Homework 02

Ihsan Kahveci

2020-04-18

Questions

Q1

Table 1: Q1 one-sex closed population

Age	Population (N_x)	Fertility Rate ($\tilde{F}_x$)	Survival Prob. (s_x)
1	18000	0.0	0.65
2	17000	0.9	0.75
3+	14000	0.2	0.15

Q1.a

The crude birth rate (CBR) is defined as the number of births over the person-years lived in the period $[T_1, T_2]$. Since our period is a single year, we can calculate CBR as:

$$CBR = \sum \frac{N_x \tilde{F}_x}{N_x s_x}$$

where we sum over all age groups. The crude birth rate for this population in the next time period is then **0.682**.

Q1.b

The total fertility rate in the population in the period $[T_1, T_2]$ is defined as the sum of the age-specific fertility rates across all age groups, multiplied by the length of the age interval, n. With $T_2 - T_1 = n = 1$, the total fertility rate represents the single-year cohort total fertility rate:

$$\text{TFR}[T_1, T_1 + 1] = \sum {}_1F_x[T_1, T_1 + 1]$$

We can convert between $\tilde{F}_x$ and ${}_1F_x$ using the equation

$$\tilde{F}_x = {}_1F_x \times \frac{1}{1 + SRB} \times \frac{1}{2}\left(1 + s_{x-1}\frac{N_{x-1,t}}{N_{x,t}}\right) \times \left(1 - \frac{q_0}{2}\right)$$

where we assume $SRB = 1.05$ and take $q_0 = 1 - s_0$. After converting to age-specific fertility rates, we calculate a total fertility rate of $0 + 2.649 + 0.52 =$ **3.17** for this population.

Q1.c

The Leslie matrix, L, for this population is defined as:

$$L = \begin{bmatrix} \tilde{F}_{A-3} & \tilde{F}_{A-2} & \tilde{F}_{A-1} \\ s_{A-3} & 0 & 0 \\ 0 & s_{A-2} & s_{A-1} \end{bmatrix} = \begin{bmatrix} 0 & 0.9 & 0.2 \\ 0.65 & 0 & 0 \\ 0 & 0.75 & 0.15 \end{bmatrix}$$

Where $(A-1)+$ is the highest age group that can be reached in this population, 3+, s_x denotes the probability of survival to the next age group for age group x, and $\tilde{F}_x$ is the expected number of female births to a woman age x, who survives to the next time interval.

Q1.d

We can project this population forward using the *cohort-component method of population projection*, which states that the age-specific populations one time period ahead (N_{t+1}) can be calculated from the matrix multiplication of the age-specific population in the current period (N_t) and the Leslie matrix (L) of the population. The population by age one period forward from our given initial population is then:

$$\begin{aligned} N_{t+1} &= LN_t \\ &= \begin{bmatrix} 0 & 0.9 & 0.2 \\ 0.65 & 0 & 0 \\ 0 & 0.75 & 0.15 \end{bmatrix} \begin{bmatrix} 18000 \\ 17000 \\ 14000 \end{bmatrix} \\ &= \begin{bmatrix} 18100 \\ 11700 \\ 14850 \end{bmatrix} \end{aligned}$$

Q1.e

This method can be extended to projecting age-specific population k periods ahead by raising the Leslie matrix to the k^{th} power (L^k). Our given population, projected 2 periods into the future is then:

$$\begin{aligned} N_{t+2} &= L^{10}N_t \\ &= \begin{bmatrix} 0 & 0.9 & 0.2 \\ 0.65 & 0 & 0 \\ 0 & 0.75 & 0.15 \end{bmatrix}^2 \begin{bmatrix} 18000 \\ 17000 \\ 14000 \end{bmatrix} \\ &= \begin{bmatrix} 0.585 & 0.15 & 0.03 \\ 0 & 0.585 & 0.13 \\ 0.488 & 0.112 & 0.022 \end{bmatrix} \begin{bmatrix} 18000 \\ 17000 \\ 14000 \end{bmatrix} \\ &= \begin{bmatrix} 13500 \\ 11765 \\ 11002.5 \end{bmatrix} \end{aligned}$$

Q1.f

The crude birth rate for this population from time period 1 to time period 2 is **0.593**.

