	122	35.25	54	83.33
1	3	1	1	100.00	2	100.00
1	3	2	13	61.54	19	68.42
1	3	3	66	77.27	45	80.00
1	3	4	183	68.85	176	68.75
1	3	5	336	73.51	397	73.55
1	3	6	588	67.69	557	83.30
1	3	7	922	60.09	776	80.03
1	4	1	81	22.22	160	47.50
1	4	2	313	45.69	1130	63.36
1	4	3	324	59.26	989	72.50
1	4	4	403	65.76	740	69.05
1	4	5	257	72.37	495	73.74
1	4	6	155	61.94	290	67.93
1	4	7	44	68.18	71	77.46
1	5	1	14959	2.89	16607	9.32
1	5	2	1846	27.63	4238	43.13
1	5	3	812	34.48	904	49.89
1	5	4	855	37.08	699	51.36
1	5	5	659	43.40	565	53.81
1	5	6	452	44.91	357	63.87
1	5	7	201	52.74	142	69.72
2	1	1	5605	.00	5058	3.44
2	1	2	10290	.00	12303	3.46
2	1	3	9756	.00	9939	2.98
2	1	4	9397	.00	8749	3.07
2	1	5	6181	.00	7404	3.32
2	1	6	2952	.00	4114	4.08
2	1	7	872	.00	1197	4.26
2	2	1	613	38.50	510	42.16
2	2	2	611	71.69	976	78.07
2	2	3	483	80.75	673	84.84
2	2	4	498	78.92	473	86.05
2	2	5	352	71.31	309	76.38
2	2	6	110	67.27	168	86.31
2	2	7	90	33.33	67	79.10
2	3	1	29	58.62	26	84.62
2	3	2	66	86.36	135	88.15
2	3	3	295	91.19	292	89.73
2	3	4	983	85.76	821	90.74
2	3	5	2071	82.04	2082	88.18
2	3	6	2948	78.22	3444	88.39
2	3	7	3248	62.75	3677	78.38

Table 5.11. Population and Household Headship Rates per 100 Persons, by Age, Sex, and Marital Status: United States, 1970 and 1980—Continued

(Population in thousands)

Sex k	Marital status j	Age i	1970 (population 1) Size N_{ijk}	1970 (population 1) Rate T_{ijk}	1980 (population 2) Size n_{ijk}	1980 (population 2) Rate t_{ijk}
2	4	1	207	39.13	401	52.37
2	4	2	563	75.67	1746	77.61
2	4	3	633	81.83	1411	89.23
2	4	4	591	81.39	1074	88.55
2	4	5	440	83.86	735	84.90
2	4	6	201	69.65	342	81.87
2	4	7	58	68.97	126	82.54
2	5	1	13222	3.68	14360	9.85
2	5	2	1098	36.89	2757	54.48
2	5	3	614	36.32	727	59.42
2	5	4	594	40.74	552	56.88
2	5	5	659	55.08	504	60.91
2	5	6	530	57.36	480	66.25
2	5	7	341	48.97	344	74.13
k = .	j = .	i = .	147,470	42.64	168,195	47.03

Source: U.S. Bureau of the Census (1971, table 6; 1981, table 6). Age i = 1, 2, ..., 7 correspond to 15-24 (14-24 for 1970), 25-34, ..., 75+. Marital status j = 1, 2, ..., 5 correspond to married (spouse present), married (spouse absent), widowed, divorced, single. Sex k= 1, 2 correspond to male and female. A married woman (husband present) could not be the head in 1970.

Table 5.12. Standardization and Decomposition of Household Headship Rates in Table 5.11

Household headship rates	Standardization: 1980 (population 2)	Standardization: 1970 (population 1)	Decomposition: Difference (effects)	Decomposition: Percent distribution of effects
(J,K,R)-standardized rates	44.93	44.44	.49 (I)	11.2
(I,K,R)-standardized rates	44.78	44.60	.18 (J)	4.1
(I,J,R)-standardized rates	44.66	44.75	−.09 (K)	−2.1
(I,J,K)-standardized rates	46.64	42.83	*3.81 (R)	86.8
Overall headship rates	47.03	42.64	*4.39 (Total effect)	100.0

*Significant at 90-percent level.

Program 5.3

The results in table 5.12 can be obtained by using Program 5.3 in which P(I,J,K,L)'s are N's and n's, and T(I,J,K,L)'s are T's and t's in table 5.11. The data file consists of 40 lines corresponding to the data in the last four columns in table 5.11 in the same order—10 lines of seven numbers for each column—with the format specified in lines 6 and 7 of the program. The eight standardized rates in table 5.12 are given by ER(L)'s and S(I,J)'s in lines 33 and 79 of the program. The four effects in table 5.12 are denoted by ERR and U(J)'s in lines 35 and 80 of the program.

The decomposition of the difference between the AIDS rates in racial groups into the effects of age, sex, and region by del Pinal (1989) is another example of a three-factor case. Also, the work by Spencer (1980) explaining the racial and ethnic differences in American fertility in terms of childlessness, nonmarriage, and age can be looked upon as a decomposition problem dealing with three factors.

Program 5.3 (Three Factors +Rate)

```
      DIMENSION P(8,6,3,2),T(7,5,2,2),R(2,3,2),U(3),S(2,3),ET(2),ER(2)
      DOUBLE PRECISION P,R,U,S,ET,ER,Q,H,A,W1,W2
      DO 1 L=1,2
      READ(5,2)  (((P(I,J,K,L),I=1,7),J=1,5),K=1,2)
    1 READ(5,8)  (((T(I,J,K,L),I=1,7),J=1,5),K=1,2)
    2 FORMAT(7F10.0)
    8 FORMAT(7F10.2)
      DO 7 L=1,2
      DO 3 J=1,5
      DO 3 K=1,2
      P(8,J,K,L)=0.0
      DO 3 I=1,7
    3 P(8,J,K,L)=P(8,J,K,L)+P(I,J,K,L)
      DO 4 I=1,8
      DO 4 K=1,2
      P(I,6,K,L)=0.0
      DO 4 J=1,5
    4 P(I,6,K,L)=P(I,6,K,L)+P(I,J,K,L)
      DO 5 I=1,8
      DO 5 J=1,6
      P(I,J,3,L)=0.0
      DO 5 K=1,2
    5 P(I,J,3,L)=P(I,J,3,L)+P(I,J,K,L)
    7 CONTINUE
      DO 6 L=1,2
      ET(L)=0.0
      ER(L)=0.0
      DO 6 I=1,7
      DO 6 J=1,5
      DO 6 K=1,2
      ET(L)=ET(L)+P(I,J,K,L)*T(I,J,K,L)/P(8,6,3,L)
      Q=(P(I,J,K,1)/P(8,6,3,1)+P(I,J,K,2)/P(8,6,3,2))*.5
    6 ER(L)=ER(L)+Q*T(I,J,K,L)
      ETT=ET(2)-ET(1)
      ERR=ER(2)-ER(1)
      DO 9 I=1,2
      DO 9 J=1,3
      DO 9 K=1,2
    9 R(I,J,K)=0.0
      DO 12 II=1,2
      DO 12 JJ=1,2
      DO 12 KK=1,2
      H=0.0
      DO 11 IS=1,7
      DO 11 JS=1,5
      DO 11 KS=1,2
      W1=1.0
      W2=1.0
      DO 16 I1=1,2
      DO 16 I2=1,2
      I3=I1+I2
      IF(I1.EQ.1) J=JS
      IF(I1.EQ.2) J=6
      IF(I2.EQ.1) K=KS
      IF(I2.EQ.2) K=3
      DO 10 NL=1,3
      GO TO (13,14,15),NL
   13 A=P(IS,J,K,II)/P(8,J,K,II)
      GO TO 17
   14 IF(I1.EQ.1) I=IS
      IF(I1.EQ.2) I=8
      A=P(I,JS,K,JJ)/P(I,6,K,JJ)
      GO TO 17
   15 IF(I2.EQ.1) J=JS
      IF(I2.EQ.2) J=6
      A=P(I,J,KS,KK)/P(I,J,3,KK)
   17 IF(I3.EQ.2.OR.I3.EQ.4) W1=W1*A
   10 IF(I3.EQ.3            ) W2=W2*A
   16 CONTINUE
   11 H=H+(T(IS,JS,KS,1)+T(IS,JS,KS,2))*.5*W1**(1./3.)*W2**(1./6.)
      DO 12 NN=1,2
      N1=NN+1
      N2=6-N1
      IF(JJ+KK.EQ.N1.OR.JJ+KK.EQ.N2) R(II,1,NN)=R(II,1,NN)+H
      IF(II+KK.EQ.N1.OR.II+KK.EQ.N2) R(JJ,2,NN)=R(JJ,2,NN)+H
   12 IF(II+JJ.EQ.N1.OR.II+JJ.EQ.N2) R(KK,3,NN)=R(KK,3,NN)+H
      DO 18 J=1,3
      DO 19 I=1,2
   19 S(I,J)=R(I,J,1)/3.+R(I,J,2)/6.
   18 U(J)=S(2,J)-S(1,J)
      WRITE(6,20) (S(2,J),S(1,J),U(J),J=1,3),ER(2),ER(1),ERR,
     1 ET(2),ET(1),ETT
   20 FORMAT(40X,3F15.2)
      STOP
      END
****
```

5.5 THE CASE OF FOUR FACTORS

We express the crude rates $T_{....}$ and $t_{....}$ in population 1 and population 2 in terms of similar notation and also express the cell proportions in population 1 as

$$\frac{N_{ijkl}}{N_{....}} = A_{ijkl}B_{ijkl}C_{ijkl}D_{ijkl}\ , \tag{5.22}$$

where (Das Gupta, 1991, equation 23)

$$A_{ijkl} = \left(\frac{N_{ijkl}}{N_{.jkl}}\right)^{\frac{1}{4}} \cdot \left(\frac{N_{ijk.}}{N_{.jk.}} \cdot \frac{N_{ij.l}}{N_{.j.l}} \cdot \frac{N_{i.kl}}{N_{..kl}}\right)^{\frac{1}{12}} \cdot \left(\frac{N_{i..l}}{N_{...l}} \cdot \frac{N_{i.k.}}{N_{..k.}} \cdot \frac{N_{ij..}}{N_{.j..}}\right)^{\frac{1}{12}} \cdot \left(\frac{N_{i...}}{N_{....}}\right)^{\frac{1}{4}}. \tag{5.23}$$

B_{ijkl}, C_{ijkl}, and D_{ijkl} are obtained from (5.23) by interchanging, respectively, i and j, i and k, and i and l. For example, $N_{.jkl}$ in (5.23) changes to $N_{i.kl}$ in the expression for B_{ijkl}. The ratio $n_{ijkl}/n_{....}$ is similarly expressed by using lower-case letters a, b, c, d, and n.

As in (5.17) through (5.21), the difference $t_{....} - T_{....}$ can be expressed as the sum of five effects: R-effect, I-effect, J-effect, K-effect, and L-effect. Each effect, again, is the difference between two standardized rates, which are given by

$R(\bar{T})$ = (I,J,K,L)-standardized rate in population 1

$$= \sum_{i,j,k,l} \frac{\dfrac{n_{ijkl}}{n_{....}} + \dfrac{N_{ijkl}}{N_{....}}}{2} T_{ijkl}\ , \tag{5.24}$$

$I(\bar{A})$ = (J,K,L,R)-standardized rate in population 1

$$= \sum_{i,j,k,l} \frac{t_{ijkl} + T_{ijkl}}{2} \text{ [Expression (2.26), i.e., (2.29) x A with subscripts ijkl in each letter]}. \tag{5.25}$$

The standardized rates $R(\bar{t})$ and $I(\bar{a})$ for population 2 are obtained, respectively, from (5.24) and (5.25) by replacing T_{ijkl} in (5.24) by t_{ijkl} and A_{ijkl} in (5.25) by a_{ijkl}. Other standardized rates $J(\bar{B})$, $J(\bar{b})$, $K(\bar{C})$, $K(\bar{c})$, $L(\bar{D})$, and $L(\bar{d})$ are obtained from (5.25) by interchanging the letters.

Example 5.7

Table 5.13 presents the data for the population and the mobility rates cross-classified by four factors: education, residence, age, and sex, for the United States, 1975-1976 and 1986-1987. Wilson (1988) studied similar data for the period 1935-1980 for decomposing the difference in the mobility rates by age and education. The mobility rate increased from 17.790 in 1975-1976 to 18.136 in 1986-1987, producing a difference of .346[3] for the 11-year period. As table 5.14 shows, this difference would have been .591 had the distributions of population by education, residence, age, and sex been identical in the two years. On the other hand, the age effect is -.575, which means that if the age structures differed as they did in the two years but all other factors and the cell-specific mobility rates were identical, then the overall mobility rate in 1975-1976 would be .575 point higher than that in 1986-1987. The factor sex appears to have played a negligible role in explaining the difference between the mobility rates in the two years.

Program 5.4

The results in table 5.14 can be obtained by using Program 5.4 in which P(I,J,K,L,M)'s are N's and n's, and T(I,J,K,L,M)'s are T's and t's in table 5.13. The data file consists of 96 lines corresponding to the data in the last four columns in table 5.13 in the same order—24 lines of six numbers for each column— with the format specified in lines 7 and 8 of the program. The 10 standardized rates in table 5.14 are given by ER(M)'s and S(I,J)'s in lines 44 and 101 of the program. The five effects in table 5.14 are denoted by ERR and U(J)'s in lines 46 and 102 of the program.

[3]Not significant at 90-percent level.

Table 5.13. Population and Mobility Rates per 100 Persons, by Age, Sex, Years of School Completed, and Residence: United States, 1975-1976 and 1986-1987

(Population in thousands)

Sex l	Age k	Residence j	Education i	1975-1976 (Popululation 1) Size N_{ijkl}	1975-1976 (Popululation 1) Rate T_{ijkl}	1986-1987 (Population 2) Size n_{ijkl}	1986-1987 (Population 2) Rate t_{ijkl}
1	1	1	1	269	40.149	460	41.957
1	1	1	2	1664	30.108	1957	27.951
1	1	1	3	3838	32.387	4226	30.549
1	1	1	4	2495	29.780	2872	27.751
1	1	1	5	630	40.952	704	35.938
1	1	1	6	157	38.854	120	40.833
1	1	2	1	202	37.129	108	29.630
1	1	2	2	902	30.266	688	24.273
1	1	2	3	1911	33.072	1221	26.863
1	1	2	4	865	31.445	574	24.216
1	1	2	5	181	41.989	79	45.570
1	1	2	6	43	53.488	20	35.000
1	2	1	1	239	33.473	352	37.216
1	2	1	2	481	36.798	800	38.125
1	2	1	3	1952	32.941	3448	30.365
1	2	1	4	1441	38.723	1849	33.207
1	2	1	5	1083	38.135	1442	41.609
1	2	1	6	670	38.209	688	40.407
1	2	2	1	173	34.682	105	37.143
1	2	2	2	292	36.986	293	36.519
1	2	2	3	1088	31.526	1091	26.673
1	2	2	4	469	33.689	367	30.518
1	2	2	5	381	39.370	177	35.028
1	2	2	6	195	40.513	82	52.439
1	3	1	1	230	22.609	368	34.511
1	3	1	2	513	24.366	692	28.468
1	3	1	3	1745	22.751	2990	23.177
1	3	1	4	912	21.272	1804	24.279
1	3	1	5	684	23.538	1528	27.029
1	3	1	6	689	24.238	973	24.460
1	3	2	1	216	20.833	95	26.316
1	3	2	2	280	24.643	245	30.612
1	3	2	3	887	20.857	1014	17.456
1	3	2	4	260	19.231	364	17.582
1	3	2	5	188	26.064	232	21.552
1	3	2	6	196	32.143	142	26.056
1	4	1	1	770	20.000	702	21.937
1	4	1	2	1024	15.527	930	20.215
1	4	1	3	2704	14.090	4217	17.240
1	4	1	4	1233	13.706	2771	19.271
1	4	1	5	983	16.887	2245	20.312
1	4	1	6	956	16.736	2186	19.350
1	4	2	1	683	17.570	276	17.029
1	4	2	2	530	23.396	393	10.941
1	4	2	3	1362	9.692	1459	14.531
1	4	2	4	390	17.179	651	17.512
1	4	2	5	222	18.018	372	19.624
1	4	2	6	250	16.800	327	19.572
1	5	1	1	2830	9.187	2107	12.150
1	5	1	2	2325	7.097	2090	9.713
1	5	1	3	4746	6.321	5591	8.871
1	5	1	4	1781	11.005	2506	10.455
1	5	1	5	1347	7.869	1949	9.800
1	5	1	6	1080	9.722	2148	9.264
1	5	2	1	2201	8.950	1072	9.795
1	5	2	2	1149	9.661	754	7.294
1	5	2	3	2000	7.450	1986	8.006
1	5	2	4	548	9.854	542	9.041
1	5	2	5	324	11.420	312	11.538
1	5	2	6	284	11.268	372	8.065

Table 5.13. Population and Mobility Rates per 100 Persons, by Age, Sex, Years of School Completed, and Residence: United States, 1975-1976 and 1986-1987—Continued

(Population in thousands)

Sex l	Age k	Residence j	Education i	1975-1976 (Popululation 1) Size N_{ijkl}	1975-1976 (Popululation 1) Rate T_{ijkl}	1986-1987 (Population 2) Size n_{ijkl}	1986-1987 (Population 2) Rate t_{ijkl}
1	6	1	1	2582	5.190	2538	4.925
1	6	1	2	792	4.419	1256	4.618
1	6	1	3	1018	4.715	2439	5.371
1	6	1	4	471	8.917	981	6.728
1	6	1	5	384	7.031	669	6.577
1	6	1	6	259	5.792	545	7.523
1	6	2	1	1908	5.398	1439	5.003
1	6	2	2	538	5.948	454	5.727
1	6	2	3	547	6.581	763	3.539
1	6	2	4	194	3.093	227	4.405
1	6	2	5	131	6.870	118	.847
1	6	2	6	89	6.742	149	5.369
2	1	1	1	290	41.724	326	41.718
2	1	1	2	1567	37.205	1647	35.762
2	1	1	3	4452	36.568	4618	35.167
2	1	1	4	2398	33.736	3147	31.045
2	1	1	5	676	50.148	833	49.220
2	1	1	6	102	52.941	110	40.909
2	1	2	1	221	37.557	99	36.364
2	1	2	2	814	41.523	562	33.630
2	1	2	3	2184	37.775	1286	32.271
2	1	2	4	842	28.147	694	28.242
2	1	2	5	210	52.381	104	42.308
2	1	2	6	24	33.333	8	87.500
2	2	1	1	277	27.798	351	36.182
2	2	1	2	640	36.406	792	35.732
2	2	1	3	2697	28.068	3540	29.096
2	2	1	4	1259	33.519	2021	30.233
2	2	1	5	916	35.590	1560	36.795
2	2	1	6	435	35.632	533	38.086
2	2	2	1	156	26.282	88	39.773
2	2	2	2	375	36.267	254	40.945
2	2	2	3	1251	26.938	1098	23.133
2	2	2	4	340	31.765	420	30.000
2	2	2	5	337	34.421	232	34.914
2	2	2	6	71	46.479	55	30.909
2	3	1	1	282	26.950	322	26.398
2	3	1	2	650	21.538	689	24.383
2	3	1	3	2274	17.942	3306	21.385
2	3	1	4	864	16.088	1901	23.567
2	3	1	5	638	18.495	1478	23.816
2	3	1	6	318	22.013	748	26.337
2	3	2	1	181	29.282	69	27.536
2	3	2	2	389	20.566	230	33.478
2	3	2	3	1026	17.057	1051	18.363
2	3	2	4	268	14.925	432	17.130
2	3	2	5	159	23.899	213	16.432
2	3	2	6	80	30.000	113	29.204
2	4	1	1	737	19.674	722	23.130
2	4	1	2	1288	14.286	1073	20.503
2	4	1	3	3792	10.443	5618	13.403
2	4	1	4	1195	12.050	2851	16.485
2	4	1	5	653	8.423	1744	15.310
2	4	1	6	417	11.751	1438	16.759
2	4	2	1	514	16.732	226	15.929
2	4	2	2	740	16.081	422	16.825
2	4	2	3	1666	11.465	1847	13.481
2	4	2	4	393	9.669	562	14.235
2	4	2	5	202	15.842	342	16.082
2	4	2	6	115	13.043	256	21.094

Table 5.13. Population and Mobility Rates per 100 Persons, by Age, Sex, Years of School Completed, and Residence: United States, 1975-1976 and 1986-1987—Continued

(Population in thousands)

Sex	Age	Residence	Education	1975-1976 (Popululation 1)		1986-1987 (Population 2)	
l	k	j	i	Size N_{ijkl}	Rate T_{ijkl}	Size n_{ijkl}	Rate t_{ijkl}
2.	5	1	1	2827	8.100	2043	10.328
2.	5	1	2	2758	8.412	2399	10.880
2.	5	1	3	6692	7.038	8202	8.400
2.	5	1	4	1731	8.319	2791	8.886
2.	5	1	5	888	6.982	1431	9.085
2.	5	1	6	514	6.420	1073	11.184
2.	5	2	1	1930	8.031	860	9.651
2.	5	2	2	1355	8.487	886	7.336
2.	5	2	3	2783	7.905	2603	7.030
2.	5	2	4	658	6.079	722	7.202
2.	5	2	5	299	8.696	273	6.960
2.	5	2	6	169	10.059	191	5.759
2.	6	1	1	3515	5.434	3606	5.435
2.	6	1	2	1385	4.332	2034	6.735
2.	6	1	3	2037	5.646	4188	5.301
2.	6	1	4	677	8.272	1327	6.179
2.	6	1	5	399	7.519	667	8.096
2.	6	1	6	199	8.040	365	5.479
2.	6	2	1	2259	6.153	1603	5.490
2.	6	2	2	703	5.832	729	6.996
2.	6	2	3	891	6.173	1217	4.437
2.	6	2	4	397	4.786	391	3.325
2.	6	2	5	180	6.111	185	2.703
2.	6	2	6	108	1.852	86	4.651
l = .	k = .	j = .	i = .	145,785	17.790	175,609	18.136

Source: U.S. Bureau of the Census (1977, table 19; 1989, table 22). Education i = 1, 2, ..., 6 correspond to elementary (0-8), high school (1-3, 4), college (1-3, 4, 5+). Residence j = 1, 2 correspond to MSA's, outside MSA's. Age k = 1, 2, ..., 6 correspond to 18-24, 25-29, 30-34, 35-44, 45-64, 65+. Sex l = 1, 2 correspond to male and female.

Table 5.14. Standardization and Decomposition of Mobility Rates in Table 5.13

Mobility rates	Standardization		Decomposition	
	1986-1987 (population 2)	1975-1976 (population 1)	Difference (effects)	Percent distribution of effects
(J,K,L,R)-standardized rates	17.928	17.724	.204 (I)	59.0
(I,K,L,R)-standardized rates	17.894	17.766	.128 (J)	37.0
(I,J,L,R)-standardized rates	17.537	18.112	*−.575 (K)	−166.2
(I,J,K,R)-standardized rates	17.832	17.834	−.002 (L)	−0.6
(I,J,K,L)-standardized rates	18.163	17.572	*.591 (R)	170.8
Overall mobility rates	18.136	17.790	.346 (Total effect)	100.0

*Significant at 90-percent level.

Technically, the four-factor decomposition problem in Example 5.7 is not different from the decomposition by Ruggles (1988) of the changes in unrelated individuals into the effects of changes in four factors, namely, age, sex and marital status, occupation, and mobility, besides the rate effect. A similar four-factor decomposition was also performed by Bachu (1981) in her study of the effects of age, age at marriage, education, and religion on the difference between the rural and urban fertility rates in India based on the 1971 census.

Program 5.4 (Four Factors +Rate)

```
      DIMENSION P(7,3,7,3,2),T(6,2,6,2,2),R(2,4,2),U(4),S(2,4),ET(2),
     1 ER(2)
      DOUBLE PRECISION P,R,U,S,ET,ER,Q,H,A,W1,W2
      DO 1 M=1,2
      READ(5,2)  ((((P(I,J,K,L,M),I=1,6),J=1,2),K=1,6),L=1,2)
    1 READ(5,9)  ((((T(I,J,K,L,M),I=1,6),J=1,2),K=1,6),L=1,2)
    2 FORMAT(6F10.0)
    9 FORMAT(6F10.3)
      DO 8 M=1,2
      DO 3 J=1,2
      DO 3 K=1,6
      DO 3 L=1,2
      P(7,J,K,L,M)=0.0
      DO 3 I=1,6
    3 P(7,J,K,L,M)=P(7,J,K,L,M)+P(I,J,K,L,M)
      DO 4 I=1,7
      DO 4 K=1,6
      DO 4 L=1,2
      P(I,3,K,L,M)=0.0
      DO 4 J=1,2
    4 P(I,3,K,L,M)=P(I,3,K,L,M)+P(I,J,K,L,M)
      DO 5 I=1,7
      DO 5 J=1,3
      DO 5 L=1,2
      P(I,J,7,L,M)=0.0
      DO 5 K=1,6
    5 P(I,J,7,L,M)=P(I,J,7,L,M)+P(I,J,K,L,M)
      DO 6 I=1,7
      DO 6 J=1,3
      DO 6 K=1,7
      P(I,J,K,3,M)=0.0
      DO 6 L=1,2
    6 P(I,J,K,3,M)=P(I,J,K,3,M)+P(I,J,K,L,M)
    8 CONTINUE
      DO 7 M=1,2
      ET(M)=0.0
      ER(M)=0.0
      DO 7 I=1,6
      DO 7 J=1,2
      DO 7 K=1,6
      DO 7 L=1,2
      ET(M)=ET(M)+P(I,J,K,L,M)*T(I,J,K,L,M)/P(7,3,7,3,M)
      Q=(P(I,J,K,L,1)/P(7,3,7,3,1)+P(I,J,K,L,2)/P(7,3,7,3,2))*.5
    7 ER(M)=ER(M)+Q*T(I,J,K,L,M)
      ETT=ET(2)-ET(1)
      ERR=ER(2)-ER(1)
      DO 10 I=1,2
      DO 10 J=1,4
      DO 10 K=1,2
   10 R(I,J,K)=0.0
      DO 11 II=1,2
      DO 11 JJ=1,2
      DO 11 KK=1,2
      DO 11 LL=1,2
      H=0.0
      DO 12 IS=1,6
      DO 12 JS=1,2
      DO 12 KS=1,6
      DO 12 LS=1,2
      W1=1.0
      W2=1.0
      DO 13 I1=1,2
      DO 13 I2=1,2
      DO 13 I3=1,2
      I4=I1+I2+I3
      IF(I1.EQ.1) J=JS
      IF(I1.EQ.2) J=3
      IF(I2.EQ.1) K=KS
      IF(I2.EQ.2) K=7
      IF(I3.EQ.1) L=LS
      IF(I3.EQ.2) L=3
      DO 14 NL=1,4
      GO TO (15,16,17,18),NL
   15 A=P(IS,J,K,L,II)/P(7,J,K,L,II)
      GO TO 19
   16 IF(I1.EQ.1) I=IS
      IF(I1.EQ.2) I=7
      A=P(I,JS,K,L,JJ)/P(I,3,K,L,JJ)
      GO TO 19
   17 IF(I2.EQ.1) J=JS
      IF(I2.EQ.2) J=3
      A=P(I,J,KS,L,KK)/P(I,J,7,L,KK)
      GO TO 19
   18 IF(I3.EQ.1) K=KS
      IF(I3.EQ.2) K=7
      A=P(I,J,K,LS,LL)/P(I,J,K,3,LL)
   19 IF(I4.EQ.3.OR.I4.EQ.6) W1=W1*A
   14 IF(I4.EQ.4.OR.I4.EQ.5) W2=W2*A
   13 CONTINUE
   12 H=H+(T(IS,JS,KS,LS,1)+T(IS,JS,KS,LS,2))*.5
     1    *W1**(1./4.)*W2**(1./12.)
      DO 11 NN=1,2
      N1=NN+2
      N2=9-N1
      IF(JJ+KK+LL.EQ.N1.OR.JJ+KK+LL.EQ.N2) R(II,1,NN)=R(II,1,NN)+H
      IF(II+KK+LL.EQ.N1.OR.II+KK+LL.EQ.N2) R(JJ,2,NN)=R(JJ,2,NN)+H
      IF(II+JJ+LL.EQ.N1.OR.II+JJ+LL.EQ.N2) R(KK,3,NN)=R(KK,3,NN)+H
   11 IF(II+JJ+KK.EQ.N1.OR.II+JJ+KK.EQ.N2) R(LL,4,NN)=R(LL,4,NN)+H
      DO 20 J=1,4
      DO 21 I=1,2
   21 S(I,J)=R(I,J,1)/4.+R(I,J,2)/12.
   20 U(J)=S(2,J)-S(1,J)
      WRITE(6,22) (S(2,J),S(1,J),U(J),J=1,4),ER(2),ER(1),ERR,
     1 ET(2),ET(1),ETT
   22 FORMAT(40X,3F15.3)
      STOP
      END
```

****

5.6 THE CASE OF FIVE FACTORS

Using analogous symbols, we can express

$$\frac{N_{ijklm}}{N_{.....}} = A_{ijklm} B_{ijklm} C_{ijklm} D_{ijklm} E_{ijklm} \,, \tag{5.26}$$

where (Das Gupta, 1991, equation 24)

$$\begin{aligned} A_{ijklm} = {} & \left(\frac{N_{ijklm}}{N_{.jklm}}\right)^{\frac{1}{5}} \cdot \left(\frac{N_{ijkl.}}{N_{.jkl.}} \cdot \frac{N_{ijk.m}}{N_{.jk.m}} \cdot \frac{N_{ij.lm}}{N_{.j.lm}} \cdot \frac{N_{i.klm}}{N_{..klm}}\right)^{\frac{1}{20}} \cdot \\ & \left(\frac{N_{ijk..}}{N_{.jk..}} \cdot \frac{N_{ij.l.}}{N_{.j.l.}} \cdot \frac{N_{ij..m}}{N_{.j..m}} \cdot \frac{N_{i.kl.}}{N_{..kl.}} \cdot \frac{N_{i.k.m}}{N_{..k.m}} \cdot \frac{N_{i..lm}}{N_{...lm}}\right)^{\frac{1}{30}} \cdot \\ & \left(\frac{N_{i...m}}{N_{....m}} \cdot \frac{N_{i..l.}}{N_{...l.}} \cdot \frac{N_{i.k..}}{N_{..k..}} \cdot \frac{N_{ij...}}{N_{.j...}}\right)^{\frac{1}{20}} \cdot \left(\frac{N_{i....}}{N_{.....}}\right)^{\frac{1}{5}} , \end{aligned} \tag{5.27}$$

and B, C, D, and E are obtained from (5.27) by interchanging the subscripts.

The difference $t_{.....} - T_{.....}$ can be expressed as the sum of six effects (including the rate effect). Each effect is the difference between two standardized rates, the two typical of them being

$R(\bar{T})$ = (I,J,K,L,M)-standardized rate in population 1

$$= \sum_{i,j,k,l,m} \frac{\frac{n_{ijklm}}{n_{.....}} + \frac{N_{ijklm}}{N_{.....}}}{2} T_{ijklm} \,, \tag{5.28}$$

$I(\bar{A})$ = (J,K,L,M,R)-standardized rate in population 1

$$= \sum_{i,j,k,l,m} \frac{t_{ijklm} + T_{ijklm}}{2} \text{ [Expression (2.33), i.e., (2.36) x A with subscripts ijklm in each letter]} . \tag{5.29}$$

The remaining 10 standardized rates may be obtained from (5.28) and (5.29) by interchanging the letters.

