Package 'DemoKin'

July 21, 2025

Title Estimate Population Kin Distribution

Description Estimate population kin counts and its distribution by type, age and sex.
The package implements one-sex and two-sex framework for studying living-death availability,
with time varying rates or not, and multi-stage model.

Version 1.0.3

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Suggests knitr, rmarkdown, testthat (>= 3.0.0), ggplot2

VignetteBuilder knitr

Imports dplyr, tidyr, purrr, progress, matrixcalc, Matrix, MASS,
stats, igraph, magrittr, data.table, lifecycle

URL <https://github.com/IvanWilli/DemoKin>

BugReports <https://github.com/IvanWilli/DemoKin/issues>

Depends R (>= 3.5.0)

Config/testthat/edition 3

NeedsCompilation no

Author Iván Williams [cre],
Diego Alburez-Gutierrez [aut],
Xi Song [ctb],
Caswell Hal [ctb]

Maintainer Iván Williams <act.ivanwilliams@gmail.com>

Repository CRAN

Date/Publication 2023-06-05 09:00:02 UTC

Contents

---

demokin_codes *DemoKin codes, Caswell (2020) codes, and useful labels.*

Description

DemoKin codes, Caswell (2020) codes, and useful labels.

Usage

```
demokin_codes
```

Format

A data.frame with codes and labels for distinction between kin types.

---

fra_asfr_sex *Fertility for France (2012) by sex in Caswell (2022).*

Description

Fertility for France (2012) by sex in Caswell (2022).

Usage

```
fra_asfr_sex
```

Format

A data.frame with age specific fertility rates by age and sex.

Source

Caswell (2022)

---

fra_surv_sex *Survival probability for France (2012) by sex in Caswell (2022).*

Description

Survival probability for France (2012) by sex in Caswell (2022).

Usage

```
fra_surv_sex
```

Format

A data.frame with survival probabilities by age and sex.

Source

Caswell (2022)

| f_caswell_2021 | *Historic and projected fertility ratios from Sweden used in Caswell (2021)* |

Description

Historic and projected fertility ratios from Sweden used in Caswell (2021)

Usage

```
f_caswell_2021
```

Format

A matrix f with years as cols and ages as rows.

Source

Caswell (2019)

| kin | *Estimate kin counts in a one-sex framework.* |

Description

Implementation of Goodman-Keyfitz-Pullum equations in a matrix framework. This produce a matrilineal (or patrilineal) kin count distribution by kin and age.

Usage

```
kin(
  p = NULL,
  f = NULL,
  time_invariant = TRUE,
  pi = NULL,
  n = NULL,
  output_cohort = NULL,
  output_period = NULL,
  output_kin = NULL,
  birth_female = 1/2.04,
  stable = lifecycle::deprecated(),
  U = lifecycle::deprecated()
)
```

Arguments

p	numeric. A vector (atomic) or matrix with probabilities (or survival ratios, or transition between age class in a more general perspective) with rows as ages (and columns as years in case of matrix, being the name of each col the year).
f	numeric. Same as p but for fertility rates.
time_invariant	logical. Constant assumption for a given year rates. Default TRUE.
pi	numeric. Same as U but for childbearing distribution (sum to 1). Optional.
n	numeric. Only for time_invariant = FALSE. Same as p but for population distribution (counts or %). Optional.
output_cohort	integer. Vector of year cohorts for returning results. Should be within input data years range.
output_period	integer. Vector of period years for returning results. Should be within input data years range.
output_kin	character. kin types to return: "m" for mother, "d" for daughter,...
birth_female	numeric. Female portion at birth. This multiplies f argument. If f is already for female offspring, this needs to be set as 1.
stable	logic. Deprecated. Use time_invariant.
U	logic. Deprecated. Use p.

Details

See Caswell (2019) and Caswell (2021) for details on formulas. One sex only (female by default).