The total fertility rate between time periods 1&2 is \$0 + 2.23 + 0.625 = \$ **2.855** for this population.

Q1.g

From the theorem that N_t converges to $\lambda^t u$ as t approaches infinity, $log(\lambda)$ is the *instantaneous rate of increase of the population.* Here, λ is defined as the dominant **right eigenvalue** of the Leslie matrix, or for the equation:

$$Lv = \lambda v$$

it is the eigenvalue λ with the largest magnitude. For our calculated Leslie matrix, the instantaneous rate of increase is **-0.161.**

Q1.h

Again, from the the formula $\lambda^t u$, u is the *stable age distribution*, and is defined as the dominant **right eigenvector** of the Leslie matrix, which is the column vector v from the eigendecomposition of L corresponding to the eigenvalue λ with the largest magnitude.

For our calculated Leslie matrix, the stable age distribution is $\begin{bmatrix} 0.666 \\ 0.509 \\ 0.545 \end{bmatrix}$.

Q1.i

The reproductive value vector (v) is a vector of expected the number of future offspring of an individual for each age group. A theorem states that v is the dominant **left eigenvector** of the Leslie matrix for the population. The left dominant eigenvector of a matrix A is equivalent to the right dominant eigenvector of the transpose of matrix, $A^\top$. So, in the formula:

$$L^\top u = \kappa u$$

the dominant eigenvector u represents the reproductive values. For our Leslie matrix, the reproductive value matrix is then $\begin{bmatrix} -0.598 \\ -0.783 \\ -0.171 \end{bmatrix}$.

Q2

Q2.a

Table 2: Age-specific Mortality Rates of Peru female population, 2015-2020

Age	$_nq_x$
0	0.012
1	0.001
5	0.000
10	0.000
15	0.000
20	0.001
25	0.001
30	0.001
35	0.002
40	0.002
45	0.003
50	0.004
55	0.005
60	0.008
65	0.013
70	0.022
75	0.036
80	0.062
85	0.108
90	0.172
95	0.251
100	0.367

Q2.b

First, we need to calculate a life table from mortality rates.

Table 3: Life table for 2015-2020 Peru female population

Age	a_x	$_5m_x$	$_1q_x$	$_1s_x$	$_5d_x$	l_x	$_5L_x$	$_5T_x$	e_x
0	0.1	0.012	0.012	0.989	1150	100000	98948	7907710	79
1-4	1.5	0.001	0.003	0.997	315	98850	394614	7808761	79
5-9	2.5	0.000	0.002	0.998	197	98535	492182	7414148	75
10-14	2.5	0.000	0.002	0.999	148	98338	491320	6921965	70
15-19	2.5	0.000	0.002	0.998	220	98190	490400	6430645	65
20-24	2.5	0.001	0.003	0.997	323	97970	489042	5940245	61
25-29	2.5	0.001	0.004	0.996	443	97647	487128	5451202	56
30-34	2.5	0.001	0.006	0.994	571	97204	484592	4964075	51
35-39	2.5	0.002	0.008	0.992	732	96633	481335	4479482	46
40-44	2.5	0.002	0.010	0.991	912	95901	477225	3998148	42
45-49	2.5	0.003	0.013	0.987	1245	94989	471832	3520922	37
50-54	2.5	0.004	0.018	0.982	1650	93744	464595	3049090	33
55-59	2.5	0.005	0.025	0.975	2350	92094	454595	2584495	28

Age	a_x	${}_5m_x$	${}_1q_x$	${}_1s_x$	${}_5d_x$	l_x	${}_5L_x$	${}_5T_x$	e_x
60-64	2.5	0.008	0.040	0.960	3614	89744	439685	2129900	24
65-69	2.5	0.013	0.061	0.939	5242	86130	417545	1690215	20
70-74	2.5	0.022	0.102	0.897	8291	80888	383712	1272670	16
75-79	2.5	0.036	0.166	0.835	12013	72597	332952	888958	12
80-84	2.5	0.062	0.270	0.730	16366	60584	262005	556005	9
85-89	2.5	0.108	0.426	0.574	18822	44218	174035	294000	7
90-94	2.5	0.172	0.601	0.399	15259	25396	88832	119965	5
95-99	2.5	0.251	0.772	0.228	7821	10137	31132	31132	3
100+	-	0.367	1.000	0.000	2316	2316	-	-	-