Example 5.8

Table 5.15 presents the population size and the mean annual earnings of Whites, and Asian and Pacific Islanders (API's) by four occupations, three age groups, three education groups, sex, and work status, as described in the footnote of the table, for the 1980 census. Das Gupta (1989) used similar data from the same source to study the race-sex inequalities in earnings. The mean earnings of Whites and API's are, respectively, \$30,998 and \$30,433, giving a difference of \$565 in favor of Whites. As table 5.16 shows, this difference would have been \$2,813 had the distributions of populations by occupation, age, education, sex, and work status been identical in the two groups. In other words, if we assume that, ideally, the mean earnings should depend only on these five factors, this difference of \$2,813 measures the inequity in mean earnings between Whites and API's. If everything else including the cell-specific mean earnings were the same for the two groups, only the difference in education structures would make the mean earnings of API's \$1,582 higher than those for Whites. Similarly, only the difference in occupation structures would produce a difference of \$1,991 in mean earnings in favor of API's. On the other hand, the differences in the other three factors, namely, age, sex, and work status, in the two groups tend to produce higher mean earnings for Whites. If there were no inequity in earnings, the rate effect (R) in table 5.16 would be 0 and the total difference would be \$ -2,248. This implies that in the absence of inequity, the mean earnings of API's would be \$2,248 higher than those for Whites.

Table 5.15. Civilian Labor Force With Earnings in 1979 and Mean Annual Earnings, by Occupation, Age, Education, Sex, and Work Status: Asian and Pacific Islander, and White, 1980

(Rate is mean annual earnings in dollars)

Work status	Sex	Education	Age	Occupation	Asian and Pacific Islander (population 1)		White (population 2)	
m	l	k	j	i	Size N_{ijklm}	Rate T_{ijklm}	Size n_{ijklm}	Rate t_{ijklm}
1	1	1	1	1	6306	20552.15	347389	23293.16
1	1	1	1	2	4899	21570.24	132196	22496.84
1	1	1	1	3	222	29108.74	4142	28504.84
1	1	1	1	4	1275	20138.09	93503	23610.01
1	1	1	2	1	6967	30446.95	292235	35928.09
1	1	1	2	2	3754	25386.35	81308	28764.83
1	1	1	2	3	296	41404.85	3169	49227.33
1	1	1	2	4	845	27496.49	63562	35899.38
1	1	1	3	1	3686	35049.83	252218	42633.10
1	1	1	3	2	2749	28243.55	89491	32475.30
1	1	1	3	3	125	48238.76	3044	50080.75
1	1	1	3	4	553	29329.05	53094	37501.52
1	1	2	1	1	3348	22017.03	166014	24854.07
1	1	2	1	2	7159	22462.81	95481	23434.66
1	1	2	1	3	1238	29735.58	7964	30445.42
1	1	2	1	4	606	25036.89	30210	26888.48
1	1	2	2	1	4155	30156.48	189982	36453.98
1	1	2	2	2	7700	26775.74	69167	31130.81
1	1	2	2	3	1495	56310.06	6681	48133.87
1	1	2	2	4	480	26541.35	26501	38175.68
1	1	2	3	1	1935	34665.08	133608	42693.57
1	1	2	3	2	2214	31988.64	53838	34883.26
1	1	2	3	3	445	64515.67	7525	49889.73
1	1	2	3	4	270	30393.07	17349	38340.14
1	1	3	1	1	2369	24322.07	48174	25888.75
1	1	3	1	2	5768	24410.36	24191	23837.90
1	1	3	1	3	6393	32904.94	91556	35288.20
1	1	3	1	4	330	24066.62	6830	28476.43
1	1	3	2	1	4153	32712.66	92193	36057.86
1	1	3	2	2	8121	28991.28	28840	32059.44
1	1	3	2	3	9732	64519.72	87882	68222.38
1	1	3	2	4	310	26471.85	8657	35407.87
1	1	3	3	1	1642	35102.54	67843	41605.08
1	1	3	3	2	1867	32160.77	16471	36384.68
1	1	3	3	3	3919	64245.75	68250	71961.60
1	1	3	3	4	163	37806.38	6006	35885.50
1	2	1	1	1	1973	14724.19	87369	15444.64
1	2	1	1	2	299	19156.99	5633	18231.18
1	2	1	1	3	77	15996.17	1194	17699.53
1	2	1	1	4	676	15173.48	24038	15569.57
1	2	1	2	1	1354	15509.48	31914	18950.33
1	2	1	2	2	145	18561.83	1053	20208.55
1	2	1	2	3	95	19613.68	362	35656.55
1	2	1	2	4	240	16413.38	9087	18442.76
1	2	1	3	1	655	16470.11	22063	18759.59
1	2	1	3	2	66	17051.06	723	20645.65
1	2	1	3	3	49	27837.65	393	34205.83
1	2	1	3	4	159	16364.37	6935	18116.46
1	2	2	1	1	1451	16425.70	40764	17186.57
1	2	2	1	2	394	18695.69	3843	19722.91
1	2	2	1	3	429	29189.21	1233	22062.73
1	2	2	1	4	154	16274.48	6720	17002.31
1	2	2	2	1	650	20683.15	22720	19825.88
1	2	2	2	2	156	20571.44	1001	22073.44
1	2	2	2	3	366	40819.64	623	27111.00
1	2	2	2	4	241	24412.66	3681	19625.39
1	2	2	3	1	298	20627.06	14084	20487.98

Table 5.15. Civilian Labor Force With Earnings in 1979 and Mean Annual Earnings, by Occupation, Age, Education, Sex, and Work Status: Asian and Pacific Islander, and White, 1980—Continued

(Rate is mean annual earnings in dollars)

Work status	Sex	Education	Age	Occupation	Asian and Pacific Islander (population 1)		White (population 2)	
m	l	k	j	i	Size N_{ijklm}	Rate T_{ijklm}	Size n_{ijklm}	Rate t_{ijklm}
1	2	2	3	2	42	19770.24	611	22686.30
1	2	2	3	3	181	39430.30	481	32101.23
1	2	2	3	4	24	31349.17	2284	18513.86
1	2	3	1	1	362	15958.55	10151	18578.88
1	2	3	1	2	198	22682.50	1342	22755.94
1	2	3	1	3	2301	25325.56	9787	22506.49
1	2	3	1	4	25	9363.40	1107	18602.58
1	2	3	2	1	436	19968.83	10938	21785.26
1	2	3	2	2	78	25682.24	473	22740.22
1	2	3	2	3	2796	43896.64	4858	40368.71
1	2	3	2	4	81	27204.26	986	20834.79
1	2	3	3	1	221	28769.00	7722	23092.75
1	2	3	3	2	38	24458.16	220	23961.32
1	2	3	3	3	852	42882.22	2964	47513.34
1	2	3	3	4	31	26834.03	634	19383.53
2	1	1	1	1	1987	13413.28	42099	14653.07
2	1	1	1	2	1511	12432.07	16273	14750.60
2	1	1	1	3	73	13939.23	1488	19063.69
2	1	1	1	4	540	16377.65	20437	17074.84
2	1	1	2	1	1669	20181.60	22153	27098.18
2	1	1	2	2	556	19703.82	5412	23016.99
2	1	1	2	3	51	40952.32	850	45744.51
2	1	1	2	4	290	18786.30	10019	30637.37
2	1	1	3	1	731	18473.15	18694	31839.66
2	1	1	3	2	281	22111.09	6123	25484.38
2	1	1	3	3	72	46188.54	1090	45535.32
2	1	1	3	4	142	19507.48	9005	29173.63
2	1	2	1	1	1165	11754.58	26966	14685.16
2	1	2	1	2	1891	14022.91	12256	14753.29
2	1	2	1	3	565	23891.12	2915	22727.08
2	1	2	1	4	230	11581.25	7972	17636.10
2	1	2	2	1	692	20360.49	22703	24958.99
2	1	2	2	2	836	21723.88	4974	23023.14
2	1	2	2	3	666	60133.12	2284	47865.35
2	1	2	2	4	247	15632.82	5201	28061.94
2	1	2	3	1	288	20753.97	18630	27919.21
2	1	2	3	2	201	24802.15	3956	26644.25
2	1	2	3	3	215	65647.89	2718	46108.45
2	1	2	3	4	60	14218.34	3685	28842.36
2	1	3	1	1	823	12653.70	11762	14008.25
2	1	3	1	2	1565	14024.68	4677	12913.80
2	1	3	1	3	2376	27815.49	36092	24057.25
2	1	3	1	4	72	15763.87	1953	17011.58
2	1	3	2	1	683	19825.25	13921	24752.96
2	1	3	2	2	980	20661.36	2694	23091.57
2	1	3	2	3	3722	64740.15	23019	62417.03
2	1	3	2	4	205	21343.28	2160	33559.61
2	1	3	3	1	287	24727.82	10132	29638.24
2	1	3	3	2	187	24564.70	1452	27095.94
2	1	3	3	3	1435	72847.97	20843	65177.10
2	1	3	3	4	65	22511.47	1602	22434.63
2	2	1	1	1	845	9992.65	31676	8463.99
2	2	1	1	2	114	7872.89	1743	10061.72
2	2	1	1	3	33	26490.14	902	10730.86
2	2	1	1	4	482	12030.86	14072	8437.87
2	2	1	2	1	560	9155.62	15754	8386.64
2	2	1	2	2	41	20344.25	427	9209.83

Table 5.15. Civilian Labor Force With Earnings in 1979 and Mean Annual Earnings, by Occupation, Age, Education, Sex, and Work Status: Asian and Pacific Islander, and White, 1980—Continued

(Rate is mean annual earnings in dollars)

Work status	Sex	Education	Age	Occupation	Asian and Pacific Islander (population 1)		White (population 2)	
m	l	k	j	i	Size	Rate	Size	Rate
					N_{ijklm}	T_{ijklm}	n_{ijklm}	t_{ijklm}
2	2	1	2	3	49	28832.97	368	16203.06
2	2	1	2	4	251	8024.56	10942	9128.05
2	2	1	3	1	255	11924.00	9966	9533.53
2	2	1	3	2	8	23640.02	201	12317.19
2	2	1	3	3	20	22605.00	229	18437.26
2	2	1	3	4	112	11460.63	7309	8550.02
2	2	2	1	1	500	8611.16	19376	9596.07
2	2	2	1	2	250	11390.69	1661	10123.12
2	2	2	1	3	377	13681.88	1205	11192.69
2	2	2	1	4	226	8561.37	4894	9713.37
2	2	2	2	1	276	12312.71	12204	11858.49
2	2	2	2	2	38	12010.27	379	10003.39
2	2	2	2	3	235	33734.46	523	16094.54
2	2	2	2	4	141	15655.46	4170	9641.37
2	2	2	3	1	164	29692.09	7862	14204.76
2	2	2	3	2	9	20991.63	297	10874.84
2	2	2	3	3	66	33692.83	237	23101.27
2	2	2	3	4	12	8864.98	2487	8845.77
2	2	3	1	1	277	9240.31	5821	10491.60
2	2	3	1	2	88	14716.03	774	8689.08
2	2	3	1	3	1705	19828.81	7107	13448.46
2	2	3	1	4	76	5480.27	924	10567.47
2	2	3	2	1	185	12951.80	5703	14099.41
2	2	3	2	2	41	11492.82	253	9873.61
2	2	3	2	3	1674	33889.79	2849	25738.88
2	2	3	2	4	50	8773.19	781	10048.33
2	2	3	3	1	114	19177.33	4312	17249.20
2	2	3	3	2	13	13420.39	92	10892.76
2	2	3	3	3	304	30910.62	1873	31360.83
2	2	3	3	4	58	13317.32	535	8998.24
m = .	l = .	k = .	j = .	i = .	162,090	30,433	3,684,673	30,998

Source: U.S. Bureau of the Census (1984c, tables 3, and 6; unpublished data for breakdown of college 5+ years into 5-6 and 7+ years and for earnings correct to cents). Occupation i = 1, 2, 3, 4 (executive and administrative occupations; engineers, architects, and surveyors; health diagnosing occupations; sales representatives, finance, and business services). Age j = 1, 2, 3 (age groups 25-34, 35-44, and 45-54). Education k = 1, 2, 3 (college 4, 5-6, and 7+ years). Sex l = 1, 2 (male and female). Work status m = 1, 2 (worked year-round full-time in 1979 and others who worked in 1979).

Table 5.16. Standardization and Decomposition of Mean Annual Earnings in Table 5.15

Mean annual earnings (dollars)	Standardization		Decomposition	
	White* (population 2)	Asian (population 1)	Difference (effects)	Percent distribution of effects
(J,K,L,M,R)-standardized mean earnings	28,745	30,736	−1,991 (I)	−352.4
(I,K,L,M,R)-standardized mean earnings	29,980	29,806	174 (J)	30.8
(I,J,L,M,R)-standardized mean earnings	28,976	30,558	−1,582 (K)	−280.0
(I,J,K,M,R)-standardized mean earnings	30,210	29,603	607 (L)	107.4
(I,J,K,L,R)-standardized mean earnings	30,168	29,624	544 (M)	96.3
(I,J,K,L,M)-standardized mean earnings	31,744	28,931	2,813 (R)	497.9
Overall mean annual earnings	30,998	30,433	565 (Total effect)	100.0

*Whites include Whites of Hispanic origin. The mean earnings for non-Hispanic Whites are higher than those shown for Whites in tables 5.15 and 5.16.

Program 5.5

The results in table 5.16 can be obtained by using Program 5.5 in which P's in line 5 denote N's and n's in table 5.15, and T's in line 6 denote T's and t's in the same table. The data file consists of 144 lines corresponding to the data in the last four columns in table 5.15 in the same order—36 lines of four numbers for each column—with the format specified in lines 7 and 8 of the program. The 12 standardized rates in table 5.16 are given by ER(N)'s and S(I,J)'s in lines 56 and 125 of the program. The six effects in table 5.16 are denoted by ERR and U(J)'s in lines 58 and 126 of the program.

5.7 THE CASE OF SIX FACTORS

In this case, we write

$$\frac{N_{ijklmn}}{N_{......}} = A_{ijklmn}B_{ijklmn}C_{ijklmn}D_{ijklmn}E_{ijklmn}F_{ijklmn}, \tag{5.30}$$

where

$$A_{ijklmn} = \left(\frac{N_{ijklmn}}{N_{.jklmn}}\right)^{\frac{1}{6}} \cdot \left(\frac{N_{ijklm.}}{N_{.jklm.}} \cdot \frac{N_{ijkl.n}}{N_{.jkl.n}} \cdot \frac{N_{ijk.mn}}{N_{.jk.mn}} \cdot \frac{N_{ij.lmn}}{N_{.j.lmn}} \cdot \frac{N_{i.klmn}}{N_{..klmn}}\right)^{\frac{1}{30}} \cdot$$

$$\left(\frac{N_{ijkl..}}{N_{.jkl..}} \cdot \frac{N_{ijk.m.}}{N_{.jk.m.}} \cdot \frac{N_{ijk..n}}{N_{.jk..n}} \cdot \frac{N_{ij.lm.}}{N_{.j.lm.}} \cdot \frac{N_{ij.l.n}}{N_{.j.l.n}} \cdot \frac{N_{ij..mn}}{N_{.j..mn}} \cdot \frac{N_{i.klm.}}{N_{..klm.}} \cdot \frac{N_{i.kl.n}}{N_{..kl.n}} \cdot \frac{N_{i.k.mn}}{N_{..k.mn}} \cdot \frac{N_{i..lmn}}{N_{...lmn}}\right)^{\frac{1}{60}} \cdot$$

$$\left(\frac{N_{i...mn}}{N_{....mn}} \cdot \frac{N_{i..l.n}}{N_{...l.n}} \cdot \frac{N_{i..lm.}}{N_{...lm.}} \cdot \frac{N_{i.k..n}}{N_{..k..n}} \cdot \frac{N_{i.k.m.}}{N_{..k.m.}} \cdot \frac{N_{i.kl..}}{N_{..kl..}} \cdot \frac{N_{ij...n}}{N_{.j...n}} \cdot \frac{N_{ij..m.}}{N_{.j..m.}} \cdot \frac{N_{ij.l..}}{N_{.j.l..}} \cdot \frac{N_{ijk...}}{N_{.jk...}}\right)^{\frac{1}{60}} \cdot$$

$$\left(\frac{N_{i....n}}{N_{.....n}} \cdot \frac{N_{i...m.}}{N_{....m.}} \cdot \frac{N_{i..l..}}{N_{...l..}} \cdot \frac{N_{i.k...}}{N_{..k...}} \cdot \frac{N_{ij....}}{N_{.j....}}\right)^{\frac{1}{30}} \cdot \left(\frac{N_{i.....}}{N_{......}}\right)^{\frac{1}{6}}, \tag{5.31}$$

and B, C, D, E, and F are obtained from (5.31) by interchanging the subscripts.

The two typical standardized rates, similar to (5.28) and (5.29), are given by

$R(\overline{T})$ = (I,J,K,L,M,N)-standardized rate in population 1

$$= \sum_{i,j,k,l,m,n} \frac{\frac{n_{ijklmn}}{n_{......}} + \frac{N_{ijklmn}}{N_{......}}}{2} T_{ijklmn}, \tag{5.32}$$

Program 5.5 (Five Factors +Rate)

```
      DIMENSION P(5,4,4,3,3,2),T(4,3,3,2,2,2),R(2,5,3),
     1 U(5),S(2,5),ET(2),ER(2)
      DOUBLE PRECISION P,R,U,S,ET,ER,Q,H,A,W1,W2,W3
      DO 1 N=1,2
      READ(5,2) (((((P(I,J,K,L,M,N),I=1,4),J=1,3),K=1,3),L=1,2),M=1,2)
    1 READ(5,9) (((((T(I,J,K,L,M,N),I=1,4),J=1,3),K=1,3),L=1,2),M=1,2)
    2 FORMAT(4F10.0)
    9 FORMAT(4F10.2)
      DO 8 N=1,2
      DO 3 J=1,3
      DO 3 K=1,3
      DO 3 L=1,2
      DO 3 M=1,2
      P(5,J,K,L,M,N)=0.0
      DO 3 I=1,4
    3 P(5,J,K,L,M,N)=P(5,J,K,L,M,N)+P(I,J,K,L,M,N)
      DO 4 I=1,5
      DO 4 K=1,3
      DO 4 L=1,2
      DO 4 M=1,2
      P(I,4,K,L,M,N)=0.0
      DO 4 J=1,3
    4 P(I,4,K,L,M,N)=P(I,4,K,L,M,N)+P(I,J,K,L,M,N)
      DO 5 I=1,5
      DO 5 J=1,4
      DO 5 L=1,2
      DO 5 M=1,2
      P(I,J,4,L,M,N)=0.0
      DO 5 K=1,3
    5 P(I,J,4,L,M,N)=P(I,J,4,L,M,N)+P(I,J,K,L,M,N)
      DO 6 I=1,5
      DO 6 J=1,4
      DO 6 K=1,4
      DO 6 M=1,2
      P(I,J,K,3,M,N)=0.0
      DO 6 L=1,2
    6 P(I,J,K,3,M,N)=P(I,J,K,3,M,N)+P(I,J,K,L,M,N)
      DO 10 I=1,5
      DO 10 J=1,4
      DO 10 K=1,4
      DO 10 L=1,3
      P(I,J,K,L,3,N)=0.0
      DO 10 M=1,2
   10 P(I,J,K,L,3,N)=P(I,J,K,L,3,N)+P(I,J,K,L,M,N)
    8 CONTINUE
      DO 7 N=1,2
      ET(N)=0.0
      ER(N)=0.0
      DO 7 I=1,4
      DO 7 J=1,3
      DO 7 K=1,3
      DO 7 L=1,2
      DO 7 M=1,2
      ET(N)=ET(N)+P(I,J,K,L,M,N)*T(I,J,K,L,M,N)/P(5,4,4,3,3,N)
      Q=(P(I,J,K,L,M,1)/P(5,4,4,3,3,1)+P(I,J,K,L,M,2)/P(5,4,4,3,3,2))*.5
    7 ER(N)=ER(N)+Q*T(I,J,K,L,M,N)
      ETT=ET(2)-ET(1)
      ERR=ER(2)-ER(1)
      DO 11 I=1,2
      DO 11 J=1,5
      DO 11 K=1,3
   11 R(I,J,K)=0.0
      DO 13 II=1,2
      DO 13 JJ=1,2
      DO 13 KK=1,2
      DO 13 LL=1,2
      DO 13 MM=1,2
      H=0.0
      DO 12 IS=1,4
      DO 12 JS=1,3
      DO 12 KS=1,3
      DO 12 LS=1,2
      DO 12 MS=1,2
      W1=1.0
      W2=1.0
      W3=1.0
      DO 14 I1=1,2
      DO 14 I2=1,2
      DO 14 I3=1,2
      DO 14 I4=1,2
      I5=I1+I2+I3+I4
      IF(I1.EQ.1) J=JS
      IF(I1.EQ.2) J=4
      IF(I2.EQ.1) K=KS
      IF(I2.EQ.2) K=4
      IF(I3.EQ.1) L=LS
      IF(I3.EQ.2) L=3
      IF(I4.EQ.1) M=MS
      IF(I4.EQ.2) M=3
```

Program 5.5 (continued)

```
      DO 20 NL=1,5
      GO TO (15,16,17,18,19),NL
   15 A=P(IS,J,K,L,M,II)/P(5,J,K,L,M,II)
      GO TO 21
   16 IF(I1.EQ.1) I=IS
      IF(I1.EQ.2) I=5
      A=P(I,JS,K,L,M,JJ)/P(I,4,K,L,M,JJ)
      GO TO 21
   17 IF(I2.EQ.1) J=JS
      IF(I2.EQ.2) J=4
      A=P(I,J,KS,L,M,KK)/P(I,J,4,L,M,KK)
      GO TO 21
   18 IF(I3.EQ.1) K=KS
      IF(I3.EQ.2) K=4
      A=P(I,J,K,LS,M,LL)/P(I,J,K,3,M,LL)
      GO TO 21
   19 IF(I4.EQ.1) L=LS
      IF(I4.EQ.2) L=3
      A=P(I,J,K,L,MS,MM)/P(I,J,K,L,3,MM)
   21 IF(I5.EQ.4.OR.I5.EQ.8) W1=W1*A
      IF(I5.EQ.5.OR.I5.EQ.7) W2=W2*A
   20 IF(I5.EQ.6           ) W3=W3*A
   14 CONTINUE
   12 H=H+(T(IS,JS,KS,LS,MS,1)+T(IS,JS,KS,LS,MS,2))*.5
     1    *W1**(1./5.)*W2**(1./20.)*W3**(1./30.)
      DO 13 NN=1,3
      N1=NN+3
      N2=12-N1
      IF(JJ+KK+LL+MM.EQ.N1.OR.JJ+KK+LL+MM.EQ.N2) R(II,1,NN)=R(II,1,NN)+H
      IF(II+KK+LL+MM.EQ.N1.OR.II+KK+LL+MM.EQ.N2) R(JJ,2,NN)=R(JJ,2,NN)+H
      IF(II+JJ+LL+MM.EQ.N1.OR.II+JJ+LL+MM.EQ.N2) R(KK,3,NN)=R(KK,3,NN)+H
      IF(II+JJ+KK+MM.EQ.N1.OR.II+JJ+KK+MM.EQ.N2) R(LL,4,NN)=R(LL,4,NN)+H
   13 IF(II+JJ+KK+LL.EQ.N1.OR.II+JJ+KK+LL.EQ.N2) R(MM,5,NN)=R(MM,5,NN)+H
      DO 22 J=1,5
      DO 23 I=1,2
   23 S(I,J)=R(I,J,1)/5.+R(I,J,2)/20.+R(I,J,3)/30.
   22 U(J)=S(2,J)-S(1,J)
      WRITE(6,24) (S(2,J),S(1,J),U(J),J=1,5),ER(2),ER(1),ERR,
     1 ET(2),ET(1),ETT
   24 FORMAT(40X,3F15.0)
      STOP
      END
****
```

$I(\bar{A})$ = (J,K,L,M,N,R)-standardized rate in population 1

$$= \sum_{i,j,k,l,m,n} \frac{t_{ijklmn} + T_{ijklmn}}{2} \quad \text{[Expression (2.41), i.e., (2.44) x A with subscripts ijklmn in each letter]} . \tag{5.33}$$

Other standardized rates and the effects are easily obtained from (5.32) and (5.33).

Example 5.9

Table 5.17 is from the 1970 U.S. Census where the women and the average number of children ever born to them are cross-classified by six factors: family income, husband's education, husband's occupation, wife's labor force status, wife's age at marriage, and race, for two education groups of women, namely, not a high school graduate and high school, 4 years (no college). Janowitz (1976) used the same data source to do similar analysis, but since she considered only wife's age at marriage and wife's labor force status as the explaining variables, her results cannot be directly compared with ours. The average number of children ever born is 3.428 for women who were not high school graduates and 3.005 for women who had 4 years of high school, the difference in these two averages being .423 child. The six-factor decomposition in table 5.18 (along with the rates as a factor) shows that each of the seven factors contributes positively towards explaining the difference of .423 in the average number of children in the two groups of women. The differences in family income, husband's education, husband's occupation, wife's labor force status, wife's age at marriage, and race explain, respectively, 1.9, 15.4, 8.7, 3.3, 13.0, and 9.0 percents of the total difference between the average number of children in the two groups of women. In other words, 48.7 percent of the total difference in the fertility between the high school graduates and non-high school graduates still remains unexplained even after standardization with respect to the six factors simultaneously. Obviously, of the six factors, husband's education plays the most important role in explaining the difference, wife's age at marriage being the next in importance. Virtually identical results were obtained by Das Gupta (1984, table 3) when a more complicated method was applied to the same set of data. This example will be discussed again in Example 6.1 (tables 6.1 and 6.2) in the context of simultaneous consideration of three populations.

Program 5.6

The results in table 5.18 can be obtained by using Program 5.6. P's in line 5 of the program denote N's and n's in table 5.17, and T's in line 7 denote T's and t's in the same table. The data file consists of 192 lines corresponding to the data in the last four columns in table 5.17 in the same order. Each column takes 48 lines, each line having seven numbers with the format specified in lines 9 and 10 of the program. The 14 standardized rates in table 5.18 are given by ER(N1)'s and S(I,J)'s in lines 74 and 154 of the program. The seven effects in table 5.18 are denoted by ERR and U(J)'s in lines 76 and 155 of the program.

5.8 THE CASE OF P FACTORS

As in (5.30) and (5.31), we express

$$\frac{N_{i_1 \text{ to } i_p}}{N_{......}} = A_{1i_1 \text{ to } i_p} A_{2i_1 \text{ to } i_p} \cdots\cdots A_{pi_1 \text{ to } i_p} , \tag{5.34}$$

where

$$A_{1i_1 \text{ to } i_p} = \left(\frac{N_{i_1 i_2 \text{ to } i_p}}{N_{.i_2 \text{ to } i_p}}\right)^{\frac{1}{P}} \cdot \left(\frac{N_{i_1 i_2 \text{ to } i_{p-1}.}}{N_{.i_2 \text{ to } i_{p-1}.}} \cdots\cdots \frac{N_{i_1 .i_3 \text{ to } i_p}}{N_{..i_3 \text{ to } i_p}}\right)^{\frac{1}{P\binom{P-1}{1}}} \cdots\cdots \cdot \left(\frac{N_{i_1 .....}}{N_{......}}\right)^{\frac{1}{P}} \tag{5.35}$$

$$= \prod_{r=0}^{P-1} (Z)^{\frac{1}{P\binom{P-1}{r}}} , \tag{5.36}$$

Table 5.17. Wives 35 to 44 Years Old and Average Number of Children Ever Born (Rate) by Family Income, Husband's Education, Husband's Occupation, Wife's Labor Force Status, Wife's Age at Marriage, and Race: Wives High School 4 Years and Wives Not a High School Graduate, United States, 1970

Race	Wife's age at marriage	Wife's labor force status	Husband's occupation	Husband's education	Family income	Wives, high school 4 years (population 1)		Wives, not a high school graduate (population 2)	
						Size	Rate	Size	Rate
n	m	l	k	j	i	N_{ijklmn}	T_{ijklmn}	n_{ijklmn}	T_{ijklmn}
1	1	1	1	1	1	1908	3.578	4197	3.496
1	1	1	1	1	2	3231	2.972	8257	3.573
1	1	1	1	1	3	3100	3.300	5519	3.394
1	1	1	1	1	4	17395	3.176	25046	3.423
1	1	1	1	1	5	25425	3.430	31399	3.574
1	1	1	1	1	6	10612	3.511	11859	3.597
1	1	1	1	1	7	9009	3.377	9269	3.540
1	1	1	1	2	1	4221	3.065	1525	2.814
1	1	1	1	2	2	5154	3.234	2551	3.397
1	1	1	1	2	3	5884	3.121	2496	3.370
1	1	1	1	2	4	42696	3.194	15638	3.309
1	1	1	1	2	5	84344	3.305	26130	3.408
1	1	1	1	2	6	40242	3.392	10816	3.426
1	1	1	1	2	7	34088	3.315	7537	3.649
1	1	1	1	3	1	3457	3.092	977	3.209
1	1	1	1	3	2	3086	3.287	1028	3.116
1	1	1	1	3	3	2999	3.263	1084	3.123
1	1	1	1	3	4	21630	3.215	6571	3.374
1	1	1	1	3	5	78263	3.284	14867	3.565
1	1	1	1	3	6	65609	3.366	10286	3.420
1	1	1	1	3	7	72352	3.452	8669	3.540
1	1	1	2	1	1	6851	3.415	32827	4.439
1	1	1	2	1	2	13914	3.326	55962	4.119
1	1	1	2	1	3	14127	3.321	40422	3.912
1	1	1	2	1	4	67259	3.419	143027	3.841
1	1	1	2	1	5	83649	3.556	151472	3.826
1	1	1	2	1	6	24032	3.830	40710	3.986
1	1	1	2	1	7	10198	3.813	14493	4.279
1	1	1	2	2	1	5842	3.351	4390	3.536
1	1	1	2	2	2	9842	3.187	7258	3.508
1	1	1	2	2	3	11074	3.185	6891	3.514
1	1	1	2	2	4	70305	3.245	34323	3.420
1	1	1	2	2	5	120211	3.445	50758	3.518
1	1	1	2	2	6	38270	3.618	14515	3.796
1	1	1	2	2	7	13535	3.785	6314	3.848
1	1	1	2	3	1	1276	3.139	439	3.784
1	1	1	2	3	2	1180	3.503	980	3.131
1	1	1	2	3	3	1529	3.534	535	3.875
1	1	1	2	3	4	9577	3.250	4391	3.557
1	1	1	2	3	5	21541	3.403	7689	3.620
1	1	1	2	3	6	9016	3.700	2834	3.849
1	1	1	2	3	7	4601	3.531	1091	3.894
1	1	2	1	1	1	1087	3.015	1759	3.437
1	1	2	1	1	2	1726	2.939	3623	3.155
1	1	2	1	1	3	2012	3.235	2893	3.287
1	1	2	1	1	4	11647	3.069	15523	3.082
1	1	2	1	1	5	33752	2.821	34998	2.945
1	1	2	1	1	6	19770	2.889	15942	2.999
1	1	2	1	1	7	10830	2.915	9188	3.135
1	1	2	1	2	1	1838	2.835	505	2.790
1	1	2	1	2	2	3395	2.935	1111	2.877
1	1	2	1	2	3	2702	2.839	1012	2.718
1	1	2	1	2	4	23040	2.877	6528	3.016
1	1	2	1	2	5	88101	2.866	21925	2.956
1	1	2	1	2	6	60315	2.783	13225	3.010
1	1	2	1	2	7	32550	2.815	6656	3.121
1	1	2	1	3	1	1544	3.071	332	4.244
1	1	2	1	3	2	1627	3.033	701	2.642
1	1	2	1	3	3	1675	2.704	444	3.011

Table 5.17. Wives 35 to 44 Years Old and Average Number of Children Ever Born (Rate) by Family Income, Husband's Education, Husband's Occupation, Wife's Labor Force Status, Wife's Age at Marriage, and Race: Wives High School 4 Years and Wives Not a High School Graduate, United States, 1970—Continued

