Homework 02

Spencer Pease

4/13/2020

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 10 periods into the future is then:

$$\begin{aligned} N_{t+10} &= L^{10}N_t \\ &= \begin{bmatrix} 0 & 0.9 & 0.2 \\ 0.65 & 0 & 0 \\ 0 & 0.75 & 0.15 \end{bmatrix}^{10} \begin{bmatrix} 18000 \\ 17000 \\ 14000 \end{bmatrix} \\ &= \begin{bmatrix} 0.102 & 0.104 & 0.022 \\ 0.056 & 0.102 & 0.022 \\ 0.084 & 0.084 & 0.018 \end{bmatrix} \begin{bmatrix} 18000 \\ 17000 \\ 14000 \end{bmatrix} \\ &= \begin{bmatrix} 3903.625 \\ 3051.594 \\ 3189.032 \end{bmatrix} \end{aligned}$$

Q1.f

The crude birth rate for this population 10 time periods ahead is **0.638**.

The total fertility rate 10 time periods ahead, after converting $\tilde{F}_x$ to ${}_1F_x$, is **3.021** 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

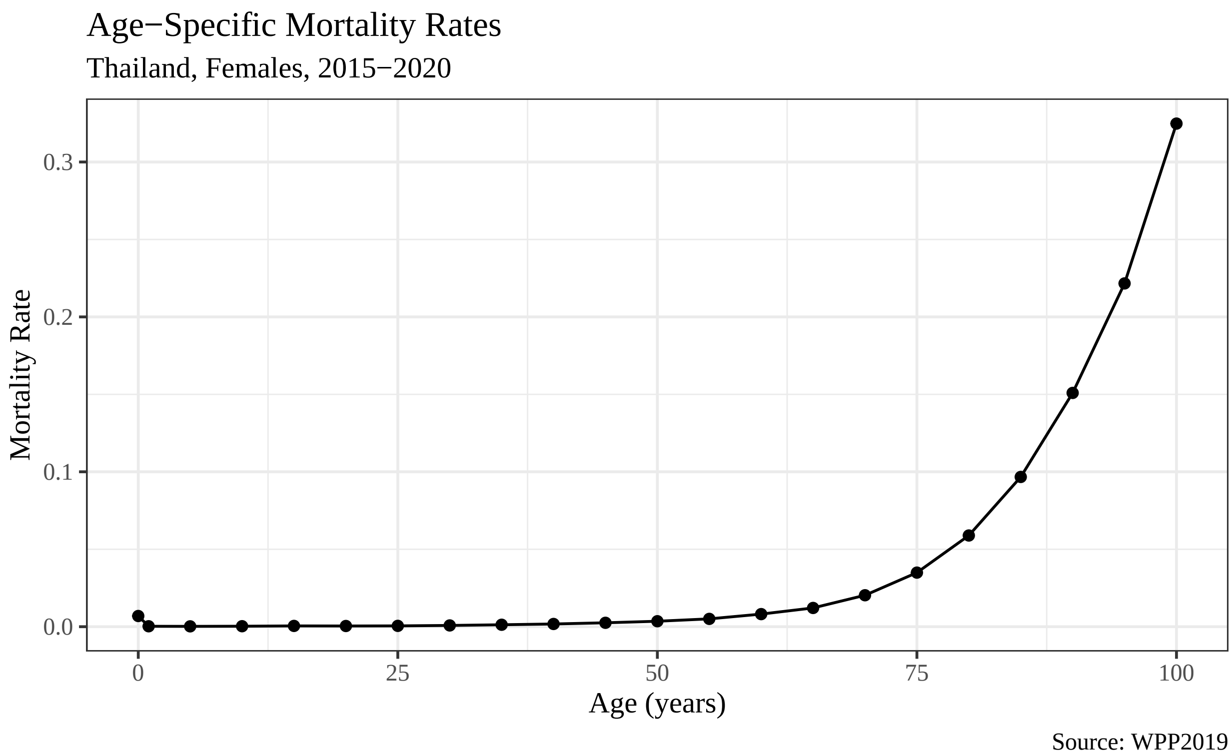
The reproductive value vector (v) is a vector of expected the number of future offspring of an individual for each age group. A theorem states that v is the dominant left eigenvector of the Leslie matrix for the population. The left dominant eigenvector of a matrix A is equivalent to the right dominant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