Table 4: Peru ${}_ns_x$: 0 and 1-4 combined

Age	${}_nq_x$	${}_ns_x$
0-4	0.01	0.99
5-9	0.00	1.00
10-14	0.00	1.00
15-19	0.00	1.00
20-24	0.00	1.00
25-29	0.00	1.00
30-34	0.01	0.99
35-39	0.01	0.99
40-44	0.01	0.99
45-49	0.01	0.99
50-54	0.02	0.98
55-59	0.03	0.97
60-64	0.04	0.96
65-69	0.06	0.94
70-74	0.10	0.90
75-79	0.17	0.83
80-84	0.27	0.73
85-89	0.43	0.57
90-94	0.60	0.40
95-99	0.77	0.23
100+	1.00	0.00

Then, we will use, we will use ${}_ns_x$ and q_0 values to convert between $\tilde{F}_x$ and ${}_1F_x$ using the equation.

To calculate ${}_5\tilde{F}_x$, we first use the provided proportional age-specific fertility rate and total fertility rate for Peru in 2015 to get age-specific fertility rate with the formula:

$$\frac{TFR[T_1, T_2] \times {}_nPASFR_x}{n} = {}_nF_x[T_1, T_2]$$

where $n = 5$ and $[T_1, T_2] = [2015, 2020]$. Then we use the provided population and previously calculated mortality rates ${}_5s_x$ and ${}_5q_0$ to calculate ${}_5\tilde{F}_x$ using the formula:

$${}_n\tilde{F}_x = {}_nF_x \times \frac{1}{1+SRB} \times \frac{1}{2}\left(1 + {}_ns_x \frac{{}_5N_{x-1,t}}{{}_5N_{x,t}}\right) \times \left(1 - \frac{{}_nq_0}{2}\right)$$

Note that ${}_5q_0$ was calculated as $1 - {}_5s_0$, which was in turn calculated from ${}_1s_0 \times {}_4s_1$.

With these calculations, our resulting ${}_5\tilde{F}_x$ is:

Table 5: Expected number of live female births per woman per five-year period in Peru, 2015-2020

Age	${}_5\tilde{F}_x$	${}_5F_x$	${}_5N_x$	${}_5s_x$
0-4	0.000	0.000	1370	0.985
5-9	0.000	0.000	1454	0.998
10-14	0.000	0.000	1361	0.999
15-19	0.000	5.688	1338	0.998
20-24	4.876	9.997	1316	0.997
25-29	5.481	10.921	1223	0.996
30-34	4.787	9.568	1141	0.994
35-39	3.115	6.266	1077	0.992
40-44	1.265	2.536	1008	0.991
45-49	0.216	0.424	899	0.987
50-54	0.000	0.000	777	0.982
55-59	0.000	0.000	639	0.975
60-64	0.000	0.000	531	0.960
65-69	0.000	0.000	395	0.939
70-74	0.000	0.000	313	0.897
75-79	0.000	0.000	233	0.835
80-84	0.000	0.000	157	0.730
85-89	0.000	0.000	73	0.574
90-94	0.000	0.000	25	0.399
95-99	0.000	0.000	5	0.228
100+	0.000	0.000	1	0.000

