

Model Complete Cases

Martin Skarzynski

2018-04-19

```
library(here)

## here() starts at /Users/marskar/gdrive/nhanes

library(readr)
library(dplyr)

##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(tidyr)
library(survey)

## Loading required package: grid
## Loading required package: methods
## Loading required package: Matrix
##
## Attaching package: 'Matrix'
## The following object is masked from 'package:tidyr':
##
##   expand
## Loading required package: survival
##
## Attaching package: 'survey'
## The following object is masked from 'package:graphics':
##
##   dotchart

library(purrr)

# this function takes in two integers as an argument
# this function returns a dataframe
get_modelstats <- function(seed, size){

  set.seed(seed)
  #move PERMTH_INT and canc_mort to the beginning
  #sample a tenth of the dataset columns
  read_rds(here('dat/3-clean-complete-cases.rds')) %>%
    select(-SEQN) %>%
```

```

select(PERMTH_INT,
       canc_mort,
       SDPPSU6,
       SDPSTRA6,
       WTPFQX6,
       everything()[sample(seq(ncol(.)),
                           round(size))]) ->
samp

# create survey design object
svydesign(ids = ~SDPPSU6,
         strata = ~SDPSTRA6,
         weights = ~WTPFQX6,
         nest = TRUE,
         data = samp) ->
des

# create left side of equations
form <- as.formula(Surv(PERMTH_INT, canc_mort) ~ x1)
# create right sides of equations
if(size == 1 & ncol(samp)==7){
vrs <- as.name(names(samp)[7])
vrs <- as.name(names(samp)[7])
} else{

vrs <- as.name(paste(names(samp)[6:ncol(samp)],
                     collapse=' + '))
vrs2 <- as.name(paste(names(samp)[6:ncol(samp)],
                      collapse=', '))
}

set.seed(seed)
#train <- sample(x = seq(nrow(samp)),
#               size = round(nrow(samp)*.7))
# generate cox models without and with penalties

cox <- try(svycoxph(update(form,
                          paste("~ ", vrs)),
              design = des, data = samp))

rid <- try(svycoxph(update(form,
                          paste("~ ridge(", vrs2, ')')),
              design = des, data = samp))

# define functions needed to create first table
get_con <- function(x) {
signif(summary(x)$concordance[1]*100, digits = 2)
}
get_HR <- function(x) {
summary(x)$conf.int[, "exp(coef)"]
}
get_HR_CI_lower <- function(x) {
summary(x)$conf.int[, "lower .95"]
}

```

```

}
get_HR_CI_upper <- function(x) {
  summary(x)$conf.int[, "upper .95"]
}
get_coef_pvalue <- function(x) {
  coefs <- summary(x)$coef
  coefs[, ncol(coefs)]
}
model_list <- try(list(cox, rid))

try(data_frame(seed = rep(seed, 2),
  size = size,
  type = c('coxph', 'ridge'),
  aic = AIC(cox, rid)[, "AIC"],
  concordance = map_dbl(model_list,
    get_con),
  hazard_ratio = map(model_list,
    get_HR),
  HR_CI_lower = map(model_list,
    get_HR_CI_lower),
  HR_CI_upper = map(model_list,
    get_HR_CI_upper),
  coef_pvalue = map(model_list,
    get_coef_pvalue)))
}

#save an object with 1000 models

map_sizes <- function(seed){
  map2_dfr(.x = seed,
    .y = seq(48),
    get_modelstats)
}
map_dfr(seq(10), map_sizes) %>%
write_rds(here(paste0("dat/6-model-diff-sizes.rds")))

```

```

## Stratified 1 - level Cluster Sampling design (with replacement)
## With (98) clusters.
## svydesign(ids = ~SDPPSU6, strata = ~SDPSTRA6, weights = ~WTPFQX6,
##   nest = TRUE, data = samp)
## Stratified 1 - level Cluster Sampling design (with replacement)
## With (98) clusters.
## svydesign(ids = ~SDPPSU6, strata = ~SDPSTRA6, weights = ~WTPFQX6,
##   nest = TRUE, data = samp)
## Stratified 1 - level Cluster Sampling design (with replacement)
## With (98) clusters.
## svydesign(ids = ~SDPPSU6, strata = ~SDPSTRA6, weights = ~WTPFQX6,
##   nest = TRUE, data = samp)
## Stratified 1 - level Cluster Sampling design (with replacement)
## With (98) clusters.
## svydesign(ids = ~SDPPSU6, strata = ~SDPSTRA6, weights = ~WTPFQX6,
##   nest = TRUE, data = samp)
## Stratified 1 - level Cluster Sampling design (with replacement)
## With (98) clusters.

```


[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]


```

## Stratified 1 - level Cluster Sampling design (with replacement)
## With (98) clusters.
## svydesign(ids = ~SDPPSU6, strata = ~SDPSTRA6, weights = ~WTPFQX6,
##      nest = TRUE, data = samp)
## Stratified 1 - level Cluster Sampling design (with replacement)
## With (98) clusters.
## svydesign(ids = ~SDPPSU6, strata = ~SDPSTRA6, weights = ~WTPFQX6,
##      nest = TRUE, data = samp)
## Stratified 1 - level Cluster Sampling design (with replacement)
## With (98) clusters.
## svydesign(ids = ~SDPPSU6, strata = ~SDPSTRA6, weights = ~WTPFQX6,
##      nest = TRUE, data = samp)

```