

Discrete-time state transition models in R

Kyu Lee, Ph.D.

HEOR 533

Jan 24, 2024



Class objectives

1. Review Markov cohort models
2. Learn how to implement a Markov cohort model in R
 - Basics in R
 - Briggs A et al model in R
3. Preview the assignment

Markov cohort model (review)

- Markov models in medical decision science simulate health transition over time (**discrete-** vs. continuous- time)
- Markov models are used to estimate **long-term** cost and health outcomes beyond the empirical study period, often attached to the end of a branch in decision trees
 - e.g. For a cohort of patients who received treatment A and experienced post-treatment complication (with a probability of 0.2), what are their 5-year health and cost outcomes?
- Building blocks of Markov models include health states, transition probabilities, cycle, rewards or pay-offs

What is R?

- R is a free programming language and environment for statistical computing and graphics
- Users can add additional functionality such as mechanistic simulation modeling by using built-in functions, user-defined functions, and packages
- Any users can contribute to creating open-sourced R packages



R studio



- R studio is an IDE (Integrated Development Environment) for R where users can edit the code, import data, visualize graphics more easily compared to when writing the codes directly to R console.

R studio

The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains an R script for a decision model. The script is enclosed in a red box and includes comments and a function definition.


```

1  #-----
2  #####           Decision Model           #####
3  #-----
4  #' Decision Model
5  #'
6  #' \code{decision_model} implements the decision model used.
7  #'
8  #' @param l_params_all List with all parameters of decision model
9  #' @param verbose Logical variable to indicate print out of messages
10 #' @return The transition probability array and the cohort trace matrix.
11 #' @export
12 decision_model <- function(l_params_all, verbose = FALSE) {
13   with(as.list(l_params_all), {
14     ##### Process model inputs #####
15     ## Transition probabilities to the Dead state
16     # compute mortality rates
      
```
- Environment Pane:** Shows the Global Environment with four data objects:

Object	Size
out_highbeta	301 obs. of 6 variables
out_highbeta_t	1505 obs. of 3 variables
out_lowbeta	301 obs. of 6 variables
out_lowbeta_t	1505 obs. of 3 variables
- Files Pane:** Displays a file tree for the project 'Markov-cohort-introduction-main'.

Name	Size	Modified
..		
.gitignore	570 B	Apr 10, 2023, 10:...
.Rhistory	15.8 KB	Dec 18, 2023, 9:...
~\$Markov_model_realworld.xlsm	165 B	Dec 21, 2023, 11:...
figures		
input_tab.csv	1.4 KB	Apr 10, 2023, 10:...
LICENSE	1 KB	Apr 10, 2023, 10:...
Markov_Intro.Rproj	205 B	Apr 16, 2023, 8:5...
Markov_model_diagram.R	1 KB	Apr 10, 2023, 10:...
Markov_model_realworld_expla...	285.4 KB	Apr 13, 2023, 10:...
Markov_model_realworld_expla...	20.1 KB	Apr 13, 2023, 12:...
Markov_model_realworld.R	15.1 KB	Apr 10, 2023, 12:...
Markov_model_realworld.xlsm	228.8 KB	Apr 10, 2023, 10:...
original_Excel_files		
output_tab.csv	192 B	Apr 10, 2023, 10:...
- Console:** Shows the R version 4.2.1 (2022-06-23) -- "Funny-Looking Kid" and copyright information.


```

R version 4.2.1 (2022-06-23) -- "Funny-Looking Kid"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: aarch64-apple-darwin20 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
      
```

R studio

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for a function named `decision_model`. The code includes comments and a function definition that takes `l_params_all` and `verbose` as arguments.
- Environment Pane:** Shows the Global Environment with four data objects:

Object	Size
<code>out_highbeta</code>	301 obs. of 6 variables
<code>out_highbeta_t</code>	1505 obs. of 3 variables
<code>out_lowbeta</code>	301 obs. of 6 variables
<code>out_lowbeta_t</code>	1505 obs. of 3 variables
- Files Pane:** Lists files in the project directory, including `.gitignore`, `.Rhistory`, `~$Markov_model_realworld.xlsm`, `figures`, `input_tab.csv`, `LICENSE`, `Markov_Intro.Rproj`, `Markov_model_diagram.R`, `Markov_model_realworld_expla...`, `Markov_model_realworld.R`, `Markov_model_realworld.xlsm`, `original_Excel_files`, and `output_tab.csv`.
- Console:** Displays the R startup message for version 4.2.1 (2022-06-23) on aarch64-apple-darwin20 (64-bit). It includes the R Foundation's warranty disclaimer and information about natural language support and collaborative project status.

