

project

```
library(markovchain)

## Package:  markovchain
## Version:  0.8.5-4
## Date:     2021-01-07
## BugReport: https://github.com/spedygiorgio/markovchain/issues

library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --

## v ggplot2 3.3.2      v purrr  0.3.4
## v tibble  3.0.4      v dplyr  1.0.2
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

dat <- read_csv('data/dtmc_data.csv', col_types = cols(
  index = col_double(),
  stay_date = col_date(format = ""),
  PTID = col_character(),
  first_covid_diag_date = col_date(format = ""),
  # hcq_usage_date = col_date(format = ""),
  medadm_drug_name_hcq_use = col_double(),
  CAREAREA = col_factor(),
  days_after_confirmation = col_double()
))

use matching before DTMC

table(dat$CAREAREA)

##
## UNKNOWN CARE AREA
## 55424
## MEDICAL OR SURGICAL UNIT
## 8589
## EMERGENCY DEPARTMENT
## 35875
## OTHER CARE AREA
## 44712
## CRITICAL CARE UNIT (CCU) / INTENSIVE CARE UNIT (ICU)
## 13431
## ANCILLARY SERVICES
## 3592
## LABORATORY
## 2300
```

##	AMBULATORY SURGERY	
##		663
##	OPERATING ROOM	
##		927
##	CARDIAC CATHETERIZATION	
##		525
##	AMBULATORY CARE	
##		4195
##	OFFICE OR CLINIC	
##		1756
##	URGENT CARE UNIT	
##		30
##	TELEMETRY UNIT	
##		172
##	HOSPICE CARE AREA	
##		34
##	PSYCHIATRIC UNIT	
##		48
##	OBSERVATION CARE	
##		332
##	POST ANESTHESIA CARE UNIT (PACU) / RECOVERY ROOM	
##		210
##	DELIVERY ROOM	
##		116
##	NURSERY	
##		1
##	PEDIATRIC UNIT	
##		18
##	REHABILITATION UNIT	
##		38
##	DAY CARE UNIT	
##		13

group those carearea to less categories

ICU: CRITICAL CARE UNIT (CCU) / INTENSIVE CARE UNIT (ICU)

high dependent unit: CARDIAC CATHETERIZATION URGENT CARE UNIT TELEMETRY UNIT
DELIVERY ROOM EMERGENCY DEPARTMENT OPERATING ROOM

Stable units: AMBULATORY CARE

AMBULATORY SURGERY MEDICAL OR SURGICAL UNIT PEDIATRIC UNIT PSYCHIATRIC UNIT
REHABILITATION UNIT SKILLED NURSING FACILITY (SNF) CARE AREA NURSERY POST ANES-
THESIA CARE UNIT (PACU) / RECOVERY ROOM HOSPICE CARE AREA LABORATORY DAY
CARE UNIT ANCILLARY SERVICES OTHER CARE AREA

Missing: UNKNOWN CARE AREA

[1] "OFFICE OR CLINIC" "UNKNOWN CARE AREA"

[3] "OTHER CARE AREA" "EMERGENCY DEPARTMENT"

[5] "LABORATORY" "ANCILLARY SERVICES"

[7] "URGENT CARE UNIT" "CRITICAL CARE UNIT (CCU) / INTENSIVE CARE UNIT (ICU)" [9]
"MEDICAL OR SURGICAL UNIT" "OBSERVATION CARE"

[11] "REHABILITATION UNIT" "PEDIATRIC UNIT"

[13] "AMBULATORY SURGERY" "TELEMETRY UNIT"

[15] "AMBULATORY CARE" "DELIVERY ROOM"

[17] "PSYCHIATRIC UNIT" "CARDIAC CATHETERIZATION"

