Fall 2021, BIOS 740: Computing Assignment 2 (C2)

The goal of this assignment is to find the optimal treatment plan for individuals under the 2-stage decision setting. You will work on simulated data this time and here is the setup: The dataset should contain 1,000 samples and two stages of treatment, denoted by subscripts 1 and 2. The covariates are X_{11}, X_{12} for stage 1 and X_{21}, X_{22} for stage 2. All covariates follow the standard normal distribution and are independent. Treatments $A_1 \sim binom(expit(X_{11}))$ and $A_2 \sim binom(expit(X_{21}))$ where $expit(x) = \frac{1}{1+e^{-x}}$. History would then be expressed as $H_1 = \{X_{11}, X_{12}\}$ for stage 1 and $H_2 = \{X_{11}, X_{12}, A_1, X_{21}, X_{22}\}$ for stage 2. The outcome is defined as

$$Y = \exp(X_{11}) + \exp(X_{21}) + X_{11}X_{21} + A_1(1 + X_{11}) + A_2(1 + X_{11} + X_{21} + A_1X_{21}) + \epsilon$$

where $\epsilon \sim N(0, 1)$. To simulate your dataset, please use the code provided at the bottom of the second page, including the randomization seed 2019. This way everyone will have the exact same dataset.

You will need to study both data points (second stage first, then first stage) to estimate the optimal individualized treatment regime $d^{opt} = \{d_1^{opt}, d_2^{opt}\} \in \mathcal{D}$ and estimate its corresponding value function

$$\hat{V}(\hat{d}^{opt}) = \frac{\sum_{i=1}^{n} Y_i 1\{A_{1i} = \hat{d}_1^{opt}(H_{1i}), A_{2i} = \hat{d}_2^{opt}(H_{2i})\}/[P(A_{1i}|H_{1i})P(A_{2i}|H_{2i})]}{\sum_{i=1}^{n} 1\{A_{1i} = \hat{d}_1^{opt}(H_{1i}), A_{2i} = \hat{d}_2^{opt}(H_{2i})\}/[P(A_{1i}|H_{1i})P(A_{2i}|H_{2i})]}.$$

Consider the following dynamic precision medicine methods and please do not assume you know the true simulation settings (described above):

1. Q-learning (the regression model is your choice) (Moodie et al 2012)
2. Backward outcome weighted learning (BOWL, Zhao et al 2015)
3. List-based dynamic treatment regime (LIST-DTR, Zhang et al 2018)

Here are the guidelines for the report:

- Perform any preprocessing or model diagnostics as you find necessary.
- Apply the three models separately to estimate the optimal regime. Same as C1, please use a one-time or repeated k-fold cross validation to split training and testing sets. Justify any model choice.
- Present results in terms of estimated value functions and their standard errors. You will have one estimated value and one SE for each model. For each model, the estimated value function is the mean of $\hat{V}$'s across all test folds and the SE is the standard deviation of the $\hat{V}$'s across all test folds.
- Compare results of each model and elaborate your observations and interesting findings.
- Briefly explain the pros (at least 2) and cons (at least 2) of **each** approach. Identify whether the approach belongs to the indirect regression-based type, the policy learning type, or a hybrid of the two.

This assignment should be submitted as a well-written technical report with sufficient explanation in the same style as Methods and Results sections in peer-reviewed journals. It should be programmed in R using packages `DynTxRegime` (hyper-linked). More specifically, you will find function `qLearn()` and `bowl()` helpful. LIST-DTR has its own package `listdtr` (hyper-linked). To help you finish this assignment, we recommend that you look up their R documentations and tutorials as well as the original paper of the model. Here (hyper-linked) is a useful tutorial of `DynTxRegime` on the multiple decision point setting.

This assignment is **due before class on October 13**. The report needs to be typed up and no longer than 5 pages (including results but not including code) and the code part should be no longer than 3 pages. Your code should have an appropriate amount of comments in between. Please turn in report together with your code (as an appendix) as one stapled hard copy. PDF files generated from LaTeX or RMarkdown are preferred. Email submission is only allowed if notified and approved by the course instructor in advance. The quality of the report is judged by (1) completion, (2) statistical correctness, (3) code presentation, and (4) explanation and report presentation.

References

1. Moodie EEM, Chakraborty B, and Kramer MS (2012). Q-learning for estimating optimal dynamic treatment rules from observation data. Canadian Journal of Statistics 40:629–645.

2. Zhao YQ, Zeng D. Laber EB, Kosorok MR (2015). New statistical learning methods for estimating optimal dynamic treatment regimes. Journal of the American Statistical Association, 110(510), 583-598.

3. Zhang Y, Laber EB, Davidian M, and Tsiatis AA (2018). Interpretable dynamic treatment regimes. Journal of the American Statistical Association 113:1541–1549.

Hints

The simulated dataset should look like this using 2019 as randomization seed:

```
set.seed(2019)
n <- 1000
expit <- function(x) {1 / (1 + exp(-x))}
X11 <- rnorm(n)
X12 <- rnorm(n)
A1 <- rbinom(n, 1, expit(X11))
X21 <- rnorm(n)
X22 <- rnorm(n)
A2 <- rbinom(n, 1, expit(X21))
gamma1 <- A1 * (1 + X11)
gamma2 <- A2 * (1 + X11 + X21 + A1 * X21)
Y <- exp(X11) + exp(X21) + X11 * X21 + gamma1 + gamma2 + rnorm(n)
dat = cbind(X11, X12, X21, X22, A1, A2, Y) %>% as.data.frame()

> head(dat)
        X11        X12       X21       X22 A1 A2          Y
1 0.7385227 1.63030574 2.1447554 0.4679061  1  1 20.0578198
```

```
2 -0.5147605  0.47104282 -1.1144256  1.2065745  0  0  2.7665145
3 -1.6401813 -0.73062473  0.4600593 -1.2276108  0  1 -0.2612055
4  0.9160368 -0.05606844  0.6045716 -0.6038704  1  0  7.9590267
5 -1.2674820  0.71840950 -0.3231821  0.6989917  1  1  0.8485606
6  0.7382478  0.80076099 -0.3802094 -0.1553459  1  1  5.4919537
```