Value

A list with:

- kin_full a data frame with year, cohort, Focal´s age, related ages and type of kin (for example d is daughter, oa is older aunts, etc.), including living and dead kin at that age.
- kin_summary a data frame with Focal´s age, related ages and type of kin, with indicators obtained processing kin_full, grouping by cohort or period (depending on the given arguments):
  - count_living: count of living kin at actual age of Focal
  - mean_age: mean age of each type of living kin.
  - sd_age: standard deviation of age of each type of living kin.
  - count_death: count of dead kin at specific age of Focal.
  - count_cum_death: cumulated count of dead kin until specific age of Focal.
  - mean_age_lost: mean age where Focal lost her relative.

Examples

```
# Kin expected matrilineal count for a Swedish female based on 2015 rates.
swe_surv_2015 <- swe_px[,"2015"]
swe_asfr_2015 <- swe_asfr[,"2015"]
# Run kinship models
swe_2015 <- kin(p = swe_surv_2015, f = swe_asfr_2015)
head(swe_2015$kin_summary)
```

Description

Implementation of two-sex matrix kinship model. This produces kin counts grouped by kin, age and sex of each relatives at each Focal´s age. For example, male cousins from aunts and uncles from different sibling's parents are grouped in one male count of cousins. Note that the output labels relative following female notation: the label m refers to either mothers or fathers, and column sex_kin determine the sex of the relative.

Usage

```
kin2sex(
  pf = NULL,
  pm = NULL,
  ff = NULL,
  fm = NULL,
  time_invariant = TRUE,
  sex_focal = "f",
  birth_female = 1/2.04,
  pif = NULL,
  pim = NULL,
  nf = NULL,
  nm = NULL,
  output_cohort = NULL,
  output_period = NULL,
  output_kin = NULL
)
```

Arguments

pf   numeric. A vector (atomic) or matrix with female probabilities (or survival ratios, or transition between age class in a more general perspective) with rows as ages (and columns as years in case of matrix, being the name of each col the year).

pm   numeric. A vector (atomic) or matrix with male probabilities (or survival ratios, or transition between age class in a more general perspective) with rows as ages (and columns as years in case of matrix, being the name of each col the year).

ff   numeric. Same as pf but for fertility rates.

fm   numeric. Same as pm but for fertility rates.

time_invariant logical. Constant assumption for a given year rates. Default TRUE.

sex_focal   character. "f" for female or "m" for male.

birth_female   numeric. Female portion at birth. This multiplies f argument. If f is already for female offspring, this needs to be set as 1.

pif	numeric. For using some specific age distribution of childbearing for mothers (same length as ages). Default NULL.
pim	numeric. For using some specific age distribution of childbearing for fathers (same length as ages). Default NULL.
nf	numeric. Only for time_invariant = FALSE. Same as pf but for population distribution (counts or %). Optional.
nm	numeric. Only for time_invariant = FALSE. Same as pm but for population distribution (counts or %). Optional.
output_cohort	integer. Vector of year cohorts for returning results. Should be within input data years range.
output_period	integer. Vector of period years for returning results. Should be within input data years range.
output_kin	character. kin types to return: "m" for mother, "d" for daughter,...

Details

See Caswell (2022) for details on formulas.

Value

A list with:

- kin_full a data frame with year, cohort, Focal´s age, related ages and type of kin (for example d could be daughter or son depending sex_kin, oa is older aunts or uncles also depending sex_kin value, etc.), including living and dead kin at that age.

- kin_summary a data frame with Focal´s age, related ages, sex and type of kin, with indicators obtained processing kin_full, grouping by cohort or period (depending on the given arguments):

  - count_living: count of living kin at actual age of Focal
  - mean_age: mean age of each type of living kin.
  - sd_age: standard deviation of age of each type of living kin.
  - count_death: count of dead kin at specific age of Focal.
  - count_cum_death: cumulated count of dead kin until specific age of Focal.
  - mean_age_lost: mean age where Focal lost her relative.