[19] “OPERATING ROOM”

```
high_dependent_units <- c(
  'CARDIAC CATHETERIZATION',
  'URGENT CARE UNIT',
  'TELEMETRY UNIT',
  'DELIVERY ROOM',
  'EMERGENCY DEPARTMENT',
  'OPERATING ROOM')

other_hospital_units <- c(
  'AMBULATORY CARE',
  'AMBULATORY SURGERY',
  'MEDICAL OR SURGICAL UNIT',
  'PEDIATRIC UNIT',
  'PSYCHIATRIC UNIT',
  'REHABILITATION UNIT',
  'HOSPICE CARE AREA',
  'LABORATORY',
  'DAY CARE UNIT',
  'POST ANESTHESIA CARE UNIT (PACU) / RECOVERY ROOM',
  'ANCILLARY SERVICES',
  'OTHER CARE AREA',
  'SKILLED NURSING FACILITY (SNF) CARE AREA',
  'NURSERY')

long_term_medical_care_facilities <- c(
  'SKILLED NURSING FACILITY (SNF) CARE AREA',
  'NURSERY'
)

dat <- dat %>% mutate(care_area=ifelse(CAREAREA == 'CRITICAL CARE UNIT (CCU) / INTENSIVE CARE UNIT (ICU)',
                                     ifelse(CAREAREA %in% high_dependent_units, 'high dependent units', 'other hospital units'),
                                     ifelse(CAREAREA %in% other_hospital_units, 'other hospital units', 'long term medical care facilities'))

dat_HCQ <- dat %>% filter(medadm_drug_name_hcq_use == 1)
dat_no_HCQ <- dat %>% filter(medadm_drug_name_hcq_use == 0)

# function used to create transition states vector for each of those patients
create_trans_states <- function(data) {
  ptids <- unique(data$PTID)
  lst <- list()
  for (ptid in ptids) {
    lst[[ptid]] <- pull(dat %>% filter(PTID == ptid) %>%
      select(care_area))
  }

  return(lst)
}

HCQ_patients <- create_trans_states(dat_HCQ)
no_HCQ_patients <- create_trans_states(dat_no_HCQ)

first_day_stay = vector()
for(i in 1:length(no_HCQ_patients)) {
```

```

first_day_stay[i] = no_HCQ_patients[[i]][1]
}

```

```
no_HCQ_patients[1:10]
```

```

## $PT254611341
## [1] NA "other hospital units or self-care"
##
## $PT413921534
## [1] "high dependent units" "high dependent units" "high dependent units"
## [4] NA NA NA
## [7] NA NA NA
## [10] NA NA NA
## [13] NA NA NA
## [16] NA NA NA
## [19] NA NA NA
## [22] NA NA NA
## [25] NA NA NA
## [28] NA NA NA
## [31] NA
##
## $PT277747789
## [1] "high dependent units" "high dependent units"
## [3] "high dependent units" "other hospital units or self-care"
## [5] "other hospital units or self-care" "other hospital units or self-care"
## [7] "other hospital units or self-care" "other hospital units or self-care"
## [9] "other hospital units or self-care" "other hospital units or self-care"
## [11] "other hospital units or self-care" "other hospital units or self-care"
## [13] "other hospital units or self-care" "other hospital units or self-care"
## [15] "other hospital units or self-care" "other hospital units or self-care"
## [17] "other hospital units or self-care" "other hospital units or self-care"
## [19] "other hospital units or self-care" "other hospital units or self-care"
## [21] "other hospital units or self-care" "other hospital units or self-care"
## [23] "other hospital units or self-care" "other hospital units or self-care"
## [25] "other hospital units or self-care" "other hospital units or self-care"
## [27] "other hospital units or self-care" "other hospital units or self-care"
## [29] "other hospital units or self-care" "other hospital units or self-care"
## [31] "other hospital units or self-care"
##
## $PT202846287
## [1] NA NA
##
## $PT249283846
## [1] "high dependent units" "high dependent units" "high dependent units"
## [4] "high dependent units" "high dependent units" "high dependent units"
## [7] "high dependent units" "high dependent units" "high dependent units"
## [10] NA NA NA
## [13] NA NA "high dependent units"
## [16] NA
##
## $PT411742036
## [1] "ICU" "ICU"
##
## $PT265624106