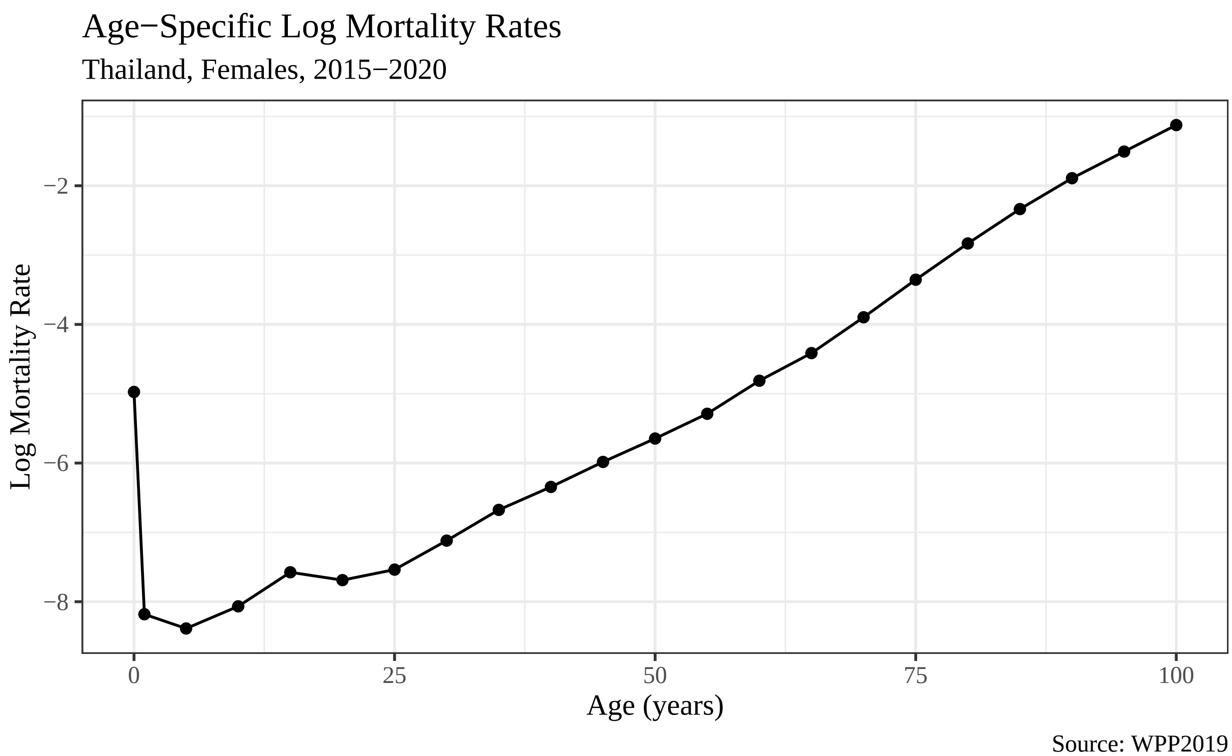
$$L^\top u = \kappa u$$

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

Q2

Q2.a





One unusual feature of these mortality rates is the the bump in log-mortality around the 15-20 year group.

This is a common feature in male populations (called the accident hump), but is not as common in female populations. Another feature to note is the incredibly high infant mortality rate, which is at the same level as older adults.

Q2.b

To derive a life table from the given ${}_nM_x$ values, we use the following formulas:

$$
\begin{aligned}
{}_nq_x &= 1 - e^{-n \cdot {}_nM_x} \\
{}_ns_x &= 1 - {}_nq_x \\
l_{x+n} &= (1 - {}_nq_x) \cdot l_x \\
{}_nd_x &= l_x - l_{x+n} \\
{}_nL_x &= l_{x+n} + a_x \cdot {}_nd_x \\
T_x &= \sum_{age=x}^{\infty} {}_nL_x \\
e_x &= T_x \div L_x
\end{aligned}
$$

Where we are using the approximation ${}_nq_x = 1 - e^{-n \cdot {}_nM_x}$ to to get ${}_nq_x$ from ${}_nM_x$, and a_x is caclulated as $\frac{n}{2}$ except for the 0-1 age group, where it is $0.07 + 1.7{}_1M_0$.