Race	Wife's age at marriage	Wife's labor force status	Husband's occupation	Husband's education	Family income	Wives, high school 4 years (population 1)		Wives, not a high school graduate (population 2)	
						Size	Rate	Size	Rate
n	m	l	k	j	i	N_{ijklmn}	T_{ijklmn}	n_{ijklmn}	T_{ijklmn}
1	1	2	1	3	4	10937	3.076	2933	2.993
1	1	2	1	3	5	54525	2.987	10732	3.025
1	1	2	1	3	6	52225	2.865	7865	2.960
1	1	2	1	3	7	40354	2.933	5064	2.970
1	1	2	2	1	1	2891	3.080	9300	3.516
1	1	2	2	1	2	6612	3.145	19812	3.619
1	1	2	2	1	3	7450	3.003	16471	3.611
1	1	2	2	1	4	46792	2.995	86584	3.336
1	1	2	2	1	5	120053	3.033	172363	3.220
1	1	2	2	1	6	55956	3.024	67635	3.324
1	1	2	2	1	7	18456	3.208	22088	3.403
1	1	2	2	2	1	2495	3.066	1294	2.931
1	1	2	2	2	2	4499	3.099	2574	3.168
1	1	2	2	2	3	4430	3.290	2795	3.455
1	1	2	2	2	4	37267	3.049	17642	3.249
1	1	2	2	2	5	127053	2.975	42958	3.178
1	1	2	2	2	6	71281	2.924	20828	3.256
1	1	2	2	2	7	24152	3.000	6780	3.367
1	1	2	2	3	1	414	2.734	249	4.406
1	1	2	2	3	2	624	3.002	490	2.986
1	1	2	2	3	3	618	3.620	403	2.628
1	1	2	2	3	4	4230	2.955	2511	3.185
1	1	2	2	3	5	19241	3.121	5948	3.251
1	1	2	2	3	6	13224	2.962	3517	3.090
1	1	2	2	3	7	5297	3.042	1205	3.232
1	2	1	1	1	1	1139	2.487	2166	2.640
1	2	1	1	1	2	1735	2.479	2825	2.384
1	2	1	1	1	3	1887	2.868	2275	2.507
1	2	1	1	1	4	9490	2.523	8721	2.772
1	2	1	1	1	5	11898	2.769	9249	2.674
1	2	1	1	1	6	3915	2.837	3269	2.941
1	2	1	1	1	7	3365	2.912	2501	2.924
1	2	1	1	2	1	2389	2.634	1114	2.066
1	2	1	1	2	2	2975	2.527	1824	2.265
1	2	1	1	2	3	4627	2.823	1195	2.542
1	2	1	1	2	4	31012	2.695	8207	2.575
1	2	1	1	2	5	50437	2.779	9969	2.789
1	2	1	1	2	6	19248	2.930	3206	2.925
1	2	1	1	2	7	15680	2.890	2161	2.556
1	2	1	1	3	1	3031	2.788	498	2.430
1	2	1	1	3	2	2656	2.689	583	2.340
1	2	1	1	3	3	2508	2.687	656	2.422
1	2	1	1	3	4	22334	2.694	5179	2.558
1	2	1	1	3	5	74947	2.851	11183	2.573
1	2	1	1	3	6	59232	2.995	6251	2.756
1	2	1	1	3	7	57384	3.100	4961	2.695
1	2	1	2	1	1	4661	2.554	17604	3.125
1	2	1	2	1	2	8893	2.784	24869	2.895
1	2	1	2	1	3	8380	2.633	18317	2.838
1	2	1	2	1	4	37558	2.793	56166	2.819
1	2	1	2	1	5	36917	2.945	45001	2.922
1	2	1	2	1	6	7949	2.946	7868	3.134
1	2	1	2	1	7	2929	3.265	2900	3.387
1	2	1	2	2	1	3896	2.784	1910	2.294
1	2	1	2	2	2	7434	2.466	3990	2.725
1	2	1	2	2	3	8038	2.670	4783	2.751
1	2	1	2	2	4	46512	2.703	17903	2.796
1	2	1	2	2	5	69189	2.928	19873	2.837

Table 5.17. Wives 35 to 44 Years Old and Average Number of Children Ever Born (Rate) by Family Income, Husband's Education, Husband's Occupation, Wife's Labor Force Status, Wife's Age at Marriage, and Race: Wives High School 4 Years and Wives Not a High School Graduate, United States, 1970—Continued

Race	Wife's age at marriage	Wife's labor force status	Husband's occupation	Husband's education	Family income	Wives, high school 4 years (population 1)		Wives, not a high school graduate (population 2)	
						Size	Rate	Size	Rate
n	m	l	k	j	i	N_{ijklmn}	T_{ijklmn}	n_{ijklmn}	T_{ijklmn}
1..........	2	1	2	2	6	15325	3.089	3840	3.250
1..........	2	1	2	2	7	4857	3.049	1537	3.082
1..........	2	1	2	3	1	858	2.618	537	2.780
1..........	2	1	2	3	2	1448	2.680	719	2.983
1..........	2	1	2	3	3	976	2.703	384	2.458
1..........	2	1	2	3	4	7473	2.630	2732	2.675
1..........	2	1	2	3	5	16470	3.082	3669	2.906
1..........	2	1	2	3	6	5101	3.094	1275	3.057
1..........	2	1	2	3	7	2564	3.125	374	2.594
1..........	2	2	1	1	1	377	1.332	808	1.749
1..........	2	2	1	1	2	849	1.984	1016	2.401
1..........	2	2	1	1	3	507	3.018	1086	2.606
1..........	2	2	1	1	4	4882	2.385	4582	2.095
1..........	2	2	1	1	5	12688	2.105	9462	2.160
1..........	2	2	1	1	6	6543	2.158	3734	2.094
1..........	2	2	1	1	7	3636	2.023	2015	1.932
1..........	2	2	1	2	1	728	1.882	376	1.213
1..........	2	2	1	2	2	1190	2.241	619	2.360
1..........	2	2	1	2	3	1123	2.346	330	2.245
1..........	2	2	1	2	4	11013	2.381	2797	2.279
1..........	2	2	1	2	5	38594	2.270	7268	2.105
1..........	2	2	1	2	6	21122	2.127	3146	2.093
1..........	2	2	1	2	7	10385	2.114	1798	2.058
1..........	2	2	1	3	1	711	2.533	168	1.917
1..........	2	2	1	3	2	941	2.491	272	2.246
1..........	2	2	1	3	3	920	2.375	226	1.934
1..........	2	2	1	3	4	8026	2.418	1606	2.300
1..........	2	2	1	3	5	33809	2.379	4756	2.088
1..........	2	2	1	3	6	29500	2.315	3451	2.132
1..........	2	2	1	3	7	22082	2.248	2269	2.345
1..........	2	2	2	1	1	1419	2.107	4066	2.388
1..........	2	2	2	1	2	2834	2.319	7252	2.555
1..........	2	2	2	1	3	3073	2.167	6237	2.553
1..........	2	2	2	1	4	18965	2.243	29953	2.386
1..........	2	2	2	1	5	42400	2.231	46051	2.216
1..........	2	2	2	1	6	17053	2.192	13535	2.226
1..........	2	2	2	1	7	4937	2.149	3829	2.589
1..........	2	2	2	2	1	996	2.213	575	2.172
1..........	2	2	2	2	2	2071	2.242	1093	2.161
1..........	2	2	2	2	3	2075	2.339	1233	2.798
1..........	2	2	2	2	4	17509	2.328	6965	2.445
1..........	2	2	2	2	5	53190	2.281	14511	2.146
1..........	2	2	2	2	6	24465	2.162	5434	2.198
1..........	2	2	2	2	7	7020	2.202	1629	2.535
1..........	2	2	2	3	1	117	2.675	208	2.159
1..........	2	2	2	3	2	461	2.269	198	1.742
1..........	2	2	2	3	3	375	1.973	271	1.900
1..........	2	2	2	3	4	2876	2.556	1405	2.510
1..........	2	2	2	3	5	9021	2.569	2568	2.428
1..........	2	2	2	3	6	6083	2.191	1331	2.574
1..........	2	2	2	3	7	1967	2.161	425	2.059
2..........	1	1	1	1	1	70	2.500	829	4.290
2..........	1	1	1	1	2	287	4.662	1101	4.621
2..........	1	1	1	1	3	114	3.447	462	5.097
2..........	1	1	1	1	4	619	4.042	1566	4.473
2..........	1	1	1	1	5	460	4.117	1084	4.785
2..........	1	1	1	1	6	152	4.020	433	6.109
2..........	1	1	1	1	7	20	2.000	64	6.000

Table 5.17. Wives 35 to 44 Years Old and Average Number of Children Ever Born (Rate) by Family Income, Husband's Education, Husband's Occupation, Wife's Labor Force Status, Wife's Age at Marriage, and Race: Wives High School 4 Years and Wives Not a High School Graduate, United States, 1970—Continued

Race	Wife's age at marriage	Wife's labor force status	Husband's occupation	Husband's education	Family income	Wives, high school 4 years (population 1)		Wives, not a high school graduate (population 2)	
						Size	Rate	Size	Rate
n	m	l	k	j	i	N_{ijklmn}	T_{ijklmn}	n_{ijklmn}	T_{ijklmn}
2..........	1	1	1	2	1	77	4.974	180	6.889
2..........	1	1	1	2	2	312	4.426	197	4.401
2..........	1	1	1	2	3	342	2.877	280	3.111
2..........	1	1	1	2	4	983	3.169	909	4.824
2..........	1	1	1	2	5	925	3.675	834	3.893
2..........	1	1	1	2	6	337	4.080	271	2.919
2..........	1	1	1	2	7	22	4.000	23	4.000
2..........	1	1	1	3	1	94	3.851	44	6.864
2..........	1	1	1	3	2	45	2.467	161	4.851
2..........	1	1	1	3	3	160	3.662	48	5.000
2..........	1	1	1	3	4	680	3.438	291	3.182
2..........	1	1	1	3	5	864	4.225	198	3.944
2..........	1	1	1	3	6	361	3.925	120	3.583
2..........	1	1	1	3	7	163	4.350	13	3.000
2..........	1	1	2	1	1	1802	4.935	11893	6.083
2..........	1	1	2	1	2	2498	4.631	15641	5.513
2..........	1	1	2	1	3	1765	4.015	7522	5.222
2..........	1	1	2	1	4	4951	4.243	18163	5.336
2..........	1	1	2	1	5	3354	4.654	9962	5.643
2..........	1	1	2	1	6	886	3.763	2415	6.059
2..........	1	1	2	1	7	182	3.571	1029	6.412
2..........	1	1	2	2	1	608	4.414	956	5.112
2..........	1	1	2	2	2	1322	3.575	1329	4.937
2..........	1	1	2	2	3	1339	3.235	956	4.812
2..........	1	1	2	2	4	3787	3.648	3147	4.362
2..........	1	1	2	2	5	2709	4.042	1726	4.301
2..........	1	1	2	2	6	627	3.191	567	5.552
2..........	1	1	2	2	7	132	5.545	250	4.580
2..........	1	1	2	3	1	108	5.750	84	1.786
2..........	1	1	2	3	2	310	5.545	224	7.188
2..........	1	1	2	3	3	103	5.146	240	5.275
2..........	1	1	2	3	4	446	3.090	241	6.469
2..........	1	1	2	3	5	469	3.578	308	4.029
2..........	1	1	2	3	6	59	2.051	116	4.362
2..........	1	1	2	3	7	61	3.180	47	3.489
2..........	1	2	1	1	1	113	3.850	424	3.976
2..........	1	2	1	1	2	185	2.703	631	4.174
2..........	1	2	1	1	3	148	4.142	481	4.534
2..........	1	2	1	1	4	861	4.416	1460	3.903
2..........	1	2	1	1	5	1312	3.091	2374	4.214
2..........	1	2	1	1	6	543	3.169	714	3.922
2..........	1	2	1	1	7	314	2.815	347	3.689
2..........	1	2	1	2	1	75	1.573	45	1.533
2..........	1	2	1	2	2	265	4.083	109	2.248
2..........	1	2	1	2	3	132	5.591	65	3.800
2..........	1	2	1	2	4	865	3.816	473	4.404
2..........	1	2	1	2	5	2162	3.230	1112	3.987
2..........	1	2	1	2	6	1375	3.231	251	2.920
2..........	1	2	1	2	7	581	3.372	156	4.199
2..........	1	2	1	3	1	24	4.000	62	2.048
2..........	1	2	1	3	2	57	3.965	101	3.604
2..........	1	2	1	3	3	77	5.649	97	3.216
2..........	1	2	1	3	4	549	3.109	288	3.073
2..........	1	2	1	3	5	1414	3.071	561	3.961
2..........	1	2	1	3	6	1003	3.518	342	3.687
2..........	1	2	1	3	7	525	2.497	219	2.411
2..........	1	2	2	1	1	1048	4.256	6796	5.257
2..........	1	2	2	1	2	2376	4.559	12424	4.972

Table 5.17. Wives 35 to 44 Years Old and Average Number of Children Ever Born (Rate) by Family Income, Husband's Education, Husband's Occupation, Wife's Labor Force Status, Wife's Age at Marriage, and Race: Wives High School 4 Years and Wives Not a High School Graduate, United States, 1970—Continued

Race	Wife's age at marriage	Wife's labor force status	Husband's occupation	Husband's education	Family income	Wives, high school 4 years (population 1)		Wives, not a high school graduate (population 2)	
						Size	Rate	Size	Rate
n	m	l	k	j	i	N_{ijklmn}	T_{ijklmn}	n_{ijklmn}	T_{ijklmn}
2..........	1	2	2	1	3	1963	4.090	7475	5.181
2..........	1	2	2	1	4	7153	4.035	21673	4.838
2..........	1	2	2	1	5	9809	3.942	20226	4.838
2..........	1	2	2	1	6	4197	3.729	6336	4.925
2..........	1	2	2	1	7	1139	3.258	2336	5.321
2..........	1	2	2	2	1	347	3.787	301	4.704
2..........	1	2	2	2	2	1059	2.589	881	4.328
2..........	1	2	2	2	3	624	3.316	564	5.394
2..........	1	2	2	2	4	3186	3.724	2075	4.470
2..........	1	2	2	2	5	5840	3.613	3411	3.944
2..........	1	2	2	2	6	2752	3.629	1162	4.731
2..........	1	2	2	2	7	1064	3.570	308	4.263
2..........	1	2	2	3	1	115	2.826	96	4.094
2..........	1	2	2	3	2	213	4.784	137	4.664
2..........	1	2	2	3	3	137	3.584	36	2.000
2..........	1	2	2	3	4	573	3.590	408	3.990
2..........	1	2	2	3	5	1364	3.488	633	3.444
2..........	1	2	2	3	6	679	3.110	199	3.935
2..........	1	2	2	3	7	192	3.781	139	3.842
2..........	2	1	1	1	1	122	2.041	334	3.150
2..........	2	1	1	1	2	138	2.188	662	2.636
2..........	2	1	1	1	3	194	2.454	435	2.430
2..........	2	1	1	1	4	544	2.364	959	4.092
2..........	2	1	1	1	5	337	2.407	378	3.418
2..........	2	1	1	1	6	63	4.365	81	1.580
2..........	2	1	1	1	7	27	3.000	23	4.000
2..........	2	1	1	2	1	35	2.629	79	1.519
2..........	2	1	1	2	2	144	2.201	144	2.542
2..........	2	1	1	2	3	214	2.542	229	1.520
2..........	2	1	1	2	4	1023	2.382	523	2.790
2..........	2	1	1	2	5	809	2.447	165	1.485
2..........	2	1	1	2	6	170	3.429	55	3.673
2..........	2	1	1	2	7	72	2.389	0	.000
2..........	2	1	1	3	1	41	1.000	0	.000
2..........	2	1	1	3	2	233	3.129	79	4.570
2..........	2	1	1	3	3	161	3.584	61	3.541
2..........	2	1	1	3	4	635	2.551	240	3.162
2..........	2	1	1	3	5	829	2.752	384	2.836
2..........	2	1	1	3	6	322	2.416	60	2.550
2..........	2	1	1	3	7	151	2.722	62	6.532
2..........	2	1	2	1	1	1495	3.418	7519	4.006
2..........	2	1	2	1	2	2377	3.109	9172	3.592
2..........	2	1	2	1	3	1583	3.683	4260	3.714
2..........	2	1	2	1	4	3212	3.030	8614	3.793
2..........	2	1	2	1	5	2095	3.261	4363	3.736
2..........	2	1	2	1	6	212	3.745	781	4.303
2..........	2	1	2	1	7	186	3.656	239	3.393
2..........	2	1	2	2	1	690	3.346	760	4.068
2..........	2	1	2	2	2	1314	2.874	1263	3.121
2..........	2	1	2	2	3	842	2.350	957	3.304
2..........	2	1	2	2	4	2435	2.623	1771	3.267
2..........	2	1	2	2	5	1531	2.830	1006	3.547
2..........	2	1	2	2	6	349	3.559	119	2.437
2..........	2	1	2	2	7	124	4.573	83	6.301
2..........	2	1	2	3	1	80	2.762	192	4.141
2..........	2	1	2	3	2	82	4.171	225	5.436
2..........	2	1	2	3	3	118	3.551	95	1.221
2..........	2	1	2	3	4	628	2.279	397	3.053
2..........	2	1	2	3	5	305	3.469	184	2.668

Table 5.17. Wives 35 to 44 Years Old and Average Number of Children Ever Born (Rate) by Family Income, Husband's Education, Husband's Occupation, Wife's Labor Force Status, Wife's Age at Marriage, and Race: Wives High School 4 Years and Wives Not a High School Graduate, United States, 1970—Continued

Race	Wife's age at marriage	Wife's labor force status	Husband's occupation	Husband's education	Family income	Wives, high school 4 years (population 1)		Wives, not a high school graduate (population 2)	
						Size	Rate	Size	Rate
n	m	l	k	j	i	N_{ijklmn}	T_{ijklmn}	n_{ijklmn}	T_{ijklmn}
2..........	2	1	2	3	6	182	4.363	15	4.000
2..........	2	1	2	3	7	53	2.000	0	.000
2..........	2	2	1	1	1	153	2.340	309	2.460
2..........	2	2	1	1	2	228	2.706	339	3.634
2..........	2	2	1	1	3	164	.927	264	3.114
2..........	2	2	1	1	4	645	2.505	939	2.440
2..........	2	2	1	1	5	998	2.462	1273	2.684
2..........	2	2	1	1	6	373	2.429	471	3.369
2..........	2	2	1	1	7	84	1.679	110	3.100
2..........	2	2	1	2	1	103	2.330	61	.754
2..........	2	2	1	2	2	42	3.095	146	2.247
2..........	2	2	1	2	3	67	1.657	100	3.040
2..........	2	2	1	2	4	596	2.250	340	1.771
2..........	2	2	1	2	5	2125	2.157	549	2.638
2..........	2	2	1	2	6	1083	2.435	248	3.931
2..........	2	2	1	2	7	231	1.476	137	1.803
2..........	2	2	1	3	1	24	.000	0	.000
2..........	2	2	1	3	2	81	1.198	60	1.050
2..........	2	2	1	3	3	53	2.604	42	4.429
2..........	2	2	1	3	4	458	3.015	168	2.292
2..........	2	2	1	3	5	1210	2.093	524	2.842
2..........	2	2	1	3	6	756	1.907	169	2.734
2..........	2	2	1	3	7	636	2.074	110	1.800
2..........	2	2	2	1	1	1213	2.607	4391	3.465
2..........	2	2	2	1	2	2093	2.972	7044	3.185
2..........	2	2	2	1	3	1175	2.698	4206	2.743
2..........	2	2	2	1	4	5254	2.670	10954	3.145
2..........	2	2	2	1	5	6804	2.544	8990	3.145
2..........	2	2	2	1	6	2121	2.429	2543	2.917
2..........	2	2	2	1	7	458	2.880	649	3.743
2..........	2	2	2	2	1	382	2.927	343	2.525
2..........	2	2	2	2	2	730	2.138	759	2.809
2..........	2	2	2	2	3	341	1.710	529	3.110
2..........	2	2	2	2	4	2454	2.395	1844	3.292
2..........	2	2	2	2	5	4531	2.089	2374	2.714
2..........	2	2	2	2	6	1612	2.441	621	3.150
2..........	2	2	2	2	7	324	1.812	184	1.304
2..........	2	2	2	3	1	23	.000	17	7.000
2..........	2	2	2	3	2	33	1.000	138	3.312
2..........	2	2	2	3	3	21	2.000	140	1.843
2..........	2	2	2	3	4	411	3.333	305	2.639
2..........	2	2	2	3	5	1070	2.271	393	2.852
2..........	2	2	2	3	6	369	3.363	117	2.462
2..........	2	2	2	3	7	80	3.137	64	1.016
n = ...	m = .	l = .	k = .	j = .	i = .	3,369,852	3.005	2,302,030	3.428

Source: U.S. Bureau of the Census (1973), Table 57. i = 1,2,...,7 (family income): less than \$4,000, \$4,000-\$5,999, \$6,000-\$7,999, \$8,000-\$9,999, \$10,000-\$14,999, \$15,000-\$19,999, greater than or equal to \$20,000. j = 1,2,3 (husband's education): not a high school graduate, high school 4 years, college 1 year or more. k = 1,2 (husband's occupation): white collar worker, blue collar or service worker. l = 1,2 (wife's labor force status): not in labor force, in labor force. m = 1,2 (wife's age at marriage): 14 to 21, 22 and over. n = 1,2 (race): White, Black.

Program 5.6 (Six Factors +Rate)

```
      DIMENSION P(8,4,3,3,3,3,2),T(7,3,2,2,2,2,2),
     1 R(2,6,3),U(6),S(2,6),ET(2),ER(2)
      DOUBLE PRECISION P,R,U,S,ET,ER,Q,H,A,W1,W2,W3
      DO 1 N1=1,2
      READ(5,2)  ((((((P(I,J,K,L,M,N,N1),I=1,7),J=1,3),K=1,2),L=1,2),
     1            M=1,2),N=1,2)
    1 READ(5,12) ((((((T(I,J,K,L,M,N,N1),I=1,7),J=1,3),K=1,2),L=1,2),
     1            M=1,2),N=1,2)
    2 FORMAT(7F10.0)
   12 FORMAT(7F10.3)
      DO 10 N1=1,2
      DO 3 J=1,3
      DO 3 K=1,2
      DO 3 L=1,2
      DO 3 M=1,2
      DO 3 N=1,2
      P(8,J,K,L,M,N,N1)=0.0
      DO 3 I=1,7
    3 P(8,J,K,L,M,N,N1)=P(8,J,K,L,M,N,N1)+P(I,J,K,L,M,N,N1)
      DO 4 I=1,8
      DO 4 K=1,2
      DO 4 L=1,2
      DO 4 M=1,2
      DO 4 N=1,2
      P(I,4,K,L,M,N,N1)=0.0
      DO 4 J=1,3
    4 P(I,4,K,L,M,N,N1)=P(I,4,K,L,M,N,N1)+P(I,J,K,L,M,N,N1)
      DO 5 I=1,8
      DO 5 J=1,4
      DO 5 L=1,2
      DO 5 M=1,2
      DO 5 N=1,2
      P(I,J,3,L,M,N,N1)=0.0
      DO 5 K=1,2
    5 P(I,J,3,L,M,N,N1)=P(I,J,3,L,M,N,N1)+P(I,J,K,L,M,N,N1)
      DO 6 I=1,8
      DO 6 J=1,4
      DO 6 K=1,3
      DO 6 M=1,2
      DO 6 N=1,2
      P(I,J,K,3,M,N,N1)=0.0
      DO 6 L=1,2
    6 P(I,J,K,3,M,N,N1)=P(I,J,K,3,M,N,N1)+P(I,J,K,L,M,N,N1)
      DO 7 I=1,8
      DO 7 J=1,4
      DO 7 K=1,3
      DO 7 L=1,3
      DO 7 N=1,2
      P(I,J,K,L,3,N,N1)=0.0
      DO 7 M=1,2
    7 P(I,J,K,L,3,N,N1)=P(I,J,K,L,3,N,N1)+P(I,J,K,L,M,N,N1)
      DO 8 I=1,8
      DO 8 J=1,4
      DO 8 K=1,3
      DO 8 L=1,3
      DO 8 M=1,3
      P(I,J,K,L,M,3,N1)=0.0
      DO 8 N=1,2
    8 P(I,J,K,L,M,3,N1)=P(I,J,K,L,M,3,N1)+P(I,J,K,L,M,N,N1)
   10 CONTINUE
      DO 9 N1=1,2
      ET(N1)=0.0
      ER(N1)=0.0
      DO 9 I=1,7
      DO 9 J=1,3
      DO 9 K=1,2
      DO 9 L=1,2
      DO 9 M=1,2
      DO 9 N=1,2
      ET(N1)=ET(N1)+P(I,J,K,L,M,N,N1)*T(I,J,K,L,M,N,N1)
     1                /P(8,4,3,3,3,3,N1)

      Q=(P(I,J,K,L,M,N,1)/P(8,4,3,3,3,3,1)+
     1   P(I,J,K,L,M,N,2)/P(8,4,3,3,3,3,2))*.5
    9 ER(N1)=ER(N1)+Q*T(I,J,K,L,M,N,N1)
      ETT=ET(2)-ET(1)
      ERR=ER(2)-ER(1)
      DO 13 I=1,2
      DO 13 J=1,6
      DO 13 K=1,3
   13 R(I,J,K)=0.0
      DO 15 II=1,2
      DO 15 JJ=1,2
      DO 15 KK=1,2
      DO 15 LL=1,2
      DO 15 MM=1,2
      DO 15 NN=1,2
      H=0.0
```

Program 5.6 (continued)

```
      DO 14 IS=1,7
      DO 14 JS=1,3
      DO 14 KS=1,2
      DO 14 LS=1,2
      DO 14 MS=1,2
      DO 14 NS=1,2
      W1=1.0
      W2=1.0
      W3=1.0
      DO 11 I1=1,2
      DO 11 I2=1,2
      DO 11 I3=1,2
      DO 11 I4=1,2
      DO 11 I5=1,2
      I6=I1+I2+I3+I4+I5
      IF(I1.EQ.1) J=JS
      IF(I1.EQ.2) J=4
      IF(I2.EQ.1) K=KS
      IF(I2.EQ.2) K=3
      IF(I3.EQ.1) L=LS
      IF(I3.EQ.2) L=3
      IF(I4.EQ.1) M=MS
      IF(I4.EQ.2) M=3
      IF(I5.EQ.1) N=NS
      IF(I5.EQ.2) N=3
      DO 19 NL=1,6
      GO TO (21,22,23,24,25,26),NL
   21 A=P(IS,J,K,L,M,N,II)/P(8,J,K,L,M,N,II)
      GO TO 20
   22 IF(I1.EQ.1) I=IS
      IF(I1.EQ.2) I=8
      A=P(I,JS,K,L,M,N,JJ)/P(I,4,K,L,M,N,JJ)
      GO TO 20
   23 IF(I2.EQ.1) J=JS
      IF(I2.EQ.2) J=4
      A=P(I,J,KS,L,M,N,KK)/P(I,J,3,L,M,N,KK)
      GO TO 20
   24 IF(I3.EQ.1) K=KS
      IF(I3.EQ.2) K=3
      A=P(I,J,K,LS,M,N,LL)/P(I,J,K,3,M,N,LL)
      GO TO 20
   25 IF(I4.EQ.1) L=LS
      IF(I4.EQ.2) L=3
      A=P(I,J,K,L,MS,N,MM)/P(I,J,K,L,3,N,MM)
      GO TO 20
   26 IF(I5.EQ.1) M=MS
      IF(I5.EQ.2) M=3
      A=P(I,J,K,L,M,NS,NN)/P(I,J,K,L,M,3,NN)
   20 IF(I6.EQ.5.OR.I6.EQ.10) W1=W1*A
      IF(I6.EQ.6.OR.I6.EQ. 9) W2=W2*A
   19 IF(I6.EQ.7.OR.I6.EQ. 8) W3=W3*A
   11 CONTINUE
   14 H=H+(T(IS,JS,KS,LS,MS,NS,1)+T(IS,JS,KS,LS,MS,NS,2))*.5
     1   *W1**(1./6.)*W2**(1./30.)*W3**(1./60.)
      DO 15 KL=1,3
      K1=KL+4
      K2=15-K1
      IT=II+JJ+KK+LL+MM+NN
      IF(IT-II.EQ.K1.OR.IT-II.EQ.K2) R(II,1,KL)=R(II,1,KL)+H
      IF(IT-JJ.EQ.K1.OR.IT-JJ.EQ.K2) R(JJ,2,KL)=R(JJ,2,KL)+H
      IF(IT-KK.EQ.K1.OR.IT-KK.EQ.K2) R(KK,3,KL)=R(KK,3,KL)+H
      IF(IT-LL.EQ.K1.OR.IT-LL.EQ.K2) R(LL,4,KL)=R(LL,4,KL)+H
      IF(IT-MM.EQ.K1.OR.IT-MM.EQ.K2) R(MM,5,KL)=R(MM,5,KL)+H
   15 IF(IT-NN.EQ.K1.OR.IT-NN.EQ.K2) R(NN,6,KL)=R(NN,6,KL)+H
      DO 17 J=1,6
      DO 16 I=1,2
   16 S(I,J)=R(I,J,1)/6.+R(I,J,2)/30.+R(I,J,3)/60.
   17 U(J)=S(2,J)-S(1,J)
      WRITE(6,18) (S(2,J),S(1,J),U(J),J=1,6),ER(2),ER(1),ERR,
     1 ET(2),ET(1),ETT
   18 FORMAT(40X,3F15.3)
      STOP
      END
****
```

Table 5.18. Standardization and Decomposition of Average Number of Children Ever Born in Table 5.17

Average number of children ever born	Standardization		Decomposition	
	Not a high school graduate (population 2)	High school 4 years (population 1)	Difference (effects)	Percent distribution of effects
(J,K,L,M,N,R)-standardized average	3.141	3.133	.008 (I)	1.9
(I,K,L,M,N,R)-standardized average	3.163	3.098	.065 (J)	15.4
(I,J,L,M,N,R)-standardized average	3.153	3.116	.037 (K)	8.7
(I,J,K,M,N,R)-standardized average	3.149	3.135	.014 (L)	3.3
(I,J,K,L,N,R)-standardized average	3.169	3.114	.055 (M)	13.0
(I,J,K,L,M,R)-standardized average	3.159	3.121	.038 (N)	9.0
(I,J,K,L,M,N)-standardized average	3.279	3.073	.206 (R)	48.7
Overall average numbers	3.428	3.005	.423 (Total effect)	100.0

where

Z = Product of all ratios with numerators having i_1 and r dots among the subscripts i_2 to i_p, *and the corresponding denominators the same as the numerators except for a dot for* i_1. (5.37)

Again, as in (5.32) and (5.33),

$R(\overline{T}) = (I_1,I_2,\ldots,I_p)$-standardized rate in population 1

$$= \sum_{i_1,i_2,\ldots,i_p} \frac{\frac{n_{i_1 \text{ to } i_p}}{n\ldots\ldots} + \frac{N_{i_1 \text{ to } i_p}}{N\ldots\ldots}}{2} T_{i_1 \text{ to } i_p}, \quad (5.38)$$

$I(\overline{A}) = (I_2,I_3,\ldots,I_p,R)$-standardized rate in population 1

$$= \sum_{i_1,i_2,\ldots,i_p} \frac{t_{i_1 \text{ to } i_p} + T_{i_1 \text{ to } i_p}}{2} \quad (5.39)$$

[Expression (2.47), i.e.,(2.50) xA_1 with additional subscripts i_1 to i_p in each letter].

5.9 THE GENERAL PROGRAM

From Programs 5.1 through 5.6 corresponding to one through six factors (+ rate), a FORTRAN program can be developed for any number of factors higher than six. However, it is not necessary to use different programs for data involving different numbers of factors. A program written for, say, six-factor cross-classified data can be used for any number of factors not exceeding six by changing basically the input and output statements. No changes are necessary in the data files previously created to be used with the specific programs.