```

```

## [1] "other hospital units or self-care" "other hospital units or self-care"
## [3] "other hospital units or self-care" "other hospital units or self-care"
## [5] "other hospital units or self-care" "other hospital units or self-care"
## [7] "other hospital units or self-care" "other hospital units or self-care"
## [9] "other hospital units or self-care" "other hospital units or self-care"
## [11] "other hospital units or self-care" "other hospital units or self-care"
## [13] "other hospital units or self-care" "other hospital units or self-care"
## [15] "other hospital units or self-care" "other hospital units or self-care"
## [17] "other hospital units or self-care" "other hospital units or self-care"
## [19] "other hospital units or self-care" "other hospital units or self-care"
## [21] "other hospital units or self-care" "other hospital units or self-care"
## [23] "other hospital units or self-care" "other hospital units or self-care"
## [25] "other hospital units or self-care" "other hospital units or self-care"
## [27] "other hospital units or self-care" "other hospital units or self-care"
## [29] "other hospital units or self-care" "other hospital units or self-care"
## [31] "other hospital units or self-care"
##
## $PT412682057
## [1] "ICU" "ICU"
## [3] "ICU" "other hospital units or self-care"
## [5] NA NA
## [7] NA NA
## [9] "other hospital units or self-care" "other hospital units or self-care"
## [11] "other hospital units or self-care" "other hospital units or self-care"
## [13] NA "ICU"
## [15] "ICU" "ICU"
## [17] NA NA
## [19] NA NA
## [21] NA NA
## [23] NA NA
## [25] NA NA
## [27] NA NA
## [29] NA NA
## [31] NA
##
## $PT349802348
## [1] "high dependent units"
##
## $PT560460142
## [1] "high dependent units" "high dependent units" "high dependent units"
## [4] "high dependent units" "high dependent units" "high dependent units"
## [7] "high dependent units" "high dependent units" "high dependent units"
## [10] "high dependent units" "high dependent units" "high dependent units"
## [13] "high dependent units" "high dependent units" "high dependent units"
## [16] "high dependent units" "high dependent units" "high dependent units"
## [19] "high dependent units" "high dependent units" "high dependent units"
## [22] "high dependent units" "high dependent units" "high dependent units"
## [25] "high dependent units" "high dependent units" "high dependent units"
## [28] "high dependent units" "high dependent units" "high dependent units"
## [31] "high dependent units"

```

```
createSequenceMatrix(HCQ_patients)
```

```

## ICU high dependent units
## ICU 5514 13

```

```

## high dependent units          62          6022
## other hospital units or self-care 148          111
##                                other hospital units or self-care
## ICU                                332
## high dependent units          281
## other hospital units or self-care 18803

fit.hcq.mle <- markovchainFit(data = HCQ_patients, method = 'mle')

(markovchain <- fit.hcq.mle$estimate)

## MLE Fit
## A 3 - dimensional discrete Markov Chain defined by the following states:
## ICU, high dependent units, other hospital units or self-care
## The transition matrix (by rows) is defined as follows:
##                                ICU high dependent units
## ICU          0.941116231          0.002218809
## high dependent units 0.009740770          0.946111548
## other hospital units or self-care 0.007764138          0.005823104
##                                other hospital units or self-care
## ICU          0.05666496
## high dependent units 0.04414768
## other hospital units or self-care 0.98641276

fit.hcq.mle$standardError

##                                ICU high dependent units
## ICU          0.0126738885          0.0006153868
## high dependent units 0.0012370790          0.0121919162
## other hospital units or self-care 0.0006382082          0.0005527045
##                                other hospital units or self-care
## ICU          0.003109894
## high dependent units 0.002633630
## other hospital units or self-care 0.007193580

fit.hcq.mle$lowerEndpointMatrix

##                                ICU high dependent units
## ICU          0.916275861          0.001012672
## high dependent units 0.007316139          0.922215826
## other hospital units or self-care 0.006513273          0.004739822
##                                other hospital units or self-care
## ICU          0.05056968
## high dependent units 0.03898586
## other hospital units or self-care 0.97231360

fit.hcq.mle$upperEndpointMatrix

##                                ICU high dependent units
## ICU          0.965956602          0.003424945
## high dependent units 0.012165401          0.970007269
## other hospital units or self-care 0.009015003          0.006906385
##                                other hospital units or self-care
## ICU          0.06276024
## high dependent units 0.04930950
## other hospital units or self-care 1.00000000

