Vienna Yearbook of Population Research 2025 (Vol. 23), pp. 1–27
https://doi.org/10.1553/p-j2pg-fm66

RESEARCH ARTICLE

Inequality, heterogeneity, and chance: Multiple factors and their interactions

Hal Caswell[1,2] and **Silke F. van Daalen**[2,3]

ABSTRACT A heterogeneous population is a mixture of groups differing in vital rates. In such a population, some of the variance in demographic outcomes (e.g., longevity, lifetime reproduction) is due to heterogeneity and some is the result of stochastic demographic processes. Many studies have partitioned variance into its between-group and within-group components, but have focused on single factors. Especially for longevity, variance due to stochasticity is far greater than variance due to heterogeneity. Here we extend such analyses to multiple-factor studies, making it possible to calculate the contributions to variance of each factor and each of the interactions among factors. We treat the population as a mixture and use the marginal mixing distributions to compute variance components. Examples are presented: longevity as a function of sex, race and U.S. state of residence; and lifetime reproduction among a set of developed countries and as a result of resource availability and pesticide exposure.

KEYWORDS Heterogeneity • Stochasticity • Variance partitioning • Longevity • Lifetime reproductive output • Markov chains with rewards

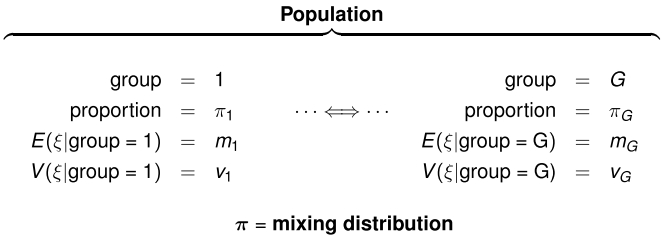
Introduction

A heterogeneous population is a mixture made up of groups of individuals that differ in the demographic rates to which they are subject (Figure 1). Each group is characterised by a mean and a variance of some demographic outcome. That heterogeneity among individuals contributes to the variance, also among individuals, at the population level. Longevity is one example of a demographic outcome. Variance in longevity is of interest to demographers as a form of inequality (e.g., Vaupel, 1988; Edwards and Tuljapurkar, 2005; Vaupel et al., 2011; van Raalte et al., 2018; Permanyer and Scholl, 2019; Permanyer et al., 2023). Longevity can be generalised to include healthy longevity (with health defined in many ways) or occupancy of medical, infection, or other kinds of states (Caswell and Zarulli, 2018; Caswell and van Daalen, 2021). Variance in longevity has implications for health systems, pensions, estate planning, etc.

Lifetime reproduction, the number of offspring produced by a female over her lifetime, is an outcome of interest to evolutionary and anthropological demographers as a measure of

✉ Hal Caswell, hcaswell@whoi.edu
1 Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands
2 Biology Department, Woods Hole Oceanographic Institution, Woods Hole MA, USA
3 Wageningen Marine Research, IJmuiden, The Netherlands

Figure 1 The structure of a heterogeneous population as a mixture of groups, each present in some proportion. Each group has a group-specific mean and variance for some demographic outcome, denoted here by ξ.



the opportunity for selection (Crow, 1958; van Daalen and Caswell, 2024). It is often reported for hunter-gatherer populations (Hill and Hurtado, 1996; Blurton-Jones, 2016; Brown et al., 2009) or historical populations (Moorad et al., 2011; Courtiol et al., 2012).

It is convenient to think of a finite number of groups, but the concepts and theory are essentially the same for continuous heterogeneity. A familiar example is the gamma-Gompertz distribution of mortality, which can be discretised to allow variances to be calculated (Caswell, 2014).

Variance and its sources

Because demographic rates contain probabilities, demographic outcomes are random variables. Because a heterogeneous population is a mixture, the distribution of the outcome is a mixture of the distributions within each group. This mixture is characterised by the moments of each group and the distribution of individuals among the groups. The latter distribution is called the *mixing distribution*.

The variance in outcome in a heterogeneous population reflects both stochasticity within each group and heterogeneity among groups (remember that the groups are homogeneous, and thus differences in outcome within a group are due to are strictly due to chance). In any outcome calculated from a life table, a Markov chain, or some equivalent machinery, individual stochasticity is the only source of within-group variance, because the calculation explicitly applies the same probabilities to every individual at every age. See Caswell (2023) and references therein for a discussion of the roles of heterogeneity and stochasticity and the issues of interpretation that arise from them.

Variances in outcomes are often called "inequalities," although that identification is more subtle than is usually appreciated (Caswell, 2023). Not all differences are inequalities.[1] In economic terms (e.g., Atkinson, 2015), variance due to heterogeneity corresponds to

1 Economists do not, as a rule, spend much time discussing what kinds of differences should be called inequalities. However, Therborn (2012, 2014) proposed three criteria that distinguish inequality from mere differences: (1) inequalities must admit at least ordinal classification so that items may be ranked as greater or lesser; (2) inequalities are not just categories, but violate some "moral norm of equality among human beings;" and (3) inequalities are potentially abolishable.

https://doi.org/10.1553/p-j2pg-fm66

inequality of opportunity. Individuals in, e.g., different income groups have different opportunities to live a long life. Variance due to stochasticity corresponds to inequality of outcome. Individuals in the same income group may differ in longevity because of the random outcomes of mortality risks, even though they all experience the same group-specific risks.

The relative magnitude of the contributions of stochasticity and heterogeneity depend on many factors, including the type and magnitude of the heterogeneity. Thus it is essential to partition the variance into within-group and between-group components.

Our intent is not to review the concepts of heterogeneity, stochasticity, and inequality, as a previous paper has explored these issues in some detail (Caswell, 2023). That paper surveyed the contributions of heterogeneity (among groups) and individual stochasticity (with groups) to: (1) variance in human longevity among latent frailty classes, income groups, education groups, and neighbourhood deprivation groups; (2) variance in adult longevity among insects differing in early life nutrition; (3) variance in longevity and lifetime reproduction in a wild seabird population; (4) variance in longevity and lifetime reproduction in laboratory populations of a rotifer, with groups defined by maternal age; (5) variance in healthy longevity with groups defined by European countries; and (6) variance in longevity and lifetime reproduction with groups defined by species and/or populations of plants and animals.

Instead, our goal is to provide the calculations necessary to extend variance partitioning from single factors to multiple interacting factors. Studies of the variance in longevity and in lifetime reproduction typically define heterogeneity in terms of a single factor. When studies examine multiple factors, they are usually treated one at a time. However, individuals are heterogeneous in multiple factors operating simultaneously, and the resulting variance is affected by all of those factors and their interactions. The inability to evaluate the contributions of interactions is a major limitation to the study of heterogeneity, and our goal here is to show how to analyse multi-factor studies, in which individuals are heterogeneous in two or more factors. We will present some two-factor and three-factor examples of the calculation. There is no limitation to the number of factors.

Variance partitioning

The variance in outcome among a heterogeneous set of individuals can be partitioned into between-group and within-group components using classical results from conditional probability. Consider a heterogeneous population that is a mixture of groups, the relative abundances of which are given by a mixing distribution $\boldsymbol{\pi}$. Let ξ denote some outcome. The variance in ξ is

$$V(\xi) = E_\pi\Big[V(\xi|\text{group})\Big] + V_\pi\Big[E(\xi|\text{group})\Big] \tag{1}$$

$$= V_{\text{within}}(\xi) + V_{\text{between}}(\xi). \tag{2}$$

The within-group variance V_{within} is the expectation of the variance within each group, weighted by the distribution $\boldsymbol{\pi}$. The between-group variance V_{between} is the variance among the group means, again weighted by the distribution $\boldsymbol{\pi}$. This variance decomposition is a

https://doi.org/10.1553/p-j2pg-fm66

standard result in conditional probability (e.g., Rényi, 1970; Frühwirth-Schnatter, 2006) and in analysis of variance (ANOVA) (e.g., Kempthorne, 1957). Indeed, the decomposition was introduced, along with the term "variance" itself, by Fisher (1918). This analysis has now been used repeatedly to examine demographic inequality among groups defined by a single factor (a "one-way" design).

Demographic outcomes: Longevity and lifetime reproductive output

We will show results in this paper for variance decomposition for longevity and lifetime reproductive output (LRO). The means and variances (and other moments) of longevity are readily calculated by expressing the demographic rates in terms of absorbing Markov chains (pioneered by Feichtinger, 1971; see Caswell, 2001, 2006, 2009). It is now known that, as a general rule, even extreme differences among groups, including those created by important socioeconomic variables, account for only a small fraction of the variance in longevity (see the overview in Caswell, 2023).

The means and variances (and other moments) of LRO can be calculated using Markov chains with rewards (Caswell, 2011; van Daalen and Caswell, 2017). Stochasticity in LRO reflects both survival (which determines how long a woman has to reproduce) and fertility (which determines whether or not she reproduces at each age). These two components can be separated, and in developed, low-mortality countries the variance in LRO is increasingly accounted for by stochasticity in fertility (van Daalen and Caswell, 2015). However, the contributions of heterogeneity and stochasticity to the variance in LRO are not yet well understood. In one case, in which the groups are defined by maternal age in a laboratory population, the variance components depend strongly on the environmental conditions (van Daalen et al., 2022). We will provide some examples below, in Section Lifetime reproductive output.