Assuming that no one is expected to deal with more than six cross-classified factors, we provide below a program (Program 5.7) for six factors that can be used as a general program for any number of factors up to six. This general program is basically the same as the specific six-factor program (Program 5.6) used for Example 5.9, except that the general program has 12 additional lines (lines 4 through 15) specifying the numbers of categories of the factors and the numbers that denote the marginal totals (i.e., the dots) of the factors. We show below the specific changes in Program 5.7 that will be needed to generate the results corresponding to Examples 5.1 through 5.9 in this chapter with the same data files used before:

Example 5.1 (one factor + rate)

Line 1:	Replace 9,8 and 8,7 in P and T by 14,5 and 13,4
Lines 4-9:	Replace 8,7,6,5,4,2 by 13,1,1,1,1,1
Lines 21,22:	Replace 8F10.0 and 8F10.3 by 13F5.1 and 13F5.1
Line 168:	Replace J=1,6 by J=1,1

Example 5.2 (one factor + rate)

Lines 4-9:	Replace 8,7,6,5,4,2 by 5,1,1,1,1,1
Lines 21,22:	Replace 8F10.0 and 8F10.3 by 5F8.0 and 5F8.3
Line 168:	Replace J=1,6 by J=1,1

Example 5.3 (two factors + rate)

Line 1:	Replace 9,8 and 8,7 in P and T by 12,6 and 11,5
Lines 4-9:	Replace 8,7,6,5,4,2 by 11,2,1,1,1,1
Lines 21,22:	Replace 8F10.0 and 8F10.3 by 11F7.0 and 11F7.3
Line 168:	Replace J=1,6 by J=1,2

Example 5.4 (two factors + rate)

Lines 4-9:	Replace 8,7,6,5,4,2 by 3,2,1,1,1,1
Lines 21,22:	Replace 8F10.0 and 8F10.3 by 6F5.0 and 6F5.2
Line 168:	Replace J=1,6 by J=1,2

Example 5.5 (two factors + rate)

Line 1:	Replace 9,7 and 8,6 in P and T by 11,6 and 10,5
Lines 4-9:	Replace 8,7,6,5,4,2 by 10,7,1,1,1,1
Lines 21,22:	Replace 8F10.0 and 8F10.3 by 10F8.0 and 10F8.2
Line 168:	Replace J=1,6 by J=1,2
Line 170:	Replace 15.3 by 15.2

Example 5.6 (three factors + rate)

Lines 4-9:	Replace 8,7,6,5,4,2 by 7,5,2,1,1,1
Lines 21,22:	Replace 8F10.0 and 8F10.3 by 7F10.0 and 7F10.2
Line 168:	Replace J=1,6 by J=1,3
Line 170:	Replace 15.3 by 15.2

Example 5.7 (four factors + rate)

Lines 4-9:	Replace 8,7,6,5,4,2 by 6,2,6,2,1,1
Lines 21,22:	Replace 8F10.0 and 8F10.3 by 6F10.0 and 6F10.3
Line 168:	Replace J=1,6 by J=1,4

Example 5.8 (five factors + rate)

Lines 4-9:	Replace 8,7,6,5,4,2 by 4,3,3,2,2,1
Lines 21,22:	Replace 8F10.0 and 8F10.3 by 4F10.0 and 4F10.2
Line 168:	Replace J=1,6 by J=1,5
Line 170:	Replace 15.3 by 15.0

Example 5.9 (six factors + rate)

Lines 4-9:	Replace 8,7,6,5,4,2 by 7,3,2,2,2,2
Lines 21,22:	Replace 8F10.0 and 8F10.3 by 7F10.0 and 7F10.3

The dimensions of P and T in line 1 of the general program (Program 5.7) are not made arbitrarily high in order to keep the total load within the capacity of the computer. It is, therefore, sometimes necessary to adjust the numbers depending on the categories of the factors in a particular example, as we did in Examples 5.1, 5.3, and 5.5 above. When the number of factors is less than six, the categories of the nonexistent factors are assumed to be 1 in lines 4 through 9 of the general program, as shown above. Instead of making the numbers of categories of the factors part of the program in lines 4 through 9, they can also be included in the data file to be read in the program.

The standardization and decomposition techniques described in this chapter for rates from cross-classified data can be conveniently used to obtain more formally the results in the two studies, *The Impact of Demographic, Social, and Economic Change on the Distribution of Income*, and *Factors Affecting Black-White Income Differentials: A Decomposition*, by Gordon Green, Paul Ryscavage, and Edward Welniak (U.S. Bureau of the Census, 1992), based on the March CPS (Current Population Survey) data for 1970, 1980, and 1990. A similar formal approach is possible for the study entitled *The Level and Trend of Poverty in the United States, 1939-1979*, by Ross, Danziger, and Smolensky (1987).

Program 5.7 (General Program for up to Six Factors +Rate)

```
      DIMENSION P(9,8,7,6,5,3,2),T(8,7,6,5,4.2,2),
     1 R(2,6,3),U(6),S(2,6),ET(2),ER(2)
      DOUBLE PRECISION P,R,U,S,ET,ER,Q,H,A,W1,W2,W3
      IA=8
      JA=7
      KA=6
      LA=5
      MA=4
      NA=2
      IB=IA+1
      JB=JA+1
      KB=KA+1
      LB=LA+1
      MB=MA+1
      NB=NA+1
      DO 1 N1=1,2
      READ(5,2)  ((((((P(I,J,K,L,M,N,N1),I=1,IA),J=1,JA),K=1,KA),L=1,LA)
     1           ,M=1,MA),N=1,NA)
    1 READ(5,12) ((((((T(I,J,K,L,M,N,N1),I=1,IA),J=1,JA),K=1,KA),L=1,LA)
     1           ,M=1,MA),N=1,NA)
    2 FORMAT(8F10.0)
   12 FORMAT(8F10.3)
      DO 10 N1=1,2
      DO 3 J=1,JA
      DO 3 K=1,KA
      DO 3 L=1,LA
      DO 3 M=1,MA
      DO 3 N=1,NA
      P(IB,J,K,L,M,N,N1)=0.0
      DO 3 I=1,IA
    3 P(IB,J,K,L,M,N,N1)=P(IB,J,K,L,M,N,N1)+P(I,J,K,L,M,N,N1)
      DO 4 I=1,IB
      DO 4 K=1,KA
      DO 4 L=1,LA
      DO 4 M=1,MA
      DO 4 N=1,NA
      P(I,JB,K,L,M,N,N1)=0.0
      DO 4 J=1,JA
    4 P(I,JB,K,L,M,N,N1)=P(I,JB,K,L,M,N,N1)+P(I,J,K,L,M,N,N1)
      DO 5 I=1,IB
      DO 5 J=1,JB
      DO 5 L=1,LA
      DO 5 M=1,MA
      DO 5 N=1,NA
      P(I,J,KB,L,M,N,N1)=0.0
      DO 5 K=1,KA
    5 P(I,J,KB,L,M,N,N1)=P(I,J,KB,L,M,N,N1)+P(I,J,K,L,M,N,N1)
      DO 6 I=1,IB
      DO 6 J=1,JB
      DO 6 K=1,KB
      DO 6 M=1,MA
      DO 6 N=1,NA
      P(I,J,K,LB,M,N,N1)=0.0
      DO 6 L=1,LA
    6 P(I,J,K,LB,M,N,N1)=P(I,J,K,LB,M,N,N1)+P(I,J,K,L,M,N,N1)
      DO 7 I=1,IB
      DO 7 J=1,JB
      DO 7 K=1,KB
      DO 7 L=1,LB
      DO 7 N=1,NA
      P(I,J,K,L,MB,N,N1)=0.0
      DO 7 M=1,MA
    7 P(I,J,K,L,MB,N,N1)=P(I,J,K,L,MB,N,N1)+P(I,J,K,L,M,N,N1)
      DO 8 I=1,IB
      DO 8 J=1,JB
      DO 8 K=1,KB
      DO 8 L=1,LB
      DO 8 M=1,MB
      P(I,J,K,L,M,NB,N1)=0.0
      DO 8 N=1,NA
    8 P(I,J,K,L,M,NB,N1)=P(I,J,K,L,M,NB,N1)+P(I,J,K,L,M,N,N1)
   10 CONTINUE
      DO 9 N1=1,2
      ET(N1)=0.0
      ER(N1)=0.0
      DO 9 I=1,IA
      DO 9 J=1,JA
      DO 9 K=1,KA
      DO 9 L=1,LA
      DO 9 M=1,MA
      DO 9 N=1,NA
      ET(N1)=ET(N1)+P(I,J,K,L,M,N,N1)*T(I,J,K,L,M,N,N1)
     1          /P(IB,JB,KB,LB,MB,NB,N1)
      Q=(P(I,J,K,L,M,N,1)/P(IB,JB,KB,LB,MB,NB,1)+
     1   P(I,J,K,L,M,N,2)/P(IB,JB,KB,LB,MB,NB,2))*.5
    9 ER(N1)=ER(N1)+Q*T(I,J,K,L,M,N,N1)
      ETT=ET(2)-ET(1)
      ERR=ER(2)-ER(1)
      DO 13 I=1,2
      DO 13 J=1,6
      DO 13 K=1,3
   13 R(I,J,K)=0.0
```

Program 5.7 (continued)

```
          DO 15 II=1,2
          DO 15 JJ=1,2
          DO 15 KK=1,2
          DO 15 LL=1,2
          DO 15 MM=1,2
          DO 15 NN=1,2
          H=0.0
          DO 14 IS=1,IA
          DO 14 JS=1,JA
          DO 14 KS=1,KA
          DO 14 LS=1,LA
          DO 14 MS=1,MA
          DO 14 NS=1,NA
          W1=1.0
          W2=1.0
          W3=1.0
          DO 11 I1=1,2
          DO 11 I2=1,2
          DO 11 I3=1,2
          DO 11 I4=1,2
          DO 11 I5=1,2
          I6=I1+I2+I3+I4+I5
          IF(I1.EQ.1) J=JS
          IF(I1.EQ.2) J=JB
          IF(I2.EQ.1) K=KS
          IF(I2.EQ.2) K=KB
          IF(I3.EQ.1) L=LS
          IF(I3.EQ.2) L=LB
          IF(I4.EQ.1) M=MS
          IF(I4.EQ.2) M=MB
          IF(I5.EQ.1) N=NS
          IF(I5.EQ.2) N=NB
          DO 19 NL=1,6
          GO TO (21,22,23,24,25,26),NL
       21 A=P(IS,J,K,L,M,N,II)/P(IB,J,K,L,M,N,II)
          GO TO 20
       22 IF(I1.EQ.1) I=IS
          IF(I1.EQ.2) I=IB
          A=P(I,JS,K,L,M,N,JJ)/P(I,JB,K,L,M,N,JJ)
          GO TO 20
       23 IF(I2.EQ.1) J=JS
          IF(I2.EQ.2) J=JB
          A=P(I,J,KS,L,M,N,KK)/P(I,J,KB,L,M,N,KK)
          GO TO 20
       24 IF(I3.EQ.1) K=KS
          IF(I3.EQ.2) K=KB
          A=P(I,J,K,LS,M,N,LL)/P(I,J,K,LB,M,N,LL)
          GO TO 20
       25 IF(I4.EQ.1) L=LS
          IF(I4.EQ.2) L=LB
          A=P(I,J,K,L,MS,N,MM)/P(I,J,K,L,MB,N,MM)
          GO TO 20
       26 IF(I5.EQ.1) M=MS
          IF(I5.EQ.2) M=MB
          A=P(I,J,K,L,M,NS,NN)/P(I,J,K,L,M,NB,NN)
       20 IF(I6.EQ.5.OR.I6.EQ.10) W1=W1*A
          IF(I6.EQ.6.OR.I6.EQ. 9) W2=W2*A
       19 IF(I6.EQ.7.OR.I6.EQ. 8) W3=W3*A
       11 CONTINUE
       14 H=H+(T(IS,JS,KS,LS,MS,NS,1)+T(IS,JS,KS,LS,MS,NS,2))*.5
         1    *W1**(1./6.)*W2**(1./30.)*W3**(1./60.)
          DO 15 KL=1,3
          K1=KL+4
          K2=15-K1
          IT=II+JJ+KK+LL+MM+NN
          IF(IT-II.EQ.K1.OR.IT-II.EQ.K2) R(II,1,KL)=R(II,1,KL)+H
          IF(IT-JJ.EQ.K1.OR.IT-JJ.EQ.K2) R(JJ,2,KL)=R(JJ,2,KL)+H
          IF(IT-KK.EQ.K1.OR.IT-KK.EQ.K2) R(KK,3,KL)=R(KK,3,KL)+H
          IF(IT-LL.EQ.K1.OR.IT-LL.EQ.K2) R(LL,4,KL)=R(LL,4,KL)+H
          IF(IT-MM.EQ.K1.OR.IT-MM.EQ.K2) R(MM,5,KL)=R(MM,5,KL)+H
       15 IF(IT-NN.EQ.K1.OR.IT-NN.EQ.K2) R(NN,6,KL)=R(NN,6,KL)+H
          DO 17 J=1,6
          DO 16 I=1,2
       16 S(I,J)=R(I,J,1)/6.+R(I,J,2)/30.+R(I,J,3)/60.
       17 U(J)=S(2,J)-S(1,J)
          WRITE(6,18) (S(2,J),S(1,J),U(J),J=1,6),ER(2),ER(1),ERR,
         1 ET(2),ET(1),ETT
       18 FORMAT(40X,3F15.3)
          STOP
          END
****
```

Chapter 6. Three Or More Populations

6.1 INTRODUCTION

The standardization and decomposition discussed in the preceding chapters involve only two populations. In many situations, however, we are interested in comparing three or more populations simultaneously. Clogg and Eliason (1988), for example, considered four parity groups of women and eliminated the effects of their age compositions to obtain the adjusted percentages desiring more children in those groups (Example 6.3). Santi (1989) compared the household headship rates for four years after eliminating the effects of age composition from these rates (Example 6.2). Again, Smith and Cutright (1988) dealt with the problem of standardizing illegitimacy ratios in the United States for five years (Example 6.5).

When there are more than two populations to be compared, we can carry out the same computations more than once by taking two populations at a time. For example, if there are three populations 1, 2, and 3, we can compute three sets of results—between 1 and 2, between 2 and 3, and between 1 and 3. Unfortunately, these three sets of results are not necessarily internally consistent (Das Gupta, 1991).

In order to illustrate the problem of internal inconsistency, let us again consider Example 5.9 discussed in tables 5.17 and 5.18. Let us add one more population, namely, college 1 year or more (say, population 1), to the two existing groups, high school 4 years (population 2) and not a high school graduate (population 3). The three pairwise comparisons, similar to the one in table 5.18, are presented in table 6.1 (which, obviously, includes the results in table 5.18).

Considering the first row in table 6.1, which represents the (J,K,L,M,N,R)-standardized rates and I-effects, we immediately notice two problems as follows:

1. For each population, there are two standardized rates. For example, for population 2, the standardized rates are 2.871 and 3.133. We would like to have only one standardized rate for a population when standardization is done with respect to the same factor or the same set of factors.

2. The I-effect in the comparisons of populations 1 and 2 and populations 2 and 3 are, respectively, .001 and .008. These two numbers add up to .009, which is different from the I-effect .035 in the comparisons of populations 1 and 3. For consistency, we would like to see that these two numbers are identical.

Table 6.1. Standardization and Decomposition of Average Number of Children Ever Born Using 2 Populations at a Time

(See Example 5.9 and tables 5.17-5.18 for the description of the factors and the interpretation of the numbers)

Standardized rates		Decomposition	Standardized rates		Decomposition	Standardized rates		Decomposition
High school 4 years (population 2)	College 1 year or more (population 1)	Difference (effects)	Not a high school graduate (population 3)	College 1 year or more (population 1)	Difference (effects)	Not a high school graduate (population 3)	High school 4 years (population 2)	Difference (effects)
2.871	2.870	.001	2.901	2.866	.035	3.141	3.133	.008
2.846	2.869	−.023	2.869	2.827	.042	3.163	3.098	.065
2.868	2.854	.014	2.893	2.819	.074	3.153	3.116	.037
2.887	2.865	.022	2.919	2.883	.036	3.149	3.135	.014
2.925	2.826	.099	2.992	2.809	.183	3.169	3.114	.055
2.877	2.877	.000	2.906	2.893	.013	3.159	3.121	.038
2.948	2.903	.045	3.154	2.956	.198	3.279	3.073	.206
3.005	2.847	.158	3.428	2.847	.581	3.428	3.005	.423

Table 6.2. Standardization and Decomposition of Average Number of Children Ever Born Using 3 Populations Simultaneously

(Based on the standardized rates in table 6.1 as data)

Standardized rates			Decomposition (effects)		
Not a high school graduate (population 3)	High school 4 years (population 2)	College 1 year or more (population 1)	(population 2) −(population 1)	(population 3) −(population 1)	(population 3) −(population 2)
2.978	2.961	2.952	.009	.026	.017
2.981	2.916	2.939	−.023	.042	.065
2.987	2.943	2.921	.022	.066	.044
2.990	2.976	2.954	.022	.036	.014
3.052	2.987	2.878	.109	.174	.065
2.989	2.960	2.968	−.008	.021	.029
3.187	2.998	2.971	.027	.216	.189
3.428	3.005	2.847	.158	.581	.423

In order to resolve the above two problems for any number of populations N, let us gradually develop formulas starting with three populations. We have used I, J, K,... to denote the factors in case of cross-classified data, and α, β, γ,... to denote them in other situations. From now on, in all cases, we will use α, β, γ,... to denote the factors as well as the factor effects. For cross-classified data, the rate effect will be treated as the effect of one of the factors, so that for a six-factor case, for example, we will have seven factor effects.

6.2 THE CASE OF THREE POPULATIONS

Regardless of how many factors are involved, let us consider only the factor α, since the formulas for other factors will be exactly the same.

Let α_{xy} denote the factor effect of α and $\alpha_{x.y}$ denote the standardized rate in population x controlled for all other factors except α, when only populations x and y are compared. Let $\alpha_{xy.z}$ and $\alpha_{x.yz}$ denote the corresponding numbers when populations x and y are compared in the presence of a third population: population z ($\alpha_{xy} = -\alpha_{yx}$, $\alpha_{xy.z} = -\alpha_{yx.z}$, $\alpha_{x.yz} = \alpha_{x.zy}$).

We already know how to compute the factor effects and the standardized rates when we compare two populations at a time. Therefore, for populations 1, 2, and 3, we can obtain the values of the nine quantities involved in the following three identities:

$$\alpha_{12} = \alpha_{2.1} - \alpha_{1.2}, \quad \alpha_{13} = \alpha_{3.1} - \alpha_{1.3}, \quad \alpha_{23} = \alpha_{3.2} - \alpha_{2.3}. \tag{6.1}$$

In order to have one standardized rate for each population and also internally consistent numbers, we want to replace the nine numbers in (6.1) by six numbers that will satisfy the following three identities:

$$\alpha_{12.3} = \alpha_{2.13} - \alpha_{1.23}, \quad \alpha_{13.2} = \alpha_{3.12} - \alpha_{1.23}, \quad \alpha_{23.1} = \alpha_{3.12} - \alpha_{2.13}. \tag{6.2}$$

One way of achieving this is to substitute

$$\alpha_{12.3} = \alpha_{12}, \quad \alpha_{13.2} = \alpha_{13}, \quad \alpha_{23.1} = \alpha_{13} - \alpha_{12},$$

$$\alpha_{1.23} = \alpha_{1.2}, \quad \alpha_{2.13} = \alpha_{2.1}, \quad \alpha_{3.12} = \alpha_{1.2} + (\alpha_{3.1} - \alpha_{1.3}). \tag{6.3}$$

There are, in fact, six possible ways we can revise the values in (6.1) in order to remove the two limitations inherent in these numbers. These six sets of numbers are shown in section A.4 in appendix A. Taking the average over the six sets, we finally obtain the standardized rate $\alpha_{1.23}$ and the factor effect $\alpha_{12.3}$ as

$$\alpha_{1.23} = \frac{\alpha_{1.2} + \alpha_{1.3}}{2} + \frac{(\alpha_{2.3} - \alpha_{2.1}) + (\alpha_{3.2} - \alpha_{3.1})}{6}, \tag{6.4}$$

$$\alpha_{12.3} = \alpha_{12} - \frac{\alpha_{12} + \alpha_{23} - \alpha_{13}}{3}. \tag{6.5}$$

Other standardized rates and factor effects can be obtained from (6.4) and (6.5) by interchanging the subscripts and/or replacing α by other factors. Equation (6.5) was given in Das Gupta (1991, equation 28).

Example 6.1

Let us again consider the expanded version of Example 5.9 presented in table 6.1. Obviously, this is a case of three populations and seven factors. We have already demonstrated in section 6.1 that the numbers in table 6.1 are not internally consistent. In order to obtain a consistent set of standardized rates and factor effects from the numbers in table 6.1, we use the formulas in (6.4) and (6.5), and present the computed values in table 6.2. For example, using the first line in table 6.1 corresponding to the factor α, we have

$$\alpha_{1.23} = \frac{2.870+2.866}{2} + \frac{(3.133-2.871)+(3.141-2.901)}{6} = 2.952,$$

$$\alpha_{12.3} = .001 - \frac{.001+.008-.035}{3} = .009. \tag{6.6}$$

Obviously, the numbers in table 6.2 do not have the two limitations mentioned in section 6.1. First, each population has now only one set of standardized rates, instead of two sets shown in table 6.1. Also, for any of the factors, the effects corresponding to populations (1, 2) and populations (2, 3) now add up to the effect corresponding to populations (1, 3), unlike the situation in table 6.1. For example, for the factor α in table 6.2, .009 + .017 = .026. We should also note that the revised numbers in table 6.2 based on the simultaneous treatment of the three populations preserve by and large the patterns and the characteristics of the unrevised numbers in table 6.1. For example, for unrevised numbers in table 6.1, the factor effects in the comparison of populations 1 and 3 are, in order of their magnitude, .198, .183, .074, .042, .036, .035, and .013. For the revised numbers in table 6.2, the corresponding values are .216, .174, .066, .042, .036, .026, and .021.

Program 6.1

The results in table 6.2 can be obtained by using Program 6.1 in which S(I,J,K)'s are the standardized rates and R(J)'s are the crude rates in table 6.1. In other words, the data file consists of seven lines. The first six lines are the six sets of standardized rates in table 6.1 in the same order, each line having seven numbers with the format specified in line 8 of the program. The last line of the data file consists of three numbers corresponding to the average numbers of children ever born in populations 1, 2, and 3, respectively, with the same format in line 8. M and N in lines 2 and 3 of the program are, respectively, the number of factors (including the rate) and the number of populations in this particular example. Program 6.1, when run with the data file described above, will generate the six columns of results shown in table 6.2.

6.3 THE CASE OF FOUR POPULATIONS

Using analogous notation, for a particular factor, there are 48 different ways the unrevised 12 standardized rates and six factor effects can be replaced to form a revised consistent set of four standardized rates and six effects. These 48 sets of consistent numbers are shown in section A.5 in appendix A. The averages over the 48 sets give us the following expressions for the standardized rate $\alpha_{1.234}$ and the factor effect $\alpha_{12.34}$:

$$\alpha_{1.234} = \frac{\alpha_{1.2}+\alpha_{1.3}+\alpha_{1.4}}{3} + \frac{(\alpha_{2.3}+\alpha_{2.4}-2\alpha_{2.1}) + (\alpha_{3.2}+\alpha_{3.4}-2\alpha_{3.1}) + (\alpha_{4.2}+\alpha_{4.3}-2\alpha_{4.1})}{12}, \tag{6.7}$$

$$\alpha_{12.34} = \alpha_{12} - \frac{(\alpha_{12}+\alpha_{23}-\alpha_{13}) + (\alpha_{12}+\alpha_{24}-\alpha_{14})}{4}. \tag{6.8}$$

Equation (6.8) was given in Das Gupta (1991, equation 30).

Program 6.1 (More than Two Populations)

```
      DIMENSION S(20,20,20),R(20),DT(20,20,20),DR(20,20),T(20,20)
      M=7
      N=3
      Z=N
      DO 1 K=1,N-1
      DO 1 J=K+1,N
    1 READ(5,2) (S(I,J,K),I=1,M),(S(I,K,J),I=1,M)
    2 FORMAT(7F8.3)
      READ(5,2) (R(J),J=1,N)
      DO 5 I=1,M
      DO 5 J=1,N
      AA=0.0
      BB=0.0
      CC=0.0
      DO 4 K=1,N
      IF(K.EQ.J) GO TO 3
      AA=AA+S(I,J,K)
      CC=CC+(Z-2.)*S(I,K,J)
    3 DO 4 JJ=1,N
      IF(JJ.EQ.J.OR.K.EQ.J.OR.K.EQ.JJ) GO TO 4
      BB=BB+S(I,JJ,K)
    4 CONTINUE
    5 T(I,J)=AA/(Z-1.)+(BB-CC)/(Z*(Z-1.))
      DO 6 K=1,N-1
      DO 6 J=K+1,N
      DR(J,K)=R(J)-R(K)
      DO 6 I=1,M
    6 DT(I,J,K)=T(I,J)-T(I,K)
      WRITE(6,7) ((T(I,N+1-J),J=1,N),I=1,M),(R(N+1-J),J=1,N)
    7 FORMAT(10X,3F8.3)
      WRITE(6,8)
    8 FORMAT(//)
      WRITE(6,9) (((DT(I,J,K),J=K+1,N),K=1,N-1),I=1,M),
     1          ((DR(  J,K),J=K+1,N),K=1,N-1)
    9 FORMAT(10X,3F8.3)
      STOP
      END
```

****

Program 6.2 (Combined Program for Example 6.5)

```
      DIMENSION W(5,4,6),V(5,4,6),R(2,4,2),S(4,5,5),U(5),DT(4,5,5),
     1 DR(5,5),T(4,5)
      READ(5,1) (((W(I,J,K),K=1,6),J=1,4),I=1,5)
    1 FORMAT(6F6.3)
      DO 10 KK=1,4
      DO 10 JJ=KK+1,5
      DO 20 J=1,4
      DO 20 K=1,6
      V(1,J,K)=W(KK,J,K)
   20 V(2,J,K)=W(JJ,J,K)
      DO 2 I=1,2
      DO 2 J=1,4
      DO 2 K=1,2
    2 R(I,J,K)=0.0
      DO 3 I=1,2
      DO 3 J=1,2
      DO 3 K=1,2
      DO 3 L=1,2
      H1=0.0
      H2=0.0
      DO 7 M1=1,6
      H1=H1+V(I,1,M1)*V(J,2,M1)*V(K,3,M1)
    7 H2=H2+V(I,1,M1)*(1.-V(J,2,M1))*V(L,4,M1)
      H=H1/(H1+H2)
      IF(I+J+K+L.EQ.4) U(KK)=H
      IF(I+J+K+L.EQ.8) U(JJ)=H
      DO 3 M=1,2
      M1=M+2
      M2=9-M1
      IF(J+K+L.EQ.M1.OR.J+K+L.EQ.M2) R(I,1,M)=R(I,1,M)+H
      IF(I+K+L.EQ.M1.OR.I+K+L.EQ.M2) R(J,2,M)=R(J,2,M)+H
      IF(I+J+L.EQ.M1.OR.I+J+L.EQ.M2) R(K,3,M)=R(K,3,M)+H
    3 IF(I+J+K.EQ.M1.OR.I+J+K.EQ.M2) R(L,4,M)=R(L,4,M)+H
      DO 4 I=1,4
      S(I,KK,JJ)=R(1,I,1)/4.+R(1,I,2)/12.
    4 S(I,JJ,KK)=R(2,I,1)/4.+R(2,I,2)/12.
   10 CONTINUE
      DO 5 I=1,4
      DO 5 J=1,5
      AA=0.0
      BB=0.0
      CC=0.0
      DO 40 K=1,5
      IF(K.EQ.J) GO TO 30
      AA=AA+S(I,J,K)
      CC=CC+3.*S(I,K,J)
   30 DO 40 JJ=1,5
      IF(JJ.EQ.J.OR.K.EQ.J.OR.K.EQ.JJ) GO TO 40
      BB=BB+S(I,JJ,K)
   40 CONTINUE
    5 T(I,J)=AA/4.+(BB-CC)/20.
      DO 6 K=1,4
      DO 6 J=K+1,5
      DR(J,K)=U(J)-U(K)
      DO 6 I=1,4
    6 DT(I,J,K)=T(I,J)-T(I,K)
      WRITE(6,70) ((T(I,6-J),J=1,5),I=1,4),(U(6-J),J=1,5)
   70 FORMAT(10X,5F10.5)
      WRITE(6,8)
    8 FORMAT(//)
      WRITE(6,9) (((DT(I,J,K),J=K+1,5),K=1,4),I=1,4),
     1          ((DR(  J,K),J=K+1,5),K=1,4)
    9 FORMAT(10X,10F10.5)
      STOP
      END
```

****

Example 6.2

Let us again consider Example 5.1 (tables 5.1 and 5.2) based on the data from Santi (1989) using four populations corresponding to the years 1970, 1975, 1980, and 1985 simultaneously. The six sets of standardized rates and factor effects from pairwise comparisons are presented in table 6.3. Table 6.4 gives the corresponding revised numbers obtained by using formulas (6.7) and (6.8). For example, the age-standardized headship rates for 1970, 1975, 1980, and 1985 are, respectively, 44.955, 46.300, 47.307, and 47.645. Santi provided two sets of these adjusted rates in table 5 of his paper. The CG-Purged rates are 44.728, 46.357, 47.526, and 46.726, and the CD-Purged rates are 46.294, 47.930, 49.103, and 48.300. All three sets of adjusted rates have very similar patterns. It is interesting to note that although the crude headship rate for 1985 is higher than that for 1980, the adjusted rate for 1980 is the highest in each of the three sets.

Table 6.3. Standardization and Decomposition of Household Headship Rates Using 2 Populations at a Time

(See Example 5.1 and tables 5.1-5.2 for the description of the factors and the interpretation of the numbers)

Standardized rates		Decomposition	Standardized rates		Decomposition
1975 (population 2)	1970 (population 1)	Difference (effects)	1980 (population 3)	1970 (population 1)	Difference (effects)
45.007	45.372	−.375	45.977	45.883	.094
45.846	44.534	1.312	47.097	44.762	2.335
45.674	44.727	.947	47.156	44.727	2.429
1985 (population 4)	1970 (population 1)	Difference (effects)	1980 (population 3)	1975 (population 2)	Difference (effects)
46.815	45.588	1.227	46.658	46.162	.496
47.071	45.331	1.740	46.903	45.917	.986
47.694	44.727	2.967	47.156	45.674	1.482
1985 (population 4)	1975 (population 2)	Difference (effects)	1985 (population 4)	1980 (population 3)	Difference (effects)
47.507	45.819	1.688	48.034	46.797	1.237
46.829	46.497	.332	47.066	47.765	−.699
47.694	45.674	2.020	47.694	47.156	.538

Table 6.4. Standardization and Decomposition of Household Headship Rates Using 4 Populations Simultaneously

(Based on the standardized rates in table 6.3 as data)

Standardized rates			
1985 (population 4)	1980 (population 3)	1975 (population 2)	1970 (population 1)
47.340	46.140	45.665	46.063
46.645	47.307	46.300	44.955
47.694	47.156	45.674	44.727

Decomposition (effects)					
(population 2) −(population 1)	(population 3) −(population 1)	(population 4) −(population 1)	(population 3) −(population 2)	(population 4) −(population 2)	(population 4) −(population 3)
−.398	.077	1.277	.475	1.675	1.200
1.345	2.352	1.690	1.007	.345	−.662
.947	2.429	2.967	1.482	2.020	.538

The results in table 6.4 can be obtained by using Program 6.1 by making the following changes in the program:

1. Replace M=7 and N=3 in lines 2 and 3 by M=2 and N=4
2. Replace 7F8.3, 3F8.3, and 3F8.3 in lines 8, 30, and 35 by, respectively, 4F8.3, 4F8.3, and 6F8.3 .

The data file consists of seven lines of which the first six lines are the six sets of four standardized rates in table 6.3. For example, line 1 has the four numbers 45.007, 45.846, 45.372, and 44.534 in this order. The last line of the data file consists of four numbers corresponding to the household headship rates for 1970, 1975, 1980, and 1985.