R studio

The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains an R script with a function definition for `decision_model`. The function takes `l_params_all` and `verbose` as arguments and returns a transition probability array and cohort trace matrix.
- Console:** Shows the R version (4.2.1) and copyright information for The R Foundation for Statistical Computing.
- Environment Pane:** Lists the objects in the Global Environment, including `out_highbeta`, `out_highbeta_t`, `out_lowbeta`, and `out_lowbeta_t`.
- Files Pane:** Displays the file structure of the project, including `.gitignore`, `.Rhistory`, `~$Markov_model_realworld.xlsm`, `figures`, `input_tab.csv`, `LICENSE`, `Markov_Intro.Rproj`, `Markov_model_diagram.R`, `Markov_model_realworld_expla...`, `Markov_model_realworld.R`, `Markov_model_realworld.xlsm`, `original_Excel_files`, and `output_tab.csv`.

```

1  #-----
2  #####          Decision Model          #####
3  #-----
4  #' Decision Model
5  #'
6  #' \code{decision_model} implements the decision model used.
7  #'
8  #' @param l_params_all List with all parameters of decision model
9  #' @param verbose Logical variable to indicate print out of messages
10 #' @return The transition probability array and the cohort trace matrix.
11 #' @export
12 decision_model <- function(l_params_all, verbose = FALSE) {
13   with(as.list(l_params_all), {
14     ##### Process model inputs #####
15     ## Transition probabilities to the Dead state
16     # compute mortality rates
17   })
18 }

```

R version 4.2.1 (2022-06-23) -- "Funny-Looking Kid"
 Copyright (C) 2022 The R Foundation for Statistical Computing
 Platform: aarch64-apple-darwin20 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
 You are welcome to redistribute it under certain conditions.
 Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.

R studio

The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains an R script for a decision model. The script includes comments and a function definition for `decision_model`.
- Environment Pane:** Shows the Global Environment with four data objects:

Object	Size
<code>out_highbeta</code>	301 obs. of 6 variables
<code>out_highbeta_t</code>	1505 obs. of 3 variables
<code>out_lowbeta</code>	301 obs. of 6 variables
<code>out_lowbeta_t</code>	1505 obs. of 3 variables
- Files Pane:** Displays the file structure of the project, including a directory listing of files and folders. The files pane is highlighted with a red box.
- Console:** Shows the R version (4.2.1) and copyright information for The R Foundation for Statistical Computing.

Basics in R

- Creating variables of scalar, vector, and matrix: an operator '`<-`' assigns a value to a variable

Scalar

```
> age <- 65
> name <- 'Mary'
> print(age); print(name)
[1] 65
[1] "Mary"
```

Vector

```
> v_age <- c(25,34,65)
> v_name <- c('Tom','Kyu','Mary')
> print(v_age)
[1] 25 34 65
> print(v_name)
[1] "Tom" "Kyu" "Mary"
```

Matrix

```
> m_prob <- matrix(c(0.1,0.9,0,0.2,0.8,0,0.3,0.5,0.2),
+                  nrow = 3,
+                  dimnames = list(c("Healthy","Sick","Dead"),
+                                  c("Healthy","Sick","Dead")))
> m_prob
```

	Healthy	Sick	Dead
Healthy	0.1	0.2	0.3
Sick	0.9	0.8	0.5
Dead	0.0	0.0	0.2

Basics in R

```
> m1 <- m2 <- matrix(seq(1,9),nrow=3,byrow=FALSE)
```

```
> m1
```

```
      [,1] [,2] [,3]
[1,]    1    4    7
[2,]    2    5    8
[3,]    3    6    9
```

```
> m3 <- m1 * m2
```

```
> m4 <- m1 %*% m2
```

- ‘*’ operator implements element-wise multiplication between two matrices

```
> m3
```

```
      [,1] [,2] [,3]
[1,]    1   16   49
[2,]    4   25   64
[3,]    9   36   81
```