Variance components

Notation

The following notation is used throughout this paper. Matrices are denoted by upper-case bold characters (e.g., $\mathbf{U}$) and vectors are denoted by lower-case bold characters (e.g., $\mathbf{a}$). Vectors are column vectors by default; $\mathbf{x}^{\mathrm{T}}$ is the transpose of $\mathbf{x}$. The vector $\mathbf{1}$ is a vector of ones, and the matrix $\mathbf{I}$ is the identity matrix. When necessary, subscripts are used to denote the size of a vector or matrix; e.g., $\mathbf{I}_{\omega}$ is an identity matrix of size $\omega \times \omega$. Matrices and vectors with a tilde (e.g., $\tilde{\mathbf{U}}$ or $\tilde{\mathbf{a}}$) are block-structured; in this paper, blocks correspond to different factors. The notation $\|\mathbf{x}\|$ denotes the 1-norm of $\mathbf{x}$ (i.e., the sum of the absolute values of the entries). The symbol $\otimes$ denotes the Kronecker product. The vec operator stacks the columns of a $m \times n$ matrix into a $mn \times 1$ column vector. When applied to an array with more than two dimensions, it stacks columns from all dimensions. We will make use of a reshape operator that is the inverse of the vec operator, changing the vector back into the

https://doi.org/10.1553/p-j2pg-fm66

array; see equation (17). On occasion, MATLAB notation will be used to refer to the entries, rows, and columns of matrices. For example, $\mathbf{F}(i, j)$ is the (i, j) entry of $\mathbf{F}$, while $\mathbf{F}(i, :)$ and $\mathbf{F}(:, j)$ refer to the ith row and jth column of the matrix.

Element-by-element operations apply to arrays of any dimension. The symbol $\circ$ denotes the Hadamard, or element-by-element product (implemented by .* in MATLAB and by * in R). The symbol $\oslash$ is used to denote the Hadamard, or element-by-element quotient. Thus, for two matrices $\mathbf{A}$ and $\mathbf{B}$,

$$\mathbf{A} \circ \mathbf{B} = (a_{ij} b_{ij}) \qquad \mathbf{A} \oslash \mathbf{B} = \left(\frac{a_{ij}}{b_{ij}}\right) \tag{3}$$

with the obvious restrictions that the objects must be of the same size and, that for the quotient, none of the entries of $\mathbf{B}$ can be zero.

Weighted means and variances

Let $\mathbf{x}$ be a vector of numbers and $\boldsymbol{\pi}$ a probability vector of the same length. The mean and variance of the entries of $\mathbf{x}$, over the mixing distribution $\boldsymbol{\pi}$ are, in matrix notation,

$$E_{\boldsymbol{\pi}}(\mathbf{x}) = \boldsymbol{\pi}^{\mathrm{T}} \mathbf{x} \tag{4}$$

$$V_{\boldsymbol{\pi}}(\mathbf{x}) = \boldsymbol{\pi}^{\mathrm{T}} (\mathbf{x} \circ \mathbf{x}) - (\boldsymbol{\pi}^{\mathrm{T}} \mathbf{x})^2 \tag{5}$$

The second of these terms, the variance of X over a mixing distribution, will appear so commonly that we define it as a function $\mathbb{V}(\mathbf{x}, \boldsymbol{\pi})$ that returns the variance, over the distribution $\boldsymbol{\pi}$, of the vector $\mathbf{x}$:

$$\mathbb{V}(\mathbf{x}, \boldsymbol{\pi}) \equiv \boldsymbol{\pi}^{\mathrm{T}} (\mathbf{x} \circ \mathbf{x}) - (\boldsymbol{\pi}^{\mathrm{T}} \mathbf{x})^2. \tag{6}$$

One-factor designs

Let us review the simplest case: a one-factor design in which a population (or some other set of individuals) is divided into groups based on a single factor that we call A (e.g., income), with levels $a = a_1, a_2, \ldots, a_{n_A}$. The population is a mixture with a mixing distribution π. Let ξ be the demographic outcome of interest. The vectors containing the means and variances of ξ are

$$\mathbf{m} = \begin{pmatrix} E(\xi|a_1) \\ \vdots \\ E(\xi|a_{n_A}) \end{pmatrix} = \begin{pmatrix} m_1 \\ \vdots \\ m_{n_A} \end{pmatrix} \tag{7}$$

$$\mathbf{v} = \begin{pmatrix} V(\xi|a_1) \\ \vdots \\ V(\xi|a_{n_A}) \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_{n_A} \end{pmatrix} \tag{8}$$

https://doi.org/10.1553/p-j2pg-fm66

In terms of these vectors, the variance decomposition in equation (2) becomes

$$V_{\text{within}} = \boldsymbol{\pi}^{\mathrm{T}}\mathbf{v} \tag{9}$$

$$V_{\text{between}} = \boldsymbol{\pi}^{\mathrm{T}}(\mathbf{m} \circ \mathbf{m}) - (\boldsymbol{\pi}^{\mathrm{T}}\mathbf{m})^2 \tag{10}$$

$$= \mathbb{V}(\mathbf{m}, \boldsymbol{\pi}). \tag{11}$$

The relative contribution of the within- and between-group components to the variance $V(\xi)$ is measured by

$$\mathcal{K} = \frac{V_{\text{between}}}{V_{\text{within}} + V_{\text{between}}} \tag{12}$$

which measures the fraction of the total variance due to heterogeneity among groups. The ratio is referred to as the intraclass correlation coefficient in quantitative genetics (Falconer, 1960). Its square root is called the correlation ratio in probability theory (Rényi, 1970).

Factorial designs: Two factors

The multi-factor analysis is based on multi-way tables of the means, variances, and mixing distribution. The calculation of variance components uses marginal distributions calculated from those tables. Consider two factors, labelled A and B (e.g., sex and race) with n_A and n_B levels, respectively. For example, for sex n_A might be two, and for race n_B could be two, or five, or some other number depending on the information collected in vital statistics registries. The means, variances, and mixing distributions are defined in the two-dimensional arrays

$$\mathbf{M} = \begin{pmatrix} m_{1,1} & \cdots & m_{1,n_B} \\ \vdots & & \vdots \\ m_{n_A,1} & \cdots & m_{n_A,n_B} \end{pmatrix} \tag{13}$$

$$\mathbf{V} = \begin{pmatrix} v_{1,1} & \cdots & v_{1,n_B} \\ \vdots & & \vdots \\ v_{n_A,1} & \cdots & v_{n_A,n_B} \end{pmatrix} \tag{14}$$

$$\mathbf{\Pi} = \begin{pmatrix} \pi_{1,1} & \cdots & \pi_{1,n_B} \\ \vdots & & \vdots \\ \pi_{n_A,1} & \cdots & \pi_{n_A,n_B} \end{pmatrix}. \tag{15}$$

Each of these arrays is of dimension $n_A \times n_B$. The mixing distribution array satisfies

$$\sum_{i,j} \mathbf{\Pi}(i,j) = 1. \tag{16}$$

The vec operator applied to any of these arrays produces a vector, of dimension $n_A n_B \times 1$, by stacking the columns of the array. The inverse of the vec operator is the reshape

https://doi.org/10.1553/p-j2pg-fm66

operator, which takes as its arguments a vector and a pair of dimensions, and produces an array of those dimensions; thus

$$\text{reshape}(\text{vec}\,\mathbf{M}, n_A, n_B) = \mathbf{M}. \tag{17}$$

Within-group and between-group variances

Treating each of the $n_A n_B$ combination of factors A and B as a group, the within- and between-group variance components are calculated from the entries of $\mathbf{M}$, $\mathbf{V}$, and $\mathbf{\Pi}$. The within-group variance is the mean, over the mixing distribution, of the variances in each group,

$$V_{\text{within}} = (\text{vec}\,\mathbf{\Pi})^{\mathrm{T}}\text{vec}\,\mathbf{V}. \tag{18}$$

The between-group variance is given by applying equation (6) to the means of all the factor combinations:

$$V_{\text{between}} = \mathbb{V}(\text{vec}\,\mathbf{M}, \text{vec}\,\mathbf{\Pi}). \tag{19}$$

This between-group variance measures the overall contribution of heterogeneity, over all the groups, of all the factors, with individuals distributed according to $\mathbf{\Pi}$. It makes no distinction between the contributions of A, B, and the AB interaction, but if the goal is a measure of how heterogeneity in factors contributes to inequality of outcomes, V_{between} is the answer.