Example 6.3

We now consider an expanded version of Example 5.2 (tables 5.3 and 5.4) based on the data from Clogg and Eliason (1988) for four parity groups 1, 2, 3, and 4+ (designated as populations 4, 3, 2, and 1, respectively). The unrevised six sets of standardized rates and factor effects using two populations at a time are presented in table 6.5. The corresponding revised numbers obtained by using the four populations simultaneously are given in table 6.6. The age-standardized percents desiring more children for the parity groups 1, 2, 3, and 4+ are, respectively, 57.805, 23.460, 18.993, and 18.512. Table 3 of the paper by Clogg and Eliason gave these adjusted numbers as 57.7, 20.1, 18.2, and 16.9, respectively. These two sets of adjusted rates are in good agreement particularly when the corresponding crude percentages are as widely different as 72.093, 26.065, 16.431, and 11.489.

The results in table 6.6 can be obtained by using Program 6.1 by making the following changes in the program (which are the same as the changes in the case of Example 6.2):

1. Replace M=7 and N=3 in lines 2 and 3 by M=2 and N=4
2. Replace 7F8.3, 3F8.3, and 3F8.3 in lines 8, 30, and 35 by, respectively, 4F8.3, 4F8.3, and 6F8.3 .

Again, the data file consists of seven lines of which the first six lines are the six sets of four standardized rates in table 6.5. For example, line 1 has the four numbers 15.418, 14.747, 12.276, and 12.947 in this order. The last line of the data file consists of four numbers corresponding to the percents desiring more children for parity groups 4+, 3, 2, and 1.

Example 6.4

Yet another example of the case of four populations is the expanded version of Example 3.5 (tables 3.9 and 3.10) based on the data from Wojtkiewicz, McLanhan, and Garfinkel (1990) for the four years 1950, 1960, 1970, and 1980. The unrevised six sets of standardized rates and factor effects using two populations at a time are presented in table 6.7. The corresponding revised numbers obtained by using the four populations simultaneously are given in table 6.8. Obviously, these numbers are internally consistent. For example, in the first line of the effects, 3.62, 2.88, and 1.97 add up to 8.47, as they should. Also the revised numbers display the same patterns as do the unrevised numbers based on pairwise comparisons. For example, the unrevised factor effects in the comparison of 1950 and 1980 are 8.72, 22.78, -0.58, -1.46, 0.34, and 2.52, which change to 8.47, 24.11, -1.37, -1.64, 0.19, and 2.56 in the revised set, the total for each set of numbers being 32.32.

The results in table 6.8 can be obtained by using Program 6.1 by making the following changes in the program:

1. Replace M=7 and N=3 in lines 2 and 3 by M=6 and N=4
2. Replace 7F8.3, 3F8.3, and 3F8.3 in lines 8, 30, and 35 by, respectively, 12F6.2, 4F8.2, and 6F8.2 .

The data file consists of seven lines of which the first six lines are the six sets of 12 standardized rates in table 6.7. For example, line 1 consists of the 12 standardized rates in the first two columns (corresponding to 1960 and 1950) in table 6.7. The last line of the data file has four numbers that are the family headship rates for 1950, 1960, 1970, and 1980, respectively.

Table 6.5. Standardization and Decomposition of Percents Desiring More Children Using 2 Populations at a Time

(See Example 5.2 and tables 5.3-5.4 for the description of the factors and the interpretation of the numbers)

Standardized rates		Decomposition	Standardized rates		Decomposition
Parity 3 (population 2)	Parity 4+ (population 1)	Difference (effects)	Parity 2 (population 3)	Parity 4+ (population 1)	Difference (effects)
15.418	12.276	3.142	22.380	13.194	9.186
14.747	12.947	1.800	20.482	15.092	5.390
16.431	11.489	4.942	26.065	11.489	14.576
Parity 1 (population 4)	Parity 4+ (population 1)	Difference (effects)	Parity 2 (population 3)	Parity 3 (population 2)	Difference (effects)
48.619	25.547	23.072	23.078	18.201	4.877
55.849	18.317	37.532	23.018	18.261	4.757
72.093	11.489	60.604	26.065	16.431	9.634
Parity 1 (population 4)	Parity 3 (population 2)	Difference (effects)	Parity 1 (population 4)	Parity 2 (population 3)	Difference (effects)
48.635	32.813	15.822	53.551	42.600	10.951
60.644	20.804	39.840	65.614	30.537	35.077
72.093	16.431	55.662	72.093	26.065	46.028

Table 6.6. Standardization and Decomposition of Percents Desiring More Children Using 4 Populations Simultaneously

(Based on the standardized rates in table 6.5 as data)

Standardized rates			
Parity 1 (population 4)	Parity 2 (population 3)	Parity 3 (population 2)	Parity 4+ (population 1)
42.154	30.471	25.304	20.843
57.805	23.460	18.993	18.512
72.093	26.065	16.431	11.489

Decomposition (effects)					
(population 2) −(population 1)	(population 3) −(population 1)	(population 4) −(population 1)	(population 3) −(population 2)	(population 4) −(population 2)	(population 4) −(population 3)
4.461	9.628	21.311	5.167	16.850	11.683
.481	4.948	39.293	4.467	38.812	34.345
4.942	14.576	60.604	9.634	55.662	46.028

Table 6.7. Standardization and Decomposition of Family Headship Rates Using 2 Populations at a Time

(See Example 3.5 and tables 3.9-3.10 for the description of the factors and the interpretation of the numbers)

Standardized rates		Decomposition	Standardized rates		Decomposition	Standardized rates		Decomposition
1960 (population 2)	1950 (population 1)	Difference (effects)	1970 (population 3)	1950 (population 1)	Difference (effects)	1980 (population 4)	1950 (population 1)	Difference (effects)
28.88	25.42	3.46	35.38	28.98	6.40	42.03	33.31	8.72
28.13	26.20	1.93	37.55	26.76	10.79	49.14	26.36	22.78
28.69	25.60	3.09	33.43	31.15	2.28	37.84	38.42	−0.58
27.53	26.85	0.68	32.29	32.51	−0.22	37.43	38.89	−1.46
27.28	27.14	0.14	32.45	32.31	0.14	38.21	37.87	0.34
27.25	27.17	0.08	32.97	31.79	1.18	39.25	36.73	2.52
32.08	22.70	9.38	43.27	22.70	20.57	55.02	22.70	32.32
1970 (population 3)	1960 (population 2)	Difference (effects)	1980 (population 4)	1960 (population 2)	Difference (effects)	1980 (population 4)	1970 (population 3)	Difference (effects)
39.06	36.27	2.79	46.19	41.40	4.79	50.19	48.40	1.79
42.76	32.67	10.09	56.09	32.16	23.93	56.54	42.32	14.22
36.96	38.54	−1.58	41.44	46.99	−5.55	47.35	51.49	−4.14
37.18	38.27	−1.09	42.72	45.30	−2.58	48.58	50.09	−1.51
37.62	37.85	−0.23	43.82	44.15	−0.33	49.37	49.25	0.12
38.32	37.11	1.21	45.22	42.54	2.68	49.93	48.66	1.27
43.27	32.08	11.19	55.02	32.08	22.94	55.02	43.27	11.75

Table 6.8. Standardization and Decomposition of Family Headship Rates Using 4 Populations Simultaneously

(Based on the standardized rates in table 6.7 as data)

Standardized rates			
1980 (population 4)	1970 (population 3)	1960 (population 2)	1950 (population 1)
41.78	39.81	36.93	33.31
53.29	39.72	30.03	29.18
35.59	39.37	40.71	36.96
36.75	38.19	39.23	38.39
38.14	38.05	38.28	37.95
39.69	38.35	37.12	37.13
55.02	43.27	32.08	22.70

Decomposition (effects)					
(population 2) −(population 1)	(population 3) −(population 1)	(population 4) −(population 1)	(population 3) −(population 2)	(population 4) −(population 2)	(population 4) −(population 3)
3.62	6.50	8.47	2.88	4.85	1.97
0.85	10.54	24.11	9.69	23.26	13.57
3.75	2.41	−.37	−1.34	−5.12	−3.78
0.84	−0.20	−1.64	−1.04	−2.48	−.44
0.33	0.10	0.19	−0.23	−0.14	0.09
−0.01	1.22	2.56	1.23	2.57	1.34
9.38	20.57	32.32	11.19	22.94	11.75

6.4 THE CASE OF FIVE POPULATIONS

Using analogous notation and proceeding as in sections A.4 and A.5 in the appendix, it is easy to show that the standardized rate $\alpha_{1.2345}$ and the factor effect $\alpha_{12.345}$ in five populations have the expressions

$$\alpha_{1.2345} = \frac{\sum_{i=2}^{5} \alpha_{1.i}}{4} + \frac{\sum_{i=2}^{5}[\sum_{j\neq 1,i}^{5} \alpha_{i.j} - 3\alpha_{i.1}]}{20}, \tag{6.9}$$

$$\alpha_{12.345} = \alpha_{12} - \frac{\sum_{j=3}^{5}(\alpha_{12}+\alpha_{2j}-\alpha_{1j})}{5}. \tag{6.10}$$

Example 6.5

Let us again consider Example 4.4 (tables 4.7 and 4.8) based on the data for illegitimacy ratios from Smith and Cutright (1988) using five populations corresponding to the years 1963, 1968, 1973, 1978, and 1983 simultaneously. The 10 sets of standardized rates and factor effects from pairwise comparisons are presented in table 6.9. Table 6.10 gives the corresponding revised numbers obtained by using formulas (6.9) and (6.10). These numbers are self-explanatory. It may be noted that the factor effects of -6.20, 48.66, 27.06, and 24.71 in table 4.8 for a comparison between 1963 and 1983 are now replaced by -8.18, 51.11, 32.00, and 19.30, respectively, in table 6.10, the total difference between the illegitimacy ratios in 1963 and 1983 being 94.23.

The results in table 6.10 can be obtained by using Program 6.1 by making the following changes in the program:

1. Replace M=7 and N=3 in lines 2 and 3 by M=4 and N=5
2. Replace 7F8.3, 3F8.3, and 3F8.3 in lines 8, 30, and 35 by, respectively, 8F7.2, 5F8.2, and 10F8.2 .

The data file consists of 11 lines of which the first 10 lines are the 10 sets of eight standardized rates in table 6.9. For example, line 1 consists of the eight standardized rates in the first two columns (corresponding to 1968 and 1963) in table 6.9. The last line of the data file has five numbers that are the illegitimacy ratios for 1963, 1968, 1973, 1978, and 1983, respectively.

6.5 THE COMBINED PROGRAM

In Examples 6.1 through 6.5, the final results are obtained in two steps by using two separate computer programs. In the first step, the basic data for several years are used as input to compute the standardized rates and the factor effects for all possible pairwise comparisons. In the second step, the computed standardized rates in the first step are used as input to finally obtain the revised set of standardized rates and factor effects. In Example 6.5, for example, the data for five years (1963, 1968, 1973, 1978, and 1983), similar to those given in table 4.7 for 1963 and 1983, are used as input in Program 4.4 to obtain 10 sets of standardized rates in table 6.9, similar to the set in table 4.8. These 10 sets of standardized rates are then used as input in Program 6.1 (for M=4 and N=5) to obtain the final results in table 6.10.

For any particular example, the two computer programs for the two steps can be easily combined into one so that the final results can be obtained directly by using the basic data as the input, without the explicit feeding of the second set of input data. For Example 6.5, Program 6.2 is such a combined program, which, obviously, is the combination of Programs 4.4 and 6.1. Program 6.2, when used with the data file created from the data for five years given in table 6.11, will generate results identical with those in table 6.10 (except that the standardized illegitimacy ratios will now be for each birth, instead of 1,000 births). The data file, based on table 6.11, consists of 20 lines, each year occupying four lines corresponding to four columns of six numbers.

6.6 THE GENERAL CASE OF N POPULATIONS (INCLUDING TIME SERIES)

It is obvious from equations (6.4) through (6.10) that the standardized rate and the factor effect for N populations can be written as

Table 6.9. Standardization and Decomposition of Illegitimacy Ratios Using 2 Populations at a Time

(See Example 4.4 and tables 4.7-4.8 for the description of the factors and the interpretation of the numbers)

Standardized rates		Decomposition	Standardized rates		Decomposition
1968 (population 2)	1963 (population 1)	Difference (effects)	1973 (population 3)	1963 (population 1)	Difference (effects)
41.84	40.87	0.97	46.61	46.44	0.17
43.74	38.71	5.03	50.26	42.29	7.97
44.77	37.65	7.12	47.52	45.63	1.89
45.75	36.60	9.15	57.14	35.15	21.99
53.22	30.95	22.27	62.97	30.95	32.02
1978 (population 4)	1963 (population 1)	Difference (effects)	1983 (population 5)	1963 (population 1)	Difference (effects)
56.80	58.37	−1.57	71.51	77.71	−6.20
69.27	44.02	25.25	96.08	47.42	48.66
61.41	53.38	8.03	86.30	59.24	27.06
68.46	44.23	24.23	84.34	59.63	24.71
86.89	30.95	55.94	125.18	30.95	94.23
1973 (population 3)	1968 (population 2)	Difference (effects)	1978 (population 4)	1968 (population 2)	Difference (effects)
58.05	59.09	−1.04	68.81	72.57	−3.76
60.00	57.08	2.92	81.86	58.99	22.87
55.09	62.48	−7.39	70.76	71.06	−0.30
66.28	51.02	15.26	77.63	62.77	14.86
62.97	53.22	9.75	86.89	53.22	33.67
1983 (population 5)	1968 (population 2)	Difference (effects)	1978 (population 4)	1973 (population 3)	Difference (effects)
84.10	94.37	−10.27	73.82	76.64	−2.82
112.35	62.94	49.41	85.18	65.10	20.08
99.04	77.30	21.74	80.12	70.13	9.99
93.47	82.39	11.08	73.57	76.90	−3.33
125.18	53.22	71.96	86.89	62.97	23.92
1983 (population 5)	1973 (population 3)	Difference (effects)	1983 (population 5)	1978 (population 4)	Difference (effects)
89.61	99.62	−10.01	102.16	109.89	−7.73
116.71	70.21	46.50	117.57	93.53	24.04
112.00	74.90	37.10	120.58	90.58	30.00
89.05	100.43	−11.38	102.08	110.10	−8.02
125.18	62.97	62.21	125.18	86.89	38.29

$$\alpha_{1.23\ldots N} = \frac{\sum_{i=2}^{N} \alpha_{1.i}}{N-1} + \frac{\sum_{i=2}^{N} \left[\sum_{j \neq 1,i}^{N} \alpha_{i.j} - (N-2)\,\alpha_{i.1} \right]}{N(N-1)}, \tag{6.11}$$

$$\alpha_{12.34\ldots N} = \alpha_{12} - \frac{\sum_{j=3}^{N} (\alpha_{12} + \alpha_{2j} - \alpha_{1j})}{N}. \tag{6.12}$$

Equation (6.12) was given in Das Gupta (1991, equation 31).

The above general formulas in (6.11) and (6.12) can be conveniently used to handle the problems of standardization and decomposition when time-series data are involved. The following two examples deal with the revision of age-sex-adjusted birth rates and age-adjusted death rates for the period 1940-1990 provided by the National Center for Health Statistics (1990a, table 1-3; 1990b, table 1-3). Curtin, Maurer,

Table 6.10. Standardization and Decomposition of Illegitimacy Ratios Using 5 Populations Simultaneously

(Based on the standardized rates in table 6.9 as data)

Standardized rates				
1983 (population 5)	1978 (population 4)	1973 (population 3)	1968 (population 2)	1963 (population 1)
64.59	71.35	73.83	74.65	72.77
104.39	79.50	59.53	56.63	53.28
94.18	68.54	60.48	69.61	62.18
74.13	79.61	81.24	64.44	54.83
125.18	86.89	62.97	53.22	30.95
Decomposition (effects)				
(population 2) –(population 1)	(population 3) –(population 1)	(population 4) –(population 1)	(population 5) –(population 1)	(population 3) –(population 2)
1.88	1.06	-1.42	-8.18	-0.82
3.35	6.25	26.22	51.11	2.90
7.43	-1.70	6.36	32.00	-9.13
9.61	26.41	24.78	19.30	16.80
22.27	32.02	55.94	94.23	9.75
(population 4) –(population 2)	(population 5) –(population 2)	(population 4) –(population 3)	(population 5) –(population 3)	(population 5) –(population 4)
-3.30	-10.06	-2.48	-9.24	-6.76
22.87	47.76	19.97	44.86	24.89
-1.07	24.57	8.06	33.70	25.64
15.17	9.69	-1.63	-7.11	-5.48
33.67	71.96	23.92	62.21	38.29

and Rosenberg (1980); Johansen (1990); and many authors have thoroughly examined whether the National Center for Health Statistics (NCHS) should continue to use the 1940 U.S. population as the standard for the computation of age-sex-adjusted birth rates and age-adjusted death rates, or replace it by the U.S. population of a more recent year. Although they have made specific recommendations on this issue, the theoretical question of the validity of the standardized rates (as computed presently by using one of the real populations as the standard) as measures of composition-controlled relative rates has not been adequately addressed.

In computing the age-adjusted death rates, for example, the age-specific death rate-adjusted death rates should also be considered side by side, and we should make sure that these two sets of adjusted death rates are internally consistent from the point of view of the decomposition of the difference between the crude death rates for any two years into the age effect and the rate effect, as explained in section 2.1 (internal inconsistencies of the type indicated in section 6.1 do not arise when there is only one age-adjusted death rate and only one rate-adjusted death rate for any year). A simple direct standardization by using a single population (say, for 1940 or for 1990) as the standard will not pass this test. The answer lies in formula (6.11) where the final standardized number is a composite of the standardized rates based on all possible pairwise comparisons of the given populations, as demonstrated in the following two examples.

Example 6.6

Table 6.12 gives the populations in thousands and the corresponding birth rates per 1,000 population in nine age-sex-groups for the 51 years 1940-1990 for the United States. The rate-adjusted birth rates and the age-sex-adjusted birth rates for these years, based on formula (6.11), are shown, along with the crude birth rates, in columns (2) through (4) of table 6.13.

The age-sex-adjusted birth rates in column (4) of table 6.13 are uniformly **lower** than the corresponding adjusted rates for all 51 years provided by the NCHS based on the 1940 population as the standard (figure 1). Since we study the relative magnitudes of the adjusted rates rather than their absolute magnitudes, the

Table 6.11. Illegitimacy Ratio as a Function of Four Vector-Factors: United States, Whites, 1963, 1968, 1973, 1978, and 1983

(For explanation of notation and source of data, see Example 4.4 and Table 4.7)

Age groups	i	W_i/W	U_i/W_i	I_i/U_i	L_i/M_i	Illegitimacy ratio (R)
		1963 (population 1)				
15 to 19	1	.200	.866	.007	.454	
20 to 24	2	.163	.325	.021	.326	
25 to 29	3	.146	.119	.023	.195	.03095
30 to 34	4	.154	.099	.015	.107	
35 to 39	5	.168	.099	.008	.051	
40 to 44	6	.169	.121	.002	.015	
		1968 (population 2)				
15 to 19	1	.215	.891	.010	.433	
20 to 24	2	.191	.373	.023	.249	
25 to 29	3	.156	.124	.023	.159	.05322
30 to 34	4	.137	.100	.015	.079	
35 to 39	5	.144	.107	.008	.037	
40 to 44	6	.157	.127	.002	.011	
		1973 (population 3)				
15 to 19	1	.218	.870	.011	.314	
20 to 24	2	.203	.396	.016	.181	
25 to 29	3	.175	.158	.017	.133	.06297
30 to 34	4	.144	.125	.011	.063	
35 to 39	5	.127	.113	.006	.023	
40 to 44	6	.133	.129	.002	.006	
		1978 (population 4)				
15 to 19	1	.205	.900	.014	.313	
20 to 24	2	.200	.484	.019	.191	
25 to 29	3	.181	.243	.015	.143	
30 to 34	4	.162	.176	.010	.069	.08689
35 to 39	5	.134	.155	.005	.021	
40 to 44	6	.118	.168	.001	.004	
		1983 (population 5)				
15 to 19	1	.169	.931	.018	.380	
20 to 24	2	.195	.563	.026	.201	
25 to 29	3	.190	.311	.023	.149	.12518
30 to 34	4	.174	.216	.016	.079	
35 to 39	5	.150	.199	.008	.025	
40 to 44	6	.122	.191	.002	.006	

fact that the NCHS rates are always higher than the present rates per se does not provide any justification for treating either of the sets more favorably than the other. However, the NCHS rates do not satisfy the criteria of internal consistencies, whereas the present rates in table 6.13 are internally consistent for any two years.

To illustrate this point, let us choose any two years, say, 1941 and 1957. From the birth rates in table 6.13, the age effect is -5.2 (the difference between the rate-adjusted rates) and the rate effect is 10.1 (the difference between the age-sex-adjusted rates), and these two effects add up to 4.9, which is the same as the difference between the crude birth rates in 1941 and 1957. On the other hand, using the 1940 population as the standard, the rate-adjusted rates in 1941 and 1957 are 19.4 and 15.5 (so that the age effect is -3.9), and the age-sex-adjusted rates are 20.3 and 32.2 (so that the rate effect is 11.9). In this case, the two effects add up to 8.0, which is different from the difference between the two crude birth rates, namely, 4.9. It is easy to show that this inconsistency will still exist if the population for 1990 or for any other

Table 6.12. Population and Birth Rates by Nine Age-Sex Groups: United States, 1940 to 1990

Year	Female								
	10 to 14	15 to 19	20 to 24	25 to 29	30 to 34	35 to 39	40 to 44	45 to 49	Remainder
	Population in thousands								
1940	5777	6145	5907	5665	5192	4823	4387	4057	90169
1941	5699	6107	5943	5731	5270	4908	4463	4119	91162
1942	5608	6059	5972	5788	5340	4986	4536	4185	92386
1943	5507	6000	5990	5839	5407	5061	4609	4252	94074
1944	5395	5930	5999	5887	5472	5137	4681	4318	95578
1945	5294	5844	5995	5938	5539	5218	4756	4382	96962
1946	5260	5735	5999	6004	5609	5306	4831	4432	98213
1947	5252	5616	5985	6080	5683	5410	4908	4468	100724
1948	5308	5497	5957	6157	5761	5524	4989	4497	102941
1949	5388	5382	5917	6227	5838	5641	5073	4528	105194
1950	5506	5294	5886	6291	5942	5762	5169	4576	107845
1951	5650	5240	5800	6248	6053	5819	5260	4670	110138
1952	5872	5235	5692	6187	6167	5873	5353	4790	112384
1953	6218	5307	5547	6148	6221	5919	5443	4902	114479
1954	6475	5410	5425	6064	6294	5961	5522	5019	116856
1955	6694	5482	5363	5977	6337	6023	5596	5131	119328
1956	6833	5628	5317	5907	6309	6132	5672	5220	121885
1957	7387	5843	5312	5812	6257	6243	5744	5309	124077
1958	7636	6186	5384	5679	6231	6296	5810	5397	126263
1959	7971	6436	5483	5563	6155	6365	5871	5473	128513
1960	8323	6639	5564	5512	6078	6402	5946	5535	129980
1961	8771	6794	5737	5476	5985	6400	6043	5585	132201
1962	8749	7376	5973	5468	5876	6362	6152	5631	134184
1963	8911	7647	6345	5533	5767	6297	6256	5683	136044
1964	9114	8008	6618	5637	5658	6212	6332	5747	137815
1965	9357	8386	6846	5727	5607	6121	6368	5827	139287
1966	9565	8842	6993	5889	5579	6030	6373	5925	140380
1967	9800	8836	7581	6105	5585	5933	6347	6038	141232
1968	9990	9013	7847	6455	5659	5829	6294	6148	142164
1969	10128	9234	8187	6696	5768	5725	6222	6230	143195
1970	10230	9517	8544	6914	5871	5679	6148	6277	144804
1971	10346	9740	9027	7061	6036	5665	6062	6280	146610
1972	10347	9988	9021	7652	6268	5688	5971	6223	148126
1973	10310	10193	9198	7918	6652	5744	5885	6178	149279
1974	10243	10349	9415	8282	6929	5836	5797	6114	150377
1975	10112	10465	9677	8660	7173	5931	5700	6072	151675
1976	9837	10582	9901	9157	7317	6075	5689	5994	153011
1977	9550	10581	10152	9157	7928	6283	5713	5910	154486
1978	9262	10555	10373	9357	8205	6651	5780	5838	156074
1979	9031	10498	10541	9597	8579	6918	5883	5766	157754
1980	8923	10377	10680	9896	8974	7159	5988	5677	159581
1981	8953	10080	10790	10132	9481	7310	6136	5643	161112
1982	8877	9779	10781	10396	9482	7918	6354	5656	162753
1983	8747	9471	10729	10607	9672	8201	6699	5754	164404
1984	8544	9231	10642	10763	9900	8584	6942	5874	165997
1985	8339	9106	10482	10869	10172	8967	7167	5968	167666
1986	8078	9128	10183	10982	10407	9467	7316	6110	169436
1987	8035	9047	9877	10971	10674	9466	7929	6325	171103
1988	8102	8923	9576	10924	10895	9660	8210	6668	172847
1989	8260	8721	9335	10837	11059	9890	8589	6920	174648
1990	8447	8525	9223	10691	11175	10167	8987	7150	176648

Table 6.12. Population and Birth Rates by Nine Age-Sex Groups: United States, 1940 to 1990—Continued

Year	Female								
	10 to 14	15 to 19	20 to 24	25 to 29	30 to 34	35 to 39	40 to 44	45 to 49	Remainder
	Birth rates per 1,000 population								
1940	.7	54.1	135.6	122.8	83.4	46.3	15.6	1.9	.0
1941	.7	56.9	145.4	128.7	85.3	46.1	15.0	1.7	.0
1942	.7	61.1	165.1	142.7	91.8	47.9	14.7	1.6	.0
1943	.8	61.7	164.0	147.8	99.5	52.8	15.7	1.5	.0
1944	.8	54.3	151.8	136.5	98.1	54.6	16.1	1.4	.0
1945	.8	51.1	138.9	132.2	100.2	56.9	16.6	1.6	.0
1946	.7	59.3	181.8	161.2	108.9	58.7	16.5	1.5	.0
1947	.9	79.3	209.7	176.0	111.9	58.9	16.6	1.4	.0
1948	1.0	81.8	200.3	163.4	103.7	54.5	15.7	1.3	.0
1949	1.0	83.4	200.1	165.4	102.1	53.5	15.3	1.3	.0
1950	1.0	81.6	196.6	166.1	103.7	52.9	15.1	1.2	.0
1951	.9	87.6	211.6	175.3	107.9	54.1	15.4	1.1	.0
1952	.9	86.1	217.6	182.0	112.6	55.8	15.5	1.3	.0
1953	1.0	88.2	224.6	184.1	113.4	56.6	15.8	1.0	.0
1954	.9	90.6	236.2	188.4	116.9	57.9	16.2	1.0	.0
1955	.9	90.5	242.0	190.5	116.2	58.7	16.1	1.0	.0
1956	1.0	94.6	253.7	194.7	117.3	59.3	16.3	1.0	.0
1957	1.0	96.3	260.6	199.4	118.9	59.9	16.3	1.1	.0
1958	.9	91.4	258.2	198.3	116.2	58.3	15.7	.9	.0
1959	.9	90.4	260.1	200.5	115.6	58.2	15.5	1.1	.0
1960	.8	89.1	258.1	197.4	112.7	56.2	15.5	.9	.0
1961	.9	88.6	251.9	197.5	113.2	55.6	15.6	.9	.0
1962	.8	81.4	241.9	191.1	108.6	52.6	14.9	.9	.0
1963	.9	76.7	229.1	185.1	105.8	51.2	14.2	.9	.0
1964	.9	73.1	217.5	178.7	103.4	49.9	13.8	.8	.0
1965	.8	70.5	195.3	161.6	94.4	46.2	12.8	.8	.0
1966	.8	70.3	185.6	148.2	85.1	41.9	11.7	.7	.0
1967	.9	67.5	172.9	142.1	78.7	38.3	10.6	.7	.0
1968	1.0	65.6	166.5	140.0	74.2	35.4	9.6	.6	.0
1969	1.0	65.5	165.7	143.0	73.5	33.1	8.8	.5	.0
1970	1.2	68.3	167.8	145.1	73.3	31.7	8.1	.5	.0
1971	1.1	64.5	150.1	134.1	67.3	28.7	7.1	.4	.0
1972	1.2	61.7	130.2	117.7	59.8	24.8	6.2	.4	.0
1973	1.2	59.3	119.7	112.2	55.6	22.1	5.4	.3	.0
1974	1.2	57.5	117.7	111.5	53.8	20.2	4.8	.3	.0
1975	1.3	55.6	113.0	108.2	52.3	19.5	4.6	.3	.0
1976	1.2	52.8	110.3	106.2	53.6	19.0	4.3	.2	.0
1977	1.2	52.8	112.9	111.0	56.4	19.2	4.2	.2	.0
1978	1.2	51.5	109.9	108.5	57.8	19.0	3.9	.2	.0
1979	1.2	52.3	112.8	111.4	60.3	19.5	3.9	.2	.0
1980	1.1	53.0	115.1	112.9	61.9	19.8	3.9	.2	.0
1981	1.1	52.7	111.8	112.0	61.4	20.0	3.8	.2	.0
1982	1.1	52.9	111.3	111.0	64.2	21.1	3.9	.2	.0
1983	1.1	51.7	108.3	108.7	64.6	22.1	3.8	.2	.0
1984	1.2	50.9	107.3	108.3	66.5	22.8	3.9	.2	.0
1985	1.2	51.3	108.9	110.5	68.5	23.9	4.0	.2	.0
1986	1.3	50.6	108.2	109.2	69.3	24.3	4.1	.2	.0
1987	1.3	51.1	108.9	110.8	71.3	26.2	4.4	.2	.0
1988	1.3	53.6	111.5	113.4	73.7	27.9	4.8	.2	.0
1989	1.4	58.1	115.4	116.6	76.2	29.7	5.2	.2	.0
1990	1.5	60.7	120.6	121.8	79.6	31.0	5.4	.2	.0

Source: For population, U.S. Bureau of the Census (1965; 1974, table 2; 1982, table 2; 1990a, table 2; 1990b, table 2; Unpublished data for 1987-1990). For rates, National Center for Health Statistics (1967a, table 1-6; 1984, table 1-6; 1991a, table B; 1991b, table 4).year is used as the standard. Thus, the present method not only removes the internal inconsistencies in the adjusted rates, but also uses a computational formula (6.11) which puts an end to the debate as to which one of the actual populations should be used as the standard.

