# Workshop

## An introduction to the analysis of population-level kinship structures

Diego Alburez-Gutiérrez[†]

[†]Kinship Inequalities Research Group,
Max Planck Institute for Demographic Research

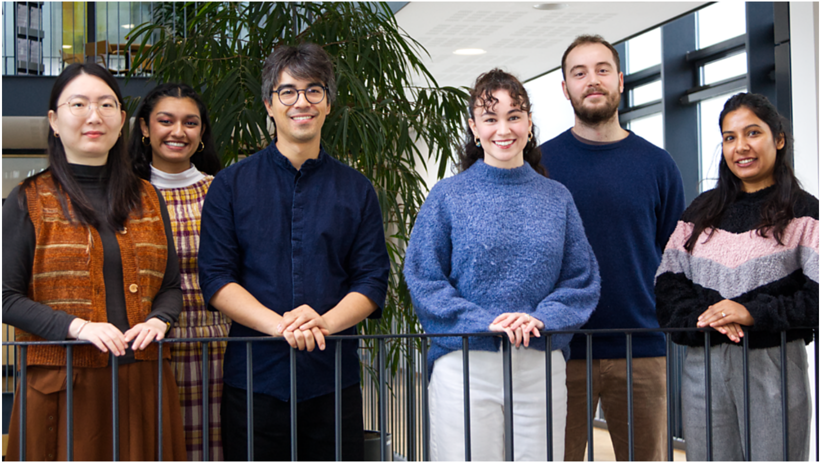
International Institute for Population Sciences,
India, November 24, 2025

MAX PLANCK INSTITUTE FOR DEMOGRAPHIC RESEARCH
MAX-PLANCK-INSTITUT FÜR DEMOGRAFISCHE FORSCHUNG

# Who am I?

1. 2012: BA in Anthropology (Del Valle University, Guatemala)
2. 2018: PhD in Demography (London School of Economics, UK)
3. 2019: Postdoc (Max Planck Institute for Demographic Research)
4. 2022: Group Leader: Kinship Inequalities Research Group (MPIDR)

# Kinship Inequalities Research Group, MPIDR



# Research focus

1. Kinship dynamics
2. Demography
3. Family bereavement
4. Armed conflict and mass violence
5. Collective memory

Rio Negro, Guatemala, 2010

# Introductions

Find someone you don't know and ask:

1. Their name
2. Where they study/work
3. Favorite food
4. What is the demography of kinship?

# This presentation

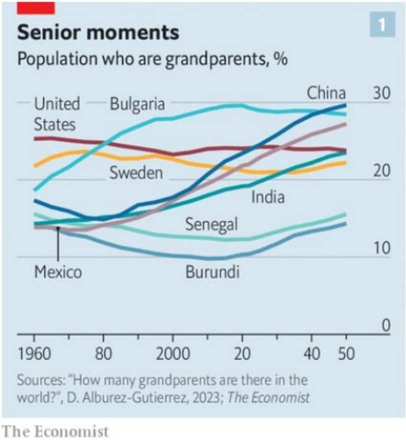
1. Introduction to kinship demography
2. Demographic models of kinship
3. Example: projections of kinship
4. Course setup
5. Appendix: The Goodman-Keyfitz-Pullum kinship equations

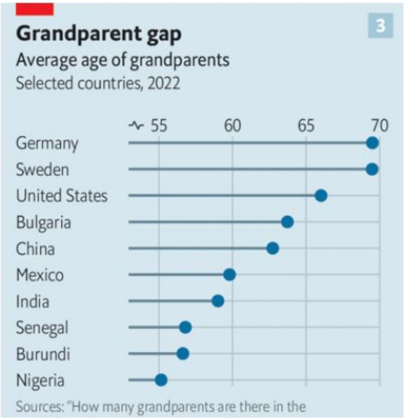
# Introduction to kinship demography

# Consider a baby born in India in 2020...

1. How old were her grandparents when she was born, on average?
2. How many living children will she have on her 70th birthday?
3. How many grandchildren?

# Kinship structure is a questions of societal interest[1]




[1] 'The age of the grandparent has arrived.' (Jan 2023). The Economist. https://www.economist.com/international/2023/01/12/the-age-of-the-grandparent-has-arrived

# Definitions (1)[2]

### Kinship

Social relationships that bind individuals together through culturally shared definitions of relatedness on biological, legal, or normative grounds, ultimately constituting family systems.

### Family

More narrow group of kin given special privilege which, among other things, organize the provision of support, socialization, and social placement of its members.

---

[2] Alburez-Gutierrez, D., Barban, N., Caswell, H., Kolk, M., Margolis, R., Smith-Greenaway, E., Song, X., Verdery, A., & Zagheni, E. (2022). Kinship, Demography, and Inequality: Review and Key Areas for Future Development. *SocArXiv*. https://doi.org/10.31235/osf.io/fk7x9

# The role of kinship in human societies

1. Socialisation, protection, and sustenance
2. Inter-generational solidarity: exchanges and bequests
3. Social structure and identity
4. Early-life conditions $\rightarrow$ later-life outcomes

# Definitions (2)[3]

**Kinship demography**

The study of family networks, their structures and dynamics from a demographic perspective and using demographic methods.