Partitioning the between-group variance into factor effects

In the two-factor case, V_{between} is due to contributions from each factor (V_{A}, V_{B}) and their interaction (V_{AB}). These components are calculated from the marginal means and marginal mixing distributions corresponding to each factor and the interaction. Use subscripts to identify the marginal means (e.g., $\mathbf{m}_{\text{A}}$ as the vector of means for each level of A, marginalising over levels of B), and let a, b denote the levels of A and B, respectively. The array of marginal means for factor A is obtained by calculating the average, weighted by the mixing distribution, over the levels of factor B. Recalling the definitions of the Hadamard product and the Hadamard quotient in Section Notation, we have

$$\mathbf{m}_{\text{A}} = \left[\sum_b (\mathbf{M} \circ \mathbf{\Pi})\right] \oslash \left[\sum_b \mathbf{\Pi}\right] \qquad n_A \times 1. \tag{20}$$

The same pattern holds for the marginal mean for B:

$$\mathbf{m}_{\text{B}} = \left[\sum_a (\mathbf{M} \circ \mathbf{\Pi})\right] \oslash \left[\sum_a \mathbf{\Pi}\right] \qquad n_B \times 1 \tag{21}$$

and for the marginal mean for AB

$$\mathbf{m}_{\text{AB}} = [\mathbf{M} \circ \mathbf{\Pi}] \oslash \mathbf{\Pi} \qquad n_A \times n_B. \tag{22}$$

https://doi.org/10.1553/p-j2pg-fm66

The marginal mean for factor A is obtained by summing over levels of B; the marginal mean for factor B is obtained by summing over levels of A. The marginal mean for the combination AB is not really marginal, because there are no other factors over which to sum. The marginal means for factors A and B are one-dimensional vectors. The marginal mean for the AB interaction is a two-dimensional array (it is, in fact, just the array $\mathbf{M}$).

The marginal mixing distributions are obtained from the array $\mathbf{\Pi}$,

$$\boldsymbol{\pi}_{\mathrm{A}} = \sum_{b} \mathbf{\Pi} \qquad n_A \times 1 \tag{23}$$

$$\boldsymbol{\pi}_{\mathrm{B}} = \sum_{a} \mathbf{\Pi} \qquad n_B \times 1 \tag{24}$$

$$\boldsymbol{\pi}_{\mathrm{AB}} = \mathbf{\Pi} \qquad n_A \times n_B. \tag{25}$$

MATLAB makes these calculations easy to implement; the command corresponding to equation (20), for example, is

$$\texttt{mA} = \texttt{sum(M.*Pi,[2])./sum(Pi,[2])}. \tag{26}$$

Finally, the variance components due to factors A and B and the interaction AB are calculated by applying the function $\mathbb{V}(\cdot,\cdot)$ defined in equation (6) to the marginal means and marginal mixing distributions:

$$V_{\mathrm{A}} = \mathbb{V}(\mathbf{m}_{\mathrm{A}}, \boldsymbol{\pi}_{\mathrm{A}}) \tag{27}$$

$$V_{\mathrm{B}} = \mathbb{V}(\mathbf{m}_{\mathrm{B}}, \boldsymbol{\pi}_{\mathrm{B}}) \tag{28}$$

$$V_{\mathrm{AB}} = \mathbb{V}(\mathrm{vec}\,\mathbf{m}_{AB}, \mathrm{vec}\,\boldsymbol{\pi}_{AB}) - V_{\mathrm{A}} - V_{\mathrm{B}}. \tag{29}$$

The marginal mean array $\mathbf{m}_{\mathrm{AB}}$ contains the effects of A and B as well as the interaction; thus the interaction variance V_{AB} is obtained by subtracting V_{A} and V_{B} from the variance among all factor combinations.

Requirements for the factorial mixing distribution

The calculation of V_{between} treats vec $\mathbf{M}$ and vec $\mathbf{V}$ as vectors containing values for each of the groups. There is no restriction on the mixing distribution $\mathbf{\Pi}$ except, of course, that its entries sum to one.

Here, however, we want to partition $\mathbf{V}_{\mathrm{between}}$ into its components, and to do this the mixing distribution $\mathbf{\Pi}$ requires some careful attention. Statistics texts are unanimous in stressing the importance of equal sample sizes in all treatment combinations in the analysis of variance (ANOVA) in a factorial experiment. The sample sizes play the role of the mixing distribution in our probability calculations. It has long been known (Yates, 1934)

https://doi.org/10.1553/p-j2pg-fm66

that unequal sample sizes make it impossible to calculate variance components. A recent text explains

> "When the sample sizes for each cell are unequal, the two-way analysis of variance for factor effects becomes complex. The component sums of squares in the analysis of variance are no longer orthogonal; that is, they do not sum to the total sum of squares. The least squares method for obtaining the best estimates of the parameters is rather complicated in the fixed effects model and the best analysis has not been and probably will not be found for the random effects models" (Sahai and Ageel, 2012).

There exist two classes of mixing distributions that permit the calculation of variance components. One is the flat, or balanced distribution, which corresponds to equal sample sizes and assigns equal weight to all factor combinations. The other is the class of rank-one, or proportional mixing distributions (Yates, 1934; Kirk, 1982, and many others). Such a mixing distribution, when written as an array as in equation (15), has proportional rows and proportional columns. The flat distribution is a special case of the rank-one distribution.

A rank-one mixing distribution array can be assembled from its marginals. Let $\boldsymbol{\pi}_{\mathrm{A}}$ and $\boldsymbol{\pi}_{\mathrm{B}}$ be the marginal distributions among the levels of factors A, B. The mixing distribution array is

$$\boldsymbol{\Pi} = \mathrm{reshape}(\boldsymbol{\pi}_{\mathrm{B}} \otimes \boldsymbol{\pi}_{\mathrm{A}}, n_A, n_B). \tag{30}$$

Every column is proportional to $\boldsymbol{\pi}_{\mathrm{A}}$ and every row is proportional to $\boldsymbol{\pi}_{\mathrm{B}}$. Note the order of the subscripts.

The mixing distribution is a useful tool

Our goal is to understand the sources of variance in some outcome within some population. But what population? The mixing distribution describes the structure of the population that we are interested in, and over which the variance is to be calculated, in terms of the proportions of the population in each of the heterogeneity groups. The proper question is not what the *correct* mixing distribution is, but rather what mixing distribution answers the question we are interested in.

A flat mixing distribution gives every group equal representation when studying the effects of heterogeneity. It is a particularly powerful tool, not because many populations are comprised of equal numbers in every group[2] but for the same reason that equal sample sizes are desirable in designed experiments. If you want to quantify the effects of some set of factors, it is wise to design your experiment with equal sample sizes in each treatment combination, because doing so maximises the ability to extract information obtained from the ranges of both variables. Thus, a variance component calculation can be thought of as a kind of numerical experiment to evaluate the effects of factors and their interactions.

2 An exception is studies in which groups are defined as quantiles of the distribution of some variable (e.g., income). By definition, every quantile contains the same fraction of the population.

https://doi.org/10.1553/p-j2pg-fm66

Alternatively, the mixing distribution may reflect the relative population sizes of groups (subject to the proportional restriction). If some groups are much larger than others, the population variance will reflect the variance within those groups to a much greater extent than the variance within the much smaller groups, as an example presented below shows.

Factorial designs: Three factors

Variance partitioning for three factors follows the same pattern as that for two factors. We briefly present the formulas here; the extension to an arbitrary number of factors should be clear.

Consider three factors that are labelled A, B, and C (e.g., sex, race, and state of residence), with levels n_A, n_B, and n_C. The arrays $\mathbf{M}$ of means, $\mathbf{V}$ of variances, and $\mathbf{\Pi}$ of mixing probabilities are now three-dimensional. Again, the mixing distribution array satisfies

$$\sum_{a,b,c} \mathbf{\Pi}(a,b,c) = 1. \tag{31}$$

The vec operator produces a vector of dimension $n_A n_B n_C \times 1$ by stacking the columns of the array in the order A, then B, and then C.

The mixing distribution array must be assembled from its marginals. Let $\boldsymbol{\pi}_{\mathrm{A}}, \boldsymbol{\pi}_{\mathrm{B}}, \boldsymbol{\pi}_{\mathrm{C}}$ be the marginal mixing distributions among the levels of factors A, B, and C, respectively. Then the rank-one mixing distribution array, generalising that for two factors in (25) is

$$\mathbf{\Pi} = \mathrm{reshape}(\boldsymbol{\pi}_{\mathrm{C}} \otimes \boldsymbol{\pi}_{\mathrm{B}} \otimes \boldsymbol{\pi}_{\mathrm{A}}, n_A, n_B, n_C). \tag{32}$$

The extension to more than three factors follows the same logic.

Within- and between-group variances

As in Section Factorial designs: Two factors, the within-group and between-group variances are calculated by treating all $n_A n_B n_C$ factor combinations as groups. Then, just as in the two-factor case,

$$V_{\mathrm{within}} = (\mathrm{vec}\,\mathbf{\Pi})^{\mathrm{T}} \mathrm{vec}\,\mathbf{V} \tag{33}$$

$$V_{\mathrm{between}} = \mathbb{V}(\mathrm{vec}\,\mathbf{M}, \mathrm{vec}\,\mathbf{\Pi}). \tag{34}$$

As in the two-factor case, V_{between} gives the contribution to variance of all the heterogeneity among groups, but no information on the contributions of the factors and their interactions.

Components of the between-group variance

The between-group variance V_{between} is partitioned into components due to the main effects of each factor ($V_{\mathrm{A}}, V_{\mathrm{B}}, V_{\mathrm{C}}$), the two-way interactions ($V_{\mathrm{AB}}, V_{\mathrm{AC}}, V_{\mathrm{BC}}$) between pairs of

https://doi.org/10.1553/p-j2pg-fm66

factors, and the three-way interaction (V_{ABC}). These components are calculated from the marginal means and marginal mixing distributions corresponding to each factor and each interaction.