Table 6.13. Crude Birth Rates and Crude Death Rates per 1,000 Population and the Corresponding Adjusted (Standardized) Rates: United States, 1940 to 1990

Year	Birth rates			Death rates		
	Crude	Rate adjusted	Age-sex adjusted	Crude	Rate adjusted	Age adjusted
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1940	19.4	22.1	17.4	10.8	6.9	13.3
1941	20.3	22.1	18.3	10.5	7.2	12.8
1942	22.2	22.1	20.2	10.3	7.4	12.4
1943	22.7	22.0	20.8	10.7	7.5	12.8
1944	21.2	21.8	19.6	10.3	7.6	12.2
1945	20.4	21.6	19.0	10.2	7.7	11.9
1946	24.1	21.6	22.6	9.9	7.8	11.6
1947	26.5	21.3	25.3	10.0	8.1	11.4
1948	24.8	20.9	24.0	9.9	8.1	11.2
1949	24.5	20.6	24.0	9.7	8.2	10.9
1950	23.9	20.3	23.8	9.6	8.3	10.8
1951	24.8	19.8	25.1	9.6	8.4	10.7
1952	25.0	19.3	25.8	9.5	8.5	10.5
1953	24.9	18.8	26.2	9.5	8.6	10.4
1954	25.1	18.3	27.0	9.1	8.7	9.9
1955	24.9	17.8	27.2	9.2	8.8	9.9
1956	25.1	17.3	27.9	9.3	8.9	9.9
1957	25.2	16.9	28.4	9.5	8.9	10.1
1958	24.4	16.7	27.9	9.4	9.0	10.0
1959	24.2	16.4	27.9	9.3	9.0	9.8
1960	23.7	16.3	27.5	9.6	9.1	10.0
1961	23.3	16.2	27.3	9.3	9.2	9.6
1962	22.4	16.4	26.1	9.4	9.2	9.8
1963	21.7	16.7	25.2	9.6	9.2	9.9
1964	21.1	16.9	24.2	9.4	9.2	9.7
1965	19.4	17.3	22.3	9.4	9.3	9.7
1966	18.4	17.6	21.0	9.5	9.3	9.7
1967	17.8	18.2	19.8	9.4	9.4	9.5
1968	17.6	18.6	19.1	9.7	9.5	9.7
1969	17.9	18.9	19.1	9.5	9.6	9.5
1970	18.4	19.3	19.3	9.4	9.7	9.2
1971	17.2	19.7	17.7	9.3	9.8	9.1
1972	15.6	20.0	15.7	9.4	9.8	9.1
1973	14.8	20.3	14.6	9.3	9.9	9.0
1974	14.8	20.7	14.2	9.1	10.0	8.6
1975	14.6	21.0	13.7	8.8	10.1	8.2
1976	14.6	21.4	13.3	8.8	10.3	8.0
1977	15.1	21.6	13.7	8.6	10.4	7.8
1978	15.0	21.7	13.4	8.7	10.5	7.7
1979	15.6	21.9	13.7	8.5	10.6	7.4
1980	16.0	22.1	14.0	8.8	10.7	7.6
1981	15.8	22.2	13.7	8.6	10.8	7.3
1982	15.9	22.2	13.8	8.5	10.9	7.1
1983	15.6	22.1	13.6	8.6	11.0	7.1
1984	15.5	22.0	13.6	8.6	11.1	7.0
1985	15.8	21.9	13.9	8.7	11.2	7.1
1986	15.6	21.8	13.9	8.7	11.3	7.0
1987	15.6	21.5	14.2	8.7	11.4	6.9
1988	15.9	21.2	14.8	8.8	11.4	6.9
1989	16.3	20.9	15.5	8.7	11.5	6.7
1990	16.7	20.6	16.2	8.6	11.6	6.6

Figure 1.
Crude Birth Rates, and Age-Sex-Adjusted Birth Rates by Three Methods: United States,1940 to 1990

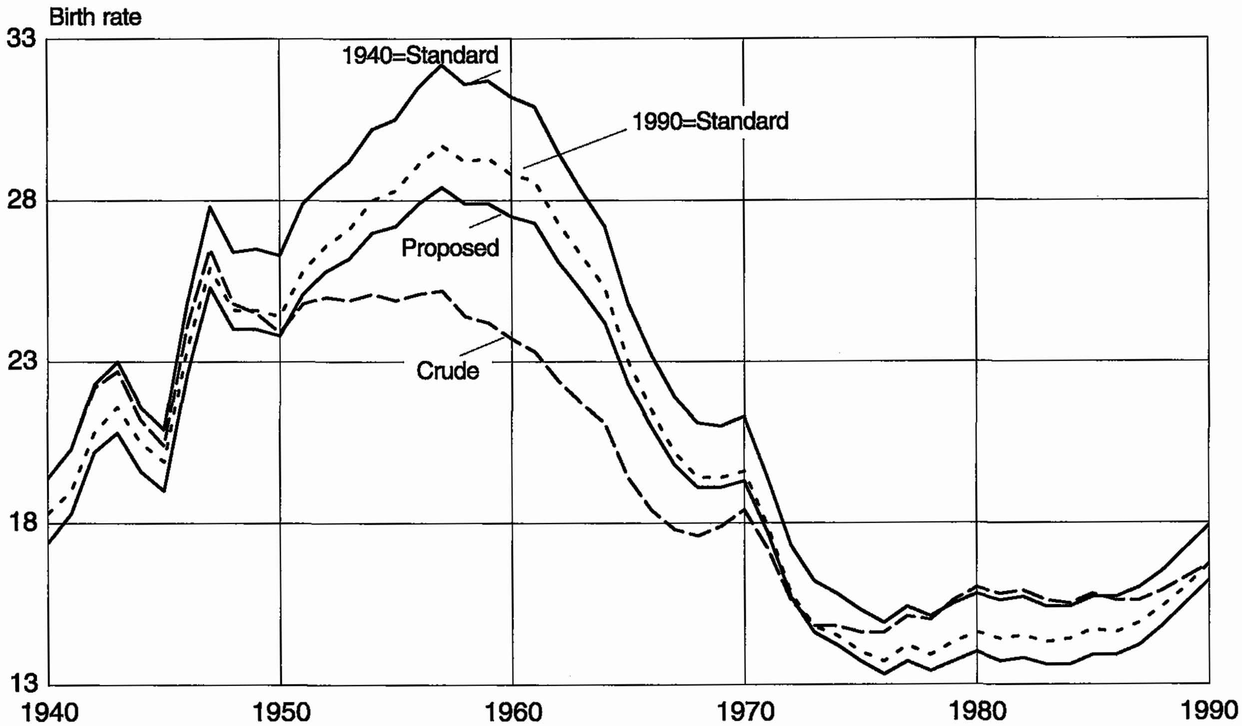
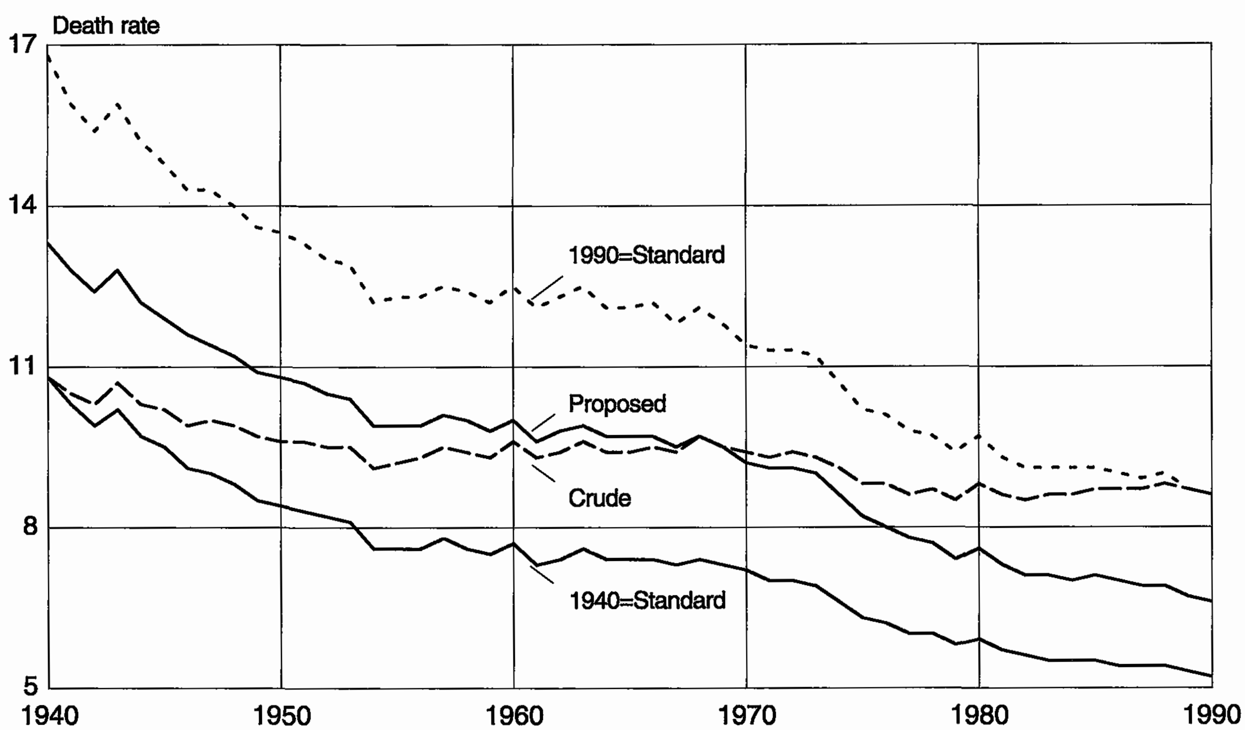


Figure 2.
Crude Death Rates, and Age-Adjusted Death Rates by Three Methods: United States,1940 to 1990



year is used as the standard. Thus, the present method not only removes the internal inconsistencies in the adjusted rates, but also uses a computational formula (6.11) which puts an end to the debate as to which one of the actual populations should be used as the standard.

Program 6.3

The results in columns (1) through (4) of table 6.13 can be obtained by using Program 6.3 in which L and N in lines 2 and 3 are the number of age-sex groups and the number of years, respectively. P(I,J)'s and U(I,J)'s in line 6 are, respectively, the populations in thousands and the birth rates per 1,000 population given in table 6.12. The data file consists of 102 lines, one pair of lines for each of the 51 years. The first and second lines, for example, give, respectively, the nine populations by age-sex groups for 1940, and the nine birth rates by age-sex groups for 1940, the formats being as shown in line 7 of the program. This data file when fed to Program 6.3 will generate an output that is identical to the first four columns in table 6.13.

As more and more years are added to the time series, the adjusted rates for earlier years will not necessarily remain the same. However, as long as Program 6.3 is available, the computation of a revised set of adjusted rates is very easy. If, for example, we want to add the year 1991 to the present time series 1940-1990, all we have to do is to add two lines to the data file giving the populations and birth rates for 1991, and run the program (Program 6.3) again with N=52 in line 3.

Example 6.7

This example is very similar to Example 6.6 except that here we adjust the death rates, instead of birth rates. Table 6.14 gives the populations in thousands and the corresponding death rates per 1,000 population in 11 age groups for the 51 years 1940-1990 for the United States. The rate-adjusted death rates and the age-adjusted death rates for these years, based on formula (6.11), and the crude death rates are shown in columns (5) through (7) of table 6.13.

The age-adjusted death rates in column (7) of table 6.13 are uniformly **higher** than the corresponding adjusted rates for all 51 years provided by the NCHS based on the 1940 population as the standard (figure 2). Here, again, the NCHS rates, unlike the rates in table 6.13, are not internally consistent, as we see from the rates of any two years, say, 1941 and 1957, again. From the death rates in table 6.13, the age effect is 1.7 (i.e., 8.9−7.2) and the rate effect is -2.7 (i.e., 10.1−12.8), and these two effects add up to -1.0, which is the same as the difference between the crude death rates in 1941 and 1957. On the other hand, using the 1940 population as the standard, the rate-adjusted rates in 1941 and 1957 are 10.9 and 13.1 (so that the age effect is 2.2), and the age-adjusted rates are 10.3 and 7.8 (so that the rate effect is -2.5). In this case, the two effects add up to -0.3, which is different from the difference of -1.0 between the two crude death rates. Again, the use of 1990 population or any other population as the standard will produce similar inconsistencies.

The results in columns (5) through (7) of table 6.13 can, again, be obtained by using Program 6.3 by making the following changes in the program:

1. Replace L=9 in line 2 by L=11
2. Replace 9F8.0/9F8.1 in line 7 by 11F7.0/11F7.1 .

The data file, again, consists of 102 lines, one pair of lines for each of the 51 years. The first and second lines, for example, give, respectively, the 11 populations by age groups for 1940, and the 11 death rates by age groups for 1940, the formats being as in line 7 with the change mentioned above. This data file when used with the revised Program 6.3 will generate columns (1) and (5) through (7) of table 6.13 as the output.

As in the case of adjusted birth rates, data for more years can be added to the data file, and the program, with a revised N in line 3, can be run again to obtain a new set of adjusted death rates.

Program 6.3 (Time Series: Birth and Death Rates)

```
      DIMENSION P(20,80),U(20,80),R(80),T(2,80),S(2,80,80)
      L=9
      N=51
      Z=N
      DO 1 J=1,N
    1 READ(5,2) (P(I,J),I=1,L),(U(I,J),I=1,L)
    2 FORMAT(9F8.0/9F8.1)
      DO 3 J=1,N
      P(L+1,J)=0.0
      DO 3 I=1,L
    3 P(L+1,J)=P(L+1,J)+P(I,J)
      DO 4 J=1,N
      R(J)=0.0
      DO 4 I=1,L
    4 R(J)=R(J)+P(I,J)*U(I,J)/P(L+1,J)
      DO 5 K=1,N-1
      DO 5 J=K+1,N
      S(1,J,K)=0.0
      S(1,K,J)=0.0
      S(2,J,K)=0.0
      S(2,K,J)=0.0
      DO 5 I=1,L
      A=(U(I,J)+U(I,K))*.5
      B=(P(I,J)/P(L+1,J)+P(I,K)/P(L+1,K))*.5
      S(1,J,K)=S(1,J,K)+A*P(I,J)/P(L+1,J)
      S(1,K,J)=S(1,K,J)+A*P(I,K)/P(L+1,K)
      S(2,J,K)=S(2,J,K)+B*U(I,J)
    5 S(2,K,J)=S(2,K,J)+B*U(I,K)
      DO 8 I=1,2
      DO 8 J=1,N
      AA=0.0
      BB=0.0
      CC=0.0
      DO 7 K=1,N
      IF(K.EQ.J) GO TO 6
      AA=AA+S(I,J,K)
      CC=CC+(Z-2.)*S(I,K,J)
    6 DO 7 JJ=1,N
      IF(JJ.EQ.J.OR.K.EQ.J.OR.K.EQ.JJ) GO TO 7
      BB=BB+S(I,JJ,K)
    7 CONTINUE
    8 T(I,J)=AA/(Z-1.)+(BB-CC)/(Z*(Z-1.))
      DO 9 J=1,N
      JJ=J+1939
    9 WRITE(6,10) JJ,R(J),T(1,J),T(2,J)
   10 FORMAT(10X,I10,3F15.1)
      STOP
      END
****
```

Table 6.14. Population and Death Rates by 11 Age Groups: United States, 1940 to 1990

Year	Less than 1	1 to 4	5 to 14	15 to 24	25 to 34	35 to 44	45 to 54	55 to 64	65 to 74	75 to 84	85+
	Population in thousands										
1940	2025	8554	22363	24033	21446	18422	15555	10694	6367	2294	370
1941	2167	8683	22089	24074	21691	18692	15759	10959	6546	2357	385
1942	2325	8976	21823	24093	21911	18950	15976	11220	6745	2437	402
1943	2693	9323	21699	24065	22194	19226	16199	11472	6941	2509	417
1944	2516	10008	21573	23999	22511	19505	16419	11719	7136	2581	430
1945	2464	10515	21599	23705	22734	19787	16642	11988	7359	2683	452
1946	2401	10843	21844	23382	22954	20073	16820	12244	7566	2792	470
1947	3452	10954	22257	23122	23236	20421	16970	12528	7776	2917	492
1948	3169	11750	23089	22866	23494	20794	17107	12824	7978	3043	517
1949	3170	12437	23770	22570	23729	21187	17260	13145	8194	3178	549
1950	3163	13247	24588	22355	24036	21637	17453	13396	8493	3314	590
1951	3315	14018	25168	22109	24190	21913	17677	13685	8742	3433	628
1952	3429	13883	26773	21885	24301	22193	17935	13949	8990	3548	665
1953	3546	14092	28003	21746	24340	22446	18227	14166	9248	3668	701
1954	3671	14386	29223	21726	24340	22662	18559	14382	9536	3802	738
1955	3777	14789	30387	21753	24283	22912	18885	14622	9808	3943	776
1956	3860	15143	31570	21956	24122	23258	19206	14852	10053	4080	804
1957	4035	15459	32669	22401	23844	23597	19579	15012	10319	4231	837
1958	4073	15814	33487	23257	23538	23798	19927	15182	10576	4365	864
1959	4097	16078	34564	23988	23169	24023	20262	15401	10819	4528	901
1960	4094	16247	35735	24194	22728	24120	20560	15624	11053	4681	940
1961	4173	16349	37031	24865	22494	24289	20856	15847	11269	4856	964
1962	4084	16385	37435	26483	22287	24413	21099	16129	11457	5017	982
1963	4013	16329	38124	27803	22196	24484	21322	16435	11611	5163	1003
1964	3947	16218	38783	29096	22195	24463	21556	16757	11759	5328	1040
1965	3770	16054	39427	30326	22266	24343	21813	17076	11887	5483	1082
1966	3555	15653	40051	31428	22483	24151	22095	17407	11989	5638	1128
1967	3450	15113	40496	32357	22896	23908	22416	17751	12082	5803	1186
1968	3366	14547	40771	33245	23700	23589	22728	18087	12179	5946	1241
1969	3413	13963	40884	34389	24406	23243	23019	18389	12301	6072	1307
1970	3508	13658	40772	35810	25109	23040	23299	18682	12493	6183	1430
1971	3601	13643	40490	37418	25769	22878	23503	18961	12684	6390	1487
1972	3306	13795	39946	38089	27463	22780	23675	19210	12922	6555	1542
1973	3128	13723	39309	38919	28788	22740	23799	19428	13247	6671	1607
1974	3065	13422	38716	39736	30072	22755	23800	19713	13574	6781	1706
1975	3152	12969	38240	40540	31314	22760	23749	20045	13917	6958	1821
1976	3115	12502	37759	41272	32605	23030	23615	20386	14237	7145	1896
1977	3279	12285	37034	41788	33841	23500	23363	20779	14638	7262	1992
1978	3326	12409	36220	42183	34803	24373	23166	21112	14996	7412	2095
1979	3426	12637	35392	42444	36038	25114	22936	21448	15338	7599	2197
1980	3561	12897	34845	42484	37451	25806	22746	21761	15653	7782	2269
1981	3620	13311	34405	42115	38986	26400	22608	21955	15915	7971	2350
1982	3666	13632	34193	41474	39561	28050	22482	22114	16198	8183	2444
1983	3684	13967	34059	40763	40413	29300	22439	22233	16495	8399	2531
1984	3617	14213	33975	40114	41231	30546	22495	22316	16740	8616	2615
1985	3736	14268	33923	39552	42027	31764	22589	22337	17010	8836	2695
1986	3770	14384	33860	39021	42778	33070	22815	22235	17334	9062	2778
1987	3785	14482	34147	38250	43312	34307	23277	22025	17674	9302	2865
1988	3852	14585	34654	37396	43670	35265	24164	21834	17915	9532	2938
1989	3947	14812	35163	36515	43836	36503	24897	21598	18193	9767	3027
1990	4137	15018	35724	35925	43771	37875	25529	21442	18469	9993	3131

Table 6.14. Population and Death Rates by 11 Age Groups: United States, 1940 to 1990

Year	Less than 1	1 to 4	5 to 14	15 to 24	25 to 34	35 to 44	45 to 54	55 to 64	65 to 74	75 to 84	85+
	Death rates per 1,000 population										
1940	54.9	2.9	1.0	2.0	3.1	5.2	10.6	22.2	48.4	112.0	235.7
1941	52.6	2.8	1.0	2.0	2.9	5.0	10.3	21.3	46.2	105.8	218.7
1942	48.8	2.4	.9	1.9	2.8	4.8	10.1	21.0	44.9	101.6	211.1
1943	44.0	2.6	1.0	2.1	2.7	4.8	10.2	21.5	46.2	107.5	230.3
1944	44.2	2.3	.9	2.0	2.7	4.6	9.7	20.8	43.9	101.7	215.3
1945	42.5	2.0	.9	1.9	2.7	4.6	9.6	20.5	42.6	98.4	209.6
1946	46.3	1.8	.8	1.7	2.3	4.2	9.2	19.8	41.2	95.1	210.6
1947	34.5	1.6	.7	1.5	2.1	4.1	9.2	20.1	42.1	97.0	216.9
1948	35.7	1.6	.7	1.4	2.0	3.9	9.0	19.7	41.4	95.1	213.2
1949	35.2	1.5	.7	1.3	1.8	3.7	8.7	19.3	40.8	93.0	203.2
1950	33.0	1.4	.6	1.3	1.8	3.6	8.5	19.0	41.0	93.3	202.0
1951	32.3	1.4	.6	1.3	1.8	3.5	8.5	18.8	40.0	93.2	192.3
1952	32.1	1.4	.6	1.3	1.7	3.4	8.3	18.7	39.2	91.9	183.0
1953	30.7	1.3	.6	1.2	1.6	3.3	8.2	18.4	39.1	92.4	183.4
1954	29.2	1.2	.5	1.1	1.5	3.1	7.8	17.4	37.6	87.5	172.6
1955	28.5	1.1	.5	1.1	1.5	3.1	7.6	17.3	37.9	89.0	179.3
1956	28.3	1.1	.5	1.1	1.5	3.0	7.5	17.5	37.8	88.5	181.8
1957	28.0	1.1	.5	1.2	1.5	3.1	7.7	17.8	38.9	88.4	188.4
1958	28.3	1.1	.5	1.1	1.5	3.0	7.5	17.4	38.3	87.8	191.1
1959	27.6	1.1	.5	1.1	1.5	2.9	7.4	17.1	37.6	85.8	191.2
1960	27.0	1.1	.5	1.1	1.5	3.0	7.6	17.4	38.2	87.5	198.6
1961	25.9	1.0	.4	1.0	1.4	2.9	7.3	16.8	36.9	84.0	196.3
1962	25.8	1.0	.4	1.0	1.4	3.0	7.4	16.9	37.4	85.0	204.9
1963	25.8	1.0	.4	1.1	1.5	3.0	7.5	17.2	37.9	86.3	209.9
1964	25.3	1.0	.4	1.1	1.5	3.1	7.5	16.9	36.6	83.2	199.2
1965	24.6	1.0	.4	1.1	1.5	3.1	7.5	16.8	36.6	83.6	200.7
1966	24.1	1.0	.4	1.2	1.5	3.1	7.5	16.9	37.0	83.5	199.8
1967	22.9	.9	.4	1.2	1.5	3.1	7.4	16.6	36.2	80.9	192.2
1968	22.7	.9	.4	1.2	1.6	3.2	7.5	17.0	37.2	82.9	195.8
1969	22.0	.9	.4	1.3	1.6	3.2	7.3	16.6	36.3	81.0	188.2
1970	21.4	.8	.4	1.3	1.6	3.1	7.3	16.6	35.8	80.0	163.4
1971	18.9	.8	.4	1.3	1.6	3.1	7.1	16.2	34.8	77.7	175.7
1972	18.2	.8	.4	1.3	1.5	3.0	7.1	16.2	35.2	78.0	175.4
1973	17.8	.8	.4	1.3	1.5	3.0	7.0	16.0	34.3	77.1	176.8
1974	17.2	.7	.4	1.2	1.4	2.8	6.8	15.3	33.2	73.8	169.0
1975	16.0	.7	.4	1.2	1.4	2.7	6.5	14.8	31.8	70.3	156.5
1976	15.5	.7	.3	1.1	1.3	2.5	6.4	14.5	31.2	69.5	160.6
1977	14.3	.7	.3	1.1	1.3	2.5	6.2	14.1	30.4	67.6	153.7
1978	13.8	.7	.3	1.1	1.3	2.4	6.1	13.9	30.2	67.1	154.8
1979	13.3	.6	.3	1.1	1.3	2.3	5.9	13.4	29.3	65.0	149.6
1980	12.9	.6	.3	1.2	1.4	2.3	5.8	13.5	29.9	66.9	159.8
1981	12.1	.6	.3	1.1	1.3	2.2	5.7	13.2	29.2	64.3	153.8
1982	11.6	.6	.3	1.0	1.3	2.1	5.5	13.0	28.9	63.3	150.5
1983	11.1	.6	.3	1.0	1.2	2.0	5.4	13.0	28.7	64.4	151.7
1984	10.9	.5	.3	1.0	1.2	2.0	5.2	12.9	28.5	64.0	152.2
1985	10.7	.5	.3	1.0	1.2	2.1	5.2	12.8	28.4	64.5	154.8
1986	10.3	.5	.3	1.0	1.3	2.1	5.0	12.6	28.0	63.5	154.0
1987	10.2	.5	.3	1.0	1.3	2.1	5.0	12.4	27.5	62.8	153.2
1988	10.1	.5	.3	1.0	1.4	2.2	4.9	12.4	27.3	63.2	155.9
1989	10.1	.5	.3	1.0	1.4	2.2	4.8	12.0	26.5	61.4	150.3
1990	9.9	.5	.3	1.0	1.4	2.2	4.7	11.9	26.1	60.5	148.1

Source: For population, the same as in table 6.12. For rates, National Office of Vital Statistics (1956, table AO); National Center for Health Statistics (1963, table 1-C; 1982, table 2; 1987, table 1-3; 1990b, table 1-4; 1991a, table F; 1992, table 5).

Example 6.8

Table 6.15 gives the populations and the corresponding census undercount rates in six race-sex groups for the United states, 50 States, and the District of Columbia for 1990. Treating race and sex as two separate factors, it is possible to compute, in addition to the crude undercount rates, the three standardized undercount rates adjusted for sex and rate, race and rate, and race and sex, for the 52 geographical areas, by using formula (6.11). These four rates for each area are shown in table 6.16.

In table 6.16, the difference between two sex-rate-adjusted rates gives the race effect. Similarly, the difference between two race-rate-adjusted rates gives the sex effect, and that between two race-sex-adjusted rates gives the rate effect (i.e., the effect of the race-sex-specific undercount rates). The rates in table 6.16 are internally consistent because, for any two geographical areas, the race effect, the sex effect, and the rate effect add up to the total difference between the crude undercount rates.

It is evident from the race-rate-adjusted rates in table 6.16 that sex does not play a significant role in explaining the differences in the undercount rates in the States. The results in table 6.16 have some implications for the synthetic method of census adjustment, which assumes that undercount rates are constant within subgroups of people with given demographic characteristics across geographical areas. If these characteristics are race and sex, then, in order for the synthetic method to work at the State level, we should expect the race-sex-adjusted undercount rates in column (5) of table 6.16 to be approximately equal. Obviously, our results indicate that the synthetic method based on race and sex is not expected to generate satisfactory undercount rates at the State level. Further research is needed to include variables that are symptomatic of coverage differences, such as house tenure (owner/non-owner), since the 1990 PES data showed consistently higher undercount rates for non-owners (Robinson and Ahmed, 1992; Hogan, 1992).

Program 6.4

The results in columns (2) through (5) of table 6.16 can be obtained by using Program 6.4. This program is basically a combination of Program 5.2 (Two Factors + Rate) when the factors I (race) and J (sex) have, respectively, three and two categories, and Program 6.3 (Time Series: Birth and Death Rates) when the number of factors (including rate) is three and the number of populations is 52. V(I,J,K)'s and U(I,J,K)'s in line 4 are, respectively, the populations and the undercount rates given in table 6.15. The data file consists of 104 lines, one pair of lines for each of the 52 geographical areas. The first and second lines, for example, give, respectively, the six populations by race-sex groups for Alabama, and the six undercount rates by race-sex groups for Alabama, the formats being as shown in line 5 of the program. This data file when fed to Program 6.4 will generate an output that is identical to the five columns in table 6.16, except that the geographical areas in column (1) are represented by serial numbers.

Table 6.15. Population and Census Undercount Rates by Race and Sex: United States, 50 States, and the District of Columbia, 1990

States	Male			Female		
	Black	Hispanic	Other	Black	Hispanic	Other
	Population					
ALABAMA	487928	13421	1473118	567679	12492	1558481
ALASKA	12784	10044	273895	10468	8507	245558
ARIZONA	62324	364575	1437701	57237	355742	1476718
ARKANSAS	180713	11235	963411	206625	9940	1020367
CALIFORNIA	1191106	4228042	9988230	1199183	3855319	10132658
COLORADO	74226	223983	1374980	70500	220357	1399311
CONNECTICUT	137710	111885	1359258	151856	112724	1434875
DELAWARE	55665	8969	264580	61326	7928	279904
DISTRICT OF COLUMBIA	192468	18288	85266	223636	17092	91559
FLORIDA	881497	827151	4695711	951357	834646	5006492
GEORGIA	853639	66442	2300583	961370	50543	2386252
HAWAII	17957	43574	518156	11594	41351	496530
IDAHO	2093	31154	480852	1396	25588	488130
ILLINOIS	824978	492074	4305368	932089	437185	4552739
INDIANA	210877	51265	2446150	235900	49572	2578476
IOWA	25035	17489	1312205	24893	16360	1392396
KANSAS	74527	51206	1101043	73763	45727	1148496
KENTUCKY	130540	12282	1673926	141998	10881	1776034
LOUISIANA	632984	48816	1398996	717069	49232	1466420
MAINE	3219	3649	597526	2108	3575	627048
MARYLAND	591508	67941	1715877	649555	65161	1792282
MASSACHUSETTS	154365	152203	2607897	165778	151993	2812926
MICHIGAN	625852	104320	3823749	709891	102683	3994835
MINNESOTA	51583	28536	2079287	48161	26721	2160392
MISSISSIPPI	440448	8288	810679	508192	8454	852838
MISSOURI	264384	31870	2189658	302897	31128	2329115
MONTANA	1475	6681	398880	992	6202	404075
NEBRASKA	29243	19958	727788	30573	18421	762714
NEVADA	43769	70235	516298	41908	60728	497736
NEW HAMPSHIRE	4163	6229	539783	3297	5718	559420
NEW JERSEY	517491	395798	2858177	578607	382638	3041700
NEW MEXICO	16869	297679	456742	15152	304530	472151
NEW YORK	1404308	1156964	6237215	1640611	1193071	6629785
NORTH CAROLINA	708238	46819	2525805	798243	35184	2638886
NORTH DAKOTA	2171	2544	316773	1497	2397	317661
OHIO	560730	71786	4642245	635577	71952	4939635
OKLAHOMA	119162	48150	1393713	124134	43476	1474095
OREGON	26037	66916	1338090	23972	53149	1387983
PENNSYLVANIA	530366	123841	5074421	610553	119196	5458253
RHODE ISLAND	20763	24366	439395	20777	23973	475538
SOUTH CAROLINA	505355	17368	1203284	572764	15027	1245120
SOUTH DAKOTA	2046	2804	342424	1336	2766	351501
TENNESSEE	378539	18145	1998011	430552	16496	2121943
TEXAS	1020663	2317674	5292263	1084083	2267517	5487048
UTAH	7447	45519	820996	5158	43121	830880
VERMONT	1168	1965	276693	862	1930	286473
VIRGINIA	582343	90742	2428295	626725	80899	2504616
WASHINGTON	87077	122813	2262094	75554	105472	2304977
WEST VIRGINIA	27199	4480	843363	30731	4474	908757
WISCONSIN	121790	50441	2241991	134375	46477	2326923
WYOMING	2047	13624	217189	1676	13227	215806
UNITED STATES	14900869	12052243	96670030	16476230	11468942	101144508

Table 6.15. Population and Census Undercount Rates by Race and Sex: United States, 50 States, and the District of Columbia, 1990—Continued

States	Male			Female		
	Black	Hispanic	Other	Black	Hispanic	Other
	Undercount rates					
ALABAMA	3.796	5.388	1.294	2.886	4.494	1.108
ALASKA	3.541	4.827	2.160	3.321	3.094	1.526
ARIZONA	7.648	4.640	2.241	7.461	4.234	1.072
ARKANSAS	3.958	6.760	1.492	3.037	5.436	1.222
CALIFORNIA	8.088	5.263	1.910	7.101	4.485	.658
COLORADO	8.259	4.774	1.824	7.728	4.241	.878
CONNECTICUT	5.253	6.165	.136	5.310	4.077	-.519
DELAWARE	4.592	6.903	1.160	3.220	5.775	1.259
DISTRICT OF COLUMBIA	5.593	8.135	.937	2.564	6.915	1.574
FLORIDA	4.800	6.148	1.053	3.260	4.409	.968
GEORGIA	4.483	7.854	1.421	3.139	5.629	1.318
HAWAII	7.877	4.730	2.378	8.124	3.564	.548
IDAHO	3.375	7.925	2.205	3.404	5.262	1.624
ILLINOIS	3.980	2.768	.551	3.214	2.561	.055
INDIANA	3.617	2.022	.464	2.992	2.041	.000
IOWA	3.868	4.033	.629	3.495	3.036	.023
KANSAS	3.757	3.795	.671	3.272	2.888	.115
KENTUCKY	4.213	5.257	1.516	2.909	4.903	1.362
LOUISIANA	4.373	5.416	1.365	3.221	4.793	1.274
MAINE	3.544	5.975	1.040	3.535	4.952	.383
MARYLAND	5.229	6.804	1.230	3.116	5.182	1.148
MASSACHUSETTS	6.206	6.400	.245	6.293	4.544	-.510
MICHIGAN	3.635	2.426	.414	2.999	2.802	.018
MINNESOTA	5.079	2.434	.525	4.526	2.538	.116
MISSISSIPPI	3.914	5.068	1.375	3.217	4.628	1.189
MISSOURI	3.750	1.776	.508	3.024	2.344	.021
MONTANA	3.485	6.262	2.707	3.534	4.699	1.893
NEBRASKA	4.357	4.008	.753	3.722	3.312	.133
NEVADA	8.416	5.593	2.094	7.689	4.308	.920
NEW HAMPSHIRE	3.709	6.069	1.130	3.272	4.120	.426
NEW JERSEY	5.232	6.235	-.558	5.565	3.631	-1.238
NEW MEXICO	5.857	3.699	3.074	5.434	3.932	1.952
NEW YORK	5.958	7.042	.123	6.229	4.571	-.875
NORTH CAROLINA	3.808	7.326	1.432	2.904	5.249	1.248
NORTH DAKOTA	3.853	6.471	.959	3.985	4.614	.247
OHIO	3.855	2.709	.536	3.125	2.914	.089
OKLAHOMA	4.547	6.491	1.554	3.283	5.385	1.392
OREGON	7.916	7.003	2.035	7.385	5.028	1.110
PENNSYLVANIA	4.275	5.358	.100	4.660	3.473	-.586
RHODE ISLAND	6.333	6.226	.045	6.565	4.465	-.865
SOUTH CAROLINA	3.960	6.252	1.362	3.182	5.042	1.264
SOUTH DAKOTA	3.486	6.448	1.321	3.961	4.954	.543
TENNESSEE	4.670	5.769	1.353	3.107	5.169	1.250
TEXAS	4.746	5.886	1.507	3.199	4.801	1.359
UTAH	8.170	4.891	1.872	8.159	4.212	1.184
VERMONT	4.042	6.608	1.503	3.669	5.378	.650
VIRGINIA	4.526	7.183	1.421	3.146	5.977	1.369
WASHINGTON	8.085	6.773	1.896	7.662	5.117	.949
WEST VIRGINIA	3.293	4.858	1.469	2.405	5.523	1.214
WISCONSIN	4.886	4.164	.591	4.223	3.493	.071
WYOMING	3.153	4.783	2.365	3.097	3.385	1.681
UNITED STATES	4.904	5.511	1.023	4.008	4.385	.450

Source: Unpublished data in the Bureau of the Census. Populations are Post Enumeration Survey (PES) estimates. Undercount rates are defined: 100 x (PES pop. - Census pop.)/PES pop. **Other** is obtained by subtracting Black and Hispanic from Total. The race categories are approximate because of some overlap betweeen Black and Hispanic.