[3] Alburez-Gutierrez, D., Barban, N., Caswell, H., Kolk, M., Margolis, R., Smith-Greenaway, E., Song, X., Verdery, A., & Zagheni, E. (2022).Kinship, Demography, and Inequality: Review and Key Areas for Future Development. *SocArXiv*. https://doi.org/10.31235/osf.io/fk7x9

# Kinship as a demographic human universal

1. All humans are born
2. All humans die
3. All humans are embedded in kinship structures[4]
4. No particular family configuration is universal or stable

[4] Caswell, H. (2019). The formal demography of kinship: A matrix formulation. *Demographic Research*, *41*, 679–712

## Demographic models of kinship

# What are kinship models?

1. Kinship is an *emergent property* of demographic systems
2. Simplified representation of interaction between reproduction, survival (and more)
3. Can be formal (mathematical) or simulation-based (computational)

# The *Strong Ergodic Principle*[6]

"A closed population with unchanging mortality and fertility rates has an **implicit age structure**."

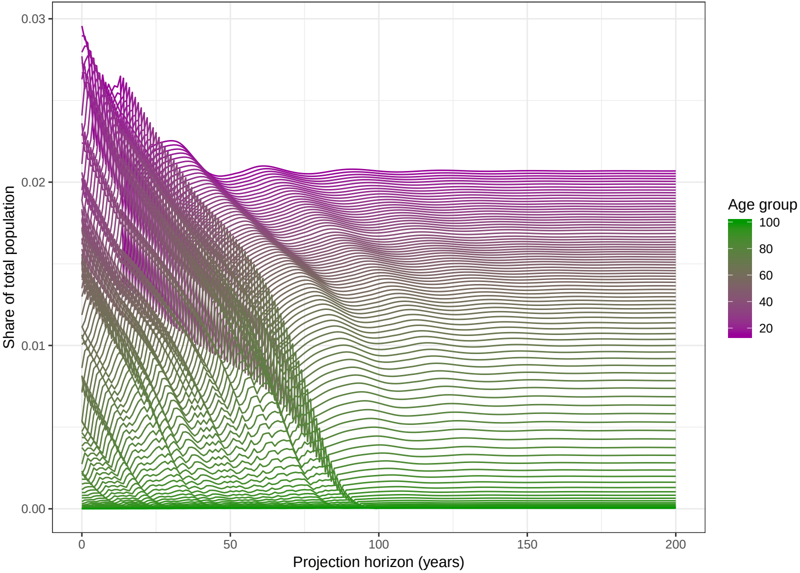
The proportion of the population aged $x$ to $x+n$ in a stable population is[5]:

$$\frac{{}_nK_x}{{}_\infty K_0} = b\frac{{}_nL_x}{l_0}e^{-rx} \tag{1}$$

[5] Wachter, K. W. (2014). *Essential demographic methods* [OCLC: 931410976]. Harvard Univ. Press

[6] Sharpe, F., & Lotka, A. (1911). A problem in age-distribution. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *21*(124), 435–438. https://doi.org/10.1080/14786440408637050

# Implied age structures using Leslie matrices



# The demographic foundations of kin structure

THEORETICAL POPULATION BIOLOGY 5, 1–27 (1974)

## Family Formation and the Frequency of Various Kinship Relationships

LEO A. GOODMAN

*The University of Chicago*

NATHAN KEYFITZ AND THOMAS W. PULLUM

*Harvard University*

Received January 19, 1970

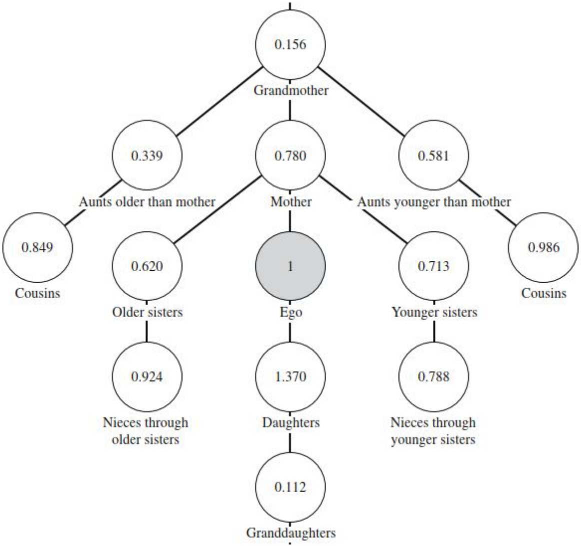
A set of age-specific rates of birth and death implies expected numbers of kin. An individual girl or woman chosen at random out of a population whose birth and death rates are specified can be expected to have a certain number of older sisters, younger sisters, nieces, cousins; expressions for these values are provided for both total kin and kin who are still living. Included also are the

# Implied kinship structures

"A fixed set of age-specific rates implies the probability that a girl aged *a* has a living mother and great-grandmother, as well as her expected number of daughters, sisters, aunts, nieces, and cousins."[7]

[7] Keyfitz, N. (1985). *Applied mathematical demography*. Springer

# Expected number of relatives (USA 1965, F, 45 yo)[8]




[8] Keyfitz, N. (1985). *Applied mathematical demography*. Springer

# Formal models of kinship

Given a set of:

- age-specific fertility rates
- survival probabilities
- simplifying assumptions

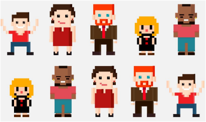
The models produce:

1. Number of (living/dead) kin
2. Age distribution of relatives
3. From the point of view of an average member of the population ('Focal')

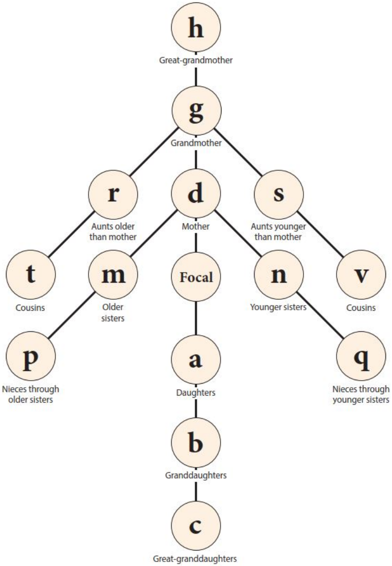
# Focal: an average member of the population (ChatGPT)

# Matrix kinship models

1. The relatives of Focal constitute a population
2. They can be modelled using traditional projection methods
3. Matrix operations provide an efficient implementation



# Kinship structure



# Implementation: time-invariant, one-sex models[9]

The models are of the general form:

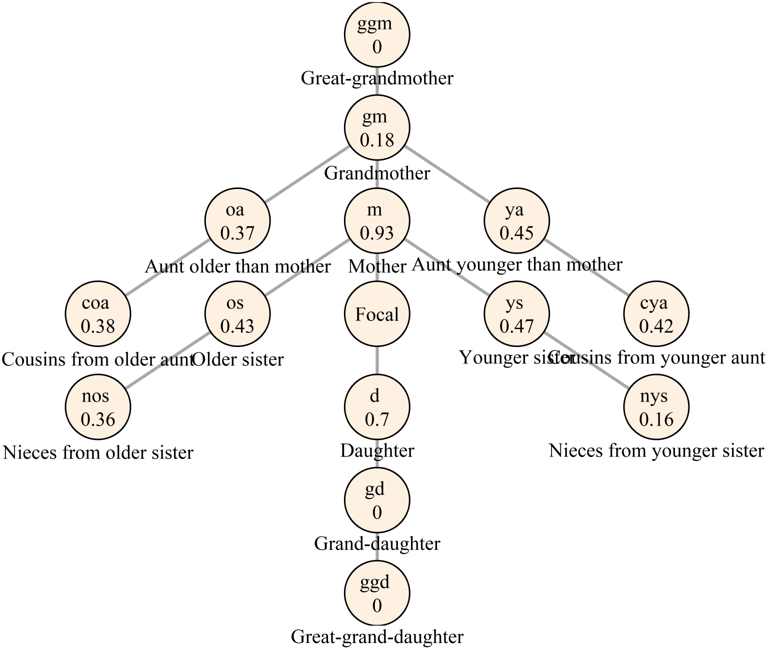
$$\underbrace{\mathbf{k}(x+1)}_{\substack{\text{age structure of kin} \\ \text{at Focal's age } x+1}} = \underbrace{\mathbf{U}\,\mathbf{k}(x)}_{\substack{\text{ageing and survival} \\ \text{of existing kin}}} + \underbrace{\begin{cases} \mathbf{0} \\ \mathbf{F}\,\mathbf{k}^{*}(x) \end{cases}}_{\substack{\text{new kin members} \\ \text{added to the population}}} .$$

where:

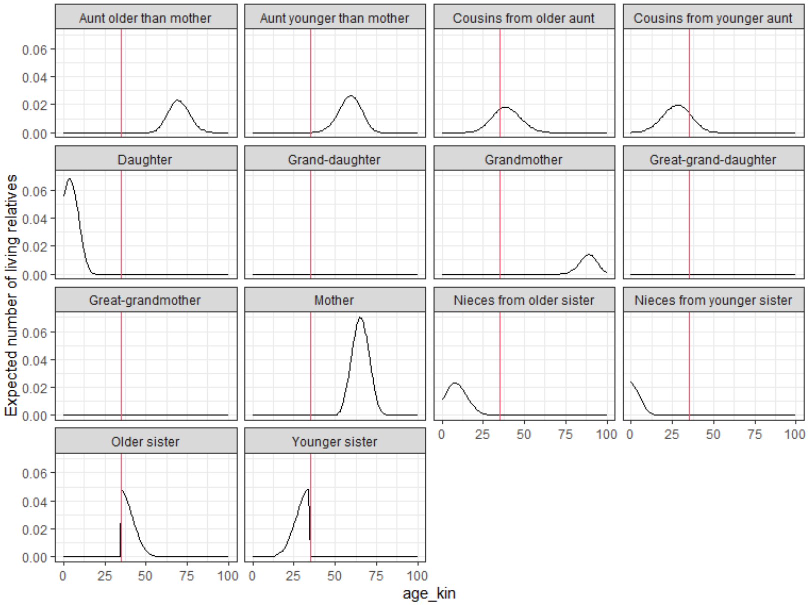
- **U** a matrix with survival probabilities in the subdiagonal
- **F** a matrix with fertility rates in the first row

[9]Caswell, H. (2019).The formal demography of kinship: A matrix formulation. *Demographic Research*, *41*, 679–712

