HW5 Q12

Zhiwei Liang

11/13/2021

```
dt = readxl::read_xlsx("HW5_Q12.xlsx")
```

1. Model for the probability of TD by time point 1 given the baseline covariates

```
fit1=glm(TD_1~CD4_0+A_0,data=dt,family = binomial(link = "logit"))
summary(fit1)
```

```
##
## Call:
## glm(formula = TD_1 ~ CD4_0 + A_0, family = binomial(link = "logit"),
##     data = dt)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.6681  -0.6681  -0.4854  -0.4854   2.0963
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.386e+00  9.597e-01  -1.444    0.149
## CD4_0       -6.931e-01  1.090e+00  -0.636    0.525
## A_0          4.833e-16  1.088e+00   0.000    1.000
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 22.967  on 27  degrees of freedom
## Residual deviance: 22.566  on 25  degrees of freedom
## AIC: 28.566
##
## Number of Fisher Scoring iterations: 4
```

2. Model for the distribution of the covariates at time point 1 given no TD and given baseline covariates

```
dt2 = dt %>% filter(TD_1==0)
fit2=glm(CD4_1~CD4_0+A_0,data=dt2,family = binomial(link = "logit"))
summary(fit2)
```

```
##
## Call:
## glm(formula = CD4_1 ~ CD4_0 + A_0, family = binomial(link = "logit"),
##     data = dt2)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.17741  -1.17741  -0.00008   1.17741   1.17741
```

```
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.957e+01  3.802e+03  -0.005    0.996
## CD4_0        1.957e+01  3.802e+03   0.005    0.996
## A_0          1.282e-16  1.000e+00   0.000    1.000
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 30.553  on 23  degrees of freedom
## Residual deviance: 22.181  on 21  degrees of freedom
## AIC: 28.181
##
## Number of Fisher Scoring iterations: 18
```

3. Model for the probability of TD by time point 2 given no TD by time point 1 and given baseline and time 1 covariates

```
dt3 = dt %>% filter(TD_1==0)
fit3=glm(TD_2~CD4_0+A_0+CD4_1+A_1,data=dt3,family = binomial(link = "logit"))
summary(fit3)
```

```
##
## Call:
## glm(formula = TD_2 ~ CD4_0 + A_0 + CD4_1 + A_1, family = binomial(link = "logit"),
##     data = dt3)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.177  -1.177   0.000   1.177   1.177
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.555e-17  9.129e-01       0        1
## CD4_0       -2.736e-16  1.000e+00       0        1
## A_0          1.813e-16  8.165e-01       0        1
## CD4_1       -2.220e-16  1.000e+00       0        1
## A_1          5.439e-16  8.165e-01       0        1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 33.271  on 23  degrees of freedom
## Residual deviance: 33.271  on 19  degrees of freedom
## AIC: 43.271
##
## Number of Fisher Scoring iterations: 2
```