- ‘%*%’ operator implements matrix multiplication between two matrices (dot product of each row vector (m1) and column vector (m2))

```
> m4
```

```
      [,1] [,2] [,3]
[1,]   30   66  102
[2,]   36   81  126
[3,]   42   96  150
```

=

```
      [,1] [,2] [,3]
[1,]    1    4    7
[2,]    2    5    8
[3,]    3    6    9
```

X

```
      [,1] [,2] [,3]
[1,]    1    4    7
[2,]    2    5    8
[3,]    3    6    9
```

Basics in R

- Array: a general data structure defined by a fixed number of dimensions; vectors (1D) and matrices (2D) are an example of array

```
pop <- array(data = rep(1,12),
             dim = c(2,3,2),
             dimnames = list(c("row1","row2"),
                             c("col1","col2","col3"),
                             c("Tx", "No Tx")))
```

> pop

	col1	col2	col3	
row1	1	1	1	Tx
row2	1	1	1	No Tx

	col1	col2	col3
row1	1	1	1
row2	1	1	1

, , No Tx

	col1	col2	col3
row1	1	1	1
row2	1	1	1

- List: a generic object consisting of an ordered collection of objects; can accommodate multiple data types

```
pop <- list( age = c(62,34,59),
             name = c("Mary","Kyu","Tom"),
             hometown = c("Seattle","Seoul","Rome"))
```

```
> print(pop)
```

```
$age
[1] 62 34 59
```

```
$name
[1] "Mary" "Kyu"  "Tom"
```

```
$hometown
[1] "Seattle" "Seoul"   "Rome"
```

```
> pop$name
```

```
[1] "Mary" "Kyu"  "Tom"
```

```
> pop$hometown[pop$name=='Kyu']
```

```
[1] "Seoul"
```

Basics in R

- A user-defined function

Name of the function

Arguments needed for the function

```
f_simple <- function(n1, n2){
```

```
  sum <- n1 + n2
```

```
  prod <- n1 * n2
```

← Operations

```
  return(c(sum,prod))
```

← Return outputs

```
}
```

```
> result <- f_simple(2,3)
```

```
> result
```

```
[1] 5 6
```

Translating an excel model in R

Category	Excel	R
Model organization	Define input parameters, run models, and compute outcomes in separate tabs	One code can streamline the works from defining input parameters, running models, to computing outcomes
Input parameters	Enter a value in a cell Link a cell to other cells	Assign a value to a variable Specify equations
Running Markov model	Repetitive calculation over rows	For-loop
Sensitivity analysis	Macro	For-loop

Translating an excel model in R

Parameters of the model				
Name	live	alpha	beta	description
Transition probabilities				
tpA2A	0.721	1251	483	Transition probability from A to A
tpA2B	0.202	350	1384	Transition probability from A to B
tpA2C	0.067	116	1618	Transition probability from A to C
tpA2D	0.010	17	1717	Transition probability from A to D
tpB2B	0.581	731	527	Transition probability from B to B
tpB2C	0.407	512	746	Transition probability from B to C
tpB2D	0.012	15	1243	Transition probability from B to D
tpC2C	0.750	1312	437	Transition probability from C to C
tpC2D	0.250	437	1312	Transition probability from C to D
Costs				
dmca	£	1,701		Direct medical costs associated with state A
dmcb	£	1,774		Direct medical costs associated with state B
dmcc	£	6,948		Direct medical costs associated with state C
ccca	£	1,055		Community care costs associated with state A
cccb	£	1,278		Community care costs associated with state B
cccc	£	2,059		Community care costs associated with state C
Drug costs				
cAZT	£	2,278		Zidovudine drug cost
cLam	£	2,087		Lamivudine drug cost
Other parameters				
RR	0.509			Treatment effect (RR)
cDR	0.06			Annual discount rate - costs (%)
oDR	0			Annual discount rate - benefits (%)

MONOTHERAPY			
A	B	C	D
0.721453	0.201845	0.066897	0.009804
	0.581081	0.406995	0.011924
		0.750143	0.249857
COMBINATION THERAPY (year 1,2)			
A	B	C	D
0.85822	0.102739	0.034051	0.00499
	0.78677	0.207161	0.006069
		0.872823	0.127177
COMBINATION THERAPY (year 3+)			
A	B	C	D
0.721453	0.201845	0.066897	0.009804
	0.581081	0.406995	0.011924
		0.750143	0.249857