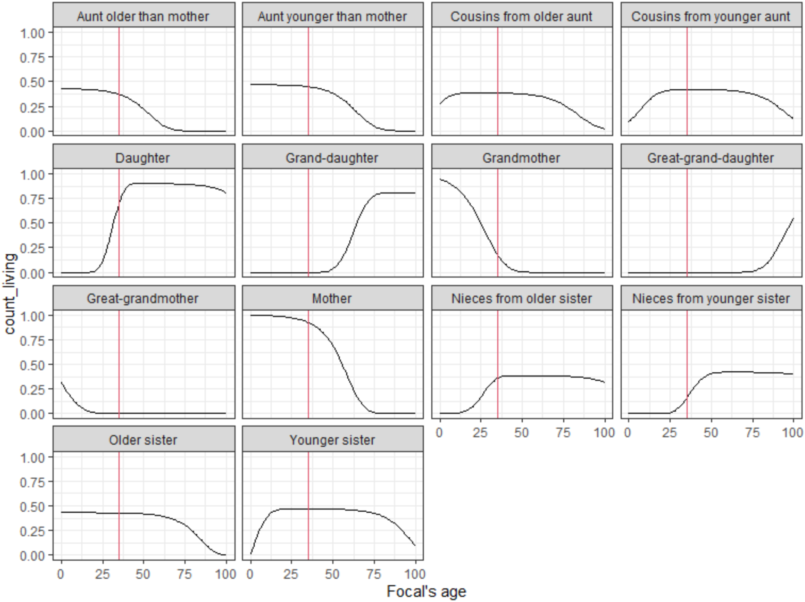
# Kinship structure



# Age distributions of kin



# Expected number of kin



# Daughters

Daughters ($\mathbf{a}$) are the result of the reproduction of Focal:

$$\underbrace{\mathbf{a}(x+1)}_{\substack{\text{age structure of daughters}\\ \text{at Focal's age } x+1}} = \underbrace{\mathbf{U}\,\mathbf{a}(x)}_{\substack{\text{ageing and survival}\\ \text{of existing daughters}}} + \underbrace{\mathbf{F}\,\mathbf{e}_x.}_{\substack{\text{new daughters}\\ \text{(subsidy)}}} \tag{2}$$

$$a(0) = \mathbf{0}.$$

where:

- $\mathbf{U}$ is a matrix with survival probabilities in the subdiagonal
- $\mathbf{F}$ is a matrix with fertility rates in the first row
- $\mathbf{F}\,\mathbf{e}_x$ is the subsidy vector
- $e_x$ is the unit vector for age $x$
- $a(0)$ is the distribution of daughters at Focal's birth

# Mothers

The population of mothers ($\mathbf{d}$) of Focal consists of at most a single individual:

$$\underbrace{\mathbf{d}(x+1)}_{\substack{\text{age structure of mothers} \\ \text{at Focal's age } x+1}} = \underbrace{\mathbf{U}\,\mathbf{d}(x)}_{\substack{\text{ageing and survival} \\ \text{of existing mothers}}} + \underbrace{\mathbf{0}.}_{\substack{\text{new mothers} \\ \text{(subsidy)}}} \tag{3}$$

$$d(0) = \pi.$$

where:

- $b(0)$ is the distribution of mothers at Focal's birth
- $\pi$ is the distribution of ages of mothers in the population

# All models[10]

**Table 1:** **Summary of the components of the kin model given in equations (4) and (5)**

| Symbol | Kin | Initial condition | Subsidy $\beta(x)$ |
|---|---|---|---|
| $\mathbf{a}$ | daughters | $\mathbf{0}$ | $\mathbf{F}\mathbf{e}_x$ |
| $\mathbf{b}$ | granddaughters | $\mathbf{0}$ | $\mathbf{F}\mathbf{a}(x)$ |
| $\mathbf{c}$ | great-granddaughters | $\mathbf{0}$ | $\mathbf{F}\mathbf{b}(x)$ |
| $\mathbf{d}$ | mothers | $\boldsymbol{\pi}$ | $\mathbf{0}$ |
| $\mathbf{g}$ | grandmothers | $\sum_i \pi_i \mathbf{d}(i)$ | $\mathbf{0}$ |
| $\mathbf{h}$ | great-grandmothers | $\sum_i \pi_i \mathbf{g}(i)$ | $\mathbf{0}$ |
| $\mathbf{m}$ | older sisters | $\sum_i \pi_i \mathbf{a}(i)$ | $\mathbf{0}$ |
| $\mathbf{n}$ | younger sisters | $\mathbf{0}$ | $\mathbf{F}\mathbf{d}(x)$ |
| $\mathbf{p}$ | nieces via older sisters | $\sum_i \pi_i \mathbf{b}(i)$ | $\mathbf{F}\mathbf{m}(x)$ |
| $\mathbf{q}$ | nieces via younger sisters | $\mathbf{0}$ | $\mathbf{F}\mathbf{n}(x)$ |
| $\mathbf{r}$ | aunts older than mother | $\sum_i \pi_i \mathbf{m}(i)$ | $\mathbf{0}$ |
| $\mathbf{s}$ | aunts younger than mother | $\sum_i \pi_i \mathbf{n}(i)$ | $\mathbf{F}\mathbf{g}(x)$ |
| $\mathbf{t}$ | cousins from aunts older than mother | $\sum_i \pi_i \mathbf{p}(i)$ | $\mathbf{F}\mathbf{r}(x)$ |
| $\mathbf{v}$ | cousins from aunts younger than mother | $\sum_i \pi_i \mathbf{q}(i)$ | $\mathbf{F}\mathbf{s}(x)$ |