The array of marginal means for factor A is obtained by averaging over the dimensions other than those for factor A:

$$\mathbf{m}_{\mathrm{A}} = \left[\sum_b \sum_c (\mathbf{M} \circ \mathbf{\Pi})\right] \oslash \left[\sum_b \sum_c \mathbf{\Pi}\right] \qquad n_A \times 1. \tag{35}$$

The same pattern holds for the other marginal means:

$$\mathbf{m}_{\mathrm{B}} = \left[\sum_a \sum_c (\mathbf{M} \circ \mathbf{\Pi})\right] \oslash \left[\sum_a \sum_c \mathbf{\Pi}\right] \qquad n_B \times 1 \tag{36}$$

$$\mathbf{m}_{\mathrm{C}} = \left[\sum_a \sum_b (\mathbf{M} \circ \mathbf{\Pi})\right] \oslash \left[\sum_a \sum_b \mathbf{\Pi}\right] \qquad n_C \times 1 \tag{37}$$

$$\mathbf{m}_{\mathrm{AB}} = \left[\sum_c (\mathbf{M} \circ \mathbf{\Pi})\right] \oslash \left[\sum_c \mathbf{\Pi}\right] \qquad n_A \times n_B \tag{38}$$

$$\mathbf{m}_{\mathrm{AC}} = \left[\sum_b (\mathbf{M} \circ \mathbf{\Pi})\right] \oslash \left[\sum_b \mathbf{\Pi}\right] \qquad n_A \times n_C \tag{39}$$

$$\mathbf{m}_{\mathrm{BC}} = \left[\sum_a (\mathbf{M} \circ \mathbf{\Pi})\right] \oslash \left[\sum_a \mathbf{\Pi}\right] \qquad n_B \times n_C \tag{40}$$

$$\mathbf{m}_{\mathrm{ABC}} = [(\mathbf{M} \circ \mathbf{\Pi})] \oslash \mathbf{\Pi} \qquad n_A \times n_B \times n_C. \tag{41}$$

The marginal mixing distributions are obtained from the array $\mathbf{\Pi}$,

$$\boldsymbol{\pi}_{\mathrm{A}} = \sum_b \sum_c \mathbf{\Pi} \qquad n_A \times 1 \tag{42}$$

$$\boldsymbol{\pi}_{\mathrm{B}} = \sum_a \sum_c \mathbf{\Pi} \qquad n_B \times 1 \tag{43}$$

$$\boldsymbol{\pi}_{\mathrm{C}} = \sum_a \sum_b \mathbf{\Pi} \qquad n_C \times 1 \tag{44}$$

$$\boldsymbol{\pi}_{\mathrm{AB}} = \sum_c \mathbf{\Pi} \qquad n_A \times n_B \tag{45}$$

$$\boldsymbol{\pi}_{\mathrm{AC}} = \sum_b \mathbf{\Pi} \qquad n_A \times n_C \tag{46}$$

$$\boldsymbol{\pi}_{\mathrm{BC}} = \sum_a \mathbf{\Pi} \qquad n_B \times n_C \tag{47}$$

$$\boldsymbol{\pi}_{\mathrm{ABC}} = \mathbf{\Pi} \qquad n_A \times n_B \times n_C. \tag{48}$$

https://doi.org/10.1553/p-j2pg-fm66

The calculations are readily expressed as MATLAB commands; for example, the command corresponding to equation (35) is

$$\texttt{mA} = \texttt{sum}(\texttt{M}.*\texttt{Pi}, [2, 3])./\texttt{sum}(\texttt{Pi}, [2, 3]). \tag{49}$$

The second argument ([2, 3]) in the `sum` command indicates the dimensions over which summation takes place.

The variance components due to each of the factors and interactions are calculated by applying the function $\mathbb{V}(\cdot, \cdot)$ in equation (6) to the vectors obtained by applying the vec operator to the marginal mean arrays

$$V_{\mathrm{A}} = \mathbb{V}(\mathbf{m}_{\mathrm{A}}, \boldsymbol{\pi}_{\mathrm{A}}) \tag{50}$$

$$V_{\mathrm{B}} = \mathbb{V}(\mathbf{m}_{\mathrm{B}}, \boldsymbol{\pi}_{\mathrm{B}}) \tag{51}$$

$$V_{\mathrm{C}} = \mathbb{V}(\mathbf{m}_{\mathrm{A}}, \boldsymbol{\pi}_{\mathrm{A}}) \tag{52}$$

$$V_{\mathrm{AB}} = \mathbb{V}(\mathrm{vec}\, \mathbf{m}_{\mathrm{AB}}, \mathrm{vec}\, \boldsymbol{\pi}_{\mathrm{AB}}) - V_{\mathrm{A}} - V_{\mathrm{B}} \tag{53}$$

$$V_{\mathrm{AC}} = \mathbb{V}(\mathrm{vec}\, \mathbf{m}_{\mathrm{AC}}, \mathrm{vec}\, \boldsymbol{\pi}_{\mathrm{AC}}) - V_{\mathrm{A}} - V_{\mathrm{C}} \tag{54}$$

$$V_{\mathrm{BC}} = \mathbb{V}(\mathrm{vec}\, \mathbf{m}_{\mathrm{BC}}, \mathrm{vec}\, \boldsymbol{\pi}_{\mathrm{BC}}) - V_{\mathrm{B}} - V_{\mathrm{C}} \tag{55}$$

$$V_{\mathrm{ABC}} = \mathbb{V}(\mathrm{vec}\, \mathbf{m}_{\mathrm{ABC}}, \mathrm{vec}\, \boldsymbol{\pi}_{\mathrm{ABC}}) - V_{\mathrm{A}} - V_{\mathrm{B}} - V_{\mathrm{C}} - V_{\mathrm{AB}} - V_{\mathrm{AC}} - V_{\mathrm{BC}}. \tag{56}$$

The arrays for the two-way interactions also include the one-factor effects, so the one-factor variances are subtracted to obtain the two-factor variances. The variance due to the three-factor interaction has the one-factor and two-factor variances subtracted.

The interpretation of interactions

The interpretation of interactions in factorial experiments has always been a challenge (e.g. Steel and Torrie, 1960; Sahai and Ageel, 2012). A large component of the variance due to an AB interaction makes it difficult to say what the effects of A and B are because the effect of A depends on the level of B, and vice versa. The situation becomes even more difficult, of course, for three-way or higher interactions. If the contributions of interactions to variance are small, they can be ignored. Note that we have no operational definition of "small" such as is provided in ANOVA by tests of the statistical significance of the interactions. However, the difficulty of interpreting interactions does not change the fact that they are substantively interesting. Knowing that two factors interact is an important substantive finding and invites further study to understand how that interaction works.

Note that if the factors have additive effects on ξ, the interaction variances are all zero (Caswell, unpublished results).

https://doi.org/10.1553/p-j2pg-fm66

Examples

Here, we present several examples of multi-factor variance decompositions: two studies of longevity and two studies of lifetime reproductive output. At this point, the results are intended to serve only as examples of choices of mixing distributions and of the kinds of results obtained. As is always the case when a new analytical method is deployed, the interpretation of the results is still developing.

Longevity and lifespan

Variance in longevity, often referred to as inequality or disparity in lifespan, has been analysed across a variety of social, economic, and biological variables. Partitioning of the variance into components has revealed that, even when group differences are very large, they contribute only a small fraction of the total variance (see an overview in Caswell, 2023). Because longevity is the outcome of a lifetime of repeated probabilistic survival events, it is subject to a large degree of individual stochasticity.

However, these analyses have been limited to single factors. Here, we report two examples of multi-factor studies: one examining variance due to sex and race, and the other examining sex, race, and state of residence. In these examples, we calculated the mean and variance of longevity using Markov chain methods (Feichtinger, 1971; Caswell, 2001, 2006), but they could equally have been calculated from a set of life tables.

Variance in longevity: Sex and race

Differences in longevity between males and females are well known; in almost every case, women live longer on average than men. Differences among racial and ethnic groups are also well known. Here we explore the variance in longevity among males and females across racial and ethnic categories in the United States. The 2020 life tables for the United States (Arias and Xu, 2022) classify individuals as male or female and into five racial and ethnic categories: Hispanic (H), Non-Hispanic American Indian and Alaska Native (NHAIAN), Non-Hispanic Asian (NHA), Non-Hispanic Black (NHB), and Non-Hispanic White (NHW). The arrays of mean longevity (life expectancy) and variance in longevity are

$$\mathbf{M} = \begin{pmatrix} 74.6 & 63.8 & 81.1 & 67.8 & 74.8 \\ 81.3 & 70.7 & 85.9 & 75.3 & 80.1 \end{pmatrix}^{\mathrm{T}} \tag{57}$$

$$\mathbf{V} = \begin{pmatrix} 289 & 408 & 217 & 371 & 294 \\ 220 & 391 & 163 & 315 & 235 \end{pmatrix}^{\mathrm{T}} \tag{58}$$

with males in the first row and females in the second row. These differences in life expectancy among ethnic groups and between the sexes are typical for those variables.