```

```

# mcDf <- as(markovchain, 'data.frame')
# mcNew <- as(mcDf, 'igraph')

fit.nohcq.mle <- markovchainFit(data = no_HCQ_patients, method = 'mle')

fit.nohcq.mle$estimate

## MLE Fit
## A 3 - dimensional discrete Markov Chain defined by the following states:
## ICU, high dependent units, other hospital units or self-care
## The transition matrix (by rows) is defined as follows:
##
##           ICU high dependent units
## ICU           0.928734235           0.005774198
## high dependent units           0.007542272           0.946327089
## other hospital units or self-care 0.005577122           0.012087858
##
##           other hospital units or self-care
## ICU                               0.06549157
## high dependent units               0.04613064
## other hospital units or self-care  0.98233502

fit.nohcq.mle$standardError

##
##           ICU high dependent units
## ICU           0.0118795443           0.0009366987
## high dependent units           0.0005143787           0.0057617246
## other hospital units or self-care 0.0003701666           0.0005449627
##
##           other hospital units or self-care
## ICU                               0.003154618
## high dependent units               0.001272116
## other hospital units or self-care  0.004912719

fit.nohcq.mle$lowerEndpointMatrix

##
##           ICU high dependent units
## ICU           0.905450751           0.003938302
## high dependent units           0.006534108           0.935034314
## other hospital units or self-care 0.004851608           0.011019751
##
##           other hospital units or self-care
## ICU                               0.05930863
## high dependent units               0.04363734
## other hospital units or self-care  0.97270627

fit.nohcq.mle$upperEndpointMatrix

##
##           ICU high dependent units
## ICU           0.952017719           0.007610095
## high dependent units           0.008550436           0.957619864
## other hospital units or self-care 0.006302635           0.013155966
##
##           other hospital units or self-care
## ICU                               0.07167451
## high dependent units               0.04862394
## other hospital units or self-care  0.99196378

#simulation

transion_matrix <- matrix(c(0.90, 0.08, 0.02,
                           0.02, 0.9, 0.08,

```

```

                                0.01, 0.01, 0.98),nrow=3,byrow=TRUE)

sim_tran_matrix <- new("markovchain", states=c("ICU","high dependent units", 'other hospital units or s
                                transitionMatrix=transion_matrix,
                                name="simpleMc")

check simulation sequence for one patient

# (sim_seq <- rmarkovchain(n = 30, object = sim_tran_matrix, t0 = 'ICU'))

#Here begins the parameter estimation part
set.seed(994)
upperlmt <- matrix(0, nrow=3, ncol=3)
lowerlmt <- matrix(0, nrow=3, ncol=3)
emat <- matrix(0, nrow=3, ncol=3)
er <- 0
count_of_within_CI <- 0
totol_checks <- 0
# do simulation for 100 times
for (i in 1:1000){
  sim_seq_list <- list()
  for (j in 1:3000) {
    # generate simulation data
    inital_status <- sample(c('ICU', 'high dependent units', 'other hospital units or self-care'), 1, p
    sim_seq_list[[j]] <- rmarkovchain(n = 30, object = sim_tran_matrix, t0 = inital_status)
  }

  # fit the current simulation data
  mkfit <- markovchainFit(data = sim_seq_list, method = 'mle', confidencelevel = 0.95)

  # get the fitted results
  upperlmt_temp <- mkfit$upperEndpointMatrix
  lowerlmt_temp <- mkfit$lowerEndpointMatrix
  emat_temp <- mkfit$estimate@transitionMatrix
  er_temp <- mkfit$standardError

  # sum the fitted result
  upperlmt <- upperlmt + upperlmt_temp
  lowerlmt <- lowerlmt + lowerlmt_temp
  emat <- emat + emat_temp
  er <- er + er_temp

  # check the fitted result of current simulation data within in the CI
  count_of_within_CI_temp <- c(lowerlmt_temp < transion_matrix & transion_matrix < upperlmt_temp)
  # transfer the boolean to numeric
  count_of_within_CI_temp_numeric <- as.numeric(count_of_within_CI_temp)
  #
  num_of_checks <- as.numeric(!is.na(count_of_within_CI_temp_numeric))
  # count the total number of check for whether CI cover the true parameters
  totol_checks <- totol_checks + num_of_checks
  # count the total number of check the CI cover the true parameters
  count_of_within_CI <- count_of_within_CI + count_of_within_CI_temp_numeric
}