[10] Caswell, H. (2019). The formal demography of kinship: A matrix formulation. *Demographic Research*, *41*, 679–712

# Typology of formal kinship models

| No | time | sex | state | reference |
|---|---|---|---|---|
| **1** | **invariant** | **female** | **age** | [11] |
| 2 | variant | female | age | [12] |
| 3 | invariant | two | age | [13] |
| 4 | invariant | female | multiple | [14] |
| 5 | variant | two | multiple | [15] |

[11] Caswell, H. (2019). The formal demography of kinship: A matrix formulation. *Demographic Research*, *41*, 679–712

[12] Caswell, H., & Song, X. (2021). The formal demography of kinship. III. kinship dynamics with time-varying demographic rates. *Demographic Research*, *45*, 517–546

[13] Caswell, H. (2022). The formal demography of kinship IV: Two-sex models and their approximations. *Demographic Research*, *47*, 359–396

[14] Caswell, H. (2020). The formal demography of kinship II: Multistate models, parity, and sibship. *Demographic Research*, *42*, 1097–1146

[15] Williams, I., Alburez-Gutierrez, D., Caswell, H., & Song, X. (2023). *DemoKin: 1.0.3*. https://CRAN.R-project.org/package=DemoKin

# Consider a baby born in India in 2020...

1. How old were her grandparents when she was born, on average?
2. How many living children will she have on her 70th birthday?
3. How many grandchildren?

# DemoKin: matrix kinship models in R[16]

- Time-(in)variant models
- One/two-sex models
- Multistate models
- Kin loss by cause of death
- More in the lab session...

[16] Williams, I., Alburez-Gutierrez, D., Caswell, H., & Song, X. (2023). *DemoKin: 1.0.3.* https://CRAN.R-project.org/package=DemoKin

# Example: projections of kinship

RESEARCH ARTICLE | DEMOGRAPHY

# Projections of human kinship for all countries

Diego Alburez-Gutierrez, Iván Williams, and Hal Caswell Authors Info & Affiliations

Edited by Cyrus Chu, Academia Sinica (Taiwan), Taipei, Taiwan; received September 12, 2023; accepted October 20, 2023

**December 19, 2023** | 120 (52) e2315722120 | https://doi.org/10.1073/pnas.2315722120

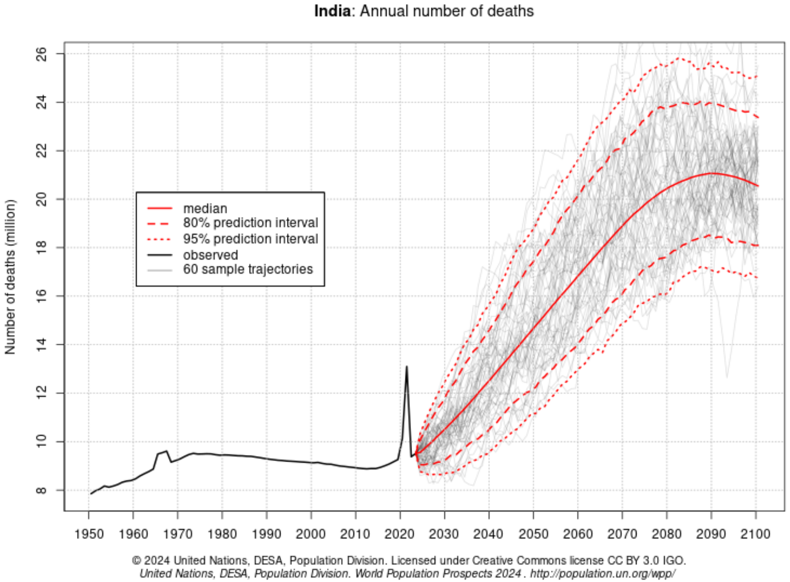


## Significance

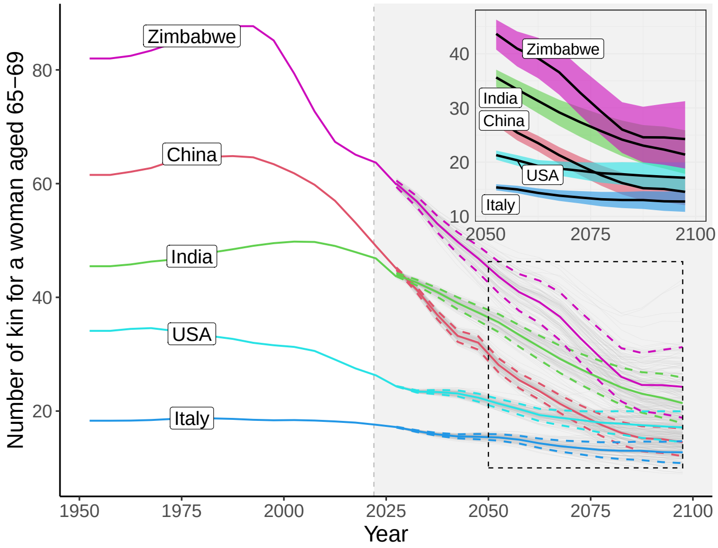
Rapid demographic change is expected to transform the supply of kin worldwide.
Changes in the size and composition of kinship networks matter because relatives