Two mixing distributions suggest themselves, asking different questions. A flat mixing distribution treats each sex–race combination equally in calculating its contribution to the variance. It provides information on how the differences in the conditions experienced by

https://doi.org/10.1553/p-j2pg-fm66

these groups contribute to the variance among individuals in length of life, and thus tells us something about the sex–race groups per se. However, the groups have quite different levels of representation in the U.S. population. A mixing distribution proportional to the population sizes of each racial group, with sexes treated as equal, provides a decomposition of the variance in longevity among a hypothetical set of individuals who are, at birth, distributed over races proportional to the racial population sizes. (A mixing distribution with racial groups proportional to the number of births for each race would be similar, but not identical.) It tells us something about how the different conditions experienced by the groups *and* the population structure of those groups contribute to the variance in length of life.

Using population figures from 2020 (U.S. Census Bureau, Population Division, 2022), and setting the abundance of the sexes as equal, the marginal distributions for the flat mixing distribution are

$$\boldsymbol{\pi}_{\text{race}} = \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \end{pmatrix} \qquad \boldsymbol{\pi}_{\text{sex}} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}. \tag{59}$$

The marginal distributions for the population-weighted mixing distribution are

$$\boldsymbol{\pi}_{\text{race}} = \begin{pmatrix} 0.1859 \\ 0.0076 \\ 0.0607 \\ 0.1300 \\ 0.6158 \end{pmatrix} \qquad \boldsymbol{\pi}_{\text{sex}} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}. \tag{60}$$

In the rank-one mixing distribution, the Hispanic and NH White groups account for 80% of the mixture, giving much less weight to the other racial categories. The rank-one mixing distribution is given by

$$\mathbf{\Pi} = \text{reshape}(\boldsymbol{\pi}_{\text{sex}} \otimes \boldsymbol{\pi}_{\text{race}}, 5, 2) \tag{61}$$

$$= \begin{pmatrix} 0.0930 & 0.0930 \\ 0.0038 & 0.0038 \\ 0.0304 & 0.0304 \\ 0.0650 & 0.0650 \\ 0.3079 & 0.3079 \end{pmatrix}. \tag{62}$$

The variance decompositions obtained from the flat and the population-weighted mixing distributions are given in Table 1. Under the flat mixing distribution, the between-group variance accounts for 12.5% of the total variance. Racial differences contribute about three times the variance of sex differences, and the interaction contributes only a small fraction. When we shift our attention to the population-weighted mixing distribution, we see that the contribution of racial-ethnic heterogeneity shrinks from 31.5 to 7.4, as a result of the

https://doi.org/10.1553/p-j2pg-fm66

Table 1 The components of variance in longevity due to race, sex, and their interaction, for the population of the United States, 2020, The population-weighted mixing distribution is constructed by setting the marginal distribution of races proportional to their representation in the U.S. population in 2020.

Flat mixing	
Component	**Variance**
A = Race	31.5
B = Sex	9.7
AB = Race × sex	0.3
(between-group)	41.5
Stochasticity	290.3
Total	331.8
$\mathcal{K}$	0.125
Population-weighted mixing	
Component	**Variance**
A = Race	7.4
B = Sex	8.5
AB = Race × sex	0.2
(between-group)	16.1
Stochasticity	269.3
Total	285.4
$\mathcal{K}$	0.056

dominance of the population by the Hispanic and Non-Hispanic White groups. When the two factors are combined, heterogeneity accounts for only 5.6% of the variance.

Variance in longevity: Sex, race, and state of residence

In a country as large and diverse as the United States, there can be appreciable regional differences in mortality. The U.S. Census Bureau provides male and female life tables for each of the 50 states and the District of Columbia (Arias et al., 2022). The differences in life expectancy among states are large, and are partly a reflection of the political affiliations and policies of the states, with occupants of more liberal states experiencing longer life expectancies (Montez and Farina, 2021; Montez et al., 2020). An earlier study, conducted at the level of U.S. counties, found such large differences that the authors suggested the existence of "eight Americas" (Murray et al., 2006).

As an example of a three-factor analysis, we consider United States life tables by sex, race, and state of residence as given in (Wei et al., 2012). Unfortunately, the results are not directly comparable with the race×sex results in Table 1 because only two racial groups, White and Black, were reported in these data. Moreover, in Wei et al. (2012), only 41 states

https://doi.org/10.1553/p-j2pg-fm66

Table 2 Components of variance in longevity due to sex, race, and U.S. state of residence.

Flat mixing	
Component	**Variance**
A = sex	9.23
B = race	7.50
C = state	1.37
AB = sex × race	0.061
AC = sex × state	0.209
BC = race × state	0.567
ABC = sex × race × state	0.170
(between-groups)	19.11
Stochasticity	303.04
Total	322.14
$\mathcal{K}$	0.059

were included because the sample sizes for the Black population in the other states were considered too small to provide reliable estimates of mortality. However, we present an analysis as an example of the potential for a three-factor interaction.

Table 2 shows the variance decomposition with a flat mixing distribution. The largest main effect variance component is due to sex, followed by race and then state of residence. The two-way interactions are small, as is the three-way interaction. Taken together, the interactions account for just over 5% of the between-group variance. Within-group stochasticity again accounts for most of the variance, and the variance ratio $\mathcal{K} = 0.059$. That is, the combined effects of the three factors and their interactions account for only 5.9% of the variance in longevity.

Lifetime reproductive output

Lifetime reproductive output (LRO), sometimes called lifetime reproductive success (LRS), is the number of offspring produced by a female over her lifetime. When offspring are measured as female children and are subject to mortality, the mean of LRO is the net reproductive rate R_0 (or NRR). When offspring are measured as children of either sex and mortality is excluded, the mean of LRO is the total fertility rate (TFR). See Keilman et al. (2014) for a discussion of measures of lifetime reproduction including different combinations of sexes, and Caswell (2009, p. 1771) for the criteria that must be met for the net reproductive rate.

Lifetime reproductive "output" is a purposely flexible term. Output could be defined as children of either sex, or of both sexes, or of any other categories. Thus, it is important to specify how output is defined. Other categories can be analysed the same way. If, for example, we were interested in left-handed blue-eyed children and had age-specific

https://doi.org/10.1553/p-j2pg-fm66

numbers of such births per female, the lifetime output of such children could be calculated. Lifetime production of left-handed blue-eyed children is purposely frivolous, but the same logic applies to, for example, children born with trisomy-21 (the chromosome duplication that causes Down syndrome). The statistics of the lifetime production of children with trisomy-21 could be very interesting indeed.

Because survival and reproduction include stochastic processes, LRO is a random variable, with a variance, skewness and higher moments. This variance could be considered a "reproductive inequality" just as variance in longevity is considered "lifespan inequality." However, this type of inequality of outcome seems to have drawn little interest (but see early work on the concentration of reproduction by Vaupel and Goodwin, 1987 and Shkolnikov et al., 2007). Variation in outcomes such as childlessness and completed parity has attracted attention, and integrating these outcomes into a variance partitioning framework is an interesting research challenge.

The variance among individuals in lifetime reproduction has been approached by population biologists in several ways (e.g., Caswell, 2011; Steiner and Tuljapurkar, 2012; van Daalen and Caswell, 2017; Snyder and Ellner, 2018; Tuljapurkar et al., 2020). Here we use Markov chains with rewards (MCWR), which provide the mean and variance (indeed, all the moments) of lifetime reproduction. See Caswell (2011); van Daalen and Caswell (2017, 2020) for details of the calculation. Briefly, individual development, survival and transitions among the possible life history states are described by an absorbing Markov chain. Rewards, in the form of offspring produced along the life history trajectory, are accumulated until death. The model provides the mean and variance of lifetime fertility as a function of the demographic rates, depending on the fate of the individual and on the outcomes of its chances of reproduction at each point in its lifetime.

While some comparisons of the variance, skewness and other statistics of LRO for humans and other species, have been reported (Caswell, 2011; van Daalen and Caswell, 2015; van Daalen and Caswell, 2017; van Daalen and Caswell, 2024; Varas Enríquez et al., 2022), only a few studies have partitioned variance in LRO into contributions from heterogeneity and stochasticity. One example is the study by van Daalen et al. (2022) of the contribution of heterogeneity in maternal age to variance in LRO in a rotifer.

As with longevity, our goal here is to extend one-factor analyses to multi-factor studies in which the variance in lifetime reproduction can be partitioned into contributions from heterogeneity and stochasticity. We show two examples. One is a comparison of lifetime fertility among human females in 31 developed countries over a 40-year time interval, treating country and time as factors. The other is a laboratory study of the effects of food limitation and pesticide exposure on a rotifer.

Variance in reproduction over time and among countries

van Daalen and Caswell (2015) used Markov chains with rewards to explore the statistics of lifetime reproduction (mean, variance, skewness, standardised variance) in a set of developed countries during the second demographic transition, using data from the Human Mortality Database and the Human Fertility Database.

https://doi.org/10.1553/p-j2pg-fm66

Table 3 Components of variance in lifetime reproduction as affected by the factors year (1960 compared with 2000) and country. Based on data from van Daalen and Caswell (2015).