Q2.c

Q2.d

Q2.e

Appendix

```r
# Prep work ---------------------------------------------------------------

# Load libraries
library(tidyverse, quietly = TRUE)
library(demogR)
library(wpp2019)

# Helper functions

write_matex <- function(x, digits = 3) {
  # From: https://stackoverflow.com/a/54088015/8866058
  x <- round(x, digits = digits)
  mat_string <- apply(x, 1, function(y) paste(y, collapse = "&"))
  paste("\\begin{bmatrix}", paste0(mat_string, collapse = "\\\\"), "\\end{bmatrix}")
```

```
}

"%^%" <- function(A, n) {
  if (n == 1) {
    A
  } else {
    A %*% (A %^% (n - 1))
  }
}

make_leslie_matrix <- function(f, s) {

  if (length(f) != length(s)) {
    stop("f and s must be the same length")
  }

  n_size <- length(f)
  l_mat <- matrix(0, nrow = n_size, ncol = n_size)

  l_mat[1, ] <- f
  diag(l_mat[-1, ]) <- s[1:(n_size - 1)]
  l_mat[n_size, n_size] <- s[n_size]

  l_mat
}


# Question 1 -------------------------------------------------------------

pop_table <- tibble(
  age = c("1", "2", "3+"),
  pop = c(18, 17, 14) * 1000,
  fr = c(0, .9, .2),
  surv = c(.65, .75, .15)
)

knitr::kable(
  pop_table,
  booktabs = TRUE,
  caption = "Q1 one-sex closed population",
  col.names = c(
    "Age",
    "Population ($N_x$)",
    "Fertility Rate ($\\tilde{F}_x$)",
    "Survival Prob. ($s_x$)"
  ),
  eval = FALSE
)


# Question 1a ------------------------------------------------------------

CBR <- pop_table %>%
```

```r
  mutate(births = pop * fr,
         person_years = pop * surv) %>%
  summarise(cbr = sum(births) / sum(person_years)) %>%
  pull(cbr)

# Question 1b ------------------------------------------------------------
## function for convertion fertility rate to expected number of female births
f_tilde_2_asfr <- function(F_tilde, srb, Sxm1, Nxm1, Nx, q0) {
  F_tilde * (1 + srb) * 2/(1 + Sxm1 * (Nxm1/Nx)) / (1 - q0/2)}

pop_asfr_1 <- 0
pop_asfr_2 <- f_tilde_2_asfr(0.9, 1.05, .65, 18000, 17000, 1 - 0.65)
pop_asfr_3 <- f_tilde_2_asfr(0.2, 1.05, .75, 17000, 14000, 1 - 0.65)

pop_asfr <- c(pop_asfr_1, pop_asfr_2, pop_asfr_3)

TFR <- sum(pop_asfr)
tfr_eqn <- paste0(round(pop_asfr, 3), collapse = " + ")


# Question 1c ------------------------------------------------------------

pop_leslie <- make_leslie_matrix(pop_table$fr, pop_table$surv)

# Question 1d ------------------------------------------------------------

pop_t0 <- matrix(pop_table$pop)
pop_t1 <- pop_leslie %*% pop_t0


# Question 1e ------------------------------------------------------------
pop_t2 <- (pop_leslie %^% 2) %*% pop_t0


# Question 1f ------------------------------------------------------------

CBR_t2 <- pop_table %>%
  mutate(
    pop = as.vector(pop_t1),
    births = pop * fr,
    person_years = pop * surv
  ) %>%
  summarise(cbr = sum(births) / sum(person_years)) %>%
  pull(cbr)

pop_asfr_t2_1 <- 0
pop_asfr_t2_2 <- f_tilde_2_asfr(0.9, 1.05, .65, pop_t1[1], pop_t1[2], 1 - 0.65)
pop_asfr_t2_3 <- f_tilde_2_asfr(0.2, 1.05, .75, pop_t1[2], pop_t1[3], 1 - 0.65)

pop_asfr_t2 <- c(pop_asfr_t2_1, pop_asfr_t2_2, pop_asfr_t2_3)

TFR_t2 <- sum(pop_asfr_t2)
tfr_eqn_t2 <- paste0(round(pop_asfr_t2, 3), collapse = " + ")
```

```r
# Question 1g ------------------------------------------------------------

pop_right_eigen <- eigen(pop_leslie)
dominant_right_index <- which.max(abs(pop_right_eigen$values))