Examples

```
# Kin expected count by relative sex for a French female based on 2012 rates.
fra_fert_f <- fra_asfr_sex[,"ff"]
fra_fert_m <- fra_asfr_sex[,"fm"]
fra_surv_f <- fra_surv_sex[,"pf"]
fra_surv_m <- fra_surv_sex[,"pm"]
fra_2012 <- kin2sex(fra_surv_f, fra_surv_m, fra_fert_f, fra_fert_m)
head(fra_2012$kin_summary)
```

---

kin_multi_stage   *Estimate kin counts by age and stage in a time invariant framework*

Description

Implementation of age-stage kin estimates (multi-state) by Caswell (2020). Stages are implied in length of input lists.

Usage

```
kin_multi_stage(
  U = NULL,
  f = NULL,
  D = NULL,
  H = NULL,
  birth_female = 1/2.04,
  output_kin = NULL,
  parity = FALSE,
  list_output = FALSE
)
```

Arguments

U	list. age elements with column-stochastic transition matrix with dimension for the state space, conditional on survival.
f	matrix. state-specific fertility (age in rows and states in columns). Is accepted also a list with for each age-class.
D	matrix. survival probabilities by state (age in rows and states in columns). Is accepted also a list for each state with survival matrices.
H	matrix. assigns the offspring of individuals in some stage to the appropriate age class (age in rows and states in columns). Is accepted also a list with a matrix for each state.
birth_female	numeric. Female portion at birth.
output_kin	character. kin to return. For example "m" for mother, "d" for daughter. See the `vignette` for all kin types.
parity	logical. parity states imply age distribution of mothers re-scaled to not have parity 0 when Focal born. Default `TRUE`.
list_output	logical. Results as a list. Default `FALSE`.

Value

A data frame with focal´s age, related ages and type of kin (for example d is daughter, oa is older aunts, etc.), living and death kin counts, and specific stage. If `list_output = TRUE` then this is a list with elements as kin types.

```
kin_svk1990_caswell2020
```
                                Output for Slovakia 1990 in Caswell (2020).

Description

Output for Slovakia 1990 in Caswell (2020).

Usage

```
kin_svk1990_caswell2020
```

Format

A list with specific kin types age-stage matrix

Source

Caswell (2021)

```
kin_time_invariant          Estimate kin counts in a time invariant framework for one-sex model
                            (matrilineal/patrilineal)
```

Description

Mtrix implementation of Goodman-Keyfitz-Pullum equations adapted by Caswell (2019).

Usage

```
kin_time_invariant(
  p = NULL,
  f = NULL,
  birth_female = 1/2.04,
  pi = NULL,
  output_kin = NULL,
  list_output = FALSE
)
```

Arguments

p	numeric. A vector of survival probabilities with same length as ages.
f	numeric. A vector of age-specific fertility rates with same length as ages.
birth_female	numeric. Female portion at birth.
pi	numeric. For using some specific non-stable age distribution of childbearing (same length as ages). Default NULL.
output_kin	character. kin to return. For example "m" for mother, "d" for daughter. See vignette for all kin types.
list_output	logical. Results as a list with output_kin elements, with focal´s age in columns and kin ages in rows (2 * ages, last chunk of ages for death experience). Default FALSE

Value

A data frame with focal´s age, related ages and type of kin (for example d is daughter, oa is older aunts, etc.), alive and death. If list_output = TRUE then this is a list.

---

kin_time_invariant_2sex

Estimate kin counts in a time invariant framework for two-sex model.

Description

Two-sex matrix framework for kin count estimates.This produces kin counts grouped by kin, age and sex of each relatives at each Focal´s age. For example, male cousins from aunts and uncles from different sibling's parents are grouped in one male count of cousins.