Table 2: Life table for 2015-2020 Thailand female population

Age	n	${}_5m_x$	${}_1q_x$	${}_1s_x$	l_x	${}_5d_x$	a_x	${}_5L_x$	${}_5T_x$	e_x
0	1	0.007	0.007	0.993	100000	689	0.08	496612	8562458	85.62
1-4	4	0.000	0.001	0.999	99311	111	2.00	496222	8065846	81.22
5-9	5	0.000	0.001	0.999	99200	113	2.50	495717	7569624	76.31
10-14	5	0.000	0.002	0.998	99087	155	2.50	495046	7073907	71.39
15-19	5	0.001	0.003	0.997	98931	253	2.50	494024	6578862	66.50
20-24	5	0.000	0.002	0.998	98678	226	2.50	492826	6084838	61.66
25-29	5	0.001	0.003	0.997	98452	262	2.50	491606	5592012	56.80
30-34	5	0.001	0.004	0.996	98190	397	2.50	489959	5100406	51.94
35-39	5	0.001	0.006	0.994	97793	615	2.50	487429	4610446	47.14
40-44	5	0.002	0.009	0.991	97178	850	2.50	483766	4123017	42.43
45-49	5	0.003	0.013	0.987	96328	1206	2.50	478626	3639251	37.78
50-54	5	0.004	0.018	0.982	95122	1665	2.50	471449	3160625	33.23
55-59	5	0.005	0.025	0.975	93458	2329	2.50	461464	2689175	28.77
60-64	5	0.008	0.040	0.960	91128	3629	2.50	446568	2227711	24.45
65-69	5	0.012	0.059	0.941	87499	5140	2.50	424646	1781143	20.36
70-74	5	0.020	0.097	0.903	82359	7950	2.50	391921	1356497	16.47
75-79	5	0.035	0.160	0.840	74409	11918	2.50	342251	964576	12.96
80-84	5	0.059	0.255	0.745	62491	15938	2.50	272612	622325	9.96
85-89	5	0.097	0.383	0.617	46553	17844	2.50	188157	349713	7.51
90-94	5	0.151	0.530	0.470	28709	15210	2.50	105522	161556	5.63
95-99	5	0.222	0.670	0.330	13499	9042	2.50	44892	56035	4.15
100+	5	0.325	1.000	0.000	4457	4457	2.50	11143	11143	2.50

Q2.c

The life expectancy at birth is **85.625**, and the life expectancy at age 10 is **71.391**

