library(rms)

N.pop <- 1000000

# Size of super-population.

target.cstat <- 0.80

# c-statistic of the logistic regression model for generating outcomes.

tolerance <- 0.001

# How close should estimated c-statistic be to the target c-statistic.

set.seed(12072022)

# Set random number seed for reproducibility.

################################################################################

# Generate 10 baseline covariates for each subject in the super-population.

# The first five are from independent standard normal distibutions.

# The last five are from independent Bernoulli distributions with parameter 0.5.

################################################################################

x1 <- rnorm(N.pop,0,1)

x2 <- rnorm(N.pop,0,1)

x3 <- rnorm(N.pop,0,1)

x4 <- rnorm(N.pop,0,1)

x5 <- rnorm(N.pop,0,1)

x6 <- rbinom(N.pop,1,0.5)

x7 <- rbinom(N.pop,1,0.5)

x8 <- rbinom(N.pop,1,0.5)

x9 <- rbinom(N.pop,1,0.5)

x10 <- rbinom(N.pop,1,0.5)

X <- cbind(1,x1,x2,x3,x4,x5,x6,x7,x8,x9,x10)

# Add a column for an intercept to the matrix of baseline covariates.

################################################################################

# Initial regression coefficients for outcomes logistic regression model.

################################################################################

B.outcome.fixed <- c(log(1.25),log(1.5),log(1.75),log(2),log(2.5),

log(1.25),log(1.5),log(1.75),log(2),log(2.5))

beta.0.outcome <- -4.368896

# Intercept: determined in Example 1.

################################################################################

# Bisection procedure to determine scalar that results in the desired

# c-statistic.

################################################################################

cstat.emp <- 0

# Initial interval for scalar that multiplies the regression coefficients.

scalar.low <- 0

scalar.high <- 10

iter.cstat <- 0

while (abs(cstat.emp - target.cstat) > tolerance){

iter.cstat <- iter.cstat + 1

scalar <- (scalar.low + scalar.high)/2

B.outcome <- scalar \* B.outcome.fixed

# Modified coefficients to increase/decrease c-statistic of regression model.

beta.outcome.modified <- c(beta.0.outcome,B.outcome)

# Set the intercept of the outcome model to the given value.

XB <- X %\*% beta.outcome.modified

p.outcome <- exp(XB)/(1 + exp(XB))

Y <- rbinom(N.pop,1,p.outcome)

remove(B.outcome,beta.outcome.modified,XB,p.outcome)

# Estimate propensity score model in super-population

psm.pop <- lrm(Y ~ x1 + x2 +x3 + x4 + x5 + x6 + x7 + x8 + x9 + x10)

cstat.emp <- psm.pop$stats["C"]

remove(Y,psm.pop)

cat(iter.cstat,target.cstat,scalar,cstat.emp,file="Table3.txt",fill=T,

append=T)

if (cstat.emp < target.cstat) scalar.low <- scalar else

scalar.high <- scalar

}