Usage

```
kin_time_invariant_2sex(
  pf = NULL,
  pm = NULL,
  ff = NULL,
  fm = NULL,
  sex_focal = "f",
  birth_female = 1/2.04,
  pif = NULL,
  pim = NULL,
  output_kin = NULL,
  list_output = FALSE
)
```

Arguments

pf	numeric. A vector of survival probabilities for females with same length as ages.
pm	numeric. A vector of survival probabilities for males with same length as ages.
ff	numeric. A vector of age-specific fertility rates for females with same length as ages.
fm	numeric. A vector of age-specific fertility rates for males with same length as ages.
sex_focal	character. "f" for female or "m" for male.
birth_female	numeric. Female portion at birth.
pif	numeric. For using some specific non-stable age distribution of childbearing for mothers (same length as ages). Default NULL.
pim	numeric. For using some specific non-stable age distribution of childbearing for fathers (same length as ages). Default NULL.
output_kin	character. kin to return, considering matrilineal names. For example "m" for parents, "d" for children, etc. See the vignette for all kin types.
list_output	logical. Results as a list with output_kin elements, with focal´s age in columns and kin ages in rows (2 * ages, last chunk of ages for death experience). Default FALSE

Details

See Caswell (2022) for details on formulas.

Value

A data frame with focal´s age, related ages and type of kin (for example d is children, oa is older aunts/uncles, etc.), sex, alive and death. If list_output = TRUE then this is a list.

---

kin_time_variant	*Estimate kin counts in a time variant framework (dynamic rates) for one-sex model (matrilineal/patrilineal)*

Description

Matrix implementation of time variant Goodman-Keyfitz-Pullum equations in a matrix framework.

Usage

```
kin_time_variant(
  p = NULL,
  f = NULL,
  pi = NULL,
  n = NULL,
  output_cohort = NULL,
```

```
    output_period = NULL,
    output_kin = NULL,
    birth_female = 1/2.04,
    list_output = FALSE
)
```

Arguments

p	numeric. A matrix of survival ratios with rows as ages and columns as years. Column names must be equal interval.
f	numeric. A matrix of age-specific fertility rates with rows as ages and columns as years. Coincident with U.
pi	numeric. A matrix with distribution of childbearing with rows as ages and columns as years. Coincident with U.
n	numeric. A matrix of population with rows as ages and columns as years. Coincident with U.
output_cohort	integer. Year of birth of focal to return as output. Could be a vector. Should be within input data years range.
output_period	integer. Year for which to return kinship structure. Could be a vector. Should be within input data years range.
output_kin	character. kin to return as output: "m" for mother, "d" for daughter,... See vignette for exahustive kin.
birth_female	numeric. Female portion at birth.
list_output	logical. Results as a list with years elements (as a result of output_cohort and output_period combination), with a second list of output_kin elements, with focal´s age in columns and kin ages in rows (2 * ages, last chunk of ages for death experience). Default FALSE

Details

See Caswell (2021) for details on formulas.

Value

A data frame of population kinship structure, with Focal's cohort, focal´s age, period year, type of relatives (for example d is daughter, oa is older aunts, etc.), living and death kin counts, and age of (living or time deceased) relatives. If list_output = TRUE then this is a list.

---

kin_time_variant_2sex   *Estimate kin counts in a time variant framework (dynamic rates) in a two-sex framework (Caswell, 2022)*

Description

Two-sex matrix framework for kin count estimates with varying rates. This produces kin counts grouped by kin, age and sex of each relatives at each Focal´s age. For example, male cousins from aunts and uncles from different sibling's parents are grouped in one male count of cousins.