```



```
upperlmt <- upperlmt/100
lowerlmt <- lowerlmt/100
emat <- emat/100
er <- er/100
#got the overall lower and upper as (0.166, 0.756, 0.198, 0.645), (0.217, 0.861, 0.314, 0.843), er as (
cp <- count_of_within_CI/totol_checks
```

Prediction

```
predict <- function(transMatrix, n, init_state) {
  status <- init_state
  # record the cumulative sum of length of stay for those three states from day 1 to day 30
  count <- matrix(status, byrow = T, nrow = 1, dimnames = list(c(), c('ICU', 'high dependent units', 'other hospital units or self-care')))
  df <- data.frame()
  df <- rbind(df, count)
  for (i in 1:(n - 1)) {
    status <- status %*% transMatrix
    count <- count + status
    df <- rbind(df, count)
  }

  df['day'] <- 1:30
  df <- df %>% select(c('day', 'ICU', 'high dependent units', 'other hospital units or self-care'))

  return(df)
}
```

```
(no_hcq_other <- predict(fit.nohcq.mle$estimate@transitionMatrix, 30, c(0, 0, 1)))
```

##	day	ICU	high dependent units	other hospital units or self-care
## 1	1	0.000000000	0.000000000	1.000000
## 2	2	0.005577122	0.01208786	1.982335
## 3	3	0.016326557	0.03543346	2.948240
## 4	4	0.031872974	0.06926382	3.898863
## 5	5	0.051868362	0.11285917	4.835272
## 6	6	0.075990040	0.16554928	5.758461
## 7	7	0.103938805	0.22671000	6.669351
## 8	8	0.135437217	0.29576015	7.568803
## 9	9	0.170228015	0.37215850	8.457613
## 10	10	0.208072644	0.45540103	9.336526
## 11	11	0.248749888	0.54501839	10.206232
## 12	12	0.292054609	0.64057348	11.067372
## 13	13	0.337796571	0.74165924	11.920544
## 14	14	0.385799359	0.84789658	12.766304
## 15	15	0.435899368	0.95893245	13.605168
## 16	16	0.487944871	1.07443807	14.437617
## 17	17	0.541795155	1.19410721	15.264098
## 18	18	0.597319718	1.31765468	16.085026
## 19	19	0.654397523	1.44481486	16.900788
## 20	20	0.712916316	1.57534039	17.711743
## 21	21	0.772771979	1.70900084	18.518227
## 22	22	0.833867945	1.84558163	19.320550
## 23	23	0.896114644	1.98488289	20.119002

```
## 24 24 0.959428996      2.12671843      20.913853
## 25 25 1.023733940      2.27091488      21.705351
## 26 26 1.088957996      2.41731072      22.493731
## 27 27 1.155034858      2.56575550      23.279210
## 28 28 1.221903022      2.71610912      24.061988
## 29 29 1.289505432      2.86824104      24.842254
## 30 30 1.357789161      3.02202969      25.620181
```

```
predict(fit.hcq.mle$estimate@transitionMatrix, 30, c(0, 0, 1))
```

```
##      day      ICU high dependent units other hospital units or self-care
## 1      1 0.000000000      0.000000000      1.000000
## 2      2 0.007764138      0.005823104      1.986413
## 3      3 0.022786461      0.017093620      2.960120
## 4      4 0.044593993      0.033460115      3.921946
## 5      5 0.072744587      0.054593844      4.872662
## 6      6 0.106824915      0.080187285      5.812988
## 7      7 0.146448587      0.109952770      6.743599
## 8      8 0.191254398      0.143621200      7.665124
## 9      9 0.240904684      0.180940846      8.578154
## 10    10 0.295083787      0.221676227      9.483240
## 11    11 0.353496623      0.265607061     10.380896
## 12    12 0.415867340      0.312527283     