# Data: empirical estimates and probabilistic projections

1. 2022 Revision of the United Nations World Population Prospects (UNWPP)
2. Empirical data (1950-2021)
3. Probabilistic projections (2021-2100): 1,000 trajectories per country

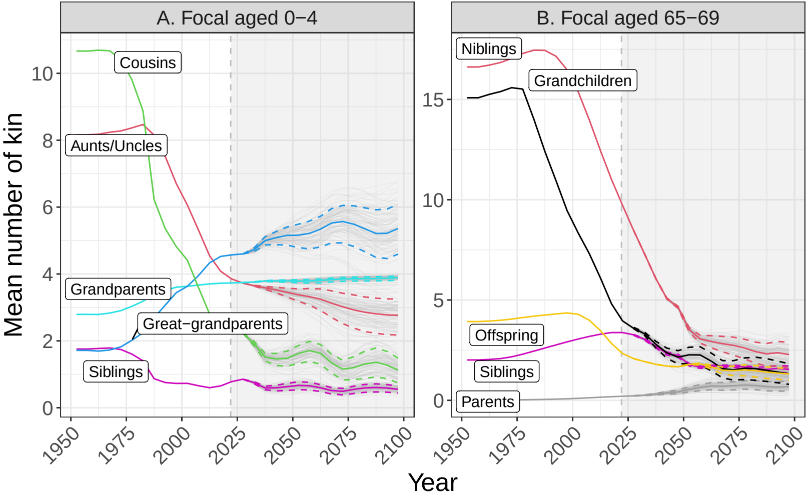

India: Annual number of deaths
Number of deaths (million)
8
10
12
14
16
18
20
22
24
26
1950
1960
1970
1980
1990
2000
2010
2020
2030
2040
2050
2060
2070
2080
2090
2100
median
80% prediction interval
95% prediction interval
observed
60 sample trajectories
© 2024 United Nations, DESA, Population Division. Licensed under Creative Commons license CC BY 3.0 IGO.
United Nations, DESA, Population Division. World Population Prospects 2024 . http://population.un.org/wpp/

# Total number of kin (all kin combined) for a 5yo woman[17]




[17] Alburez-Gutierrez, D., Williams, I., & Caswell, H. (2023). Projections of human kinship for all countries. *Proceedings of the National Academy of Sciences*, *120*(52), e2315722120. https://doi.org/10.1073/pnas.2315722120

# Number of living kin in China



# Kinship: An active area of study

Theoretical Population Biology
Volume 163, June 2025, Pages 1-12

## A mathematical framework for time-variant multi-state kinship modelling

Joe W.B. Butterick, Peter W.F. Smith, Jakub Bijak, Jason Hilton

Show more

Add to Mendeley Share Cite

https://doi.org/10.1016/j.tpb.2025.02.002 Get rights and content



PNAS

RESEARCH ARTICLE | POPULATION BIOLOGY

## The Kinship Formula: Inferring the numbers of all kin from any structured population projection model

Christophe F. D. Coste Authors Info & Affiliations

Edited by Nils Stenseth, Universitetet i Oslo, Oslo, Norway; received June 13, 2024; accepted September 7, 2025
November 14, 2025 | 122 (46) e2411888122 | https://doi.org/10.1073/pnas.2411888122

# Course setup

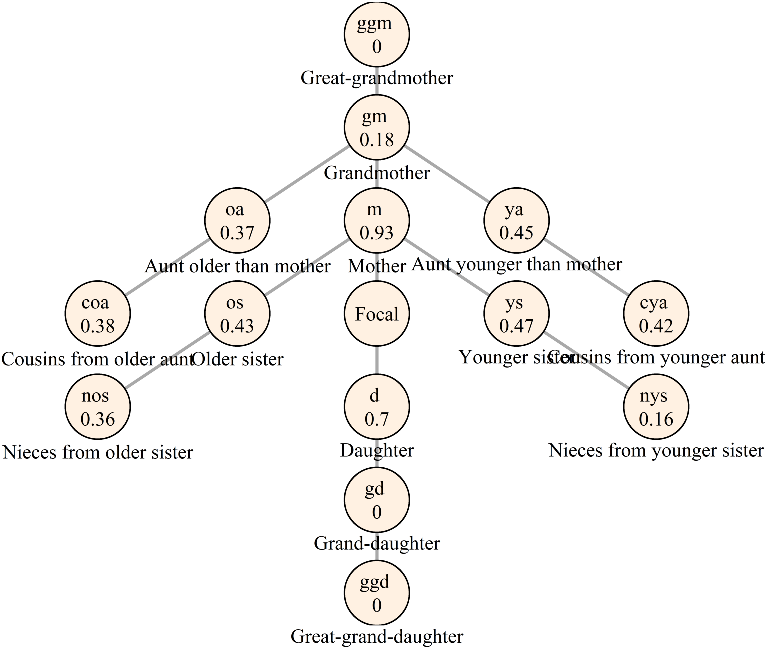
# Course overview

1. Mon - "Introduction and one-sex time-invariant models"
2. Tue - "Two-sex time-variant models"
3. Wed - "Multistate kinship models"
4. Thu - Group Work