Usage

```
kin_time_variant_2sex(
  pf = NULL,
  pm = NULL,
  ff = NULL,
  fm = NULL,
  sex_focal = "f",
  birth_female = 1/2.04,
  pif = NULL,
  pim = NULL,
  nf = NULL,
  nm = NULL,
  output_cohort = NULL,
  output_period = NULL,
  output_kin = NULL,
  list_output = FALSE
)
```

Arguments

pf	numeric. A vector (atomic) or matrix with probabilities (or survival ratios, or transition between age class in a more general perspective) with rows as ages (and columns as years in case of matrix, being the name of each col the year).
pm	numeric. A vector (atomic) or matrix with probabilities (or survival ratios, or transition between age class in a more general perspective) with rows as ages (and columns as years in case of matrix, being the name of each col the year).
ff	numeric. Same as pf but for fertility rates.
fm	numeric. Same as pm but for fertility rates.
sex_focal	character. "f" for female or "m" for male.
birth_female	numeric. Female portion at birth. This multiplies f argument. If f is already for female offspring, this needs to be set as 1.
pif	numeric. For using some specific age distribution of childbearing for mothers (same length as ages). Default NULL.
pim	numeric. For using some specific age distribution of childbearing for fathers (same length as ages). Default NULL.
nf	numeric. Same as pf but for population distribution (counts or %). Optional.
nm	numeric. Same as pm but for population distribution (counts or %). Optional.
output_cohort	integer. Vector of year cohorts for returning results. Should be within input data years range.

output_period	integer. Vector of period years for returning results. Should be within input data years range.
output_kin	character. kin types to return: "m" for mother, "d" for daughter,...
list_output	logical. Results as a list with years elements (as a result of output_cohort and output_period combination), with a second list of output_kin elements, with focal´s age in columns and kin ages in rows (2 * ages, last chunk of ages for death experience). Default FALSE

Details

See Caswell (2022) for details on formulas.

Value

A data.frame with year, cohort, Focal´s age, related ages, sex and type of kin (for example d is daughter, oa is older aunts, etc.), including living and dead kin at that age and sex.

---

output_period_cohort_combination

APC combination to return

Description

define APC combination to return in kin and kin2sex.

Usage

```
output_period_cohort_combination(
  output_cohort = NULL,
  output_period = NULL,
  age = NULL,
  years_data = NULL
)
```

Arguments

output_cohort	integer. A vector with selected calendar years.
output_period	integer. A vector with selected cohort years.
age	integer. A vector with ages from the kinship network to be filtered.
years_data	integer. A vector with years from the time-varying kinship network to be filtered.

Details

Because returning all period and cohort data from a huge time-series would be hard memory consuming, this function is an auxiliary one to deal with selection from inputs output_cohort and output_period.

Value

data.frame with years and ages to filter in `kin` and `kin_2sex` functions.

`pi_caswell_2021`	*Historic and projected mother´s age distribution of childbearing from Sweden used in Caswell (2021)*

Description

Historic and projected mother´s age distribution of childbearing from Sweden used in Caswell (2021)

Usage

```
pi_caswell_2021
```

Format

A matrix pi with years as cols and ages as rows.

Source

Caswell (2019)

`plot_diagram`	*plot a Kin diagram (network)*

Description

Draws a Keyfitz-style kinship diagram given a kinship object created by the `kin` function. Displays expected kin counts for a Focal aged 'a'.

Usage

```
plot_diagram(kin_total, rounding = 3)
```

Arguments

`kin_total`	data.frame. values in column `kin` define the relative type - see `demokin_codes()`. Values in column `count` are the expected number of relatives.
`rounding`	numeric. Number of decimals to show in diagram.

Value

A Keyfitz-style kinship plot.

---

rename_kin   *rename kin*

Description

Add kin labels depending the sex

Usage

```
rename_kin(df, sex = "f")
```

Arguments

df   data.frame. A data frame with variable `kin` with `DemoKin` codes to be labelled.

sex   character. "f" for female, "m" for male or "2sex" for both sex naming.

Details

See table `demokin_codes` to know label options.

Value

Add a column with kin labels in the input data frame.

---

svk_fxs   *Female Slovakian fertility rates by parity stage in 1990 (Caswell, 2021)*

Description

Female Slovakian fertility rates by parity stage in 1990 (Caswell, 2021)

Usage

```
svk_fxs
```

Format

A matrix of fx with stages as cols and ages as rows.