Each cell in the transition probability matrix refers to a parameter value or a value calculated based on multiple parameter values

```

tpA2B <- 0.202
tpA2C <- 0.067
tpA2D <- 0.010
tpA2A <- 1-tpA2B-tpA2C-tpA2D
tpB2C <- 0.407
tpB2D <- 0.012
tpB2B <- 1-tpB2C-tpB2D
tpC2D <- 0.250
tpC2C <- 1-tpC2D

mP_mono <- matrix(c(tpA2A, tpA2B, tpA2C, tpA2D,
                    0, tpB2B, tpB2C, tpB2D,
                    0, 0, tpC2C, tpC2D,
                    0, 0, 0, 1), byrow = TRUE, nrow=4,
                  dimnames = list(c('A', 'B', 'C', 'D'), c('A', 'B', 'C', 'D')))

rr <- 0.5
mP_comb1 <- matrix(c(1-tpA2B*rr-tpA2C-tpA2D, tpA2B*rr, tpA2C, tpA2D,
                    0, tpB2B, tpB2C, tpB2D,
                    0, 0, tpC2C, tpC2D,
                    0, 0, 0, 1), byrow=TRUE, nrow=4,
                  dimnames = list(c('A', 'B', 'C', 'D'), c('A', 'B', 'C', 'D')))

```

In R, parameters and transition probability matrix can be defined as a variable of scalar or matrix

The code makes the calculation more explicit, reducing the risk of transcription errors

Translating an excel model in R

Markov model					
Year	MONOTHERAPY				
	A	B	C	D	
0	1				
1	0.721	0.202	0.067	0.010	
2	=C8*(1-tpA2B-tpA2C-tpA2D)				0.036
3	0.376	0.258	0.277	0.089	
4	0.271	0.226	0.338	0.165	
5	0.195	0.186	0.364	0.255	
6	0.141	0.147	0.361	0.350	
7	0.102	0.114	0.341	0.444	
8	0.073	0.087	0.309	0.531	
9	0.053	0.065	0.272	0.610	
10	0.038	0.049	0.234	0.679	
11	0.028	0.036	0.198	0.739	
12	0.020	0.026	0.165	0.789	
13	0.014	0.019	0.136	0.830	
14	0.010	0.014	0.111	0.865	
15	0.007	0.010	0.090	0.893	
16	0.005	0.008	0.072	0.915	
17	0.004	0.005	0.057	0.933	
18	0.003	0.004	0.045	0.948	
19	0.002	0.003	0.036	0.959	
20	0.0015	0.0021	0.0282	0.9682	

Pop distribution at t=1

A	B	C	D

=

A	B	C	D
1	0	0	0

×

	A	B	C	D
A	0.75	0.2	0.06	0.01
B	0	0.59	0.4	0.01
C	0	0	0.7	0.3
D	0	0	0	1

Pop distribution at t=0

Transition probability matrix

Repeat until t = 20

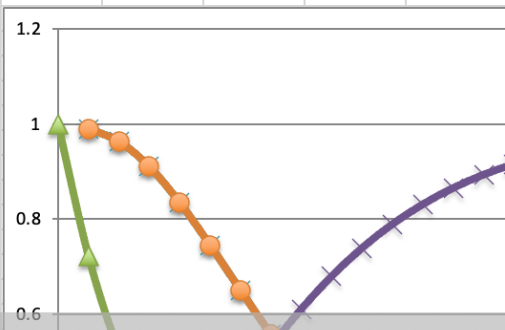
```
for (j in range(1,20)){
```

```
  pop[, cycle = j] <-  
    pop[, cycle = j - 1] %*% p_matrix
```