Table 6.16. Crude Undercount Rates and the Corresponding Three Adjusted (Standardized) Rates: United States, 50 States, and the District of Columbia, 1990

States	Undercount rates			
	Crude	Sex-rate adjusted	Race-rate adjusted	Race-sex adjusted
(1)	(2)	(3)	(4)	(5)
ALABAMA	1.763	1.815	1.543	1.535
ALASKA	1.998	1.334	1.571	2.222
ARIZONA	2.372	1.669	1.551	2.282
ARKANSAS	1.738	1.529	1.544	1.794
CALIFORNIA	2.727	2.213	1.557	2.088
COLORADO	2.051	1.512	1.552	2.117
CONNECTICUT	.641	1.538	1.544	.688
DELAWARE	1.799	1.641	1.546	1.741
DISTRICT OF COLUMBIA	3.407	3.349	1.530	1.658
FLORIDA	1.962	1.969	1.544	1.579
GEORGIA	2.125	1.903	1.547	1.806
HAWAII	1.854	1.221	1.571	2.192
IDAHO	2.183	1.308	1.554	2.450
ILLINOIS	.986	1.809	1.545	.762
INDIANA	.504	1.376	1.545	.713
IOWA	.417	1.106	1.545	.896
KANSAS	.689	1.355	1.548	.916
KENTUCKY	1.612	1.289	1.545	1.907
LOUISIANA	2.169	2.074	1.544	1.681
MAINE	.744	1.024	1.546	1.303
MARYLAND	2.066	1.935	1.545	1.716
MASSACHUSETTS	.475	1.273	1.541	.791
MICHIGAN	.705	1.588	1.546	.702
MINNESOTA	.446	1.098	1.548	.929
MISSISSIPPI	2.118	2.138	1.543	1.566
MISSOURI	.621	1.460	1.543	.747
MONTANA	2.352	1.180	1.552	2.749
NEBRASKA	.649	1.210	1.547	1.023
NEVADA	2.343	1.567	1.566	2.340
NEW HAMPSHIRE	.837	1.068	1.549	1.349
NEW JERSEY	.569	1.957	1.544	.198
NEW MEXICO	3.074	2.144	1.549	2.511
NEW YORK	1.487	2.241	1.541	.835
NORTH CAROLINA	1.844	1.707	1.547	1.720
NORTH DAKOTA	.660	1.006	1.553	1.230
OHIO	.685	1.446	1.543	.826
OKLAHOMA	1.784	1.367	1.546	2.000
OREGON	1.859	1.074	1.551	2.363
PENNSYLVANIA	.294	1.373	1.542	.509
RHODE ISLAND	.134	1.187	1.541	.537
SOUTH CAROLINA	2.029	1.957	1.546	1.656
SOUTH DAKOTA	.978	1.036	1.549	1.523
TENNESSEE	1.743	1.550	1.544	1.779
TEXAS	2.763	2.428	1.548	1.917
UTAH	1.727	1.039	1.552	2.267
VERMONT	1.113	1.024	1.549	1.670
VIRGINIA	1.999	1.700	1.548	1.881
WASHINGTON	1.842	1.154	1.554	2.264
WEST VIRGINIA	1.403	1.186	1.543	1.804
WISCONSIN	.615	1.223	1.547	.974
WYOMING	2.153	1.347	1.555	2.381
UNITED STATES	1.584	1.789	1.546	1.379

Program 6.4 (Census Undercount Rates for States)

```
      DIMENSION V(4,3,52),P(4,3,52),U(3,2,52),T(3,2,52),ET(52),ER(52),
     1 S(3,52,52),R(2,2),Z(3,52)
      DO 1 K=1,52
    1 READ(5,2) ((V(I,J,K),I=1,3),J=1,2),((U(I,J,K),I=1,3),J=1,2)
    2 FORMAT(6F12.0/6F12.3)
      DO 6 K=1,52
      DO 3 J=1,2
      V(4,J,K)=0.0
      DO 3 I=1,3
    3 V(4,J,K)=V(4,J,K)+V(I,J,K)
      DO 4 I=1,4
      V(I,3,K)=0.0
      DO 4 J=1,2
    4 V(I,3,K)=V(I,3,K)+V(I,J,K)
    6 CONTINUE
      DO 5 K=1,52
      ET(K)=0.0
      DO 5 I=1,3
      DO 5 J=1,2
    5 ET(K)=ET(K)+V(I,J,K)*U(I,J,K)/V(4,3,K)
      DO 12 K1=1,51
      DO 12 J1=K1+1,52
      DO 16 I=1,4
      DO 16 J=1,3
      P(I,J,1)=V(I,J,K1)
   16 P(I,J,2)=V(I,J,J1)
      DO 17 I=1,3
      DO 17 J=1,2
      T(I,J,1)=U(I,J,K1)
   17 T(I,J,2)=U(I,J,J1)
      DO 18 K=1,2
      ER(K)=0.0
      DO 18 I=1,3
      DO 18 J=1,2
      Q=(P(I,J,1)/P(4,3,1)+P(I,J,2)/P(4,3,2))*.5
   18 ER(K)=ER(K)+Q*T(I,J,K)
      DO 7 I=1,2
      DO 7 J=1,2
    7 R(I,J)=0.0
      DO 11 II=1,2
      DO 11 JJ=1,2
      H=0.0
      DO 10 IS=1,3
      DO 10 JS=1,2
      W=1.0
      DO 15 I1=1,2
      IF(I1.EQ.1) J=JS
      IF(I1.EQ.2) J=3
      DO 8 NL=1,2
      GO TO (13,14),NL
   13 A=P(IS,J,II)/P(4,J,II)
      GO TO 8
   14 IF(I1.EQ.1) I=IS
      IF(I1.EQ.2) I=4
      A=P(I,JS,JJ)/P(I,3,JJ)
    8 W=W*A
   15 CONTINUE
   10 H=H+(T(IS,JS,1)+T(IS,JS,2))*.5*W**(1./2.)
      R(II,1)=R(II,1)+H
   11 R(JJ,2)=R(JJ,2)+H
      DO 9 J=1,2
      S(J,K1,J1)=R(1,J)/2.
    9 S(J,J1,K1)=R(2,J)/2.
      S(3,K1,J1)=ER(1)
      S(3,J1,K1)=ER(2)
   12 CONTINUE
      DO 19 I=1,3
      DO 19 J=1,52
      AA=0.0
      BB=0.0
      CC=0.0
      DO 20 K=1,52
      IF(K.EQ.J) GO TO 21
      AA=AA+S(I,J,K)
      CC=CC+50.*S(I,K,J)
   21 DO 20 JJ=1,52
      IF(JJ.EQ.J.OR.K.EQ.J.OR.K.EQ.JJ) GO TO 20
      BB=BB+S(I,JJ,K)
   20 CONTINUE
   19 Z(I,J)=AA/51.+(BB-CC)/(52.*51.)
      DO 22 J=1,52
   22 WRITE(6,23) J,ET(J),(Z(I,J),I=1,3)
   23 FORMAT(10X,I10,4F20.3)
      STOP
      END
****
```

Appendix A. Derivation and Summary of Formulas

A.1 DERIVATION OF FORMULAS (3.18) THROUGH (3.20)

$$R = F(\alpha,\beta,\gamma)\,.$$

α, β, and γ assume values A,B,C in population 1 and a,b,c in population 2, so that the difference R_2 - R_1 is

$$F(a,b,c)-F(A,B,C)=\alpha\text{-effect} + \beta\text{-effect} + \gamma\text{-effect}\,. \tag{A1}$$

We write the three effects as

$$\begin{aligned}\alpha\text{-effect} = {} & w[F(a,b,c) - F(A,b,c)] + x[F(a,b,C) - F(A,b,C)] \\ & + y[F(a,B,c) - F(A,B,c)] + z[F(a,B,C) - F(A,B,C)]\,, \end{aligned} \tag{A2}$$

$$\begin{aligned}\beta\text{-effect} = {} & w[F(a,b,c) - F(a,B,c)] + x[F(a,b,C) - F(a,B,C)] \\ & + y[F(A,b,c) - F(A,B,c)] + z[F(A,b,C) - F(A,B,C)]\,, \end{aligned} \tag{A3}$$

$$\begin{aligned}\gamma\text{-effect} = {} & w[F(a,b,c) - F(a,b,C)] + x[F(A,b,c) - F(A,b,C)] \\ & + y[F(a,B,c) - F(a,B,C)] + z[F(A,B,c) - F(A,B,C)]\,, \end{aligned} \tag{A4}$$

where w, x, y, z are suitably chosen constants.

Substituting (A2) through (A4) on the right-hand side of (A1) and then equating the coefficients from both sides, we have

$$w = z = 1/3\,, \quad x = y = 1/6\,.$$

Substituting these values in (A2) through (A4), we obtain the formulas in (3.18) through (3.20).

A.2 THREE FACTORS WITH INTERACTIONS

$$F(\alpha,\beta,\gamma) = K + E_\alpha + E_\beta + E_\gamma + E_{\alpha\beta} + E_{\alpha\gamma} + E_{\beta\gamma} + E_{\alpha\beta\gamma}\,, \tag{A5}$$

where

$$\sum_\alpha E_\alpha = \sum_\beta E_\beta = \sum_\gamma E_\gamma = 0\,, \tag{A6}$$

$$\sum_\alpha E_{\alpha\beta} = \sum_\beta E_{\alpha\beta} = \sum_\alpha E_{\alpha\gamma} = \sum_\gamma E_{\alpha\gamma} = \sum_\beta E_{\beta\gamma} = \sum_\gamma E_{\beta\gamma} = 0\,, \tag{A7}$$

$$\sum_\alpha E_{\alpha\beta\gamma} = \sum_\beta E_{\alpha\beta\gamma} = \sum_\gamma E_{\alpha\beta\gamma} = 0\,. \tag{A8}$$

There are 27 unknowns in (A5), which can be solved from 27 independent equations (8 in A5, 3 in A6, 9 in A7, and 7 in A8). Using these solutions, we have

$$\begin{aligned}F(a,b,c) - F(A,B,C) &= (E_a - E_A) + (E_b - E_B) + (E_c - E_C) + (E_{abc} - E_{ABC}) \\ &= \alpha\text{-effect} + \beta\text{-effect} + \gamma\text{-effect} + \alpha\beta\gamma\text{-interaction effect,}\end{aligned}$$

where

$$\alpha\text{-effect} = \frac{\begin{array}{c}[F(a,b,c) - F(A,b,c)] + [F(a,B,C) - F(A,B,C)] \\ + [F(a,b,C) - F(A,b,C)] + [F(a,B,c) - F(A,B,c)]\end{array}}{4},$$

$$\alpha\beta\gamma\text{-interaction effect} = \frac{\begin{array}{c}[F(a,b,c) - F(A,b,c)] + [F(a,B,C) - F(A,B,C)] \\ - [F(a,b,C) - F(A,b,C)] - [F(a,B,c) - F(A,B,c)]\end{array}}{4}$$

β-effect and γ-effect have expressions similar to that for α-effect. If we distribute the $\alpha\beta\gamma$-interaction effect equally among the three main effects, we obtain the formulas in (3.18) through (3.20). All three two-factor interaction effects in the difference R_2-R_1 corresponding to the model (A5) turn out to be zero.

For any number of factors, the difference F(a,b,c,...) − F(A,B,C,...) involves two-factor interaction effects such as (E_{ab}-E_{AB}), each of which vanishes because of the conditions similar to those in (A7). For example, the two equations $E_{AB}+E_{aB}=0$ and $E_{aB}+E_{ab}=0$ together give $E_{AB}=E_{ab}$. Thus, the two-factor interaction effects are always zero regardless of the number of factors involved. This provides a justification for writing the difference R_2-R_1 in terms of only the main effects, as in (A1) above.

A.3 DERIVATION OF FORMULAS IN (5.16)

$$\frac{N_{ijk}}{N_{...}} = A_{ijk}\, B_{ijk}\, C_{ijk}\,, \tag{A9}$$

where A_{ijk}, B_{ijk}, and C_{ijk} involve ratios which represent, respectively, the I-effect, the J-effect, and the K-effect.

We write these three quantities as

$$A_{ijk} = \left(\frac{N_{ijk}}{N_{.jk}}\right)^x \cdot \left(\frac{N_{ij.}}{N_{.j.}} \cdot \frac{N_{i.k}}{N_{..k}}\right)^y \cdot \left(\frac{N_{i..}}{N_{...}}\right)^z, \tag{A10}$$

$$B_{ijk} = \left(\frac{N_{ijk}}{N_{i.k}}\right)^x \cdot \left(\frac{N_{ij.}}{N_{i..}} \cdot \frac{N_{.jk}}{N_{..k}}\right)^y \cdot \left(\frac{N_{.j.}}{N_{...}}\right)^z, \tag{A11}$$

$$C_{ijk} = \left(\frac{N_{ijk}}{N_{ij.}}\right)^x \cdot \left(\frac{N_{i.k}}{N_{i..}} \cdot \frac{N_{.jk}}{N_{.j.}}\right)^y \cdot \left(\frac{N_{..k}}{N_{...}}\right)^z, \tag{A12}$$

where x, y, z are suitably chosen exponents corresponding to the ratios with 0, 1, and 2 dots in the numerators, respectively.

Substituting (A10) through (A12) on the right-hand side of (A9) and then equating the exponents from both sides, we have

$$x = z = 1/3\,, \quad y = 1/6\,.$$

Substituting these values in (A10) through (A12), we obtain the formulas in (5.16).

A.4 DERIVATION OF FORMULAS (6.4) AND (6.5)

Six Consistent Sets for Three Populations

Set no.	Effects of factor α			Standardized rates controlled for all factors except α		
	$\alpha_{12.3}$	$\alpha_{13.2}$	$\alpha_{23.1}$	$\alpha_{1.23}$	$\alpha_{2.13}$	$\alpha_{3.12}$
1	α_{12}	α_{13}	$\alpha_{13}-\alpha_{12}$	$\alpha_{1.2}$	$\alpha_{2.1}$	$\alpha_{1.2}+(\alpha_{3.1}-\alpha_{1.3})$
2				$\alpha_{1.3}$	$\alpha_{1.3}+(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{3.1}$
3	α_{12}	$\alpha_{12}+\alpha_{23}$	α_{23}	$\alpha_{1.2}$	$\alpha_{2.1}$	$\alpha_{2.1}+(\alpha_{3.2}-\alpha_{2.3})$
4				$\alpha_{2.3}-(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{2.3}$	$\alpha_{3.2}$
5	$\alpha_{13}-\alpha_{23}$	α_{13}	α_{23}	$\alpha_{1.3}$	$\alpha_{3.1}-(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{3.1}$
6				$\alpha_{3.2}-(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{2.3}$	$\alpha_{3.2}$

A.5 DERIVATION OF FORMULAS (6.7) AND (6.8)

Forty-Eight Consistent Sets for Four Populations

Set no.	Effects of factor α						Standardized rates controlled for all factors except α			
	$\alpha_{12.34}$	$\alpha_{13.24}$	$\alpha_{14.23}$	$\alpha_{23.14}$	$\alpha_{24.13}$	$\alpha_{34.12}$	$\alpha_{1.234}$	$\alpha_{2.134}$	$\alpha_{3.124}$	$\alpha_{4.123}$
1 . . .	α_{12}	α_{13}	α_{14}	$\alpha_{13}-\alpha_{12}$	$\alpha_{14}-\alpha_{12}$	$\alpha_{14}-\alpha_{13}$	$\alpha_{1.2}$	$\alpha_{2.1}$	$\alpha_{1.2}+(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{1.2}+(\alpha_{4.1}-\alpha_{1.4})$
2 . . .							$\alpha_{1.3}$	$\alpha_{1.3}+(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{3.1}$	$\alpha_{1.3}+(\alpha_{4.1}-\alpha_{1.4})$
3 . . .							$\alpha_{1.4}$	$\alpha_{1.4}+(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{1.4}+(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{4.1}$
4 . . .	α_{12}	α_{13}	$\alpha_{12}+\alpha_{24}$	$\alpha_{13}-\alpha_{12}$	α_{24}	$\alpha_{12}+\alpha_{24}-\alpha_{13}$	$\alpha_{1.2}$	$\alpha_{2.1}$	$\alpha_{1.2}+(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{2.1}+(\alpha_{4.2}-\alpha_{2.4})$
5 . . .							$\alpha_{1.3}$	$\alpha_{1.3}+(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{3.1}$	$\alpha_{1.3}+(\alpha_{2.1}-\alpha_{1.2})+(\alpha_{4.2}-\alpha_{2.4})$
6 . . .							$\alpha_{2.4}-(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{2.4}$	$\alpha_{2.4}-(\alpha_{2.1}-\alpha_{1.2})+(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{4.2}$
7 . . .	α_{12}	α_{13}	$\alpha_{13}+\alpha_{34}$	$\alpha_{13}-\alpha_{12}$	$\alpha_{13}+\alpha_{34}-\alpha_{12}$	α_{34}	$\alpha_{1.2}$	$\alpha_{2.1}$	$\alpha_{1.2}+(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{1.2}+(\alpha_{3.1}-\alpha_{1.3})+(\alpha_{4.3}-\alpha_{3.4})$
8 . . .							$\alpha_{1.3}$	$\alpha_{1.3}+(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{3.1}$	$\alpha_{3.1}+(\alpha_{4.3}-\alpha_{3.4})$
9 . . .							$\alpha_{3.4}-(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{3.4}+(\alpha_{2.1}-\alpha_{1.2})-(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{3.4}$	$\alpha_{4.3}$
10 . .	α_{12}	$\alpha_{12}+\alpha_{23}$	α_{14}	α_{23}	$\alpha_{14}-\alpha_{12}$	$\alpha_{14}-\alpha_{12}-\alpha_{23}$	$\alpha_{1.2}$	$\alpha_{2.1}$	$\alpha_{2.1}+(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{1.2}+(\alpha_{4.1}-\alpha_{1.4})$
11 . .							$\alpha_{1.4}$	$\alpha_{1.4}+(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{1.4}+(\alpha_{2.1}-\alpha_{1.2})+(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{4.1}$
12 . .							$\alpha_{2.3}-(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{2.3}$	$\alpha_{3.2}$	$\alpha_{2.3}-(\alpha_{2.1}-\alpha_{1.2})+(\alpha_{4.1}-\alpha_{1.4})$
13 . .	α_{12}	$\alpha_{14}-\alpha_{34}$	α_{14}	$\alpha_{14}-\alpha_{12}-\alpha_{34}$	$\alpha_{14}-\alpha_{12}$	α_{34}	$\alpha_{1.2}$	$\alpha_{2.1}$	$\alpha_{1.2}+(\alpha_{4.1}-\alpha_{1.4})-(\alpha_{4.3}-\alpha_{3.4})$	$\alpha_{1.2}+(\alpha_{4.1}-\alpha_{1.4})$
14 . .							$\alpha_{1.4}$	$\alpha_{1.4}+(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{4.1}-(\alpha_{4.3}-\alpha_{3.4})$	$\alpha_{4.1}$
15 . .							$\alpha_{4.3}-(\alpha_{4.1}-\alpha_{1.4})$	$\alpha_{4.3}-(\alpha_{4.1}-\alpha_{1.4})+(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{3.4}$	$\alpha_{4.3}$
16 . .	α_{12}	$\alpha_{12}+\alpha_{23}$	$\alpha_{12}+\alpha_{24}$	α_{23}	α_{24}	$\alpha_{24}-\alpha_{23}$	$\alpha_{1.2}$	$\alpha_{2.1}$	$\alpha_{2.1}+(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{2.1}+(\alpha_{4.2}-\alpha_{2.4})$
17 . .							$\alpha_{2.3}-(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{2.3}$	$\alpha_{3.2}$	$\alpha_{2.3}+(\alpha_{4.2}-\alpha_{2.4})$
18 . .							$\alpha_{2.4}-(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{2.4}$	$\alpha_{2.4}+(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{4.2}$
19 . .	α_{12}	$\alpha_{12}+\alpha_{23}$	$\alpha_{12}+\alpha_{23}+\alpha_{34}$	α_{23}	$\alpha_{23}+\alpha_{34}$	α_{34}	$\alpha_{1.2}$	$\alpha_{2.1}$	$\alpha_{2.1}+(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{2.1}+(\alpha_{3.2}-\alpha_{2.3})+(\alpha_{4.3}-\alpha_{3.4})$
20 . .							$\alpha_{2.3}-(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{2.3}$	$\alpha_{3.2}$	$\alpha_{3.2}+(\alpha_{4.3}-\alpha_{3.4})$
21 . .							$\alpha_{3.4}-(\alpha_{3.2}-\alpha_{2.3})-(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{3.4}-(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{3.4}$	$\alpha_{4.3}$
22 . .	α_{12}	$\alpha_{12}+\alpha_{24}-\alpha_{34}$	$\alpha_{12}+\alpha_{24}$	$\alpha_{24}-\alpha_{34}$	α_{24}	α_{34}	$\alpha_{1.2}$	$\alpha_{2.1}$	$\alpha_{2.1}+(\alpha_{4.2}-\alpha_{2.4})-(\alpha_{4.3}-\alpha_{3.4})$	$\alpha_{2.1}+(\alpha_{4.2}-\alpha_{2.4})$
23 . .							$\alpha_{2.4}-(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{2.4}$	$\alpha_{4.2}-(\alpha_{4.3}-\alpha_{3.4})$	$\alpha_{4.2}$
24 . .							$\alpha_{4.3}-(\alpha_{4.2}-\alpha_{2.4})-(\alpha_{2.1}-\alpha_{1.2})$	$\alpha_{4.3}-(\alpha_{4.2}-\alpha_{2.4})$	$\alpha_{3.4}$	$\alpha_{4.3}$
25 . .	$\alpha_{13}-\alpha_{23}$	α_{13}	α_{14}	α_{23}	$\alpha_{23}+\alpha_{14}-\alpha_{13}$	$\alpha_{14}-\alpha_{13}$	$\alpha_{1.3}$	$\alpha_{3.1}-(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{3.1}$	$\alpha_{1.3}+(\alpha_{4.1}-\alpha_{1.4})$
26 . .							$\alpha_{1.4}$	$\alpha_{1.4}+(\alpha_{3.1}-\alpha_{1.3})-(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{1.4}+(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{4.1}$
27 . .							$\alpha_{3.2}-(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{2.3}$	$\alpha_{3.2}$	$\alpha_{3.2}-(\alpha_{3.1}-\alpha_{1.3})+(\alpha_{4.1}-\alpha_{1.4})$
28 . .	$\alpha_{14}-\alpha_{24}$	α_{13}	α_{14}	$\alpha_{13}+\alpha_{24}-\alpha_{14}$	α_{24}	$\alpha_{14}-\alpha_{13}$	$\alpha_{1.3}$	$\alpha_{1.3}+(\alpha_{4.1}-\alpha_{1.4})-(\alpha_{4.2}-\alpha_{2.4})$	$\alpha_{3.1}$	$\alpha_{3.1}+(\alpha_{4.1}-\alpha_{1.4})$
29 . .							$\alpha_{1.4}$	$\alpha_{4.1}-(\alpha_{4.2}-\alpha_{2.4})$	$\alpha_{1.4}+(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{4.1}$
30 . .							$\alpha_{4.2}-(\alpha_{4.1}-\alpha_{1.4})$	$\alpha_{2.4}$	$\alpha_{4.2}-(\alpha_{4.1}-\alpha_{1.4})+(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{4.2}$
31 . .	$\alpha_{1}\ \alpha_{23}$	α_{13}	$\alpha_{13}+\alpha_{24}-\alpha_{23}$	α_{23}	α_{24}	$\alpha_{24}-\alpha_{23}$	$\alpha_{1.3}$	$\alpha_{3.1}-(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{3.1}$	$\alpha_{3.1}-(\alpha_{3.2}-\alpha_{2.3})+(\alpha_{4.2}-\alpha_{2.4})$
32 . .							$\alpha_{3.2}-(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{2.3}$	$\alpha_{3.2}$	$\alpha_{2.3}+(\alpha_{4.2}-\alpha_{2.4})$
33 . .							$\alpha_{2.4}+(\alpha_{3.2}-\alpha_{2.3})-(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{2.4}$	$\alpha_{2.4}+(\alpha_{3.2}-\alpha_{2.3}$	$\alpha_{4.2}$
34 . .	$\alpha_{13}-\alpha_{23}$	α_{13}	$\alpha_{13}+\alpha_{34}$	α_{23}	$\alpha_{23}+\alpha_{34}$	α_{34}	$\alpha_{1.3}$	$\alpha_{3.1}-(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{3.1}$	$\alpha_{3.1}+(\alpha_{4.3}-\alpha_{3.4})$
35 . .							$\alpha_{3.2}-(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{2.3}$	$\alpha_{3.2}$	$\alpha_{3.2}+(\alpha_{4.3}-\alpha_{3.4})$
36 . .							$\alpha_{3.4}-(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{3.4}-(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{3.4}$	$\alpha_{4.3}$
37 . .	$\alpha_{13}+\alpha_{34}-\alpha_{24}$	α_{13}	$\alpha_{13}+\alpha_{34}$	$\alpha_{24}-\alpha_{34}$	α_{24}	α_{34}	$\alpha_{1.3}$	$\alpha_{3.1}+(\alpha_{4.3}-\alpha_{3.4})-(\alpha_{4.2}-\alpha_{2.4})$	$\alpha_{3.1}$	$\alpha_{3.1}+(\alpha_{4.3}-\alpha_{3.4})$
38 . .							$\alpha_{4.2}-(\alpha_{4.3}-\alpha_{3.4})-(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{2.4}$	$\alpha_{4.2}-(\alpha_{4.3}-\alpha_{3.4})$	$\alpha_{4.2}$
39 . .							$\alpha_{3.4}-(\alpha_{3.1}-\alpha_{1.3})$	$\alpha_{4.3}-(\alpha_{4.2}-\alpha_{2.4})$	$\alpha_{3.4}$	$\alpha_{4.3}$

Forty-Eight Consistent Sets for Four Populations—Continued

Set no.	Effects of factor α						Standardized rates controlled for all factors except α			
	$\alpha_{12.34}$	$\alpha_{13.24}$	$\alpha_{14.23}$	$\alpha_{23.14}$	$\alpha_{24.13}$	$\alpha_{34.12}$	$\alpha_{1.234}$	$\alpha_{2.134}$	$\alpha_{3.124}$	$\alpha_{4.123}$
40 ..	$\alpha_{14}-\alpha_{24}$	$\alpha_{14}+\alpha_{23}-\alpha_{24}$	α_{14}	α_{23}	α_{24}	α_{34}	$\alpha_{1.4}$	$\alpha_{4.1}-(\alpha_{4.2}-\alpha_{2.4})$	$\alpha_{4.1}-(\alpha_{4.2}-\alpha_{2.4})+(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{4.1}$
41 ..							$\alpha_{2.3}+(\alpha_{4.2}-\alpha_{2.4})-(\alpha_{4.1}-\alpha_{1.4})$	$\alpha_{2.3}$	$\alpha_{3.2}$	$\alpha_{2.3}+(\alpha_{4.2}-\alpha_{2.4})$
42 ..							$\alpha_{4.2}-(\alpha_{4.1}-\alpha_{1.4})$	$\alpha_{2.4}$	$\alpha_{2.4}+(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{4.2}$
43 ..	$\alpha_{14}-\alpha_{34}-\alpha_{23}$	$\alpha_{14}-\alpha_{34}$	α_{14}	α_{23}	$\alpha_{23}+\alpha_{34}$	α_{34}	$\alpha_{1.4}$	$\alpha_{4.1}-(\alpha_{4.3}-\alpha_{3.4})-(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{4.1}-(\alpha_{4.3}-\alpha_{3.4})$	$\alpha_{4.1}$
44 ..							$\alpha_{3.2}+(\alpha_{4.3}-\alpha_{3.4})-(\alpha_{4.1}-\alpha_{1.4})$	$\alpha_{2.3}$	$\alpha_{3.2}$	$\alpha_{3.2}+(\alpha_{4.3}-\alpha_{3.4})$
45 ..							$\alpha_{4.3}-(\alpha_{4.1}-\alpha_{1.4})$	$\alpha_{3.4}-(\alpha_{3.2}-\alpha_{2.3})$	$\alpha_{3.4}$	$\alpha_{4.3}$
46 ..	$\alpha_{14}-\alpha_{24}$	$\alpha_{14}-\alpha_{34}$	α_{14}	$\alpha_{24}-\alpha_{34}$	α_{24}	α_{34}	$\alpha_{1.4}$	$\alpha_{4.1}-(\alpha_{4.2}-\alpha_{2.4})$	$\alpha_{4.1}-(\alpha_{4.3}-\alpha_{3.4})$	$\alpha_{4.1}$
47 ..							$\alpha_{4.2}-(\alpha_{4.1}-\alpha_{1.4})$	$\alpha_{2.4}$	$\alpha_{4.2}-(\alpha_{4.3}-\alpha_{3.4})$	$\alpha_{4.2}$
48 ..							$\alpha_{4.3}-(\alpha_{4.1}-\alpha_{1.4})$	$\alpha_{4.3}-(\alpha_{4.2}-\alpha_{2.4})$	$\alpha_{3.4}$	$\alpha_{4.3}$

A.6 SUMMARY OF FORMULAS IN CHAPTER 2

α, β, γ, ... are the factors that assume values A, B, C, ... in population 1 and a, b, c, ... in population 2. The rate $R = \alpha\beta\gamma\ldots$, so that in population 1 and population 2, $R_1 = ABC\ldots$ and $R_2 = abc\ldots$. We define Q corresponding to the number of factors 2, 3, 4, 5, and 6, respectively, as

$$Q = \frac{b+B}{2},$$

$$Q = \frac{bc+BC}{3} + \frac{bC+Bc}{6},$$

$$Q = \frac{bcd+BCD}{4} + \frac{bcD+bCd+Bcd+BCd+BcD+bCD}{12},$$

$$Q = \frac{bcde+BCDE}{5} + \frac{bcdE+bcDe+bCde+Bcde+BCDe+BCdE+BcDE+bCDE}{20} + \frac{bcDE+bCdE+bCDe+BCde+BcDe+BcdE}{30},$$

$$Q = \frac{bcdef+BCDEF}{6} + \frac{bcdeF+bcdEf+bcDef+bCdef+Bcdef+BCDEf+BCDeF+BCdEF+BcDEF+bCDEF}{30} + \frac{\begin{matrix}bcdEF+bcDeF+bcDEf+bCdeF+bCdEf+bCDef+BcdeF+BcdEf+BcDef+BCdef\\+BCDef+BCdEf+BCdeF+BcDEf+BcDeF+BcdEF+bCDEf+bCDeF+bCdEF+bcDEF\end{matrix}}{60}.$$

The β-standardized rate, $\beta\gamma$-standardized rate, $\beta\gamma\delta$-standardized rate, $\beta\gamma\delta\epsilon$-standardized rate, and $\beta\gamma\delta\epsilon\eta$-standardized rate in population 1 corresponding to, respectively, 2, 3, 4, 5, and 6 factors are given by QA, when the appropriate Q is chosen from the above. The corresponding standardized rates in population 2 are Qa. The numbers in the denominators of the above expressions are P, $P\binom{P-1}{1}$, $P\binom{P-1}{2}$, ..., where P is the number of factors.