Q3

Q3.a

Q3.b

Q3.c

Q3.d

Q3.e

Q3.f

Q3.g

Appendix

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

# Load libraries
library(dplyr)
library(tidyr)
library(magrittr)
library(ggplot2)
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)
```

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


# Question 2b -------------------------------------------------------------

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(.9, 1.05, .65, 18000, 17000, 1-.65)
pop_asfr_3 <- f_tilde_2_asfr(.2, 1.05, .75, 17000, 14000, 1-.65)

pop_asfr <- c(pop_asfr_1, pop_asfr_2, pop_asfr_3)
```

```r
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_t10 <- (pop_leslie %^% 10) %*% pop_t0


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

CBR_t10 <- pop_table %>%
  mutate(
    pop = as.vector(pop_t10),
    births = pop * fr,
    person_years = pop * surv
  ) %>%
  summarise(cbr = sum(births) / sum(person_years))

pop_asfr_1_t10 <- 0
pop_asfr_2_t10 <- f_tilde_2_asfr(.9, 1.05, .65, pop_t10[1], pop_t10[2], 1-.65)
pop_asfr_3_t10 <- f_tilde_2_asfr(.2, 1.05, .75, pop_t10[2], pop_t10[3], 1-.65)

pop_asfr_t10 <- c(pop_asfr_1_t10, pop_asfr_2_t10, pop_asfr_3_t10)

TFR_t10 <- sum(pop_asfr_t10)


# 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 -------------------------------------------------------------
```

```r
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 ------------------------------------------------------------

data(mxF)

thailand_mx <- mxF %>%
  filter(name == "Thailand") %>%
  select(age, mx = `2015-2020`)


# Question 2a -----------------------------------------------------------

plot_mx <- function(data, x, y) {
  ggplot(data, aes(x = {{ x }}, y = {{ y }})) +
    geom_point() +
    geom_line() +
    theme_bw() +
    theme(text = element_text(family = "serif")) +
    labs(
      title = "Age-Specific Mortality Rates",
      subtitle = "Thailand, Females, 2015-2020",
      caption = "Source: WPP2019",
      x = "Age (years)",
      y = "Mortality Rate"
    )
}

thailand_mx_plot <- plot_mx(thailand_mx, age, mx)
thailand_log_mx_plot <- plot_mx(thailand_mx, age, log(mx)) +
  labs(title = "Age-Specific Log Mortality Rates", y = "Log Mortality Rate")

plot(thailand_mx_plot)
plot(thailand_log_mx_plot)


# Question 2b -----------------------------------------------------------

thailand_l0 <- 100000
lx = numeric(nrow(thailand_mx))
lx[1] <- thailand_l0
n <- c(1, 4, rep(5, 20))

for (i in 2:length(lx)) {
  lx[i] <- (1 - (1 - exp(-1 * n[i - 1] * thailand_mx$mx[i - 1]))) * lx[i - 1]
}

thailand_lt <- thailand_mx %>%
  mutate(
```

```r
    age_str =
      paste0(age, "-", age + 4) %>%
      inset(c(1, 2, length(age)), c("0", "1-4", "100+")),
    n = n,
    qx = 1 - exp(-1 * n * mx),
    sx = 1 - qx,
    lx = lx,
    dx = lx - if_else(!is.na(lead(lx)), lead(lx), 0),
    ax = n / 2
  ) %>%
  select(age, age_str, n, everything())

thailand_lt[["ax"]][1] <- 0.07 + 1.7 * thailand_lt[["mx"]][1]
thailand_lt[["qx"]][nrow(thailand_lt)] <- 1
thailand_lt[["sx"]][nrow(thailand_lt)] <- 0

thailand_lt <- thailand_lt %>%
  mutate(
    Lx = 5 * if_else(!is.na(lead(lx)), lead(lx), 0) + ax * dx,
    Tx = rev(cumsum(rev(Lx))),
    ex = Tx / lx
  )

thailand_lt_names <- c(
  "Age",
  "$n$",
  "$_{5}m_x$",
  "$_{1}q_x$",
  "$_{1}s_x$",
  "$l_x$",
  "$_{5}d_x$",
  "$a_x$",
  "${}_{5}L_x$",
  "${}_{5}T_x$",
  "$e_x$"
)

knitr::kable(
  select(thailand_lt, -age),
  booktabs = TRUE,
  col.names = thailand_lt_names,
  eval = FALSE,
  digits = c(0, 1, 3, 3, 3, 0, 0, 2, 0, 0, 2),
  caption = "Life table for 2015-2020 Thailand female population"
)

# Question 3 --------------------------------------------------------------

data(tfr)
data(percentASFR)
data(sexRatio)
data(popF)
data(migration)
```

```r
# Question 3a -------------------------------------------------------------

thailand_pop <- popF %>% filter(name == "Thailand") %>% select(age, pop = `2015`)
thailand_srb <- sexRatio %>% filter(name == "Thailand") %>% pull(`2015-2020`)
thailand_tfr <- tfr %>% filter(name == "Thailand") %>% pull(`2015-2020`)

thailand_fertilty <- percentASFR %>%
  filter(name == "Thailand") %>%
  select(age, pasfr = `2015-2020`) %>%
  mutate(asfr = pasfr * thailand_tfr / 5) %>%
  right_join(thailand_pop, by = "age") %>%
  replace_na(list(pasfr = 0, asfr = 0))


# Question 3e -------------------------------------------------------------

# Net migration is per 1000 person-years
thailand_net_mig <- migration %>%
  filter(name == "Thailand") %>%
  pull(`2015-2020`)

thailand_mig <- thailand_pop %>%
  mutate(migrants = (thailand_net_mig / 1000) * pop)
```