Source

Caswell (2021)

---

svk_Hxs	*Age where assign offspring of individuals in each partity stage (Caswell, 2021). All to zero age in this case.*

Description

Age where assign offspring of individuals in each partity stage (Caswell, 2021). All to zero age in this case.

Usage

    svk_Hxs

Format

A matrix of ones in ages where assign offspring individuals, with stages as cols and ages as rows.

Source

Caswell (2021)

---

svk_pxs	*Female Slovakian survival probabilities by parity stage in 1990 (Caswell, 2021)*

Description

Female Slovakian survival probabilities by parity stage in 1990 (Caswell, 2021)

Usage

    svk_pxs

Format

A matrix of px with stages as cols and ages as rows.

Source

Caswell (2021)

svk_Uxs	*Probability of transition among parity stage for Slovakia in 1990, for each age, conditional on survival (Caswell, 2021).*

Description

Probability of transition among parity stage for Slovakia in 1990, for each age, conditional on survival (Caswell, 2021).

Usage

```
svk_Uxs
```

Format

A list of column-stochastic matrix with probabilities of transition among parity stage, for each age, conditional on survival.

Source

Caswell (2021)

swe_asfr	*Swedish age-specific fertility rates from 1900 to 2015*

Description

Swedish age-specific fertility rates from 1900 to 2015

Usage

```
swe_asfr
```

Format

A matrix with years as cols and ages (0 to 100 as OAG) as rows.

Source

HMD/HFD

---

swe_pop *Female swedish population from 1900 to 2015*

Description

Female swedish population from 1900 to 2015

Usage

```
swe_pop
```

Format

A matrix with years as cols and ages (0 to 100 as OAG) as rows.

Source

HMD/HFD

---

swe_px *Female swedish survival probabilities from 1900 to 2015*

Description

Female swedish survival probabilities from 1900 to 2015

Usage

```
swe_px
```

Format

A matrix with years as cols and ages (0 to 100 as OAG) as rows.

Source

HMD/HFD

---

swe_Sx   *Female swedish survival ratios from 1900 to 2015*

Description

Female swedish survival ratios from 1900 to 2015

Usage

    swe_Sx

Format

A matrix with years as cols and ages (0 to 100 as OAG) as rows.

Source

HMD/HFD

---

 timevarying_kin   *one time projection kin*

Description

one time projection kin. internal function.

Usage

    timevarying_kin(Ut, ft, pit, ages, pkin)

Arguments

Ut	numeric. A matrix of survival probabilities (or ratios).
ft	numeric. A matrix of age-specific fertility rates.
pit	numeric. A matrix with distribution of childbearing.
ages	numeric.
pkin	numeric. A list with kin count distribution in previous year.

Value

A list of 14 types of kin matrices (kin age by Focal age) projected one time interval.

`timevarying_kin_2sex` *one time projection kin*

Description

one time projection kin. internal function.

Usage

```
timevarying_kin_2sex(Ut, Ft, Ft_star, pit, sex_focal, ages, pkin)
```

Arguments

Ut	numeric. A matrix of survival probabilities (or ratios).
Ft	numeric. A matrix of age-specific fertility rates.
Ft_star	numeric. Ft but for female fertility.
pit	numeric. A matrix with distribution of childbearing.
sex_focal	character. "f" for female or "m" for male.
ages	numeric.
pkin	numeric. A list with kin count distribution in previous year.

Value

A list of 14 types of kin matrices (kin age by Focal age, blocked for two sex) projected one time interval.

`U_caswell_2021` *Historic and projected survival ratios from Sweden used in Caswell (2021)*

Description

Historic and projected survival ratios from Sweden used in Caswell (2021)

Usage

```
U_caswell_2021
```

Format

A matrix U with years as cols and ages as rows.

Source

Caswell (2019)

Index