A.7 SUMMARY OF FORMULAS IN CHAPTER 3

Using notation as in section A.6, the rate $R = F(\alpha,\beta,\gamma,\ldots)$, so that $R_1 = F(A,B,C,\ldots)$, and $R_2 = F(a,b,c,\ldots)$. We define Q(A) corresponding to the number of factors 2, 3, 4, 5, and 6, respectively, as

$$Q(A) = \frac{F(A,b) + F(A,B)}{2},$$

$$Q(A) = \frac{F(A,b,c) + F(A,B,C)}{3} + \frac{F(A,b,C) + F(A,B,c)}{6},$$

$$Q(A) = \frac{F(A,b,c,d) + F(A,B,C,D)}{4} + \frac{F(A,b,c,D) + F(A,b,C,d) + F(A,B,c,d) + F(A,B,C,d) + F(A,B,c,D) + F(A,b,C,D)}{12},$$

$$Q(A) = \frac{F(A,b,c,d,e) + F(A,B,C,D,E)}{5} + \frac{\begin{matrix}F(A,b,c,d,E) + F(A,b,c,D,e) + F(A,b,C,d,e) + F(A,B,c,d,e)\\ + F(A,B,C,D,e) + F(A,B,C,d,E) + F(A,B,c,D,E) + F(A,b,C,D,E)\end{matrix}}{20}$$

$$+\frac{F(A,b,c,D,E)+F(A,b,C,d,E)+F(A,b,C,D,e)+F(A,B,C,d,e)+F(A,B,c,D,e)+F(A,B,c,d,E)}{30},$$

$$Q(A)=\frac{F(A,b,c,d,e,f)+F(A,B,C,D,E,F)}{6}$$

$$+\frac{\begin{gathered}F(A,b,c,d,e,F)+F(A,b,c,d,E,f)+F(A,b,c,D,e,f)+F(A,b,C,d,e,f)+F(A,B,c,d,e,f)\\+F(A,B,C,D,E,f)+F(A,B,C,D,e,F)+F(A,B,C,d,E,F)+F(A,B,c,D,E,F)+F(A,b,C,D,E,F)\end{gathered}}{30}$$

$$+\frac{\begin{gathered}F(A,b,c,d,E,F)+F(A,b,c,D,e,F)+F(A,b,c,D,E,f)+F(A,b,C,d,e,F)+F(A,b,C,d,E,f)\\+F(A,b,C,D,e,f)+F(A,B,c,d,e,F)+F(A,B,c,d,E,f)+F(A,B,c,D,e,f)+F(A,B,C,d,e,f)\\+F(A,B,C,D,e,f)+F(A,B,C,d,E,f)+F(A,B,C,d,e,F)+F(A,B,c,D,E,f)+F(A,B,c,D,e,F)\\+F(A,B,c,d,E,F)+F(A,b,C,D,E,f)+F(A,b,C,D,e,F)+F(A,b,C,d,E,F)+F(A,b,c,D,E,F)\end{gathered}}{60}.$$

The β-standardized rate, $\beta\gamma$-standardized rate, $\beta\gamma\delta$-standardized rate, $\beta\gamma\delta\epsilon$-standardized rate, and $\beta\gamma\delta\epsilon\eta$-standardized rate in population 1 corresponding to, respectively, 2, 3, 4, 5, and 6 factors are given by Q(A), when the appropriate Q(A) is chosen from the above. The corresponding standardized rates in population 2 are Q(a). Obviously, the formulas in section A.6 can be derived as special cases of those in this section by substituting $\alpha\beta\gamma$... for $F(\alpha,\beta,\gamma,...)$.

A.8 SUMMARY OF FORMULAS IN CHAPTER 4

Using the vector notation for the scalars in section A.7, the rate $R = F(\overline{\alpha},\overline{\beta},\overline{\gamma},...)$, so that $R_1 = F(\overline{A},\overline{B},\overline{C},...)$, and $R_2 = F(\overline{a},\overline{b},\overline{c},...)$. We define $Q(\overline{A})$ corresponding to the number of factors 2, 3, 4, 5, and 6 exactly the same way as in section A.7 except that the scalars A, B, C, D, E, F, a, b, c, d, e, and f in the equations are now replaced by the corresponding vectors $\overline{A}$, $\overline{B}$, $\overline{C}$, $\overline{D}$, $\overline{E}$, $\overline{F}$, $\overline{a}$, $\overline{b}$, $\overline{c}$, $\overline{d}$, $\overline{e}$, and $\overline{f}$. As shown in section A.7, the standardized rates in population 1 are given by $Q(\overline{A})$'s and those in population 2 are given by $Q(\overline{a})$'s.

A.9 SUMMARY OF FORMULAS IN CHAPTER 5

When there is only one factor I, N_i and T_i are the number of persons and the rate for the ith category of I, and $N_.$ and $T_.$ are the total number of persons and the crude rate, in population 1. When there are two factors I and J, N_{ij} and T_{ij} are the number of persons and the rate for the (i,j)-category of I and J, $N_{i.}$ and $T_{i.}$ are the number of persons and the rate for the ith category of I, $N_{.j}$ and $T_{.j}$ are the number of persons and the rate for the jth category of J, and $N_{..}$ and $T_{..}$ are the total number of persons and the crude rate, in population 1. Analogous symbols are used for population 2 with lower-case letters n and t. For higher number of factors I, J, K, L, ..., the symbols are extended along the same lines.

For number of factors 1, 2, 3, 4, 5, and 6, the A's are defined, respectively, as follows:

$$A_i = \frac{N_i}{N_.},$$

$$A_{ij} = \left(\frac{N_{ij}}{N_{.j}}\cdot\frac{N_{i.}}{N_{..}}\right)^{\frac{1}{2}},$$

$$A_{ijk} = \left(\frac{N_{ijk}}{N_{.jk}}\right)^{\frac{1}{3}}\cdot\left(\frac{N_{ij.}}{N_{.j.}}\cdot\frac{N_{i.k}}{N_{..k}}\right)^{\frac{1}{6}}\cdot\left(\frac{N_{i..}}{N_{...}}\right)^{\frac{1}{3}},$$

$$A_{ijkl} = \left(\frac{N_{ijkl}}{N_{.jkl}}\right)^{\frac{1}{4}}\cdot\left(\frac{N_{ijk.}}{N_{.jk.}}\cdot\frac{N_{ij.l}}{N_{.j.l}}\cdot\frac{N_{i.kl}}{N_{..kl}}\right)^{\frac{1}{12}}\cdot\left(\frac{N_{i..l}}{N_{...l}}\cdot\frac{N_{i.k.}}{N_{..k.}}\cdot\frac{N_{ij..}}{N_{.j..}}\right)^{\frac{1}{12}}\cdot\left(\frac{N_{i...}}{N_{....}}\right)^{\frac{1}{4}},$$

$$A_{ijklm} = \left(\frac{N_{ijklm}}{N_{.jklm}}\right)^{\frac{1}{5}} \cdot \left(\frac{N_{ijkl.}}{N_{.jkl.}} \cdot \frac{N_{ijk.m}}{N_{.jk.m}} \cdot \frac{N_{ij.lm}}{N_{.j.lm}} \cdot \frac{N_{i.klm}}{N_{..klm}}\right)^{\frac{1}{20}} \cdot$$

$$\left(\frac{N_{ijk..}}{N_{.jk..}} \cdot \frac{N_{ij.l.}}{N_{.j.l.}} \cdot \frac{N_{ij..m}}{N_{.j..m}} \cdot \frac{N_{i.kl.}}{N_{..kl.}} \cdot \frac{N_{i.k.m}}{N_{..k.m}} \cdot \frac{N_{i..lm}}{N_{...lm}}\right)^{\frac{1}{30}} \cdot$$

$$\left(\frac{N_{i...m}}{N_{....m}} \cdot \frac{N_{i..l.}}{N_{...l.}} \cdot \frac{N_{i.k..}}{N_{..k..}} \cdot \frac{N_{ij...}}{N_{.j...}}\right)^{\frac{1}{20}} \cdot \left(\frac{N_{i....}}{N_{.....}}\right)^{\frac{1}{5}},$$

$$A_{ijklmn} = \left(\frac{N_{ijklmn}}{N_{.jklmn}}\right)^{\frac{1}{6}} \cdot \left(\frac{N_{ijklm.}}{N_{.jklm.}} \cdot \frac{N_{ijkl.n}}{N_{.jkl.n}} \cdot \frac{N_{ijk.mn}}{N_{.jk.mn}} \cdot \frac{N_{ij.lmn}}{N_{.j.lmn}} \cdot \frac{N_{i.klmn}}{N_{..klmn}}\right)^{\frac{1}{30}} \cdot$$

$$\left(\frac{N_{ijkl..}}{N_{.jkl..}} \cdot \frac{N_{ijk.m.}}{N_{.jk.m.}} \cdot \frac{N_{ijk..n}}{N_{.jk..n}} \cdot \frac{N_{ij.lm.}}{N_{.j.lm.}} \cdot \frac{N_{ij.l.n}}{N_{.j.l.n}} \cdot \frac{N_{ij..mn}}{N_{.j..mn}} \cdot \frac{N_{i.klm.}}{N_{..klm.}} \cdot \frac{N_{i.kl.n}}{N_{..kl.n}} \cdot \frac{N_{i.k.mn}}{N_{..k.mn}} \cdot \frac{N_{i..lmn}}{N_{...lmn}}\right)^{\frac{1}{60}} \cdot$$

$$\left(\frac{N_{i...mn}}{N_{....mn}} \cdot \frac{N_{i..l.n}}{N_{...l.n}} \cdot \frac{N_{i..lm}}{N_{...lm}} \cdot \frac{N_{i.k..n}}{N_{..k..n}} \cdot \frac{N_{i.k.m.}}{N_{..k.m.}} \cdot \frac{N_{i.kl..}}{N_{..kl..}} \cdot \frac{N_{ij...n}}{N_{.j...n}} \cdot \frac{N_{ij..m.}}{N_{.j..m.}} \cdot \frac{N_{ij.l..}}{N_{.j.l..}} \cdot \frac{N_{ijk...}}{N_{.jk...}}\right)^{\frac{1}{60}} \cdot$$

$$\left(\frac{N_{i....n}}{N_{.....n}} \cdot \frac{N_{i...m.}}{N_{....m.}} \cdot \frac{N_{i..l..}}{N_{...l..}} \cdot \frac{N_{i.k...}}{N_{..k...}} \cdot \frac{N_{ij....}}{N_{.j....}}\right)^{\frac{1}{30}} \cdot \left(\frac{N_{i.....}}{N_{......}}\right)^{\frac{1}{6}}.$$

Like the numbers in the denominators in the expressions in sections A.6 and A.7, the exponents in the above expressions are the reciprocals of P, $P\binom{P-1}{1}$, $P\binom{P-1}{2}$, ... , where P is the number of factors. Similar expressions for B's, C's, D's, ... are obtained from those for A's above by interchanging, respectively, i and j, i and k, i and l, ... a's, b's, c's, d's, ... are obtained from A's, B's, C's, D's, ... by using n's in place of N's.

The I-standardized rate, (I,J)-standardized rate, (I,J,K)-standardized rate, (I,J,K,L)-standardized rate, (I,J,K,L,M)-standardized rate, and (I,J,K,L,M,N)-standardized rate in population 1 corresponding to 1, 2, 3, 4, 5, and 6 factors are denoted by $R(\overline{T})$ which are, respectively,

$$R(\overline{T}) = \sum_{i} \frac{\frac{n_i}{n.} + \frac{N_i}{N.}}{2} T_i,$$

$$R(\overline{T}) = \sum_{i,j} \frac{\frac{n_{ij}}{n..} + \frac{N_{ij}}{N..}}{2} T_{ij},$$

$$R(\overline{T}) = \sum_{i,j,k} \frac{\frac{n_{ijk}}{n...} + \frac{N_{ijk}}{N...}}{2} T_{ijk},$$

$$R(\overline{T}) = \sum_{i,j,k,l} \frac{\frac{n_{ijkl}}{n....} + \frac{N_{ijkl}}{N....}}{2} T_{ijkl},$$

$$R(\overline{T}) = \sum_{i,j,k,l,m} \frac{\frac{n_{ijklm}}{n.....} + \frac{N_{ijklm}}{N.....}}{2} T_{ijklm},$$

$$R(\overline{T}) = \sum_{i,j,k,l,m,n} \frac{\frac{n_{ijklmn}}{n......} + \frac{N_{ijklmn}}{N......}}{2} T_{ijklmn}.$$

The corresponding standardized rates in population 2 are $R(\overline{t})$, which are obtained from the above expressions by replacing T's by the corresponding t's.

The R-standardized rate, (J,R)-standardized rate, (J,K,R)-standardized rate, (J,K,L,R)-standardized rate, (J,K,L,M,R)-standardized rate, and (J,K,L,M,N,R)-standardized rate in population 1 corresponding to 1, 2, 3, 4, 5, and 6 factors are denoted by $I(\bar{A})$ which are, respectively,

$$I(\bar{A}) = \sum_i \frac{t_i + T_i}{2} A_i ,$$

$$I(\bar{A}) = \sum_{i,j} \frac{t_{ij} + T_{ij}}{2} \text{ [QA for 2 factors in section A.6 with subscripts ij in each letter]},$$

$$I(\bar{A}) = \sum_{i,j,k} \frac{t_{ijk} + T_{ijk}}{2} \text{ [QA for 3 factors in section A.6 with subscripts ijk in each letter]},$$

$$I(\bar{A}) = \sum_{i,j,k,l} \frac{t_{ijkl} + T_{ijkl}}{2} \text{ [QA for 4 factors in section A.6 with subscripts ijkl in each letter]},$$

$$I(\bar{A}) = \sum_{i,j,k,l,m} \frac{t_{ijklm} + T_{ijklm}}{2} \text{ [QA for 5 factors in section A.6 with subscripts ijklm in each letter]},$$

$$I(\bar{A}) = \sum_{i,j,k,l,m,n} \frac{t_{ijklmn} + T_{ijklmn}}{2} \text{ [QA for 6 factors in section A.6 with subscripts ijklmn in each letter]}.$$

The corresponding standardized rates in population 2 are $I(\bar{a})$, which are obtained from the above expressions by replacing A's by the corresponding a's.

A.10 SUMMARY OF FORMULAS IN CHAPTER 6

When there are two populations 1 and 2, α_{12} denotes the factor effect of α and $\alpha_{1.2}$ denotes the standardized rate in population 1 controlled for all other factors except α. When there are three populations 1, 2, and 3, $\alpha_{12.3}$ and $\alpha_{1.23}$ denote the corresponding numbers when populations 1 and 2 are compared (in the presence of population 3). For four and higher number of populations, analogous symbols are used.

The standardized rates in population 1 controlled for all other factors except α in 3, 4, 5, and N populations are, respectively, given by

$$\alpha_{1.23} = \frac{\sum_{i=2}^{3} \alpha_{1.i}}{2} + \frac{\sum_{i=2}^{3} [\sum_{j\neq 1,i}^{3} \alpha_{i.j} - \alpha_{i.1}]}{6},$$

$$\alpha_{1.234} = \frac{\sum_{i=2}^{4} \alpha_{1.i}}{3} + \frac{\sum_{i=2}^{4} [\sum_{j\neq 1,i}^{4} \alpha_{i.j} - 2\alpha_{i.1}]}{12},$$

$$\alpha_{1.2345} = \frac{\sum_{i=2}^{5} \alpha_{1.i}}{4} + \frac{\sum_{i=2}^{5} [\sum_{j\neq 1,i}^{5} \alpha_{i.j} - 3\alpha_{i.1}]}{20},$$

$$\alpha_{1.23...N} = \frac{\sum_{i=2}^{N} \alpha_{1.i}}{N-1} + \frac{\sum_{i=2}^{N} [\sum_{j\neq 1,i}^{N} \alpha_{i.j} - (N-2)\alpha_{i.1}]}{N(N-1)}.$$

When there are 3, 4, 5, and N populations, the factor effects of α in the comparison of populations 1 and 2 are, respectively, given by

$$\alpha_{12.3} = \alpha_{12} - \frac{\sum_{j=3}^{3} (\alpha_{12} + \alpha_{2j} - \alpha_{1j})}{3},$$

$$\alpha_{12.34} = \alpha_{12} - \frac{\sum_{j=3}^{4} (\alpha_{12} + \alpha_{2j} - \alpha_{1j})}{4},$$

$$\alpha_{12.345} = \alpha_{12} - \frac{\sum_{j=3}^{5} (\alpha_{12}+\alpha_{2j}-\alpha_{1j})}{5},$$

$$\alpha_{12.34...N} = \alpha_{12} - \frac{\sum_{j=3}^{N} (\alpha_{12}+\alpha_{2j}-\alpha_{1j})}{N}.$$

Appendix B. References

Arriaga, Eduardo E. 1984. "Measuring and Explaining the Change in Life Expectancies." *Demography* 21(1): 83-96.

Bachu, Amara. 1981. *Urban and Rural Differentials in Fertility in India*. Ph.D. Dissertation, Howard University, Washington, D.C.

Bianchi, Suzanne M. and Nancy Rytina. 1986. "The Decline in Occupational Sex Segregation During the 1970s: Census and CPS Comparisons." *Demography* 23(1): 79-86.

Blake, Judith and Prithwis Das Gupta. 1976. "Components of the Decline in American Marital Fertility Between 1960 and 1970." Manuscript, University of California: Berkeley.

Bongaarts, John. 1978. "A Framework for Analyzing the Proximate Determinants of Fertility." *Population and Development Review* 4(1): 105-132.

Cho, Lee-Jay and Robert D. Retherford. 1973. "Comparative Analysis of Recent Fertility Trends in East Asia." *Proceedings of IUSSP International Population Conference* 2: 163-181.

Clogg, Clifford C. 1978. "Adjustment of Rates Using Multiplicative Models." *Demography* 15(4): 523-539.

Clogg, Clifford C. and Scott R. Eliason. 1988. "A Flexible Procedure for Adjusting Rates and Proportions, Including Statistical Methods for Group Comparisons." *American Sociological Review* 53: 267-283.

Curtin, Lester R., Jeffrey D. Maurer, and Harry M. Rosenberg. 1980. "On the Selection of a Standard Population for Computing Age-Adjusted Death Rates." *Proceedings of the Social Statistics Section, American Statistical Association:* 218-223.

Das Gupta, Prithwis. 1978. "A General Method of Decomposing a Difference Between Two Rates into Several Components." *Demography* 15(1): 99-112.

_______. 1984. "Contributions of Other Socio-Economic Factors to the Fertility Differentials of Women by Education: A Multivariate Approach." *Genus* 40(3-4): 117-127.

_______. 1987. "Comment on Bianchi and Rytina's 'The Decline in Occupational Sex Segregation During the 1970s: Census and CPS Comparisons.'" *Demography* 24(2): 291-295.

_______. 1988. "Methods of Decomposing the Difference Between Two Rates with Applications to the Study of Race-Sex Inequality in Earnings in the U.S." Paper presented at the 1988 Joint Statistical Meetings in New Orleans, Louisiana.

_______. 1989. "Methods of Decomposing the Difference Between Two Rates with Applications to Race-Sex Inequality in Earnings." *Mathematical Population Studies* 2(1): 15-36.

_______. 1990. "Decomposition of the Difference Between Two Rates When the Factors Are Nonmultiplicative with Applications to the U.S. Life Tables." Paper presented at the 1990 Annual Meeting of the Population Association of America in Toronto, Canada.

_______. 1991. "Decomposition of the Difference Between Two Rates and Its Consistency When More than Two Populations are Involved." *Mathematical Population Studies* 3(2): 105-125.

_______. 1992. "The Links Between Standardization of Rates and Decomposition of Rate Differences." Paper presented at the 1992 Joint Statistical Meetings in Boston, Massachusetts.

del Pinal, Jorge. 1989. "AIDS, Blacks and Hispanics: What is the Connection?" Paper presented at the 1989 Public Health Conference on Records and Statistics: National Center for Health Statistics.

Gibson, Campbell. 1976. "The U.S. Fertility Decline 1961-1975: The Contribution of Changes in Marital Status and Marital Fertility." *Family Planning Perspectives* 8: 249-252.

Hernandez, Donald J. 1984. *Success or Failure? Family Planning Programs in the Third World*. Westport, Connecticut: Greenwood Press.

Hoem, Jan M. 1987. "Statistical Analysis of a Multiplicative Model and Its Application to the Standardization of Vital Rates: A Review." *International Statistical Review* 55: 119-152.

Hogan, Howard. 1992. "The 1990 Post-Enumeration Survey: Operations and New Estimates." Paper presented at the 1992 Joint Statistical Meetings in Boston, Massachusetts.

Janowitz, Barbara S. 1976. "An Analysis of the Impact of Education on Family Size." *Demography* 13(2): 189-198.

Johansen, Robert J. 1990. "Proposed New Standard Population." *Proceedings of the Social Statistics Section, American Statistical Association:* 176-181.

Keyfitz, Nathan. 1968. *Introduction to the Mathematics of Population*. Reading, MA: Addison-Wesley.

Kim, Young J. and Donna M. Strobino. 1984. "Decomposition of the Difference Between Two Rates with Hierarchical Factors." *Demography* 21(3): 361-372.

Kitagawa, Evelyn M. 1955. "Components of a Difference Between Two Rates." *Journal of the American Statistical Association* 50(272): 1168-1194.

______. 1964. "Standardized Comparisons in Population Research." *Demography* 1: 296-315.

Kuczynski, Robert R. 1935. *The measurement of Population Growth: Methods and Results*. New York: Gordon and Breach.

Liao, Tim Futing. 1989. "A Flexible Approach for the Decomposition of Rate Differences." *Demography* 26(4): 717-726.

Little, R.J.A. and T.W. Pullum. 1979. "The General Linear Model and Direct Standardization: A Comparison." *Sociological Methods and Research* 7: 475-501

Moreno, Lorenzo. 1991. "An Alternative Model of the Impact of the Proximate Determinants on Fertility Change: Evidence from Latin America." *Population Studies* 45(2): 313-337.

Myers, George C. 1991. Presentation at the Annual Meeting of the Southern Demographic Association in Jacksonville, Florida.

Nathanson, Constance A. and Young J. Kim. 1989. "Components of Change in Adolescent Fertility, 1971-1979." *Demography* 26(1): 85-98.

National Center for Health Statistics. 1962. *Vital Statistics of the United States, 1960, Vol. I-Natality.*

______. 1963. *Vital Statistics of the United States, 1960, Vol. II-Mortality, Part A*: Washington.

______. 1964. *Vital Statistics of the United States, 1962, Vol. II-Mortality, Part A:* Washington.

______. 1967a. *Vital Statistics of the United States, 1965, Vol. I-Natality:* Washington.

______. 1967b. *Vital Statistics of the United States, 1965, Vol. II-Mortality, Part A*: Washington.

______. 1972. *A Study of Infant Mortality from Linked Records: Comparison of Neonatal Mortality from Two Cohort Studies, United States, January-March, 1950 and 1960*, Series 20, No. 13: Rockville, Maryland.

______. 1982. *Monthly Vital Statistics Report: Advance Report of Final Mortality Statistics, 1979*, Vol. 31, No. 6, Supplement: Hyattsville, Maryland.

______. 1984. *Vital Statistics of the United States, 1980, Vol. I-Natality*: Hyattsville, Maryland.

______. 1985. *U.S. Decennial Life Tables for 1979-81, Vol. I, No. 1, United States Life Tables*: Hyattsville, Maryland.

______. 1987. *Vital Statistics of the United States, 1983, Vol. II-Mortality, Part A*: Hyattsville, Maryland.

_______. 1990a. *Vital Statistics of the United States, 1988, Vol. I-Natality*: Hyattsville, Maryland.

_______. 1990b. *Vital Statistics of the United States, 1987, Vol. II-Mortality, Part A:* Hyattsville, Maryland.

_______. 1991a. *Monthly Vital Statistics Report: Annual Summary of Births, Marriages, Divorces, and Deaths:* United States, 1990 (Provisional Data), Vol. 39, No. 13: Hyattsville, Maryland.

_______. 1991b. *Monthly Vital Statistics Report: Advance Report of Final Natality Statistics*, 1989, Vol. 40, No. 8, Supplement: Hyattsville, Maryland.

_______. 1992. *Monthly Vital Statistics Report: Advance Report of Final Mortality Statistics*, 1989, Vol. 40, No. 8, Supplement 2: Hyattsville, Maryland.

National Office of Vital Statistics. 1956. *Vital Statistics of the United States, 1954*, Vol. I: Washington.

Pollard, J.H. 1988. "On the Decomposition of Changes in Expectation of Life and Differentials in Life Expectancy." *Demography* 25(2): 265-276.

Pullum, Thomas W., Lucky M. Tedrow, and Jerald R. Herting. 1989. "Measuring Change and Continuity in Parity Distributions." *Demography* 26(3): 485-498.

Robinson, J. Gregory and Bashir Ahmed. 1992. "Utility of Synthetic Estimates of Census Coverage for States Based on National Demographic Analysis Estimates." Paper presented at the 1992 Annual Meeting of the Population Association of America, Denver.

Ross, Christine, Sheldon Danziger, and Eugene Smolensky. 1987. "The Level and Trend of Poverty in the United States, 1939-1979." *Demography* 24(4): 587-600.

Ruggles, Steven. 1988. "The Demography of the Unrelated Individual: 1900-1950." *Demography* 25(4): 521-536.

Santi, Lawrence L. 1989. "Partialling and Purging: Equivalencies Between Log-Linear Analysis and the Purging Method of Rate Adjustment." *Sociological Methods and Research* 17(4): 376-397.

Scarborough, James B. 1962. *Numerical Mathematical Analysis*. Baltimore: The Johns Hopkins Press.

Smith, Herbert L. and Phillips Cutright. 1988. "Thinking About Change in Illegitimacy Ratios: United States, 1963-1983." *Demography* 25(2): 235-247.

Spencer, Gregory. 1980. "The Contributions of Childlessness and Non-marriage to Racial and Ethnic Differences in American Fertility." Paper presented at the 1980 Annual Meeting of the Population Association of America, Denver.

Spiegelman, M. and H.H. Marks. 1966. "Empirical Testing of Standards for the Age-Adjustment of Death Rates by the Direct Method." *Human Biology* 38: 280-292.

Suchindran, C.M. and Helen P. Koo. 1992. "Age at Last Birth and Its Components." *Demography* 29(2): 227-245.

Sweet, James A. 1984. "Components of Change in the Number of Households: 1970-1980." *Demography* 21(2): 129-140.

United Nations. 1988. *Demographic Yearbook 1986*, Department of International Economic and Social Affairs, Statistical Office: New York.

_______. 1989. *Demographic Yearbook 1987*, Department of International Economic and Social Affairs, Statistical Office: New York.

U.S. Bureau of the Census. 1946. *United States Life Tables and Actuarial Tables*, 1939-1941: Washington.

_______. 1965. *Estimates of the Population of the United States, by Single Years of Age, Color, and Sex, 1900 to 1959*, P-25, No. 311: Washington, D.C.

_______. 1971. *Marital Status and Family Status: March 1970*, P-20, No. 212: Washington, D.C.

_______. 1973. Women by Number of Children Ever Born, Census of Population 1970, Subject Report PC(2)-3A: Washington, D.C.

______. 1974. *Estimates of the Population of the United States, by Age, Sex, and Race: April 1, 1960 to July 1, 1973*, P-25, No. 519: Washington, D.C.

______. 1977. *Geographical Mobility: March 1975 to March 1976*, P-20, No. 305: Washington, D.C.

______. 1981. *Marital Status and Living Arrangements: March 1980*, P-20, No. 365: Washington, D.C.

______. 1982. *Preliminary Estimates of the Population of the United States, by Age, Sex, and Race: 1970* to 1981, P-25, No. 917: Washington, D.C.

______. 1984a. 1980 Census of Population: *Detailed Population Characteristics*, United States Summary, PC80-1-D1-A: Washington, D.C.

______. 1984b. *Detailed Occupation of the Experienced Civilian Labor Force by Sex for the United States and Regions: 1980 and 1970*, Supplementary Report PC80-S1-15: Washington, D.C.

______. 1984c. *Earnings by Occupation and Education*, 1980 Census of Population, Subject Report PC80-2-8B: Washington, D.C.

______. 1989. *Geographical Mobility: March 1986 to March 1987*, P-20, No. 430: Washington, D.C.

______. 1990a. *United States Population Estimates, by Age, Sex, Race, and Hispanic Origin: 1980 to 1988*, P-25, No. 1045: Washington, D.C.

______. 1990b. *U.S. Population Estimates, by Age, Sex, Race, and Hispanic Origin: 1989*, P-25, No. 1057: Washington, D.C.

______. 1992. *Studies in the Distribution of Income*, P-60, No. 183: Washington, D.C.

Wilson, Franklin D. 1988. "Components of Change in Migration and Destination-Propensity Rates for Metropolitan and Nonmetropolitan Areas: 1935-1980." *Demography* 25(1): 129-139.

Wojtkiewicz, Roger A., Sara S. McLanahan, and Irwin Garfinkel. 1990. "The Growth of Families Headed by Women: 1950-1980." *Demography* 27(1): 19-30.

Woolsey, T.D. 1959. "Adjusted Death Rates and Other Indices of Mortality," Chapter 4 in *Vital Statistics Rates in the United States, 1900-1940*. Washington, D.C.: Government Printing Office.

Xie, Yu. 1989. "An Alternative Purging Method: Controlling the Composition-Dependent Interaction in an Analysis of Rates." *Demography* 26(4): 711-716.

Appendix C. Author Index

U.S. Department of Commerce
BUREAU OF THE CENSUS
Washington, D.C. 20233

Official Business
Penalty for Private Use, $300

FIRST-CLASS MAIL
POSTAGE & FEES PAID
CENSUS
PERMIT No. G–58