Flat mixing	
Component	**Variance**
A = year	0.330
B = country	0.135
AB = year × country	0.046
(between-groups)	0.513
Stochasticity	1.976
Total	2.487
$\mathcal{K}$	0.206
Population-weighted	
Component	**Variance**
A = year	0.348
B = country	0.229
AB = year × country	0.035
(between-groups)	0.612
Stochasticity	2.170
Total	2.782
$\mathcal{K}$	0.220

Strictly as an example of the calculations, we present an analysis of a two-factor variance decomposition, choosing as factors years (the two years 1960 and 2000) and countries (the 31 countries for which data were available for both years).[3] We computed the within-group and between-group components of variance, comparing two mixing distributions. One is flat, treating each country-year combination as equally important. The other creates marginal mixing distributions proportional to population size measured at age zero. This accounts for the distribution of population sizes of individuals beginning their "lifetime reproduction" that would result from randomly selecting newborn females in proportion to their abundance.

Table 3 shows the components of variance due to years, countries, and the interaction of countries and years. It also shows the within-factor variance due to stochasticity.

With a flat mixing distribution, the variance due to years is more than twice the variance due to countries. The interaction of years and countries is very small. The within-factor variance due to stochasticity is large, and the variance ratio $\mathcal{K} = 0.21$. That is, only 21%

3 The countries are Austria, Bulgaria, Canada, Switzerland, Czech Republic, East Germany, West Germany, Estonia, Finland, France, Scotland, England and Wales, Hungary, Japan, Lithuania, Netherlands, Portugal, Russia, Slovakia, Sweden, Ukraine, United States, Australia, Belgium, Belarus, Denmark, Spain, Ireland, Italy, New Zealand, and Poland.

https://doi.org/10.1553/p-j2pg-fm66

of the variance in lifetime reproduction is due to the historical changes over this 40-year period and the political and social differences among countries over these two years. Almost 80% of the variance is due to stochasticity from the random outcome of survival and fertility. Within that 21% of the variance due to heterogeneity, 91% is due to main effects.

Using the population-weighted rank-one mixing distribution makes only small changes in the variance components. The component due to years is similar to that with a flat mixing distribution. The component due to countries is larger with the population-weighted mixing distribution than with the flat mixing distribution. This contrasts with the results for longevity in Table 1, in which weighting the races by population size reduced, rather than increased, the variance due to race. The interaction is again very small, and the variance ratio $\mathcal{K} = 0.22$ is very similar to that with the flat mixing distribution.

Variance in reproduction due to diet and pesticide exposure

There is a rich biodemographic literature that applies demographic methods to laboratory populations of animals in order to evaluate the population effects of exposure to toxic substances. We present one such example here to demonstrate how exposure to conditions that are known to affect human fertility can be analysed in a case where experimentation is real, not imaginary.

Rotifers are microscopic invertebrate animals commonly used as bioindicator species for water quality, and are increasingly used as model organisms for the study of ageing, maternal effects, and maternal investment (Bock et al., 2019). They have a relatively short lifespan ($\approx$ 2 weeks), but high maternal investment into a small number of offspring.

In a classic experiment by Rao and Sarma (1986), the rotifer *Brachionus patulus* was exposed to five different concentrations of the pesticide DDT under low and high food resource levels. Age-specific survival and fertility schedules were reported for all combinations of DDT treatment and food levels. The resulting arrays of means and variances of LRO are

$$\mathbf{M} = \begin{pmatrix} 4.59 & 4.87 & 0.79 & 0.04 & 0.01 \\ 12.64 & 11.38 & 9.50 & 3.79 & 1.49 \end{pmatrix} \tag{63}$$

$$\mathbf{V} = \begin{pmatrix} 12.95 & 12.45 & 1.69 & 0.08 & 0.01 \\ 34.93 & 39.89 & 28.10 & 16.11 & 5.48 \end{pmatrix}. \tag{64}$$

The two rows represent low and high food levels, respectively. The columns represent the five increasing levels of DDT exposure (0, 15, 30, 45, 60 μg/l of DDT).

In this experiment, both DDT exposure and low food levels reduce mean lifetime reproduction. The effect of the pesticide exposure is greater at low food levels than at high food levels. The variance in LRO decreases with increasing DDT exposure and at low food levels. The variance is much larger than the mean, implying that the distribution of LRO is overdispersed relative to the Poisson distribution. This is a frequent pattern in analyses of LRO; it often results from some portion of the population failing to reproduce at all (e.g., Tuljapurkar et al., 2020; van Daalen and Caswell, 2017).

https://doi.org/10.1553/p-j2pg-fm66

Table 4 Components of variance in lifetime reproduction as affected by food level and DDT exposure. Data from the experimental study of Rao and Sarma (1986).

Flat mixing	
Component	**Variance**
A = food level	8.12
B = DDT exposure	10.09
AB = food × DDT	1.84
(between-groups)	20.06
Stochasticity	15.17
Total	35.22
$\mathcal{K}$	0.57

The uniform mixing distribution is appropriate to a designed experiment like this. The resulting variance decomposition is shown in Table 4. The biggest component is that due to pesticide exposure, while the food level contribution is slightly smaller. The between-group variance, including the two main effects and their interaction, accounts for 57% of the variance in the experiment, which is higher than any of the percentages found to date for variance in longevity. The contribution of the food×DDT interaction is small, making up 9% of the total between-group variance.

Discussion

Despite the obvious importance of variation, demography has often focused on expected values. Life expectancy is the expected value of longevity, the net reproductive rate R_0 is the expected value of lifetime reproduction, and the total fertility rate TFR is the expected value of lifetime reproduction conditional on survival to the end of reproduction. The focus is expanding as questions of inequality become more important. Addressing these questions naturally leads to partitioning variation into components due to differences among individuals in the rates to which they are subject (heterogeneity) and components due to the stochastic outcomes of those rates among individuals subject to the same rates.

The current state of the art focuses on contributions of factors treated one at a time (sex, education, income, nutrition, etc.). Many studies of variance in longevity have found that the contribution of heterogeneity, even in factors known to have important effects on individuals, is dwarfed by the contribution of stochasticity. While variance in lifetime reproduction has been studied less, it appears that heterogeneity may make larger contributions to this demographic outcome.

Variance decomposition requires the means and variances of the outcome for all combinations of the factors. Although the means and variances can be obtained in many different ways, Markov chains with rewards are a particularly powerful method.

https://doi.org/10.1553/p-j2pg-fm66

They can be applied to lifetime reproduction, to longevity in single- or multi-state models, and to healthy longevity in both prevalence-based (e.g., Caswell and Zarulli, 2018 and Zarulli and Caswell, 2022 for disability, Owoeye et al., 2020 for malnutrition) and incidence-based (Caswell and van Daalen, 2021 for stages of cancer) models.

In this paper we have provided the results needed to apply these means and variances in studies of multiple factors operating simultaneously. The protocol for the variance decomposition follows a simple series of steps, independent of the number of factors considered.

A protocol for factorial variance decomposition

1. Define the factors that characterise the heterogeneity among individuals (e.g., sex, race, income, education, resource levels) and the levels of each factor.
2. Choose a demographic outcome of interest (e.g., longevity, lifetime reproductive output).
3. Compute the means and variances of this demographic outcome for all combinations of the factors, using Markov chains, Markov chains with rewards, or life tables as deemed appropriate.
4. Create the arrays $\mathbf{M}$ and $\mathbf{V}$ containing the means and variances, as in equations (13) and (14). The dimension of these arrays is the number of factors in the study (i.e., for n factors, $\mathbf{M}$ and $\mathbf{V}$ are n-dimensional).
5. Think about the question of interest and specify a set of individuals over which the variances are to be computed:
   (a) a set consisting of equal representation of all factor combinations (the flat mixing distribution); or
   (b) a set defined by a rank-one combination of marginal distributions.
6. Create the array $\mathbf{\Pi}$, containing the n-dimensional mixing distribution, as in (15).
7. Treating all the factor combinations as groups, calculate the overall within- and between-factor variance components V_{within} and V_{between} using equations (9) and (10).
8. Partition the between-group variance into components:
   (a) compute marginal mean arrays for each factor and each factor interaction, as in equations (20)–(22);
   (b) create the corresponding marginal mixing distributions, as in equations (23)–(25); and
   (c) compute the variance components for each factor and each interaction, as in equations (27)–(29).

Variance as a measure of inequality

There are many ways to measure the variation ("inequality" in a broad sense) of some quantity. Economists have developed many indices to address specifically economic issues related to income, transfers, and so on (e.g., Sen, 1997; Jenkins and van Kerm, 2009; Atkinson, 2015). In demographic contexts, these measures are highly correlated

https://doi.org/10.1553/p-j2pg-fm66

(Van Raalte and Caswell, 2013). Thus if the aim is to focus only on the value of the index, there is little basis on which to choose among them. But if the aim is to go beyond the values, there are properties of the variance that make it an attractive choice.

As is well known, the variance is decomposable into within- and between-group components (as are indices based on entropy). Additive decomposability is more than a mathematical nicety; it is fundamental to our attempts to understand how heterogeneity contributes to inequality of outcomes. It was for this purpose that Fisher introduced it as the basis for the analysis of experimental data (Fisher, 1936). Statistical ANOVA has developed into an enormous variety of study designs, referred to as "experimental designs" in the experimental sciences. In our context, these designs correspond to arrangements of factors within and across populations. The factorial study design we explore here only scratches the surface of the possibilities.