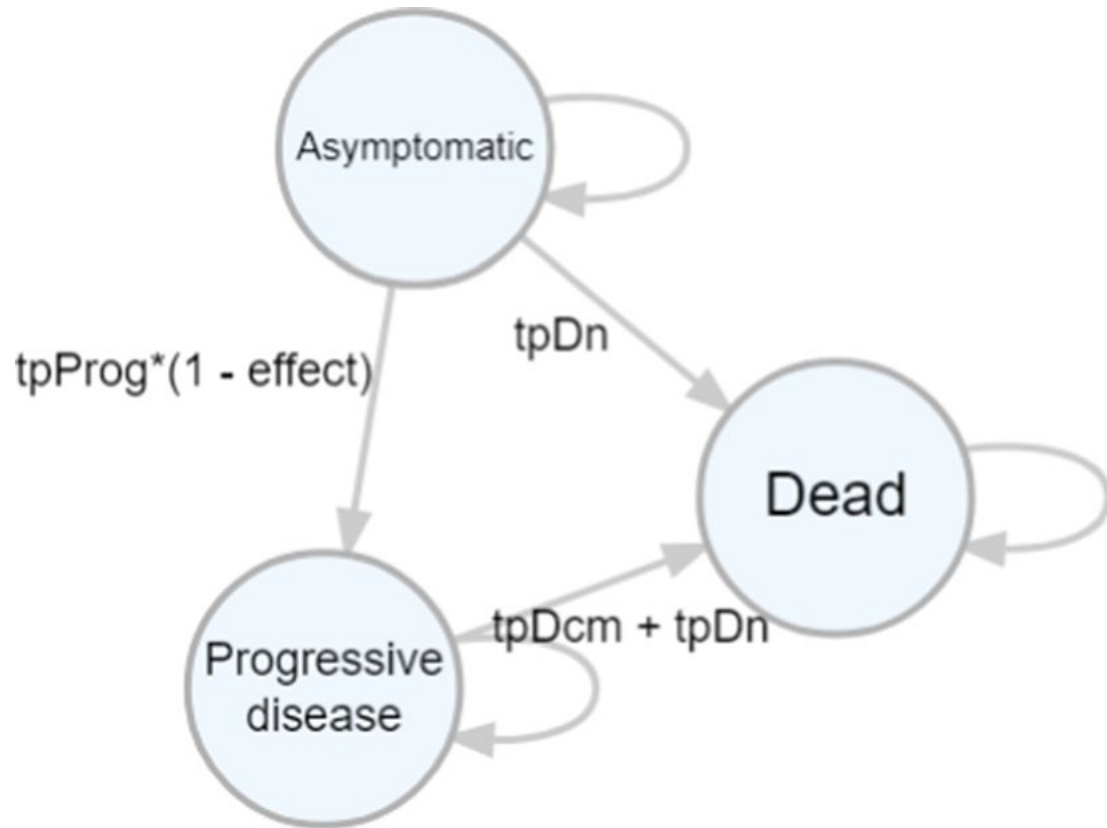
<- matrix
multiplication

Repeat until
t = 20

```
}
```



Review of Briggs A. et al. Model



- 3-state Markov cohort model
- The risk of progressing from Asymptomatic to Progressive Disease decreases by 50% with treatment
- Disease progression accelerates with age
- Progressive disease has excess mortality ($tpDcm$)
- There is cost for excess mortality among progressive diseases (i.e. transitional cost from progressive to dead)

Questions?



R session – Review and run the code

Open the R code (Markov_model_realworld_explain2_2024.Rmd)

R session – class exercise (optional)

- Make the following changes and discuss how ICER of treatment changes
 - Reduce the treatment efficacy from 0.5 to 0.2
 - Increase the starting age from 55 to 65

Assignments

- A new end-of-life care was approved based on the clinical trial result that it can reduce the excess mortality due to progressive disease by 50%. It costs \$500 as one-time cost.
- Decision makers (e.g. clinicians) requested information on whether they should provide this end-of-life care in addition to the original treatment to reduce the disease progression.
- Using the same health state transition model, conduct a cost-effectiveness analysis considering three strategies: 1) without treatment 2) with treatment but no end-of-life care, 3) with both treatment and end-of-life care
- Provide a table of cost and QALY of three strategies and report ICER. It is optional to generate ICER graph.

Wrapping up

1. Review Markov cohort models

- Discrete-time health transition models simulate health transitions over time and estimate long-term health and cost outcomes

2. Learn how to implement a Markov cohort model in R

- R is a free programming language and has flexible functionality to build simulation models such as health transition models
- Compared to excel models, Rmodels can streamline the process of building the model in an explicit way in a single code file with lower risk of transcription error.