Causal Inference Assignment #1: 4.1

Janelle Downing

October 14, 2014

Defining the target causal parameter with a working MSM

4.1 A specific data generating process

1a. For n = 5000 children, generate the exogenous factors U and the pre-intervention covariates (V;W1;W2).

```
set.seed(252)
n = 5000
#Setting the exogenous factors:
U.V = runif(n, 0, 3)
U.W1 = runif(n, 0, 1)
U.W2 = runif(n, 0, 1)
U.A = runif(n, 0, 1)
U.Y = rnorm(n, mean=0, sd=0.1)

#Setting the endogenous factors:
V = 2 + U.V
W1 = as.numeric(U.W1 < 0.2)
W2 = as.numeric(U.W2 < (plogis(W1*0.5)))
A = as.numeric(U.A < (plogis(W1*W2 + (V/5))))
Y = 2*A + 0.3*W1 + 2*A*W2 + 0.5*A*V + U.Y
```

1b. Then set A = 1 to generate the counterfactual weight gain under RUTF Y1. Likewise, set A = 0 to generate the counterfactual weight gain under the standard supplement Y0.

```
Y.0 <-2*0 + 0.3*W1 + 2*0*W2 + 0.5*0*V + U.Y
Y.1 <-2*1 + 0.3*W1 + 2*1*W2 + 0.5*1*V + U.Y
```

2. Create a data frame X.msm consisting of age V , the set treatment levels a and the corresponding outcomes Ya, where V (i) and Ya(i) denote the age and counterfactual outcome for the ith subject.

```
#Combines into a vector
Y.a <- c(Y.0, Y.1)
#Get vector of corresponding treatments
a<- c( rep(0,n), rep(1, n))

X.msm <-data.frame(a, Y.a, V)
head(X.msm)
```

```
##   a         Y.a        V
## 1 0 -0.04796384 4.692824
## 2 0  0.25979353 4.136408
## 3 0 -0.21512580 2.982279
## 4 0  0.01580981 4.303568
## 5 0 -0.18891568 4.049335
## 6 0  0.04662055 3.114801
```

```
summary(X.msm)
```

```
##        a            Y.a                 V
##  Min.   :0.0   Min.   :-0.41888   Min.   :2.001
##  1st Qu.:0.0   1st Qu.: 0.03361   1st Qu.:2.788
##  Median :0.5   Median : 1.70139   Median :3.550
##  Mean   :0.5   Mean   : 2.47173   Mean   :3.527
##  3rd Qu.:1.0   3rd Qu.: 5.09851   3rd Qu.:4.270
##  Max.   :1.0   Max.   : 6.91742   Max.   :5.000
```

3. Evaluate the target causal parameter. Without an interaction term

```
workMSM <- glm(formula = Y.a ~ a + V, data=X.msm)
summary(workMSM)
```

```
##
## Call:
## glm(formula = Y.a ~ a + V, data = X.msm)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.71102  -0.45289   0.01894   0.57048   1.74273
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.792595   0.033153  -23.91   <2e-16 ***
## a            4.816777   0.015280  315.23   <2e-16 ***
## V            0.242670   0.008886   27.31   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.5837092)
##
##     Null deviance: 64274.0  on 9999  degrees of freedom
## Residual deviance:  5835.3  on 9997  degrees of freedom
## AIC: 23000
##
## Number of Fisher Scoring iterations: 2
```

With an interaction term

```
workMSM <- glm(formula = Y.a ~ a + V*a, data=X.msm)
summary(workMSM)
```

```
##
## Call:
## glm(formula = Y.a ~ a + V * a, data = X.msm)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.4318  -0.2472  -0.0213   0.7318   1.4832
```

```
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.07170    0.04385   1.635    0.102
## a            3.08819    0.06202  49.793   <2e-16 ***
## V           -0.00237    0.01208  -0.196    0.844
## a:V          0.49008    0.01708  28.687   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.5393628)
##
##     Null deviance: 64274.0  on 9999  degrees of freedom
## Residual deviance:  5391.5  on 9996  degrees of freedom
## AIC: 22211
##
## Number of Fisher Scoring iterations: 2
```

4. Interpret the results. The MSM without the interaction term yields coefficents of ~ B0 = -0.793, B1 = 4.817, and B2 = 0.243. If we believe that the MSM has a linear form, then the coefficent of B1 would suggest that receipt of RUTF results in a 4.817 lb higher mean counterfactual weight gain after adjusting for age. The intercept would be interpreted as the expected counterfactual weight gain is -0.793 lbs (i.e. weight loss of 0.793 lbs), regardless of treatment(A) or age(V). On average, for each year of age, children will weigh 0.243 lbs more.

The MSM with the interaction term yields coefficents of ~ ~ B0 = 0.072, B1 = 3.088, B2= -0.002, and B3 = 0.490. B0 and B2 are insignificant, no different from zero. This model appears to fit the data a bit better than the non-interacted model based on the AIC. This can be interpreted as the receipt of RUTF (taking out the effect of age) results in a 3.09 lb higher mean counterfactual weight gain. There is no difference in mean expected counterfactual weight gain by age for those who did not receive RUTF. For each additional year of age, those on RUTF can expect to have a mean counterfactual weight gain of 0.49lbs. For example, a child on RUTF who is 3 is expected to have a 3.09 + 0.49*3 (or ~ 4.60 lb) weight gain while a child of 1 would have a 3.09 + 0.49 (3.60lb) weight gain.