The most popular alternative to the variance is the Gini coefficient. The Gini coefficient is based on the mean of the absolute values of the deviations between pairs of variates (e.g., individuals) selected from the distribution. Because of this, Permanyer et al. (2023) call the Gini coefficient an individual measure. They contrast it with the variance, which they call a group measure because the variance is usually written as the mean of the squared deviations from the mean, rather than as a difference among individual values. However, the variance and the Gini coefficient are both individual measures in this sense. The mean difference between two individuals from the distribution $f(\cdot)$ is

$$\Delta_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x-y| f(x) f(y) dxdy. \tag{65}$$

The Gini coefficient is a standardised, dimensionless version of the mean difference

$$G = \frac{\Delta_1}{2\mu_1} \tag{66}$$

where μ_1 is the mean.

The variance can also be written as a difference among individuals, "without reference to deviations from a central value, the mean" (Kendall and Stuart, 1969, p. 47) as

$$V = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-y)^2 f(x) f(y) dxdy. \tag{67}$$

The similarity between the mean difference and the variance is clear; the former uses absolute values of the differences between individuals, while the latter uses the squares of the differences. Just as the Gini coefficient is a standardised mean difference, the variance can be standardised as the familiar coefficient of variation

$$CV = \frac{\sqrt{V}}{\mu_1} \tag{68}$$

or the standardised variance

$$V_s = \frac{V}{\mu_1^2}. \tag{69}$$

https://doi.org/10.1553/p-j2pg-fm66

The standardised variance is known in evolution and anthropology as Crow's index of the opportunity for selection (Crow, 1958; Courtiol et al., 2012; van Daalen and Caswell, 2024). Like the variance, it is additively decomposable into between-group and within-group components as a measure of inequality (Rosenbluth, 1951).

The variance also has the advantage of being a central moment of the distribution, inviting connections to other moments. These moments highlight other aspects of the distribution (e.g., skewness), and at least some of them can also be decomposed into within- and between-group components. There seems to be no comparable linkage of absolute deviations to such other properties of the distribution.

Remarks

The variance partitioning presented here is related to, but is not the same as, statistical analysis of variance (ANOVA). Our results do not provide hypothesis tests, because they lack an underlying sampling theory to obtain distributions of the variance components under a null hypothesis. If they did, most of the between-group components would be statistically insignificant. However, the factors investigated in demographic studies are usually known a priori to be of social or biological importance. The observation that the contribution of heterogeneity to variance is small does not imply that the factors are not worthy of attention; rather, it merely implies that stochasticity is itself an important factor and is worthy of study (see Caswell 2023, Section 8.5).

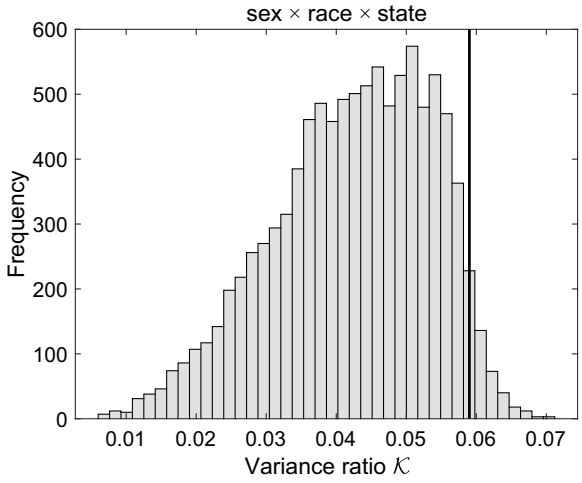
The mixing distribution plays a central role in variance partitioning. It is tempting to wonder if the results are somehow an artefact of that distribution. Insight into the potential effect of the mixing distribution can be obtained by generating marginal distributions at random, combining them to form a rank-one distribution, and then computing the variance ratio. Figure 2 shows the distribution of the variance ratio $\mathcal{K}$ obtained from a sample of 10,000 mixing distributions distributed uniformly over the simplex, for the case of the U.S. sex, race, and states (Table 2). This is a sample from the set of all possible rank-one mixing distributions (including many that are demographically absurd). Even so, the largest variance ratio in this sample is only 0.075. Since the variance ratio with a flat mixing distribution, is $\mathcal{K} = 0.059$, it seems unlikely from Figure 2 that this ratio could be increased very much by any other choice of mixing distribution.

A rank-one mixing distribution is required in order to partition the between-group variance into contributions from main effects and interactions. However, the calculation of the between-group variance itself, as in equation (19), places no restriction on the mixing distribution. It simply treats all the treatment combinations as levels of a single factor, reducing that portion of the calculation to a one-way design.

A recent survey of single-factor studies found that heterogeneity usually explained only 5%–10% of the variance in longevity (Caswell, 2023). The results of the present paper suggest that including multiple factors and their interactions may not increase $\mathcal{K}$ very much. Further comparative research is needed. The variance ratio for lifetime reproduction appears to be larger than that for longevity, and the variance components for LRO appear to be strongly influenced by conditions. van Daalen et al. (2022) found that heterogeneity in

https://doi.org/10.1553/p-j2pg-fm66

Figure 2 Frequency distribution of the variance ratio $\mathcal{K}$, from a sample of 10,000 random, uniformly distributed samples of all possible rank-one mixing distributions over sex, race, and U.S. states. The vertical line indicates the observed value.



maternal age in a rotifer explained about 26% of the variance in LRO under laboratory conditions, but as little as 2% under most conditions that would lead to a stationary population. The sensitivity analysis of variance components (van Daalen and Caswell, 2020) may be helpful for exploring these relationships.

We encourage the use of the methods we present here to explore the contributions of multiple factors and their interactions to variance in demographic outcomes. Patterns await discovery.

Acknowledgements

This research has been supported by the European Research Council through the European Union's Horizon 2020 research and innovation program under ERC Advanced Grant 788195 (FORMKIN). SFvD was also supported by the Postdoctoral Scholar Program at Woods Hole Oceanographic Institution, with funding provided by the Doherty Foundation. We thank two anonymous reviewers for helpful comments. HC acknowledges with gratitude the late Dr. John L. Gill of Michigan State University.

ORCID iDs

Hal Caswell https://orcid.org/0000-0003-4394-6894

Silke F. van Daalen https://orcid.org/0000-0002-2034-8763

References

Arias, E., and Xu, J. (2022). United States life tables, 2020. *National Vital Statistics Reports*, *71*(1). https://doi.org/10.15620/cdc:118055

Arias, E., Xu, J., Tejada-Vera, B., Murphy, S.L., and Bastian, B. (2022). U.S. state life tables, 2020. *National Vital Statistics Reports*, *71*(2). http://doi.org/10.15620/cdc:118271

Atkinson, A.B. (2015). *Inequality: What can be done?* Harvard University Press.

Blurton-Jones, N. (2016). *Demography and evolutionary ecology of Hadza hunter-gatherers*. Cambridge University Press.

Bock, M.J., Jarvis, G.C., Corey, E.L., Stone, E.E., and Gribble, K.E. (2019). Maternal age alters offspring lifespan, fitness, and lifespan extension under caloric restriction. *Scientific Reports*, *9*(1), 3138. https://doi.org/10.1038/s41598-019-40011-z

Brown, G.R., Laland, K.N., and Mulder, M.B. (2009). Bateman's principles and human sex roles. *Trends in Ecology & Evolution*, *24*(6), 297–304. https://doi.org/10.1016/j.tree.2009.02.005

Caswell, H. (2001). *Matrix population models: Construction, analysis, and interpretation* (2nd ed.). Sinauer Associates.

Caswell, H. (2006). Applications of Markov chains in demography. In *MAM2006: Markov anniversary meeting*, (pp. 319–334), Boson Books.

Caswell, H. (2009). Stage, age and individual stochasticity in demography. *Oikos*, *118*(12), 1763–1782. https://doi.org/10.1111/j.1600-0706.2009.17620.x

Caswell, H. (2011). Beyond R_0: Demographic models for variability of lifetime reproductive output. *PloS ONE*, *6*(6), e20809. https://doi.org/10.1371/journal.pone.0020809

Caswell, H. (2014). A matrix approach to the statistics of longevity in heterogeneous frailty models. *Demographic Research*, *31*, 553–592. https://doi.org/10.4054/demres.2014.31.19

Caswell, H. (2023). The contributions of stochasticity and social inequality to lifespan variability. *Demographic Research*, *49*, 309–354. https://doi.org/10.4054/demres.2023.49.13

Caswell, H., and van Daalen, S.F. (2021). Healthy longevity from incidence-based models: More kinds of health than stars in the sky. *Demographic Research*, *45*, 397–452. https://doi.org/10.4054/demres.2021.45.13

Caswell, H., and Zarulli, V. (2018). Matrix methods in health demography: A new approach to the stochastic analysis of healthy longevity and DALYs. *Population Health Metrics*, *16*, 8. https://doi.org/10.1186/s12963-018-0165-5

Courtiol, A., Pettay, J.E., Jokela, M., Rotkirch, A., and Lummaa, V. (2012). Natural and sexual selection in a monogamous historical human population. *Proceedings of the National Academy of Sciences*, *109*(21), 8044–8049. https://doi.org/10.1073/pnas.1118174109

Crow, J.F. (1958). Some possibilities for measuring selection intensities in man. *Human Biology*, *30*, 1–13.