pop_iroi <- log(pop_right_eigen$values[dominant_right_index])


# Question 1h ------------------------------------------------------------

pop_sad <- matrix(pop_right_eigen$vectors[, dominant_right_index])


# Question 1i ------------------------------------------------------------

pop_left_eigen <- eigen(t(pop_leslie))
dominant_left_index <- which.max(abs(pop_left_eigen$values))

pop_repv <- matrix(pop_left_eigen$vectors[, dominant_left_index])

# Question 2 -------------------------------------------------------------
## Question 2a -----------------------------------------------------------
data(mxF)
peru_mx <- mxF %>%
  filter(name == "Peru") %>%
  select(age, mx = `2015-2020`)

# Builds a life table by using the mortality rate
peru_LT =
  LifeTables::lt.mx(peru_mx$mx, age = peru_mx$age)$lt %>%
  as_tibble() %>%
  mutate(age =
           paste0(Age, "-", Age + 4) %>%
           magrittr::inset(c(1, 2, length(Age)), c("0", "1-4", "100+")), .before=1)

peru_mx %>%
  knitr::kable(
  booktabs = TRUE,
  escape = TRUE,
  digits = 3,
  col.names = c("Age", "$_nq_x$"),
  eval = FALSE,
  caption = "Age-specific Mortality Rates of Peru female population, 2015-2020")

# Question 2b ------------------------------------------------------------
LT_colnames <- c(
  "Age",
  "$a_x$",
  "$_{5}m_x$",
  "$_{1}q_x$",
  "$_{1}s_x$",
```

```
  "$_{5}d_x$",
  "$l_x$",
  "${}_{5}L_x$",
  "${}_{5}T_x$",
  "$e_x$"
)

peru_LT %>%
  select(-Age)  %>%
  knitr::kable(
    booktabs = TRUE,
    col.names = LT_colnames,
    eval = FALSE,
    digits = c(0, 1, 3, 3, 3, 0, 0, 2, 0, 0, 2),
    caption = "Life table for 2015-2020 Peru female population")


# Create standard 5-year age goup npx and nqx mortality
peru_0_to_5 =peru_LT %>%
  filter(Age < 5) %>%
  summarise(age = "0-4", nqx = 1 - prod(npx), npx = prod(npx))

peru_q0 <- peru_0_to_5$nqx[1]

peru_nqx <- peru_LT %>%
  filter(Age >= 5) %>%
  select(age, nqx, npx) %>%
  bind_rows(peru_0_to_5, .)


peru_nqx%>%
  knitr::kable(
    booktabs = TRUE,
    col.names = c("Age", "$_nq_x$", "$_ns_x$"),
    eval = FALSE,
    digits = c(0, 2, 2),
    caption = "Peru $_ns_x$: 0 and 1-4 combined")
data(popF)
data(sexRatio)
data(tfr)
data(percentASFR)

peru_pop <- popF %>%
  filter(name == "Peru") %>%
  select(age, pop = `2015`) %>%
  mutate(order = row_number())

peru_srb <- sexRatio %>% filter(name == "Peru") %>% pull(`2015-2020`)
peru_tfr <- tfr %>% filter(name == "Peru") %>% pull(`2015-2020`)

asfr_2_f_tilde <- function(asfr, srb, Sxm1, Nxm1, Nx, q0) {
  asfr * (1 / (1 + srb)) * .5 * (1 + Sxm1 * (Nxm1 / Nx)) * (1 - q0 / 2)}
```

```r
peru_Fx <- percentASFR %>%
  filter(name == "Peru") %>%
  select(age=age, pasfr = `2015-2020`) %>%
  mutate(asfr = pasfr * peru_tfr / 5) %>%
  right_join(peru_pop, by = "age", ) %>%
  replace_na(list(pasfr = 0, asfr = 0)) %>%
  left_join(peru_nqx, by = "age") %>%
  mutate(
    f_tilde =
      asfr_2_f_tilde(asfr, peru_srb, lag(npx), lag(pop), pop, peru_q0)
  ) %>%
  replace_na(list(f_tilde = 0)) %>%
  arrange(order) %>%
  select(-order)

knitr::kable(
  select(peru_Fx, age, f_tilde, asfr, pop, npx),
  booktabs = TRUE,
  col.names =
    c("Age", "${}_{5}\\tilde{F}_x$", "${}_{5}F_x$", "${}_{5}N_x$", "${}_{5}s_x$"),
  eval = FALSE,
  digits = c(0, 3, 3, 0, 3),
  caption = paste(
    "Expected number of live female births per woman per five-year period",
    "in Peru, 2015-2020"
  )
)
```