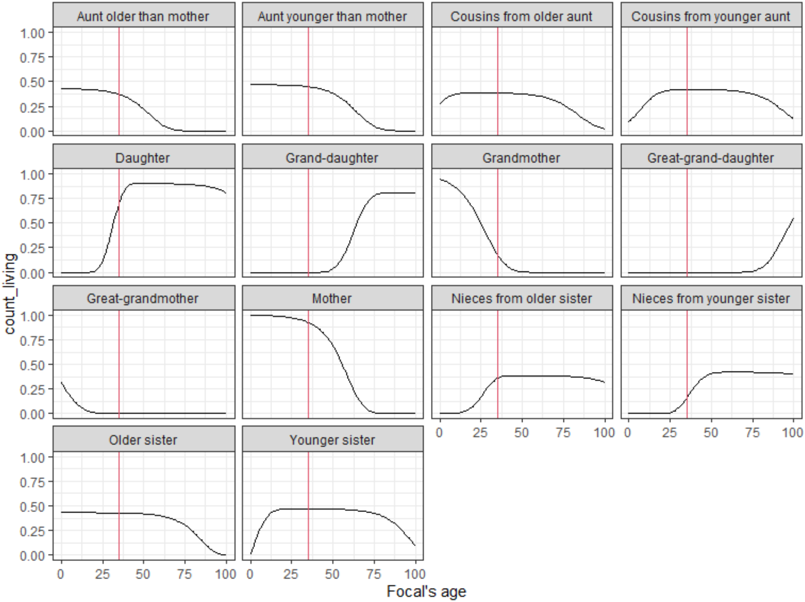
Website for lab sessions:
https://alburez.me/kinship_workshop_iips/

# Appendix: The Goodman-Keyfitz-Pullum kinship equations

# The tree of life



# Expected number of kin



# Daughters

$B_1(a)$ is the expected number of living daughters in a time-invariant female-only population[18]:

$$B_1(a) = \int_{\alpha}^{a} m(x) l(a-x)\, dx \tag{4}$$

where:

- $m(x)$ are fertility rates of mothers
- $l(a-x)$ are survival probabilities of daughters

[18] Goodman, L. A. (1974). Family Formation and the Frequency of Various Kinship Relationships. *Theoretical Population Biology*, 27

# Daughters

If $a = 20$ and $\alpha = 15$; then:

$$B_1(20) \approx \sum_{15}^{20} m(x)l(20 - x)$$

So...

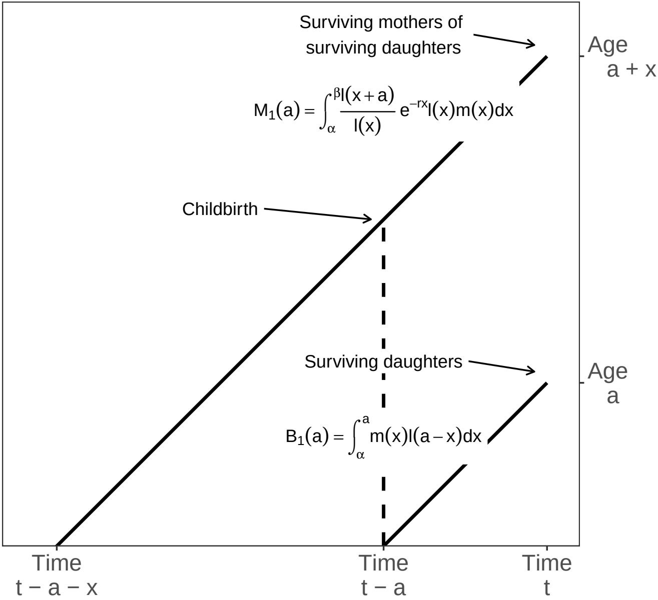
$$B_1(20) \approx m(15)l(5) + m(16)l(4) + m(17)l(3) \ldots$$

# Granddaughters

$B_1(a)$ is the expected number of living granddaughters in a time-invariant female-only population[19]:

$$B_2(a) = \int_{\alpha}^{a} m(x) \int_{\alpha}^{a-x} l(y)m(y)l(a-x-y)\, dy\, dx \qquad (5)$$

[19] Goodman, L. A. (1974). Family Formation and the Frequency of Various Kinship Relationships. *Theoretical Population Biology*, 27



Surviving mothers of surviving daughters
$M_1(a) = \int_\alpha^\beta \frac{l(x+a)}{l(x)} e^{-rx} l(x) m(x) dx$
Childbirth
Surviving daughters
$B_1(a) = \int_\alpha^a m(x) l(a-x) dx$
Age a + x
Age a
Time t − a − x
Time t − a
Time t

# Mothers

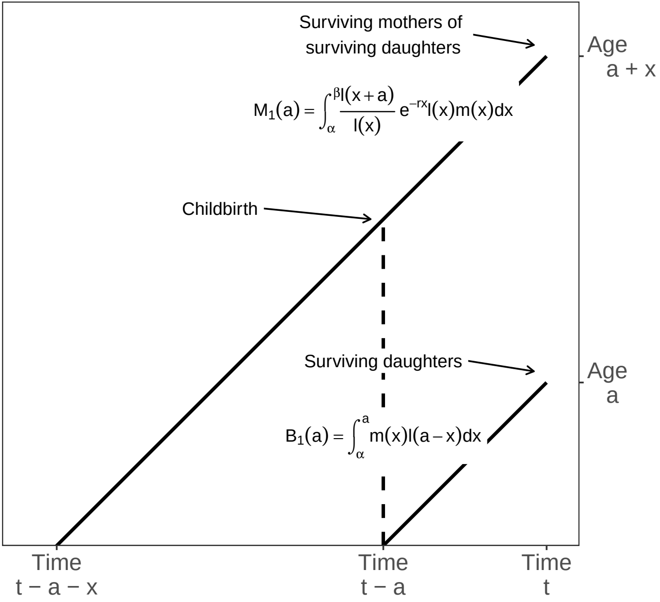
$M_1(a)$ is the probability of having a living mother in a time-invariant female-only population[20]:

$$M_1(a) = \int_{\alpha}^{\beta} \underbrace{\frac{l(x+a)}{l(x)}}_{\substack{\text{prob of surviving} \\ \text{from } x \text{ to } a+x}} \times \underbrace{W(x)}_{\substack{\text{age distribution of} \\ \text{mothers}}} dx. \tag{6}$$

where:

- $W(x) = e^{-rx} l(x) m(x)$ is the age distribution of mothers
- $l(x)$ are survival probabilities
- $m(x)$ are fertility rates
- $r$ is the population growth rate
- $\alpha$-$\beta$ is the reproductive period

[20] Goodman, L. A. (1974). Family Formation and the Frequency of Various Kinship Relationships. *Theoretical Population Biology*, 27


Surviving mothers of
surviving daughters
$M_1(a) = \int_\alpha^\beta \frac{l(x+a)}{l(x)} e^{-rx} l(x)m(x)dx$
Childbirth
Surviving daughters
$B_1(a) = \int_\alpha^a m(x)l(a-x)dx$
Age
a + x
Age
a
Time
t − a − x
Time
t − a
Time
t

# Grandmothers

$M_2(a)$ is the expected number of living grandmothers in a time-invariant female-only population[21]:

$$M_2(a) = \int_{\alpha}^{\beta} \underbrace{M_1(a)}_{\substack{\text{prob of having}\\ \text{living mother}}} \times \underbrace{W(x)}_{\substack{\text{age distribution of}\\ \text{mothers}}} \, dx. \qquad (7)$$

[21] Goodman, L. A. (1974). Family Formation and the Frequency of Various Kinship Relationships. *Theoretical Population Biology*, 27

# Great-grandmothers

$M_3(a)$ is the expected number of living great-grandmothers in a time-invariant female-only population[22]:

$$M_3(a) = \int_{\alpha}^{\beta} \underbrace{M_2(a)}_{\text{number of grandmother}} \times \underbrace{W(x)}_{\text{age distribution of mothers}} \, dx. \tag{8}$$

[22] Goodman, L. A. (1974). Family Formation and the Frequency of Various Kinship Relationships. *Theoretical Population Biology*, 27

# Sisters

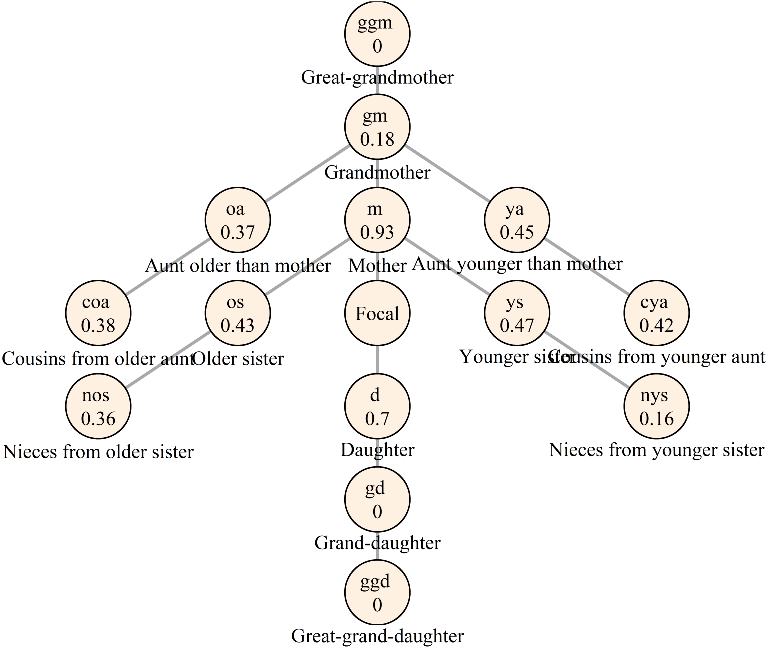
$B_1(a)$ is the expected number of living older sisters in a time-invariant female-only population[23]:

$$S^{old}(a) = \int_{\alpha}^{\beta} \int_{\alpha}^{x} m(y) l(a + x - y) W(x) \, dy \, dx \tag{9}$$

$$S^{young}(a) = \int_{\alpha}^{\beta} \int_{0}^{a} \left[ \frac{l(x+u)}{l(x)} \right] m(x+u) l(a-u) \, du \, W(x) \, dx \tag{10}$$

[23] Goodman, L. A. (1974). Family Formation and the Frequency of Various Kinship Relationships. *Theoretical Population Biology*, 27

# Why do we model younger and older sisters/kin separately?



# Demographic subsidy

"New members of the population arise not from reproduction of current members, but from elsewhere"[24]

- Can you think of other instances of 'subsidy' in demography?

[24] Caswell, H. (2019). The formal demography of kinship: A matrix formulation. *Demographic Research*, *41*, 679–712