Edwards, R.D., and Tuljapurkar, S. (2005). Inequality in life spans and a new perspective on mortality convergence across industrialized countries. *Population and Development Review*, *31*(4), 645–674. https://doi.org/10.1111/j.1728-4457.2005.00092.x

Falconer, D.S. (1960). *Introduction to quantitative genetics*. Oliver and Boyd.

Feichtinger, G. (1971). *Stochastische Modelle demographischer Prozesse*, Lecture Notes in Economics and Mathematical Systems. Springer Verlag.

Fisher, R.A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, *52*(2), 399–433. https://doi.org/10.1017/s0080456800012163

Fisher, R.A. (1936). *Statistical methods for research workers* (6th ed.). Oliver and Boyd.

Frühwirth-Schnatter, S. (2006). *Finite mixture and Markov switching models*. Springer Verlag.

Hill, K., and Hurtado, A.M. (1996). *Aché life history: The ecology and demography of a foraging people*. Aldine de Gruyter.

Jenkins, S.P., and van Kerm, P. (2009). The measurement of economic inequality. In W. Salverda, B. Nolan, and T.M. Smeeding (Eds.) *The Oxford handbook of economic inequality*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199606061.013.0003

Keilman, N., Tymicki, K., and Skirbekk, V. (2014). Measures for human reproduction should be linked to both men and women. *International Journal of Population Research*, *2014*, 908385. https://doi.org/10.1155/2014/908385

Kempthorne, O. (1957). *An introduction to genetic statistics*. John Wiley & Sons.

Kendall, M.G., and Stuart, A. (1969). *The advanced theory of statistics. Volume 1: Distribution theory* (3rd ed.). Charles Griffin.

Kirk, R.E. (1982). *Experimental design: Procedures for the behavioral sciences* (2nd ed.). Brooks/Cole.

Montez, J.K., Beckfield, J., Cooney, J.K., Grumbach, J.M., Hayward, M.D., Koytak, H.Z., Woolf, S.H., and Zajacova, A. (2020). US state policies, politics, and life expectancy. *The Milbank Quarterly*, *98*(3), 668–699. https://doi.org/10.1111/1468-0009.12469

Montez, J.K., and Farina, M.P. (2021). Do liberal U.S. state policies maximize life expectancy? *Public Policy and Aging Report*, *31*(1), 7–13. https://doi.org/10.1093/ppar/praa035

Moorad, J.A., Promislow, D.E., Smith, K.R., and Wade, M.J. (2011). Mating system change reduces the strength of sexual selection in an American frontier population of the 19th century. *Evolution and Human Behavior*, *32*(2), 147–155. https://doi.org/10.1016/j.evolhumbehav.2010.10.004

Murray, C.J.L., Kulkarni, S.C., Michaud, C., Tomijima, N., Bulzacchelli, M.T., Iandiorio, T.J., and Ezzati, M. (2006). Eight Americas: Investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Medicine*, *3*(9), e260. https://doi.org/10.1371/journal.pmed.0030260

Owoeye, S.M., Oseni, B.M., and Gayawan, E. (2020). Estimating lifetime malnourished period and its statistics based on the concept of Markov chain with reward. *Heliyon*, *6*(5), e04 073. https://doi.org/10.1016/j.heliyon.2020.e04073

Permanyer, I., Sasson, I., and Villavicencio, F. (2023). Group- and individual-based approaches to health inequality: Towards an integration. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *186*(2), 217–240. https://doi.org/10.1093/jrsssa/qnac001

Permanyer, I., and Scholl, N. (2019). Global trends in lifespan inequality: 1950–2015. *PLoS ONE*, *14*(5), e0215742. https://doi.org/10.1371/journal.pone.0215742

Rao, T.R., and Sarma, S. (1986). Demographic parameters of *Brachionus patulus* Muller (Rotifera) exposed to sublethal DDT concentrations at low and high food levels. *Hydrobiologia*, *139*, 193–200. https://doi.org/10.1007/bf00028292

Rényi, A. (1970). *Probability theory*. North-Holland.

Rosenbluth, G. (1951). Note on Mr. Schutz's measure of income inequality. *The American Economic Review*, *41*(5), 935–937.

Sahai, H., and Ageel, M.I. (2012). *The analysis of variance: Fixed, random, and mixed models*. Springer Science & Business Media.

Sen, A. (1997). *On economic inequality* (2nd ed.). Oxford University Press.

Shkolnikov, V.M., Andreev, E.M., Houle, R., and Vaupel, J.W. (2007). The concentration of reproduction in cohorts of women in Europe and the United States. *Population and Development Review*, *33*(1), 67–99. https://doi.org/10.1111/j.1728-4457.2007.00159.x

Snyder, R.E., and Ellner, S.P. (2018). Pluck or luck: Does trait variation or chance drive variation in lifetime reproductive success? *The American Naturalist*, *191*(4), E90–E107. https://doi.org/10.1086/696125

Steel, R.G.D., and Torrie, J.H. (1960). *Principles and procedures of statistics*. McGraw-Hill.

Steiner, U.K., and Tuljapurkar, S. (2012). Neutral theory for life histories and individual variability in fitness components. *Proceedings of the National Academy of Sciences*, *109*(12), 4684–4689. https://doi.org/10.1073/pnas.1018096109

Therborn, G. (2012). The killing fields of inequality. *International Journal of Health Services*, *42*(4), 579–589. https://doi.org/10.2190/hs.42.4.a

Therborn, G. (2014). *The killing fields of inequality*. John Wiley & Sons.

Tuljapurkar, S., Zuo, W., Coulson, T., Horvitz, C., and Gaillard, J.M. (2020). Skewed distributions of lifetime reproductive success: Beyond mean and variance. *Ecology Letters*, *23*(4), 748–756. https://doi.org/10.1111/ele.13467

U.S. Census Bureau, Population Division (2022). *Annual estimates of the resident population by sex, race, and Hispanic origin for the United States: April 1, 2020 to July 1, 2021* (NC-EST2021-SR11H). https://www2.census.gov/programs-surveys/popest/tables/2020-2021/national/asrh/nc-est2021-sr11h.xlsx

van Daalen, S., and Caswell, H. (2015). Lifetime reproduction and the second demographic transition: Stochasticity and individual variation. *Demographic Research*, *33*, 561–588. https://doi.org/10.4054/demres.2015.33.20

van Daalen, S.F., and Caswell, H. (2017). Lifetime reproductive output: Individual stochasticity, variance, and sensitivity analysis. *Theoretical Ecology*, *10*, 355–374. https://doi.org/10.1007/s12080-017-0335-2

https://doi.org/10.1553/p-j2pg-fm66

van Daalen, S.F., and Caswell, H. (2020). Variance as a life history outcome: Sensitivity analysis of the contributions of stochasticity and heterogeneity. *Ecological Modelling*, *147*, 101856. https://doi.org/10.1016/j.ecolmodel.2019.108856

van Daalen, S.F., and Caswell, H. (2024). Demographic sources of variation in fitness. In O. Burger, R. Lee, and R. Sear (Eds.) *Human evolutionary demography*. Open Book Publishers. https://doi.org/10.11647/obp.0251.15

van Daalen, S.F., Hernández, C.M., Caswell, H., Neubert, M.G., and Gribble, K.E. (2022). The contributions of maternal age heterogeneity to variance in lifetime reproductive output. *The American Naturalist*, *199*(5), 603–616. https://doi.org/10.1086/718716

van Raalte, A.A., and Caswell, H. (2013). Perturbation analysis of indices of lifespan variability. *Demography*, *50*(5), 1615–1640. https://doi.org/10.1007/s13524-013-0223-3

van Raalte, A.A., Sasson, I., and Martikainen, P. (2018). The case for monitoring life-span inequality. *Science*, *362*(6418), 1002–1004. https://doi.org/10.1126/science.aau5811

Varas Enríquez, P.J., Van Daalen, S., and Caswell, H. (2022). Individual stochasticity in the life history strategies of animals and plants. *PLoS ONE*, *17*(9), e0273 407. https://doi.org/10.1371/journal.pone.0273407

Vaupel, J.W. (1988). Inherited frailty and longevity. *Demography*, *25*(2), 277–287. https://doi.org/10.2307/2061294

Vaupel, J.W., and Goodwin, D.G. (1987). The concentration of reproduction among US women, 1917–80. *Population and Development Review*, *13*(4), 723–730. https://doi.org/10.2307/1973030

Vaupel, J.W., Zhang, Z., and van Raalte, A.A. (2011). Life expectancy and disparity: An international comparison of life table data. *BMJ Open*, *1*(1), e000 128. https://doi.org/10.1136/bmjopen-2011-000128

Wei, R., Anderson, R.N., Curtin, L.R., and Arias, E. (2012). U.S. decennial life tables for 1999–2001: State life tables. *National Vital Statistics Reports*, *60*(9), 1–67.

Yates, F. (1934). The analysis of multiple classifications with unequal numbers in the different classes. *Journal of the American Statistical Association*, *29*(185), 51–66. https://doi.org/10.2307/2278459

Zarulli, V., and Caswell, H. (2022). Longer healthy life, but for how many? Insights on healthy lifespan inequality from the Global Burden of Disease Study. *Annals of Operations Research* https://doi.org/10.1007/s10479-024-06203-1

https://doi.org/10.1553/p-j2pg-fm66