11.271605
## 13    13 0.481938060      0.362244127     12.155818
## 14    14 0.551467716      0.414577266     13.033955
## 15    15 0.624230946      0.469358010     13.906411
## 16    16 0.700017080      0.526428554     14.773554
## 17    17 0.778629171      0.585641275     15.635730
## 18    18 0.859883111      0.646858074     16.493259
## 19    19 0.943606788      0.709949764     17.346443
## 20    20 1.029639303      0.774795489     18.195565
## 21    21 1.117830246      0.841282192     19.040888
## 22    22 1.208039005      0.909304108     19.882657
## 23    23 1.300134131      0.978762295     20.721104
## 24    24 1.393992741      1.049564190     21.556443
## 25    25 1.489499959      1.121623204     22.388877
## 26    26 1.586548393      1.194858329     23.218593
## 27    27 1.685037648      1.269193783     24.045769
## 28    28 1.784873873      1.344558671     24.870567
## 29    29 1.885969330      1.420886668     25.693144
## 30    30 1.988241996      1.498115728     26.513642
```

```
no_hcq_other <- predict(fit.nohcq.mle$estimate@transitionMatrix, 30, c(0, 0, 1))
hcq_other <- predict(fit.hcq.mle$estimate@transitionMatrix, 30, c(0, 0, 1))
comb <- cbind(no_hcq_other, hcq_other)
colnames(comb) <- c('day', 'ICU no hcq', 'high dependent units no hcq', 'other hospital units or self-care',
                    'high dependent units', 'other hospital units or self-care')
comb <- comb %>% select('day', 'ICU no hcq', 'ICU', 'high dependent units no hcq', 'high dependent units',
                       'other hospital units or self-care', 'other hospital units or self-care')
write.csv(comb, 'data/other.csv')
```

```
no_hcq_icu <- predict(fit.nohcq.mle$estimate@transitionMatrix, 30, c(1, 0, 0))
hcq_icu <- predict(fit.hcq.mle$estimate@transitionMatrix, 30, c(1, 0, 0))
comb <- cbind(no_hcq_icu, hcq_icu)
colnames(comb) <- c('day', 'ICU no hcq', 'high dependent units no hcq', 'other hospital units or self-care',
                    'high dependent units', 'other hospital units or self-care')
```

```

      'high dependent units', 'other hospital units or self-care')
comb <- comb %>% select('day', 'ICU no hcq', 'ICU', 'high dependent units no hcq', 'high dependent units')

write.csv(comb, 'data/icu.csv')

no_hcq_high <- predict(fit.nohcq.mle$estimate@transitionMatrix, 30, c(0, 1, 0))
hcq_high <- predict(fit.hcq.mle$estimate@transitionMatrix, 30, c(0, 1, 0))
comb <- cbind(no_hcq_high, hcq_high)
colnames(comb) <- c('day', 'ICU no hcq', 'high dependent units no hcq', 'other hospital units or self-care',
      'high dependent units', 'other hospital units or self-care')
comb <- comb %>% select('day', 'ICU no hcq', 'ICU', 'high dependent units no hcq', 'high dependent units')

write.csv(comb, 'data/high.csv')

(sim_seq <- rmarkovchain(n = 30, object = sim_tran_matrix, t0 = 'ICU'))

## [1] "ICU" "ICU" "ICU"
## [4] "ICU" "ICU" "ICU"
## [7] "ICU" "ICU" "ICU"
## [10] "high dependent units" "high dependent units" "high dependent units"
## [13] "high dependent units" "ICU" "ICU"
## [16] "ICU" "ICU" "ICU"
## [19] "ICU" "ICU" "ICU"
## [22] "ICU" "ICU" "ICU"
## [25] "ICU" "ICU" "ICU"
## [28] "ICU" "ICU" "ICU"

```